

Washback Effects of School-Based Assessment (SBA) on Malaysian Secondary School Students' English Language Learning

Nur Farisya Amylia Mohd Salleh^a, Vahid Nimehchisalem^{a*}, Ilyana Jalaluddin^a and
Jayakaran Mukundan^b

^a Universiti Putra Malaysia, Serdang, Malaysia

^b School of Education, Taylor's University, Subang Jaya, Malaysia

*Corresponding author: vahid@upm.edu.my; nimechie@gmail.com

Article information

Abstract

This paper reports on the washback effects of Malaysia's School-Based Assessment (SBA) on Form 4 students after almost a decade of its implementation in the country. It critically examines how SBA has affected Form 4 students' overall perceptions about the learning of the English language at school. This study also investigates the challenges that the students are facing with SBA in their school. In this sequential explanatory study, the data were collected via three methods: questionnaires, interviews, and classroom observations. It was found that there was a presence of positive washback in the students' perceptions of and attitudes towards SBA. Themes, such as engaging activities and constructive feedback, emerged, which showed the students' positive perceptions of SBA. However, the major challenge that the students were facing was with regard to the availability of the right resources at school for SBA. The findings of this study contribute to a better understanding of the status quo of the phenomenon of SBA and its impact on language learning and teaching in the context of secondary schools in Malaysia.

Keywords	washback, School-Based Assessment, perceptions, Form 4 students, sequential explanatory
APA citation:	Mohd Salleh, N. F. A., Nimehchisalem, V., Jalaluddin, I., & Mukundan, J. (2023). Washback effects of school-based assessment (SBA) on Malaysian secondary school students' English language learning [Special issue]. <i>PASAA</i> , <i>66</i> , 168–201.

1. Introduction

Washback has been defined as the influences that language assessment has on curriculum design, teaching practices, and learning behavior as well as the learners, teachers, and stakeholders' choices in the process of language learning (Mckinley & Thompson, 2018). For instance, teachers may teach about a test and students might focus on the aspects of language learning that are likely to be assessed in the future. Some scholars, such as Cohen (1984), narrow the term 'washback' on teachers and learners in classroom settings, while others, such as Alderson and Wall (1993) view it as the influences of tests on educational systems and even on society in general (Bailey, 1996). Jim (2018) and Dong (2021) point out that 'washback' can be defined as the influence that language testing has on curriculum design, teaching practices, and learning behavior.

Washback effect in language teaching and learning could either be positive or negative to the extent that it can either support or hinder the accomplishment of educational goals (Mckinley & Thompson, 2018). These positive and negative types of washback would depend on whether assessment has beneficial or harmful impacts on educational practice (Hughes, 1989). A positive washback is present if the test or assessment is seen to encourage the positive teaching-learning process (Pearson, 1991). Alternatively, negative washback is present if learners improve by studying for that particular assessment only, and the scores achieved do not reflect students' actual progress in learning the English language (Wall, 1999). In fact, washback plays a pivotal role on students' language development. For instance, a positive washback of assessment could encourage students to study more or may promote awareness of the connections between standards and instruction. A negative washback, on the other hand, could affect students' motivation and lack of confidence to succeed (Pan & Newfield, 2011). There has been a constant debate on the effect of washback in language testing and assessment. This area of study is seen as a complex entity that warrants considerably more research to be conducted in the future (Ha, 2019).

The phenomenon of washback is not a new area in the field of language assessment. This area has, in fact, become one of the significant areas among researchers in this field for the past 30 years (Cheng et al., 2015; Alqahtani, 2021). The awareness on the importance of test consequences have existed ever since the emergence of modern testing (Alderson & Wall, 1993). Due to the increasing number of large-scale and high-stake testing worldwide, there has been a growing focus on the impact of the test on its stakeholders, the educational system, and society in general (Cheng et al., 2015; Dawadi, 2021). As a result, there has been a growing number of studies in language education that examine the intended and unintended relationships between testing, teaching, and learning (Majid, 2011; Liu & Yu, 2021). Given the range and extent of testing consequences reported worldwide, it is crucial that testing practices yield well-grounded data about students' achievement and performances throughout the years.

A number of researchers emphasized the importance of washback effect to be investigated from time to time. Shohamy et al. (1996), for instance, have discovered that washback varies over time, owing to many factors such as the status of the language and the uses of the test. In Malaysia, Alla et al. (2016) studied the washback effect of School-Based Assessment (SBA) on lower secondary school students. They revealed that students were unclear about the requirements and the potential benefits of the newly introduced SBA upon a year of its implementation. Hence, this paper aimed to explore secondary students' perceptions of SBA, as well as the factors that promoted or hindered this washback effect after a decade of its implementation in Malaysia.

1.1 Research Questions

The following research questions were addressed in this study:

1. What are the Malaysian secondary school students' perceptions on the SBA of English language learning?

2. What are the student-centered English learning activities practiced inside the classroom during the current implementation of SBA in Malaysian secondary schools?
3. What are the students' English activities outside the classroom during the current implementation of SBA?
4. What are the main challenges that students face in SBA?

2. Literature Review

According to Cheng and Curtis (2004), washback is bidirectional, either positive or negative, and could be affected by several contextual factors. Hence, in order to examine washback of a specific test or assessment, a good knowledge of the educational context is necessary.

Traditionally in Malaysia, students undergo 11 years of formal education system (school) and sit for three important examinations which are (i) the Primary School Achievement Test (*Ujian Penilaian Sekolah Rendah, UPSR*), carried out at the end of Standard 6; (ii) the Form 3 Assessment (*Pentaksiran Tingkatan 3, PT3*); and (iii) the Malaysian Certificate of Education (*Sijil Pelajaran Malaysia, SPM*) at the end of Form 5, the equivalent of the GCE O-levels (Alla et al., 2016). In the last few decades, there has been a surge of interest in the field of the quality of education that Malaysia has to offer (Woo, 2019). Recent events have reported that despite going to school for an average of 12 years, the system has failed in developing basic skills such as reading, mathematics, and science in school students. Malaysia's education system has been regarded as 'unnaturally low in quality' according to OECD cross-country surveys based on the scores of primary and secondary school students in these basic skills (Woo, 2019).

There have been changes of late in the Malaysian education policy. This could be reflected by transformations in the content, format, and structure of the public examinations and assessment delivery procedures. After more than six

decades of independence, there had been a paradigm shift in the area of assessment within the Malaysian Education System, where SBA became more dominant starting in 2012 (Ministry of Education, Malaysia, 2013).

Over the years before the implementation of SBA, assessment in Malaysian schools was rather centralized and summative in nature (Fook & Sidhu, 2006). The Malaysian Examination Syndicate was the body that prepared tests, administered them, and marked exam papers. However, after years of its implementation, the Ministry of Education (MoE) stipulated that the exams were impractical and taxing on students (Joachim & Hashim, 2021). According to Hashim et al. (2013), the previous exam-oriented assessment failed to portray students' true competence. Correspondingly, Stiggins (2005) supported this notion as he mentioned that placing high importance on public examination as the core assessment in education could negatively affect students' emotional strength which could be an obstacle for their success in their learning. Therefore, with the 21st century of learning and globalization, it is crucial for students to be able to receive relevant knowledge and skills in order to learn effectively (Joachim & Hashim, 2021). The skills that are required for learners in 21st century include critical thinking, creativity, meta-cognition, communication, digital and technological literacy, civic responsibility, and global awareness (Kim et al., 2019). Therefore, the central idea for assessment now is learner-orientedness, and its primary objective is to promote learning.

The term SBA is revolutionary and was introduced by the Ministry of Education of Malaysia in 2012 (Majid, 2011; Wilson & Narasuman, 2020). Policy makers believed the Malaysian education system failed to produce students who could compete at international levels (Hunter, 2022). According to international assessment systems, such as The Program for International Student Assessment (PISA), Malaysia has been in the bottom third of the group (PISA, 2018).

In Malaysia, SBA was introduced in order to replace the centralized examination and to promote a combination of SBA and centralized examination (Talib et al., 2014). Chew and Muhamad (2017) explained that the assessments in SBA should be designed, built, managed, checked, recorded, and reported by the assigned school teachers. In contrast with previous assessment system, SBA is supposedly empowering teachers in assessing their students holistically. Because there is no urgency towards preparing students for exit examinations, most people believe that such formative features of assessment like diagnostics and feedback would feature widely. It is also believed that activity-based learning-teaching with its emphasis on learner-centeredness will be given emphasis. This would then lead to change in classroom dynamics where chalk and talk, repetitive drills, and teacher-controlled teaching are replaced with peer work, group work, and project-based learning (PBL) where the teacher acts as a facilitator rather than a dictator.

SBA requires teachers to monitor students' learning progress, to evaluate students' performance, to record students' progress, and to provide feedback from time to time (Talib et al., 2014). In contrast to the former national and standardized examination, SBA provides greater diagnostic information on the students' performance and actual learning (Jamal et al., 2015). Consequently, with SBA, students could be assessed both formatively and summatively (Handbook for the Implementations of School-Based Assessment in Malaysia, 2014). It is widely believed that there will be activities outside classrooms (widely known in the literature as out-of-class English (OOCE) (Mukundan, 2016). This is due to the introduction of such long-term learning-teaching episodes like PBL which usually takes place during co-curricular hours and in informal settings outside classrooms (Mukundan, 2014).

There are several papers that have attempted to investigate washback. However, most papers look into the aspect of washback in teaching and learning in quantitative ways, while its qualitative aspects remain limited (Dong, 2020).

There are several researchers that explore the relationships between the variables quantitatively, particularly through Structural Equation Model (SEM) (Xie & Andrews, 2013; Dong, 2020). However, a quantitative study is insufficient in order to have a deeper understanding of the mechanism of washback. It would require the methodological triangulation of data which consists of questionnaires, interviews, and classroom observations (Alla et al., 2016; Alla, 2021).

A few studies have also reported that stakeholders' behaviors affect their products and outcomes (Xie & Andrews, 2013). Nevertheless, due to separate studies conducted by most of the researchers, the results do not clearly reflect the mechanism of washback holistically. According to Dong (2020), few empirical studies have examined the relationship among stakeholder's perceptions, their learning behaviors, and learning outcomes within a triangulated model.

Several studies have centralized on teachers' perception pertaining to the washback effect of SBA in Malaysia (Azid, 2022; Rahman et al., 2021; Ghazali, 2016; Majid, 2011). Despite students being the ultimate stakeholders for the assessment, students have been investigated less frequently than other stakeholders in the area of assessment even though they are the direct stakeholders in any form of assessment (Hamp-Lyons, 1997; Cheng et al., 2015; Xie & Andrews, 2013; Alla et al., 2016; Dong, 2020). However, in recent years, researchers have shown greater interest in learners and learning, and washback studies on learners and learning have proliferated (Damankesh & Babaii, 2015). Considering that students are the primary stakeholders of the assessment, studies with regard to their perceptions and attitudes are essential. A study by Chin et al. (2019) reported that the transformation of the Malaysian national assessments was driven by the decline in school students' performances in international-large scale assessments. The findings of this research inferred the necessity of refining the monitoring and coaching system to ensure that the student assessment implementation does not diverge from the missions of education policy reform

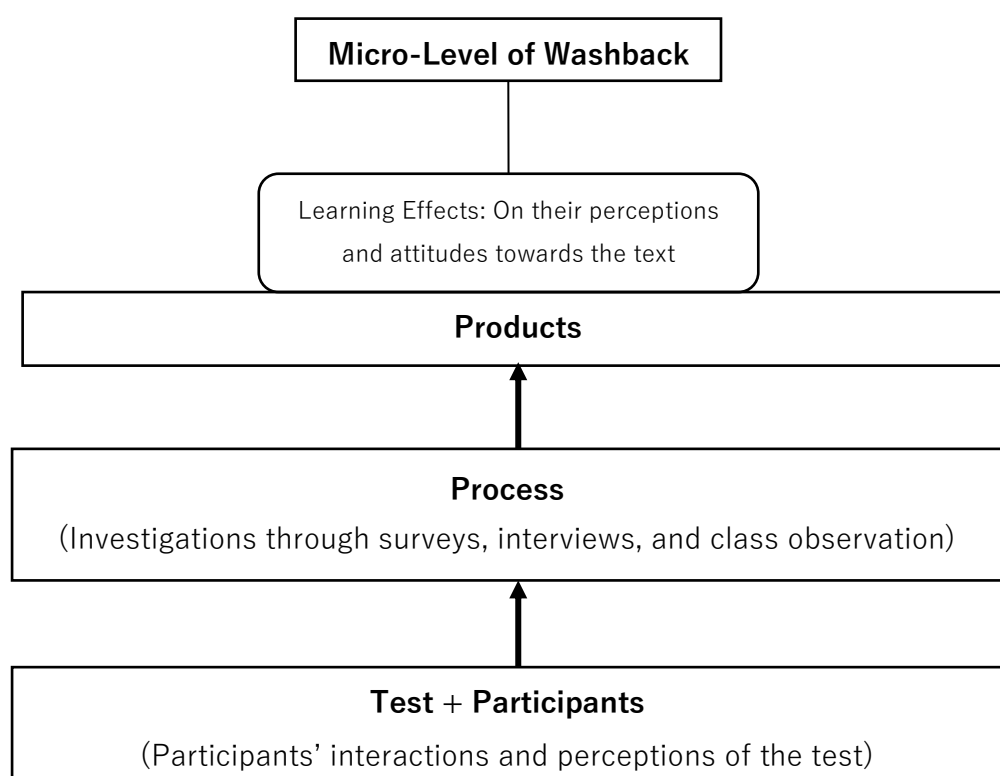
(Chin et al., 2019). Consequently, this washback study is vital to ensure that the implementation of the system is parallel to the designated education policy.

2.1 Conceptual Framework

In order to have a bigger picture of the mechanism of washback among learners, or in this case, the stakeholders, this study explored Pan's (2008) holistic model of washback instead of Hughes' (1993). Pan's (2008) model of washback was initially based on Hughes's (1993) and Bachman and Palmer's (1996) ideas. Figure 1 shows the washback mechanism that has been proposed by Pan (2008):

Figure 1

Pan's (2008) Holistic Model of Washback



The common characteristic in washback models is that they tend to highlight what washback looks like and who is affected, but do little to address the factors that it contributes to the phenomenon (Pan, 2008). In other words, less

focus is given to 'process' than 'participants' and 'products' in these models of washback.

Therefore, Pan's (2008) model of washback would reflect a holistic balance of both micro and macro levels. Washback in micro level illustrates the effects of teaching, learning, score gain, and material selection, whereas, in macro level, this model dives into the effects of innovation and social sequences.

3. Methodology

This study generally adapted to a mixed mode, specifically sequential explanatory design of the study. In order to dissect the nature of the current affairs in students' perception of SBA, this study adopted a cross-sectional design, with the aim to find out the prevalence of perceptions by taking a cross-section of the selected population. Parallel to a case study conducted by Yu (2010) on the washback effect, a statistical questionnaire design was adopted to obtain information about the participants' attitudes, practices, and concerns. In order to complement and further verify those findings obtained in the questionnaire, class observations and interview sessions were conducted.

3.1 Participants

The participants of this study were Form 4 students (upper secondary), 16 years of age, and experienced the Malaysian Education curriculum. The samples were from a residential school in Selangor. The sampling method was purposive sampling. This paper could deliver a different perspective of washback by using a different type of schools in Malaysia. Hence, a residential school was selected instead of a regular public school. This school was listed as the one of the best schools (top 25) residential schools in Selangor. A total of 142 students participated in this study and this was the total number of Form 4 students in this school.

3.2 Ethical Matters

The study was reviewed by The Ethics Committee for Research Involving Human Subjects Universiti Putra Malaysia. We started data collection after the ethical clearance (Reference number: JKEUPM-2019-360) had been granted.

The school and students were assured that their names would remain anonymous, and the data would be used for research purposes only. In our report, we used pseudonyms rather than the participants' actual names to ensure anonymity. Participation in this study was voluntary, and participants were informed that they had the freedom to withdraw from the study whenever they wished.

3.3 Instruments and Data Collection Methods

The data were collected using a questionnaire, observations, and interviews. More details on these instruments and data collection methods are presented in the following sections.

3.3.1 Questionnaire

Both quantitative and qualitative data were collected using a questionnaire adapted from Annie (2011). The questionnaire elicited data related to the students' perceptions and attitudes towards SBA and how it impacted their learning of the English language. Modifications were made to the questionnaire considering the demographic context of this study. The questionnaire was then administered physically by the first author. The face-to-face administration of the questionnaire allowed the researcher to aid the participants whenever further clarifications and explanations were required on any of the items. The questionnaire consisted of five sections and 56 items. All the items consisted of a 6-point Likert scale. The first section (items i-iv) elicited data on participants' demographic background. The second section focused on the students' views about SBA, and more

specifically on the students' understanding of SBA (items 1-2), fairness of SBA (items 3-8), motivation to learn English (items 9-11), and autonomous learning promoted by SBA (items 12-22). The third section investigated student-centered activities inside the classroom (items 24-36) as well as out-of-class activities (items 37-44). The fourth section (items 45-48) covered the challenges of implementing SBA. The final section (items 49-56) looked into the students' views about External Examinations and SBA.

3.3.2 Observation

We conducted class observations until we reached the saturation point; that is, we could confirm that the analysis of our data would no longer result in new themes. In qualitative research, theoretical saturation refers to the gathering of more data about a theoretical construct until it reveals no new themes will emerge if we collect and analyze more data (Hennink & Kaiser, 2022). The total number of hours for class observations for this research was 72 hours, (six hours per week, three days a week). There were five classes all together, and random sampling was conducted for each observational session. The non-participant observation was selected to level out researcher biases in another method (participant observation) and investigate the differences between what people said and what they actually practiced (Salmon, 2015).

3.3.3 Interview

In order to have an in-depth understanding of how SBA influences students' perceptions and learning activities, structured interview sessions were arranged individually at different times, after school hours. Interviews are crucial as they allow researchers to analyze the actual attitude and perception of the participants on the phenomenon under study (Mwita, 2022). As for participant selection for the interview session ($n = 10$), criterion sampling was selected as it was more conclusive to complement

the data obtained from the survey. This sampling method was also recommended by Yu (2010) to gain a deeper insight into the perspectives of the quantitative data. This interview session would enable the researchers to conduct a follow-up session on the response delivered in the questionnaire survey. Based on the data obtained in the first pilot study, the responses from the interview would 'seem' saturated as the students delivered similar feedback/comments due to a lack of skills in their speaking component of the English language. Therefore, criterion sampling was selected to ensure that the students' response was delivered effectively. The sampling criteria for interview were as follows:

1. students who obtained a minimum of B1 of the CEFR level/an 'A' for their speaking component of their PT3 result; and
2. students who selected "somewhat agree," "agree," and "strongly agree" on the items that inquired on the sufficiency of resources on the implement of SBA and on the validity of SBA as an assessment (items 49 and 50 of the questionnaire).

3.4 Data Analysis

The data retrieved from the questionnaire survey were analyzed using Software Package for Social Science (SPSS) 25. The data were then reported into the mode, median, mean score, and standard deviation for each item of the questionnaire. In order to interpret the mean scores, this study adopted Pimentel (2019) 6-point Likert scale to classify the respective interpretation. The following demarcation points were followed in the interpretations of the mean scores: mean values of 1.00-1.82 signified a 'very bad' level, 1.83-2.65, a 'rather bad' level, 2.66-3.48, a 'bad' level, 3.49-4.31, a 'slightly good' level, 4.32-5.14, a 'good' level, and 5.15-6.00, a 'very good' level.

On the other hand, for the interview session and class observation, the data recorded manually by the researchers were analyzed using NVivo (Version 11).

Repeating ideas led us to emerging themes by analyzing the participants' data. We continued the circular collection and analysis of the data until our data reached saturation period (Hennink et al., 2017; Saunders et al., 2018). Thematic analysis of data involved categorizing it into nodes and sub-nodes to gain a deeper analysis on the questionnaire. These emerging nodes and sub-nodes were then listed as the themes and sub-themes in order to address the research questions. Also, the data from class observation sessions were imported to NVivo 11 and further categorized into their respective themes. The data that had been analyzed was reported and triangulated according to the research questions.

4. Results and Discussion

This section reports the results following the order of the research questions.

4.1 Demographics

Out of a total of 142 students who participated in this study, 68 (47.9%) were male, while 74 (52.1%) were female. The participants were all 16 years of age. Most of the participants were exposed to SBA in Year 4 (2013) when the school assessment was implemented in primary schools in Malaysia. These participants had also experienced SBA in Forms 1- 4. In total, 78 students used Malay (f = 78, 55%) as their medium of interaction at home, while 58 students (41%) used English, and the rest used Tamil (f = 3, 2%), or other languages (f = 3, 2%).

4.2 Perceptions on SBA

As Table 1 indicates, the survey data revealed that the students had generally positive perceptions of SBA. Approximately, two in three students agreed or somewhat agreed that they understood SBA functions, procedures, and requirements. Additionally, the majority of the students (about 70%) agreed or strongly agreed that SBA was fair, created learning opportunities, motivated them, and made them autonomous learners. These results confirmed the findings of our previous study on an international school in Malaysia where the stakeholders

showed awareness of the importance of continuous and formative assessment (Nimehchisalem et al., 2023). As the results in Table 1 indicate, the shift to SBA seemed to be successful; however, the students who rated the domains negatively should not be ignored.

Table 1

Students' Perceptions of SBA

Domains	Frequency (%)					
	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
1. Understanding of SBA functions, procedures, and requirements	3 (2.0)	6 (4.3)	21 (15.1)	48 (33.7)	48 (33.5)	16 (11.4)
2. Fairness of SBA	1 (0.7)	3 (2.1)	12 (8.5)	27 (19.0)	50 (35.0)	49 (34.5)
3. Learning opportunities created by SBA	1 (0.4)	1 (0.6)	8 (5.3)	19 (13.6)	57 (40.3)	53 (36.8)
4. Motivation to learn English	4 (2.5)	5 (3.5)	12 (8.5)	33 (23.1)	53 (37.7)	35 (24.7)
5. Autonomous learning promoted by SBA	1 (0.7)	2 (1.7)	7 (4.8)	23 (16.5)	54 (37.7)	55 (38.6)

Table 1 summarizes the results for the first research question that focused on the students' perceptions of SBA.

These results were also confirmed by the interview data, which revealed that the students had positive perceptions of SBA. When we asked the students about the reason for their positive perceptions, the theme 'feedback' frequently emerged. The students found the constructive feedback from their teacher and peers very useful in developing their language skills. For example, according to Akmal (this and the following names are all pseudonyms):

For the writing part, teachers give comments and write what I need to improve. So, I think, yes. It does help me to know my strength and weakness in English...I think if I'm comparing to my old self, when I was like 12 or 13, I think especially my writing skills, they've improved a lot because the teacher has helped me in giving feedbacks. [Akmal, male, aged 16]

Emmanuel and Areena had similar ideas:

In each class, my English teacher always gives feedbacks on our performance throughout the whole class or even performance throughout the assignment. Her feedback helps me to know my knowledge, weaknesses, and strengths in the English language. (Emmanuel, male, aged 16)

Personally, I believe that I am really good at speaking, but when I write, sometimes I make mistakes and my teachers will correct me on-the-spot, and they will give me verbal feedback, which would definitely help me in improving my English skills. [Areena, female, aged 16]

Finally, the data from class observations also supported what participants had said about immediate feedback for most of the activities and tasks:

- The teacher listed out common mistakes (writing and reading) conducted by the students and ways to improve on them.
- The teacher provided feedback and comments based on the rubric marks of C (content), O (organization), L (language), and CA (communicative accuracy) in their essay writing (directed writing).
- The teacher provided immediate corrections on the students' pronunciations during speaking activities.

Observation data also provided evidence of positive perceptions and attitudes of students with regard to the implementation of SBA when they commented on interesting activities provided in class as well as constructive feedback provided by the teacher and their peers in class. Regarding English language skills, SBA motivated and improved students' English language skills particularly in writing and speaking. According to research by Burgers et al. (2015), there is a correlation between positive and negative feedback and motivation. Negative feedback encourages learners to work on short-term performances, while positive feedback is powerful in fostering long-term motivation. In other words, positive feedback could boost intrinsic motivation among learners.

4.3 Learning Activities inside the Classroom

The second research question dealt with the student-centered English learning activities practiced inside the classroom during the current implementation of SBA (Table 2).

Table 2

Student-Centered Learning Activities inside the Classroom

Domains	Frequency (%)					
	Never	Very Rarely	Seldom	Sometimes	Often	Always/ All the time
1. Group discussions	0 (.1)	2 (1.4)	4 (2.8)	23 (16.2)	54 (38.0)	59 (41.5)
2. Individual presentations	6 (4.2)	10 (7.0)	14 (9.9)	23 (16.2)	39 (27.5)	50 (35.2)
3. Constructive feedback in language skills	2 (1.2)	2 (1.6)	6 (4.1)	19 (13.6)	43 (30.5)	70 (49.1)
4. Self-assessments and peer-assessments	1 (.4)	4 (2.5)	6 (4.2)	42 (29.3)	54 (37.7)	35 (24.5)
5. Follow-up activities (post-activities)	4 (2.8)	17 (12.0)	31 (21.8)	40 (28.2)	35 (24.6)	15 (10.6)

According to Table 2, most participants (about 60-80%) agreed that student-centered learning activities were ‘often’ or ‘always’ practiced during the implementation of SBA. This means that group discussions, self-assessment, peer feedback, and teacher feedback now constitute an important part of language assessment at schools in Malaysia. Even though these results are promising, the students’ relatively lower ratings of ‘follow-up activities’ (Domain 5) seem alarming. A possible explanation for the remarkably high number of students who responded ‘never’ to ‘seldom’ (about one in three students) to this domain could be that teachers are in a rush to cover the syllabus on time. They are in such a hurry that they forget to create a feedback-feed forward loop; that is, creating a new learning opportunity with every round of their feedback by showing their students what they expect from them to do to be able to take their work to the next level (Farhady, 2021).

These results were further confirmed by the interview data. According to the participants, with the implementation of SBA, the activities that were prepared and organized by the teachers were engaging and motivated them to learn. For example, Iffah described:

I do think they [the activities] motivate me to learn English because personally I think the work gets more international. [Iffah, female, aged 16]

These activities were part of their SBA which would further be graded in their report card, as Jazmin pointed out:

For speaking skills, my teachers like to pair us up to put us in groups. She will then assign a task to us that will encourage us to speak in English with each other. We will then exchange our opinions about

the topic during the time given to us...I was in the Research and Development Competition. This competition required us to do a video presentation and the video presentation had to be in English. [Jazmin, female, aged 16]

Marhana also confirmed that she commonly participated in activities related to SBA and enjoyed doing so:

...mostly class activities such as presentations and group discussions and those sorts. I do enjoy all of those as well. [Marhana, female, aged 16]

In addition our class observations indicated that group discussions were held at least twice weekly (Mondays and Wednesdays). These were the days when speaking activities were held for 30 minutes in class (1 period). During these group discussions, the students were in groups of 4-5. Then, they shared their own experiences or opinions on the provided topics. During these group discussions, the students shared their own experiences or opinions on the provided topics in groups of 4-5.

These results were generally promising. However, teachers seemed to be spending relatively little time on follow-up activities—those that could provide further learning opportunities for slow learners while aiding quick learners to overlearn. A possible explanation for the lower level of follow-up activities could be that due to their heavy workload, large classes, and heavy materials with too many items to be covered in their syllabuses, teachers may have been unable to find sufficient time to do such activities. Indeed, English language teachers at Malaysian schools complain about the excessive changes that the Ministry of Education consistently makes to the educational system and materials

(Nimehchisalem et al., 2023). Such changes leave little time for teachers to familiarize themselves with the new rules, systems, and/or materials.

4.4 Learning Activities outside the Classroom

The third research question focused on the students' English learning activities outside the classroom during the current implementation of SBA. Table 3 presents the results of this research question.

Table 3

Students' Out-of-Class Activities

Domains	Frequency (%)					
	Never	Very Rarely	Seldom	Sometimes	Often	Always/ All the time
1. Out-of-class reading	8 (5.8)	24 (16.6)	22 (15.5)	42 (29.8)	29 (20.4)	17 (11.7)
2. Out-of-class listening	3 (2.1)	6 (4.2)	8 (5.9)	20 (13.8)	46 (32.1)	59 (41.8)
3. Out-of-class writing	36 (25.4)	36 (25.4)	18 (12.4)	23 (16.2)	16 (11.0)	13 (9.1)
4. Out-of-class speaking	9 (6.6)	18 (12.9)	25 (17.4)	44 (31.0)	29 (20.4)	17 (11.7)
5. English extra-curricular activities	4 (2.8)	17 (12.0)	31 (21.8)	40 (28.2)	35 (24.6)	15 (10.6)
6. Use of school resources outside class hours	32 (22.5)	44 (31.0)	34 (23.9)	19 (13.5)	7 (4.9)	6 (4.2)

According to the results of the third research question in Table 3, during the implementation of SBA, excluding the 'out-of-class listening' activities (Domain 2), the rest of the domains turned out to be practiced not often. Among these, 'use of school resources outside class hours' was 'never' (22.5%), 'very rarely' (31%), or 'seldom' (29.9%) practiced. Likewise, close to two-thirds of the students (63.2%) 'never' to 'seldom' performed 'out-of-class writing.'

The data from the interview session revealed that some of the activities were recommended by the teachers before the holidays in order to improve their English language skills. A further class discussion was then held as a follow-up to this activity. As the students were enrolled in a residential school, the activities of ‘outside classroom’ could only be conducted over the holidays as they had no access to their devices during school days. The excerpts below from the interview sessions reflect some of the out-of-class activities commonly performed by our participants:

- I read English books, and I communicate with others, including my family in English. Sometimes I even communicate with my friends in English. I can say that I use English excessively in my life especially after PBS (SBA) and its after-school activities. English is also one of my favorite subjects in school. [Muhsin, male, aged 16]

- I listen to English songs. The teacher would talk about them in class as well. I also watch a couple of Netflix series. My favorite is “Stranger Things.” Not from the teacher’s recommendation though. [Marhana, Female, aged 16]

- Most of the things that I do in my daily life are mostly in English. I watch English movies, I listen to English songs, and I read English books. And you know, usually full with English. Based on PBS (SBA), I remember that the teacher would sometimes gave us some list of songs for us to listen to over the holidays. [Emmanuel, male, aged 16]

Previous studies have indicated the benefits of out-of-class activities. For examples, Mukundan (2016) and Mukundan and Sha (2015) investigated the use of PBL in out-of-class activities in fully residential schools (MARA Junior Science Colleges–MRSMS). As they reported, their participants found utilizing prep hours which were plentiful in these settings for PBL in out-of-class activities extremely

beneficial. Students, together with their facilitators (undergraduates working on service-learning), reported in their journals that a lot more student-teacher rapport was observed outside classrooms when compared to regular class hours, which they believed contributed to more spontaneous use of English.

4.4 Challenges of SBA

The fourth research question covered the challenges that students faced in SBA. Table 4 summarizes the results of this research question.

Table 4

Challenges of SBA

Domains	Frequency (%)					
	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
1. Available resources to conduct SBA	14 (9.9)	30 (21.0)	38 (26.8)	37 (26.1)	19 (13.4)	4 (2.8)
2. Teachers' knowledge and skills to implement SBA	0 (0.0)	0 (0.0)	4 (2.8)	14 (9.9)	55 (38.7)	69 (48.6)
3. Sufficient class time for SBA	23 (16.2)	27 (19.0)	36 (25.4)	28 (19.7)	22 (15.5)	6 (4.2)

The data from the questionnaire (Table 4) showed that students mostly (87.3%) agreed or strongly agreed that their teachers were knowledgeable and skilled in implementing SBA. However, very few (about 15%) agreed or strongly agreed that there were sufficient resources for conducting SBA or sufficient class time for SBA (about 20%). Therefore, the lack of resources and shortage of time seem to be the main challenges in the implementation of SBA in schools.

Some of the findings for the interview showed that books were provided to prepare students for activities, but students considered this to be insufficient.

Some opined that there was a mismatch between the activities provided in class and the resources available for them to explore before the lessons, as can be seen in the excerpts below:

I wouldn't say really sufficient because there are not as many resources to prepare for PBS compared to during class or after class, where the teachers ask you to do some research about what you have learned today... Also, we are not using the resources provided all the time. Teachers would provide us with some newspapers. But most of the time, PBS does not provide exercises or activities that have anything to do with newspapers. So, we rarely use newspapers as a reference or something like that. [Muhsin, male, aged 16]

Though there are some books that my teacher will bring to class, I think that's all. I think that not having the Internet makes it difficult for us to do some of the tasks. Maybe we're in a fully residential school. That's why we can't use the Internet. I'll just use whatever knowledge that I have in mind to complete the assignments given to me. [Marhana, female, aged 16]

Well, I don't think my school has that sufficient resources for us to prepare for our lessons because at a certain point, it's just going to be me and my classmates brainstorming together. For example, if Iffah reads something on economics and I read about politics, we would exchange ideas. If the teacher would be like "okay today, we're going to do something on economy," then Iffah will help me to brainstorm. This is because we don't have enough resources. We don't have things like laptops, the Internet, or our phones to do our read-ups. [Areena, Female, aged 16]

The findings from class observations explained the claims made by the participants. Firstly, most of the external materials were outdated. The external materials that were referred to were on newspapers. Besides that, magazines were provided for the students to explore; nevertheless, only outdated publications in Bahasa Melayu were available. English magazines were subscribed but only one copy was available in the library for the whole school's use for SBA. In terms of devices, since this is a residential school, students were not allowed to bring their devices to school. The only devices that were available for use were the computers at the school computer lab. Students would use these computers for research purposes. During class observations, it was also discovered that when the students were instructed to write a poem, except for dictionaries, they did not have any other materials to refer to. To conclude, the list of findings for the class observations were as follows:

- External materials to assist students with their lessons were absent (newspapers were eight months old).
- Magazines provided were only in Bahasa Melayu and were in reading corners in classrooms. An outdated copy of an English magazine was provided in the library for the whole school's use. Students were 'reading' dictionaries to write a poem.
- No devices were brought by the students.
- The computers malfunctioned. The computer lab needed maintenance. Only a few computers worked.

Even though our results showed that most of the students who participated in our study had positive perceptions of SBA, we should not ignore the minority who rated the questionnaire items low. This suggested that there were students who were confused with SBA. Therefore, it is necessary that teachers make sure that students understand the concept of SBA as students' appreciation of the fact that SBA is meant to motivate them and improve their learning should help improve their attainment.

The investigation of SBA would bring great contribution to the area of teaching and learning. The perspectives of students in this area of research could provide feedback for instructional strategies, a guidance in grading system, and policy making. They can also be utilized to motivate students to learn the acquired skill in school (Issaka et al., 2020). Nonetheless, any change in the education system that might be proposed to be executed in the future would be a complex phenomenon in this field. The change in the implementation of assessments would require much resources and manpower (Hock et al., 2022).

5. Conclusion

Notwithstanding the limitations, this study is one of the few that gauged the upper secondary. The students' perspectives of SBA in Malaysia were brought to light. Many changes in the Malaysian Education policy require more research to be conducted in the future in relation to SBA in Malaysia. Ultimately, the students' perceptions of and attitudes towards SBA are also shaped by other stakeholders such as parents, teachers, and policymakers.

The findings of this study have theoretical contributions to the washback literature and practical implications for further improvement and implementation of SBA in Malaysia as well as in other parts of the world. In terms of methodological contributions, the nature of this sequential explanatory study could be an added value to the existing body of literature when it comes to the study of washback, perception, and language learning. The study findings also shed more light on the nature of the existing system after a decade of its implementation in Malaysia. Although the findings of this study could not be generalized to a broad context of Malaysian secondary schools, the study findings help contribute to the theoretical prepositions of washback, especially to our understanding of washback in relation to the context of summative and formative assessments (Mulianti, 2021).

Our study is not without its limitation. Firstly, the findings of this research worked mainly on a case study basis and the data that had been obtained could not be generalized to the population in general. Future research could work on different types of schools and inferential statistical tests could be conducted in order to compare the results. Research such as the one carried out by Alla et al. (2016) looked into the difference of perceptions pertaining to SBA between male and female students. The same method could be implemented in different school contexts. Secondly, this study focused on the micro level of washback. In future research, researchers could focus on both micro and macro levels of washback by examining students' perceptions of and attitudes toward SBA alongside teachers' perception and attitudes to better determine the extent of washback throughout wider stakeholders' point of the view. This method of study has been conducted in Hong Kong by Yu (2010), but none in relation to the context of Malaysia. Third, we excluded other types of schools in Malaysia such as private schools and vernacular schools. Moreover, due to the limited nature of the data whose focus was placed on the aspect of students as the stakeholder, studies that provide a richer understanding of the relationship between students' perceptions and policymakers' perceptions would be highly significant. Hence, a document analysis or an analysis of the interaction between different stakeholders is greatly needed to achieve this goal.

6. About the Authors

Nur Farisya Amylia Salleh teaches English at Tenby Schools Johor (Setia Eco Gardens). She completed her Master's degree at the Faculty of Modern Languages and Communication, Universiti Putra Malaysia (UPM). Her area of research interest is in language assessment.

Vahid Nimehchisalem started teaching English in 1996. He is currently an Associate Professor at the Faculty of Modern Languages and Communication, Universiti Putra Malaysia, and an Adjunct Professor at Asia e University Malaysia.

He's the chief editor of the *International Journal of Education and Literacy Studies* and the managing editor of the *Journal of Language and Communication*.

Ilyana Jalaluddin is a Senior Lecturer at the Department of English, Faculty of Modern Languages, Universiti Putra Malaysia. Her areas of specialization include TESL, ESL Writing, and Language Studies.

Jayakaran A/L A.P Mukundan holds a doctoral degree in English language from Universiti Putra Malaysia. He is currently a Senior Research Fellow at the School of Education, Faculty of Social Sciences & Leisure Management, Taylor's University, Malaysia. His main area of research interest is ELT materials evaluation.

7. Acknowledgement

We appreciate the journal editors and reviewers for their constructive feedback and professional support.

8. References

- Alderson, J. C., & Wall, D. (1993). Does washback exist?. *Applied linguistics*, 14(2), 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Alla K., & Hassan, N. (2021). Implementing learning-oriented assessment in Malaysia. In A. Gebril (Ed.), *Learning-oriented language assessment: Putting theory into practice* (pp. 140–161). Routledge.
- Baksh, A., Mohd Sallehudin, A. A., Tayeb, Y. A., & Norhaslinda, H. (2016). Washback effect of school-based English language assessment: a case-study on students' perceptions. *Pertanika Journal of Social Sciences & Humanities*, 24(3), 1087–1104.
- Alqahtani, F. (2021). The impact of language testing washback in promoting teaching and learning processes: A theoretical review. *English Language Teaching*, 14(7), 21–26. <https://doi.org/10.5539/elt.v14n7p21>

- Annie, T. S. Y. (2011). Exploring students' perception of and reaction to feedback in school-based assessment. *Malaysian Journal of ELT Research*, 7(2), 104–117.
- Azid, M. A. (2022). Higher order thinking skills, school-based assessment and students' mathematics achievement: Understanding teachers' thoughts. *International Journal of Evaluation and Research in Education (IJERE)*, 11(1), 290–302. <http://doi.org/10.11591/ijere.v11i1.22030>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Bailey, K. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257–279. <https://doi.org/10.1177/026553229601300303>
- Burgers, C., Eden, A., van Engelenburg, M. D., & Buningh, S. (2015). How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior*, 48, 94–103.
- Cheng, L., & Curtis, A. (2004). *Washback or backwash: A review of the Impact of testing on teaching and learning*. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3–17). Lawrence Erlbaum Associates.
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching*, 48(4), 436–470. <https://doi.org/10.1017/S0261444815000233>
- Chew, F. P., & Muhamad, N. (2017). Readiness of implementation of school-based assessment among the Malay language teachers in national schools. *Advanced Science Letters*, 23(3), 2169–2173. <https://doi.org/10.1166/asl.2017.8589>
- Chin, H., Thien, L. M., & Chiew, C. M. (2019). The reforms of national assessments in Malaysian education system. *Journal of Nusantara Studies (JONUS)*, 4(1), 93–111. <https://doi.org/10.24200/jonus.vol4iss1pp93-111>

- Cohen, A. D. (1984). On taking tests: What the students report. *Language Testing* 1(1), 70–81. <https://doi.org/10.1177/026553228400100106>
- Damankesh, M., & Babaii, E. (2015). The washback effect of Iranian high school final examinations on students' test-taking and test-preparation strategies. *Studies in Educational Evaluation*, 45, 62–69. <https://doi.org/10.1016/j.stueduc.2015.03.009>
- Dawadi, S. (2021). Factors Influencing Washback of a high-stakes English-as-a-Foreign-Language test. *Teaching English as a Second or Foreign Language (TESL-EJ)*, 25(3). 1–16 Retrieved from <https://www.tesl-ej.org/wordpress/issues/volume25/ej99/e99a1/>
- 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>
- Dong M., F. X. (2021). Differential washback effects of a high-stakes test on students' English learning process: Evidence from a large-scale stratified survey in China. *Asia Pacific Journal of Education*, 43(1), 252–269. <https://doi.org/10.1080/02188791.2021.1918057>
- Farhady, H. (2021). Learning-oriented assessment in virtual classroom contexts. *Journal of Language and Communication*, 8(2), 121–132.
- Fook, C. Y., & Sidhu, G. K. (2012). School-based assessment among ESL teachers in Malaysian secondary schools. *Journal of the Malaysian Education Dean's Council*, 9, 1–18.
- Ghazali, N. H. (2016). A Reliability and Validity of an Instrument to Evaluate the school-based assessment system: A pilot study. *International Journal of Evaluation and Research in Education (IJERE)*, 148–157. <http://doi.org/10.11591/ijere.v5i2.4533>
- Ha, N. T. T. (2019). A literature review of washback effects of assessment on language learning. *Ho Chi Minh City Open University Journal of Science*, 9(2), 3–16. <https://doi.org/10.46223/HCMCOUJS.soci.en.9.2.257.2019>

- Hashim, C. N., Ariffin, A., & Hashim, N. M. (2013). Ideal vs. reality: Evidences from senior teachers' experiences on the Malaysian school-based assessment system (SBA). *Proceedings of the Malaysian Education Deans' Council*, 770–781. <http://iium.edu.my/medc2013/>
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14(3), 295–303. <https://doi.org/10.1177/026553229701400306>
- Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: how many interviews are enough?. *Qualitative Health Research*, 27(4), 591–608. <https://doi.org/10.1177/1049732316665344>
- Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>
- Hock, T. T., Ayub, A. F. M., Shah, M. M., & Ahamed, A. B. (2022). Implementation of classroom-based assessment in Malaysia. *International Journal of Academic Research in Progressive Education and Development*, 11(4), 76–87. <https://doi.org/10.6007/IJARPED/v11-i4/14621>
- Hughes, A. (1989). *Testing for language teachers*. Cambridge University Press.
- Hughes, R. I. G. (1993). Theoretical explanation. *Midwest Studies in Philosophy*, 18(1), 132–153. <https://doi.org/10.1111/j.1475-4975.1993.tb00261.x>
- Hunter, M. (2022, April 26). *Turning Malaysia's failure in education into great success*. Free Malaysia Today (FMT). <https://www.freemalaysiatoday.com/category/opinion/2022/04/26/reimagining-malaysian-higher-education/>
- Issaka, J., Hammond, D. K., Yeyie, P., & Agroh, P. K. (2020). Benefits of school-based assessment in the learning of social studies. *Social Education Research*, 1(2), 219–228. <https://doi.org/10.37256/ser.122020458>
- Jamal, S. N., Abd Rahman, S. N. S., Sujak, N. N., Saravanan, A., Maniam, L., Obeng, S. N. W., Ching, T. Y., & Bathumalai, V. (2015). *What is School-*

based Assessment (SBA)? How SBA is implemented in classroom?.

<https://www.studocu.com/my/u/54133772?sid=0169824998>

- McKinley, J. & Thomson, G. (2018). Washback effect in teaching English as an international language. In J. I. Liontas, T. International Association & M. DelliCarpini (Eds.). *The TESOL Encyclopedia of English Language Teaching*, (pp. 1–12). <https://doi.org/10.1002/9781118784235.eelt0656>
- Joachim, A., & Hashim, H. (2021). ESL Teacher's Knowledge and Readiness on the Implementation of School-Based Assessment (SBA) in Malaysian Primary School. *Creative Education*, 12(5), 1066–1078.
- Kim, S., Raza, M., & Seidman, E. (2019). Improving 21st-century teaching skills: The key to effective 21st-century learners. *Research in Comparative and International Education*, 14(1), 99–117.
- Liu, X., & Yu, J. (2021). Relationships between learning motivations and practices as influenced by a high-stakes language test: The mechanism of washback on learning. *Studies in Educational Evaluation*, 68, 100967.
- Majid, F. A. (2011). School-based assessment in Malaysian schools: The concerns of the English teachers. *US-China Education Review*, 393–402.
- Ministry of Education, Malaysia. (2013). Executive summary: *Malaysia education blueprint 2013-2025 (preschool to post-secondary education)*. Ministry of Education Malaysia. <https://www.moe.gov.my/muat-turun/penerbitan-dan-jurnal/dasar/1207-malaysia-education-blueprint-2013-2025/>
- Mukundan, J. (2014). Why we must reeducate teachers on the way we teach writing. In N. Z. H.P Widodo (Ed.), *Recent Issues in English Language Education: Challenges and Directions* (pp. 180–190). UNS Press. <https://lib.atmajaya.ac.id/default.aspx?tabID=61&id=310091&src=a>
- Mukundan, J. (2016). Incorporating out-of-class English language development in pre-service teacher preparation. In I. B. Fuad Abdul Hamied (Ed.), *Developing indigenous models of English language teaching and assessment* (pp. 111–118). Udayana University Press.

- Mukundan, J., & Sha, I. H. P. (2015). A new classroom emerging: What role for materials?. *The European Journal of Applied Linguistics and TEFL*, 4(2), 129–141.
<https://link.gale.com/apps/doc/A528616277/AONE?u=anon~acd5e835&sid=googleScholar&xid=3e99157d>
- Mulianti, M. (2021). *An analysis of washback effect of English formative assessment to the teacher and students at Smait Ukhuwah Banjarmasin*. [Bachelor's thesis, Antasari state islamic university]. <https://idr.uin-antasari.ac.id/15358/2/AWAL.pdf>
- Mwita, K. (2022). Factors influencing data saturation in qualitative studies. *International Journal of Research in Business and Social Science (2147-4478)*, 11(4), 414–420.
- Nimehchisalem, V., Hosseini, M., Cortazzi, M., & Jin, L. (2023). Multiple perspectives of stakeholders towards young learners' language assessment in an international school in Malaysia. *Language Teaching Research*, <https://doi.org/10.1177/13621688231154440>
- Pan, Y. C. (2008). A critical review of five language washback studies from 1995-2007: Methodological considerations. *JALT Testing & Evaluation SIG Newsletter*, 12(2), 2–16.
- Pan, Y., & Newfields, T. (2011). Teacher and student washback on test preparation evidenced from Taiwan's English certification exit requirements. *International Journal of Pedagogies and Learning*, 6(3), 260–272.
- Pearson, P. D. (1991). Literacy Assessment in a Diverse Society, *CSR Technical Report* (525), 1–23. <https://core.ac.uk/download/pdf/4826362.pdf>
- Pimentel, J. L. (2019). Some biases in likert scaling usage and its correction. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, 45(1), 183–191.

- PISA. (2018). *PISA 2018 results executive summaries volume I, II and III*.
https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf
- Rahman, K. A., Seraj, P. M. I., Hasan, M. K., Namaziandost, E., & Tilwani, S. A. (2021). Washback of assessment on English teaching-learning practice at secondary schools. *Language Testing in Asia*, 11(1), 1-23.
- Salmon, J. (2015). Using observational methods in nursing research. *Nursing Standard*, 45, 36–41.
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity*, 52, 1893–1907.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language testing*, 13(3), 298–317.
- Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87(4), 324–328.
- Talib, R., Naim, H. A., Ali, N. S. M., & Hassan, M. A. M. (2014). School-based assessment: A study on teacher's knowledge and practices. *PalArch's Journal of Archaeology of Egyptology*, 17(9), 1–14.
https://www.researchgate.net/publication/347946206_school-based_assessment_practices_among_primary_school_mathematics_teachers_based_on_teaching_experience
- Wall, D. (1999). *The Impact of High-Stakes Examinations on Classroom Teaching: A case study using insights from testing and innovation theory*. [Unpublished Doctoral dissertation, Lancaster University].
- Wilson, D. M., & Narasuman, S. (2020). Investigating teachers' implementation and strategies on higher order thinking skills in school based assessment instruments. *Asian Journal of University Education*, 16(1), 70–84.

- Woo, W. T. (2019, March 27). Decentralising Malaysia's education system. *East Asia Forum*. <https://www.eastasiaforum.org/2019/03/27/decentralising-malysias-education-system/>
- 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.
- Yu, Y. (2010). *The washback effects of school-based assessment on teaching and learning - A case Study* [Doctoral dissertation]. Hong Kong University.