Journal of Kesidang
eISSN 2600-7851
Volume 9 (Bil.1) 2024: 75-94

# *VALIDATION OF THE MEASUREMENT MODEL (LEARNING AND FACILITATION-LAF) OF ROLE-PERFORMANCE MODEL SCALE (RPMS) FOR MANDARIN LECTURERS AT MALAYSIAN LOCAL UNIVERSITIES.*

**Yee chin Yip***
**Nurhasma Muhamad Saad**
**Wan Moharani Mohammad**
Faculty of Major Language Studies
Universiti Sains Islam Malaysia

**Nurul Ain Chua Binti Abdullah**
Language and Communication Department
Universiti Malaysia Terengganu

**Ng Siew Yin**
Putra Business School
Universiti Putra Malaysia

**Corresponding Author's Email cyyee@usim.edu.my / ain,chua@umt.edu.my*

***Article History:***
*Received* : 17 May 2024
*Revised* : 20 October 2024
*Published* : 31 December 2024


***To cite this article:***
Yee,CY., Nurul Ain., Nurhasma., Wan Moharani & Ng S. Y. (2024). VALIDATION OF THE MEASUREMENT MODEL (LEARNING AND FACILITATION-LAF) OF ROLE-PERFORMANCE MODEL SCALE (RPMS) FOR MANDARIN LECTURERS AT MALAYSIAN LOCAL UNIVERSITIES. *Jurnal Kesidang*, *9*(1), 75-94.

## *ABSTRACT*

*The role performance of lecturers is pivotal in transferring knowledge, especially in soft skills such as Mandarin communication. However, concerns have been raised about the qualifications and quality of some Mandarin lecturers in Malaysian local universities. This study addresses these issues by validating the Learning and Facilitation (LaF) measurement model to examine factors influencing lecturers' quality and role performance in Mandarin language teaching. The Ministry of Education (MOE) has established the Role-Performance Model Scale (RPMS), which is grounded in various educational theories and models, including Educational Quality and Standards Theories (EQST), Developmental and Constructivist Theories (DCT), and Systems Theory in Education (STE). Developed for the 2013-2025 period by the MOE, it encompasses strategies to enhance the national education system's quality, aligning it with current needs. This cross-sectional study involved 36 Mandarin lecturers in a pilot study, analysing the data through exploratory factor analysis (EFA) using SPSS. The EFA results led to the removal of some items. The actual study collected responses from 109 lecturers, with the data undergoing confirmatory factor analysis (CFA) via AMOS. The findings confirm that the RPMS LaF measurement model meets CFA requirements and validly measures factors influencing lecturers' role performance in Mandarin language teaching. The LaF is instrumental in identifying factors affecting Mandarin teaching performance, enabling stakeholders to address unsatisfactory teaching outcomes and foster a more effective educational environment.*

***Keywords:*** *Learning and Facilitation (Laf), Role-Performance Model Scale (RPMS), Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), Mandarin.*

## 1.0 INTRODUCTION

Mandarin has gained increasing significance in the 21st century, particularly due to China's growing influence in global economics and politics, motivating many non-native speakers to pursue the language at the tertiary level. In response, the Malaysian government introduced Mandarin as a foreign language in local universities, systematically implementing courses and recruiting qualified lecturers who play a pivotal role in students' success in learning the language (Abdul et al., 2023). Research consistently highlights the significant impact lecturers have on learner engagement and academic outcomes (Mercer & Dörnyei, 2020; Pishghadam et al., 2021). Huang (2022) further emphasizes that lecturers are instrumental in reducing students' negative emotions and enhancing positive experiences in foreign language learning, fostering strong lecturer-student relationships and a supportive learning environment. To improve the national education system, the Malaysian Ministry of Education launched the Pelan Pembangunan Pendidikan Malaysia (PPPM) 2013-2025, which led to the development of the Role-Performance Model Scale (RPMS). This scale was designed to guide lecturers in teaching methodologies, pedagogical skills, and effective syllabus design (Abdul et al., 2023). Although the Ministry of Education has implemented the scale for several years, clearly outlining lecturers' tasks and responsibilities, numerous challenges, obstacles, and criticisms regarding its implementation have surfaced. Concerns about the feasibility, practicality, and overall suitability of the RPMS, especially in university settings, have been raised (Chong & Graham, 2013; Mukherjee & Singh, 2016; Ormilla & Abrojena, 2023). These issues underscore the need for continuous evaluation and improvement of the RPMS to better align with the practical realities of teaching Mandarin as a foreign language in Malaysian universities.

In Mandarin language courses, lecturers play a crucial role, as Mandarin is a foreign language for most learners (Hoe et al., 2017; Tan, 2010). Their ability to communicate effectively and engage with students is essential for fostering understanding and generating interest in the language (Chua et al., 2022; Djamas, 2005; Hoe & Liaw, 2013; Liu, 2022; Tan, 2010). Lecturers need to be not only proficient in Mandarin but also skilled in teaching methodologies that are tailored to non-native speakers (Ali & Rizka, 2021; Pan et al., 2023; Rathakrishnan et al., 2024). However, the role of lecturers is often overly ambiguous and overloaded. In addition to teaching, they are frequently tasked with supervision, research, consultation, publication, and administrative duties. This multi-faceted workload can be overwhelming, negatively impacting their performance and overall well-being (Karnine et al., 2023; Rathakrishnan et al., 2024). Although the Pelan Pembangunan Pendidikan Malaysia (PPPM) clearly outlines the responsibilities of educators through the Role-Performance Model Scale (RPMS), its practical implementation often diverges from the initial planning. This discrepancy may be attributed to feasibility issues or concerns about the validity of the scale (Ormilla & Abrojena, 2023). Similar findings have been reported by Noordin & Jusoff (2009), Yusof & Abdullah (2014), and Zainal & Elias (2018), who noted significant differences between the planned and actual roles of educators. These variations are often linked to issues of practicality, feasibility, or concerns regarding the validity of the scale (Chong & Graham, 2013; Mukherjee & Singh, 2016). The validity, feasibility, and practicality of the RPMS have never been fully demonstrated or empirically tested in prior research, leaving room for potential bias and a validation gap that must be addressed.

Although the creation of the Role-Performance Model Scale (RPMS) has been met with criticism and lacks thorough verification in previous studies, some local research has adapted elements of the model. However, it still remains empirically under-validated. Alarcón & Blanca (2020), Mohammadi et al. (2024), and Ntumi & Twum (2022) emphasize that many educational research studies, particularly in fields like psychology, nursing, and counseling, have overlooked two fundamental factors: validity and reliability. These are critical in ensuring that any measurement instrument used for quality research meets appropriate standards (Kimberlin & Winterstein, 2008). Many prior studies have

adopted educational instruments without appropriate adaptation, validation, or consideration of contextual differences, which is problematic. Ensuring the validity and reliability of a tool is key to maintaining its integrity and measurement accuracy (Flake et al., 2017; Kimberlin & Winterstein, 2008). As Forza (2002) notes, without properly assessing these aspects, it becomes difficult to measure errors or establish any theoretical relationships among the studied concepts (Kennedy, 2014; Mukherjee & Singh, 2016). This concern extends to RPMS, raising questions about its effectiveness and applicability in practice. To ensure that RPMS fulfills its intended purpose, comprehensive validation studies must be undertaken to assess the framework and its implementation strategies. Without this, the reliability and relevance of the RPMS model may remain questionable, undermining its broader applicability in educational contexts. This study, therefore, seeks to address these gaps and cultivate the appropriate processes needed to bridge the practical void often overlooked in previous research.

To address this issue, the current study aims to validate the measurement model (Laf) of the RPMS, designed to investigate factors influencing lecturers' role performance. Ultimately, the findings of this study can offer a better understanding of how multiple roles might influence their quality of performance. Additionally, it provides a constructive framework for addressing these issues, thereby enhancing the overall quality of Mandarin language education.

## 2.0 LITERATURE REVIEW

The RPMS was developed to align the country's education level with international standards. This framework guides educational management in all level of institutions and serves as a reference for organizational tasks such as planning, implementing, reviewing, monitoring, and improving methods. It is action-oriented and emphasizes quality, outlining lecturers' roles in leadership practices and facilitation (Abdul Rahman et al., 2023). According to the Ministry of Education (MOE), the RPMS is supported by five key standards: Leadership, Organizational Management, Curriculum Management, Co-curriculum and Student Affairs, Learning and Facilitation, and Student Achievement. This study primarily focuses on learning and facilitation (Laf), pivotal to lecturers' roles as effective facilitators in Teaching and Learning (PdP). This standard aims to develop learners' abilities comprehensively and enhance their performance continuously (KPM, 2013). According to Zaaim et al. (2019), Standard 4 highlights five main roles of a lecturer: planner, guardian, mentor, motivator, and assessor. However, Harmer (1990), Rindu and Ariyanti (2017), and Sanjaya (2007) contend that an education lecturer assumes six roles in classroom management, namely controller, assessor, organizer, prompter, participant, and resource provider, whereas, learning and facilitation should be allocated in resources provider. In other words, this presents contradiction opinions or outcomes, thereby suggesting the need for further exploration and clarification of the lecturer's roles in the educational process. The discrepancy between the identified roles underlines a potential gap in the RPMS framework or in its interpretation and application in the context of Teaching and Learning (PdP). This divergence highlights the complexity of educators' functions and the importance of a comprehensive understanding that encompasses both traditional and innovative pedagogical roles.

The Laf was initially designed for school teachers, its application has extended to university-level education in numerous studies (Abdul et al., 2023). This broader use has brought the framework's adaptability and suitability under scrutiny. The Laf, originally intended to enhance school teaching practices, is now being applied in diverse educational contexts, including higher education. This shift raises questions about the framework's universal applicability and effectiveness across different educational levels. As a result, the validity and reliability of the Laf when applied beyond its intended scope have become topics of academic discussion and research. The debate centers around whether a model tailored for school environments can accurately and effectively be adapted for university settings, a question that continues to inspire further research and analysis in the field of education. According to Hoe and Lim (2013), a significant number of Mandarin lecturers in Malaysian local universities hold degrees in Chinese studies, linguistics, or other disciplines. A notable proportion of these lecturers commence their teaching careers without having acquired any formal government

certification specific to teaching. This situation presents a challenge in terms of standardizing teaching quality and methodologies. The absence of a uniform certification or training process for university-level Mandarin lecturers means that teaching practices can vary widely (Chan et al., 2022; Pan et al., 2023; Zhao, 2007). As a result, the application of standardized models or frameworks for teaching assessment and development, such as the Laf, may not fully align with the actual on-ground practices and diverse educational backgrounds of these lecturers. This discrepancy highlights the need for more study to identify the suitability of the model, and ensure the consistency and quality in Mandarin language education at the university level.

Additionally, several studies have presented strong evidence suggesting that the approach to teaching Mandarin as a foreign language course may differ significantly from the standard school (Chua et al., 2023; Hoe, 2013; Tan, 2010). This discrepancy arises because Mandarin as foreign language courses for learners often require specialized methodologies tailored to the unique challenges of acquiring a new language (Pan et al., 2023), whereas the Laf is typically designed for fundamental courses aligned with school curricula (Abdul et al., 2023). Besides, foreign language learners are typically mature and adult, and their expectations and learning attitudes may differ from those of school students. This adult learner demographic often has more specific, practical goals and may require a more flexible and contextually relevant approach to learning (Oztürk; 2020; Wlosowicz, 2016). Therefore, it is essential to justify the LaF model to accommodate these differences, ensuring that it addresses the distinct needs and motivations of adult learners in language education. Validating the model can ensure that lecturers truly perform, leading to more effective teaching strategies that enhance engagement, facilitate deeper learning, and ultimately improve language proficiency among adult students.

In educational research, the importance of measurement validity cannot be overstated, as it is critical for ensuring the accuracy and consistency of research findings (Barry et al., 2014; Kimberlin & Winterstein, 2008; Mohajan, 2017; Ntumi & Twum, 2022). Ntumi and Twum (2022) note that in their review of education articles published from 2010 to 2020, more than half failed to report on measurement validity. Although the Laf was meticulously developed to assess the performance of roles, there is a notable scarcity of studies conducted to verify the model's validity. This oversight reveals a significant practical knowledge gap, underscoring the necessity for rigorous validation studies. Such research is vital not only for confirming the Laf's reliability and applicability in educational settings but also for enhancing its utility in evaluating educational outcomes. Especially, the teaching and learning of foreign languages like Mandarin might present unique challenges and dynamics not fully encapsulated by current models (Chan et al., 2022; Tao & gao, 2022; Pan et al., 2023). The cultural nuances, linguistic complexities, and specific pedagogical needs associated with teaching Mandarin necessitate a tailored approach to role performance measurement. This highlights the need to conduct validation studies in diverse educational settings, including those focused on foreign language instruction. Conducting such studies across different contexts, including online and traditional classroom environments, with learners of varying proficiency levels, can provide comprehensive insights into the RPMS's versatility and effectiveness.

Although this model has been established for years and utilised in various educational studies, very few have applied factor analysis techniques to explore its validity across different contexts, especially within the Mandarin teaching setting. Given the inconsistencies and questions surrounding the model's validity, this study has been conducted to examine the validity of the RPMS’s Learning and Facilitation (LaF) measurement model using EFA and CFA techniques, with data collected from Mandarin lecturers at a local Malaysian university. This approach will be able to confirm the validity, reliability, workability, suitability, and generalizability of the LaF measurement model (Bandalos & Enders, 1996; Costello & Osborne, 2005). Thus, validating the LaF model presents an opportunity to deepen our understanding of the role factors impacting lecturer quality in Mandarin teaching, especially at the university level.

## 3.0 METHODOLOGY

This study employed a cross-sectional research design, collecting data at a single point in time from 15 different local universities in Malaysia. Simple random sampling was utilized to select respondents among Mandarin lecturers at these universities. The questionnaire was supported by five sub-sections: organizer, controller, guider, prompter, and evaluator, excluding the demographic respondent profile. All the items were adapted and adopted from previous studies. However, certain items required specific clarification and grammar correction. Subsequently, the researcher restructured the questionnaire items based on previous literature and studies, a method recommended by Francis et al. (2004) and Sutton et al. (2017) for reconstructing survey instruments.

After amendments and corrections, the questionnaire required final approval from experts. A pre-test was conducted to ensure the content validity, face validity, and criterion validity of the LaF for the actual fieldwork. Content validity was assessed by five content experts, all of whom were academicians with over 20 years of experience teaching and coordinating Mandarin courses. Criterion validity was evaluated by a statistical expert, a retired professor with extensive experience in data analysis, teaching, and writing, to confirm the appropriateness of the scale used. Following this, the LaF underwent back-to-back translation from English to Malay (Bahasa Malaysia) by a certified translator to ensure face validity. Once the validation procedure was completed, the LaF was pre-tested on six randomly selected respondents, all Mandarin language lecturers from different universities. This pre-test aimed to assess the consistency of their responses and gather feedback on any ambiguous terms, question clarity, and overall questionnaire design. Identified issues were addressed before proceeding to the pilot study and the actual fieldwork (Zikmund & Babin, 2010).

After revising the instrument based on expert comments and pre-test results, a pilot study was conducted, yielding 36 valid responses, fulfilling the required minimum sample size of 30 (Ghazali, 2016; Hill, 1998; Isaac & Michael, 1995; Saunders, 2007; Sekaran, 2003). The pilot study data underwent EFA before the actual survey (Awang, 2015; Bahkia et al., 2019). The finalized version of the LaF instrument consisted of 34 items, excluding questions on the respondents' demographic profile. A 5-point interval scale ranging from 1 (strongly disagree) to 5 (strongly agree) was employed for the instrument. This interval scale was recommended by Awang et al. (2016), Ghazali (2021), and Lee and Cherner (2015) to ensure that the data obtained from LaF are more independent. Even a 5-point Likert scale can be considered more suitable for testing a questionnaire regarding attitude, practice, and perception, and is easily understood by respondents (Ibrahim et al., 2019; Zaujan et al., 2021).

The actual survey obtained 116 responses, of which 109 were valid for analysis, while 7 responses were eliminated due to normality issues. According to Boomsma (1985), Hair et al. (2010), and Kline (2010), the minimum sample size requirement for Structural Equation Modelling (SEM) is at least 100 respondents. This aligns with established rules of thumb, as the measurement model in this study, with fewer than 50 items, can be considered relatively simple (Akter et al., 2010; MacCallum et al., 1999). The data were analysed using the Statistical Package for Social Sciences (SPSS version 26) for data screening, including checks for outliers, normality, and missing data, followed by Exploratory Factor Analysis (EFA). To validate the measurement model, the Analysis of Moment Structures (AMOS) was utilized, ensuring constructs' unidimensionality, validity, and reliability through Confirmatory Factor Analysis (CFA) (Afthanorhan et al., 2019; Awang et al., 2015; Awang et al., 2018; Mahfouz et al., 2019; Mohamad et al., 2018; Rahlin et al., 2019).

## 4.0 RESULTS

### 4.1 Exploratory Factor Analysis (EFA)

The primary objective of EFA was to explain the data by grouping variables that are correlated (Zikmund & Babin, 2010). EFA utilised data from the pilot study to disclose the underlying dimensions of organizer, controller, guider, prompter, and evaluator. Several key requirements are

deemed essential for EFA. Firstly, the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (MSA) should greater 0.50. Secondly, the results of Bartlett's test of sphericity should be significant at $p < 0.001$, as suggested by Hair et al. (2014), Awang (2015), and Bahkia et al. (2019).

Table 1 outlines the results of the KMO and Bartlett's test of sphericity for organizer, controller, guider, prompter, and evaluator. The values of KMO for all constructs exceeded 0.5. The Bartlett's test of sphericity results for all constructs were significant ($p < 0.001$) as recommended by Hair et al. (2014), Bahkia et al. (2019), Rahlin et al. (2019), and Shkeer and Awang (2019).

**Table 1: Results of KMO and Bartlett's Test of Sphericity**

| Construct | KMO (>0.50) | Bartlett's Test of Sphericity (p<0.001) |
|---|---|---|
| Organizer | 0.865 | 0.00 |
| Controller | 0.919 | 0.00 |
| Guider | 0.867 | 0.00 |
| Prompter | 0.941 | 0.00 |
| Evaluator | 0.878 | 0.00 |

During the exploratory factor analysis (EFA), the principal component analysis method was utilized to extract factors and determine which factors should be retained or eliminated. To enhance the interpretability of the factor analysis, varimax rotation, an orthogonal rotation method widely employed in factor analysis, was applied (Hair et al., 2014). Factor loadings with absolute values below 0.5 were deemed insignificant and excluded from further analysis, while those exceeding 0.5 were considered significant and retained for measurement (Hair et al., 2014). The result of the rotated component matrix is displayed in Table 2 as follows:

**Table 2: Rotated Component Matrix - Learning and Facilitation (Laf)**

| No | Sub-Construct | Item Label | Rotated Component Matrix Item Statement | Component | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | 1 | 2 | 3 | 4 | 5 |
| 1 | Organizer | Or1 | Prepare the lesson that includes objectives | .591 | | | | |
| 2 | | Or 2 | Prepare the suitable learning activities | .692 | | | | |
| 3 | | Or3 | Clearly state the course's evaluation components | .735 | | | | |
| 4 | | Or4 | Prepare the teaching aids | .666 | | | | |
| 5 | | Or5 | Prepare the ICT learning aids | .655 | | | | |
| 1 | Controller | Ct1 | Manage the teaching content effectively. | | .686 | | | |
| 2 | | Ct2 | Manage the learning activities effectively | | .669 | | | |
| 3 | | Ct3 | Manage the P&P period effectively according to the activities | | .569 | | | |
| 4 | | Ct4 | Provide opportunities for the students to take part in the learning activities | | .610 | | | |
| 5 | | Ct5 | Monitor students' communication throughout the P&P | | .585 | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 6 | | Ct6 | Monitor students' behavior throughout the P&P | .572 | | | |
| 7 | | Ct7 | Create a conducive environment for P&P | .612 | | | |
| 1 | Guider | Gu1 | Assist students to master the content. | | .666 | | |
| 2 | | Gu2 | Assist students to master the skills in learning activities. | | .694 | | |
| 3 | | Gu3 | Assist students to make decisions and solve learning-related issues. | | .660 | | |
| 4 | | Gu4 | Assist students in effectively utilizing academic materials. | | .679 | | |
| 5 | | Gu5 | Integrate learning content with daily life. | | .605 | | |
| 1 | Prompter | Pr1 | Encourage students to communicate. | | | .556 | |
| 2 | | Pr2 | Encourage collaboration within students. | | | .685 | |
| 3 | | Pr3 | Ask critical and creative skill-related questions. | | | *.466 | |
| 4 | | Pr4 | Ask problem-solving skill related questions. | | | .689 | |
| 5 | | Pr5 | Create opportunities for students to experience leadership with appropriate activities. | | | .630 | |
| 6 | | Pr6 | Encourage learning-related questions from students. | | | .773 | |
| 7 | | Pr7 | Encourage students' independence in acquiring knowledge and skills. | | | *.491 | |
| 8 | | Pr8 | Reward positive behavior. | | | .721 | |
| 9 | | Pr9 | Appreciate great works. | | | .771 | |
| 10 | | Pr10 | Enhance students' confidence in responding. | | | .698 | |
| 11 | | Pr11 | Concern towards students' wellbeing. | | | .733 | |
| 1 | Evaluator | Ev1 | Utilize various evaluation methods in P&P. | | | | .820 |
| 2 | | Ev2 | Conduct remedial activities for low-achieving students. | | | | .795 |
| 3 | | Ev3 | Conduct enrichment activities for high-achieving students. | | | | *.456 |
| 4 | | Ev4 | Provide reinforcement activities. | | | | .691 |
| 5 | | Ev5 | Conduct a reflection session after class/tutorial. | | | | .796 |
| 6 | | Ev6 | Review students' assignments by providing appropriate assessment. | | | | .769 |

According to Table 2, the rotated component matrix results for the LaF (Table 2) divided all 34 items into five sub-components. Most of the items have factor loadings greater than 0.5, and the items align

neatly with the stated sub-constructs of LaF. However, three items (Pr 3, Pr 7 and Ev 3) with factor loadings lower than 0.5 have been eliminated.

At the same time, the researcher has presented the outcomes of the scree plot. The scree plot for LaF sorted out the 34 items into five components in Figure 1.

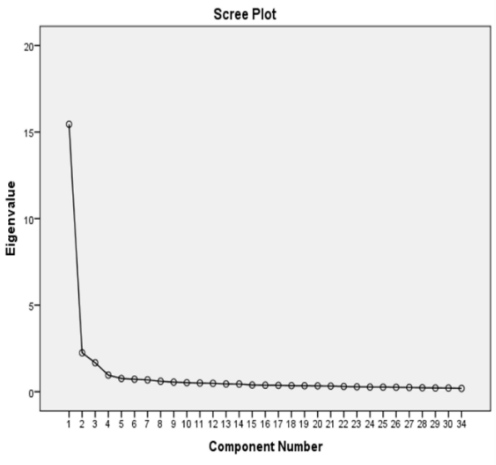


Figure 1: Scree Plot of LaF

Based on the result of the scree plot, the inspection revealed a clear break after the five factors for LaF. This suggests that the five factors extracted adequately represent the underlying structure of the data, supporting the decision to retain these factors in the analysis (Hair et al., 2014).

Next, Table 3 illustrates the EFA results and the number of items for each construct before and after the analysis. Essentially, most of the items surpassed the minimum threshold of 0.5; however, there are three items that need to be removed due to low factor loading (>0.50).

**Table 3: Item Retention Result after EFA**

| No | Construct | Items before EFA | Number of Items Dropped | Number of Items Retained after EFA |
|---|---|---|---|---|
| 1 | Organizer | 5 | 0 | 5 |
| 2 | Controller | 7 | 0 | 7 |
| 3 | Guider | 5 | 0 | 5 |
| 4 | Prompter | 11 | 2 | 9 |
| 5 | Evaluator | 6 | 1 | 5 |
| | **Total** | **34** | **3** | **31** |

Thus, the final 31 items in the LaF will remain. The EFA results also indicated that all items neatly aligned with the stated constructs without mixing values. Subsequently, the sub-constructs and items are directed to proceed to the next level of analysis in this research.

### 4.2 Confirmatory Factor Analysis (Pooled-CFA)

This study validated the measurement models of latent constructs from three crucial perspectives: unidimensionality, validity, and reliability (Afthanorhan et al., 2017; Awang, 2015; Hair et al., 2014; Mohamad et al., 2018). This essential procedure, known as 'confirmatory factor analysis' (CFA), subjected the measurement model of latent constructs to three types of validity tests: convergent,

construct, and discriminant (Awang, 2015; Hair et al., 2014; Yusof et al., 2017). Convergent validity was assessed by calculating the average variance extracted (AVE), while construct validity was evaluated by reviewing the fit indices of the measurement model. Discriminant validity was established through a Discriminant Validity Index Summary.

For reliability assessment, Composite Reliability (CR) was employed, considered a better and advance alternative to the traditional Cronbach Alpha method in analysing validity of model LaF (Awang, 2015; Aziz et al., 2016; Hair et al., 2014; Yusof et al., 2017). The outcomes of the pooled CFA are displayed as Figure 2 below:

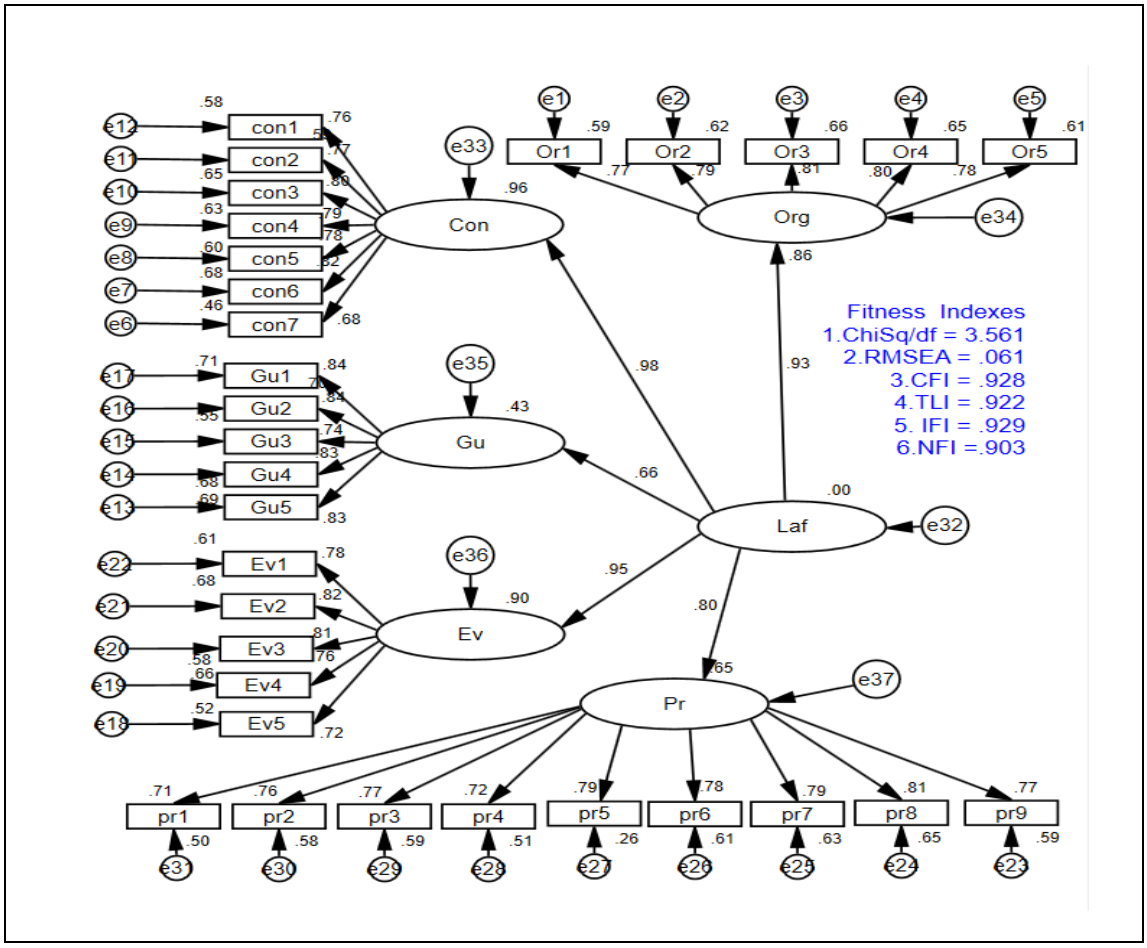


Figure 2: Pooled CFA of LaF measurement model

Figure 2 illustrates that all sub-constructs in this measurement model were pooled for simultaneous validation using a Pooled-CFA, and the fitness indices proved the fitness of the LaF model. According to Awang (2015) and Awang et al. (2018), the function of pooled CFA is its ability to assess the measurement model across different groups simultaneously, allowing for direct comparison of model fit and parameter estimates between sub-groups. This approach provides more robust and comprehensive insights into the validity and reliability of the measurement model across diverse populations or conditions.

### 4.3 Unidimensionality

Unidimensionality refers to a set of variables that can be attributed to one construct (Hair et al., 2014). As stated by Awang (2015), unidimensionality is obtained when all measurement items for the respective constructs show acceptable factor loadings. Items in CFA with low factor loadings should be eliminated immediately from the measurement model until satisfactory fit indices are achieved (Afthanorhan et al., 2017; Asnawi et al., 2019; Awang, 2015; Hair et al., 2014; Kashif et al., 2016).

Awang (2015) and Awang et al. (2018) outlines two prerequisites that must be met before considering the removal of an item:

1. Newly developed items should have a factor loading of 0.5 or higher.
2. Established items should exhibit a factor loading of 0.6 or higher.

**Table 3. Factor Loading of All Items**

| No | Construct/Item | Factor Loading |
|---|---|---|
| | **Organizer (Org)** | |
| 1 | Prepare the lesson that includes objectives. | 0.77 |
| 2 | Prepare the suitable learning activities. | 0.79 |
| 3 | Clearly state the course's evaluation components. | 0.81 |
| 4 | Prepare the teaching aids. | 0.80 |
| 5 | Prepare the ICT learning aids. | 0.78 |
| | **Controller (Con)** | |
| 1 | Manage the teaching content effectively. | 0.76 |
| 2 | Manage the learning activities effectively. | 0.77 |
| 3 | Manage the P&P period effectively according to the activities. | 0.88 |
| 4 | Provide opportunities for the students to take part in the learning activities. | 0.79 |
| 5 | Monitor students' communication throughout the P&P. | 0.78 |
| 6 | Monitor students' behavior throughout the P&P. | 0.82 |
| 7 | Create a conducive environment for P&P. | 0.68 |
| | **Guider (Gu)** | |
| 1 | Assist students to master the content. | 0.78 |
| 2 | Assist students to master the skills in learning activities. | 0.82 |
| 3 | Assist students to make decisions and solve learning-related issues. | 0.81 |
| 4 | Assist students in effectively utilizing academic materials. | 0.76 |
| 5 | Integrate learning content with daily life. | 0.72 |
| | **Prompter (Pr)** | |
| 1 | Encourage students to communicate. | 0.71 |
| 2 | Encourage collaboration within students. | 0.76 |
| 3 | Ask problem-solving skill related questions. | 0.77 |
| 4 | Create opportunities for students to experience leadership with appropriate activities. | 0.72 |
| 5 | Encourage learning-related questions from students. | 0.79 |
| 6 | Reward positive behavior. | 0.78 |
| 7 | Appreciate great works. | 0.79 |
| 8 | Enhance students' confidence in responding. | 0.81 |
| 9 | Concern towards students' wellbeing. | 0.77 |
| | **Evaluator (Ev)** | |
| 1 | Utilize various evaluation methods in P&P. | 0.78 |
| 2 | Conduct remedial activities for low-achieving students. | 0.82 |
| 3 | Provide reinforcement activities. | 0.81 |
| 4 | Conduct a reflection session after class/tutorial. | 0.76 |
| 5 | Review students' assignments by providing appropriate assessment. | 0.72 |

**Factor loading (>.50)**

Table 3 displays that all items from all construct have surpassed the minimum threshold (>.50) factor loading values recommended by Awang (2015) and Awang et al. (2018). Thus, no any item needs to be removed from this study.

### 4.4 Convergent Validity

Convergent validity refers to a set of indicators presumed to measure a construct (Hair et al., 2014; Kline, 2011; Awang, 2015; Awang et al., 2018). As noted by Brown (2006), it represents the strength of relationships among items that are expected to represent a single latent construct. Convergent validity can be verified by computing the average variance extracted (AVE). A construct achieves convergent validity if its AVE exceeds the threshold value of 0.5 (Awang et al., 2018; Awang, 2015; Fornell & Larcker, 1981; Hair et al., 2014). The results of this study are displayed in Table 4 below:

**Table 4: Average Variance Extracted for All Constructs**

| Codes | Construct | AVE (Above 0.5) |
|---|---|---|
| 1 | Organiser | 0.624 |
| 2 | Controller | 0.616 |
| 3 | Guider | 0.607 |
| 4 | Prompter | 0.589 |
| 5 | Evaluator | 0.607 |

Based on the table, the AVE for all sub-constructs exceeded the minimum value of 0.5. The sub-construct of Organizer received the highest AVE (0.624), while the sub-construct of Prompter received the lowest factor loading (0.589). Thus, it can be concluded that the model has achieved convergent validity.

### 4.5 Construct Validity

Generally, construct validity is attained when all the fitness indices for a model meet the required criteria (Awang, 2015; Awang et al., 2018). The three model fit categories-absolute fit indices, incremental fit indices, and parsimonious fit indices-are adequate for establishing construct validity (Awang et al., 2015, 2018; Kashif et al., 2016; Yusof et al., 2018; Asnawi et al., 2019). The most commonly used indicators include root mean square of approximation (RMSEA), comparative fit index (CFI), and normed Chi-Square ($\chi 2$/df) (Awang, 2015; Awang et al., 2018). Table 5 summarises the fitness indices categories and the corresponding acceptance levels according to the literature.

**Table 5: Fitness Indices**

| Name of category | Name of index | Level of acceptance | Result | Status |
|---|---|---|---|---|
| Absolute Fit Index | RMSEA | RMSEA < 0.08 (Hu & Bentler, 1999) | 0.061 | Achieved |
| Incremental Fit Index | CFI | CFI > 0.90 | 0.928 | Achieved |
| | TLI | TLI > 0.90 | 0.922 | |
| | IFI | IFI > 0.90 | 0.929 | |
| | NFI | NFI > 0.90 (Awang, 2012) | 0.903 | |
| Parsimonious Fit Index | Chi-Square/df | Chi-Square/df < 5.0 (Hu & Bentler, 1990) | 3.561 | Achieved |

Table 5 indicates that the LaF met all three categories of fitness indices: (1) the RMSEA value was less than 0.08 (specifically, 0.061), confirming the absolute fit index; (2) the LaF achieved an incremental fit index by securing a CFI value (0.928) above the recommended threshold of 0.90; (3) regarding the parsimonious fitness index, which was gauged using the Chisq/df, a value of 3.561 was

obtained. This value is below the recommended threshold of 5.0 as suggested by Bentler (1990), Awang (2015, 2018). Consequently, this study successfully established the construct validity of the LaF.

### 4.6 Discriminant Validity

Discriminant validity ensures that the measurement model of a construct is free from redundant items, meaning that each item provides unique information without overlapping with others (Awang, 2015, 2018). Redundancy arises when two constructs in the model show a high correlation (Hair et al., 2014). When evaluating discriminant validity, the correlation between exogenous constructs should not surpass 0.85. If this threshold is exceeded, it suggests redundancy and highlights a significant multicollinearity problem (Lewis et al., 2004).

**Table 6: Discriminant Validity Index Summary**

| Construct/ Codes | Org | Con | Gu | Pr | Ev |
|---|---|---|---|---|---|
| **Org** | ***0.624*** | | | | |
| **Con** | 0.130 | ***0.616*** | | | |
| **Gu** | 0.386 | 0.217 | ***0.607*** | | |
| **Pr** | 0.312 | 0.317 | 0.239 | ***0.589*** | |
| **Ev** | 0.361 | 0.359 | 0.262 | 0.151 | ***0.607*** |

Discriminant validity for each construct was successfully established, as indicated by the square root of its AVE exceeding its correlation value with other constructs in the model (Table 6) (Awang et al., 2018; Awang, 2015; Hair et al., 2014). The achievement of discriminant validity is further confirmed by observing that the diagonal values (in bold) are higher than any other value in both their respective rows and columns. Therefore, discriminant validity is confirmed for all sub-constructs in the LaF, meeting the required threshold as demonstrated in Table 6.

### 4.7 Composite Reliability (CR)

Composite reliability is employed to estimate reliability in structural equation models (Awang et al., 2018; Awang, 2015; Hair et al., 2014). Composite reliability estimates of 0.7 or higher suggest good reliability, while a score between 0.6 and 0.7 is considered acceptable (Awang, 2015; Hair et al., 2014). The results of this study are displayed in Table 7 below:

**Table 7: Composite Reliability (CR)**

| Codes | Construct | CR (Above 0.6) |
|---|---|---|
| Org | Organizer | 0.893 |
| Con | Controller | 0.918 |
| Gu | Guider | 0.885 |
| Pr | Prompter | 0.928 |
| Ev | Evaluator | 0.885 |

Table 7 reveals that the analysis demonstrates composite reliability scores for all constructs in the LaF exceeding the minimum threshold of 0.6. The construct with the highest composite reliability is Prompter (0.928), while Controller and Evaluator exhibit the lowest composite reliability (0.885). Therefore, the LaF achieves satisfactory composite reliability.

### 4.8 Normality Assessment

The normal distribution of all items measuring the constructs in LaF was assessed. According to Asnawi et al. (2019), Awang (2015), Hair et al. (2014), Kashif et al. (2015, 2016), and Mohamad et al. (2016, 2018), skewness values should ideally fall within the acceptable range of -2 to 2, ensuring no significant deviation from normality. Skewness is important because it provides insights into the symmetry (or lack thereof) in the distribution of data, allowing researchers to identify whether the data are skewed towards the left or right tail (Hair et al., 2014). This information is crucial for the validity of statistical analyses and the interpretation of the results.

In this study, given the relatively small sample size of 100 to 150 respondents, skewness is particularly sensitive and informative for detecting data issues (Hair et al., 2010) Skewness is often more critical than kurtosis in such situations, as small sample sizes may limit the reliability of kurtosis measurements (Curran et al., 1996; DeCarlo, 1997; Kim, 2013). Since skewness directly addresses the asymmetry in the data distribution, it was deemed sufficient for assessing normality, and the analysis of kurtosis was considered unnecessary (Kline, 2011; Tabachnick & Fidell, 2013). The results of normality data are displayed in Table 8 below:

**Table 8: Normality Assessment Results**

| No | Construct/Item | Skewness |
|---|---|---|
| | **Organizer (Org)** | |
| 1 | Prepare the lesson that includes objectives. | -0.662 |
| 2 | Prepare the suitable learning activities. | -0.966 |
| 3 | Clearly state the course's evaluation components. | -0.168 |
| 4 | Prepare the teaching aids. | 0.606 |
| 5 | Prepare the ICT learning aids. | -0.909 |
| | **Controller (Con)** | |
| 1 | Manage the teaching content effectively. | -0.669 |
| 2 | Manage the learning activities effectively. | -0.921 |
| 3 | Manage the P&P period effectively according to the activities. | -0.928 |
| 4 | Provide opportunities for the students to take part in the learning activities. | -0.828 |
| 5 | Monitor students' communication throughout the P&P. | -0.626 |
| 6 | Monitor students' behavior throughout the P&P. | -0.747 |
| 7 | Create a conducive environment for P&P. | -0.681 |
| | **Guider (Gu)** | |
| 1 | Assist students to master the content. | -0.926 |
| 2 | Assist students to master the skills in learning activities. | -0.269 |
| 3 | Assist students to make decisions and solve learning-related issues. | -0.703 |
| 4 | Assist students in effectively utilizing academic materials. | -0.857 |
| 5 | Integrate learning content with daily life. | -0.639 |
| | **Prompter (Pr)** | |
| 1 | Encourage students to communicate. | -0.689 |
| 2 | Encourage collaboration within students. | -0.753 |
| 3 | Ask problem-solving skill related questions. | -0.836 |
| 4 | Create opportunities for students to experience leadership with appropriate activities. | -1.116 |
| 5 | Encourage learning-related questions from students. | -0.601 |
| 6 | Reward positive behavior. | -1.128 |
| 7 | Appreciate great works. | -0.992 |
| 8 | Enhance students' confidence in responding. | -1.166 |
| 9 | Concern towards students' wellbeing. | -1.560 |

| | Evaluator (Ev) | |
|---|---|---|
| 1 | Utilize various evaluation methods in P&P. | -1.009 |
| 2 | Conduct remedial activities for low-achieving students. | -1.398 |
| 3 | Provide reinforcement activities. | -1.201 |
| 4 | Conduct a reflection session after class/tutorial. | -1.229 |
| 5 | Review students' assignments by providing appropriate assessment. | -0.601 |

The skewness values for all components in the model fall within the acceptable range of -2 to 2, indicating that their distribution does not significantly deviate from normality (Table 8) (Awang, 2015; Kashif et al., 2016; Mohamad et al., 2019; Asnawi et al., 2019). This suggests that the data distribution in the LaF conforms to the requirement of normality distribution, which is crucial for ensuring the robustness and validity of statistical analyses conducted in this study.

## 5.0 CONCLUSION

This study aimed to validate the measurement model of the RPMS for investigating educators' quality in teaching. Findings from both the EFA and CFA suggest that the measurement model is a good fit and suitable for examining the factors affecting educator quality. Based on the EFA results, two items from the prompter sub-construct and one item from the evaluator sub-construct were recommended for removal. These items did not meet the minimum factor loading criterion of 0.50 based on data from the actual study. Subsequently, the CFA confirmed that the LaF met the requirements for convergent validity, construct validity, and discriminant validity. Findings from the assessments of unidimensionality and normality also support the validity of the items in the LaF sub-constructs. Therefore, the results from both the EFA and CFA indicate that the LaF is reliable for examining the factors influencing lecturers' quality and role performance in language teaching.

## 6.0 IMPLICATIONS

Based on the findings of this study, several key implications arise. First and foremost are the academic and theoretical implications. This study contributes to the growing body of research on educational assessment by validating a specific measurement tool, the Role-Performance Model Scale (RPMS), tailored for Mandarin lecturers. It expands the scope of educational measurement and performance models within the context of language teaching in higher education institutions. By focusing specifically on Mandarin lecturers, the study highlights the pedagogical needs, strategies, and teaching methodologies essential for effectively teaching Mandarin as a foreign language in Malaysian universities. Additionally, the research provides valuable insight into how to measure teaching performance in foreign language instruction, which has often been an underexplored area. The study addresses the critical gap of validating the RPMS, which had previously not been empirically tested extensively. By filling this validation gap, the research ensures that the measurement model accurately reflects the role and performance of Mandarin lecturers, enhancing the reliability and robustness of the scale for future use in educational settings. In other words, this validation strengthens the theoretical framework for evaluating foreign language teaching performance and sets a precedent for the development of similar performance scales in other academic disciplines.

Secondly, this study has significant practical implications for educators and universities. It offers improvements in lecturer assessment: The validation of the LaF model allows educators to be evaluated with greater accuracy, providing better insights into their performance. Universities can leverage this tool to assess the effectiveness of Mandarin

lecturers and implement more targeted professional development programs, thus enhancing their teaching methods. In addition, the study promotes enhanced teaching and learning practices: With a validated tool for evaluating lecturer performance, universities can adopt more effective strategies for improving teaching methods and student learning outcomes. The RPMS provides structured feedback to lecturers, helping them identify specific areas for improvement in their facilitation and pedagogical skills, thus elevating the quality of Mandarin instruction. In other words, this practical approach not only helps in the betterment of teaching standards but also aligns with strategic goals for higher education institutions striving to offer quality foreign language instruction.

Thirdly, the policy implications of this study hold significant value for both informing educational policy and promoting standardization in performance assessments. The findings can guide policymakers in the Malaysian Ministry of Education regarding the effectiveness of the RPMS, particularly in the context of foreign language teaching. With the validated model, more robust policies can be developed to improve lecturer performance evaluation, helping ensure fair and comprehensive assessments that enhance the quality of Mandarin language instruction. Furthermore, the validation of the LaF model provides a framework for standardizing performance evaluation tools across other foreign language departments. By adapting this model, policymakers and educational institutions can create uniform, consistent methods for assessing language educators' performance. This standardization will not only contribute to the fair evaluation of lecturers but also support the development of targeted policies that promote continuous improvement and excellence in foreign language education.

## 7.0 RECOMMENDATION FOR FUTURE STUDY

Based on the outcomes of this study, several recommendations can be made for future research on the Role-Performance Model Scale (RPMS). First, future studies could explore the application of RPMS in other foreign language departments or across various academic disciplines. Testing the scale in diverse educational contexts would allow researchers to determine its adaptability and effectiveness for evaluating teaching performance in different subject areas beyond Mandarin language instruction.

Second, researchers can conduct longitudinal studies that track lecturer development over time using the RPMS to provide deeper insights into its effectiveness in enhancing teaching practices. Such studies could reveal long-term trends in lecturer performance improvement and highlight the impact of RPMS on student learning outcomes.

Lastly, future research should investigate the integration of student feedback with the RPMS. Incorporating students' perspectives on the effectiveness of teaching methodologies could lead to a more comprehensive evaluation of teaching performance. This approach would allow future studies to refine the RPMS and better align lecturer assessments with student learning experiences and satisfaction.

## AUTHOR CONTRIBUTIONS

Yee.C.Y: Conceptualization, Methodology, Validation; Nurul Ain: Writing-Original Draft Preparation; Nurhasma, Wan Moharani and Ng: Writing-Reviewing, Editing and questionnaire distribution.

## CONFLICT OF INTEREST

The manuscript has not been published elsewhere and is not under consideration by other journals. All authors have approved the review, agree with its submission and declare no conflict of interest on the manuscript.

## REFERENCES

Abdul Rahman, M. A., Mohd Akbar, A. J., Raja Abdullah, R. R., Amat, N., Yoon, L. S., Mohd Adzhar, M. A., & Osman, S. (2023). Assessing Mandarin lecturers' quality on role-performance in MARA education institutions (IPMA). *Journal of Nusantara Studies, 8*(1), 140-157.

Abrojena, V. L., & Ormilla, R. C. G. (2023). Challenges and Opportunities in the Implementation of Results-Based Performance Management System. *Management Research Journal*, 12 (1), 1-11.

Afthanorhan, A., Awang, Z., & Fazella, S. (2017). Developing the patients' loyalty model for medical tourism industry: The case of Malaysia. *International Journal Society Systems Science*, 9(2), 139–164.

Afthanorhan, A., Awang, Z., Rashid, N., Foziah, H., & Ghazali, P. L. (2019). Assessing the effects of service quality on customer satisfaction. *Management Science Letters, 9*(1), 13–24.

Akter, S., D'Ambra, J., & Ray, P. (2010). Service quality of mHealth: Development and validation of a hierarchical model using PLS. *Electronic Markets*, 20(3), 209-227.

Alarcon, R., & Blanca, M. J. (2020). Development and psychometric properties of the questionnaire for assessing educational podcasts (QAEP). *Frontiers in Psychology*, 11, 454-579.

Ali Wira and Rizka Indahyanti. (2021). Factor influencing the lecturers' teaching performance. *Journal of Education, Language Teaching and Science*, 3 (3), 96-105.

Asnawi, A. A., Awang, Z., Afthanorhan, A., Mohamad, M., & Karim, F. (2019). The influence of hospital image and service quality on patients' satisfaction and loyalty. *Management Science Letters, 9* (6), 911–920.

Asnawi, N., Awang, Z., & Rahlin, N. A. (2019). Factor analysis in business research: A step-by-step procedure using SmartPLS. *Journal of Management Development*, 38(5), 387-403.

Awang, Z. (2015). *SEM Made Simple: A gentle approach to learning structural equation modeling*. Malaysia: MPWS Rich Resources.

Awang Z., Asyraf Afthanorhan & Mustafa Mamat, (2016). Computational Methods in Social Sciences (CMSS), *Journal Economic Sciences*, 4 (1), 13-21.

Awang, Z., Lim, S. H., & Zainudin, N. (2018). *Pendekatan mudah SEM - Structural equation modeling*. Malaysia: MPWS Rich Resources.

Aziz, M. I., Afthanorhan, A., & Awang, Z. (2016). Talent development model for a career in Islamic banking institutions: A SEM approach. *Cogent Business & Management, 2* (1), 1–11.

Bahkia, A. S., Awang, Z., Afthanorhan, A., Ghazali, P. L., & Foziah, H. (2019). Exploratory factor analysis on occupational stress in context of Malaysian sewerage operations. *AIP Conference Proceedings, 2138*(1), 050006. doi: 10.1063/1.5121111.

Bandalos, D. L., & Enders, C. K. (1996). The effects of nonnormality and number of response categories on reliability. *Applied Measurement in Education*, 9 (2), 151–160.

Barry, A. E., Chaney, B., Piazza-Gardner, A. K., & Chavarria, E. A. (2014). Validity and reliability reporting practices in the field of health education and behaviour: A review of seven journals. *Health Education & Behaviour, 41*(1), 12-18.

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika*, 50(2), 229–242.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York City: The Guilford Press.

Chan, Jessica & Woore, Robert & Molway, Laura & Mutton, Trevor. (2022). Learning and teaching Chinese as a foreign language: A scoping review. *Review of Education*. 10. 10.1002/rev3.3370.

Chong, S., & Graham, L. (2013). The 'Chinese dream' in the educational reform context in Malaysia: Implementation and challenges. *Asia Pacific Journal of Education*, 33(3), 355-369.

Chong, S. T., & Graham, R. (2013). Assessing the feasibility of teacher performance scales: A critique of practical implementation. *International Journal of Education*, 17(4), 329-346.

Chua, H. W., Lee, H. L., Omar, S., Tohleheng, M., & Chin, S. M. (2022). Malaysian Non-native Chinese Students' Challenges and Learning Strategies in Chinese Characters Learning. *EDUCATUM Journal of Social Sciences*, 8(1), 67-82.

Costello, A. B., & Osborne, J. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. *Practical Assessment Research & Evaluation*, 10, 1-9.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16–29.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. Psychological Methods, 2(3), 292-307.

Djamas, Nurhayati. (2005). *Strategi Peningkatan Mutu Dosen PTAI*. Jakarta: Puslitbang Pendidikan Agama dan Keagamaan.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

Francis, J., Eccles, M. P., Johnston, M., Walker, A. E., Grimshaw, J. M., Foy, R., & Bonetti, D. (2004). *Constructing questionnaires based on the theory of planned behaviour:* A manual for health services researchers.

Ghazali Darusalam. (2021). *Metodologi Penyelidikan Dalam Penyelidikan. Amalan Dan Analisis Kajian*. University of Malaya Press.

Isaac, S., & Michael, W. B. (1995). *Handbook in research and evaluation*. San Diego, CA: Educational and Industrial Testing Services.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*: A global perspective (7th ed.). Pearson.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Harlow: Pearson Education Limited.

Harmer, J. (1990). *The practice of English language teaching. Twelfth impression*. Longman Inc.

Hill, R. (1998). What sample size is "enough" in internet survey research? Interpersonal Computing and Technology. *An Electronic Journal for the 21st Century*, 6, 3-4.

Hoe Foo Terng, Liaw Shun Chone, Lim Teck Heng and Rasaya A/L Marimuthu. (2017). Pertinent Information For Potential Mandarin Teachers In Malaysia. *International Academic Research Journal of Social Science*, 3 (1), 37-41.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6* (1), 1–55.

Huang J (2022) The Role of English as a Foreign Language Teachers' Mindfulness and Compassion in Fostering Students' Foreign Language Enjoyment. *Front. Psychol*. 13:899298. doi: 10.3389/fpsyg.2022.899298

Ibrahim, N., Che Din, N., Amit, N., Ghazali, S. E., & Mohd Safien, A. (2019). Development and validation of Yatt Suicide Attitude Scale (YSAS) in Malaysia. *PloS one*, *14* (2), e0209971.

Karnine, Seeni Mehraj Begam Bt V.K.s & Amzat, Ismail & Preece, Abdul & Kiziltas, Ercan & Khandakar, Ayesha & Muhammad, Siti Salmiah. (2023). "Turnover intention" among foreign lecturers in one of the universities in Malaysia: An investigation into the leading factors. *Environment and Social Psychology*. 9. 10.54517/esp.v9i2.1894.

Kashif, M., Awang, Z., Walsh, J., & Altaf, U. (2016). I'm loving it but hating US: Understanding consumer emotions and satisfaction/dissatisfaction in trans-national fast-food brands. *British Food Journal*, 118(9), 2279-2293.

Kashif, M., Samsi, S. Z. M., Awang, Z., & Mohamad, M. (2016). EXQ: Measurement of healthcare experience quality in Malaysian settings. *International Journal of Pharmaceutical and Healthcare Marketing*, *10*(1), 27–47.

Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, 38(1), 52-54.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacists, 65*(1), 2276-2284.

Kennedy, A. (2014). Models of continuing professional development: A framework for analysis. *Professional Development in Education*, 40 (3), 336-351.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York City: Guilford Press

KPM. (2013). *Pelan pembangunan pendidikan Malaysia 2013-2025.* Kementerian Pendidikan Malaysia.

Lee, C.-Y. & Cherner, T. S. (2015). A comprehensive evaluation rubric for assessing instructional apps. *Journal of Information Technology Education Research*, 14, 21-53.

Lewis-Beck, M.S., A.E. Bryman and T.F. Liao. (2004). *The Sage encyclopedia of social science research method*. Thousand Oaks, CA: International Educational and Professional Publisher. Sage Publication.

Liu N (2022) Integrating intercultural activities into teaching Mandarin for international students in China. *Front. Psychol.* 13:972733. doi: 10.3389/fpsyg.2022.972733

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–99.

Mahfouz, S. A., Awang, Z., & Muda, H. (2019). The impact of transformational leadership on employee commitment in the construction industry. *International Journal of Innovation, Creativity and Change*, *7*(10), 151–167.

Mercer, S., and Dörnyei, Z. (2020). *Engaging Language Learners in Contemporary Classrooms*. Cambridge: Cambridge University Press.

Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University, 17*(4), 59-82.

Mohammadi, F., Kouhpayeh, S. A., Bijani, M., & Hassani, A. (2024). Development and psychometric testing of a questionnaire for the assessment of medical science educators' adherence to ethical principles in virtual education: Exploratory sequential mixed methods study. *BMC Medical Education*, 24, 56-69

Mohamad, M., Mohammad, M., Azman, N., Ali, M., & Awang, Z. (2016). Measuring positive youth development: Confirmatory factor analysis. *International Journal of Applied Business and Economic Research, 14*(13), 9441–9451

Mohamad, M., Mohammad, M., Mat Ali, N. A., & Awang, Z. (2018). The impact of life satisfaction on substance abuse: Delinquency as a mediator. *International Journal of Adolescence and Youth*, *23*(1), 25–35.

Mukherjee, H., & Singh, J. S. (2016). Challenges in implementing education policies in Southeast Asia: The Malaysian experience. *International Journal of Education Development*, 50, 89-95.

Noordin, F., & Jusoff, K. (2009). Levels of job satisfaction amongst Malaysian academic staff. *Asian Social Science*, 5(5), 122-128.

Ntumi, S., & Twum Antwi-Agyakwa, K. (2022). A Systematic Review of Reporting of Psychometric Properties in Educational Research. *Mediterranean Journal of Social & Behavioral Research, 6*(2), 53-59.

Ormilla, A., & Abrojena, E. (2023). Challenges in the implementation of teacher evaluation systems: A review of feasibility and practicality issues. *Educational Assessment Journal*, 29 (2), 45-60.

Oztürk, Altan. (2020). The perception of an adult learner toward foreign language learning and its effects on speaking production language.10.13140/RG.2.2.22489.88168.

Pan., L., Sun , D., Zou, Y. et al. (2023). Psycho-linguistic and educational challenges in Teaching Chinese (Mandarin) Language: voices from None-Chinese teachers of Mandarin language. *BMC Psychol*, 11, 390-406.

Pan, Lei & Sun, Dan & Zou, Yumei & Cao, Yi & Zhang, Jingxian & Li, Fangfang. (2023). Psycho-linguistic and educational challenges in Teaching Chinese (Mandarin) Language: voices from None-Chinese teachers of Mandarin language. *BMC Psychology.* 11. 10.1186/s40359-023-01432-8.

Pishghadam, R., Derakhshan, A., Zhaleh, K., and Al-Obaydi, L. H. (2021). Students' willingness to attend EFL classes with respect to teachers' credibility, stroke, and success: a cross-cultural study of Iranian and Iraqi students' perceptions. *Curr. Psychol.* 40, 1–15.

Rahlin, N. A., Awang, Z., Afthanorhan, A., & Aimran, N. (2019a). Antecedents and consequences of employee safety climate in the small manufacturing enterprises: Translation, validation and application of the generic safety climate questionnaire. *International Journal of Innocation, Creativity and Change*, *7*(10), 307–328.

Rahlin, N. A., Awang, Z., Afthanorhan, A., & Aimran, N. (2019). The art of covariance based analysis in behaviour-based safety performance study using confirmatory factor analysis: Evidence from SMES. *International Journal of Innovation, Creativity and Change*, *7* (10), 351–370.

Rathakrishnan, Thanuja & Ng, Siew & Kok, T.K. (2016). Turnover intentions of lecturers in private universities in Malaysia. Pertanika *Journal of Social Sciences and Humanities*. 24. 129-146.

Rindu, I., & Ariyanti, A. (2017). Teacher's role in managing the class during the teaching and learning process. Script Journal: *Journal of Linguistic and English Teaching*, 2(1), 83–95.

Sanjaya, W. (2007). *Strategi pembelajaran* (3rd ed.). Prenada Media Group.

Saunders, M.N., (2007). Research methods for business students, 5/e. Pearson Education India.

Sekaran, U., (2003). Research methods for business: A skill building approach. John Wiley & Sons.

Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill-building Approach* (6th ed.). United Kingdon: John Wiley and Sons Ltd.

Shkeer, A. S., & Awang, Z. (2019). Exploring the items for measuring the marketing information system construct: An exploratory factor analysis. *International Review of Management and Marketing*, *9*(6), 87–97.

Sutton, S., French, D. P., Hennings, S. J., Mitchell, J., Wareham, N. J., Griffin, S., Hardeman, W., & Kinmonth, A. L. (2017). Eliciting salient beliefs in research on the theory of planned behavior: The effect of question wording. *Planned Behavior: The Relationship between Human Thought and Action*, *22*(3), 63–84.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Tao, J., & Gao, X. (2022). Teaching and learning languages online: Challenges and responses. *System*, *107*, 102819.

Wlosowicz, Teresa. (2016). Adult Learners' Expectations Concerning Foreign Language Teachers and the Teaching-Learning Process. 10.1007/978-3-319-32954-3_15.

Yusof, N. M., & Abdullah, S. (2014). Educational policy in Malaysia: The Malaysian Education Blueprint 2013-2025. *Asian Education Studies*, 1(1), 1-10.

Yusof, Y., Awang, Z., Jusoff, K., & Ibrahim, Y. (2017). The influence of green practices by non-green hotels on customer satisfaction and loyalty in hotel and tourism industry. *International Journal of Green Economics*, *11*(1), 1–14.

Zaaim, K., Mohd Nor, M. Y., & Surat, S. (2019). Tahap Kompetensi Guru (pdpc) SKPMG2 dan Tahap Kemenjadian Murid. International Journal of Education, Psychology and Counseling, 4(27), 51-62.

Zainal, Z., & Elias, S. (2018). Challenges in Implementing School-Based Assessment in Malaysia: Educators' Perspectives. *Journal of Education and Social Sciences*, 10 (2), 100-108.

Zaujan, Nur & Ali, Asma & Raihana, N & Chee, H-Y & Misni, Norashiqin & Sukeri, Surianti & Chin, Christie Pei-Yee & Osman, Malina. (2021). Validity and reliability of a questionnaire on knowledge, attitude, practice and perception (KAP2) towards food poisoning and its

prevention during dining out among consumers in Terengganu. Food Research. 5. 10.26656/fr.2017.5(4).603.

Zhao Jinming. (2007). Hanyu Zuowei Waiyu Jiaoxue Nengli Biaozhun Shishuo (Chinese as a Foreign Language Teaching Competency Standards). *Language Teaching and Research*, 2, 1-10.

Zikmund, W. G., & Babin, B. J. (2010). *Essentials of marketing research* (4th ed.).