

WASHBACK STUDY OF AN OUTCOME BASED  
ENGLISH LANGUAGE ASSESSMENT ON STUDENT  
LEARNING

BY

NORHASLINDA BINTI HASSAN

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 2022

## ABSTRACT

The assessment reform around the world has culminated in the adoption of various assessment systems that focus on student learning, and in the case of higher Malaysia, Outcome Based Assessment (OBA) was adopted. Since OBA was claimed to be an effective system to enhance student learning, it is interesting to note that learning may be debilitated by poor assessment practices. Thus, the present study investigated how OBA has thus far influenced student learning and their learning outcomes by means of sequential exploratory mixed methods with item development. Specifically, this study investigated the intended and unintended washback of an OBA, i.e., Integrated Language Skills III test battery, focusing on its washback effect on students' learning and learning outcomes. A survey named WOLOS was developed and validated by means of qualitative (semi-structured interview) and quantitative analysis (Item Objective Congruence (IOC) and Rasch Measurement Model). Drawing on an argument-based validation approach, i.e., Assessment Use Argument (AUA), the evidence about the washback effect of Integrated Language Skills III test battery were linked to construct validity. To better understand how the students experienced Integrated Language Skills III test battery, students who have taken this test battery were selected as the participants of this study. Document analysis, semi-structured interviews, focus group discussion and survey data were collected. The document analysis provides information on the intended washback of Integrated Language Skills III test battery. Semi-structured interviews and focus group discussions involving 22 students were analysed using NVivo12, by means of phenomenology approach. Rasch Measurement Model was employed to analyse the survey data from 309 respondents. A range of factors were found to have mediated the intended washback, both positively and negatively. In short, it was found that Integrated Language Skills III test battery brought about more positive washback owing to its 10 years of implementation. Further directions may involve studies which include various stakeholders' perspectives to perceive the consequences of a test and focusing on the impact of technology in teaching and learning.

## خلاصة البحث

تُوجت عملية إصلاح التقييم حول العالم، بتبني أنظمة تقييم مختلفة، تركز على تعلم الطلاب، وفي السياق الماليزي، تم اعتماد التقييم المستند إلى النواتج؛ وذلك نظرًا لأن الأوساط المعنية تزعم أنه نظام فعال لتعزيز تعلم الطلاب. ومن المثير للاهتمام ملاحظة أن عملية التعلم قد تضعف بسبب ممارسات التقييم السيئة. وبالتالي، فقد بحثت الدراسة الحالية في كيفية تأثير التقييم المستند إلى النواتج في عملية تعلم الطلاب، ونواتجهم التعليمية، وذلك عن طريق الأساليب الاستكشافية المختلطة المتسلسلة مع تطوير العناصر. وعلى وجه التحديد، بحثت هذه الدراسة في التأثير المقصود وغير المقصود لعملية التقييم المستند إلى النواتج، على سبيل المثال باستخدام البطارية المتكاملة لاختبار مهارات اللغة (3)، مع التركيز على تأثيرها على عملية تعلم الطلاب، وكذلك على نواتج تعلمهم. تم تطوير دراسة استقصائية باسم (WOLOS-وُلوس)، مع التحقق من صحتها عن طريق التحليل الكيفي (مقابلة شبه منظمة)، والتحليل الكمي (تطابق هدف العنصر)، ونموذج "راش" للقياس. وبناءً على نهج التحقق القائم على الحجج، على سبيل المثال حجة استخدام التقييم، تم الاعتماد على دليل تأثير البطارية المتكاملة لاختبار مهارات اللغة (3)، من أجل بناء الصلاحية. ولفهم كيفية اختبار الطلاب للبطارية المتكاملة لاختبار مهارات اللغة (3) بشكل أفضل، تم اختيار الطلاب الذين خضعوا لهذا الاختبار، ليكونوا مشاركين في هذه الدراسة. وقد تم جمع البيانات باستخدام تحليل الوثائق، والمقابلات شبه المنظمة، ومناقشة مجموعة التركيز، والاستبانة. حيث يوفر تحليل الوثائق معلومات عن البطارية المتكاملة لاختبار مهارات اللغة (3). أما تحليل المقابلات شبه المنظمة، ومناقشات المجموعة المركزة، التي ضمت (22) طالبًا، فقد تم تحليلها باستخدام برنامج (NVivo12)، عن طريق استخدام منهج دراسة الظواهر. كما تم استخدام نموذج "راش" للقياس، من أجل تحليل بيانات الاستبانة الواردة من (309) مشاركين. وقد تم العثور على مجموعة من العوامل التي كان لها دور في التأثير المقصود، إيجابًا وسلبيًا. وخلاصة القول، أنه قد وُجد أن البطارية المتكاملة لاختبار مهارات اللغة (3)، قد أحدثت تأثيرًا إيجابيًا بشكل أكبر، بسبب مدة الـ(10) سنوات التي قُضيت في عملية تنفيذها. ويُرجى أن تتضمن الدراسات المستقبلية اتجاهات إضافية تشمل وجهات نظر مختلفة لأصحاب المصلحة، وذلك من أجل بناء تصور حول عواقب استخدام اختبار ما، وكذلك التركيز على تأثير التكنولوجيا في عملية التعليم والتعلم.

## **APPROVAL PAGE**

The thesis of Norhaslinda binti Hassan has has been approved by the following:

---

Ainol Madziah Zubairi  
Supervisor

---

Ismail Sheikh Ahmad  
Co-Supervisor

---

Zainurin Abdul Rahman  
Co-Supervisor

---

Noor Lide Abu Kassim  
Internal Examiner

---

Nurahimah Mohd Yusoff  
External Examiner

---

Mohd Ramzi Mohd Hussain  
Chairman

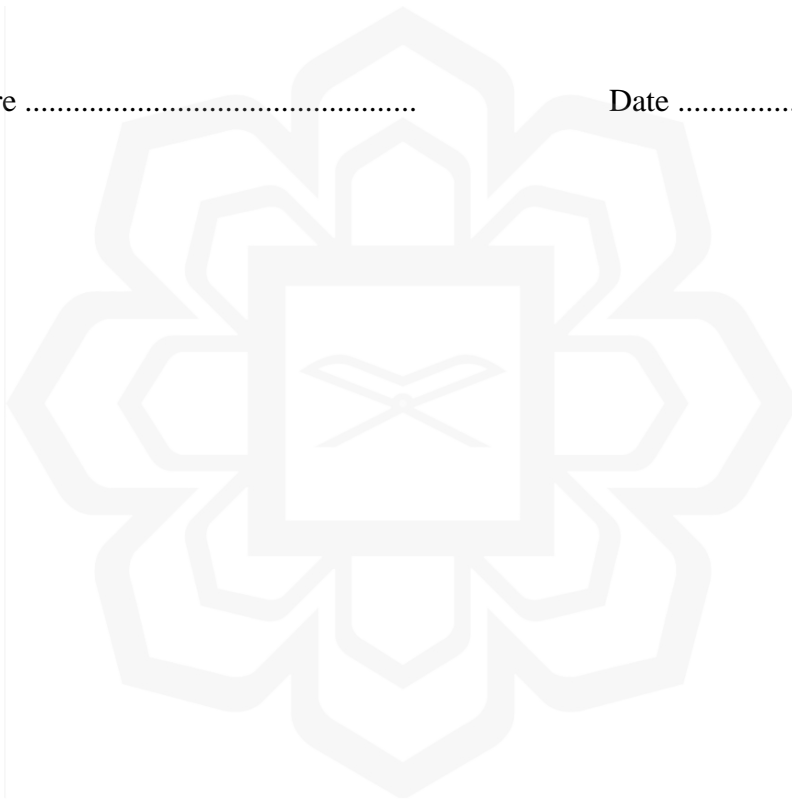
## DECLARATION

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.

Norhaslinda binti Hassan

Signature .....

Date .....



**INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA**

**DECLARATION OF COPYRIGHT AND AFFIRMATION OF  
FAIR USE OF UNPUBLISHED RESEARCH**

**WASHBACK STUDY OF AN OUTCOME BASED ENGLISH  
LANGUAGE ASSESSMENT ON STUDENT LEARNING**

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

Copyright © 2022 Norhaslinda binti Hassan and International Islamic University Malaysia. All rights reserved.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.
2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.
3. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Norhaslinda binti Hassan

.....  
Signature

.....  
Date

## ACKNOWLEDGEMENTS

In the name of Allah, the most merciful, the most compassionate. Alhamdulillah.

I would like to express my gratitude for being able to complete my PhD journey. In the course of finishing this manuscript, I am indebted to a lot of people, my supervisors, family and friends. It is my utmost pleasure to dedicate this work to my dear parents, my husband and my three little ‘monsters’, who granted me encouragement, support and patience.

Firstly, I would like to express my sincere gratitude to my dear supervisor, Professor Ainol Madziah Zubairi, for her continuous guidance that kept me on the track. Her insights: emotionally, theoretically, are the things that keep me moving in my PhD journey. To Professor Ismail Sheikh Ahmad, my co-supervisor, your insights on qualitative analysis have lit up my understanding on handling qualitative data. Thank you for your time and support, Prof. My second co-supervisor, Assoc. Prof. Dr. Zainurin, has always been very supportive and encouraged me to keep on track and stay focused. I appreciate it very much and I pray that I will be as good as my supervisors; knowledgeable, supportive, and very down to earth.

Secondly, to Mak (Yah Binti Man) and Ayah (Hassan Bin Noor), thank you is not enough as I am indebted to you for always being my biggest supporters. Throughout my PhD journey, I’m indebted to my parents for always being there to take care of my kids and giving me the allowances to stay focused on my work. I cannot thank you much and I love you so much. My gratitude also goes to my siblings who always encouraged me to move further and always be there when I need them.

My dearest love to my other half, Abbas for supporting me from the day I intend to pursue my PhD. Pouring his knowledge and love to me, I am out of this world! Thank you for sharing your experiences, as well as being my number one critic, without which, I would not know my weaknesses and help me to be better. Most importantly, your sense of humour has always kept me smiling. To my dear Waels: Farhan, Harith and Emran, Mama is truly indebted to you guys. My profuse apologies for always being missing, no vacation during school holidays and not enough quality time with you guys. Thank you very much for being very understanding.

A few colleagues helped in the data collection and analysis of this study, especially my dearest friends: Mal, Pjoy, Kak Nani and Marni. A special thanks also to all the participants of the study, without which, this study would not be accomplished.

This journey would not be accomplished without the support of Universiti Teknologi Mara and the Ministry of Education Malaysia. I am forever grateful for the funding I received for my PhD program.

# TABLE OF CONTENTS

Abstract .....	ii
Abstract in Arabic .....	iii
Approval Page.....	iv
Declaration .....	v
Copyright .....	vi
Acknowledgements.....	vii
List of Tables .....	xi
List of Figures .....	xiii
<b>CHAPTER ONE: INTRODUCTION.....</b>	<b>1</b>
1.1 Introduction.....	1
1.2 Context of the Study .....	3
1.3 Statement of Problem .....	10
1.4 Research Objectives.....	15
1.5 Research Questions.....	15
1.6 Significance of the Study.....	16
1.7 Operationalized Definition of Terms.....	17
1.8 Thesis Overview .....	19
<b>CHAPTER TWO: LITERATURE REVIEW.....</b>	<b>21</b>
2.1 Introduction.....	21
2.2 Outcome Based Assessment.....	21
2.2.1 Defining OBA .....	21
2.2.2 Features of OBA .....	24
2.2.3 Reliability, Validity and Fairness in OBA .....	27
2.2.4 Issues with the Implementation of OBA.....	30
2.2.5 OBA in UiTM .....	33
2.3 Washback.....	34
2.3.1 The Origin and Definition of Washback.....	34
2.3.2 Washback in language testing.....	36
2.3.3 The Nature and Scope of Washback .....	43
2.3.3.1 Washback and Impact.....	43
2.3.3.2 Dimensions of Washback .....	44
2.3.3.2.1 Specificity.....	45
2.3.3.2.2 Intensity .....	45
2.3.3.2.3 Length.....	46
2.3.3.2.4 Intentionality.....	46
2.3.3.2.5 Value.....	47
2.3.3.3 Washback and Validity.....	48
2.3.3.4 Test Validation Framework .....	50
2.3.3.5 Defining washback in the present study .....	59
2.3.4 Rasch Measurement Model.....	59
2.3.4.1 Rasch Model Analysis for Survey Validation .....	60
2.3.4.2 Rasch Analysis for analysing survey.....	64
2.3.5 Review of Empirical Studies.....	67

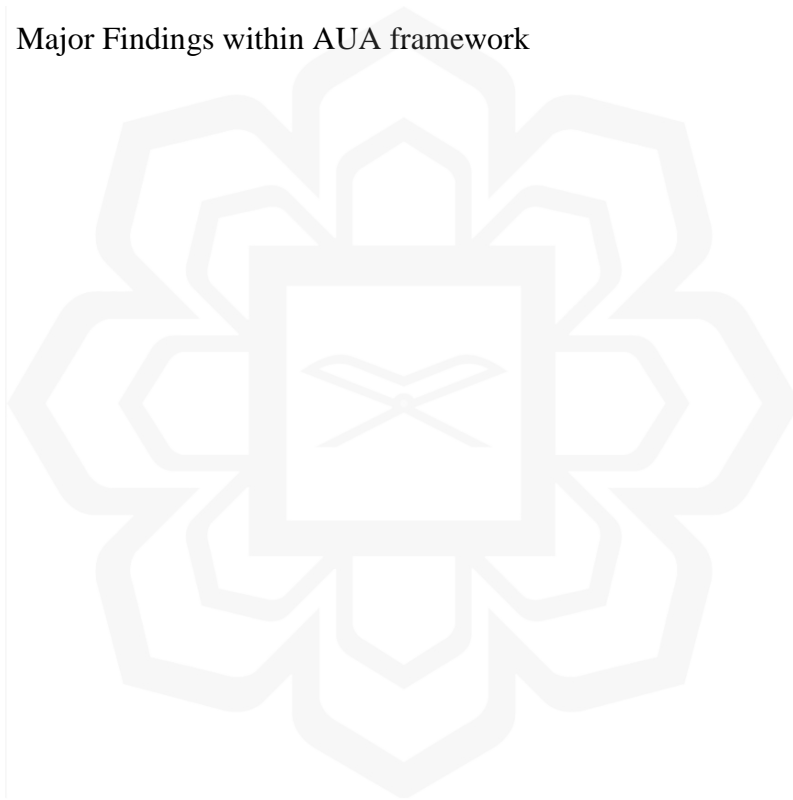
2.3.5.1 Washback on Students' Perceptions (Participants) .....	70
2.3.5.2 Washback on Learning Processes (Process).....	75
2.3.5.3 Washback on Learning Outcome (Product) .....	80
2.3.6 Conceptual Framework .....	84
2.4 Chapter Summary .....	85
<b>CHAPTER THREE: METHODOLOGY.....</b>	<b>86</b>
3.1 Introduction.....	86
3.2 Restatement of the Research Questions.....	86
3.3 Researcher's Role in this Study.....	88
3.4 Sequential Exploratory Mixed Method with Item Development .....	90
3.4.1 Triangulation .....	93
3.5 Selection of Informants .....	95
3.6 Instruments .....	97
3.6.1 Document Analysis .....	98
3.6.2 Semi-Structured Interviews.....	98
3.6.3 Questionnaire .....	100
3.7 Data Collection .....	110
3.8 Data Analysis.....	112
3.9 Issues of Credibility and Trustworthiness .....	113
3.10 Ethical Issues .....	114
3.11 Chapter Summary .....	114
<b>CHAPTER FOUR: RESULTS .....</b>	<b>115</b>
4.1 Introduction.....	115
4.2 Demographic Information .....	115
4.2.1 Quantitative respondents.....	116
4.2.2 Qualitative informants.....	117
4.3 Intended Washback of Integrated Language Skills III Test Battery.....	118
4.4 Research Question 1: Students' Conception of the Integrated Language Skills III Test Battery .....	119
4.4.1 Psychometric properties of the conception section.....	120
4.4.2 The importance of Integrated Language Skills III test battery .....	122
4.4.3 Pre-requisite assessments .....	126
4.4.4 Perceived difficulty .....	128
4.4.5 Summary of Research Question 1 .....	132
4.5 Research Question 2: The Influence of Integrated Language Skills III Test Battery on Students' Learning.....	134
4.5.1 In-class learning .....	134
4.5.1.1 Psychometric properties of the instrument to measure the in- class learning experiences .....	134
4.5.1.2 In-class learning practices.....	137
4.5.1.3 Assessment .....	146
4.5.1.4 Factors promoting washback (in-class learning) .....	151
4.5.1.5 Factors inhibiting washback (in-class learning) .....	153
4.5.2 Summary of Research Question 2a .....	157
4.5.3 Out-of-class learning.....	159
4.5.3.1 Psychometric properties of the instrument to measure out-of- class learning experiences .....	159

4.5.3.2 Out-of-class learning practices .....	161
4.5.3.3 Out-of-class test preparation.....	172
4.5.3.4 Factor promoting out-of-class learning .....	181
4.5.3.5 Factors inhibiting out-of-class learning.....	182
4.5.4 Summary of Research Question 2b.....	185
4.5.5 Research Question 3: The Products of Learning Integrated Language Skills III Test Battery.....	187
4.5.5.1 Psychometric properties of the instrument to measure the learning outcome .....	187
4.5.5.2 Promotion of Language Skills (Writing).....	189
4.5.5.3 Promotion of Language Skills (Speaking).....	195
4.5.5.4 Promotion of Language Skills (Reading) .....	198
4.5.5.5 Promotion of Language Skills (Overall).....	200
4.5.5.6 Learner Affect (Motivation and self-confidence).....	203
4.5.6 Summary of Research Question 3.....	204
4.6 Chapter summary.....	205
<b>CHAPTER FIVE: DISCUSSION AND CONCLUSIONS.....</b>	<b>206</b>
5.1 Introduction.....	206
5.2 Intended Washback of Integrated Language Skills III .....	207
5.3 Main Findings .....	208
5.3.1 Major Findings of Research Question 1 .....	208
5.3.2 Major Findings of Research Question 2 .....	212
5.3.3 Major findings of research question 3.....	221
5.4 Contributions and Implications of the Study .....	229
5.5 Limitations of the Study .....	231
5.6 Further Directions.....	232
<b>REFERENCES.....</b>	<b>234</b>
<b>APPENDIX A: INTEGRATED LANGUAGE SKILLS III OBE TEMPLATE .....</b>	<b>249</b>
<b>APPENDIX B: ETHICAL APPROVAL .....</b>	<b>251</b>
<b>APPENDIX C: CONSENT FORM .....</b>	<b>252</b>
<b>APPENDIX D: INTERVIEW PROTOCOL .....</b>	<b>254</b>
<b>APPENDIX E: INTER-RATER RELIABILITY .....</b>	<b>256</b>
<b>APPENDIX F: WOLOS .....</b>	<b>269</b>
<b>APPENDIX G: IOC .....</b>	<b>280</b>
<b>APPENDIX H: RASCH ANALYSIS OF PILOT ITEMS .....</b>	<b>282</b>
<b>APPENDIX I: PERSONAL COMMUNICATION WITH SCHOLARS .....</b>	<b>291</b>

## LIST OF TABLES

<u>Table No.</u>		<u>Page No.</u>
1.1	The development of Integrated Language Skills III	6
1.2	Integrated Language Skills III Entrance and Exit Survey	8
2.1	Criteria and Statistical Info for Validation	61
2.2	Empirical studies on washback on learners and learning	68
3.1	Models of saturation and their principal foci in the research process	96
3.2	Outline of moderator's guide	99
3.3	Methods and outcomes of WOLOS development	101
3.4	Cut-off point for IOC	102
3.5	Rasch analysis on the criteria of WOLOS	105
3.6	Summary of Rasch Analysis on the criteria of WOLOS	107
3.7	Model S.E for every section	109
3.8	Informants and duration of interviews	111
4.1	Demographic information of quantitative respondents	116
4.2	Demographic information of qualitative informants	117
4.3	Psychometric properties of conception	120
4.4	Person reliability and item reliability for conception	122
4.5	Item measure of perceived importance	123
4.6	Perceived Importance Frequency	124
4.7	Psychometrics properties of in-class learning	135
4.8	Person and item reliability of in-class learning section	137
4.9	Item measure for in-class learning practices	137
4.10	Item measure for in-class test-related activities	146
4.11	Psychometric properties of out-of-class learning	159

4.12	Person and item reliability of out-of-class learning	161
4.13	Item measure for out-of-class learning practices	161
4.14	Item measure for out-of-class test preparation	172
4.15	Psychometric properties of learning outcome	187
4.16	Person reliability of learning outcome	189
4.17	Item measure of learning outcome on writing skill	190
4.18	Item measures for speaking skill	195
4.19	Learning outcome on reading skill	198
5.1	Major Findings within AUA framework	207



## LIST OF FIGURES

<u>Figure No.</u>		<u>Page No.</u>
1.1	OBE-SCL timeline in UiTM (RoZIAH & Zainab, 2011)	4
2.1	Teaching, learning and assessment as a continuous four-step cycle	23
2.2	A basic model of washback (Bailey, 1996, p. 264)	39
2.3	Booth's sociocultural washback model	41
2.4	Washback as separate but related to impact	43
2.5	Washback as a subset of impact	44
2.6	The structure of AUA (Bachman, 2005 p.25)	56
2.7	Conceptual framework of the study	84
3.1	Research Design of the present study	93
3.2	Triangulation framework of the study	95
4.1	Category probability curves of conception	121
4.2	The pre-requisite assessments	127
4.3	Frequency count of perceived difficulty	129
4.4	Informants' perceived difficulty	130
4.5	Category probabilities for in-class learning section	136
4.6	Ask friends vs ask teacher	138
4.7	Skills focused	140
4.8	The usage of WhatsApp and video & slides in classroom	142
4.9	Technology used in the classroom	143
4.10	In-class test-related activities	148
4.11	In-class feedback	150
4.12	Factors promoting in-class learning	151
4.13	Factors inhibiting in-class learning (quantitative results)	153

4.14	Factors inhibiting in-class learning (qualitative findings)	154
4.15	Category probability of out-of-class learning	160
4.16	Technology usage in out-of-class learning	162
4.17	Songs	165
4.18	Discussion with friends and speak English	167
4.19	Movie subtitle	169
4.20	Reading out-of-class	170
4.21	Test prepared (out-of-class)	175
4.22	Out-of- class test preparation	179
4.23	Factors promoting out-of-class learning	181
4.24	Factors inhibiting out-of-class learning (quantitative)	182
4.25	Factors inhibiting out-of-class learning (qualitative)	183
4.26	Category probability curves of learning outcome	188
4.27	Essay writing	191
4.28	Evaluative commentary	193
4.29	Oral commentary	196
4.30	Reading test	199
4.31	Overall	201
5.1	The AUA structure of Integrated Language Skills III test battery	225
5.2	The AUA structure of Integrated Language Skills III speaking assessment	226
5.3	The AUA structure of Integrated Language Skills III reading assessment	227
5.4	The AUA structure of Integrated Language Skills III writing assessments	228

# CHAPTER ONE

## INTRODUCTION

### 1.1 INTRODUCTION

Assessment has been proven to be an influential tool for exercising change, as evidenced by the assessment reform that has in recent years enveloped the globe through the implementation of new paradigms in educational institutions with the objective of meeting national economics imperatives, i.e., producing quality graduates. Educational institutions have started adopting systems that are using ‘standards’, ‘benchmark’, ‘competencies’, to name a few, to inform learners’ achievement and progress (Brindley, 1998, 2001). There is a paradigm shift from teacher-centred to student-centred approach, which focuses on student learning and “what students do to achieve learning, rather than what the teacher does” (O’Neil & McMahon, 2005, p35). To enhance student learning, the adoption of assessment systems such as learning oriented assessment (Carless, 2007) and classroom-based teacher-led formative assessments (Cheng, 2013; Brookhart, 2004) is evident throughout the world, i.e., School-based Assessment, Outcome-Based Assessment, etc. Malaysia is not exempt from this reform as in 2007, its government announced the implementation of Outcome-based Education (OBE) at all Higher Learning Institutions (HLIs) throughout the country to enhance student learning. Its assessment system, Outcome-based Assessment (OBA) has been used as a tool to measure student learning. It has to be noted that students are at stake when it comes to assessment practices, since poor assessment practices within efficient assessment may culminate in debilitating learning (Crusan, Plakans & Gebiril, 2016). This has been the impetus to

the researcher in investigating how the adoption of OBA; the widely acclaimed effective assessment, has thus far influenced student learning.

In the area of language testing, the phenomenon of tests/examinations influencing teaching and learning practices is known as washback. The burgeoning studies for washback mushroomed following the flagship paper of Alderson and Wall (1993) 'Does washback exist?'. It has been over two decades and studies on washback keep evolving due to the complexity of this phenomenon. Upon reviewing the literature, it was found that the first wave of washback studies witnessed scholarly works addressing the question of whether washback does exist. Moving a step further, studies on washback provide answers to questions like "what does washback look like?", "what brings washback about?" and "why does washback exist?" (Alderson, 2004: ix). Recently, washback studies tended to focus on promoting positive washback and mitigating negative washback.

OBA is introduced to promote learning, i.e., OBA facilitates in enhancing student learning by means of formative and summative assessments. Upon reviewing the literature and reflecting on the researcher's experience as a language teacher, the present study focuses on investigating the washback effects of an OBA on learning outcome. In doing so, the present study sets out to examine the consequences of using an OBA in the Malaysian educational context, particularly its washback, i.e., its influences on student learning and learning outcomes by means of drawing on Bachman and Palmer's (2010) assessment use argument validation model. Specifically, this study investigates how students experience washback from enrolling in an OBA, i.e., Integrated Language Skills III test battery in Universiti Teknologi MARA (UiTM), Penang Branch Campus.

## 1.2 CONTEXT OF THE STUDY

In recent years, governments in many countries worldwide have announced the implementation of a new paradigm in educational institutions to meet national economic imperatives through a demonstration of ‘efficiency and cost effectiveness’ with more rigorous reporting mechanisms of program outcomes (Brindley, 1998, 2001). According to Blackmore (1998:35),

‘the primary concern of economic accountability is not *process*, but the *outcome*. It is driven by utilitarian goals in relation to maximising the benefits with the limited resources at disposal. It is market-oriented.’

In line with the emphasis on outcomes, government policymakers and bureaucrats in many countries have put major efforts into the development of a system that allows the performance of educational institutions and systems to be closely monitored and evaluated (Fitz-Gibbon, 1990; Singh, 1990; Riley and Nuttall, 1994 as cited in Brindley, 1998). This has led to the widespread adoption of systems that are using ‘standards’, ‘benchmarks’, ‘competencies’ and ‘attainment targets’ as a basis for assessing and reporting learners’ progress and achievement which have increasingly become the cornerstone of assessment and reporting systems in Europe, North America, Australia, New Zealand and Asia (Brindley, 1998).

Malaysia has in recent years, announced its decision to implement OBE at all HLIs in line with the National Higher Education Action Plan 2007-2010 and the ‘Strategic Plan for Higher Education: Laying the foundation Beyond 2020’ (Country Report - Malaysia, 2008). This strategic plan was initiated to align with the Ninth Malaysia Plan (2006-2010), which focuses on developing world-class human capital. The recent call made by the government in propagating an outcome-based approach has culminated in only some institutions immediately adapting or adopting this

assessment system across faculties, while the others have decided to observe the pioneer institutions, possibly with the aim of modelling the best practices.

In this regard, UiTM is among the institutions, which decided to implement OBE across all the faculties of its campuses, and this includes the Academy of Language Studies (ALS), the focus of the present study. UiTM introduced Outcome-Based Education and Student-Centred Learning (OBE-SCL) 6+1 modules in July 2007 with the purpose of maximizing the learning engagement via various approaches in pedagogical knowledge. The modules include outcome-based education and student-centred learning, effective course design & planning: constructive alignment of learning outcomes (LOs), instruction & assessment, delivering courses in making learning effective, assessing course outcomes (COs), assessing program outcomes (POs), measurement and monitoring & closing the loop and aligning assessment in OBE-SCL (<http://www.obescl.uitm.edu.my/>).

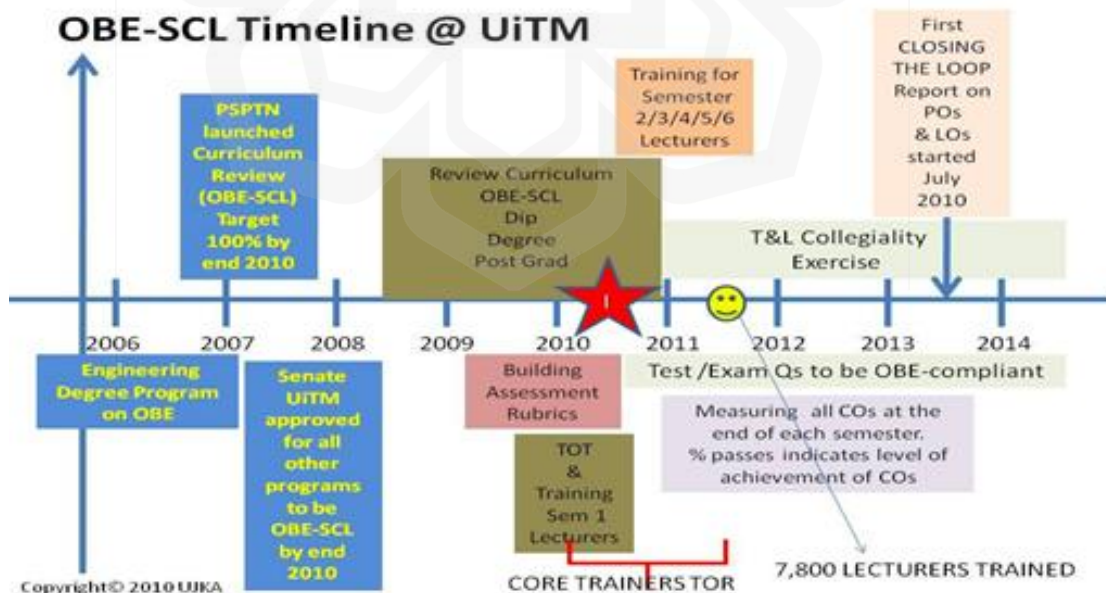


Figure 1.1 OBE-SCL timeline in UiTM (Roziyah & Zainab, 2011)

Figure 1.1 shows the OBE-SCL timeline applicable for all faculties in UiTM. Apart from the engineering degree program, the senate of UiTM has encouraged all the other programs to employ OBE-SCL by the end of 2010. In mid of 2008, UiTM started to alter its curriculum at diploma, degree, and postgraduate levels in line with the government's policy of OBE, which includes ALS. Consequently, UiTM has started training its academic staff members on the newly introduced OBE-SCL in all its faculties in campuses around the country. The first Closing the Loop Report on POs and LOs started in July 2010. Closing the loop (CDL) is a mechanism to measure the effectiveness of the OBE-SCL curriculum of all courses under a program. CDL indicates whether there are alignments among CO-PO- LO-KI (Course Outcome- Program Outcome-Learning Outcome and soft skills). The results of the CDL analysis can be used for the quality improvement purpose. The tool consists of six sections: CDL outcomes attainment report (automatically generated from other sections), Course CGPA, CO-PO-LO-KI (lecturer's perception), frequency gap of Exit-Entrance Survey, SuFO (Students' Feedback On-line) and OBE-SCL implementation section. The Continuous Quality Improvement (CQI) tool, on the other hand, is a report that accompanies the CDL 8 version to address possible issues and suggestions which leads to improving the quality of a program.

It is noteworthy that the most important aspect in OBE is the LOs and hence, a curriculum has to start 'with a clear picture of what is important for students to be able to do' (Spady, 1994, p.1). Since the focus of planning and instruction of education is the outcomes, the extent to which the outcomes have been realised is measured by means of a tool, i.e., assessment (Killen, 2000). Hence, within the OBE system, OBA is the tool used to observe and measure the extent to which the learners achieve the LOs by means of valid, reliable, and fair assessment.

For graduation purposes, students at diploma level in UiTM must pass 3 levels of proficiency courses and thus, making these proficiency courses high stakes for them. The final or third level of the proficiency courses is the concern of this study. Previously, this course was known as English for Academic Purpose or BEL311. With the implementation of OBE/OBA in UiTM, this course has been improved by means of the CDL and CQI. In 2014, the revised version of proficiency courses in UiTM were known as ELC. In the case of BEL311, it was revised as ELC230 (Integrated Language Skills: Writing). With different course codes and names, the COs and syllabus changed. In 2016, another revision was done, and the course changed to ELC231 (Integrated Language Skills III). For ease of reference, Table 1.1 is tabulated to demonstrate the changes in the test battery of Integrated Language Skills III:

Table 1.1 The development of Integrated Language Skills III

	<b>On-going Assessments</b>	<b>Marks</b>	<b>Final Exam</b>	<b>Marks</b>
BEL311	A written assignment (400-700 words)	30%	Reading	20%
	Group Discussion	20%	Writing (300-400 words)	20%
	On-line Assignment	10%		
ELC230	Writing (Paragraph Writing)	10%	Writing <ul style="list-style-type: none"> <li>● Outline</li> <li>● Argumentative / Discussion-type essay (300-350 words)</li> </ul>	10%
	Evaluative Commentary	10%		30%
	Mid Sem Test (Reading & Writing Test)	20%		
	Speaking	15%		
	Online Practices	5%		
Integrated Language	Essay Writing (Expository)	15%	Writing <ul style="list-style-type: none"> <li>● Outline</li> </ul>	10%

Skills III	Evaluative Commentary (Writing)	10%	● Argumentative / Discussion-type essay (300-350 words)	30%
	Commentary (Oral)	15%		
	Mid Sem test (Reading & Report Writing)	20%		
	Essay Writing	20%		
Integrated Language Skills III September 2017 onwards	Evaluative Commentary	30%	NA	NA
	Mid-Semester Test (Reading Test)	25%		
	Oral Commentary	25%		

The present study zoomed into the latest version of Integrated Language Skills III, i.e., September 2017 onwards. It is a core course with 3 credit units and 4 contact hours. Therefore, students have 2 face to face sessions per week (2 hours per session). This course is undertaken by third semester students and the prerequisite is Integrated Language Skills II, the course taken by students during their second semester. The main objective of this course is to equip students with necessary writing skills, in which speaking, reading, and writing skills are integrated with the emphasis on writing skills.

The course outcomes of Integrated Language Skills III are:

At the end of Integrated Language Skills III, students should be able to:

1. Demonstrate the ability to speak confidently based on non-academic issues in a social setting.
2. Demonstrate the ability to respond to questions by applying a variety of reading strategies based on authentic and non-authentic discourses.

3. Demonstrate the ability to write an expository essay and evaluative commentary in academic /non- academic context individually.

Grounded on the learning outcomes, the course Entrance and Exit survey is formulated and presented in Table 1.2 below:

Table 1.2 Integrated Language Skills III Entrance and Exit Survey

CO	NO	ITEMS
CO1	1.	I can produce responses orally to a given situation using appropriate speaking strategies.
	2.	I can produce an oral commentary of an English movie/film/TV program.
CO2	3.	I can apply general reading strategies.
	4.	I can paraphrase and summarize information.
CO3	5.	I can analyse and evaluate reading texts.
	6.	I can identify the structure of a paragraph.
	7.	I can identify the structure of an essay.
	8.	I can prepare an outline of an essay and an evaluative commentary.
	9.	I can select and develop ideas and supporting details to write an effective expository essay.
	10.	I can write an evaluative commentary of an argumentative article.
	11.	I can develop a point of view, ideas and supporting details to write an argumentative/discussion type essay.
	12.	I can revise, refine, and edit a draft of an argumentative /a discussion type essay.

The intended learning outcome of Integrated Language Skills III test battery is to enhance the English written skill among diploma students by means of equipping them with necessary writing skills. As this is an integrated English proficiency assessment, reading and speaking skills are integrated with emphasis given on writing skill. Further, Integrated Language Skills III test battery focuses on enabling students to discuss arguments and issues effectively, as well as exploiting a variety of materials

in varied situations in enhancing their abilities to use language. This assessment is also intended to increase students' lexical density by exposing them to the development of higher-level grammatical construction and vocabulary expansion.

Focusing on proficiency, grammar items such as parts of speech, subject-verb agreement, verb tenses, clauses, passive voice, etc are taught incidentally within the writing, reading, and speaking skills. To reiterate, OBA is student-centred and hence, the students are encouraged to be independent learners. Teachers play a role in enhancing student learning, especially to inculcate independent learning by means of student-centred activities, i.e., direct instructions, role-play, discussions, group activities, grammar exercises, authentic reading texts, etc. Hence, students should be encouraged to use the language to present arguments and discuss issues effectively both in writing and speaking.

As can be seen in Table 1.1, the Integrated Language Skills III test battery consists of four assessments and no final exam is included. Two assessments (essay writing and evaluative commentary) focus on assessing the students' writing skill, while the other two assessments assess the students' reading skill (reading test) and speaking skill (oral commentary). The writing assessments accumulate to fifty percent and reading and speaking assessment is another half of the entire assessment.

Taking such a paradigm shift in HLIs among the nations around the world along with Malaysia in recent years, the present study aims to investigate how an OBA at ALS in UiTM has thus far affected the student learning. UiTM has been identified due to one unique feature: it is prevalent in every state around the country in comparison with the rest of the HLIs. Specifically, its language centres are existent in all campuses of UiTM and on practicality grounds, one particular language centre in one of the branch campuses has been identified.

### **1.3 STATEMENT OF PROBLEM**

Alderson & Wall (1993) define washback as the phenomenon of how tests influence teaching and learning, which is restricted to classroom behaviours of teachers and learners (Alderson & Wall, 1993). It is noteworthy that a washback study is partly an evaluation study, in which it evaluates teaching and learning practices. Cheng (2013) contested that the core issue of washback resides in the use (or misuse) of test scores and the values and stakes attached to a test within the society and within the teaching and learning context where a particular test exists. To reiterate, educational reforms have enveloped around the world with the intention to enhance educational quality and hence, it is deemed necessary to evaluate these kinds of decision in terms of its consequences (washback) because the consequences associated with the policy and practice are the criteria for evaluating the success of the policy and practice (Kane, 2012). Moreover, it is also pertinent for testing practices to yield valid data about students' achievement and performance. Scholars like Messick (1989), Cheng (1997), McNamara (1996) and Bachman (2005), among others, deemed that it is vital for washback studies to look into the consequences of a test because this information serves as evidence to the success or failure of any curriculum change. To better understand the relationship that exists between washback and test validity, Booth (2018) urged for more research into washback which is guided by contemporary test validation approach as these studies may provide practical support to methods of investigation in the future. Considering Integrated Language Skills III test battery was introduced to enhance education quality, evidence about washback on students' learning is necessary and crucial part of the test evaluation.

A number of significant washback studies ranging from standardized large-scale tests at international levels carried out in high-stakes contexts are inter alia,