Validity and Reliability of a Questionnaire on the Knowledge, Attitude and Challenges among Malay Language Teachers in Implementing Classroom-Based Assessment (PBD)

Vinoothini Mannogaran & Nor Azwahanum Nor Shaid

Validity ansd Reliability of a Questionnaire on the Knowledge, Attitude and Challenges among Malay Language Teachers in Implementing Classroom-Based Assessment (PBD)

Vinoothini Mannogaran[1] & Nor Azwahanum Nor Shaid[2]

[1]Fakulti Pendidikan, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia,
[2]Fakulti Pendidikan, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia.
Email: vinoothinimanogaran@gmail.com, azwahanum@ukm.edu.my

Abstract
The purpose of this survey study is to design a research tool to assess the degree of knowledge, attitudes, and difficulties among Malay teachers in Malaysia during the implementation of the classroom-based assessment (PBD). By evaluating the validity and reliability of the instrument, the accuracy of each research item can be guaranteed. Two experts' feedbacks are taking into consideration to evaluate the face validity and content validity in the questionnaire. Meanwhile, Cronbach's Alpha value is used to determine the reliability rating. For this study, a number of 30 Malay teachers from the Seberang Perai Utara National School were chosen as participants. There are four elements in the questionnaire such as respondent demographics, knowledge, attitudes, and issues facing by Malay teachers must be filled by respondents. The Scale-Content Validity Index (S-CVI/Ave) for this study is 1.000, and the Content Validity Ratio (CVR) has a value of +1. The required construct domain has thus been represented properly by each study item. The Cohen Kappa value study's findings demonstrate that this research instrument has a high validity value (k = 0.680, p = 0.001) as well as high assessor reliability (α = 0.813). This indicates that this research instrument is excellent for conducting research. This study's implications will make it easier for schools, PPD, JPN, and KPM to implement PBD effectively.
Keywords: Validity, Reliability, Classroom-Based Assessment, Knowledge, Attitudes, Challenges

Introduction
The School-Based Assessment was introduced in 2011 as a replacement for the exam-oriented education system in Malaysia. The purpose of School-Based Assessment is to assess pupils' performances in terms of various aspects such as cognitive (intellectual), affective (emotional and spiritual), and psychomotor (physical) which align with the National Education Philosophy. However, in 2018 Classroom-Based Assessment, which is known as '*Pentaksiran Bilik Darjah*' (PBD) was first introduced and aligned with the Standard Secondary School Curriculum. PBD was conducted to evaluate and interpret the pupils' performance in all the subjects throughout the year. Climie & Henley (2016), revealed that now we are in an era that

requires measurement based on the ability owned by the pupils and not comparing the difference in scores.

According to Halimah & Rozita (2019), the purpose of PBD helps to develop pupil's learning and overall proficiency in Malay Language. Pupils can recognize their strengths and weaknesses in the learning session if this assessment process is carried out. Meanwhile, teachers have the opportunity to plan and improve their teaching methods to be more effective (KPM, 2018). This is because teachers are responsible for evaluating and making overall scoring for pupils at the school level (Zamri et al., 2010). Furthermore, PBD depends entirely on teachers who handle it so that the success of this transformation can be seen clearly. Although PBD has been drafted based on the concept of formative assessment that teachers have been practicing in schools for a long time, there is still a concern for teachers to use this assessment method to evaluate pupils (Keddie, 2018). According to Kalai & Joohari (2021), this situation became livelier when the Ministry Of Education (MOE) announced the abolition of UPSR and the full implementation of PBD. Therefore, this preliminary survey was conducted to identify the knowledge, attitudes, and challenges faced by Malay teachers during the implementation of PBD in teaching and learning processes.

Literature Review
Classroom-Based Assessment (PBD)
Usually, the teaching and learning process requires planning and implementation followed by assessment by respective teachers. The assessment process is practiced in Malaysian schools, especially in the context of Malay Language Education. Learners' ability to extract a piece of information and share creative and innovative opinions by using correct pronunciation and intonation is the purpose of the assessment (Sukatan Pelajaran Bahasa Melayu, 2000). The Malay language learning process emphasizes writing skills to improve the implementation of assessments from time to time. The teacher will conduct the assessment according to the appropriate school location and set a date for its implementation (Aniza & Zamri, 2014).

According to Sahari (2002), the process of collecting pupils' information for interpretation and the data used to improve the effectiveness of the teaching and learning process is defined as PBD. This process is managed by subject teachers to measure and evaluate the quality of the teaching process. Mark (2018) argues that PBD not only aims to improve pupils' abilities and achievements in lessons but is also carried out for teachers to recognize the strengths and weaknesses of pupils to improve the teaching process based on pupils' feedback. Assessment helps a teacher to plan and design quality lesson plans according to the pupils' proficiency. This situation will help the pupils achieve their desired level of proficiency (Maki, 2002). Therefore, teachers should be more focused on the assessment to build excellent character and practice among pupils (KPM, 2018).

Knowledge, attitudes and challenges of implementing Classroom-Based Assessment
In fact, from the past studies, it states the scoring criterion which prepared by Malay teachers is very important in the implementation of PBD. Heejeong (2015) believes that scoring rules are the conditions required by evaluators when conducting PBD. These scores are determined by the teacher through the evaluation process during the teaching session according to grade from the easiest level to the most difficult. Evaluators use scoring criteria as a guide to assess and track pupils' progress and enable pupils to improve their academic performance (Arter,

2000). This is because the teaching process can be done successfully by teacher's knowledge and the way the teacher handles it. Therefore, teachers need to plan their activities after teaching session such as assessment to see the quality of their teaching.

Next, there is a study by Masfarizan & Yusoff (2020) which discusses the attitude of Malay Language teachers in providing appropriate assessment materials based on the teaching session. It also discusses teachers' own steps or initiatives to assess pupils during teaching and learning processes. Huzairi (2017) also stated that scores of PBD should be taken efficiently to evaluate pupils' proficiency level during the implementation PBD in school for long term. Malaysian education system will be affected among others if a teacher is not playing their roles properly (Norshafinaz & Faridah, 2018). A teacher should not rely on a piece of paper to evaluate pupils' proficiency level. Thus, PBD is also considered as an overall assessment which helps to develop pupils' potential.

In addition, Arumugham's (2020) proves that teachers face various challenges in the implementation of PBD such as lack of equipment and teaching materials. Norafizah (2018) asserted that the main challenge in the implementation of PBD is the use of ICT equipment as a medium in teaching and assessment by teachers who lacked ICT skills. He also stated that teachers face challenges in preparing assessment materials that are suitable for pupil's ability levels and that can attract their interest. In addition, teachers also face challenges in terms of the attitude of pupils who do not know the objectives of PBD implementation. The pupils cannot appreciate the benefits and will not try to develop their potential in order to get a good proficiency level score (Jacob & Parkinson, 2015). Therefore, it is proven that the implementation of PBD needs to be improved so that the issues mentioned can be solved.

Therefore Vlachou & King (2018), have suggested a few steps for teachers to carry out assessment correctly. For example, the teacher needs to determine the pupil's achievement target, identify the level of the pupil's existing knowledge, determine what needs to be taught so that the pupils can achieve the learning objectives and take appropriate actions in teaching. Teachers must have evidence of complete assessment implementation so that the success can be seen. As a result of these steps, the teacher can determine the proficiency level score sheet that should be given to a pupil. Therefore, Koh's study (2011), has suggested that the Ministry of Education needs to conduct a teacher empowerment program so that teachers' competence level can be improved.

Research Objectives
This study aims to test the validity and reliability of research instruments on the knowledge, attitudes, and challenges of Malay teachers in implementing PBD.

Methodology
The researcher used a quantitative approach to conduct this preliminary survey. According to David et al (2010), a researcher can measure the relationship between variables based on data received through quantitative research. The researcher used descriptive and inferential survey method to conduct this research. According to Noraini (2010), this survey method makes it easier for the researcher to obtain the necessary information from the study respondents. The findings of the study have been obtained from the constructed questionnaire. The respondents for this study were a total of 30 Malay teachers from

Seberang Perai Utara district. There are four sections in this research instrument that must be answered by each respondent. However, this questionnaire was sent to two experts to calculate the validity value while the reliability value was calculated through the findings received from the study respondents.

Questionnaire Instrument

A questionnaire was used as an instrument in this study. According to David et al (2010), a questionnaire is defined as a set of printed or online questions that help researchers to obtain or collect research data individually. Several journals and articles from past studies have been consulted by the researcher in the process of producing this questionnaire instrument. The instruments that have been built by other researchers are also referred as a guide in the construction of questionnaire items. This questionnaire includes four parts. Each section contains 10 questions. Questions in the form of a Likert scale with five answer options were given to identify the level of respondents' agreement with the items presented. Researchers often use this scale to measure the views or responses of research participants about the knowledge, attitudes, or challenges (Budiaji, 2013). So, the survey respondents have to choose one of the numbers starting from 1 to 5 according to their level of agreement.

In order to test the validity of the research instrument, the researcher has obtained content validity and face validity from two experts in the field related to this study. Validity is defined as the appropriateness of the item with the purpose of the study to enable inferences to be made by the researcher (Fraenkel & Norman, 1996). The accuracy of a research instrument is tested and measured through validity by experts in their respective fields. So, two members of the Research Instrument Review Panel were appointed to review the research instrument by conducting the validity test. These panellists are lecturers at public universities and experts in the field of language and research. These two panels were asked to review and score the research instrument based on face validity and content validity. They have been given two options to evaluate each question item, which is 'agree' and 'disagree'. Comments for items that disagree need to be corrected and need to include in the instruments returned by the review panel. Overall, both review panels stated that appropriate and understandable grammar and language was used in the questionnaire.

However, in order to avoid the occurrence of bias elements, the researcher has measured the validity of the content by using the Content Validity Ratio (CVR), Content Validity Index (CVI) and also the Cohen Kappa test. CVR is one of the measurements that helps to maintain or drop the items in the study instrument. According to Yusoff (2019), CVR helps researchers to filter each item empirically to ensure that it truly represents the content of the construct domain. Mohamed et al (2017), stated that the CVR value will be in the range of -1 to +1. A value close to +1 indicates that the expert reviewer has confirmed that the item is very important and closely related to the topic of the study. While a value below 0.50 means that the item has doubtful elements, needs improvement, or needs to be dropped. The CVI value will be calculated once the CVR value is identified. According to Shrotryia and Dhanda (2019), a CVI value in the range of 0.8 and above means that the items in the research instrument have clear and high content validity. The method of calculating CVR and CVI is as follows:

$$CVRi = \frac{[ne-(N/2)]}{N/2}$$

ne = the number of expert panels that rated the item as important
N = total number of expert panels involved

I-CVI = Approved item (Score 3 or 4)
        Number of experts

S-CVI/Ave = Total score of I–CVI
        Number of items

As a result of calculating the CVR value, the researcher has successfully proved that all the items in this research instrument are valid and suitable for use. This is because, the result of the CVR calculation shows a value of +1 and it means that each study item has represented the desired construct domain. The value of CVI shows a value of 1.000 as found in Table 1. This value is calculated based on the agreement of the two expert reviewers (N=2). Therefore, the researcher found that this research instrument has appropriate content validity. The results of the calculation of CVR and CVI values are as below:

Table 1
Calculation of CVI and CVR from questionnaire

Item	Expert 1	Expert 2	Expert Consent	I-CVI	CVR
1	1	1	2	1.000	1
2	1	1	2	1.000	1
3	1	1	2	1.000	1
4	1	1	2	1.000	1
5	1	1	2	1.000	1
6	1	1	2	1.000	1
7	1	1	2	1.000	1
8	1	1	2	1.000	1
9	1	1	2	1.000	1
10	1	1	2	1.000	1
11	1	1	2	1.000	1
12	1	1	2	1.000	1
13	1	1	2	1.000	1
14	1	1	2	1.000	1
15	1	1	2	1.000	1
16	1	1	2	1.000	1
17	1	1	2	1.000	1
18	1	1	2	1.000	1
19	1	1	2	1.000	1
20	1	1	2	1.000	1
21	1	1	2	1.000	1
22	1	1	2	1.000	1
23	1	1	2	1.000	1
24	1	1	2	1.000	1
25	1	1	2	1.000	1
26	1	1	2	1.000	1

27	1	1	2	1.000	1
28	1	1	2	1.000	1
29	1	1	2	1.000	1
30	1	1	2	1.000	1
31	1	1	2	1.000	1
32	1	1	2	1.000	1
33	1	1	2	1.000	1
34	1	1	2	1.000	1
35	1	1	2	1.000	1
36	1	1	2	1.000	1
37	1	1	2	1.000	1
			S-CVI/Ave	1.000	

Next, the evaluation of the Cohen Kappa test is calculated based on the responses given by the review panel. The purpose of this evaluation is to express agreement between the instrument reviewers in determining a high reliability value for a question item. So, the researcher has used the Statistical Package for Social Sciences (SPSS) version 26.0 to obtain the Cohen Kappa value. The findings have been analysed according to the Kappa Coefficient value interpretation scale suggested by Bernard and Ryan (2010) as found in table 2. According to table 2, a Kappa Coefficient value greater than 0.61 is classified as good and acceptable.

Table 2
Kappa Coefficient value interpretation scale

Interpretation	Kappa Values
Very good	$0.81 - 0.99$
Good	$0.61 - 0.80$
Medium	$0.41 - 0.60$
Moderately weak	$0.21 - 0.40$
Weak	$0.00 - 0.20$
Very weak	< 0.00

The researcher also calculated the reliability value through the feedback given by the respondents. Reliability is a research tool that tests the consistency of accurate item statements, creating uniformity on each item and the instructions of each item clearly (Sabitha, 2005). A total of 30 primary school Malay teachers in Seberang Perai Utara district were made as respondents in this survey study. The researcher has measured the level of reliability of the research instrument through the findings obtained from the respondents. The researcher has used the Croanbach's Alpha Test in SPSS software and the value is measured through the scale in table 3. According to Bond and Fox (2007), Croanbach's Alpha value greater than 0.61 is classified as good and acceptable.

Table 3
Interrestation of Croanbach Values

Croanbach Alpha Value	Reliability
0.9 – 1.0	Very good and effective
0.7 – 0.9	Good and acceptable
0.6 – 0.7	Acceptable
< 0.5	Needs to be improved

Finding and Discussion

The researcher has evaluated the instrument in terms of validity and reliability so that it can be used in the actual study. Sabitha (2006) thinks that a researcher is able to measure the accuracy of his research instrument by using validity. Researchers are also able to identify all the features or elements that must be present in a research instrument through validity measurement. While according to Chua (2012), the researcher can identify the ability of the study in obtaining the same value even if the same measurement is repeated through reliability measurement. The findings of the study have proven the validity and reliability of the questionnaire to be at a high level.

Kappa Coefficient Test

Validity values were analysed through 37 items that were given to two expert reviewers of the research instrument. The evaluation of the validity of the 37 items is divided into two parts, which is face validity and content validity. Validity values were calculated using the Cohen Kappa calculation through SPSS software. Table 4 shows that the two instrument reviewers had the same agreement on 34 question items and had differences of opinion on 3 question items. The Kappa Coefficient value also shows a value of 0.680 where it is at a good level of validity and suitable for collecting research data (Bernard and Ryan, 2010). This means there is significant agreement between the first and second instrument reviewers ($p = 0.001$ < alpha = 0.05).

Table 4
Findings of Kappa Coefficient value calculation

Rating Item	Total
Similarities	34
Difference	3
Kappa Coefficient Value	0.680
Approx. Sig (*p*) Value	0.001

However, comments and suggestions were given by both reviewers for items that were not agreed upon and needed to be improved. For example, the first reviewer (R1) and the second reviewer (R2) have suggested that the researcher not use abbreviations in the research instrument. The word *'P&P'* found in the original questionnaire was asked to change to '*proses pengajaran dan pembelajaran'*. In addition, there are also some errors in the use of adverbs that need to be corrected by the researcher. For example, the word *'memperbaiki'* is wrong in terms of the use of affixes and it is suggested to change it to *'membaiki'*. The researcher also made a mistake in use of few terms such as '*tertinggal banyak masa'* which was proposed

to be changed to *'kekurangan masa'*. However, both experts have stated that all the items provided have met the requirements of the study.

Table 5
Comments from instrument review experts

Section	Item	Comment
Knowledge	B1	P&P must be written as full (R1/R2)
	B8	Sentence structure error (R2)
Attitude	C8	Grammatical error (R1)
	C9	Grammatical error (R2)
Challenge	D2	Wrong terms (R2)
	D3	Spelling mistakes (R1/R2)
	D7	Wrong terms (R1/R2)
	D8	Sentence structure error (R2)

Note: R1 = First Reviewer, R= Second Reviewer

Reliability
As a result of the data analysis of the pilot study, the researcher got a Croanbach Alpha value of part B of 0.609, the value of part C was 0.756 while the value of part D was 0.870. The overall reliability value of this study's questionnaire items is 0.748. Based on the total alpha value of 0.813, the reliability value of the questionnaire used is considered good and acceptable (Bond and Fox, 2007).

Table 6
Croanbach Alpha value of the pilot study

Variable	Number of items	Alpha Values
Section B	10	0.609
Section C	10	0.756
Section D	10	0.870
Total	**30**	**0.813**

Conclusion
In conclusion, in a very short period of time, Ministry of Education Malaysia made statement that PBD will be replaced to UPSR examination in the national education system. The school teachers are able to implement PBD very well without proper courses or training. Keddie (2018) also believes that if policy makers feel that teachers are capable of implementing the changes that will be introduced, the changes can be made at any time. As PBD is a new inspiration to all Malaysian educators, it's essential for the educators to understand, apply and adopt as well as implement during teaching and learning processes. Thus, this PBD should be studied in depth. The researcher successfully conducted this initial survey using final result of this study.

This research also contributes theoretically and contextually for Malay Language teachers and researchers. This research instrument helps teachers to recognize their existing knowledge, the behavior that need to be have and the challenges will be faced through the

implementation of PBD. Teachers also play an important role in the implements of PBD. So, this research instrument is able to give exposure to Malay Language teachers about their field of work. In terms of context, this study helps teachers to recognize new updates in the education system. Therefore, the researcher hopes to provide a quality research instrument about PBD through this study. Future researchers can use the data of this study as reference materials to conduct more researches on PBD as it's still new. All the educators still need guidance on how to implement the PBD using correct procedures. Thus, educators should recognize the importance in implementing PBD in their teaching sessions. The results of this preliminary survey also proved that the research instruments used by the researcher had high validity and reliability. It shows that researchers could use this research instruments to obtain research data in the future. With this, the quality of assessment in Malaysia can be improved to an international level.

References

Abdullah, H. (2017). Pentaksiran dalam proses pengajaran dan pembelajaran Bahasa Arab di sebuah sekolah menengah. *Kertas Projek Sarjana Pendidikan yang tidak diterbitkan*. Fakulti Pendidikan, Universiti Malaya, Kuala Lumpur.

Ahmad, A., & Mahamod, Z. (2015). Pengetahuan, kefahaman dan kemahiran guru Bahasa Melayu terhadap pelaksanaan pentaksiran berasaskan sekolah (PBS) di sekolah menengah. *Kertas Projek Sarjana Pendidikan*. Fakulti Pendidikan, Universiti Kebangsaan Malaysia.

Arter, J. (2000). *Rubrics, scoring guides, and performance criteria: Classroom tools for assessing and improving student learning*. Pembentangan di mesyuarat tahunan New Orleans, LA: American Educational Research Association.

Arumugham, K. S. (2020). Kurikulum, Pengajaran Dan Pentaksiran Dari Perspektif Pelaksanaan Pentaksiran Bilik Darjah. *Asian People Journal*, 3(1), 152–161.

Bernard, H. R., & Ryan, G. W. (2010). Analyzing Qualitative Data. Los Angeles, Thousand Oaks, CA: Sage Publications.

Bond, T. & Fox, C. (2007). Applying the Rasch model: Fundamental measurement in the human sciences (2nd). Mahwah, NJ: LEA.

Budiaji, W. (2013). Skala Pengukuran dan Jumlah Respon Skala Likert (The Measurement Scale and The Number of Responses in Likert Scale). Ilmu Pertanian Dan Perikanan, 2(2), 127–133.

Chua, Y. P. (2012). *Kaedah dan statistik penyelidikan buku 2: Asas statistik penyelidikan.* Edisi ke-2. Kuala Lumpur: Mc Graw Hill Education.

Climie, E., & Henley, L. (2016). A renewed focus on strengths-based assessment in schools. *British Journal of Special Education*, *43*(2), 108-121.

David, L., Salleh, N. A., Yasin, S. M., Shahran, K. M., Yatim, A. M., & Habib, A. R. (2010). *HBEF2503: Kaedah Penyelidikan dalam Pendidikan.* Open University of Malaysia, Kuala Lumpur.

Fraenkel, R.J, & Wallen, N.E. (1996). *How to Design and Evaluate Research*. USA : Mc. Fraw-Hill Inc.

Heejeong, J. (2015). What is your teacher rubric? Extracting teachers' assessment constructs ? Practical Assessment, *Research & Evaluation*, 20 (6): 1-13.

Idris, N. (2010). *Penyelidikan dalam Pendidikan.* Mc Graw Hill, Kuala Lumpur.

Jacob, R. T., & Parkinson, J. (2015). The Potential for School-Based Interventions that Target Executive Function to Improve Academic Achievement: *A Review: Review of Educational Research*, 512-552.

Jamil, H. & Radhiah, R. (2019). Pelaksanaan Penskoran Pentaksiran Lisan Bahasa Melayu Dalam Pentaksiran Bilik Darjah. *Jurnal Pendidikan Bahasa Melayu*, Vol 9 (2), 25-39.

Keddie, A. (2018). *Adult education: An ideology of individualism. In Adult education for a change.* 45-64. Routledge.

KPM. (2018). *Panduan Pelaksanaan Pentaksiran Bilik Darjah.* Putrajaya: Bahagian Pembangunan Kurikulum.

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, 22(3), 255-276.

Maki, P. L. (2002). Developing an assessment plan to learn about student learning, *Journal of Academic Librarianship*, 28(1), 8-13.

Masfarizan, M., & Yusoff, N. M. R. N. (2020). Kebolehlaksanaan Pentaksiran Bilik Darjah (PBD) secara atas talian sepanjang perintah kawalan pergerakan (PKP) di daerah Sentul, Kuala Lumpur. *Prosiding Seminar Nasional FIP* 2020, 213–218.

Mohamed, Z., Lebar, O., & Shamsuddin, S. (2017). Pembinaan Dan Penilaian Instrumen Ujian Aptitud Kemasukan Ke Institut Pengajian Tinggi Malaysia. *Journal of Science and Mathematics Letters, 5*, 16-27.

Mahamod, Z., Embi, M. A., Yusoff, N. M. R. N., Badusah, J., & Sharala. (2014). *Strategi Pembelajaran Bahasa Melayu dalam kalangan Pelajar Warganegara Asing Berdasarkan Kemahiran Bahasa dan Gred.* Bangi: Penerbitan Fakulti Pendidikan, Universiti Kebangsaan Malaysia.

Norafizah. (2018). The Washback Effect of Primary School Evaluation Test (UPSR) On Teaching and Learning: A Case Study Of An English Teacher In Kuala Terengganu, Malaysia. *International Research Journal of Education and Sciences*, 2 (2) : 15 – 25,

Sabitha Marican. (2005). *Kaedah penyelidikan Sains Sosial*. Prentice Hall/ Pearson: Malaysia.

Sabitha Marican. (2006). *Penyelidikan Sains Sosial Pendekatan Pragmatik*. Batu Caves: Edusyste.

Sahari. (2002). *Pengujian Dan Penaksiran Di Bilik Darjah*. Kuala Lumpur: Universiti Islam Antarabangsa Malaysia.

Sani, N. A., & Yunus, F. (2018). Amalan Perancangan, Pelaksanaan dan Pentaksiran dalam Proses Pengajaran dan Pembelajaran Pranumerasi di Tadika Swasta. *Jurnal Pendidikan Malaysia*, 101-110.

Selvan, K., & Ariffin, J. (2021). Pemansuhan Ujian Penilaian Sekolah Rendah Dan Pelaksanaan Pentaksiran Bilik Darjah: Satu Tinjauan Dalam Kalangan Guru Besar. *Asian People Journal*, 4 (2): 80 -89.

Shrotryia, V. K., & Dhanda, U. (2019). Content Validity of Assessment Instrument for Employee Engagement. SAGE Open, 9(1), 1–7.

Sukatan Pelajaran Bahasa Melayu. (2000). *Huraian sukatan pelajaran Bahasa Melayu edisi penyesuaian*. Kuala Lumpur: Pusat Perkembangan Kurikulum, Kementerian Pelajaran Malaysia.

Vlachou, M., & King, S. (2018). Classroom assessment practices in middle school science lessons: A study among Greek science teachers. *Cogent Education*, *5*(1), 1455633.

Wilson, M. (2018). Making measurement important for education: The crucial role of classroom assessment. Educational Measurement: Issues and Practice, 37(1), 5-20.

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. Resource, 11(2), 49-54.