

[THE TEACHER KNOWLEDGE INSTRUMENT FOR AL-LUGHAH AL-‘ARABIAH AL-MU‘ASIRAH (I-PGLAM): A RASCH MEASUREMENT MODEL ANALYSIS]

INSTRUMEN PENGETAHUAN GURU DALAM MATA PELAJARAN AL-LUGHAH AL-‘ARABIAH AL-MU‘ASIRAH (I-PGLAM): ANALISIS MENGGUNAKAN MODEL PENGUKURAN RASCH

Juwairiah Mohamad Ali
p124047@siswa.ukm.edu.my
Research Center for Arabic Language and Islamic Civilization,
Faculty of Islamic Studies, National University of Malaysia

Lily Hanefarezan Asbulah
lilyhane@ukm.edu.my (Corresponding Author)
Research Center for Arabic Language and Islamic Civilization,
Faculty of Islamic Studies, National University of Malaysia

Hakim Zainal
haza@ukm.edu.my
Research Center for Arabic Language and Islamic Civilization,
Faculty of Islamic Studies, National University of Malaysia

Abstrak

Keterjaminan dalam menghasilkan sebuah pengajaran yang berkesan di sesebuah institusi sama ada di sekolah rendah, menengah mahupun di Institut Pengajian Tinggi (IPT) adalah bermula dengan pengetahuan guru. Pengetahuan guru yang jitu terhadap sesuatu matapelajaran, akan berperanan sebagai elemen utama terhadap keberhasilan proses pengajaran dan pembelajaran (PdP). Maka, instrumen yang sahih dan boleh dipercayai amat diperlukan bagi menilai tahap pengetahuan guru. Kajian ini memfokuskan subjek bahasa Arab di Sekolah Agama Bantuan Kerajaan (SABK) yang dikenali dengan *al-Lughah al-‘Arabiah al-Mu‘asirah* (LAM). Justeru, kajian ini bertujuan untuk menilai kesahan dan kebolehpercayaan instrumen i-PGLAM yang dibangunkan untuk mengukur tahap pengetahuan guru (LAM). Kajian berbentuk tinjauan ini melibatkan 214 orang guru LAM di SABK seluruh Malaysia. Instrumen kaji selidik yang dibangunkan ini mempunyai 32 item yang berskala Likert empat mata. Justeru itu, perisian Winstep versi 3.72.3 telah digunakan bagi mendapatkan nilai kesahan dan kebolehpercayaan menggunakan Model Pengukuran Rasch. Hasil kajian mendapati bahawa instrumen tersebut hanya mencatatkan noise pada kadar 7%, jauh lebih rendah daripada tahap maksimum yang dibenarkan iaitu sebanyak 15%. Dapatan juga merekodkan tiga item yang berada di luar julat kesesuaian Mean Square, namun item tersebut telah melalui proses pemurnian dan ia dikekalkan. Di samping itu, analisis juga menunjukkan bahawa indeks kebolehpercayaan item berada pada 0.97, iaitu dalam kategori sangat tinggi, sementara nilai pengasingan item pula dicatatkan sebanyak 5.95 yang mengesahkan kestabilan instrumen. Maka, ini jelas membuktikan bahawa dapatan yang diperoleh mengesahkan bahawa instrumen i-PGLAM mampu mengukur konstruk pengetahuan guru dengan baik. Melalui analisis Model Rasch, instrumen ini terbukti sahih dan konsisten, sekali gus mengukuhkan justifikasi bahawa ia wajar digunakan bagi menilai tahap pengetahuan guru terhadap mata pelajaran LAM di SABK.

Kata Kunci: *pengetahuan; guru bahasa Arab; penilaian instrumen; model pengukuran Rasch.*

Abstract

The assurance of delivering effective teaching in any institution be it primary school, secondary school, or higher education—fundamentally begins with teachers' knowledge. A solid grasp of the subject matter by teachers functions as a key element in determining the success of the teaching and learning process (T&L). Therefore, a valid and reliable instrument is essential to assess the level of teachers' knowledge. This study focuses on the Arabic subject at Government-Aided Religious Schools (SABK), known as al-Lughah al-'Arabiah al-Mu'asirah (LAM). Accordingly, the aim of this research is to evaluate the validity and reliability of the i-PGLAM instrument developed to measure teachers' knowledge of LAM. The study, designed as a survey, involved 214 LAM teachers from SABK across Malaysia. The questionnaire comprised 32 items measured using a four-point Likert scale. To determine validity and reliability, Winstep software version 3.72.3 was employed, applying the Rasch Measurement Model. The findings revealed that the instrument recorded only 7% noise, which is well below the permissible maximum of 15%. The results also indicated three items falling outside the acceptable Mean Square range, however, these items were refined and subsequently retained. Moreover, the analysis showed an item reliability index of 0.97, categorized as very high, with an item separation value of 5.95, confirming the stability of the instrument. These outcomes clearly demonstrate that the i-PGLAM instrument is capable of effectively measuring teachers' knowledge construct. Through the Rasch analysis, the instrument has been proven to be valid and consistent, thus strengthening the justification for its use in assessing teachers' knowledge of LAM in SABK.

Keywords: knowledge; Arabic language teachers; instrument evaluation; Rasch Measurement Model

Article Received:
3 October 2025

Article Reviewed:
14 November 2025

Article Published:
5 December 2025

PENGENALAN

Bahasa Arab diiktiraf sebagai salah satu bahasa utama dunia yang memainkan peranan penting dalam pelbagai aspek kehidupan manusia, khususnya dalam konteks agama, ilmu pengetahuan, tamadun dan warisan peradaban Islam. Kedudukannya sebagai bahasa wahyu menjadikan bahasa Arab unik kerana ia bukan sahaja digunakan untuk komunikasi harian, tetapi juga sebagai medium utama dalam penyebaran ilmu, pembinaan tamadun, dan pemeliharaan identiti umat Islam (Bachtiar 2022; Daud 2023). Bahasa Arab juga menjadi bahasa rasmi lebih 25 negara, bahasa utama dalam organisasi antarabangsa seperti PBB, dan alat penyatuan umat Islam di seluruh dunia (Tom 2025; Bachtiar 2022).

Di Malaysia, bahasa Arab memainkan peranan signifikan kerana turut dianggap sebagai sebahagian daripada usaha mencapai pendidikan bertaraf dunia (Abdul Razif Zaini et al. 2020). Dengan komitmen yang konsisten, bahasa Arab diyakini akan terus mempunyai kedudukan yang bermakna di Malaysia dan bukan hanya terbatas pada aspek keagamaan, malah juga dalam memperkukuh pelbagai bidang seperti ekonomi, pelancongan, pendidikan dan kebudayaan (Sabri & Ilyani 2019). Tambahan pula, kini bahasa Arab dilihat telahpun mendapat tempat yang signifikan dalam sistem pendidikan, khususnya di sekolah-sekolah agama. Kedudukan bahasa Arab ini telah diperkukuh melalui dasar pendidikan Islam yang dirangka oleh Kementerian Pendidikan Malaysia (KPM) sejak beberapa dekad lalu. Manakala dalam konteks Sekolah Agama Bantuan Kerajaan (SABK), bahasa Arab diberi kedudukan yang istimewa melalui pelaksanaan Kurikulum Bersepadu Dini (KBD). Kurikulum ini menjadikan bahasa Arab sebagai salah satu subjek teras yang wajib dipelajari, menandakan komitmen Kementerian Pendidikan Malaysia dalam memperkukuh penguasaan bahasa tersebut dalam kalangan pelajar. Langkah ini bukan sahaja bertujuan memastikan kesinambungan tradisi ilmu Islam yang berasaskan sumber Arab, tetapi juga mempersiapkan pelajar dengan kemahiran berbahasa yang relevan untuk diaplikasikan

dalam pembelajaran, komunikasi seharian, serta cabaran global semasa (Firdaus et. al 2019; Fadzli & Nik Rahimi 2021).

Di SABK, KBD baru dilaksanakan pada tahun 2015 bermula di tingkatan 1 hingga 5. Namun pada tahun 2023 kurikulum ini telah ditambah baik bagi tingkatan empat dan lima agar ia setanding dengan standard antarabangsa. Kurikulum yang digunakan ini disesuaikan dengan kurikulum *Maahad Bu'uth Al-Azhar*. Hasil Inisiatif ini adalah berdasarkan pemeteraian Perjanjian Persefahaman Akademik dan Kebudayaan antara Kerajaan Malaysia dengan Universiti *Al-Azhar Al-Syarif*, Republik Arab Mesir pada 6 Jun 2022. Oleh itu, LAM ini bertujuan melahirkan murid yang bukan sahaja mahir dalam aspek teknikal bahasa Arab, tetapi juga mampu menggunakannya secara praktikal dalam pelbagai konteks. Pendekatan ini menekankan pembentukan individu yang berkemahiran tinggi dalam bahasa Arab dan memiliki peribadi yang positif, selaras dengan aspirasi pembangunan negara (KPM 2023).

Disebabkan kurikulum ini masih baru dilaksanakan, maka terdapat beberapa cabaran yang dihadapi dalam kalangan guru. Di antara cabaran yang dihadapi oleh guru dalam melaksanakan pengajaran dan pembelajaran (PdP) bahasa Arab, adalah tahap kesediaan guru untuk melaksanakan kurikulum ini. Kajian mendapati bahawa sebahagian guru masih berpegang pada pendekatan tradisional yang terlalu menekankan hafalan dan terjemahan literal, sedangkan LAM menuntut pendekatan komunikatif (Abdul Razif Zaini et al. 2019). Di samping itu, faktor kekangan masa, beban kerja pentadbiran, serta kurangnya latihan profesionalisme berterusan turut memberi kesan kepada keupayaan guru untuk melaksanakan kurikulum ini secara optimum (Awatif et al. 2020).

Selain faktor guru, terdapat juga cabaran daripada segi tahap penguasaan pelajar yang tidak seimbang. Seseengah pelajar mempunyai asas bahasa Arab yang kuat, manakala sebahagian lain pula masih lemah dalam asas tatabahasa dan kosa kata. Hal ini menyebabkan guru perlu melaksanakan pengajaran yang berbeza tahap kesukaran, yang kadangkala sukar dilakukan dalam kelas yang besar (Khairulaizam & Harun 2022). Tambahan pula, kekangan sumber pembelajaran, seperti bahan bantu mengajar (BBM) dalam bahasa Arab, juga menjejaskan keberkesanan pelaksanaan LAM (Fariyah Hussain et al. 2020).

Para sarjana pendidikan bersepakat bahawa pengetahuan guru adalah faktor utama yang mempengaruhi keberkesanan sesuatu proses pengajaran (Abiola & Yusuf 2023; Fauth et al. 2019; Lena 2024; Nathan & Jacqueline 2019). Shulman (1986), melalui kerangka *Pedagogical Content Knowledge* (PCK), menekankan bahawa keberkesanan guru bukan sahaja bergantung kepada penguasaan isi kandungan, malah kepada keupayaan menyesuaikan kaedah penyampaian dengan keperluan pelajar. Dalam konteks pendidikan moden, teori *Technological Pedagogical Content Knowledge* (TPACK) oleh Koehler, Mishra & Cain (2009) pula memperluas konsep ini dengan menekankan integrasi teknologi dalam pengajaran. Bagi mata pelajaran LAM, pengetahuan guru terhadap kurikulum, isi kandungan, serta kaedah penilaian merupakan faktor kritikal yang mempengaruhi pencapaian pelajar. Justeru itu, pengetahuan guru bahasa Arab terhadap kurikulum memberi kesan langsung terhadap pemahaman pelajar untuk menguasai isi kandungan (Halim Tamuri & Hussin 2017).

MODEL PENGUKURAN RASCH

Dalam menilai kesahan dan kebolehpercayaan instrumen, pelbagai pendekatan statistik boleh digunakan. Namun, kaedah klasik seperti ujian kebolehpercayaan Cronbach Alpha sahaja tidak mencukupi. Menurut Bond & Fox (2015) model pengukuran Rasch memberikan analisis yang lebih komprehensif kerana ia menilai item dan responden secara serentak dalam skala logit. Rasch juga membolehkan penyelidik mengenal pasti item yang tidak sesuai, mengukur unidimensionaliti, serta memetakan hubungan antara tahap kesukaran item dengan tahap kebolehan responden (Boone 2014; Linacre 2002).

Berbanding teori ujian klasik (Classical Test Theory, CTT), Rasch mempunyai kelebihan kerana ia menghasilkan ukuran berskala interval yang lebih objektif. Selain itu, Rasch dapat mengasingkan kebolehpercayaan item dan responden, sekali gus membantu penyelidik menentukan sama ada masalah terletak pada instrumen atau sampel kajian (Bambang & Wahyu 2014a). Tambahan pula, Rasch membolehkan analisis polariti item dilakukan melalui *Point Measure Correlation* (PTMEA CORR), serta mengukur keseragaman dimensi melalui *Principal Component Analysis* (PCA) of Residuals.

Beberapa kajian lepas telah menggunakan model pengukuran Rasch dalam kajian mereka, seperti kajian Rafidah & Mohd Effendi (2019) mengesahkan instrumen Cabaran Guru Pelatih Praktikum (I-CGPKM) dengan kebolehpercayaan item 0.96 dan responden 0.91. Seterusnya, Hasnurol et. al (2022) membangunkan Ujian Struktur Dalam Kata Imbuhan (USDai) bagi mengukur morfologi Arab dengan kebolehpercayaan individu 0.95 dan item 0.87. Selain itu, kajian Nur Afifah & Lily Hanefarezan (2022) melalui instrumen i-SPPBA turut melaporkan kebolehpercayaan item 0.96 dan responden 0.92. Berdasarkan kajian-kajian ini, jelas membuktikan bahawa Model Rasch mampu memastikan kesahan konstruk serta kebolehpercayaan instrumen dalam pelbagai konteks pendidikan bahasa Arab.

PENYATAAN MASALAH

Kajian terdahulu menunjukkan bahawa pelaksanaan mata pelajaran KBD masih mendepani pelbagai cabaran, terutama unsur linguistik yang terdapat di dalam buku teks kurikulum dini agak kompleks. Hal ini dikukuhkan oleh kajian Ahmad Sabri (2017) yang menjalankan kajian terhadap kurikulum dini, yang memfokuskan unsur-unsur linguistik dalam kebolehbacaan buku teks berbahasa Arab. Kajian beliau mendapati bahawa tahap penguasaan dalam penggunaan ayat mudah adalah pada kategori rendah. Manakala ayat kompleks dan ayat kata nama berada pada kategori sederhana. Namun, tahap ayat kompleks lebih tinggi daripada ayat mudah. Selain itu, kajian Hamidzah & Isa (2023) mengenai kesediaan guru bukan ikhtisas bahasa Arab pula menyatakan bahawa tahap pengetahuan guru berada pada tahap yang rendah. Justeru itu, bagi merealisasikan hasrat KPM terhadap kurikulum ini, guru seharusnya dipersiapkan dengan pengetahuan yang jitu bagi menjamin keberkesanan dan kejayaan di dalam PdP. Ini kerana, kekurangan pengetahuan guru terhadap ilmu pedagogi akan mempengaruhi kualiti pengajaran bahasa Arab (Hazrul 2020).

Walaupun banyak kajian telah dijalankan untuk menilai sikap atau motivasi pelajar terhadap pembelajaran bahasa Arab, masih kurang kajian yang menumpukan kepada pengetahuan guru, khususnya dalam konteks LAM di SABK. Ini kerana, pengetahuan guru merupakan sumber kekuatan bagi seseorang guru dalam menjamin prestasi terbaik dalam profesion mereka. Guru yang tidak mempunyai pengetahuan yang jelas dan menyeluruh akan mempunyai masalah dalam penyampaian kepada pelajar (Zaffi 2021). Oleh itu, jurang yang wujud ini menyebabkan tahap sebenar pengetahuan guru tidak dapat dikenal pasti secara menyeluruh.

Walau bagaimanapun, sorotan literatur mendapati kajian mengenai kesahan dan kebolehpercayaan instrumen khusus untuk mengukur tahap pengetahuan guru LAM masih sangat terhad. Terdapat beberapa penyelidikan berkaitan kesahan dan kebolehpercayaan instrumen, namun fokusnya lebih kepada aspek sikap guru Matematik di sekolah menengah (Norkumalasari et al. 2020), dan sikap pelajar terhadap kursus bahasa arab di IPTA (Nur Afifah & Lily Hanefarezan 2022) serta pengesahan instrumen ujian struktur dalaman kata imbuhan (USDai) bagi mengukur keupayaan morfologi Arab (Hasnurol et.al 2022). Maka, ini jelas menunjukkan bahawa skop kajian kesahan dan kebolehpercayaan instrumen soal selidik khusus untuk pengetahuan guru bahasa Arab di SABK masih kurang dijalankan.

Sehubungan itu, kajian ini dijalankan dengan objektif utama untuk menilai kesahan konstruk dan kebolehpercayaan instrumen i-PGLAM dengan menggunakan model pengukuran

Rasch. Analisis ini melibatkan pemeriksaan terhadap kebolehppercayaan item dan responden, indeks kesesuaian, polariti item, unidimensionaliti, serta peta item-responden. (Bond dan Fox 2015). Maka, dengan adanya sebuah alat pengukuran yang jitu, sudah tentu akan mengukur tahap pengetahuan guru dengan baik. Ia bertepatan dengan kenyataan (Azizi et.al 2017) yang mengatakan bahawa sesebuah instrumen yang dibangunkan mestilah mempunyai kesahan, iaitu berfungsi untuk menilai aspek yang sepatutnya diukur.

METODOLOGI

Kajian ini menggunakan reka bentuk kuantitatif berbentuk tinjauan untuk menilai kesahan serta kebolehppercayaan instrumen i-PGLAM. Seramai 214 orang guru yang mengajar mata pelajaran *al-Lughah al-'Arabiah al-Mu'asirah* (LAM) di Sekolah Agama Bantuan Kerajaan (SABK) seluruh Malaysia telah dijadikan sampel kajian. Pemilihan sampel dibuat melalui kaedah persampelan rawak, melibatkan guru-guru yang sedang mengajar mata pelajaran LAM di tingkatan empat dan lima.

Instrumen soal selidik i-PGLAM ini dibahagikan kepada dua bahagian. Bahagian A memfokuskan kepada maklumat latar belakang responden, merangkumi jantina, gred jawatan, pengkhususan ijazah sarjana muda, opsyen pengajian, pengalaman mengajar bahasa Arab, pengalaman mengajar berasaskan kurikulum, serta latar belakang pembelajaran kitab turath seperti *Syarah Ibn 'Aqil* dan *Sarf al-Muyassar*. Bahagian B pula mengandungi 32 item yang menilai tahap pengetahuan guru terhadap mata pelajaran LAM, meliputi tiga aspek utama: pengetahuan terhadap isi kandungan kurikulum LAM, kefahaman mengenai KBD, serta pengetahuan berkaitan penilaian dan peperiksaan.

Disamping itu, kajian ini menggunakan skala Likert 4 mata iaitu: 1 (sangat tidak setuju), 2 (tidak setuju), 3 (setuju), dan 4 (sangat setuju) bagi mengukur tahap persetujuan responden. Pemilihan skala tanpa opsyen neutral bertujuan mengelakkan kecenderungan responden untuk memilih jawapan tengah tanpa menggambarkan pendirian mereka yang sebenar (Bambang & Wahyu 2014).

Manakala proses pengumpulan data dijalankan sepenuhnya secara atas talian menggunakan borang soal selidik berbentuk *Google Form*. Bagi memastikan kelancaran serta keabsahan prosedur, kebenaran rasmi diperoleh terlebih dahulu daripada Bahagian Perancangan dan Penyelidikan Dasar Pendidikan (EPRD) KPM, Bahagian Pendidikan Islam (BPI) KPM, Jabatan Pendidikan Negeri (JPN) seluruh Malaysia, serta pengetua sekolah-sekolah terlibat. Justeru, bagi mendapatkan kesahan dan kebolehppercayaan instrumen i-PGLAM, data yang dikumpulkan telah dianalisis menggunakan model pengukuran Rasch dengan bantuan perisian Winsteps versi 3.72.3.

Kesahan Instrumen Kajian

Bagi memastikan ketepatan sesuatu instrumen, ia perlu melalui semakan kesahan muka (*face validity*) dan kesahan kandungan (*content validity*), yang menilai dari segi bahasa, susunan, serta struktur ayat agar sepadan dengan konstruk pengukuran dan maksud item (Mohd Najib 2003). Ia bertujuan menjamin kualiti instrumen dan sejauh mana item-item yang digunakan memenuhi objektif kajian dan dapat mengukur apa yang hendak diukur (Abdull Sukor 2022). Proses berkenaan sangat signifikan kerana ia menentukan ketepatan data penyelidikan, seterusnya memastikan kesemua objektif dan persoalan kajian dapat dicapai.

Dalam kajian ini, instrumen soal selidik yang digubal telah melalui proses kesahan kandungan bersama-sama dengan tujuh orang panel pakar yang dipilih. Panel pakar tersebut terdiri daripada pensyarah dari Fakulti Pengajian Bahasa Utama (USIM), Fakulti Bahasa dan

Linguistik (UM), Fakulti Pendidikan (UIAM), Guru Cemerlang LAM dan Panel Penggubal DSKP LAM KBD. Pemilihan panel pakar ini adalah berdasarkan kepakaran mereka dalam bidang pedagogi pengajaran bahasa Arab terutamanya aspek pengetahuan, kemahiran, sikap dan amalan guru dalam PdP bahasa Arab serta pembinaan soal selidik.

Setelah melalui proses penilaian pakar, pengkaji menilai kesahan kandungan instrumen menggunakan Nisbah Kesahan Kandungan (*Content Validity Ratio*, CVR) dan Indeks Kesahan Kandungan (*Content Validity Index*, CVI). Kedua-dua kaedah ini merupakan prosedur kuantitatif yang diperkenalkan oleh Lawshe (1975) bagi mengukur kesahan kandungan. CVR digunakan untuk menilai setiap item secara empirikal dengan menentukan tahap kepentingannya dalam instrumen. Kaedah ini membantu pengkaji menentukan sama ada sesuatu item perlu dikekalkan atau digugurkan (Mohd Effendi et al. 2017). Julat nilai CVR adalah antara -1 hingga +1, dan semakin hampir sesuatu nilai kepada +1, semakin tinggi tahap persetujuan panel pakar terhadap kepentingan item berkenaan dalam memastikan kesahan kandungan. Menurut Lawshe (1975), nilai CVR yang melebihi sifar ($CVR > 0$) menandakan majoriti panel pakar bersetuju bahawa item tersebut adalah penting serta menepati kriteria kesahan kandungan. Secara umumnya, analisis mendapati bahawa setiap item dalam instrumen i-PGLAM memperoleh pekali persetujuan lebih besar daripada sifar, menandakan kesemua item tersebut sah mewakili domain konstruk yang ingin diukur dan mempunyai tahap kerelevanan yang tinggi.

Selepas memperoleh nilai CVR bagi setiap item, pengiraan dibuat untuk mendapatkan nilai CVI keseluruhan instrumen. Menurut Lynn (1986) bilangan panel pakar yang disarankan adalah di antara tiga hingga sepuluh orang. Nilai CVI yang diterima adalah >0.80 iaitu berdasarkan purata tahap persetujuan antara panel pakar yang dipilih (Lynn 1986). Secara keseluruhannya, nilai CVI bagi setiap item dalam instrumen i-PGLAM melepasi pekali persetujuan 0.80. Hal ini menunjukkan bahawa instrumen tersebut telah diiktiraf oleh panel pakar sebagai sesuai dan mampu menilai kandungan yang ingin diuji. Di samping itu, pandangan pakar terhadap kandungan soal selidik dapat disimpulkan seperti berikut:

1. Item perlu disusun semula bagi memastikan aliran dan kesinambungan yang lebih baik antara satu sama lain.
2. Terdapat item yang bertindih kerana mengukur konstruk yang sama.
3. Terdapat item yang perlu di semak semula agar selari dan jelas dengan konstruk yang ingin diukur.

Sehubungan itu, pengkaji telah menambahbaik instrumen i-PGLAM berdasarkan saranan serta pandangan pakar. Instrumen yang telah dimurnikan ini kemudiannya diedarkan kepada guru, dan data yang dikumpul dianalisis untuk menentukan kebolehpercayaan item.

ANALISIS DAN PERBINCANGAN

Indeks Kebolehpercayaan dan Pemisahan Item-Responden

Melalui analisis Rasch, kebolehpercayaan item dan responden dapat dinilai bagi menentukan sejauh mana item mematuhi keperluan model pengukuran serta menunjukkan tahap pemisahan antara item dan responden. Melalui indeks pemisahan, ia dapat menggambarkan jumlah strata kesukaran item bersama tahap kebolehan individu yang berjaya dikesan dalam kalangan sampel kajian. Bagi instrumen i-PGLAM, nilai Alpha Cronbach dicatatkan pada 0.97, yang mengesahkan kesahan kebolehpercayaan dengan konsistensi dalaman yang sangat kukuh. Pandangan Bond & Fox (2015) dalam Jadual 1 turut menyokong dapatan ini, dengan menyatakan bahawa kebolehpercayaan pada tahap berkenaan menggambarkan instrumen yang baik serta mempunyai tahap konsistensi yang mantap.

JADUAL 1. Interpretasi Skor Alpha Cronbach

Skor Alpha Cronbach	Interprestasi Nilai kebolehpercayaan
0.9-1.0	Sangat baik dan efektif dengan tahap konsistensi yang tinggi
0.7-0.8	Baik dan boleh diterima
0.6-0.7	Boleh diterima
<0.6	Item perlu dibaiki
<0.5	Item perlu digugurkan

Berdasarkan analisis Rasch, kebolehpercayaan item untuk 32 item i-PGLAM adalah 0.97 dengan nilai pemisahan item 5.95, sebagaimana dipaparkan dalam Jadual 2. Menurut Bond dan Fox (2015), nilai kebolehpercayaan item yang melebihi 0.80 menunjukkan tahap konsistensi yang sangat baik dan berkesan. Nilai pemisahan item 6.15 pula boleh ditafsirkan sebagai tinggi serta memenuhi kriteria yang dikemukakan oleh Linacre (2002), iaitu pemisahan item melebihi 2.00 menandakan kualiti instrumen berada pada tahap yang memuaskan.

Sementara itu, kebolehpercayaan responden bagi i-PGLAM mencatatkan nilai 0.94 dengan indeks pemisahan 4.03 seperti yang dirumuskan dalam Jadual 3. Kebolehpercayaan responden yang melebihi 0.80 membuktikan bahawa jawapan yang diberikan adalah konsisten dan boleh dipercayai (Bond & Fox, 2015). Selain itu, nilai pemisahan responden yang melebihi 2.00 (Linacre, 2002) turut menegaskan bahawa lebih daripada empat strata keupayaan responden dapat dikenal pasti melalui kajian ini. Secara keseluruhannya, analisis Rasch yang dijalankan memperlihatkan bahawa instrumen i-PGLAM mencapai standard kesahan dan kebolehpercayaan yang tinggi, justeru memberi jaminan terhadap kesesuaian dan kesahannya dalam mengukur aspek tahap pengetahuan guru dalam matapelajaran LAM.

JADUAL 2. Nilai Kebolehpercayaan dan Pemisahan Item terhadap Instrumen i-PGLAM

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	711.5	214.0	.00	.18	.98	-.2	.97	-.3
S.D.	34.9	.0	1.11	.00	.20	1.8	.27	1.6
MAX.	805.0	214.0	2.44	.20	1.61	4.3	1.68	3.5
MIN.	631.0	214.0	-3.12	.16	.57	-4.3	.47	-4.2
REAL RMSE	.18	TRUE SD	1.10	SEPARATION	5.95	ITEM	RELIABILITY	.97
MODEL RMSE	.18	TRUE SD	1.10	SEPARATION	6.15	ITEM	RELIABILITY	.97
S.E. OF ITEM MEAN = .20								

JADUAL 3. Nilai Kebolehpercayaan dan Pemisahan Responden terhadap Instrumen i-PGLAM

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	104.8	32.0	3.70	.46	.96	-.3	.97	-.3
S.D.	11.4	.0	2.15	.11	.62	1.9	.78	1.9
MAX.	127.0	32.0	8.92	1.04	5.00	5.6	5.01	5.1
MIN.	48.0	32.0	-4.20	.34	.14	-3.4	.10	-3.4
REAL RMSE	.52	TRUE SD	2.08	SEPARATION	4.03	PERSON RELIABILITY		.94
MODEL RMSE	.48	TRUE SD	2.09	SEPARATION	4.40	PERSON RELIABILITY		.95
S.E. OF PERSON MEAN = .15								

Polariti Item

Pemeriksaan awal terhadap kesahan konstruk dijalankan melalui analisis polariti item menggunakan indeks *Point Measure Correlation* (PTMEA CORR), yang bertujuan mengenal pasti sama ada semua item bergerak secara konsisten dalam kerangka konstruk yang sama. (Bond & Fox, 2015). Berdasarkan pandangan Linacre (2005), item dianggap telah mengukur konstruk yang disasarkan sekiranya indeks menunjukkan nilai positif, dengan arah pengukuran yang konsisten dengan objektif yang telah ditetapkan. Sebaliknya, item dengan nilai negatif perlu diberi perhatian khusus, sama ada melalui proses pemurnian atau digugurkan daripada instrumen. Tambahan pula, Nunnally & Bernstein (1994) menegaskan bahawa nilai PTMEA CORR di bawah 0.30 mencerminkan item yang lemah atau tidak berfungsi dengan baik.

Berdasarkan Jadual 4, nilai PTMEA CORR bagi instrumen i-PGLAM berada dalam julat 0.45 hingga 0.78. Justeru, dapatan tersebut jelas memperlihatkan bahawa setiap item mencapai nilai positif melepasi 0.30, sekali gus menandakan instrumen ini sah dan konsisten. Hal ini membuktikan bahawa item-item dalam i-PGLAM berupaya membezakan tahap pengetahuan guru terhadap mata pelajaran LAM dengan baik, serta menyokong kesahan konstruk instrumen secara keseluruhan.

JADUAL 4. Nilai Polariti Item

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	ITEM
13	745	214	-1.06	.18	1.03	.3	1.68	2.7	A .66	.66	77.4	78.5	B13
30	675	214	1.17	.18	1.61	4.3	1.62	3.5	B .64	.72	70.4	79.1	D09
31	631	214	2.44	.16	1.29	2.4	1.48	2.9	C .67	.74	69.8	75.8	D10
2	750	214	-1.22	.18	1.23	2.1	1.40	1.7	D .63	.65	76.9	78.6	B02
1	805	214	-3.12	.20	1.38	3.5	1.18	.6	E .45	.51	77.4	81.1	B01
32	648	214	1.97	.17	1.03	.3	1.18	1.2	F .71	.73	76.4	77.3	D11
27	675	214	1.17	.18	1.05	.4	1.13	.9	G .66	.72	80.9	79.1	D06
22	702	214	.32	.18	1.08	.7	1.11	.7	H .65	.71	76.4	78.8	D01
14	741	214	-.93	.18	1.11	1.0	1.05	.3	I .65	.67	73.4	78.3	B14
20	726	214	-.45	.18	1.07	.6	1.03	.2	J .67	.69	78.4	78.3	C06
9	746	214	-1.09	.18	1.06	.6	1.06	.4	K .67	.66	81.9	78.5	B09
5	748	214	-1.16	.18	1.06	.6	.95	-.2	L .68	.66	79.4	78.6	B05
6	718	214	-.19	.18	1.04	.4	.98	-.1	M .70	.69	78.9	78.3	B06
7	743	214	-1.00	.18	1.03	.4	.89	-.4	N .68	.66	79.9	78.4	B07
8	710	214	.06	.18	1.03	.3	.94	-.3	O .73	.70	73.4	78.5	B08
3	749	214	-1.19	.18	1.02	.2	.97	-.1	P .67	.65	80.4	78.6	B03
29	693	214	.60	.18	1.01	.2	1.00	.0	p .69	.71	80.9	79.1	D08
10	724	214	-.39	.18	.97	-.2	.86	-.7	o .72	.69	77.4	78.3	B10
4	728	214	-.51	.18	.89	-1.0	.94	-.3	n .72	.68	79.9	78.3	B04
11	705	214	.22	.18	.92	-.6	.83	-1.0	m .72	.70	78.9	78.7	B11
21	714	214	-.07	.18	.91	-.8	.85	-.8	l .72	.70	80.4	78.4	C07
15	716	214	-.13	.18	.90	-.9	.81	-1.1	k .72	.70	83.4	78.4	C01
19	727	214	-.48	.18	.88	-1.2	.73	-1.5	j .70	.68	81.4	78.3	C05
16	716	214	-.13	.18	.86	-1.3	.76	-1.5	i .73	.70	82.4	78.4	C02
12	736	214	-.77	.18	.86	-1.4	.86	-.7	h .70	.67	82.4	78.3	B12
28	658	214	1.68	.17	.82	-1.6	.80	-1.4	g .73	.73	84.4	78.3	D07
24	687	214	.79	.18	.81	-1.7	.78	-1.4	f .74	.72	83.9	79.2	D03
18	715	214	-.10	.18	.79	-2.0	.77	-1.4	e .71	.70	85.4	78.4	C04
23	691	214	.67	.18	.76	-2.2	.70	-2.0	d .75	.71	85.9	79.2	D02
17	679	214	1.04	.18	.74	-2.4	.74	-1.8	c .75	.72	85.4	79.2	C03
26	678	214	1.08	.18	.63	-3.6	.55	-3.5	b .77	.72	86.4	79.2	D05
25	688	214	.76	.18	.57	-4.3	.47	-4.2	a .78	.72	90.5	79.2	D04
MEAN	711.5	214.0	.00	.18	.98	-.2	.97	-.3			80.0	78.6	
S.D.	34.9	.0	1.11	.00	.20	1.8	.27	1.6			4.5	.8	

Kesesuaian Item (Item Fit)

Melalui analisis berasaskan model pengukuran Rasch, tahap kesesuaian item dapat dikenal pasti dalam mengukur konstruk yang ingin dikaji. Pemeriksaan kesesuaian item biasanya dimulakan dengan melihat nilai Mean-Square (MNSQ) sama ada INFIT atau OUTFIT, di mana julat yang boleh diterima ialah antara 0.5 hingga 1.5. Item yang mempunyai nilai MNSQ kurang daripada 0.5 menandakan ia bertindih dengan item lain, manakala item yang melebihi 1.5 pula dianggap tidak seiring dengan item-item lain dalam skala yang sama. Secara umumnya, nilai MNSQ yang berada di luar julat ini turut disertai dengan skor ZSTD yang melebihi julat -2.0 hingga +2.0.

Menurut Bond et.al (2021) ketidakselarasan pada statistik INFIT MNSQ lazimnya lebih membimbangkan berbanding OUTFIT MNSQ kerana ia menggambarkan sejauh mana responden memberikan jawapan yang konsisten dengan tahap kesukaran item. Walaupun begitu, OUTFIT MNSQ masih penting untuk mengesan sama ada terdapat respon luar jangka terhadap sesuatu item.

Berdasarkan Jadual 5, didapati tiga item yang keluar daripada julat kesesuaian. Item yang mencatatkan nilai MNSQ melebihi 1.5 ialah B13 (1.68) dan D09 (1.62), manakala item yang lebih rendah daripada 0.5 ialah D04 (0.47). Berdasarkan dapatan tersebut, tiga item ini memerlukan analisis lanjut bagi menentukan keperluan sama ada untuk dikekalkan selepas pemurnian atau digugurkan daripada instrumen. Namun begitu, berdasarkan saranan Boone et al. (2014), kesesuaian item seharusnya ditentukan melalui gabungan tiga kriteria iaitu MNSQ, ZSTD dan PTMEA CORR. Daripada pemerhatian, ketiga-tiga item tersebut (B13, D09 dan D04) telah

memenuhi satu kriteria yang melepasi julat iaitu PTMEA CORR. Selaras dengan pandangan Bambang dan Wahyu (2014), item yang masih memenuhi sekurang-kurangnya satu daripada tiga kriteria adalah wajar untuk dikekalkan. Malah, (Azrilah et. al 2013) turut menekankan bahawa sesuatu item hanya dianggap benar-benar tidak sesuai sekiranya gagal pada ketiga-tiga kriteria tersebut. Justeru, kesemua item ini telah dimurnikan dan dikekalkan dalam i-PGLAM ini.

JADUAL 5. Senarai Ketidaksesuaian Item (Item Misfit)

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFI MNSQ	Z TD	OUTFIT MNSQ	Z TD	PT-MEASURE CORR.	EXACT EXP.	EXACT OBS%	MATCH EXP%	ITEM
13	745	214	-1.06	.18	1.03	.3	1.68	2.7	A .66	.66	77.4	78.5	B13
30	675	214	1.17	.18	1.61	.3	1.62	3.5	B .64	.72	70.4	79.1	D09
31	631	214	2.44	.16	1.29	2.4	1.48	2.9	C .67	.74	69.8	75.8	D10
2	750	214	-1.22	.18	1.23	2.1	1.40	1.7	D .63	.65	76.9	78.6	B02
1	805	214	-3.12	.20	1.38	3.5	1.18	.6	E .45	.51	77.4	81.1	B01
32	648	214	1.97	.17	1.03	.3	1.18	1.2	F .71	.73	76.4	77.3	D11
27	675	214	1.17	.18	1.05	.4	1.13	.9	G .66	.72	80.9	79.1	D06
22	702	214	.32	.18	1.08	.7	1.11	.7	H .65	.71	76.4	78.8	D01
14	741	214	-.93	.18	1.11	1.0	1.05	.3	I .65	.67	73.4	78.3	B14
20	726	214	-.45	.18	1.07	.6	1.03	.2	J .67	.69	78.4	78.3	C06
9	746	214	-1.09	.18	1.06	.6	1.06	.4	K .67	.66	81.9	78.5	B09
5	748	214	-1.16	.18	1.06	.6	.95	-.2	L .68	.66	79.4	78.6	B05
6	718	214	-.19	.18	1.04	.4	.98	-.1	M .70	.69	78.9	78.3	B06
7	743	214	-1.00	.18	1.03	.4	.89	-.4	N .68	.66	79.9	78.4	B07
8	710	214	.06	.18	1.03	.3	.94	-.3	O .73	.70	73.4	78.5	B08
3	749	214	-1.19	.18	1.02	.2	.97	-.1	P .67	.65	80.4	78.6	B03
29	693	214	.60	.18	1.01	.2	1.00	.0	p .69	.71	80.9	79.1	D08
10	724	214	-.39	.18	.97	-.2	.86	-.7	o .72	.69	77.4	78.3	B10
4	728	214	-.51	.18	.89	-1.0	.94	-.3	n .72	.68	79.9	78.3	B04
11	705	214	.22	.18	.92	-.6	.83	-1.0	m .72	.70	78.9	78.7	B11
21	714	214	-.07	.18	.91	-.8	.85	-.8	l .72	.70	80.4	78.4	C07
15	716	214	-.13	.18	.90	-.9	.81	-1.1	k .72	.70	83.4	78.4	C01
19	727	214	-.48	.18	.88	-1.2	.73	-1.5	j .70	.68	81.4	78.3	C05
16	716	214	-.13	.18	.86	-1.3	.76	-1.5	i .73	.70	82.4	78.4	C02
12	736	214	-.77	.18	.86	-1.4	.86	-.7	h .70	.67	82.4	78.3	B12
28	658	214	1.68	.17	.82	-1.6	.80	-1.4	g .73	.73	84.4	78.3	D07
24	687	214	.79	.18	.81	-1.7	.78	-1.4	f .74	.72	83.9	79.2	D03
18	715	214	-.10	.18	.79	-2.0	.77	-1.4	e .71	.70	85.4	78.4	C04
23	691	214	.67	.18	.76	-2.2	.70	-2.0	d .75	.71	85.9	79.2	D02
17	679	214	1.04	.18	.74	-2.4	.74	-1.8	c .75	.72	85.4	79.2	C03
26	678	214	1.08	.18	.63	-3.6	.55	-3.5	b .77	.72	86.4	79.2	D05
25	688	214	.76	.18	.57	-4.3	.47	-4.2	a .78	.72	90.5	79.2	D04
MEAN	711.5	214.0	.00	.18	.98	-.2	.97	-.3			80.0	78.6	
S.D.	34.9	.0	1.11	.00	.20	1.8	.27	1.6			4.5	.8	

Nilai Korelasi Residual Terpiawai

Bagi mengenal pasti kemungkinan pertindihan atau gangguan antara item dalam mengukur konstruk, ujian korelasi residual terpiawai telah dijalankan. Apabila dua item mempunyai korelasi residual yang tinggi, ia memberi isyarat bahawa kedua-duanya tidak berdiri secara bebas, sebaliknya berkongsi persamaan atau mengukur dimensi tambahan di luar konstruk asal. Menurut Linacre (2005), nilai korelasi residual maksimum yang masih boleh diterima ialah 0.70; melebihi paras ini menandakan salah satu item perlu dipertimbangkan untuk digugurkan.

Berdasarkan Jadual 6, sepuluh pasangan item dalam instrumen i-PGLAM menunjukkan nilai korelasi residual yang berada dalam julat 0.38 hingga 0.54. Ini membuktikan bahawa semua pasangan item mematuhi syarat kebebasan setempat (*local independence*) dan tidak wujud isu pertindihan yang serius. Dengan demikian, boleh disimpulkan bahawa kesemua item dalam i-PGLAM adalah tidak bertindan satu sama lain dan tidak mengelirukan responden.

JADUAL 6. Nilai Korelasi Terpiawai Terbesar pada Item

CORREL- ATION	ENTRY NUMBER ITE	ENTRY NUMBER ITE
.54	5 B05	7 B07
.50	6 B06	8 B08
.48	3 B03	5 B05
.47	4 B04	6 B06
.47	8 B08	10 B10
.45	23 D02	24 D03
.42	4 B04	8 B08
.39	20 C06	21 C07
.38	27 D06	28 D07
.38	3 B03	7 B07

Keseragaman Dimensi

Keseragaman dimensi merupakan aspek penting dalam memastikan instrumen benar-benar mengukur konstruk yang sama tanpa dipengaruhi oleh faktor sampingan. Instrumen yang bersifat multidimensi berpotensi menghasilkan keputusan yang mengelirukan serta menjejaskan kesahan pengukuran (Azrilah et al., 2013). Oleh itu, pemeriksaan unidimensionaliti dalam model Rasch dilaksanakan menggunakan *Principal Component Analysis (PCA) of Residuals* untuk menentukan sama ada instrumen benar-benar berfungsi secara konsisten dalam satu dimensi dengan tahap gangguan (*noise*) yang minimum.

Berdasarkan Jadual 7, nilai varians kasar yang diterangkan oleh ukuran dalam instrumen i-PGLAM ialah 51.3%. Walaupun nilai ini belum mencapai tahap ideal 60% seperti yang dicadangkan oleh Linacre (2002), namun ia tetap berada pada kategori baik mengikut pandangan Bambang dan Wahyu. Mereka menyatakan bahawa nilai melebihi 20% sudah mencukupi, 40% adalah lebih baik, dan 60% ke atas dianggap istimewa. (Bambang & Wahyu 2014). Oleh itu, dapatan ini mengesahkan bahawa instrumen i-PGLAM mempunyai unidimensionaliti yang kukuh dan layak digunakan bagi menilai pengetahuan guru terhadap LAM. Tambahan pula, nilai kontras pertama yang direkodkan hanyalah 7.0%, jauh di bawah ambang 15% yang ditetapkan oleh Fisher (2007). Dapatan ini menandakan bahawa gangguan yang wujud tidak cukup kuat untuk membentuk dimensi baharu, sebaliknya semua item masih bergerak dalam arah konstruk yang sama.

		-- Empirical --		Modeled
Total raw variance in observations	=	65.7	100.0%	100.0%
Raw variance explained by measures	=	33.7	51.3%	50.4%
Raw variance explained by persons	=	21.2	32.2%	31.6%
Raw Variance explained by items	=	12.6	19.1%	18.8%
Raw unexplained variance (total)	=	32.0	48.7%	100.0%
Unexplnd variance in 1st contrast	=	4.6	7.0%	14.5%
Unexplnd variance in 2nd contrast	=	2.4	3.7%	7.6%
Unexplnd variance in 3rd contrast	=	2.1	3.3%	6.7%
Unexplnd variance in 4th contrast	=	2.0	3.1%	6.3%
Unexplnd variance in 5th contrast	=	1.7	2.6%	5.4%

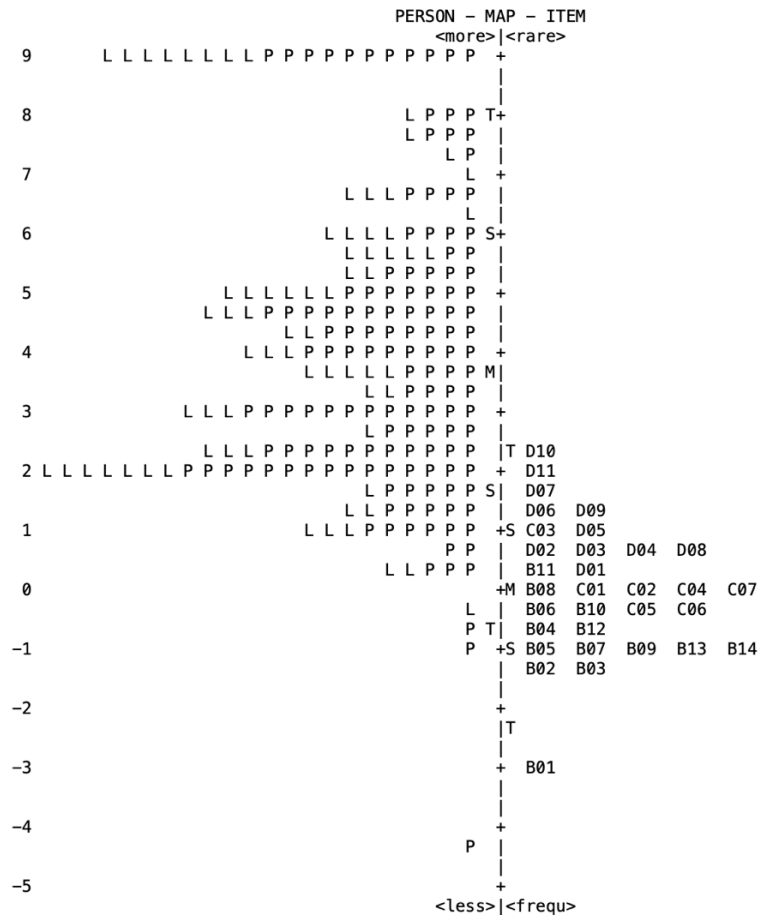
Peta Item-Responden

Dalam peta item–responden, taburan responden digambarkan di sebelah kiri manakala item ditempatkan di sebelah kanan pada skala logit. Peta ini bertujuan menunjukkan hubungan antara tahap kesukaran item dengan keupayaan responden (Rafidah & Effendi 2019). Susunan skala dalam peta ini berbentuk menurun, iaitu item dengan tahap kesukaran tertinggi berada di atas, manakala item dengan tahap persetujuan paling mudah diletakkan di bawah.

Berdasarkan Rajah 1, didapati min keupayaan responden berada pada tahap yang lebih tinggi berbanding min kesukaran item. Hal ini menunjukkan bahawa secara keseluruhan, instrumen i-PGLAM sesuai digunakan kerana tahap kesukaran item seimbang dengan keupayaan responden untuk menjawabnya. Dari segi keupayaan, sebahagian besar responden tertumpu pada logit antara +2 hingga +6, menandakan kebolehan yang sederhana hingga tinggi. Terdapat juga beberapa responden yang melepasi had atas skala (melebihi +8 logit), yang mencerminkan kumpulan individu dengan kebolehan yang sangat tinggi berbanding responden lain. Sementara itu, terdapat responden yang berada pada bahagian bawah skala (sekitar -5 logit), menunjukkan mereka mempunyai kebolehan yang paling rendah dalam kumpulan kajian.

Di sebelah kanan peta, item yang berada di bahagian atas seperti D10 dilihat sebagai item paling sukar untuk dipersetujui oleh responden. Sebaliknya, item B01 terletak di bahagian paling bawah skala, menandakan ia adalah item yang paling mudah dipersetujui. Taburan ini membuktikan wujud keseimbangan antara aras kesukaran item dengan tahap keupayaan responden, sekali gus mengesahkan kesesuaian instrumen i-PGLAM untuk mengukur konstruk yang dikaji.

RAJAH 1. Peta Item-Responden



KESIMPULAN

Kajian ini secara keseluruhan memberikan bukti bahawa instrumen i-PGLAM yang dibangunkan untuk mengukur tahap pengetahuan guru LAM di SABK menepati piawaian kesahan dan kebolehpercayaan melalui analisis menggunakan model pengukuran Rasch. Di samping itu, kajian ini juga mengesahkan bahawa item dan responden terletak pada tahap ketekalan yang tinggi, dengan nilai kebolehpercayaan meyakinkan serta kesahan konstruk yang mantap. Ini memberikan keyakinan bahawa instrumen ini sesuai dijadikan alat ukur bagi menilai sejauh mana pengetahuan guru terhadap isi kandungan, kefahaman kurikulum, dan penilaian dalam konteks KBD. Instrumen ini bukan sekadar memenuhi keperluan akademik, malah menjadi sumbangan praktikal dalam usaha memperkukuh kompetensi guru bahasa Arab di sekolah agama. Lebih penting lagi, kajian ini telah mengisi jurang penyelidikan yang sebelum ini kurang diberi perhatian, iaitu penilaian khusus terhadap tahap pengetahuan guru bahasa Arab melalui instrumen yang sistematik dan teruji. Dengan adanya dapatan ini, diharapkan pihak berkepentingan seperti pembuat dasar, penyelidik, dan guru sendiri boleh menggunakan instrumen ini untuk menilai dan menambah baik kualiti PdP bahasa Arab di SABK. Pada akhirnya, pengukuhan pengetahuan guru akan memberi kesan langsung kepada keberkesanan kurikulum, pencapaian pelajar, dan matlamat besar pendidikan Islam di Malaysia.

PENGHARGAAN

Penghargaan kepada Fakulti Pengajian Islam, Universiti Kebangsaan Malaysia di bawah program Tabung Agihan Pelajar.

RUJUKAN

- Abdul Razif Zaini, Noor Shamshinar Zakaria, Mohd Rofian Ismail, Mohd Redzaudhin Ghazali & Hasmadi Hamdan. (2020). Kompetensi Guru Bahasa Arab Sekolah Rendah di Malaysia Competency of Arabic Teachers in Primary School in Malaysia. *Jurnal Sultan Alauddin Sulaiman Shah* 7(1): 103–113
- Abdul Razif Zaini, Noorshamsinar Zakaria, Hasmadi Hamdan, Muhammad Redzaudin & Mohd Rofian Ismail. (2019). Pengajaran Bahasa Arab di Malaysia : Permasalahan dan Cabaran: 47–57.
- Abdull Sukor Shaari. (2022). *Kaedah Penyelidikan : Pendidikan Kuantitatif*.
- Ahmad Sabri. (2017). Ciri-Ciri Linguistik Dalam Kebolehbacaan Buku Teks Berbahasa Arab Dalam Kurikulum Dini. Tesis Doktor Falsafah, Universiti Kebangsaan Malaysia, Bangi.
- Awatif Abdul Rahman, Mohd Shafie Zulkifli, Muhammad Hashimee, Ainon Wazir, Muafah Sayed. (2020). Kompetensi guru bahasa Arab di Malaysia: Satu implikasi. *Persidangan Antarabangsa Sains Sosial dan Kemanusiaan ke-5 (PASAK5)*, hlm. 1119–1131.
- Azizi Yahya, Peter Voo, Ismail Maakip, M.D.A. Malek. (2017). *Kaedah Penyelidikan dalam Pendidikan*. Tanjung Malim. Universiti Pendidikan Sultan Idris (UPSI Press).
- Azrilah Abdul Aziz, Mohd Saidfudin Masodi, A. Zaharim. (2013). *Asas Model Pengukuran Rasch: Pembentukan Skala & Struktur Pengukuran*. Penerbit Universiti Kebangsaan Malaysia.: Bangi.
- Bambang Sumintono, Wahyu Widhiarso. (2014). *Aplikasi model rasch untuk penelitian ilmu-ilmu sosial (Rev. ed.)*. Trim Komunikata Publishing House, Cimahi.
- Bond, T. G. & Fox, C.M. (2015). *Applying the Rasch Model Fundamental Measurement in the Human Sciences*. London: Routledge.
- Bond, T. G., Yan, Z. & Heene, M. (2021). *Applying The Rasch Model Fundamental Measurement in the Human Sciences*. London: Routledge.
- Daud Lintang Al Yamin. (2023). Bahasa Arab Sebagai Identitas Budaya Islam dan Pemersatu Keberagaman Suku. *Ta'limi | Journal of Arabic Education and Arabic Studies* 2(1): 73–86.
- Fariyah Hussain, Nik Mohd Rahimi Nik Yusoff, Maimun Aqsha Lubis Abdin Lubis & Mohamad Zamri Jusoh. (2020). Amalan bahan bantu mengajar pengajaran guru dalam pengajaran Nahu Kurikulum Bersepadu Dini Sekolah Agama. *Jurnal 'Ulwan* 5(1): 134–145.
- Fauth, B., Decristan, J., Decker, A.T., Büttner, G., Hardy, I., Klieme, E. & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education* 86.
- Firdaus Abdul Fatah, Khadijah Abdul Razak & Maimun Aqsha Lubis. (2019). Syariah Subject Using Integrated Dini Curriculum: A Renewal in Government Aided Educational Religious School. *International Journal of Academic Research in Business and Social Sciences* 9(2): 1068–1077.
- Fisher, J.W.P. (2007). *Rating Scale Instrument Quality Criteria. Rasch Measurement Transactions*. 21(1).
- Halim Tamuri, A. & Hussin, N.H. (2017). Cabaran Pendidikan Islam di Era Media Sosial. *Malaysian Online Journal of Education*, hlm.
- Hasnurol Hashim, Kaseh Abu Bakar & Maheram Ahmad. (2022). Kesahan Pengukuran dan Kebolehppercayaan Instrumen Ujian Struktur Dalaman Kata imbuhan (USDAi) bagi Penguasaan Morfologi Arab (June).

- Hazrul Affendi Mohmad Razali. (2020). Elemen pengajaran berkesan bahasa Arab. *Journal of Asian Islamic Higher Institutions (JAIHI)journal* 5.
- John M.Linacre. (2002). Understanding Rasch Measurement: Optimizing Rating Scale Category Effectiveness. *Journal of Applied Measurement*, v3 n1 p85-106 2002.
- Kementerian Pendidikan Malaysia. (2023). *Dokumen standard kurikulum dan pentaksiran (Al-Lughah al-'Arabiah al-Mu'asirah Tingkatan 5)*.
- Khairulaizam Abdul Rahman & Harun Baharudin. (2022). Persekitaran Bahasa dalam Pembelajaran Kosa Kata Arab Pelajar. *Proceeding of the 9th International Conference on Management and Muamalah 2022 (ICoMM 2022)*, hlm. 195–208.
- Koehler, M.J., Mishra, P. & Cain, W. (2009). What is Technological Pedagogical Content Knowledge (TPACK)? *Journal of Education* 193(3): 13–19.
- Lawshe, C.H. (1975). *A Quantitative Approach to Content Validity*. *Personnel Psychology*. *Personnel Psychology* 28(4): 563–575.
- Lena Hollenstein, C.B. (2024). The importance of teachers' pedagogical-psychological teaching knowledge for successful teaching and learning. *Journal of Curriculum Studies* 56(4): 480–495.
- Lynn, M.R. (1986). *Determination and quantification of content validity*. *Nursing Research* 35(6): 382–385.
- Mohd Effendi@Ewan Mohd Matore, Hisyamsani Idris, Normawati Abdul Rahman & Ahmad Zamri Khairani. (2017). Kesahan Kandungan Pakar Instrumen IKBAR Bagi Pengukuran AQ Menggunakan Nisbah Kesahan Kandungan. *Proseeding of International Conference On Global Education V (ICGE V), Padang Indonesia [10-11 April 2017](April)*: 979–997.
- Mohd Fadzli Ismail & Nik Mohd Rahimi NikYusoff. (2021). The Relationship between Linguistic Features and Readability of Dini Integrated Curriculum Arabic Textbook in Government-Aided Religious School. *Islamiyyat* 43(2): 117–129.
- Mohd Najib Abdul Ghafar. (2003). *Reka bentuk tinjauan soal selidik pendidikan*. Universiti Teknologi Malaysia: Skudai, Johor Bahru.
- Mohd Zaffi Alias. (2021). Pengetahuan Pedagogi Isi Kandungan dan Amalan Pengajaran Guru Bahasa Arab di Sekolah Agama Bantuan Kerajaan Dalam Kurikulum Bersepadu Dini. Tesis Doktor Falsafah, Universiti Kebangsaan Malaysia, Bangi, Selangor.
- Muhammad Bachtiar Syamsuddin. (2022). The Future Of The Arabic Language. *Ijazarabi Journal of Arabic Learning* 5(3): 913–921.
- Muhammad Sabri Sahrir & Ilyani Syiham Muhammad. (2019). Tinjauan Respons Pelajar Terhadap Program Akademik Sarjana Muda Pendidikan Pengajaran Bahasa Arab Sebagai Bahasa Kedua di Universiti Islam Antarabangsa Malaysia (UIAM). *Jurnal Sultan Alauddin Sulaiman Shah* 6: 248–262.
- Nathan Burroughs, Jacqueline Gardner, Y.L. (2019). A Review of the Literature on Teacher. *Teaching for Excellence and Equity*(2007): 7–17.
- Norkumalasari Othman, Nor Hasnida Che Md Ghazali, Mohd Nazir Md Zabir & Nazaruddin Abdul Hadi. (2020). Penerokaan dan Pengesahan Instrumen Sikap Terhadap Matematik dalam kalangan Guru Matematik Sekolah Menengah. *Jurnal Kurikulum & Pengajaran Asia Pasifik* 8(4): 1–11.
- Nunnally, J. C. & Bernstein, I.H. (1994). The Assessment of Reliability. *Psychometric Theory*. 3: 248–292.
- Nur Afiqah Athirah Mohd Rushdi & Lily Hanefarezan Asbullah. (2022). Validity and Reliability of the Student Attitudes towards Arabic Language Learning Instrument (I-SPPBA) using the Rasch Measurement Model. *International Journal of West Asian Studies* 14(December): 98–110.
- Nurul Hamidzah Mohd Zafari & Mohd Isa Hamzah. (2023). Tahap kesediaan guru bukan pengkhususan bahasa Arab dalam P&P subjek bahasa Arab di Sekolah Rendah Daerah Kota Kinabalu, Sabah. *Jurnal Dunia Pendidikan* 5(3).

- Rafidah Mohd Adnan & Mohd Effendi Ewan Mohd Matore. (2019). Kesahan dan Kebolehpercayaan Instrumen I-CGPKM Menggunakan Model Rasch. *Jurnal Pengukuran Kualiti dan Analisis* 15(1): 1–14.
- Tawa Ahmed, A. & Olayinka Shogbesan, Y. (2023). Exploring Pedagogical Content Knowledge of Teachers: a Paradigm For Measuring Teacher's Effectiveness. *Pedagogi: Jurnal Ilmu Pendidikan* 23(1): 64–73.
- Tom Leijnse. (2025). The Arabic language, nationalism, and nation-building in the Mashriq and the Maghrib: the case of Iraq and Morocco. *Taylor & Francis* 27(4): 393–412.
- William J. Boone, John R. Staver & Melissa S. Yale. (2014). *Rasch Analysis in the Human Sciences*. Dordrecht: Springer.