

*International Journal of Learning, Teaching and Educational Research*  
 Vol. 24, No. 10, pp. 493-508, October 2025  
<https://doi.org/10.26803/ijlter.24.10.23>  
 Received Jul 17, 2025; Revised Sept 10, 2025; Accepted Sept 15, 2025

## Face Validity Assessment of Malaysian Teachers' Global Competency Instrument using Face Validity Index Analysis from Potential Test-takers' Perspective

Siti Nurul Aqiedah Ibrahim\* 

Faculty of Education, The National University of Malaysia (UKM),  
 Bangi, Selangor, Malaysia

Mohd Effendi@Ewan Mohd Matore 

Research Centre of Education Leadership and Policy, Faculty of Education,  
 The National University of Malaysia (UKM), Bangi, Selangor, Malaysia

**Abstract.** In the VUCA era (Volatile, Uncertain, Complex, Ambiguous), global competency is a key proficiency that teachers must possess to produce students who can think openly, act responsibly, and interact across cultural boundaries. Developing an instrument to measure the level of global competency among teachers in Malaysia requires validity, particularly from non-expert groups such as potential test-takers. This study reported the findings of a quantitative survey design to evaluate the face validity of the global competency instrument by ten teachers who represented the target population of potential test takers. The clarity and comprehensibility of the items were assessed using a four-point Likert scale and analyzed with the face validity index (FVI) through the calculation of item-level FVI (I-FVI), scale-level FVI average (S-FVI/Ave), and scale-level FVI universal agreement (S-FVI/UA). A total of 68 items achieved an I-FVI value of  $\geq 0.83$ , while the remaining eight items recorded values between 0.70 and 0.80. The overall S-FVI/Ave and S-FVI/UA values were 0.97 and 0.84, respectively, indicating a good level of clarity and comprehensibility of the items. This study emphasized the crucial role of face validity in ensuring the clarity, comprehensibility, and usability of an instrument by end-users. The systematic use of the FVI approach enhances transparency and precision in evaluating perceptions of item clarity and comprehensibility, hence supporting the recommendation of FVI as a fundamental procedure in instrument development studies, before proceeding to more complex stages of content and construct validity. It also underlines the importance of face validity evaluation from the perspective of potential test takers.

---

\*Corresponding author: Mohd Effendi @ Ewan Mohd Matore; [effendi@ukm.edu.my](mailto:effendi@ukm.edu.my)

**Keywords:** face validity index; instrument development; global competency; teachers; potential test takers

## 1. Introduction

In today's VUCA world where change is no longer a distant possibility and global challenges have become increasingly demanding, almost all sectors are affected, including the education sector. As educators, it is essential to equip ourselves with the necessary skills to develop globally competitive students. Teachers play a vital role in equipping students with global competency to face this fast-paced world (Putman & Byker, 2020). Global competency refers to a teacher's ability to nurture and develop students with noble values, an appreciation and openness towards cultural diversity, an understanding of multiple perspectives, and the ability to communicate and collaborate with individuals from different backgrounds to ensure global sustainability (Ibrahim & Mohd Matore, 2024).

Although global competency has been identified as a key element in building teachers' capacity to meet the challenges of 21st-century education, previous studies reveal a lack of contextually developed instruments to assess the level of global competency among teachers in Malaysia (Jiaxin et al., 2024; Yaccob et al., 2022). In the era of the Fourth Industrial Revolution (4IR), the global education system is undergoing significant changes due to the emergence of new technologies such as artificial intelligence (AI), the Internet of Things (IoT), virtual reality (VR), and innovative digital learning (Masdoki et al., 2021).

This development not only demands technological proficiency among educators but also reinforces the need for global competency as a prerequisite for teachers to interact, adapt, and perform their roles in increasingly open, interconnected, and cross-cultural educational environments. Xuanyun et al. (2024) support the notion that the effectiveness of technology integration in education requires not only technical skills but also intercultural communication, collaboration, and ethical digital citizenship.

These are core elements of global competency and essential for navigating today's digitally interconnected classrooms. Hussain et al. (2025) reported that technology-driven transformations, such as the shift from traditional subjective evaluations to objective, AI-enhanced systems, leveraging big data analytics, are useful for predicting and improving faculty performance. Institutions adopting these approaches aim to develop more transparent, evidence-based, and dynamic evaluation frameworks that align with technological advancements in education.

According to Mansilla and Jackson (2011), global competency enables teachers to understand transboundary issues, communicate effectively in multicultural settings, and instill a sense of global responsibility in their students. Follow-up research by Tichnor-Wagner et al. (2019) emphasized that globally competent teachers practice teaching that incorporates intercultural dispositions,

international knowledge, communication skills, and collaboration within the classroom. Furthermore, Aydarova and Marquardt (2016) noted that teacher education should internationalize its curriculum to meet the demands of 21st-century globalized education.

Therefore, the need arises to develop a specific, contextually relevant instrument to assess teachers' global competency aligns with the aspiration to produce educators capable of playing an effective role within an educational landscape shaped by technological advancement and globalisation. Moreover, Nadmilail et al. (2022) emphasized that broad traits are enduring features of humans that persist across time and contexts, as well as incorporating several lower-level personality traits. In contrast, narrow traits are precise personality dispositions toward specific behaviors, theoretically defined constructs that are continuous and accountable for more behavioral criteria in terms of diversity and specificity.

Sürücü and Maslakçı (2020) reported that face validity involves the researchers' subjective judgment based on their perceptions, ideas, and intuition regarding the effectiveness of an instrument. Face validity is often overlooked in scale development, despite being a crucial early step to ensure the clarity and appropriateness of items for potential test takers (Allen et al., 2023). Face validity assessment helps ensure that each item not only appears clear but is also consistently understood by respondents.

Without a high level of clarity, item interpretation can become ambiguous, thus undermining the reliability and accuracy of research findings. In this regard, quantitative approaches such as the FVI offer a systematic and transparent mechanism for assessing face validity based on empirical data (Yusoff, 2019). Moreover, although FVI has been widely applied in the medical field, its use in the education sector remains limited, despite its significant potential to strengthen the educational instrument development process.

## **2. Objectives**

The objectives of this study were as follows:

1. To assess the face validity of Malaysian teachers' global competency instrument using FVI analysis; and
2. To identify the role of potential test-takers' perceptions in evaluating the clarity and comprehensibility of the items.

## **3. Literature Review**

### **3.1 Global Competency**

Global competency is a multidimensional ability that encompasses knowledge, attitudes, values, and skills enabling individuals to understand and evaluate local, international, and intercultural issues. This competency includes the ability to appreciate diverse perspectives, interact effectively and respectfully across cultural backgrounds, and actively contribute to global sustainability. In other words, global competency not only involves the ability to know and understand the world but also the ability to engage actively and responsibly within a diverse global society (OECD, 2018; Parmigiani et al., 2022).

In the educational context, global competency plays a vital role in developing the holistic capacity of both teachers and students, including intellectual, emotional, spiritual, and adversity-quotient aspects. By fostering these dimensions, it prepares learners to address global issues, engage in meaningful intercultural interactions, and contribute to inclusive and sustainable communities (Jiaxin et al., 2024). This indirectly supports the development of holistic students and teachers. This also aligns with the aims of the National Education Philosophy (NEP), which aspires to produce balanced and harmonious individuals in terms of physical, emotional, spiritual, and intellectual development (Mohd Matore et al., 2018).

Studies have shown that mastering both emotional and spiritual quotients has a significant relationship with improving the adversity quotient, which refers to an individual's ability to overcome challenges and remain competitive in demanding environments (Mohd Matore et al., 2018). Global competency indirectly strengthens these four dimensions by fostering intercultural understanding, communication skills, self-awareness, and high levels of adaptability among teachers and students.

Therefore, strong global competency among teachers is not only crucial for managing culturally diverse classrooms but also contributes to shaping quality individuals who can compete holistically at the international level, in line with the aspirations of the NEP. A sustainable level of global competency among teachers is essential in nurturing students capable of competing at the international level (Jiaxin et al., 2024), ultimately producing proactive global citizens.

The development of global competency among teachers in Malaysia is essential to align educational practices with the demands of the 21st century. Preliminary findings reveal significant gaps in research and implementation, highlighting the need for comprehensive strategies to enhance teachers' global competency. Although extensive research on teachers' global competency in Malaysia remains limited, early findings indicate an urgent need to ensure that teachers achieve a level of global competency that meets the expectations of global education and adequately prepares students for a globalized world (Guler, 2024; Yacob et al., 2022).

At both national and international levels, research focusing on the global competency of in-service teachers remains underexplored (Ibrahim & Mohd Matore, 2024). This lack of empirical evidence makes it difficult for stakeholders, such as policymakers and educational institutions, to assess and strengthen teachers' global competency comprehensively within the local context.

### **3.2 Face Validity**

Face validity is a fundamental aspect in determining the appropriateness and clarity of an instrument's items from the perspective of potential test takers. It refers to the extent to which a test or an instrument appears suitable, clear,

relevant, and sensitive to the context and characteristics of the target respondents before it is used in large-scale studies (Allen et al., 2023). Although face validity is often considered the most basic and subjective form of validity, it remains an important initial indicator of an instrument's item quality, particularly in social, educational, and psychological research (Zhang & Garcia, 2023).

Face validity assessment is conducted by collecting responses from individuals who share similar demographic characteristics to the actual study sample (Baco & Ishak, 2020). According to Mohd Matore et al. (2021), developing items with a clear logical structure, careful word selection, and contextually appropriate language is essential to ensure the overall validity and quality of the measurement instrument. Therefore, the involvement of the target teachers as reviewers is crucial. In Malaysia, several previous studies have also implemented face validity assessments in the development of instruments, albeit through differing methods.

For instance, Mohd Matore et al. (2021) conducted a face validity assessment of item clarity and appropriateness by consulting expert reviews before further psychometric testing in the development of the Malaysian Youth Adversity Quotient Instrument (MY-AQi). Similarly, the study by Kamaruddin and Mohd Matore (2021), which developed the 10 IB Learner Profile Instrument (10IBLP-I), implemented face validity assessment through expert reviews involving academics and experienced teachers to ensure language accuracy, clarity of terms, and appropriateness of item content. However, both studies relied solely on qualitative approaches through expert discussions or narrative reviews without quantitative calculation of face validity indices.

In contrast, this study adopts a more systematic and evidence-based approach by employing FVI calculations to assess the instrument's face validity. Unlike expert evaluations, the FVI method in this study involves respondents from the actual target group, namely potential test takers, to assess the clarity and comprehensibility of the instrument items. The calculation of FVI at the item level (I-FVI) and for the overall scale (S-FVI/Ave and S-FVI/UA) provides objective numerical indicators of item face validity, allowing for more transparent, data-driven decisions regarding item elimination or improvement (Baharuddin et al., 2024). This approach aligns with current recommendations in instrument development research, which emphasize the importance of involving potential test takers in the initial review process (Polit & Beck, 2006; Zamanzadeh et al., 2015).

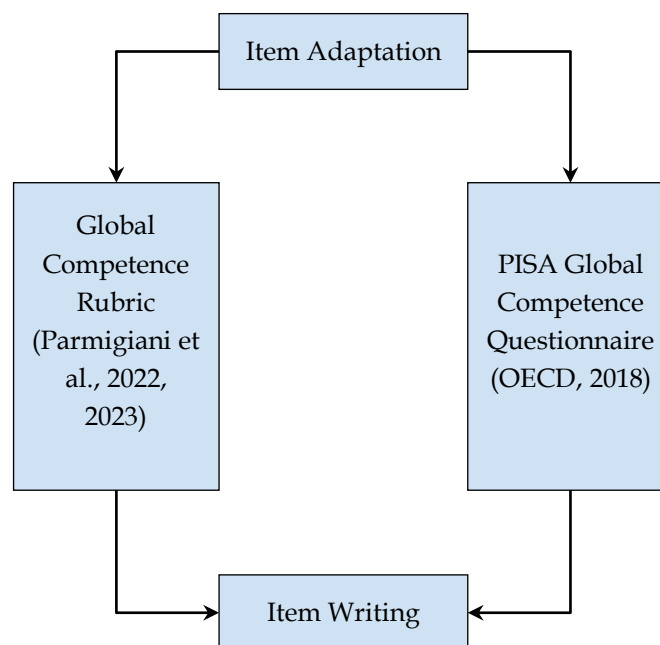
Therefore, the use of the FVI method in this study not only strengthens the content accuracy of the instrument from the end-users' perspective but also complements the item quality assessment process before further psychometric testing. This approach is seen as an enhancement to the face validity review methods applied in previous Malaysian studies, which primarily focused on qualitative expert evaluations.

## 4. Methodology

This study employed a quantitative survey study by applying the Malaysian Educators' Global Competency Instrument (MAGNET) questionnaire to test for its face validity from the perspective of potential test takers. The main benefits of using surveys include scalability and cost-effectiveness for collecting data from large populations, speed and efficiency in data collection, the ability to conduct robust statistical analysis because of structured data, and the capacity to collect a wide range of information on different topics such as opinions, attitudes, and behaviors (Pressbooks, 2025). The face validity of this paper is one of the main efforts to enhance the quality of items measured. Establishing face validity in this study serves as a key effort to improve the quality of the measured items, whereas reliability will be examined during the pilot testing stage. The identities of the potential test takers remained anonymous.

### 4.1 Developing Instrument

In developing this instrument, a literature review process was conducted to identify the definition and constructs of global competency. Consequently, operational definitions and conceptualizations for each construct were established, and suitable items were selected to meet the requirements of the constructs. The items in this instrument were adapted from the Global Competence Rubric by Parmigiani et al. (2022, 2023) and the Programme for International Student Assessment (PISA) Global Competence Questionnaire by the OECD (2018). The adaptation process is illustrated in Figure 1:



**Figure 1: The Process of Item Adaptation**

In the early stages of instrument development, a total of 73 items were systematically selected from two established sources: 32 items from the Global Competence Rubric and 41 items from the PISA Global Competence Questionnaire. The selection process was guided by a conceptual framework

encompassing four key constructs: Self Awareness, Global Awareness, Attitudes and Values, and Skills. Each item was carefully reviewed for relevance and theoretical alignment, then categorized accordingly. This results in 16 items under Self-Awareness, 12 under Global Awareness, 21 under Attitudes and Values, and 24 under Skills.

To ensure content validity, the items underwent expert evaluation by a panel of subject matter specialists in global competencies, as well as measurement and evaluation experts. These experts assessed the clarity, relevance, and appropriateness of each item in relation to the defined constructs and the target population. The adaptation process primarily focused on ensuring conceptual and contextual alignment, laying the groundwork for a theoretically sound and contextually appropriate instrument for assessing global competence.

The Face Validity Form – Potential Test Takers’ Review, developed and used in this study, contained an explanation of the research and guidelines for completing the form. The explanation included the purpose and objectives of the study, while the response guidelines covered the criteria being assessed and the measurement scale. The form contained explicit instructions directing reviewers to evaluate each item in terms of clarity and comprehensibility. Reviewers assigned their ratings using a four-point Likert scale, as recommended by Yusoff (2019): 1 = Very unclear, 2 = Unclear, 3 = Clear, and 4 = Very clear.

#### **4.2 Potential Test Takers**

A total of 10 teachers from fully residential schools (SBP) under the Residential School Management Division were selected to review the face validity of this teacher global competency measurement instrument. These teachers were drawn from the same population that the instrument intends to measure, ensuring that the face validity sample possessed the same demographic characteristics as the actual study respondents. They represented teachers from different genders, age groups, subjects taught, and variations in teaching experience span. From these 10 test takers, two are males and eight are females; nine are Malays and one is Chinese. They have served for between one to 24 years and are aged between 28 and 43 years old.

#### **4.3 Data Collection and Analysis**

The potential test takers were contacted via WhatsApp to serve as reviewers for the Malaysian Teacher Global Competency Instrument. After obtaining their consent, the Google Forms link for the Potential Test Takers Face Validity Form was shared via WhatsApp. The reviewers evaluated the instrument items through this link.

The face validity findings from the candidate review were analyzed using the FVI approach, a quantitative method that assesses potential test-takers' perceptions of item clarity. There are two main types of FVI analysis: the item-level FVI (I-FVI) (Yusoff, 2019), which evaluates the clarity of each item individually, and the scale-level FVI (S-FVI), which provides an overall picture of the scale's clarity. For S-FVI, analysis was conducted by construct as well as in aggregate for all items. Two indices were used: the scale-level FVI average (S-

FVI/Ave), representing the average clarity score across items, and the scale-level FVI universal agreement (S-FVI/UA), indicating the percentage of respondents who unanimously provided high scores. The combination of these three indices allowed for a systematic, comprehensive, and evidence-based assessment of face validity.

The formulas for calculating I-FVI, S-FVI/Ave, and S-FVI/UA are as follows:

*I-FVI Calculation Formula*

$$I - FVI = \frac{e}{N}$$

Where:

e = number of respondents who gave a score of 3 or 4 (Clear or Very clear)

N = total number of respondents

*S-FVI/Ave Calculation Formula*

$$S - FVI/Ave = \frac{\sum I - FVI}{k}$$

Where:

$\sum I - FVI$  = total I-FVI value for all items

k = total number of items

*S-FVI/UA Calculation Formula*

$$S - FVI/UA = \frac{ne}{k}$$

Where:

ne = number of items given a score of 3 or 4 by all respondents

k = total number of items

The FVI value ranged from 0 to 1, with a score of 1 indicating perfect clarity and comprehensibility. An acceptable I-FVI value is 0.83 or above, based on the recommendations of Mohamad Marzuki et al. (2018). Items scoring below 0.83 should be eliminated as they are considered unclear or poorly understood (Mohamad Marzuki et al., 2018; Yusoff, 2019). For S-FVI/Ave, a value of 0.9 or above is considered acceptable (Dalawi et al., 2023). As for S-FVI/UA, no studies have explicitly defined an acceptable value. However, Yusoff (2019) reported that a value of 0.83 is considered satisfactory. Therefore, in this study, an S-FVI/UA value of 0.83 or higher was deemed acceptable.

## 5. Results and Findings

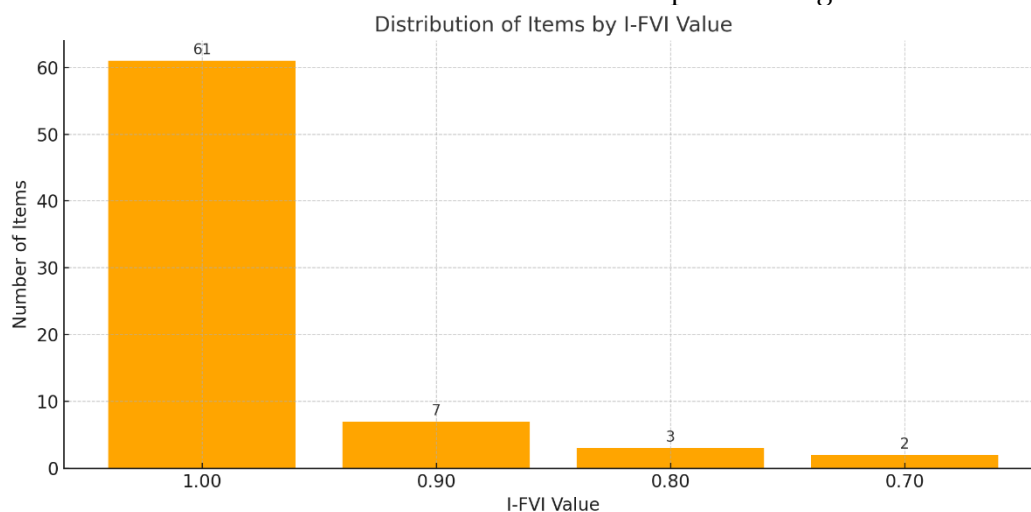
Based on the FVI analysis calculated using Google Sheets, the I-FVI values for each item are shown in Table 1:

**Table 1: I-FVI values for each item representing the constructs of Self-Awareness (KD), Global Awareness (KG), Attitudes & Values (SN), and Skills (KM)**

I-FVI Value	Item	No. of item	Decision

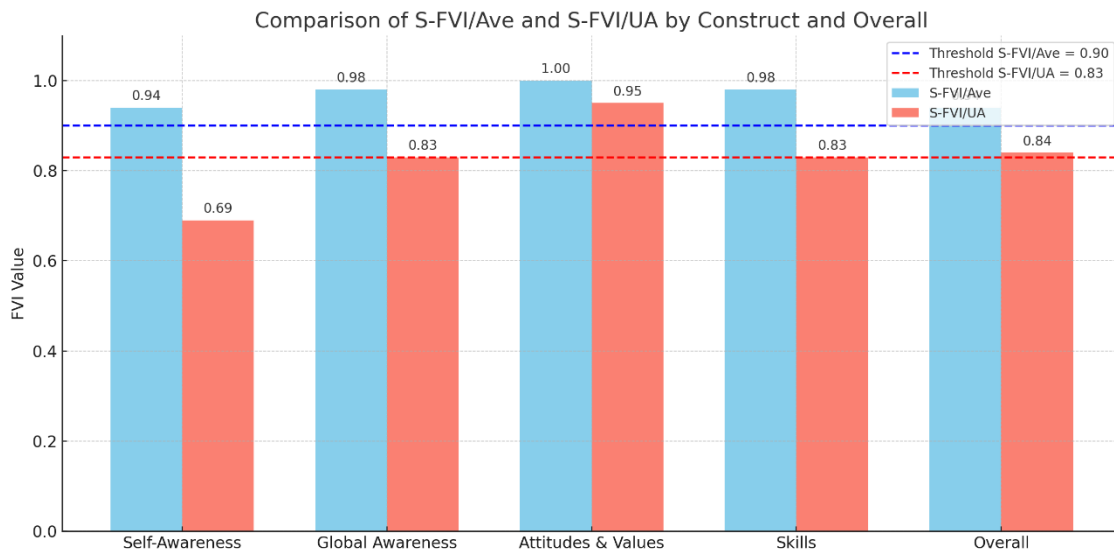
1.00	KD1, KD2, KD3, KD4, KD5, KD6, KD7, KD8, KD9, KD13, KD14, KG1, KG2, KG3, KG4, KG5, KG6, KG7, KG10, KG11, KG12, SN1, SN2, SN3, SN5, SN6, SN6, SN7, SN8, SN9, SN10, SN11, SN12, SN13, SN14, SN15, SN16, SN17, SN18, SN19, SN20, SN21, KM1, KM2, KM3, KM4, KM6, KM7, KM8, KM9, KM10, KM12, KM13, KM14, KM16, KM17, KM18, KM19, KM21, KM22, KM23, KM24	61	Retained
0.90	KD10, KD11, KD15, KG8, SN4, KM15, KM20	7	Retained
0.80	KG9, KM5, KM11,	3	Removed
0.70	KD12, KD16	2	Removed

The total number of items for each I-FVI value is depicted in Figure 2:



**Figure 2: Distribution of Items by I-FVI Value**

Based on Table 1 and Figure 2, a total of 68 items recorded an I-FVI value greater than 0.83, while five items recorded an I-FVI value below 0.83. From these I-FVI values, the S-FVI/Ave and S-FVI/UA values for each construct and the overall instrument were obtained, as shown in Figure 3:



**Figure 3: S-FVI/Ave and S-FVI/UA Values by Construct and Overall**

Figure 3 illustrates the S-FVI/Ave and S-FVI/UA values by construct and overall. Both indices exceeded the recommended thresholds of 0.90 for S-FVI/Ave and 0.83 for S-FVI/UA, with the exception of the S-FVI/UA value for the Self-Awareness construct, which fell below the cut-off.

## 6. Discussion

The findings of this study directly addressed the main objectives, namely to evaluate the face validity of the teacher global competency instrument using the FVI approach, and to identify the role of potential test-takers' perspectives in assessing item clarity and comprehensibility.

### 6.1 Objective 1: To evaluate the face validity of the teacher global competency instrument using FVI analysis

The findings indicated that the use of the FVI is an effective and systematic approach to assess face validity quantitatively. The majority of items in this instrument, namely 68 out of 73 items, exceeded the recommended I-FVI threshold of  $\geq 0.83$  (Mohamad Marzuki et al., 2018). The findings indicated that the majority of items were effectively understood by the potential test-takers, reflecting satisfactory clarity and comprehensibility. The overall values of S-FVI/Ave (0.97) and S-FVI/UA (0.84) also exceeded the minimum thresholds ( $\geq 0.90$  and  $\geq 0.83$ , respectively), further supporting the clarity and comprehensibility of the instrument as a whole.

However, five items, namely KG9, KM5, KM11, KD12, and KD1, did not meet the minimum I-FVI threshold, as shown in Table 2:

**Table 2: Items to be Refined**

No	Item	I-FVI Value	Refinement
1	KG9	0.80	Language refinement based on expert linguist recommendation
2	KM5	0.80	Language refinement based on expert linguist recommendation
3	KM11	0.80	Language refinement based on expert linguist recommendation
4	KD12	0.70	Language refinement based on expert linguist recommendation
5	KD16	0.70	Language refinement based on expert linguist recommendation

Although no qualitative feedback was collected from respondents, the low FVI values for these five items suggested weaknesses in clarity and comprehensibility. This aligns with the recommendations by Dalawi et al. (2023) and Yusoff (2019), who demonstrated that the FVI can quantitatively identify potentially problematic items based on clarity and comprehensibility ratings, especially when used alongside qualitative feedback from the intended respondents.

Based on the operational method underpinned by Yusoff (2019), all five items underwent language refinement, which involved improving sentence structure, selecting more appropriate vocabulary, and reducing the use of potentially confusing terms that may hinder respondents' understanding and affect the validity of the instrument (DeVellis, 2017). The refinements were carried out under the guidance of linguistic experts to enhance item clarity and ensure contextual relevance to the target population, consistent with the recommendation to involve experts in establishing content validity (Polit & Beck, 2006). This is similar to the steps taken by a previous research study conducted by Dalawi et al. (2019).

In short, the use of FVI in this study has proven to be an effective quantitative approach for identifying and improving weaknesses in item clarity and comprehension, thereby enhancing the overall quality of the instrument before it is used in large-scale studies.

## **6.2 Objective 2: To identify the role of potential test-takers' perceptions in assessing item clarity and comprehensibility in face validity**

In instrument development, the perception of potential test takers regarding item clarity and understanding is a crucial aspect that cannot be overlooked. Involving potential test takers directly in the assessment process ensures that the instrument is suitable for real-world contexts (Baharuddin et al., 2024). Moreover, it provides early insights into the strengths and weaknesses of the items developed (Dalawi et al., 2023; Yusoff, 2019). This aligns with the contention of Allen et al. (2023), who emphasize that user-end face validity evaluation is a critical yet often neglected component in scale development, despite its importance in determining instrument quality.

In this study, teachers from the target group were selected as representatives of potential test-takers to evaluate the clarity and comprehensibility of items in the teacher global competency instrument. FVI analysis revealed that the target group's perception of the developed items was generally satisfactory. The S-FVI/Ave score of 0.97 and S-FVI/UA score of 0.84 surpassed the minimum thresholds suggested in previous studies (Dalawi et al., 2023; Yusoff, 2019), indicating that the instrument is considered appropriate for use based on test-takers' evaluations.

However, more detailed analysis based on test-taker perceptions identified five items (KG9, KM5, KM11, KD12, and KD16) with I-FVI values ranging between 0.70 and 0.80. These findings demonstrated the critical role of potential test-takers' perceptions in identifying items that are unclear or potentially misleading. This supports the views of Awan et al. (2022), who stress that test-taker perception is a vital data source in determining whether an instrument can be confidently and effectively applied in real settings.

Based on the quantitative data gathered from test-taker perceptions, improvements were made focusing on the linguistic structure, sentence clarity, and terminology of the five items. The refinement process was based on expert linguistic recommendations. The items were retained, as the identified issues were limited to surface presentation rather than construct content. Most of the previous research only focuses on content specifically and not fully on face validity. The focus of this paper supports the content validity by providing empirical evidence of face validity to enhance the quality of content validity. The FVI applied was suited to the objective and the results to empower the items as well.

These findings affirmed that potential test-taker perception is an important indicator that should be considered in the development and refinement of instruments. The test-taker quality as a rater can also be improved, as mentioned by Mohd Noh and Mohd Matore (2022), who highlighted that raters' ratings and teaching experience significantly affect rating severity in speaking assessments, while training experience may have less impact. They emphasize the importance of selecting qualified raters to ensure assessment fairness and suggest that unexamined rater biases could potentially impact test takers' psychological well-being.

The FVI analysis in this study not only provided objective quantitative data but also served as a systematic means of integrating the perspectives and experiences of the target group to improve item quality. This approach is in line with the work of Baharuddin et al. (2023), who emphasized the importance of involving test-takers directly in the validation process, especially in ensuring item clarity and comprehension in educational contexts. Furthermore, the emphasis on face validity evaluation in this study supports the conclusions of Ibrahim and Mohd Matore (2025), who argued that quality instrument development must involve comprehensive validation processes, including both face and content validity assessments that are systematic and contextually

grounded. Such an approach is essential to ensure that developed instruments are truly appropriate, valid, and effective in the intended setting.

### **6.3 Implications and Recommendations**

The FVI analysis conducted in this study indicates a high level of clarity and comprehensibility from the perspective of the potential test takers, thereby confirming the preliminary usability of this instrument in the local educational context. These findings underscore the importance of involving potential test-takers directly in the instrument evaluation process, especially for abstract constructs such as global competency, which are strongly shaped by cultural factors, professional experience, and educational context. The quality of the test takers as raters should be considered carefully, with reference to their experience and rating performance. Additionally, the rater bias that may impact the psychological well-being of certain groups of test takers can be addressed.

Furthermore, the systematic application of the FVI approach in this study demonstrates that face validity assessment is not merely a formality but a critical component that contributes significantly to improving item quality before proceeding to more complex stages of content and construct validation. This approach also enhances transparency and credibility in instrument development, as it provides clear quantitative evidence of item clarity from the perspective of intended end-users.

This study supports the recommendation that FVI should be established as a standard procedure in the development of educational instruments, particularly those designed to assess complex competencies such as global competency. These implications not only strengthen more systematic and transparent research practices but also have the potential to improve the effectiveness of measuring teachers' global competency in the effort to produce educators who can meet the demands of a VUCA world.

### **6.4 Limitations**

However, limitations such as the small sample size and an evaluation that focused solely on clarity, without accounting for cultural factors, usage context, or respondents' interpretation levels, may affect the generalizability of the findings. Therefore, it is recommended that future studies not only involve a larger and more diverse group of respondents, but also explore more holistic and innovative approaches to face validity evaluation. These may include integrating both quantitative and qualitative methods, utilizing technologies such as automated linguistic analysis to detect language ambiguity, or developing more comprehensive face validity indices that encompass cultural and contextual elements.

## **7. Conclusion**

The use of the FVI proved to be an effective, systematic, and evidence-based approach in evaluating the face validity of the Teacher Global Competency Instrument. The majority of items demonstrated good clarity and comprehensibility as perceived by the potential test takers. These findings

underscore the importance of involving test takers in the initial stages of instrument evaluation as a key step in enhancing item clarity and comprehension before broader application. This study provides a meaningful early contribution toward ensuring a more valid, reliable and contextually appropriate teacher global competency instrument for the Malaysian education system. At the same time, it supports the need to strengthen and modernize face validity evaluation methods to be more transparent, systematic, and user-focused, in line with current developments in educational instrument design.

## 8. Acknowledgments

We are grateful for the assistance that the Federal Training Prize (HLP) scholarship from the Ministry of Education (MOE) has given us. This research was funded by the Faculty of Education, Universiti Kebangsaan Malaysia (UKM) with grant number GG-2024-024 (Program STEM dan Minda). We acknowledge the supervisor's and Kumpulan Penyelidikan Universiti (KPU), UKM for the knowledge and insights, which greatly aided this study. We would also like to express our gratitude to our colleagues in the Measurement and Evaluation course for their support and helpful criticism, which has helped to make this manuscript better.

## 9. References

- Allen, M. S., Robson, D. A., & Iliescu, D. (2023). Face validity: A critical but ignored component of scale construction in psychological assessment [Editorial]. *European Journal of Psychological Assessment* 39(3), 153–156. <https://doi.org/10.1027/1015-5759/a000777>
- Awan, W. A., Ameen, K., & Soroya, S. H. (2022.) Content validity and reliability of an instrument to investigate serendipitous retrieval of information: Information encountering and its subsequent keeping and sharing behaviour scale (IEKSBS). *Journal of Information Science* 51(1), 31-39. <https://doi.org/10.1177/01655515221128756>
- Aydarova, E., & Marquardt, S. K. (2016). The global imperative for teacher education: Opportunities for comparative and international education. *FIRE: Forum for International Research in Education*, 3(1).
- Baco, S., & Ishak, M. Z. (2021). Kesahan dan kebolehpercayaan instrumen penilaian sendiri pembelajaran ungkapan algebra tingkatan dua. *Malaysian Journal of Social Sciences and Humanities (MJSSH)* 6(11), 127–137. <https://doi.org/10.47405/mjssh.v6i11.1136>
- Baharuddin, I. H., Ismail, N., Naing, N. N., Ibrahim, K., Yasin, S. M., & Patterson, M. S. (2024). Content and face validity of Workplace COVID-19 Knowledge & Stigma Scale (WoCKSS). *BMC Public Health* 24, 83. <https://doi.org/10.1186/s12889-023-17614-3>
- Dalawi, I., Isa, M. R., Xin, W.C., Azhar, Z. I., & Aimran, N. (2023). Development of the Malay Language of understanding, attitude, practice and health literacy questionnaire on COVID-19 (MUAPHQ C-19): Content validity & face validity analysis. *BMC Public Health* 23, 1636. <https://doi.org/10.1186/s12889-023-16044-5>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE Publications, Inc.
- Guler, N. (2024). Global competency for teacher educators: Exploring the needs of teacher education faculty in a professional development project. *Teacher Development*, 29(2), 324–342. <https://doi.org/10.1080/13664530.2024.2399075>

- Hussain, N. A. S., Mohd Matore, M. E. E., & Hamzah, M. I. M. (2025). Performance evaluation for educators in higher education from bibliometric analysis views. *International Journal of Learning, Teaching and Educational Research*, 24(5), 283–303. <https://doi.org/10.26803/ijlter.24.5.15>
- Ibrahim, S. N. A., & Mohd Matore, M. E. E. (2025). The validity of the Malaysian teachers' global competency level instrument using Cohen kappa, content validity ratio and content validity index analyses. *Journal of Applied Science, Engineering, Technology, and Education*, 7(1), 82-93. <https://doi.org/10.35877/454RI.asci3848>
- Ibrahim, S. N. A., & Mohd Matore, M. E. E. (2024). How does global competency look like in Malaysian teachers? Issues and suggestions. *International Journal of Academic Research in Progressive Education and Development* 13(2), 837-844. <https://doi.org/10.6007/ijarped/v13-i2/21375>
- Jiaxin, G., Huijuan, Z., & Hasan, M. H. (2024). Global competence in higher education: A ten-year systematic literature review. *Frontiers in Education* 9, Article 1404782. <https://doi.org/10.3389/feduc.2024.1404782>
- Kamaruddin, M., & Mohd Matore, M. E. E. (2021). Development and validation of psychometric properties of the 10 IB learner profile instrument (10IBLP-I): A combination of the Rasch and classical measurement model. *International Journal of Environmental Research and Public Health*, 18(12), 6455. <https://doi.org/10.3390/ijerph18126455>
- Mansilla, V. B., & Jackson, A. (2011). *Educating for global competence: Preparing our youth to engage the world*. Asia Society.
- Masdoki, M., Din, R., & Mohd Matore, M. E. E. (2021). Teaching 4.0 competency in higher learning institutions: A systematic mapping review. *International Journal of Learning, Teaching and Educational Research*. 20, 217-231. <https://doi.org/10.26803/ijlter.20.10.12>
- Mohamad Marzuki, M. F., Yaacob, N. A., & Yaacob, N. M., (2018). Translation, cross-cultural adaptation and validation of system usability scale (Malay Version) questionnaire for the assessment of mobile application. *JMIR Human Factors* 5(2), Article e10308. <https://doi.org/10.2196/10308>
- Mohd Matore, M. E. E., Khairani, A. Z., Maat, S. M., Ahmad, N. A., & Mohd Matore, E. R. (2018). The influence of intellectual quotient (IQ), emotional quotient (EQ) and spiritual quotient (SQ) against adversity quotient (AQ) on polytechnic students in Malaysia. *Journal of Engineering Science and Technology, Special Issue on ICETVESS, 13*(Special Issue on ICETVESS 2017), 83–91.
- Mohd Matore, M. E. E., Zainal, M. A., Mohd Noh, M. F., Khairani, A. Z., & Abd Razak, N. (2021). The development and psychometric assessment of Malaysian Youth Adversity Quotient Instrument (MY-AQi) by combining Rasch model and confirmatory factor analysis. *IEEE Access*. 9, 13314–13329. <https://doi.org/10.1109/ACCESS.2021.3050311>
- Mohd Noh, M. F., & Mohd Matore, M. E. E. (2022). Rater severity differences in English language as a second language speaking assessment based on rating experience, training experience, and teaching experience through many-faceted Rasch measurement analysis. *Frontiers in Psychology*, 13, 941084. <https://doi.org/10.3389/fpsyg.2022.941084>
- Nadmilail, A. I., Mohd Matore, M. E. E., & Maat, S. M. (2022). Measurement of non-academic attributes in the situational judgment test as part of school teacher selection: A systematic literature review. *International Journal of Learning, Teaching and Educational Research*, 21(5), 263–280. <https://doi.org/10.26803/ijlter.21.5.14>

- Organisation for Economic Co-operation and Development (OECD). (2018). *PISA 2018 global competence*. Retrieved from <https://www.oecd.org/en/topics/sub-issues/global-competence/pisa-2018-global-competence.html>
- Organisation for Economic Co-operation and Development (OECD). (2018). *Preparing our youth for an inclusive and sustainable world: The OECD PISA Global Competence Framework*. OECD Publishing.
- Parmigiani, D., Jones, S. L., Kunnari, I., & Nicchia, E. (2022). Global competence and teacher education programmes: A European perspective. *Cogent Education*, 9(1), Article 2022996, 1–17. <https://doi.org/10.1080/2331186X.2021.2022996>
- Parmigiani, D., Bar Nir, A., Ferguson-Patrick, K., Forkosh-Baruch, A., Heddy, E., Impedovo, M. A., et al. (2023). Assessing the development of global competence in teacher education programmes: Internal consistency and reliability of a set of rubrics. *Higher Education Pedagogies*, 8(1), Article 2216190, 1–30. <https://doi.org/10.1080/23752696.2023.2216190>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in nursing & health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Pressbooks. (2025, 10 September). 11.2 *Strengths and weaknesses of survey research*. In *Scientific Inquiry in Social Work*. <https://pressbooks.pub/scientificinquiryinsocialwork/chapter/11-2-strengths-and-weaknesses-of-survey-research/>
- Putman, S. M., & Byker, E. J. (2020). Global citizenship 1-2-3: Learn, think, and act. *Kappa Delta Pi Record* 56(1), 16–21. <https://doi.org/10.1080/00228958.2020.1696088>
- Sürücü, L. & Maslakçı, A. (2020). Validity and reliability in quantitative research. *Business & Management Studies: An International Journal* 8(3), 2694–2726.
- Tichnor Wagner, A., Parkhouse, H., Glazier, J., & Cain, J. M. (2019). *Becoming a globally competent teacher*. ASCD
- Xuanyun, Z., Noh@Seth, N. H., & Pairan, M. R. (2024). Global perspectives on teacher ICT competencies: Diversity and commonalities in eight representative frameworks. *Journal of e-Learning and Knowledge Society*, 20(2), 42-54. <https://doi.org/10.20368/1971-8829/1135977>
- Yacob, N. S., Md. Yunus, M., & Hashim, H. (2022). Globally competent teachers: English as a second language teachers' perceptions on global competence in English lessons. *Frontiers in Psychology*, 13, 925160. <http://doi.org/10.3389/fpsyg.2022.925160>
- Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal* (11), 49–54. <https://doi.org/10.21315/eimj2019.11.2.6>
- Zhang, Z., & Garcia, L. (2023). Examining dimensionality and validity of the academic integrity survey instrument. *Journal of Education and Development* 7(1), 46–54. <https://doi.org/10.20849/jed.v7i1.1326>