




Content Validity of Teachers' Efficacy Instruments Toward Implementing of Classroom Assessment in Primary Schools Using Content Validity Ratio (CVR)

***Rodhiana Rosli, Siti Eshah Mokhsein, Zahari Suppian**

Faculty of Human Development, Universiti Pendidikan Sultan Idris, Malaysia

Article Info	ABSTRACT
<p>Article history: Received: 8 June 2023 Revised: 25 July 2023 Accepted: 7 August Published: 1 September 2023</p>	<p>Teachers' efficacy can influence how well they perform their responsibilities, including assessing pupils' learning. As a result, a specific instrument must be used to measure teachers' efficacy. Content validity is one factor that is necessary to create an effective instrument. This study will use the content validity ratio (CVR) to assess the instrument's content validity to gauge teachers' efficacy in implementing classroom assessments in primary schools in Malaysia. There are seventy-five items in this instrument, with four primary constructs. Twenty-one experts were chosen by purposive sampling to take part in this study. Eleven are professional experts in psychometrics, assessment, educational psychology, languages, and efficacy. Ten lay experts were also involved in this study, including teachers as classroom assessment practitioners. The results of the study show that the instrument constructed has relevant content validity and has the potential to be an instrument that can measure teachers' efficacy in classroom assessment implementation. In the future, researchers should concentrate their efforts on a Rasch measurement model-based construct analysis to improve the quality of the instrument.</p>
<p>Keywords: Content Validity Ratio (CVR), Instrument, Teachers' Efficacy, Classroom Assessment</p> <p></p>	

Corresponding Author:

*Rodhiana Rosli,
Faculty of Human Development, Universiti Pendidikan Sultan Idris, Malaysia.
Email: rodhianarosli29@gmail.com



This is an open access article under the CC BY-SA license.

INTRODUCTION

The goal of raising the standard of education should be achieved both nationally and internationally. Shift one of the Malaysian Education Development Plan (PPPM) 2013–2025 and the Sustainable Development Goals (SDGs) both explicitly declare this goal (Jamaludin et al., 2019; Mokhsein, 2018). Teachers now have responsibilities beyond merely instructing students to fulfill these aspirations. In addition, teachers must make sure that students are given relevant information and training. This will allow students to put their newly acquired information and skills to use in practical settings.

Teachers typically try to change their methods of instruction to give their students the knowledge and skills they need to succeed. However, not all teachers are open to change. Only teachers with high efficacy could do it. This is because high efficacy aims to increase teachers' confidence in fulfilling their responsibilities, particularly when teachers must change current practices (Bruun & Evans 2020; Mood et al. 2017; Tschannen-Moran & Hoy 2001). Studies have also revealed that teachers with high efficacy are more inclined to pick up new skills to enhance their instruction and better fulfill the needs of their students (Lin et al., 2019; Shamsudin & Majid, 2018). Additionally, teachers with high efficacy are more inclined to embrace any adjustments in the education system.

The confidence a teacher has in their ability to complete assignments is referred to as efficacy (Baharin et al., 2017; Bruun & Evans, 2020; Hoi et al., 2017). Since the teacher must perform this action to evaluate the effectiveness of education in the classroom, the teacher's efficacy is critical from the perspective of addressing this activity (Perera et al., 2019). Teachers with high efficacy levels are better equipped to conduct assessments with confidence and effectiveness (Bruun & Evans, 2020; Green, 2019). However, the survey found that teachers are still not prepared to properly execute the classroom assessment (Sulaiman & Hassan, 2019). In actuality, there are many difficulties that teachers must overcome during the teaching and learning processes. Teachers are less likely to properly execute the classroom assessment due to time constraints and the huge number of students (Fauzi & Ahmad, 2022).

It is necessary to evaluate teachers' efficacy in using the classroom assessment because it is well-known how crucial teachers' efficacy is to the quality of education, particularly in the context of the implementation of the classroom assessment. By doing this, the assessment process will be appropriately conducted. Therefore, instruments that satisfy the requirements of validity and reliability need to be developed to assess how effective teachers are at implementing classroom assessment. The instrument's content must gauge teachers' efficacy in implementing classroom assessments. This study will focus on the elements of content validity testing for instruments that were developed using the Content Validity Ratio (CVR) to achieve this goal.

The CVR method is used to find the content validity of EfiG-PBD because this method has advantages over other methods. In addition, the CVR method is more practical in terms of time and cost (Matore et al., 2017). Also, the CVR method is more open, clear, and easy to use because it is simple to calculate the CVR, and the critical cut-off value is easy to find (Cohen et al., 2013). As said by scholars, content validity is the first step in deciding if instrument items align with the original purpose (Bond & Fox, 2015; DeVellis, 2017; Zainal et al., 2020). In addition, content validity can also figure out the representativeness of items to measure what the researcher wants to measure (Zainal et al., 2020). Therefore, this study focused more on how CVR can ensure that each item in the instrument accurately reflected teachers' efficacy in implementing classroom assessments,

especially in Malaysian primary schools.

LITERATURE REVIEW

To provide a comprehensive overview of the ideas covered in this article, a literature review was done. This section will include literature reviews on classroom assessment, teachers' efficacy, and Content Validity Ratio (CVR).

Classroom Assessment

Assessment is a process conducted by the teacher to obtain information to decide the level of mastery of students. Through assessment, teachers can use the information obtained to design further learning (Brookhart, 2003; Fives & Barnes, 2020). Besides, judgmental decisions are also important in deciding the effectiveness of a teacher's teaching. Therefore, assessments should be conducted continuously in the classroom under the auspices of the teacher (Ghazali et al., 2018), and the intended measurement is known as classroom assessment.

Most researchers define classroom assessment as the process of collecting, synthesizing, interpreting, and using classroom assessment results to make decisions related to student learning and teacher teaching performance (Brookhart, 1997; Fives & Barnes, 2020; Russell & Airasiaan, 2012). The information obtained during the classroom assessment process is very important in helping teachers improve student learning mastery. This concept is also applied in the implementation of classroom assessment in Malaysia, as described in the Classroom Assessment Implementation Guidelines (Bahagian Pembangunan Kurikulum, 2019). Since classroom assessment has been identified as the most effective assessment tool for influencing student learning and motivation (Fives & Barnes, 2020; McMillan, 2018), teachers need to have the motivation to conduct classroom assessments (Bruun & Evans, 2020). At the same time, teachers' efficacy is an important aspect that can help teachers increase their confidence in implementing classroom assessments effectively and properly.

Teachers' Efficacy

Teachers' efficacy in conducting assessments refers to the teacher's belief in the planning, preparation, and implementation of assessment-related activities (Hartell, 2018). This definition coincides with the concept of self-efficacy proposed by Bandura (1997). According to Bandura (1997), the teachers' efficacy is the confidence of the teacher in performing his duties successfully. In the context of the implementation of the classroom assessment, teachers need to have the confidence to conduct the evaluation process based on the standards set out in the Standard Document Curriculum and Assessment (DSKP).

Teachers' efficacy will guide and motivate them to act. Teachers who have high efficacy will act on the basis of confidence in their ability to carry out the assessment process and continue educational activities that will definitely affect students' learning (Hartell, 2018). Although the self-efficacy of teachers is important in encouraging them to perform classroom assessments effectively, studies relating to the teachers' efficacy in implementing classroom assessments are rarely done (Green, 2019). These independent studies focus more on the study of teachers' efficacy in other aspects such as teaching, leadership, job satisfaction, and emotional intelligence. Therefore, the implementation of self-efficacy studies of teachers in aspects of classroom assessment, especially related to the development of instruments, should be done in order to measure the extent to which the teacher's

efficacy in implementing classroom assessment. In addition, the content of the instrument must gauge teachers' efficacy in implementing classroom assessments.

Content Validity Ratio (CVR)

CVR is a quantitative approach developed by Lawshe in 1975. CVR is a method to measure consensus among experts on the importance of an item in an instrument. The CVR approach is more correct, user-friendly, and straightforward than other alternatives for proving content validity, such as Cohen's Kappa, Tinsley-Weiss, James, Demaree, and Wolf Index. Additionally, CVR offers a table for showing critical values that must be met (Chong et al., 2021; Wilson et al., 2012). Local and foreign researchers have widely used the CVR method as a first step in the instrument development process (Aziz et al., 2018; Zainal et al., 2020; Matore et al., 2017; Surip et al., 2019; Arifain et al., 2021).

The procedure for conducting a CVR analysis requires an assessment by a panel of experts appointed to decide whether a measurement item is essential to keep based on the theoretical operationalization of the construct (Johnston & Wilkinson, 2009). The CVR implementation procedure begins with the determination and selection of a panel of experts who have recognized ability in the field under study (Chong et al., 2021). The experts evaluated each item by using a three-point scale for each item, which is (1) important (essential), (2) useful but not essential, and (3) not necessary (Lawshe, 1975). Content validity was decided based on the professional judgment of a panel of experts by asking for their views and comments on the instrument's compatibility with constructs, items, and measurement scales. If more than the number of experts involved evaluate the item as important, then the item is considered to have met content validity (Lawshe, 1975). Sekaran and Bougie (2016) found that the content validity is higher when the item scale is more like the construct.

CVR values range from -1 to +1, where a value close to +1 shows that experts agree that the item is essential in content validity. A CVR value of less than zero ($CVR < 0$) suggests that less than half of the expert panel believes the measurement item is essential means that half of the expert panels considered the item unimportant. A CVR value equal to zero ($CVR = 0$) means that half of the expert panel's sample size considers the item measurement important, and the other half agrees it is unimportant. A CVR value greater than zero ($CVR > 0$) shows that half of the expert panel believes the measurement items meet content validity. In this study, the minimum value of CVR for twenty-one experts needs to be at the critical CVR value of 0.429 (Colin & Andrew, 2014). The formula for deciding the value of CVR is as follows:

$$\text{Content Validity Ratio (CVR)} = [n_e - (N / 2)] / (N / 2)$$

n_e = The number of experts who agree is important

N = The number of research expert panels that were involved.

METHODOLOGY

The study uses a quantitative approach in the form of a questionnaire survey. Two expert categories, professional and lay, include the research sample. Researchers choose samples by using a purposeful sampling technique (judgment sampling). This type of sampling chooses the sample based on the expert's knowledge of the studied subject (Sekaran & Bougie, 2016). This technique gets information from people who know a lot about the subject. So, judgment sampling is the best method for this

study because the expert panel was chosen for one reason: to confirm the content of items related to the instrument.

In this study, the researcher looks to Rubio et al.'s (2003) definition of an expert, which says that there should be at least three expert panels for each group, including professional expert groups and lay expert groups, with a total of more than ten experts or accepted between six and twenty people. This study recruited eleven professional expert panels and ten lay expert panels. The number of experts exceeds the range suggested by Rubio et al. (2003), aiming to obtain more robust and more convincing content validity results. Using more expert panels will produce more information on the things to be measured.

RESEARCH FINDINGS AND DISCUSSION

A total of twenty-one expert panels, which included eleven professional experts and ten lay experts, took part in validating the instrument. The response rate received from all the professional and lay expert panels was 100 percent. All the expert panels (professional and lay experts) completed their assessment within the given period. The eleven professional expert panels were academicians who worked as lecturers or researchers in the education sector. In contrast, the ten lay expert panels were the research subject directly involved with the assessment process in the primary schools.

The list of professional and lay expert panels' fields of expertise and years of experience may be seen in Tables 1 and 2. According to Table 1, the 11 professional experts have expertise ranging from 10 to 20 years. The ten lay expert panels' years of expertise varied from 10 to 23 years, based on Table 2.

Table 1 : List of Professional Expert Panels

Professional Expert Code	Expertise	Experience (Years)
P1	Psychometrics	16
P2	Psychometrics	10
P3	Psychometrics	13
P4	Efficacy and Psychometrics	16
P5	Assessment	20
P6	Assessment	18
P7	Assessment	10
P8	Educational Psychology	20
P9	Educational Psychology	20
P10	Educational Psychology	15
P11	Linguistics	10

Table 2 : List of Lay Expert Panels

Lay Expert Code	Expertise	Experience (Years)
P1	Assessment	10
P2	Educational Psychology	20
P3	Bahasa Melayu	10
P4	Pendidikan Islam	23
P5	Mathematics	15
P6	Science	13

P7	Pendidikan Islam	11
P8	English	15
P9	Arabic	12
P10	Bahasa Melayu	10

Twenty-three items are below the value of 0.429 due to a combination of eleven professional experts. A total of ten items were below the value of 0.800 through the evaluation of ten lay experts. The findings show that only three items are below the critical value of 0.429, which are items 8, 52, and 59 (table 3). It is necessary to either remove or revise items from the instrument if the expert panels cannot agree on them in a minimum agreement (DeVon, 2007). In this case, three things that fell below the threshold of 0.429 have been removed, and the researcher will revise the other items before the pilot test. Table 4 displays the item distribution following expert validation.

Table 3 : Items that need to be removed based on the type of experts (N=21)

Item Number	Item	The CVR Category Expert			Item Status
		Professional Panel	Field	Total	
		(N=11)	(N=10)	(N=21)	
		≥ 0.636	≥ 0.800	≥ 0.429	
8	I can provide assessment activities that stimulate students to produce ideas.	0.091	0.600	0.333	removed
52	I can give feedback by giving marks and written comments on the student's work.	0.273	0.400	0.333	removed
59	I can write down the student's performance if I am happy with what they have done.	-0.091	0.800	0.333	removed

Table 4 : Item distribution after verifications by experts

Construct	Sub-Construct	Number of Items	Number of revised/removed items
Teachers' Planning (12 items)	Efficacy in Identifying learning objectives	5	Six items were revised. Two items were added.
	Selection of assessment strategies	2	Three items removed
	Construction of Assessment Instruments	5	
Teachers' Use of Strategies (22 items)	Efficacy in the Use of Assessment Methods	10	Eleven items were revised.
	Implementation of Assessment Activities	6	Two items were removed.
	Use of Assessment Instruments	2	
	Self and Peer Assessment	4	
Teachers' Use of Results (13 items)	Efficacy in the Student needs	5	Eight items were revised.
	Teacher needs	5	Four items were removed.
	Giving feedback	3	
Teachers' Reporting (13 items)	Efficacy in Administration of Performance Records	4	Ten items were revised. Seven items were removed.
	Determining the Student Mastery Level	2	
	Using professional judgment	2	
	Reporting decision	5	

CONCLUSION AND RECOMMENDATION

Using Lawshe's CVR model, this study shows how to conduct a content validity analysis, a crucial step in instrument development. Expert panels reviewed and judged the instrument items. The EfiG-PBD instrument's content validity is adequate and acceptable. According to the computed CVR score, only twenty-three out of the 75 items from the 21 expert panels failed to meet the standards. However, certain things have changed because of the expert panels' review. In the end, sixty of the original 75 items were remaining. The CVR is a well-known measurement tool that measures expert agreement using statistical analysis. The decision of which information to include or remove was made clearly and appropriately.

Additionally, a thorough analysis of the psychometric properties of measurement equipment is needed. Future research should therefore support the proper validity analysis of each instrument to increase the applicability of the measurement instrument. Therefore, a pilot study will use the upgraded instrument version to examine its added psychometric properties and validity. Overall, by showing how to evaluate an instrument's content validity using Lawshe's CVR approach method, the content validity study on the EfiG-PBD instrument has added to the knowledge of instrument validity research. Since classroom assessment has been used in primary schools in Malaysia recently, the EfiG-PBD instrument was made to measure how well teachers implement it.

REFERENCES

- Arifain, S.M.K., Sulaiman, W.S.W., Badayai, A.R.A., & Abdullah, N. (2021). Pembangunan dan penilaian instrumen nilai sosial (INS) remaja. *Jurnal 'Ulwan*, 6(I), 1–21.
- Aziz, A.A, Yusof, Z.M., Mokhtar, U.A., & Jambari, D.I. (2018). Penerimgunaan sistem pengurusan dokumen dan rekod elektronik: protokol pembangunan instrumen dan kesahan kandungan menggunakan nilai ketetapan kandungan. *Jurnal Pengurusan*, 53(2018), 119–132. <https://doi.org/10.17576/pengurusan-2018-53-11>
- Bahagian Pembangunan Kurikulum. (2019). *Garis panduan pelaksanaan pentaksiran bilik darjah (PBD) bagi murid Tahap 1* (Edisi Kedua). Kementerian Pendidikan Malaysia
- Baharin, B, Adnan, M.H., Zin, M.H.M, Kamaludin, M.N., & Mansor, A.N. (2017). Gaya kepimpinan guru besar dan tahap efikasi guru. *Journal of Personalized Learning*, 2(1), 9–17. <http://spaj.ukm.my/jplearning/index.php/jplearning/article/view/38/63>
- Bandura, A. (1997). Theoretical Perspectives: the nature of human agency. In *Self-efficacy: The exercise of control* (p. 3). https://doi.org/10.1007/SpringerReference_223312
- Bond, T. G., & Fox, C. M. (2015). Applying the Rasch model: fundamental measurement in the human sciences. In *Routledge* (Vol. 3). <https://doi.org/10.4324/9781410614575>
- Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education*, 10(2), 161–180. https://doi.org/10.1207/s15324818ame1002_4
- Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and

- uses. *Educational Measurement: Issues and Practice*, 22(4), 5–12. <https://doi.org/10.1111/j.1745-3992.2003.tb00139.x>
- Bruun, J., & Evans, R. H. (2020). Network analysis of survey data to identify nonhomogeneous teacher self-efficacy development in using formative assessment strategies. *Education Sciences*, 10(3), 1–20. <https://doi.org/10.3390/educsci10030054>
- Chong, J., Mokshein, S. E., & Mustapha, R. (2021). A content validity study for vocational teachers' assessment literacy instrument (VoTAL). *International Journal of Academic Research in Business and Social Sciences*, 11(4), 868–883.
- Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). Psychological testing and assessment: An introduction to tests and measurement (8th ed.). In *McGraw-Hill*. https://doi.org/10.1007/978-1-4899-0222-1_8
- DeVellis, R. F. (2017). Scale Development: Theory and Applications. In *Sage Publications* (Vol. 4).
- DeVon, H.A., Block, M.E., Moyle-Wright, P., Ernst, D.M., Hayden, S.J., Lazzara, D.J., Savoy, S.M. & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*. 39(2), pp. 155-164.
- Faudzi, N. H. M., & Ahmad, A. (2022). Kesiediaan guru sejarah melaksanakan pentaksiran bilik darjah dalam pengajaran dan pembelajaran. *International Conference on Business Studies and Education (ICBE), March 2022*.
- Fives, H., & Barnes, N. (2020). Navigating the complex cognitive task of classroom assessment. *Teaching and Teacher Education*, 92, 1–17. <https://doi.org/10.1016/j.tate.2020.103063>
- Ghazali, N.H.C.M., Rabi, N.M., Hassan, N.M. & Wahab N.A. (2018). A confirmatory factor analysis of classroom assessment practises scale in a Malaysian context. *International Journal of Academic Research in Progressive Education and Development*, 7(3), 516–529. <https://doi.org/10.6007/IJARPED/v7-i3/4570>
- Green, F. (2019). An exploration into the value of formative assessment and the barriers associated with the implementation of formative strategies. In *Rethinking Teacher Education for the 21st Century* (pp. 203–222). <https://doi.org/10.2307/j.ctvvp3xhh.17>
- Hartell, E. (2018). Teachers' self-efficacy in assessment in technology education. In *Handbook of Technology Education* (pp. 785–800). <https://doi.org/10.1007/978-3-319-44687-5>
- Hoi, C. K. W., Zhou, M., Teo, T., & Nie, Y. (2017). Measuring efficacy sources: Development and validation of the sources of teacher efficacy questionnaire (STEQ) for Chinese teachers. *Psychology in the Schools*, 54(7), 756–769. <https://doi.org/10.1002/pits.22025>
- Jamaludin, N.A.F., Azman, M.A.D., & Zakaria, W.F.A.W. (2019). Matlamat pembangunan mapan (SDGs) dan agenda perbandaran baru 2030 dalam konteks Malaysia. *Persidangan Siswa Usuluddin, Falsafah Dan Pengajian Islam*, April, 226–236.
- Johnston, P., & Wilkinson, K. (2009). Enhancing validity of critical tasks selected for college and university program portfolios. *National Forum of Teacher Education Journal*, 19(3), 1–6.

- Lawshe, C. H. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563–575.
- Matore, M.E.E.M., Idris, H., Rahman, N.A., & Khairani, A.Z. (2017). Kesahan kandungan pakar instrumen IKBAR bagi pengukuran AQ menggunakan nisbah kesahan kandungan. *Proceeding of International Conference on Global Education V (ICGE V)*, April, 979–997.
- McMillan, J. H. (2018). *Classroom assessment: principles and practice that enhance student learning and motivation* (7th ed.). Pearson Education, Inc. <https://doi.org/August 3, 2017>
- Mokhsein, S.E. (2018). Education for sustainable development (ESD) in Malaysia: Policy, program and evaluation. *International Conference on Current Issues in Education (ICCIE 2018)*, 326, 6–12. <https://doi.org/10.2991/iccie-18.2019.2>
- Mood, M. Z. A. C., Rahman, M. A. A., Cheong, L. S., & Hussin, Z. (2017). Hubungan kecerdasan emosi terhadap efikasi sendiri guru pelatih pendidikan khas di Institut Pendidikan Guru Kampus Ilmu Khas. *Jurnal Penyelidikan Tempawan*, XXXIV, 1–15.
- Perera, H. N., Calkins, C., & Part, R. (2019). Teacher self-efficacy profiles: Determinants, outcomes, and generalizability across teaching level. *Contemporary Educational Psychology*, 58(February), 186–203. <https://doi.org/10.1016/j.cedpsych.2019.02.006>
- Rubio, D. M., Berg-weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, 27(2), 94–104.
- Russell, K. M., & Airasian, W. P. (2012). *Classroom assessment concepts and applications* (7th ed). McGraw-Hill
- Shamsudin, A., & Majid, R. M. (2018). Efikasi sendiri guru sekolah dalam hospital di Malaysia. *Journal of Social Sciences and Humanities*, 15(3), 20–28.
- Sekaran, U., & Bougie, R. (2016). Research methods for business: A skill-building approach. In *John Wiley & Sons Inc.* (7th ed.). <https://doi.org/10.1108/lodj-06-2013-0079>
- Sulaiman, A. R., & Hassan, H. (2019). Persepsi Guru Pendidikan Islam Dalam Pentaksiran Dokumen Standard Kurikulum Dan Pentaksiran Sekolah Rendah. *International Journal of Humanities Technology and Civilization*, 2, 13-19.
- Surip, N.A., Razak, K.A., & Tamuri, A.H. (2019). Kesahan dan kebolehpercayaan instrumen pengurusan pendidikan dakwah di sekolah. *Journal of Islamic, Social, Economics, and Development (JISED)*, 4(19), 108–118.
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783–805. [https://doi.org/10.1016/S0742-051X\(01\)00036-1](https://doi.org/10.1016/S0742-051X(01)00036-1)
- Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, 45(3), 197–210. <https://doi.org/10.1177/0748175612440286>

Zainal, M.A., Matore, M.E.E.M., Musa, W.N.W., & Hashim, N.H. (2020). Kesahan kandungan instrumen pengukuran tingkah laku inovatif guru menggunakan kaedah nisbah kesahan kandungan (CVR). *Akademika*, 90(3), 43–54.