A RASCH APPROACH TO VALIDATION OF CEFR-ALIGNED READING TESTS FOR TESTING UNDERGRADUATE READING COMPREHENSION

BY

MOHAMED ISMAIL FOUZUL KAREEMA

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy in Education

Kulliyyah of Education
International Islamic University Malaysia

MARCH 2023

ABSTRACT

In the ESL context, reading is an important skill necessary for academic success. Similarly, reading tests commonly are conducted in order to find out the students' ability in comprehending texts so that appropriate teaching and learning instructions are provided to enhance the skill. Applying the latest developments in testing reading and test validation, this study focused on three important objectives. The first was to produce valid and reliable instruments to measure the academic reading comprehension ability of university students in Sri Lanka by adapting the CEFR-aligned tests. The second was to examine the reading ability of students of the four faculties at SEUSL, using these validated instruments. The third objective was to investigate the students' achievement level in the cognitive processes of reading based on Khalifa and Weir's (2009) model of reading. To achieve these three objectives, 13 texts were adapted along with their (127) items from the CEFR-aligned LRN materials, and four testlets were produced. Eight cognitive processes of reading, namely Word Recognition (WR), Lexical Access (LA), Syntactic Parsing (SP), Establishing Prepositional Meaning (EPM), Inferencing (I), Building a Mental Model (BMM), Creating Text Level Structure (CTLS), and Creating Inter-Textual Representation (CITR), which are arranged hierarchically, were measured. A single test had 40 selected-response objective items including eleven common items, which had been used as anchoring items to horizontally equate four tests. The concurrent analysis of the Rasch measurement model was used to examine the psychometric properties of the tests. The findings revealed the validity and reliability of the tests and the strength of using the Rasch model for test equating. The findings also discovered that, while there was inconsistency in the hierarchical order of the cognitive processes of reading, there was consistency among the LOT (except for EPM) and the HOT processes, and the items within the same process did not have the same difficulty level, which indicates that certain cognitive processes can be used across different difficulty levels. The results also showed that 843 students, 93.5% out of 902, scored the CEFR B1 and B2 levels, which were identified as the minimum requirement for academic success in the ESL context. In addition, students' reading performance was measured according to their degree programmes with English as a-medium of instruction, and the results showed that students from the FE outperformed their counterparts in FAS, FMC, and FAC in the reading test. The study had several theoretical and practical implications in language testing and validation, and testing reading.

ملخص البحث

تعد القراءة مهارة ضرورية للنجاح الأكاديمي. ولذا يتم إجراء اختبارات القراءة من أجل معرفة قدرة الطلاب على فهم النصوص بحيث يتم توفير التعليمات المناسبة لتعليم تنمية المهارة. لذا ركزت هذه الدراسة على ثلاثة أهداف مهمة مواكبة في ذلك أحدث التطورات في مجال اختبار اللغة وفاعليتها. الهدف الأول هو إنتاج أدوات صالحة وموثوقة لقياس قدرة فهم القراءة لدى طلاب جامعيين سريلنكيين. والهدف الثاني هو فحص القدرة على طلاب جامعة الجنوب الشرقي بسريلنكا من خلال تكييف الاختبارات المتوافقة مع معيار القراءة لطلاب الكليات الأربع باستخدام التكييف، بعد أن تم قياسها بواسطة هذه الأدوات التي تم التحقق من صحتها. وأما الهدف الثالث فقد استكشف مستوى تحصيل الطلاب في العمليات المعرفية للقراءة بناءًا على نموذج خليفة وير (٢٠٠٩) للقراءة. ولتحقيق هذه الأهداف الثلاثة فإن الدراسة قد تبنت ١٣ نصًا مع الأسئلة ١٢٧ مما سمح بإعداد أربعة اختبارات. وقد تم ذلك عن طريق تبني ثمان عمليات للقراءة المعرفية وهي: التعرف على الكلمات، واسخدام المعجم، والتحليل النحوي، وإنشاء معان الجر، والاستدلال، وبناء نموذج عقلي، وإنشاء بنية لمستوى النص، وانشاء تمثيل لما بين النصوص. وقد تم ترتيب هذه العمليات بشكل هرمي. وقد احتوى كل اختبار على ٤٠ عنصرًا لإجابات عن أسئلة متعددة الخيارات بما في ذلك أحد عشر عنصرًا مشتركًا والتي تم استخدامها كعناصر إرساء لمعادلة أربعة اختبارات أفقيًّة. وقد تم استخدام التحليل المتزامن (RASCH) لنموذج قياس فحص الخصائص السيكومترية للاختبارات. كشفت النتائج عن صحة وموثوقية الاختبارات وقوة استخدامه لنموذج التحليل المتزامن (RASCH) لمعادلة الاختبار. كما أسفرت النتائج على أنه وفي ظل وجود تضارب في الترتيب الهرمي للعمليات المعرفية للقراءة، وجد أيضا اتساق في مستوى التفكير البسيط باستثناء عملية إنشاء معاني الجر ومستوى التفكير الأعلى، ولم يكن للعناصر الموجودة في العملية نفس مستويات الصعوبة، مما يشير إلى أن بعض العمليات المعرفية يمكن استخدامها عبر مستويات مختلفة الصعوبة. كما أظهرت النتائج أيضًا أن ٨٤٣ طالبًا (٩٣,٥٪) من أصل ٩٠٢ طالبًا حصلوا على مستويات في مواد شبكة مصادر التعلم (CEFR) في المستويين (B1) و (B2) واللذين يعتبران الحد الأدنى من متطلبات النجاح الأكاديمي في اللغة الإنجليزية باعتبارها لغة ثانية. بالإضافة إلى ذلك، فقد تم قياس أداء الطلاب في القراءة وفقًا للبرامج الأكاديمية التي يزاولونها والتي تبنت اللغة الإنجليزية كأداة للتعلم. وقد أظهرت النتائج بأن الطلاب الذين ينتمون إلى كلية الهندسة تفوقوا في اختبار القراءة على نظرائهم في كلية العلوم التطبيقية وكذلك كلية الإدارة والتجارة وكلية الآداب والثقافة. كما توصلت الدراسة الحالية إلى مجموعة من النتائج ذات الطابع النظري والعملي المتعلق بالاختبارات اللغوية والتصديق واختبارات القراءة.

APPROVAL PAGE

The thesis of Mohamed Ismail Fouzul Kareema has has been approved by the following:

_____
Ainol Madziah Zubairi
Supervisor

_____
Noor Lide Abu Kassim
Co-Supervisor

_____
Kamal J I Badrasawi
Co-Supervisor

_____
Mohamad Sahari Nordin
Internal Examiner

_____
Raja Safinas Raja Harun
External Examiner

_____
Noor Mohammad Osmani
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Mohamed Ismail Fouzul Kareema

Signature .... Date ....

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

A RASCH APPROACH TO VALIDATION OF CEFR-ALIGNED READING TESTS FOR TESTING UNDERGRADUATE READING COMPREHENSION

I declare that the copyright holders of this thesis are jointly owned by the student and IIUM.

Affirmed by Mohamed Ismail Fouzul Kareema

………..……………….. ………..……………..
Signature Date

This thesis is dedicated to my loving parents for the foundation of what I turned out to be in life.

To my dedicated husband MCM. Sathif

And to my wonderful children: Abdullah, Fathih Sadhaf, and Fikrath Sadhafa

May this serve as an inspiration to all of you!

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 INTRODUCTION

This introductory chapter discusses concisely the importance of reading and academic reading skills for learning, and how they are important in English as a Second Language (ESL) classes, and for university students, generally. Avowedly, reading comprehension is integral in English as a medium of instruction (EMI). Assessment of reading ability along the baseline of the Common European Frameworks of Reference (CEFR) is presented, followed by the problem statement, research objectives, rationale, and significance of the study. It also outlines the limitations of the research, operational definitions, as well as overall organization of the study.

1.2 BACKGROUND OF THE STUDY

Reading, in addition to writing, speaking, and listening, is one of the core skills in language mastery. Perfetti (1985) defined reading as the skill of decoding printed words into spoken words. However, Fries (1963) embellished the definition of reading as a process of stimulating, cultivating, and evaluating the techniques of thinking; in fact, he later mentioned that reading is thinking guided by print. Widdowson (1979) stated that reading is the process of getting linguistic information via print. This perspective has been further illustrated by the latest definition provided by Urquhart and Weir (1998), that "Reading is the process of receiving and interpreting information encoded in language from via the medium of print" (p.22).

According to Grabe and Stoller (2011), the above single-sentenced definition has four deficiencies. Firstly, it does not convey the purpose of reading; second, the nature of reading abilities was not emphasized; thirdly, it does not connect reading with the cognitive processes; and fourthly, it does not address the social context in which reading takes place.

Further, reading is viewed as a cognitive process that engages the mind, as well as eye-movement, sub-vocalisation, etc. Since the 1960s, reading has been a major focus of interest among cognitive psychologists (Urquhart & Weir, 1998). They constructed reading models on the premise that reading happens in the human mind.

Reading models are built on the assumption that reading is a process as well as a product. According to Alderson (2000), the process approach emphasizes the interaction between the reader and the text, comprising several stages. The Reading-as-a-Process model is mainly classified into the bottom-up, top-down, and interactive approaches (Birch, 2007; Birch & Fulop, 2020; Urquhart & Weir, 1998). The reader uses cultural and world knowledge and generalized cognitive strategies in the top-down approach to creating meaning for the text by prediction and inferencing. On the other hand, the bottom-up model contains the precise bits of linguistic knowledge of the text from orthographic, phonological, syntactic, and semantic perspectives, which enable the mind to squiggle the page into meaningful symbols (Birch, 2007; Birch & Fulop, 2020). Due to severe criticisms of the aforesaid models, a resultant balanced model, known as the interactive model, combining the best of both approaches, emerged. Stanovich (1980) and Rumelhart (1977), as cited by Urquhart and Weir (1998), stated that in the interactive (a balanced) model, "a pattern is synthesized based on information 'provided simultaneously from several sources" (Urquhart & Weir, 1998, p.45).

Urquhart and Weir (1998) characterized reading as a product or componential approach, in which many components are involved in the process of reading comprehension. Hoover and Tunmer (1993) mentioned that the componential model "is to understand reading as a set of theoretically distinct and empirically isolable constituents" (p. 4). Word recognition, language background, world knowledge, and literacy are among the components involved in reading (Hoover & Tunmer, 1993; Urquhart and Weir, 1998). Based on this approach, numerous reading taxonomies consisting of sub-skills of reading emerged (Grabe, 1991; Munby, 1978; Vacca & Vacca, 2008).

Reading comprehension in the first language (L1) is different from that in the second language (L2) (Birch, 2007; Grabe, 2009; Jiang, 2011). Grabe (2009) indicated three major sets of differences: linguistic and processing differences, cognitive and educational differences, and sociocultural and institutional differences; whereas Birch (2007) differentiates the six stages of L1 reading development from three types of L2 reading development procedures, such as incomplete knowledge of English, inferencing, and missing English processing strategies. However, to better understand L2 reading, the role of L1 literacy in the development of L2 reading is essential (Carrell et al., 2000; Hudson, 2007; Wade-Woolley, 1999).

1.2.1 Reading in the Second Language

Reading in L2 is a gateway to enhancing the other skills to be succeeded in a particular language. Anderson (1999) highlights that:

> Reading is an essential skill for English as a second/foreign language (ESL/EFL) students; and for many, reading is the most important skill to master. With strengthened reading skills, ESL/EFL readers will make greater progress and attain greater development in all academic areas. (p.1)

Similarly, Mikulecky (2008) mentions that reading is the key to acquiring a second language, which means that reading is the most significant fundamental instruction in all aspects of language learning. Additionally, Carrell et al. (2000) stated, "For many students, reading is by far the most important of the four skills in a second language, particularly in English as a second or foreign language" (p. 1).

Reading is recognized as a receptive skill, according to Aebersold and Field (1997), and has long been considered a prerequisite for learning a foreign language, because it serves as a critical source of input for the development of other skills. Improving one's reading activity can certainly develop one's writing and speaking skills. In other words, students who are good readers improve vocabulary, and write more grammatically compared to those who do not read much (Hafiz & Tudor, 1989). Conversely, "The studies are fairly consistent in showing that learners with

inconsequential exposure to the second language have difficulty in reading" (Hudson, 2007, p. 74) also concurred in this regard with other reading researchers.

Brown (2001) stated that reading comprehension is essentially a matter of acquiring adequate, effective comprehension skills for most second language learners who are already literate in a prior language. He suggested that both top-down and bottom-up strategies may need to be emphasized, depending on individual needs and proficiency levels.

1.2.2 Influence of Reading for Academic Success

In higher education, reading is regarded to be one of the essential skills for successful academic study (Hermida, 2009). Howard et al. (2018) mentioned that 83% of faculty members in California institutions of higher education believe that students' reading skills play a vital role in academic success. Therefore, academic reading is crucial for the L2 learners at tertiary levels while they learn a discipline through English. Academic reading has been defined as "purposeful and critical reading of a range of lengthy academic reading texts for completing the study of specific major subject areas" (Sengupta, 2002, p. 3). Further, this reading draws students into a discourse within their major studies, as well as enhancing their writing and critical thinking skills (Paul & Elder, 2008). Rather than the surface reading approach, deep reading is more effective for academic success at the university level, because university-level reading is different from school-level reading (Hermida, 2009). Internationally, reading is considered to be crucial for higher academic achievement.

To have academic success, a learner needs to be a competent comprehender (Snowling et al., 2010). According to the simple-view formula presented by Gough and Tunmer (1986), reading comprehension (RC) is equal to decoding (D) multiplied by linguistic comprehension (LC), (RC= D x LC). In the simple view, language comprehension becomes reading comprehension when word meaning is decoded or derived from print. Even if a reader has strong language comprehension, if there is difficulty with decoding, there is a possibility that the reader might be a poor comprehender. Kamhi (2007) elaborated that comprehension "is not a skill; it is a

complex of higher-level mental processes that include thinking, reasoning, imagining, and interpreting" (p. 28).

1.2.3 Reading Skill for English Medium Instruction (EMI)

Reading is a needed skill for students to master because information exists in text form in the world (Cimmiyotti, 2013). Much information is heaped in books, websites, magazines, newspapers, notice boards, notes, notices, brochures, leaflets, and sometimes pictures for visual reference for readers. Students must heavily focus on information in text formats to achieve better performance since the educational systems depend more on it. Carrell et al. (1989) highlighted that the ability to read is deliberated as an important feature to comprehend written material and to become successful in higher educational institutions, like universities.

Reading is exceedingly crucial for undergraduate students because they do not depend only on teachers, as the higher education system highly fosters self- or student-centred learning. Hence, they get themselves prepared for the new subjects by reading and understanding diverse sources alone or in groups. Therefore, it is evident that one's reading ability, especially English-related reading, fosters one's academic achievement, as was further confirmed by many research studies (Alkialbi, 2015; Anderson, 1999; Bernhardt, 2005; Grabe & Stoller, 2011; Li & Munby, 1996).

At present, English has been a medium of instruction in many countries around the world. According to Rogier (2012), Macaro et al. (2018), and Chalmers (2019), English Medium Instruction (EMI) uses English to teach curriculum subjects to students whose mother tongue or first language is not English. The popularity of EMI in school education around the globe has dramatically increased in recent decades; traditionally, this has been mainly in higher education. To compete in the international education market, universities started to offer courses, modules, and entire degree programmes in English to attract foreign students. To prepare the children to enter such universities, parents demanded the EMI approach in the "secondary", "primary" and "preschool" curricula (Chalmers, 2019, p. 8).

If EMI is to be practised at the higher education level, students have to read and comprehend enormous amounts of texts to gain knowledge, listen to lectures, interact in the classroom, take notes, present on given topics, and write assignments and final exams in English. Thus, as it is required by many foreign universities for university admission, students must attain the C1 level of the CEFR, which illustrates the ability to use English fluently and flexibly in a wide range of contexts (Cambridge University Press, 2013).

1.2.4 Assessing Reading

Assessing reading is an intricate procedure similar to defining the nature of reading comprehension. Alderson (2000) illustrates that there are various ways of looking at how reading is developed and assessed. Using reading scales with a detailed description of each level, point, or band is one of the ways to assess reading. ACTFL proficiency guidelines, ALTE framework of language tests, Master and Forster scales, DIALANG, and CEFR can-do descriptors are some of such scales. Using language tests with different levels or bands is another way of assessing reading. These include Cambridge ESOL main suite exams like Key English Test (KET), Preliminary English Test (PET), the First Certificate in English (FCE), the Certificate in Advanced English (CAE), Certificate of Proficiency in English (CPE), and TOEFL, International English Language Testing System (IELTS), Learning Resource Network (LRN) ESOL exams; and International English Language Competency Assessment (IELCA).

American Council for the Teaching of Foreign Languages (ACTFL)'s reading definitions focus on text type, reading skill, and task-based performance. These guidelines are commonly used and influential in the USA. The guidelines lack familiarity as they are based on *a priori* definitions of levels and there is no empirical validation (Alderson, 2000).

The Association of Language Testers in Europe (ALTE) has developed a framework of levels, particularly for ALTE member language tests. It presents a general description of what a learner can do at each level before describing each skill separately (ALTE, 2002). According to the ALTE context, text type, language, and

6

reader's knowledge about the content are needed to be considered when developing reading, while it improved confidence, speed, awareness, length and amount of text, nature of the text, and text practicability (Alderson, 2000)

1.2.5 CEFR

The Common European Framework of Reference (CEFR) is a modified version of ALTE (Council of Europe, 2001a). ALTE's five levels have been aligned with A2 to C2 levels of the CEFR Framework (ALTE, 2002). It has three main groups comprising two stages each. It is intended to provide a common basis for describing "levels of proficiency required by existing standards, tests, and examinations in order to facilitate comparisons between different systems of qualification" (Cambridge University Press, 2013; Council of Europe, 2001, p.21).

Researchers advocate that a university student following the EMI system should be at the C1 level of CEFR (Council of Europe, 2001a; Jiménez-Muñoz, 2014). The Common European Framework of Reference for Languages (CEF or CEFR) is a way of standardizing the levels of language exams in different regions, introduced by the Council of Europe in 1996. Though it was intended to apply to European countries, as the CEFR descriptors have been translated into 40 European languages, including sign language, its influence is unquestionable in language teaching, learning and assessment beyond Europe (Figueras, 2012).

CEFR has been extensively utilized by many organizations and educational institutions as a reference tool for teaching, learning, and assessment for the last decade (North, 2014a; Waluyo, 2019; Wu & Wu, 2007).. In accordance with CEFR, language users are clustered into three main groups: Proficient users (levels C1 & C2), Independent users (levels B1 & B2), and Basic users (levels A1 & A2) (Council of Europe, 2001; Cambridge University Press, 2013). The CEFR levels represent a 'conceptual grid' of illustrative *can-do* descriptors of language competence, which was intended to be applied equally across different European languages since the 1980s (North, 2014b). A comprehensive Swiss research project scaled the levels through empirical Rasch analysis (North & Schneider, 1998).