

*International Journal of Learning, Teaching and Educational Research*

*Vol. 20, No. 8, pp. 1-17, August 2021*

<https://doi.org/10.26803/ijlter.20.8.1>

*Received May 31, 2021; Revised Aug 16, 2021; Accepted Aug 30, 2021*

# Washback Impact of the MUET: The Before and After Effect of a High-stake University English Test in Malaysia

**Najihah Mahmud and Nor Hazwani Munirah Lateh**

Universiti Malaysia Kelantan, Malaysia

<https://orcid.org/0000-0001-7938-1499>

<https://orcid.org/0000-0002-6968-0513>

**Nazirah Mahmud**

Universiti Sultan Azlan Shah, Malaysia

<https://orcid.org/0000-0001-6343-6555>

**Ariezal Afzan Hassan, Amaal Fadhlini Mohamed  
and Siti Amirah Ahmad Tarmizi**

Universiti Malaysia Kelantan, Malaysia

<https://orcid.org/0000-0002-9806-0874>

<https://orcid.org/0000-0001-8559-4618>

<https://orcid.org/0000-0002-4932-6007>

**Abstract.** This study seeks to explore the washback effect of the Malaysian University English Test (MUET), a high-stakes compulsory university entry test in the context of Malaysia. As simple and linear as it commonly appears, washback has been found to be far more complex than simply looking at the impact that a test might or might not have on the stakeholders. Therefore, this study aims to fill in this knowledge gap by systematically re-examining the beliefs on washback by investigating the relationship between the students' perceptions of the MUET in terms of its importance and difficulty, with their language learning strategies whilst preparing for the test and after sitting the test. Using a mixed methods approach, a student questionnaire and student interview were utilised to elicit data from 30 male and 46 female students. The students were further divided into two groups, specifically those who were preparing for the MUET and those who had already sat the MUET. The findings suggest that preparing for the MUET encouraged the students to utilise a certain language learning strategy more compared to the others, specifically focusing on language skills that had not been formally tested before. Although the students' perception did not significantly shape the students' course of action when preparing for the test, it still had an impact on their overall view of the whole test-taking matter. This study is

intended to add more insights to the less explored areas of washback, specifically the students' perceptions and washback over time.

**Keywords:** washback; language assessment; high-stake test; language testing; backwash

## 1. Introduction

In the field of education, it is well known that tests, especially high stakes tests, have a huge influence on the teaching and learning process. Specifically, in language education, the influence that the tests exert over teaching and learning is known as 'washback' or 'backwash'. Washback has always been associated with the negative consequences of tests. However, this phenomenon was not empirically investigated until the 1980s. Since then, language education researchers have been looking at washback differently as the findings from the previous research on washback continually reveal how complex it is (Dong, 2020; Khan et al., 2019; Hughes, 2021).

Alderson and Wall (1993), in their pioneering study on washback in Sri Lanka, hypothesised that "Tests that have important consequences will have washback" and conversely "Tests that do not have important consequences will have no washback" (p. 120). In other words, the higher the stakes of a test, the more of an impact it will have on the teaching and learning process. According to Qi (2007), the authorities are always tempted to resort to manipulating high stakes testing in the name of 'curriculum innovation' for an immediate outcome as it is claimed to be a "quick and most cost-effective way to improve education" (p.52). Studies on the washback of high-stakes testing reported that the test could change the students' learning behaviour by motivating them to put more effort into their learning (Cho, 2004; Pan & Newfields, 2012; Thomas, 2005) while also promoting learner autonomy (Pan, 2014; Stecher, 2002) when preparing for the test if the stakes are sufficiently high.

## 2. Literature Review

The connection between (1) testing, and (2) the teaching and learning practices has been commonly explored based on the research in the field (Barrows et al., 2013; Cheng et al., 2011; Gebril & Eid, 2017; Luong-Phan & Effeney, 2015) sparked by the ground-breaking study on washback by Alderson and Wall in the late 1980s. Alderson and Wall (1993), in their significant publication on washback, raised the notion of the complexity of washback and the needs for more in-depth research attempting to not only describe what washback looks like but also to account for what occurs. They also pointed out that although it is widely known that a relationship between testing, teaching and learning practices does exist, the complexity of the washback concept itself makes it difficult to prove how directly tests affect the teaching and learning practices without considering other mediators or variables that may or may not contribute to the change. Similarly, Stoneman (2006) observed that the washback phenomenon has yet to be clearly explained and deeply explored despite the abundant literature on language assessments focusing on the impact of testing on teaching and learning.

In a doctoral study by Mahmud (2018), it was hypothesised that the students' perception of the test has an impact when it comes to determining the students' course of action when preparing for the test as opposed to their actual proficiency level. For instance, even if a student has a weak command of the English language, if the students think that they can perform well in the test, they will be very motivated when preparing for the test and vice versa. In a more recent study undertaken in China involving 3,105 high school students, Dong (2020) found that the students' positive perception of the test increased the positive washback of the test. Test preparation activities may help to increase the students' probability of success but it only works in the short-term. If it is too excessive, it can be counterproductive regarding the students' achievements in terms of their test score. Aside from being a waste of the students' time and energy, Dong (2020) emphasised that it defeats the purpose of learning and in turn, creates negative washback.

The MUET was first introduced in 1999 with the aim (1) of "bridge the gap in English language needs between secondary and tertiary education (Malaysian Examination Council); and (2) to consolidate and enhance the English proficiency of students preparing to enter Malaysian public universities" (Lee, 2004, p. 1). The four language skills of reading, writing, listening and speaking are tested in the MUET. It is designed and administered by the Malaysian Examination Council and it is recognised in Malaysia and Singapore (Othman & Nordin, 2013). This test is significant for pre-degree students as it serves as an indicator of their English language proficiency which enables them to enrol on their desired course (Kaur & Nordin, 2006). This study aimed to investigate the washback effect of a high-stake language test, the Malaysian University English Test (MUET), while preparing for the test and after sitting the test.

This study investigated the Malaysian students' perception of the MUET by exploring its relationship with the language learning strategies that they employed when preparing for the test. In addition, it also explored how long the washback effect of MUET lasted after the students sat the test. The research questions guiding this study were as follows:

1. What are the students' perceptions of the MUET?
2. To what extent do the students' perceptions seem to have a washback impact on the students' language learning strategies?
3. Is there a difference in the washback impact before and after sitting for the MUET?

### **3. Methodology**

The use of both quantitative and qualitative data gathering techniques assisted the researcher in illuminating different aspects of the same issue and providing a more complete picture of the study (Denscombe, 2014). As this study aimed to generate data on the students' perception of the test and their language learning strategies in relation to the washback effect of the MUET over time, an analytical comparison was carried out between the students who were preparing for the MUET and the students who had already sat the test. To provide multiple perspectives and methods, the data was gathered through the collection of the

students' perceptions using a questionnaire and through interviews. Two groups of students were recruited for this study, specifically students who were preparing for the MUET (Group A, n=30) and students who had already sat the MUET and who were currently undertaking a general English language course at a public university in Malaysia (Group B, n=46).

**Table 1: Participants**

Gender	Group A	Group B	Total
Male	11	13	24
Female	19	33	52

Data from Group A was used to describe the washback effect of the MUET whilst data from Group B was used to explore the washback impact over time, also known as the washback length.

For this study, the questionnaire was adapted from established questionnaires from relevant previous studies as detailed in Table 2.

**Table 2: Questionnaire**

Sections	Sources	Cronbach's alpha ( $\alpha$ )
Background information	N/A	N/A
Section 1 : Perception of the Test	Xie & Andrews (2013)	N/A
Section 2 : Perception of the Test Importance	Bodas (2006)	.833
Section 3 : Perception of the Test Consequences	Bodas (2006)	.805
Section 4 : Language Learning Strategies	Oxford (1990) Green (2007)	.890

In general, the questionnaire consisted of two main sections. The first section covered the demographic questions such as gender, field of study, English proficiency level etc. Section two consisted of four sub-sections, dealing with the students' (1) perception of the test (MUET) in general, (2) their perception of the test importance, (3) their perception of the test consequences and (4) Language Learning strategies (Oxford, 1990). All of the variables in section two were assessed using a Likert scale.

This study also utilised the students' individual interviews to gather the qualitative data. As pointed out by Atkins and Wallace (2012), interviews not only allow the researchers to engage with the participants individually, but it also allows them to collect various types of in-depth data, for example, factual data, views and opinions, personal narratives and histories. According to Ary et al. (2013), there is no general rule for determining the number of participants for the purpose of collecting qualitative data. For this study, 12 open-ended questions with several probing questions on the students' experiences of learning English were prepared for the student interviews. The questions were adapted based on

the previous studies on washback (Hsu, 2010; Mahmud 2018; Shih, 2013) focusing on the learners' general perception of their English language learning and the test under investigation, the MUET.

#### 4. Findings

Item analysis of the mean scores of the students' perceptions for all constructs was carried out to better understand how the students responded to each item individually before analysing the items according to their grouping. The student questionnaire consisted of four constructs: (i) the perception of the MUET in general, (ii) the perception of the test importance, (iii) the perception of the test consequences and (iv) the language learning strategies. In the item analysis, the constructs were treated as the dependent variables, with gender as fixed factors, and English language proficiency and experience within the test as covariates. Taking into consideration that the data was not normally distributed, non-parametric analyses, the Mann-Whitney U-test and the Kruskal-Wallis test were employed to analyse the aforementioned constructs both as individual items and as an overall scale. To ensure a greater level of reliability, only significant differences in the mean scores ( $p < .05$ ) were taken into consideration. The findings were arranged according to the sub-sections as follows.

##### 4.1 Students' Perception of the MUET

###### 4.1.1 Perception of the MUET in general

The mean scores for all items were calculated and arranged in descending order as illustrated in Table 3. A Likert scale (comprised 5 for strongly agree, 4 for agree, 3 for undecided, 2 for disagree and 1 for strongly disagree) was used to score the items.

**Table 3: The MUET in general**

Items		Mean	Std. Deviation
1.4 The MUET made me practise my listening skills more than before.	Strongly Agree	4.30	.611
1.12 More speaking activities should be conducted in MUET preparation class.		4.29	.629
1.5 The MUET made me practise my speaking skills more than before.		4.24	.671
1.11 More listening activities should be conducted in MUET preparation class.	Agree	4.16	.590
1.6 The MUET made me practise my reading skills more than before.		4.11	.723
1.13 More writing activities should be conducted in MUET preparation class.		4.11	.759
1.10 More reading activities should be conducted in MUET preparation class.		4.03	.765
1.3 The MUET made me practise my writing skills more than before.	Agree	3.99	.721
1.2 My English language learning was improved by practising MUET past year questions.		3.89	.665
1.9 I think that the MUET preparation class that I took was not very helpful.	Disagree	2.46	1.113

The trend seems to be that the highest-ranking items with mean scores that are within the range of 'Strongly Agree' are those related to the listening and speaking skills. These two skills are seldom tested in other standardised English language test in Malaysia as opposed to reading and writing. This clearly indicates the washback that the MUET had on these students, especially towards the two skills which were not tested before. The mean scores for the set of items regarding the students' increase in effort connected to their reading and writing skills ranges from 3.99 to 4.11, which are ranked very high. Item 1.9 has the lowest mean score (2.46) which falls under the 'Disagree' range. This indicates that most of the students perceived the MUET preparation class that they took as being beneficial to them.

Next, the analysis of the overall scale was carried out. The Mann-Whitney U-test was run and the analysis revealed there to be no significant differences in terms of the students' general perception of the MUET in relation to their (i) gender,  $p = .906$ , (ii) experience with the test,  $p = .074$ , and (iii) proficiency level,  $p = .442$ . The result suggests that the students' general perception of the MUET did not change even after they had already taken the test. The students' proficiency level in English did not affect their general perception of the MUET as well.

#### 4.1.2 Perception of Test Importance

The perception of the test importance construct consisted of five items as shown in Table 4.

**Table 4: Test Importance**

Items		Mean	Std. Deviation
2.3 It is very important for me that I do well in the MUET	Strongly Agree	4.42	.753
2.4 It is very important for my future undertakings that I do well in the MUET		4.38	.816
2.5 Every student who wants to get into the university should pass the MUET	Agree	4.13	.900
2.2 It is very important for my teacher that I do well in the MUET		3.97	1.045
2.1 It is very important for my parents that I do well in the MUET		3.70	.980

The top two items are items relating to the students' priority of their own self, as opposed to external factors like their teachers (item 2.2) and parents (item 2.1), which were ranked as the bottom two.

At the item level, both the Mann-Whitney U-test and the Kruskal-Wallis test showed there to be no significant differences across gender ( $p = .506$ ), English proficiency level ( $p = .619$ ), and experience with the test ( $p = .619$ ) for each item in the perceived test importance construct. This shows that the students of all English proficiency levels regarded the MUET as an important test. This finding confirms the status of the MUET as a high-stakes test. The students were also

asked about their perception of the consequences that they had to face if they failed the MUET, which has been presented in the next section.

#### 4.1.3 Perception of the Consequences of Exam Failure

The perception of the test consequences construct consisted of six items as shown in Table 5. A Likert scale (comprising 5 for extremely to 1 for not at all) was used to score the items.

**Table 5: Consequences of Exam Failure**

Items		Mean	Std.
			Deviation
3.2 How upset would you be for letting your teachers down?	Very	4.14	.948
3.1 How upset would you be for letting your parents down?		3.99	1.052
3.3 Will your chance to get into top universities be affected?		3.93	1.037
3.4 Will your chance to enrol on your desired course be affected?		3.91	1.048
3.6 Will your ability to communicate in the English language affected?	Undecided	3.03	1.376
3.5 Will your motivation to learn the English language be affected?		3.03	1.366

Interestingly, the top two items with the highest mean score for the test consequences construct are the items that are extrinsic in nature, pertaining to living up to others' expectations, namely their teachers (item 3.2) and parents (item 3.1). Following very closely are the items related to the main objective of the MUET, which is where it acts as an entry test into university. The students agreed and were aware that their application to their desired university (item 3.3) and courses (item 3.4) would be affected if they performed poorly on the MUET. However, they appeared to be unsure whether their ability to communicate and their motivation to learn English would be affected if they did poorly in the MUET. Based on the overall ranking, it is safe to assume that the severity of the consequences of exam failure for the MUET is perceived to be quite threatening to the students.

Next, the Mann-Whitney U-test was run to determine whether there were any differences in the perception of the test consequences between the three independent variables of gender ( $p = .699$ ), experience with the test ( $p = .321$ ) and English language proficiency ( $p = .496$ ). The findings revealed that there were differences between the variables, but they were not statistically significant. Similar to the findings for perceived test importance, the students' proficiency level in English did not seem to affect their perception of the test consequences.

The findings thus far show that there were no significant differences reported between the students' perception of the test and the three independent variables. However, at the item level, the Kruskal-Wallis test revealed there to be significant differences across the English language proficiency levels for item 3.6 (Will your ability to communicate in the English language affected?,  $p=.042$ ) with the mean

ranking across the grouping in an ascending manner being as follows: Low (22.83), Average (39.60) and High (47.06). Interestingly, the findings reveal that the students might regard the MUET as one of the main determinants of their English language ability, especially for high proficiency students.

#### 4.2 Relationship between the students' perceptions and washback impact on language learning strategies

In order to explore the washback impact of the MUET on the students' learning, item analysis was carried out to determine the students' pattern of responses based on the frequency of their perceived language learning strategy usage. The mean scores for all items were calculated and arranged in descending order as illustrated in Table 6. A Likert scale (comprising 5 for all the time, 4 for most of the time, 3 for undecided, 2 for sometimes and 1 for never) was used to score the items.

**Table 6: Language Learning Strategies**

Items	Std.	
	Mean	Deviation
4.15 When writing in English, I tried to translate from my language.	4.08	1.043
4.10 When reading in English, I tried to translate it into my language to help me understand.	4.00	1.007
4.6 If I couldn't think of an English word, I used a word or phrase that means the same thing.	3.99	1.026
4.16 I thought about the goals that I wanted to achieve in this English language course.	3.97	1.078
4.18 When I received corrected work from the teacher, I thought about how to improve next time.	3.89	1.078
4.14 To understand unfamiliar English words, I tried to guess their meaning.	3.82	1.186
4.8 I encouraged myself to use English even when I was afraid of making a mistake.	3.82	1.055
4.17 I tried to improve my writing by analysing the work of other writers.	3.71	1.198
4.2 I tried to find better ways of learning English.	3.70	1.178
4.3 I tried to improve my English by asking others to correct my mistakes.	3.66	1.250
4.12 I used new English words in sentences so then I could remember them.	3.54	1.259
4.13 When I learned a grammar rule, I tested myself to make sure that I really knew it.	3.51	1.301
4.1 I memorised English words by saying or writing them several times.	3.09	1.246
4.11 I was NOT sure how to improve my English skills.	2.92	1.393
4.4 I did the MUET practice tests in my free time.	2.83	1.320
4.7 I reviewed my English class notes or textbook in my free time.	2.82	1.230
4.9 I read English without looking up every new word.	2.75	1.406
4.5 I studied extra English outside of my MUET preparation class.	2.67	1.258



The findings revealed that the students reported a moderate-to-high frequency of using language learning strategies to prepare for the MUET. Specifically, the students used meta-cognitive strategies (items 4.2, 4.11, and 4.18), often associated with learning success, at a high frequency. A high frequency of usage was also reported for cognitive strategies related to rote-translation (item 4.15 and 4.10). As depicted in Table 6, in terms of the individual learning strategies for all of the strategy items belonging to the meta-cognitive (planning, organising and evaluating learning) and compensation (to overcome difficulties in communication) categories, the students reported a high frequency of use with a mean score ranging from 3.70 to 3.99. Meanwhile, none of the language learning strategies were reported at a low frequency of use. For the strategies used at moderate frequencies, the findings show that most of the strategies were related to the learning activities that the students did in their free time, for example, item 4.4 (I did MUET practice tests in my free time) and item 4.7 (I reviewed my English class notes or textbook in my free time). Finding opportunities to practice outside of the classroom (item 4.5) was the least used strategy among the students, which is not surprising. This strategy requires self-initiation and courage, which many young learners in a second language context might not possess.

To determine the relationship between the students' perception of the test importance and their language learning strategies, Kendall's tau-b correlation was carried out. There was found to be a weak positive association that was not statistically significant between perceived test importance and language learning strategies as a whole,  $\tau_b = .140$ ,  $p = .089$ , and direct learning strategies,  $\tau_b = .149$ ,  $p = .076$ , and a very weak positive association between perceived test importance and indirect language learning strategies,  $\tau_b = .091$ ,  $p = .283$ . There was no statistically significant association between the students' perceived test importance and their use of language learning strategies.

Next, there was found to be a weak positive association, which was not statistically significant, between the perceived test consequences with language learning strategies,  $\tau_b = .154$ ,  $p = .058$  and direct learning strategies,  $\tau_b = .153$ ,  $p = .065$ , and a very weak positive association between the perceived test consequences and indirect language learning strategies,  $\tau_b = .082$ ,  $p = .326$ . There was no statistically significant association found between the students' perceived test consequences and their use of language learning strategies. Therefore, the null hypothesis cannot be rejected, and the alternative hypothesis cannot be accepted.

The data suggests that the students' perception of test importance and the test consequences did not seem to affect the students' usage of language learning strategies.

#### **4.3 The before and after effect of the MUET**

This section presents the analysis of the individual interviews with two students from Group B, which consisted of students who had already taken the MUET. The two students were labelled R1 (Female, MUET Band 2) and R2 (Male, MUET Band 3) to ensure anonymity. Students R1 and R2 were randomly chosen from a list of the students who volunteered to be interviewed.

When it comes to English language learning, it was evident from the qualitative data that both students seemed to be particularly concerned with grammar:

*"For me it's difficult, in terms of grammar. And also, the usage of the words." (R2)*

*"Even when we enter the university, we still have to learn grammar because it's easy to forget. We have to really focus in the class." (R1)*

This finding can be attributed to the washback effect of the 11 years of formal English language education in Malaysia. This is because grammar comes first in the syllabus before any of the four language skills. Therefore, most Malaysian students are under the impression that before they work on their language skills, they must make sure that their grammar is good.

Another skill that they were concerned with when it comes to English language learning was speaking skills. Both agreed that for them to improve their speaking skills, they had to practice:

*"We have to practice it a lot. We have to make it a habit. For example, when communicating with the teacher, we are supposed to use the English language." (R2)*

*"For speaking skills, we have to always use it. If we want to improve our speaking skills, we should mix around with the Chinese students, that'll help." (R1)*

The interview participants were also asked about their current English language learning experience. When asked to describe how their current English language class is, the following are their descriptions:

*"Normally the teacher would emphasize on the syllabus from the textbook. For example, we were involved in drama production for an event called 'Drama Night', so the teacher asked us to develop our script which took months to finished. Then, after the event, we continued with the syllabus from the textbook." (R1)*

*"The teacher taught us based on the textbook, and if there was any question, we would discuss it together. Then, the teacher also asked us to present in front of the class." (R2)*

Obviously, emphasis was given to the activities in the textbook. However, these particular students prefer something more interactive as opposed to being constrained or restricted by the syllabus. They mentioned the activities that were conducted in their previous semester's English class, which was English I. They seemed to like it and perceived it as helping them to improve their English language skills:

*"I remember last semester's English teacher, for example, when we were learning about grammar, the teacher would make up a story about it or the teacher would create songs about the grammar that we were learning."*

*That way, we could memorize it better. Well actually we memorized the song, but at the same time, we were learning grammar as well.” (R1)*

*“My teacher last semester conducted some sort of game. In my opinion, after we played that game, it has somehow helped improve my English because throughout the game, the teacher made it compulsory for us to use English language, forcing us to speak in English. I think that’s quite helpful.” (R2)*

Moving on to the MUET, the students were asked to recall their experience preparing for the high-stakes language test. Questions pertaining to what they did inside and outside of their classes were asked to see what kind of learning activities and strategies they used, and to determine the intensity of the washback of the MUET. The participants were first asked about their perception of the MUET in general compared to the other English language tests that they had taken before. The following were their responses:

*“For me, MUET is important because it helps us improve especially our speaking skills, it encouraged the students to speak in English more because they definitely did not want to get low marks, so they would try to speak with other people to practice. And then for listening skills, for example my friends, if previously they were mostly listening to Malay songs, but because listening was tested in MUET, they started to listen to English songs more.” (R1)*

R1 described the MUET as having more challenging questions for the writing component and more questions for the reading component. She also mentioned that the MUET encourages students to speak more in English because they do not want to get low marks for the speaking component. Being tested on all four language components forced the students to work on all four skills. Improvising the four skills required different learning strategies as mentioned by R2 below:

*“MUET does not consist of only one test, but it has different papers for listening, speaking, writing, and reading, so for each of these skills, different preparation strategies are required. For example, when I took SPM in the past, it was more on the writing skills only, but for MUET, we must be able to grasp all of the skills, and for speaking, we have to be able to communicate well.” (R2)*

R2 mentioned that the previous English language test that he took mainly focused on writing skills. Since the MUET also tested his speaking and listening skills, he had to apply different learning strategies to prepare himself for it. He stated that his teacher devoted one whole period of the English lesson to teaching each language skill separately:

*“For example, if for this period, the teacher decided to focus only on listening skills, for other period, the teacher will focus only on speaking skills and so on.” (R2)*

He also mentioned the use of textbooks in the MUET preparation class. In R1’s case, her teacher emphasised the listening and speaking activities more in the classroom:

*"The teacher focuses more on the listening and speaking skills. For listening skills, the teacher provided us with MUET text book. The teacher would play the CD in the classroom for us, then the teacher would ask us to answer the questions in the text book. For speaking skills, the teacher divided us into groups of four, because in MUET speaking test, the students are divided in the same way, so the teacher would imitate the real situation in MUET speaking test." (R1)*

Similar to R2's MUET preparation class, R1's teacher also utilised textbooks to prepare the students for the test. The textbook that R1 mentioned is a commercialised MUET preparation textbook sold at 'Popular', one of the main bookstore chains in Malaysia. In her comment, R1 mentioned that for speaking, her teacher mimicked the real MUET situation to familiarise the students with the format. However, her teacher only did this when the date of the MUET drew closer, which was two weeks before the actual test took place:

*"The teacher would act as the examiner. But, during that time, it was just two weeks before the actual test took place. Before that, the teacher just asked us to practice on our own." (R1)*

The researcher also asked the students to comment on the importance of the MUET to them. R1 initially was not aware of the purpose of the MUET or the use of the MUET result, not until after she had taken the test:

*"When I sat for MUET at the matriculation college, I did not even know that MUET result would affect my university admission result. My teacher did not say anything about it. My teacher just mentioned that it is important, that is all." (R1)*

R1 stated that the reason why her teacher did not say anything on the importance of the MUET was because her teacher did not want to pressure her students. Her teacher thought that there would not be any problem for the students when it came performing well in the MUET as most of them managed to score above Band 3 in the mock-MUET practice test.

*"My teacher did not tell us about it because he did not want us to be stressed out. It was also because when we did a pre-test for MUET, he told us that all of us would be able to get Band 3 and above. That was why my teacher did not say anything." (R1)*

For R2, the only thing that he knew about the objective of the MUET was that it was one of the requirements of applying to tertiary education in Malaysia.

*"All I know is MUET is compulsory to gain entry into the university, that's all." (R2)*

R1 specifically described her difficulties when applying for her desired course as most of the degree courses required at least Band 3 results in the MUET:

*"When I was applying for the university, there were a lot of courses that I could not apply because these courses require at least Band 3. It was quite difficult for me. So, I just chose Band 2 courses because I only got Band 2 in MUET." (R1)*

She ended up choosing courses with a set minimum requirement of MUET Band 2, as those were her only options. However, despite getting a low MUET Band, she did not let her MUET Band define her English language proficiency:

*“For speaking skills, the percentage for the overall score is less than reading and writing, hence, those who like to read books, or overly focused when answering the questions, they would be able to perform well. I mean, unless if the percentage for all skills are equal, then maybe MUET can be used to really measure the overall English language ability. For me, no. Imagine those who manage to get Band 4, but are not able to speak fluently, it’s still the same.” (R1)*

She stated that since the MUET has different weightages for the different components, like the reading component contributes the highest percentage to the overall Band, students who have good reading skills might be able to perform better. For her, getting a higher Band in the test with no ability to speak the language well would still be pointless.

## **5. Discussion**

### **5.1 Students’ Perception of the MUET**

As can be seen in the findings of this study, although limited in scope, there is clear evidence of the washback effect of the MUET especially for the listening and speaking skills as indicated by the quantitative data. Although the MUET consists of all four language skills, reading and listening have not been formally tested before in any standardised English language test in Malaysia other than the MUET. Therefore, once these two skills were added to the test, the learners started to pay more attention to those two skills as well. A scrutiny of the qualitative data showed that the students were particularly concerned with their speaking and listening skills as well as opposed to their writing and reading skills. They also mentioned that their teachers focused more on these two skills during the English lessons. As suggested by Nambiar and Ransirini (2012), both the teachers and students tend to focus more on the tasks that they consider to be imperative to the outcome of the test. Different washback effects depend on the perceived task importance. Although the MUET tested all four skills, it is safe to assume that due to the novelty effect, the students and teachers decided to focus more on the listening and speaking skills because they have been dealing with reading and writing skills for the past 11 years of formal education in Malaysia. This was also the case in Dong’s (2020) study where the students were found to rarely engage in communicative learning, for instance speaking, as it was not tested.

The quantitative findings also suggest that the students’ general perceptions about the MUET were not influenced by either their gender or their English proficiency level. Their perceptions of the MUET also did not change even after they had taken the test. Similarly, when it comes to the students’ perception of the importance of the MUET to them, their gender, English proficiency level, and their experience with the MUET did not seem to affect their perception as well. This shows that regardless of their English language proficiency, the students regard the MUET as an important test. This confirms the status of it as a high-stakes test. The findings also revealed that failing the test threaten the students’ chances of enrolling on their desired course and to their intended university. This frustration

was clearly expressed by one of the interview respondents as her choices when applying to tertiary education were limited due to her low MUET score.

However, when it comes to the test consequences, the students who were preparing for the MUET appeared to be more worried about the consequences of failing the test and how it would affect their motivation to learn and their ability to use the English language. This can be attributed to a phenomenon known as test anxiety, as they had yet to sit for the MUET and did not know what to expect. Those who had already sat the MUET seemed less concerned. This might be due to the fact that they have gone already through the whole experience of the MUET and managed to pull through well in the tertiary level as they were currently studying in a university at the time that this study was conducted.

## **5.2 Language Learning Strategies**

One of the objectives of this study was to see if the MUET encourages students to employ more language learning strategies when preparing for it. The quantitative findings revealed that the students reported a moderate-to-high frequency of language learning strategy use to prepare for the MUET. None of the language learning strategies were reported at a low frequency of use. A closer analysis of the quantitative data disclosed that specifically, the students used cognitive strategies related to rote-translation at a high frequency. The qualitative data revealed similar findings in that the students utilised translation techniques to help them learn. Cognitive strategies like rote-translation are forms of direct learning strategies that, according to Pan (2014), are not deep learning strategies that can really help the students acquire the necessary language skills. In his study, he found that most of the students reported to have frequently used traditional language learning activities such as reading textbooks, memorising vocabulary and idioms, and practicing sentence patterns to name a few. A moderate obsession with grammar was also observed in the responses given by the interview respondents in this study. Similar findings were reported by Shih (2013). He found that most of the students in his study seemed to employ more surface strategies in their English language learning process rather than deep strategies. This type of strategy is mostly geared towards scoring on the test and will not benefit the students in the long run.

To see whether the students' perception of the MUET could be associated with their language learning strategy use, correlation analysis was carried out on perceived test importance and perceived test consequences in relation to the language learning strategies. The results suggest that the students' perception of the test importance and test consequences are not statistically significantly associated with the students' usage of language learning strategies.

## **5.3 The before and after effect of the MUET**

As reported by the qualitative data in this study (see section 4.3), it can be seen that the students were made to practice the skills that had not been tested before in their formal education in Malaysia more, which were speaking and listening skills. This is because the MUET tests all four language skills. However, they perceived that their speaking skills had particularly improved since they practiced both at school and in real life. Although the students expressed a

reservation when speaking in English, it was encouraging to see a slight self-realisation regarding the need to speak the language more in order to help them improve their skills.

As a high-stakes test, the MUET results affect the students significantly, as the MUET is one of the requirements for university entry in Malaysia. Certain courses like medicine, engineering and TESL require a slightly higher MUET band compared to others. Not being able to score the minimum MUET band requirement for their desired course would result in the student having to take another course at the university. This explains the strong washback impact that the MUET can impose on the students due to its high stakes. If utilised correctly by the stakeholders, especially the educators, the problems related to encouraging the students to practice their language skills more can be tackled in due time.

#### **5.4 Limitation of the study**

It should be noted that the number of participants in this study was fairly small and that the grouping was not normally distributed. The findings from the analysis should not be taken at face value. More holistic data needs to be collected when attempting to explore a complex phenomena like washback, hence, more student interviews need to be carried out for both groups, not only two Group B students.

#### **5.5 Implications of the study**

It can be seen in this study that the perception of the test could be one of the important factors involved in determining test washback. According to Dong (2020), among the stakeholders, the teacher is determined to be the most important individual affecting the teaching and learning process. Hence, the teacher can promote the proper perspective of the test among the students which in return could help them to promote positive washback and improve the students' performance in the long run. Several studies tapping into washback have indicated that washback changes over time, hence more research should investigate this. The present study attempted to explore a part of the washback length element by comparing the before and after effect of washback. This study is hoped to add more insights to the less explored areas on washback, specifically the students' perceptions and washback over time.

### **6. Conclusion**

It was apparent from the findings that the students' actual proficiency in the English language did not have a washback impact on the students' course of action when preparing for a high-stakes test like the MUET. Although the quantitative data revealed a similar result in relation to the students' perceptions, the qualitative data appeared to reveal a glimpse into the relationship between the students' perceptions and their influence on their selected language learning strategies when preparing for a test. This calls for further exploration regarding the washback impact of a test utilising a much bigger sample with more diverse participants and instruments.

## Acknowledgments

This work was supported by a research grant provided by Universiti Malaysia Kelantan (R/SGJP/A0400/01060A/001/2019/00596).

## 7. References

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*. <https://doi.org/10.1093/applin/14.2.115>
- Ary, D., Jacobs, L., Sorensen, C., & Walker, D. (2013). *Introduction to research in education*: Cengage Learning.
- Atkins, L., & Wallace, S. (2012). *Qualitative research in education*. Sage Publications.
- Barrows, J., Dunn, S., & Lloyd, C. A. (2013). Anxiety, self-efficacy, and college exam grades. *Universal Journal of Educational Research*, 1(3), 204-208.
- Bodas, J. (2006). *Intra-individual and Extra-individual predictors of text anxiety in Indian children: A cross-cultural perspective* (Doctoral dissertation, Virginia Tech).
- Cheng, L., Andrews, S., & Yu, Y. (2011). Impact and consequences of school-based assessment (SBA): Students' and parents' views of SBA in Hong Kong. *Language Testing*, 28(2), 221-249.
- Cho, D. (2004). Use of standardized tests as university graduation requirement. *English Teaching*, 59(1), 251-266.
- Denscombe, M. (2014). *The Good Research Guide: For Small-scale Social Research Projects* (5<sup>th</sup> Edition). UK: McGraw-Hill Education.
- Dong, M. (2020). Structural relationship between learners' perceptions of a test, learning practices, and learning outcomes: A study on the washback mechanism of a high-stakes test. *Studies in Educational Evaluation*, 64, 100824. <https://doi.org/10.1016/j.stueduc.2019.100824>
- Gebril, A., & Eid, M. (2017). Test Preparation Beliefs and Practices in a High-Stakes Context: A Teacher's Perspective. *Language Assessment Quarterly*, 14(4), 360-379. <https://doi.org/10.1080/15434303.2017.1353607>
- Green, A. (2007). *IELTS washback in context: Preparation for academic writing in higher education* (Vol. 25). Cambridge University Press.
- Hsu, H. F. (2010). *The impact of implementing English proficiency tests as a graduation requirement at Taiwanese universities of technology* (Doctoral Dissertation, University of York).
- Hughes, L. (2021). Washback and the assessment practices of ESL instructors at Japanese university. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, 5(1), 1-9. <https://doi.org/10.30743/ll.v5i1.3238>
- Kaur, N., & Nordin, R. (2006). A case for reconstruction of the pedagogy of the Malaysian University English test (MUET) through thematic units instruction. *Journal of Institutional Research South East Asia*, 4(1), 5-16.
- Khan, A. B. M. A., Aziz, M. S. A., & Stapa, S. H. (2019). Examining the factors mediating the intended washback of the English language school-based assessment: Pre-service ESL teachers' accounts. *Pertanika Journal of Social Sciences & Humanities*, 27(1), 51-68.
- Lee, K. S. (2004). Exploring the connection between the testing of reading and literacy: The case of the MUET. *GEMA Online Journal of Language Studies*, 4(1), 41-50.
- Luong-Phan, N. H., & Effeney, G. (2015). TOEFL iBT and language learning motivation: An investigation into teaching styles and influential factors for Vietnamese adolescents. *International Journal of Research Studies in Language Learning*, 4(3), 3-18.
- Mahmud, N. (2018). *Investigating the Washback Effect of the MUET as a University Entry Test on Students in Malaysia* (Doctoral dissertation, University of York).



- Nambiar, M., & Ransirini, S. (2012). Teaching MUET, not English: A study of the washback effect of the Malaysian University English Test (MUET). *English in Multicultural Malaysia. Pedagogy and Applied Research*, 49-62.
- Othman, J., & Nordin, A. B. (2013). MUET as a predictor of academic achievement in ESL teacher education. *GEMA Online Journal of Language Studies*, 13(1), 99-111.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle Publishers.
- Pan, Y. C. (2014). Learner washback variability in standardized exit tests. *TESL-EJ: Teaching English as a Second or Foreign Language*, 18(2), 1-30.
- Pan, Y. C., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching*, 9(1), 108-122.
- Qi, L. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education: Principles, Policy & Practice*, 14(1), 51-74. <https://doi.org/10.1080/09695940701272856>
- Shih, P. C. (2013). *The English benchmark policy for graduation: An investigation of perception, motivation, and approaches to learning at a university of technology in Central Taiwan* (Doctoral dissertation, Durham University).
- Stecher, B. M. (2002). Consequences of large-scale, high stakes testing on school and classroom practice. *Making Sense of Test-Based Accountability in Education*, 79-100.
- Stoneman, B. W. H. (2006). *The impact of an exit English test on Hong Kong undergraduates: a study investigating the effects of test status on students' test preparation behaviours* (Doctoral dissertation, The Hong Kong Polytechnic University).
- Thomas, R. M. (2005). *High-stakes testing: Coping with collateral damage*. Routledge.
- Xie, Q., & Andrews, S. (2013). Do test design and uses influence test preparation? Testing a model of washback with Structural Equation Modeling. *Language Testing*, 30(1), 49-70.