



INTERNATIONAL JOURNAL OF
MODERN EDUCATION
(IJMOE)
www.ijmoe.com



**PEMBINAAN DAN PENGESAHAN INSTRUMEN KOMPETENSI
GURU DALAM PELAKSANAAN PENTAKSIRAN BILIK
DARJAH BAGI MATA PELAJARAN BAHASA MELAYU**

*DEVELOPMENT AND VALIDATION OF A TEACHER COMPETENCY
INSTRUMENT IN THE IMPLEMENTATION OF CLASSROOM-BASED
ASSESSMENT FOR THE MALAY LANGUAGE SUBJECT*

Muhammad Syafiq Majid¹, Mohd Nazri Abdul Rahman^{2*}

¹ Faculty of Education Universiti Malaya, Malaysia
Email: 22097226@siswa.um.edu.my

² Faculty of Education Universiti Malaya, Malaysia
Email: mohdnazri_ar@um.edu.my

* Corresponding Author

Article Info:

Article history:

Received date: 10.08.2025

Revised date: 24.08.2025

Accepted date: 09.09.2025

Published date: 15.10.2025

To cite this document:

Majid, M. S., & Abdul Rahman, M. N.
(2025). Pembinaan dan Pengesahan
Instrumen Kompetensi Guru dalam
Pelaksanaan Pentaksiran Bilik Darjah
bagi Mata Pelajaran Bahasa Melayu.
*International Journal of Modern
Education*, 7 (27), 537-557.

DOI: 10.35631/IJMOE.727033

This work is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)



Abstrak:

Kajian ini bertujuan membina dan mengesahkan instrumen bagi mengukur tahap kompetensi guru dalam pelaksanaan Pentaksiran Bilik Darjah (PBD) bagi mata pelajaran Bahasa Melayu sekolah rendah di Wilayah Persekutuan Kuala Lumpur. Walaupun PBD telah menjadi dasar penting dalam sistem pendidikan, kajian lepas tentang pelaksanaannya dalam konteks mata pelajaran Bahasa Melayu masih terhad. Selain itu, kajian tentang kompetensi guru dalam melaksanakan PBD juga belum banyak diterokai secara menyeluruh. Kajian ini menggunakan pendekatan kuantitatif berasaskan model pembinaan instrumen Wilson (2004) yang dikenali sebagai *four building blocks*. Instrumen awal telah dinilai oleh lima orang pakar bagi memastikan kesahan kandungan, manakala Model Pengukuran Rasch pelbagai faset digunakan untuk mendapatkan bukti empirikal. Kajian rintis dijalankan melibatkan 120 guru, manakala kajian lapangan melibatkan 172 guru. Hasil analisis menunjukkan instrumen akhir mengandungi 34 item yang memenuhi ciri unidimensi dan mempunyai kebolehpercayaan yang tinggi. Instrumen yang dinamakan sebagai IKGPBD BM ini mempunyai implikasi praktikal yang penting. Instrumen ini boleh digunakan oleh guru sebagai alat penilaian sendiri untuk mengenal pasti kekuatan dan kelemahan kompetensi mereka dalam pelaksanaan PBD bagi mata pelajaran Bahasa Melayu. Pentadbir sekolah serta pegawai pendidikan di peringkat daerah dan negeri pula dapat memanfaatkannya untuk merancang program pembangunan profesional yang

lebih bersasar. Selain itu, dapatan kajian ini juga boleh dijadikan rujukan oleh institusi latihan guru dan Kementerian Pendidikan Malaysia dalam usaha memperkukuh dasar dan latihan berkaitan pelaksanaan PBD. Secara keseluruhannya, kajian ini menyumbang kepada pengukuhan amalan pentaksiran bilik darjah dalam konteks mata pelajaran Bahasa Melayu serta menyediakan asas yang kukuh untuk penambahbaikan latihan guru dan dasar pendidikan negara.

Kata Kunci:

Kompetensi Guru, Pentaksiran Bilik Darjah, Bahasa Melayu, Model Rasch, Pembinaan Instrumen

Abstract:

This This study aims to develop and validate an instrument to measure teachers' competency in implementing Classroom-Based Assessment (PBD) for the Malay Language subject in primary schools within the Federal Territory of Kuala Lumpur. Although PBD has become an important policy in the education system, previous studies in the context of the Malay Language remain limited. Moreover, teachers' competency in implementing PBD has not been extensively explored. The study employed a quantitative approach based on Wilson's (2004) instrument development model, known as the four building blocks. The initial instrument was evaluated by five experts to ensure content validity, while the Multifaceted Rasch Measurement Model was applied to obtain empirical evidence. A pilot study involving 120 teachers was conducted, followed by a field study with 172 teachers. The findings showed that the final instrument consists of 34 items, which demonstrated unidimensionality and high reliability. The instrument, named IKGPBD BM, carries significant practical implications. It can serve as a self-assessment tool for teachers to identify their strengths and weaknesses in implementing PBD for the Malay Language subject. School administrators and education officers at the district and state levels may also use it to design more targeted professional development programs. In addition, the findings of this study can inform teacher training institutions and the Ministry of Education in strengthening policies and training related to PBD implementation. Overall, this study contributes to the enhancement of classroom-based assessment practices in the Malay Language subject and provides a solid foundation for improving teacher training and education policy.

Keywords:

Teacher Competency, Classroom-Based Assessment, Malay Language, Rasch Model, Instrument Development

Pendahuluan

Pendidikan di Malaysia terus melalui proses transformasi yang berterusan untuk memastikan keberkesanan sistem pengajaran dan pembelajaran seiring dengan tuntutan global. Salah satu aspek penting dalam reformasi ini ialah pembaharuan sistem pentaksiran, yang telah dilaksanakan melalui Pelan Pembangunan Pendidikan Malaysia (PPPM) 2013–2025. PPPM menekankan kepentingan pembangunan modal insan secara menyeluruh berdasarkan Falsafah Pendidikan Kebangsaan, termasuk dalam aspek pentaksiran pembelajaran yang lebih autentik dan berpusatkan murid.

Sebagai salah satu komponen utama dalam pendekatan Pentaksiran Berasaskan Sekolah (PBS), Pentaksiran Bilik Darjah (PBD) telah diperkenalkan secara meluas bermula tahun 2019 untuk murid tahap satu sekolah rendah. PBD merupakan bentuk pentaksiran formatif yang dilaksanakan secara berterusan bagi menilai keupayaan murid dalam pelbagai aspek seperti kognitif, afektif dan psikomotor melalui kaedah pemerhatian, tugas, projek serta maklum balas guru secara langsung. Langkah ini sejajar dengan amalan pendidikan negara maju seperti Finland, Australia dan Sweden yang telah lama menekankan pentaksiran holistik dalam bilik darjah (Darling-Hammond & McCloskey, 2008; Mercurio, 2008).

Dalam konteks mata pelajaran Bahasa Melayu yang menjadi teras kepada sistem pendidikan kebangsaan, pelaksanaan PBD memerlukan pendekatan pentaksiran yang menyeluruh merangkumi kemahiran mendengar, bertutur, membaca dan menulis. Keupayaan guru melaksanakan PBD secara berkesan menjadi tunjang kepada kejayaan dasar ini. Justeru, penilaian terhadap kompetensi guru dalam pelaksanaan PBD adalah kritikal dan perlu dilakukan secara sistematik serta berasaskan instrumen yang sah dan boleh dipercayai.

Penyataan Masalah

Walaupun PBD telah dilaksanakan secara rasmi dalam sistem pendidikan negara, pelaksanaannya masih menghadapi pelbagai cabaran. Antara isu utama ialah kelemahan guru dalam menguasai konsep dan prinsip PBD seperti pentaksiran formatif, autentik, serta penggunaan standard prestasi (KPM, 2021; Jamil & Said, 2019; Villasamy & Mohammad, 2023). Kekurangan pengetahuan ini menyebabkan penskoran dibuat berdasarkan interpretasi peribadi guru, tanpa rujukan kepada standard yang ditetapkan.

Selain itu, tahap kemahiran guru dalam merancang dan melaksanakan pentaksiran turut menjadi isu kritikal. Kajian terdahulu menunjukkan guru cenderung untuk memilih kaedah pentaksiran yang tidak selari dengan kemahiran yang ditaksir, seperti menilai kemahiran lisan melalui aktiviti bertulis (Annaliza et al., 2020; Arumugham, 2020; Said & Jamil, 2021). Walaupun terdapat guru yang memahami keperluan kurikulum, mereka masih menghadapi kesukaran dalam prosedur penskoran dan pengurusan data pentaksiran (Nor S. M. & Haron, 2023).

Dari segi sikap pula, kajian melaporkan bahawa masih terdapat guru yang menunjukkan tahap keyakinan dan profesionalisme yang sederhana dalam membuat pertimbangan pentaksiran (Dorin & Yasin, 2019; Faudzi & Ahmad, 2022). Terdapat percanggahan dalam dapatan beberapa kajian terdahulu, di mana ada kajian yang menunjukkan sikap guru terhadap pelaksanaan PBD adalah positif dan proaktif (Sh. Siti Hauzimah, 2019).

Di samping itu, instrumen yang digunakan dalam kajian lepas didapati terlalu umum dan tidak menilai secara menyeluruh ketiga-tiga aspek utama kompetensi guru iaitu pengetahuan, kemahiran dan sikap dalam konteks pelaksanaan PBD bagi mata pelajaran Bahasa Melayu (Norlizam et al., 2023; Khairuddin & Yaakub, 2022). Kebanyakan instrumen juga tidak dibina berdasarkan konteks sebenar bilik darjah dan kurang menggunakan pendekatan pengukuran yang sah seperti Model Rasch.

Justeru, kajian ini mengambil inisiatif untuk membina dan mengesahkan satu instrumen yang mengukur kompetensi guru Bahasa Melayu dalam pelaksanaan PBD secara menyeluruh dengan menggunakan pendekatan Model Rasch. Instrumen ini diharapkan dapat memberikan

data empirikal yang sah dan boleh dipercayai untuk tujuan penilaian dan intervensi profesional guru.

Objektif Kajian

Kajian ini dijalankan untuk membina serta menilai kualiti item-item yang terkandung dalam Instrumen Kompetensi Guru dalam Pelaksanaan Pentaksiran Bilik Darjah bagi mata pelajaran Bahasa Melayu (IKGPBD BM) dengan menggunakan analisis permodelan Rasch. Hasil analisis ini dapat mengesahkan kesahan dan kebolehpercayaan instrumen serta menilai item-item yang digunakan dalam instrumen mempunyai statistik yang sepadan dengan Model Rasch. Bagi mencapai tujuan kajian, berikut merupakan objektif khusus kajian ini:

1. Menenal pasti item-item bagi kompetensi guru dalam pelaksanaan pentaksiran bilik darjah bagi mata pelajaran Bahasa Melayu mempunyai bukti kesahan kandungan.
2. Menguji dan mengesahkan kesahan konstruk dan kebolehpercayaan instrumen dengan menggunakan model Rasch.
3. Menilai kualiti item-item dalam instrumen berdasarkan standard pengukuran Model Rasch.
4. Menganalisis perbezaan signifikan tahap kompetensi guru berdasarkan min logit mengikut tempoh pengalaman mengajar.

Sorotan Kajian

Sorotan kajian ini disusun secara tematik merangkumi empat bidang utama yang berkait rapat dengan fokus penyelidikan, iaitu pentaksiran bilik darjah, kompetensi, pembinaan instrumen, dan model pengukuran Rasch. Huraian bagi setiap tema akan membincangkan dapatan penyelidikan terdahulu serta mengenal pasti jurang yang wujud dalam kajian sedia ada.

Pentaksiran Bilik Darjah

Pentaksiran bilik darjah (PBD) memainkan peranan penting dalam pelaksanaan Transformasi Pendidikan Nasional di Malaysia. Pelaksanaan ini mengubah pendekatan penilaian yang sebelum ini terlalu bergantung kepada peperiksaan berpusat, kepada sistem yang lebih holistik dengan menggabungkan pentaksiran berpusat dan pentaksiran formatif secara berterusan dalam bilik darjah (Kementerian Pendidikan Malaysia, 2016). Pendekatan ini bertujuan mengurangkan tekanan peperiksaan, meningkatkan minat belajar, dan memastikan pembangunan potensi murid secara menyeluruh (Mokhtar, 2016; Veloo et al., 2015).

Di Malaysia, PBD telah diimplementasi dari peringkat prasekolah sehingga ke sekolah menengah dan memberikan autonomi kepada guru untuk menilai prestasi murid berdasarkan standard yang ditetapkan (Mohd Isa et al., 2020). Guru berperanan dalam merancang, membina, merekod dan mentafsir tugas pentaksiran yang berbentuk formatif dan sumatif. Pelaporan pencapaian murid berasaskan tahap penguasaan membolehkan guru merancang intervensi pengajaran yang lebih berkesan (Ibrahim & Reduan, 2002; Black & Wiliam, 2012).

Penekanan kepada pentaksiran formatif di dalam bilik darjah mendapat sokongan daripada kajian antarabangsa yang terkini. Sebagai contoh, El Zerk (2025) menekankan keperluan guru menggunakan pelbagai teknik pentaksiran bilik darjah untuk mengembangkan kemahiran komunikasi dan literasi dalam kalangan pelajar EFL (*English as a Foreign Language*) di Maghribi. Selain itu, Han et al. (2025) pula menunjukkan bagaimana model bahasa generatif

boleh membantu dalam pentaksiran berasaskan bilik darjah dengan ketepatan tinggi dan keberkesanan dalam pentaksiran.

Pelaksanaan PBD yang efektif memerlukan guru memiliki kecekapan profesional tinggi. Ini termasuk keupayaan untuk mereka bentuk instrumen pentaksiran autentik, mentafsir bukti pembelajaran dengan teliti, serta menyampaikan maklum balas yang dapat membina kefahaman dan motivasi murid (Yates & Johnston, 2018). KPM (2018) turut menegaskan keperluan guru untuk menguasai ilmu pentaksiran bagi menjayakan hasrat pembelajaran berpusatkan murid.

Kompetensi

Konsep kompetensi telah menjadi tumpuan dalam banyak kajian berkaitan pembangunan modal insan dan prestasi kerja. Kompetensi sering ditakrifkan sebagai satu set pengetahuan, kemahiran, dan sifat peribadi yang membolehkan seseorang individu melaksanakan tugas atau tanggungjawab secara berkesan (*United Nations Industrial Development Organization*, 2002; Sutton, 2004). Ia merangkumi bukan sahaja kebolehan teknikal, tetapi juga nilai, sikap dan ciri psikologi dalaman yang membentuk tingkah laku kerja seseorang (Sampson & Fytros, 2008).

Dalam konteks pendidikan, kompetensi guru merupakan faktor utama yang mempengaruhi kualiti pengajaran dan pembelajaran. Spencer dan Spencer (1993) melalui *Iceberg Competency Model* memperkenalkan kerangka yang relevan dalam memahami struktur kompetensi individu. Menurut model ini, kemahiran dan pengetahuan berada di bahagian permukaan (atas gunung ais) dan mudah diperhatikan, manakala sifat peribadi seperti motivasi, nilai, dan ciri keperibadian berada di bawah permukaan dan lebih sukar dikenalpasti tetapi sangat berpengaruh terhadap prestasi sebenar individu.

Vargas et al. (2023) dalam satu kajian berkaitan latihan guru di Amerika Latin menekankan perlunya pendekatan pembangunan kompetensi yang holistik melibatkan pembelajaran berterusan, pembangunan kemahiran sosial, dan peningkatan kompetensi pedagogi. Sementara itu, Yong dan Tuan (2022) menunjukkan bahawa latihan berasaskan model kompetensi yang menekankan aspek refleksi sendiri dan kesedaran nilai diri mampu memberi impak positif terhadap pencapaian profesional guru dalam jangka panjang.

Penilaian terhadap kompetensi juga memerlukan instrumen yang mampu menilai elemen tersirat seperti sikap dan nilai. Dalam kajian oleh Jain dan Narayanan (2021), pembangunan rubrik dan sistem pentaksiran kompetensi yang lebih menyeluruh dapat mengenalpasti kekuatan dan jurang dalam prestasi guru secara lebih adil dan menyeluruh.

Pembinaan Instrumen Pengukuran

Dalam bidang pendidikan, pembinaan instrumen pengukuran memainkan peranan yang penting untuk menilai ciri-ciri yang tidak dapat dilihat secara langsung seperti pengetahuan, sikap, atau kompetensi. Berbeza dengan pengukuran kuantiti fizikal yang bersifat nyata dan terukur secara langsung, pengukuran kuantiti bukan fizikal memerlukan pendekatan yang lebih kompleks kerana ia melibatkan konstruk yang bersifat abstrak (Wilson, 2004).

Secara umum, pengukuran kuantiti bukan fizikal boleh dibahagikan kepada dua kategori iaitu pengukuran kognitif (seperti pemahaman dan pemikiran logik) dan pengukuran bukan kognitif (seperti sikap, minat dan nilai). Instrumen yang digunakan untuk pengukuran tersebut perlu

direka bentuk secara sistematik agar dapat menghasilkan data yang sahih, boleh dipercayai dan konsisten.

Mark Wilson (2004) telah memperkenalkan pendekatan Four Building Blocks yang melibatkan empat komponen utama dalam proses pembangunan instrumen, iaitu:

1. Peta konstruk (Construct Map) – Penerangan jelas tentang ciri yang ingin diukur dan bagaimana ia berkembang dari tahap rendah ke tinggi.
2. Reka bentuk item (Item Design) – Pembangunan soalan atau tugas yang boleh menimbulkan bukti langsung tentang tahap penguasaan konstruk.
3. Ruang keluaran (Outcome Space) – Kategori respons atau jawapan yang menunjukkan tahap penguasaan responden.
4. Model pengukuran (Measurement Model) – Model statistik seperti Rasch yang digunakan untuk menganalisis jawapan dan menempatkan responden pada skala konstruk.

Sehubungan itu, pembinaan instrumen bukan hanya soal membina soalan semata-mata, tetapi memerlukan pemahaman konstruk dan kebolehan mentafsir respons secara objektif. Pendekatan sistematik Wilson dapat meningkatkan keberkesanan pentaksiran serta memberi maklumat yang lebih tepat kepada guru dan pembuat dasar

Model Pengukuran Rasch

Model Rasch merupakan satu pendekatan dalam Teori Respons Item (IRT) yang memainkan peranan penting dalam analisis dan pembangunan instrumen dalam bidang pendidikan. Ia dikenali juga sebagai model *1-Parameter Logistic (1PL)* kerana hanya mempertimbangkan satu parameter sahaja bagi setiap item, iaitu kesukaran item (Bond & Fox, 2015; Boone, 2016). Melalui transformasi logit, Model Rasch membolehkan data ordinal seperti skala Likert ditukar kepada data sela (interval), menjadikannya lebih sesuai untuk analisis statistik lanjutan dan pengukuran saintifik (Linacre, 2010).

Model Rasch menawarkan kelebihan signifikan berbanding pendekatan klasik seperti Teori Ujian Klasik (CTT). Model ini membolehkan penyelidik menyusun responden mengikut tahap keupayaan serta item mengikut tahap kesukaran, melalui peta item-responden (Wright Map). Kaedah ini memberikan visualisasi yang membantu dalam mengenal pasti ketidaksepadanan antara responden dan item, serta menilai keberkesanan dan keadilan instrumen (Azrilah Abdul Aziz et al., 2015).

Menurut Siti Eshah Mokhsein (2018), model ini sangat berguna dalam mengukur kesahan konstruk kerana item yang tidak sepadan dengan model ini dianggap sebagai indikator multidimensi iaitu satu ciri yang boleh merendahkan kesahan instrumen. Dalam konteks ini, pengesahan menggunakan Model Rasch membolehkan pengkaji memperkemarkan konstruk dan memastikan hanya item berkualiti tinggi dikekalkan.

Kajian antarabangsa terkini terus menyokong kelebihan penggunaan Model Rasch. Sebagai contoh, Li et al. (2025) dalam kajian mereka terhadap skala literasi fizikal untuk pelajar sekolah rendah di China, menggunakan Rasch untuk membuktikan kesahan konstruk dan menyusun kesukaran item dengan tepat. Kajian mereka menunjukkan bahawa peta item-responden

membantu memadankan tahap pelajar dengan kerumitan item, sekali gus meningkatkan keberkesanan pentaksiran.

Berdasarkan kelebihan ini, Model Rasch terus menjadi pilihan utama dalam pengukuran psikometrik, terutamanya dalam bidang pendidikan yang memerlukan penilaian konstruk secara terperinci, tepat, dan objektif.

Jurang Kajian

Meskipun pelaksanaan Pentaksiran Bilik Darjah (PBD) telah banyak dibincangkan dalam pelbagai kajian terdahulu, kebanyakan tumpuan diberikan kepada mata pelajaran seperti Sains, Sejarah, dan Matematik. Walaupun Bahasa Melayu merupakan bahasa kebangsaan dan elemen teras dalam kurikulum, kajian yang memfokuskan kepada pelaksanaan PBD dalam mata pelajaran Bahasa Melayu masih lagi terhad. Kekurangan ini mewujudkan keperluan untuk meneliti secara khusus keberkesanan pelaksanaan PBD dalam mata pelajaran Bahasa Melayu bagi memastikan pendekatan pentaksiran yang digunakan menyokong penguasaan bahasa dalam kalangan pelajar.

Selain itu, kajian terdahulu juga menunjukkan bahawa kompetensi guru dalam melaksanakan PBD masih belum diterokai secara menyeluruh, khususnya dalam konteks Bahasa Melayu. Kebanyakan instrumen yang digunakan hanya menilai aspek tertentu seperti pengetahuan atau sikap, tanpa merangkumi keseluruhan dimensi kompetensi secara holistik. Tambahan pula, kaedah analisis yang digunakan selalunya terbatas kepada statistik asas, sedangkan pendekatan yang lebih tepat seperti Model Pengukuran Rasch dapat memberikan penilaian yang lebih mendalam dan sah. Oleh itu, terdapat keperluan jelas untuk membangunkan instrumen yang komprehensif serta menggunakan model analisis yang lebih mantap untuk menilai kompetensi guru dalam melaksanakan PBD.

Metodologi Kajian

Kajian ini menggunakan pendekatan kuantitatif dengan reka bentuk pembangunan instrumen berdasarkan model Four Building Blocks yang diperkenalkan oleh Wilson (2004). Reka bentuk kajian dilaksanakan secara berperingkat dalam tiga fasa utama, bermula dengan pembinaan dan pembangunan instrumen, diikuti fasa pengujian instrumen, dan akhirnya fasa pengesahan instrumen melalui analisis data menggunakan Model Pengukuran Rasch.

Fasa Pembinaan Item

Fasa pertama melibatkan proses pembangunan item instrumen yang merangkumi penentuan konstruk, pembentukan domain dan subdomain kompetensi, serta penggubalan item-item berdasarkan sorotan literatur dan kajian lepas. Item yang dibina merangkumi tiga konstruk utama iaitu pengetahuan, kemahiran dan sikap guru dalam melaksanakan Pentaksiran Bilik Darjah (PBD). Item-item yang dibangunkan dinilai dari aspek kesahan kandungan oleh lima orang panel pakar dalam bidang pentaksiran dan pendidikan Bahasa Melayu. Dapatan tersebut dianalisis dengan menggunakan perisian Minifac.

Fasa Pengujian Instrumen

Fasa kedua ialah pengujian awal instrumen melalui kajian rintis yang melibatkan 120 orang guru Bahasa Melayu sekolah rendah kebangsaan di Wilayah Persekutuan Kuala Lumpur. Pemilihan sampel dilakukan menggunakan kaedah persampelan rawak kluster berperingkat, dengan syarat guru yang dipilih mengajar mata pelajaran Bahasa Melayu dan terlibat secara

langsung dalam pelaksanaan PBD. Dapatan kajian rintis ini dianalisis menggunakan Model Rasch pelbagai faset dengan bantuan perisian Winsteps bagi mengenal pasti item yang sesuai serta menilai kesahan dan kebolehpercayaan instrumen dari aspek teknikal.

Fasa Pengesahan Instrumen

Fasa ketiga merupakan fasa pengesahan konstruk instrumen. Instrumen akhir yang terdiri daripada 34 item ditadbir kepada 172 orang guru Bahasa Melayu sekolah rendah kebangsaan di Wilayah Persekutuan Kuala Lumpur. Sama seperti fasa sebelumnya, sampel dipilih berdasarkan kriteria kelayakan iaitu guru yang mengajar Bahasa Melayu dan terlibat secara langsung dalam pelaksanaan PBD. Pengesahan instrumen dilaksanakan melalui analisis kesahan konstruk, kesepadanan item, serta analisis kategori skala. Hasil analisis ini membolehkan pemetaan kompetensi guru secara lebih empirikal dan objektif.

Instrumen Kajian

Instrumen yang digunakan dalam kajian ini ialah soal selidik yang dibangunkan untuk menilai kompetensi guru dalam pelaksanaan PBD bagi mata pelajaran Bahasa Melayu. Instrumen ini merangkumi tiga konstruk utama: pengetahuan, kemahiran, dan sikap, dan menggunakan skala Likert empat mata untuk menilai tahap kompetensi guru. Pembinaan item instrumen disandarkan kepada teori kompetensi guru dan kerangka Standard Guru Malaysia (SGM 2.0), manakala analisis instrumen dilakukan menggunakan Model Rasch untuk memastikan kesahan konstruk dan kebolehpercayaan item.

Aspek Etika Kajian

Kajian ini telah dilaksanakan dengan mematuhi garis panduan etika penyelidikan yang ditetapkan. Kebenaran bertulis diperoleh daripada Kementerian Pendidikan Malaysia (KPM), Jabatan Pendidikan Negeri (JPN), serta pihak sekolah sebelum kajian dijalankan. Selain itu, setiap responden diberikan borang persetujuan penyertaan yang menerangkan tujuan kajian, kerahsiaan data, serta hak mereka untuk menarik diri pada bila-bila masa tanpa sebarang implikasi.

Bagi memastikan kerahsiaan maklumat terjamin, identiti responden dirahsiakan sepenuhnya dan data yang diperoleh hanya digunakan untuk tujuan akademik. Segala maklumat peribadi guru tidak didedahkan dalam laporan kajian, manakala analisis dan dapatan hanya dipersembahkan dalam bentuk agregat tanpa menjejaskan individu. Langkah ini diambil bagi memastikan integriti kajian serta melindungi hak dan kebajikan semua responden yang terlibat.

Penggunaan Model Rasch Pelbagai Faset

Penggunaan Model Rasch pelbagai faset dalam kajian ini adalah bertepatan dengan keperluan untuk menilai kesahan dan kebolehpercayaan instrumen secara lebih menyeluruh. Berbanding dengan analisis tradisional, Model Rasch pelbagai faset memberikan kelebihan kerana ia bukan sahaja menilai kesesuaian item, tetapi juga dapat mengawal pelbagai faktor lain yang mungkin mempengaruhi keputusan. Faktor lain yang dimaksudkan ialah tahap kesukaran item, kebolehan responden dan konsistensi penilaian pakar. Kaedah ini membolehkan analisis dilakukan secara objektif berdasarkan bukti empirikal dan memastikan instrumen yang dibina dapat mengukur konstruk yang dimaksudkan secara unidimensi (Bond & Fox, 2015).

Selain itu, Model Rasch pelbagai faset juga membolehkan penentuan kualiti item dengan lebih tepat melalui laporan statistik seperti nilai infit dan outfit *mean square*, indeks kebolehppercayaan item dan individu serta pemetaan logit item. Analisis ini penting untuk memastikan setiap item dalam instrumen menyumbang secara konsisten terhadap pengukuran kompetensi guru dalam pelaksanaan PBD. Kelebihan ini menjadikan model ini lebih sesuai berbanding kaedah tradisional seperti analisis faktor penerokaan (EFA) atau analisis faktor pengesahan (CFA) yang terhad dari segi penilaian kesepadanan item secara individu. Oleh itu, pemilihan Model Rasch pelbagai faset dalam kajian ini adalah wajar kerana ia dapat memberikan justifikasi statistik yang lebih kukuh terhadap kualiti instrumen dan sekaligus meningkatkan kesahan dan kebolehppercayaan hasil kajian.

Kriteria Pemilihan Pakar

Pemilihan pakar dalam kajian ini dibuat berdasarkan beberapa kriteria penting. Pertama, pakar yang dipilih mestilah mempunyai pengetahuan mendalam serta pengalaman dalam bidang yang berkaitan dengan kajian, khususnya Bahasa Melayu dan pentaksiran bilik darjah. Kepakaran ini diperlukan supaya mereka dapat menilai kesesuaian serta ketepatan kandungan instrumen yang dibangunkan. Kedua, pakar yang dipilih mempunyai pengalaman luas dalam bidang pengajaran dan penilaian. Pengalaman ini termasuk pengajaran Bahasa Melayu dan penglibatan dalam penyediaan atau pelaksanaan instrumen pentaksiran. Berdasarkan pengalaman tersebut, pakar dapat memberikan pandangan yang relevan dan membina terhadap kesahan instrumen. Ketiga, pakar juga perlu mempunyai latar belakang dalam pembinaan instrumen atau soal selidik. Pengetahuan dan pengalaman dalam proses pembangunan instrumen seperti pemilihan item, pengukuran, dan validasi membolehkan mereka menilai instrumen ini dengan lebih berkesan. Secara keseluruhannya, kriteria ini memastikan panel pakar yang dipilih benar-benar berkelayakan untuk menilai instrumen IKGPBD BM dari segi kesahan kandungan.

Penggunaan Skala Likert Empat Mata

Kajian ini menggunakan skala Likert empat mata sebagai skala pengukuran dalam instrumen soal selidik. Pemilihan skala empat mata dibuat bagi mengelakkan kewujudan pilihan neutral. Hal ini penting kerana responden dikehendaki memberikan jawapan yang jelas sama ada cenderung kepada positif atau negatif. Fisher (2006) menegaskan bahawa penggunaan skala neutral boleh menyebabkan responden mengambil jalan tengah tanpa memberikan pendirian sebenar.

Skala empat mata juga lebih sesuai dalam konteks kajian pendidikan kerana ia memerlukan responden membuat keputusan yang lebih tegas terhadap pernyataan yang diberikan. Dalam kajian ini, skala 1 mewakili sangat tidak setuju, skala 2 tidak setuju, skala 3 setuju, dan skala 4 sangat setuju. Bilangan skala yang digunakan adalah seimbang, tidak terlalu banyak dan tidak terlalu sedikit, sekali gus memudahkan responden untuk membuat pilihan. Oleh itu, skala empat mata dipilih kerana ia dianggap lebih sesuai untuk mendapatkan gambaran yang tepat mengenai tahap kompetensi guru dalam pelaksanaan PBD.

Dapatan Kajian

Kajian ini bertujuan untuk membangunkan dan mengesahkan instrumen bagi mengukur kompetensi guru dalam pelaksanaan Pentaksiran Bilik Darjah (PBD) bagi mata pelajaran Bahasa Melayu di sekolah rendah. Seramai 172 orang guru Bahasa Melayu dari sekolah rendah kebangsaan di Wilayah Persekutuan Kuala Lumpur telah terlibat sebagai responden dalam

kajian ini. Kesemua responden dipilih menggunakan kaedah persampelan rawak kluster berperingkat. Berdasarkan jumlah ini, majoriti responden terdiri daripada guru wanita, sejajar dengan trend kebangsaan yang menunjukkan peratusan guru wanita di sekolah rendah adalah lebih tinggi. Responden kajian ini mewakili pelbagai kumpulan umur dan tempoh pengalaman mengajar, sekaligus memberikan gambaran yang lebih menyeluruh tentang kompetensi guru dalam pelaksanaan Pentaksiran Bilik Darjah (PBD).

Ketidaksepadanan Item

Analisis awal menggunakan perisian Winsteps mendapati bahawa daripada 35 item yang dibangunkan untuk mengukur kompetensi guru dalam pelaksanaan PBD, hanya satu item dikenal pasti tidak menepati kriteria kesepadanan model Rasch berdasarkan nilai Mean Square Infit dan Outfit melebihi julat piawai yang ditetapkan, iaitu antara 0.5 hingga 1.5 (Bond & Fox, 2015). Item tersebut didapati memberikan tindak balas yang tidak konsisten dengan pola keseluruhan dan oleh itu telah digugurkan daripada analisis seterusnya.

Pengguguran item ini penting bagi memastikan kesahan konstruk dapat dikekalkan dan hanya item yang menunjukkan unidimensi dan kebolehpercayaan tinggi digunakan dalam penilaian akhir. Maka, instrumen versi akhir mengandungi 34 item yang memenuhi ciri psikometrik yang baik untuk pengukuran kompetensi guru berdasarkan tiga konstruk utama iaitu pengetahuan, kemahiran dan sikap.

Dapatan kajian dianalisis berdasarkan empat persoalan utama yang telah digariskan.

Persoalan Kajian 1

Apakah tahap kesepakatan pakar terhadap item-item instrumen yang dibangunkan bagi mengukur kompetensi guru dalam melaksanakan Pentaksiran Bilik Darjah (PBD) bagi mata pelajaran Bahasa Melayu?

Dapatan persoalan kajian pertama diperoleh melalui analisis menggunakan Model Rasch pelbagai faset dengan bantuan perisian *Minifac* bagi menilai ketegasan dan kesepakatan penilai terhadap item-item dalam instrumen IKGPBD BM. Lima orang penilai pakar terlibat dalam pentaksiran ini dan masing-masing menunjukkan tahap ketegasan dan gaya pemarkahan yang berbeza seperti yang ditunjukkan dalam Jadual 1.

Jadual 1: Laporan Tahap Ketegasan Penilai

Penilai	Nilai Logit	Infit MnSq	Outfit MnSq	Purata Skor	Interpretasi
B	1.73	1.01	0.89	3.61	Paling tegas
A	-1.10	0.96	1.06	3.84	Sederhana
E	-2.02	0.84	0.69	3.91	Longgar
C	-2.47	1.07	1.12	3.93	Lebih Longgar
D	-2.65	0.94	1.27	3.94	Paling Longgar

Penilai B dikenal pasti sebagai penilai paling tegas dengan nilai logit tertinggi (1.73), manakala penilai D paling longgar dengan nilai logit -2.65. Walaupun wujud perbezaan tahap ketegasan, nilai Infit MnSq dan Outfit MnSq bagi kesemua penilai masih berada dalam julat yang diterima (0.5 - 1.5), menandakan penilaian mereka adalah konsisten dan tidak menyimpang secara ekstrem daripada jangkaan model.

Ketekalan pemarkahan penilai disokong oleh analisis kebolehppercayaan yang menunjukkan indeks 0.95. Nilai ini melebihi ambang 0.90 yang disarankan oleh Bond dan Fox (2015) bagi tahap kebolehppercayaan yang sangat tinggi. Berdasarkan jadual 1.2 di bawah, nilai pengasingan penilai adalah 4.59, dengan strata sebanyak 6.46, menunjukkan bahawa wujud sekurang-kurangnya empat tahap ketegasan pemarkahan yang berbeza dalam kalangan penilai. Selain itu, keputusan ujian Chi-Square yang signifikan ($\chi^2 = 87.9$, $df = 4$, $p < 0.00$) mengesahkan bahawa perbezaan dalam pemarkahan bukan berlaku secara rawak tetapi konsisten dengan pola tertentu dalam kalangan penilai.

Jadual 1.2: Statistik Kebolehppercayaan Penilai

Parameter	Nilai
Indeks Kebolehppercayaan	0.95
Pengasingan Penilai	4.59
Strata	6.46
Chi-Square	87.9
Darjah Kebebasan (df)	4
Signifikan (p-value)	0.00
Kesepakatan Penilai (%)	79.3
Kesepakatan Jangkaan (%)	79.7

Nilai kesepakatan penilai yang diperoleh adalah 79.3%, hampir menyamai jangkaan model sebanyak 79.7%. Ini menunjukkan bahawa meskipun terdapat perbezaan tahap ketegasan antara penilai, secara keseluruhannya mereka masih mencapai tahap kesepakatan yang baik. Perbezaan ini memberikan gambaran realistik tentang gaya pemarkahan individu, tetapi masih terkawal dalam julat kesahan Model Rasch.

Selain daripada analisis terhadap tahap kesepakatan penilai, penilaian terhadap kesesuaian item dalam instrumen IKGPBD BM juga dijalankan menggunakan Model Rasch. Item-item dalam instrumen ini telah dikategorikan kepada empat tahap berdasarkan nilai logit pengukuran. Jadual 1.3 menunjukkan item yang memperoleh nilai logit melebihi 1.50 diklasifikasikan sebagai sangat baik, manakala item dengan nilai antara 0.00 hingga 1.50 dikategorikan sebagai baik. Item dengan logit antara -1.50 hingga 0.00 dikategorikan sebagai boleh diterima, manakala item dengan nilai logit kurang daripada -1.50 diklasifikasikan sebagai lemah.

Jadual 1.3: Ringkasan Pengukuran Item Berdasarkan Nilai Logit

Kategori	Julat Nilai Logit	Bilangan Item
Sangat Baik	> 1.50	17
Baik	0.00 – 1.50	4
Boleh Diterima	-1.50 – 0.00	8
Lemah	< -1.50	1

Secara keseluruhan, sebanyak 17 item berada dalam kategori sangat baik, mencerminkan keupayaan tinggi item-item ini dalam membezakan individu berdasarkan tahap kecekapan yang kompleks. Empat item dikategorikan sebagai baik dan masih menunjukkan sifat

diskriminatif yang kukuh. Lapan item pula diklasifikasikan sebagai boleh diterima, dengan tahap kesukaran yang sederhana, dan boleh menyumbang kepada keseimbangan dalam instrumen. Hanya satu item berada dalam kategori lemah, yang menunjukkan ia terlalu mudah untuk responden dan berpotensi memberi maklumat yang kurang signifikan dalam pembezaan keupayaan.

Dapatan ini menunjukkan bahawa penilai memberikan penilaian secara konsisten dan stabil terhadap item-item dalam instrumen, justeru menyokong kesahan kandungan instrumen dari perspektif pemarkahan pakar. Variasi yang wujud juga mencerminkan realiti pemarkahan yang pelbagai tanpa menjejaskan kebolehpercayaan keseluruhan instrumen.

Persoalan Kajian 2

Adakah instrumen yang telah dibina mempunyai eviden berkaitan kesahan konstruk dan kebolehpercayaan dengan menggunakan Model Rasch?

Bagi menjawab persoalan ini, empat analisis utama telah dijalankan menggunakan Model Pengukuran Rasch, iaitu; Analisis Komponen Prinsipal, Indeks Kebolehpercayaan Item dan Individu, Nilai Pengasingan Item dan Individu, serta Analisis Skala Penarafan. Hasil analisis ini memberikan bukti empirikal terhadap kesahan konstruk dan kebolehpercayaan instrumen IKGPBD BM dalam mengukur kompetensi guru pelaksanaan PBD bagi mata pelajaran Bahasa Melayu.

Analisis Komponen Prinsipal

Jadual 2: Ringkasan Analisis Komponen Prinsipal (PCA)

Komponen	Nilai
Varians dijelaskan oleh pengukuran Rasch	44.3%
Varians dijelaskan oleh individu	21.6%
Varians dijelaskan oleh item	22.7%
Varians tidak dijelaskan (kontras pertama)	3.1% (eigen = 1.92)

Jadual 2 menunjukkan hasil dapatan Analisis Komponen Prinsipal terhadap Residual (PCA) ini memenuhi syarat unidimensionaliti, iaitu hanya mengukur satu konstruk utama secara konsisten. Berdasarkan jadual di atas, nilai varians mentah yang dijelaskan oleh dimensi Rasch adalah sebanyak 44.3%, melebihi had minimum 40% yang disarankan (Fisher, 2007). Sementara itu, nilai eigen kontras pertama dalam residual ialah 1.92, bersamaan 3.1% daripada jumlah varians, yang berada di bawah ambang 3.0 seperti disarankan oleh Linacre (2010).

Nisbah antara varians dijelaskan oleh item (22.7%) dan kontras pertama (3.1%) ialah 7.3:1, yang jauh melebihi nisbah minimum 3:1. Selain itu, ujian kebebasan setempat mendapati korelasi residual antara pasangan item berada dalam julat -0.44 hingga 0.48, iaitu di bawah ambang 0.70. Ini menunjukkan tiada pelanggaran terhadap andaian kebebasan setempat dan mengesahkan kesahan unidimensi bagi data yang dianalisis.

Indeks Kebolehpercayaan Individu dan Item

Jadual 2.1: Statistik Kebolehpercayaan Instrumen IKGPBD BM

Aspek	Nilai
Indeks Kebolehpercayaan Individu	0.91
Indeks Kebolehpercayaan Item	0.98
Alpha Cronbach	0.92
Outfit MNSQ Individu	1.00
Outfit Zstd Individu	0.00
Outfit MNSQ Item	1.00
Outfit Zstd Item	0.01

Berdasarkan jadual 2.1, instrumen IKGPBD BM menunjukkan tahap kebolehpercayaan yang tinggi. Analisis mendapati nilai kebolehpercayaan individu dan item masing-masing melebihi 0.90, sekali gus mengatasi nilai minimum 0.80 yang lazim dijadikan penanda aras bagi kebolehpercayaan yang baik (Bond & Fox, 2015). Nilai Alpha Cronbach juga berada pada tahap tinggi (0.92), menandakan konsistensi dalaman yang kukuh. Secara keseluruhannya, dapatan ini membuktikan bahawa instrumen mempunyai kestabilan pengukuran yang baik serta mampu menilai keupayaan responden dan kesukaran item dengan tepat.

Nilai Pengasingan Individu dan Item

Nilai pengasingan individu dan item digunakan untuk mengukur keupayaan instrumen membezakan antara tahap kompetensi yang berbeza. Nilai pengasingan individu dalam kajian ini ialah 3.26, yang menunjukkan bahawa instrumen mampu mengklasifikasikan responden kepada sekurang-kurangnya empat kumpulan keupayaan. Nilai pengasingan item pula ialah 7.87, membuktikan bahawa item-item dalam instrumen mampu mengukur tahap kesukaran yang sangat pelbagai, dari paling mudah hingga paling sukar. Ini menunjukkan bahawa instrumen yang dibangunkan bukan sahaja sah, tetapi juga sensitif terhadap perbezaan tahap kompetensi guru.

Analisis Skala Penarafan

Instrumen ini menggunakan skala penarafan empat mata, dan ketepatannya telah dinilai melalui analisis struktur skala. Berdasarkan jadual 2.2 di bawah, semua kategori menunjukkan nilai Outfit MNSQ yang hampir kepada 1.0 (0.92–1.07), menunjukkan responden menggunakan skala ini dengan tepat dan konsisten. Nilai Andrich threshold juga menunjukkan urutan kategori yang logik, dari -1.71 hingga 1.72, dan jarak antara kategori adalah memadai (1.70–1.73 logit).

Jadual 2.2: Kalibrasi Struktur Skala Penarafan IKGPBD BM

Kategori	Bil. Pemerhatian	Purata Pemerhatian	Outfit MNSQ	Andrich Threshold
1	938	-1.35	1.04	none
2	2051	-0.51	0.96	-1.71
3	2006	0.44	0.92	-0.01
4	853	1.30	1.07	1.72

Jadual 2.3: Jarak antara Kategori Skala Penarafan

Beza antara kategori	Jarak antara kategori (Andrich threshold)	Justifikasi
1 – 2	1.71	Skala 1,2,3 dan 4 dikekalkan kerana jarak antara kategori dalam lingkungan 1.40 logit hingga 5.00 logit
2 – 3	1.70	
3 – 4	1.73	

Jadual 2.3 di atas menunjukkan bahawa jarak antara kategori berada dalam julat yang diterima dan menandakan tiada keperluan untuk menggabungkan kategori. Berdasarkan hasil analisis, jarak antara kategori berada dalam lingkungan 1.40 logit hingga 5.00 logit iaitu 1.70 logit hingga 1.73 logit.

Persoalan Kajian 3

Adakah item-item dalam instrumen memenuhi standard pengukuran berdasarkan Model Pengukuran Rasch?

Persoalan kajian ketiga bertujuan menilai sama ada item-item dalam instrumen yang dibangunkan memenuhi standard pengukuran berdasarkan Model Pengukuran Rasch. Untuk menjawab persoalan ini, beberapa analisis telah dijalankan, termasuk analisis kesepadanan item, analisis kesukaran item, serta pemetaan kesukaran item dan kebolehan individu. Kesemua analisis ini memberikan gambaran menyeluruh terhadap kesahan dan keberkesanan item dalam mengukur tahap kompetensi guru dalam pelaksanaan PBD bagi mata pelajaran Bahasa Melayu.

Statistik Kesepadanan Item

Analisis kesepadanan item menunjukkan bahawa secara keseluruhannya, majoriti item dalam instrumen ini menepati jangkaan Model Rasch. Sebanyak 34 daripada 35 item mencatatkan nilai Outfit Mean Square (MNSQ) dalam lingkungan 0.76 hingga 1.23, yang berada dalam julat ideal antara 0.50 hingga 1.50 seperti yang disarankan oleh Linacre (2002) dan Boone et al. (2014). Hal ini membuktikan bahawa item-item tersebut mempunyai tahap kesepadanan yang baik dengan model dan menyumbang secara sah terhadap pengukuran konstruk.

Namun begitu, item 10 dikenal pasti sebagai item yang menyimpang secara signifikan daripada jangkaan model. Item tersebut mencatatkan nilai Outfit MNSQ setinggi 3.18 dan nilai Z-standardized (Zstd) 9.90, menunjukkan ketidaksepadanan yang jelas. Kandungan item ini berkemungkinan terlalu umum dan pelbagai tafsiran oleh responden menyebabkan berlaku ketidakselarasan dalam respons yang diberikan. Justeru, item ini telah digugurkan daripada analisis bagi memastikan konsistensi dan kesahan model dikekalkan.

Semakan terhadap nilai ralat piawai model (Model S.E.) turut menunjukkan bahawa semua item mempunyai nilai antara 0.11 hingga 0.12, jauh lebih rendah daripada ambang 0.25 seperti yang dicadangkan oleh Fisher (2007), menandakan ketepatan tinggi dalam pengukuran item.

Nilai Point-Measure Correlation (PTMea Corr) bagi semua 34 item juga didapati positif dan berada dalam julat 0.40 hingga 0.62. Ini bermakna setiap item dalam instrumen ini menyumbang secara konsisten kepada pengukuran konstruk yang diukur, serta menyokong kesahan struktur instrumen secara keseluruhan.

Kesukaran Item

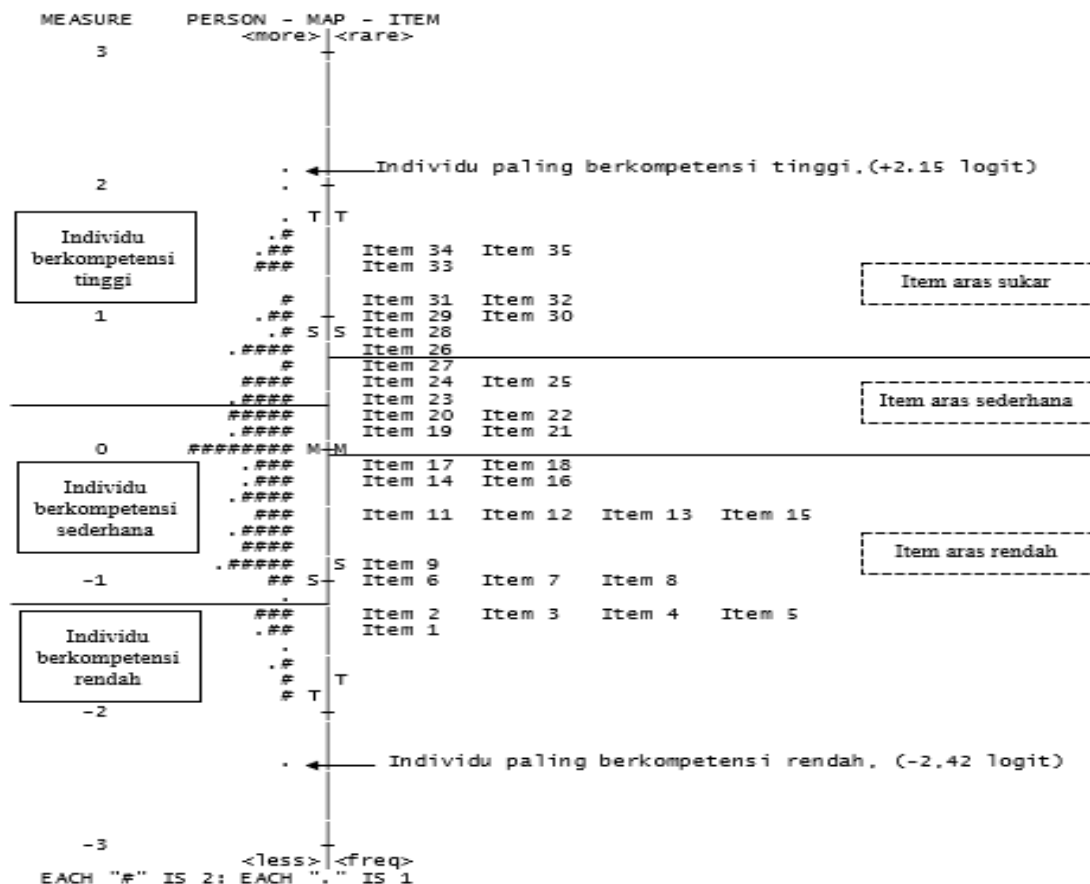
Tahap kesukaran item dianalisis berdasarkan nilai logit yang diperoleh. Nilai logit yang tinggi menandakan item yang lebih sukar, manakala nilai yang rendah menandakan item yang lebih mudah. Berdasarkan analisis, nilai logit dalam IKGPBD BM bertabur dari -1.43 logit (Item 1) hingga 1.54 logit (Item 34), menunjukkan julat kesukaran yang luas dan seimbang. Ini membolehkan pengukuran dilakukan terhadap pelbagai tahap kompetensi guru.

Semua item dalam konstruk pengetahuan (kecuali item 10 yang digugurkan) berada dalam kategori aras mudah. Ini menunjukkan bahawa aspek pengetahuan asas berkaitan PBD umumnya telah dikuasai oleh majoriti guru. Dalam konstruk kemahiran pula, item terbahagi kepada aras mudah dan sederhana. Item 13 hingga 18 diklasifikasikan sebagai mudah, manakala item seperti item 25 berada pada tahap sederhana, mencerminkan cabaran dalam kemahiran penggunaan sistem dalam talian.

Konstruk sikap pula menunjukkan taburan item yang cenderung kepada aras sukar, terutamanya item-item berkaitan komitmen berterusan, penyediaan bahan bantu mengajar, dan pelaksanaan sistematik. Item 34, yang memerlukan guru menyediakan bahan bantu mengajar secara konsisten, dikenal pasti sebagai item paling sukar dalam keseluruhan instrumen. Dapatan ini memperlihatkan aspek sikap sebagai cabaran utama dalam pelaksanaan PBD.

Pemetaan Kesukaran Item-Individu

Analisis pemetaan antara kesukaran item dan keupayaan individu melalui peta item-individu memberikan gambaran yang jelas tentang kesepadanan antara dua komponen penting dalam pengukuran. Rajah 2 di bawah menunjukkan peta-individu yang dihasilkan melalui analisis pengukuran Model Rasch.



Rajah 1: Peta Item-Individu IKGPBD BM

Peta dalam rajah 1 di atas menunjukkan bahawa skala pengukuran item merangkumi keseluruhan julat dari -1.43 logit (item paling mudah) hingga 1.54 logit (item paling sukar), manakala skala kebolehan individu merangkumi julat lebih luas dari -2.42 logit hingga 2.15 logit. Perbezaan julat ini membuktikan bahawa instrumen dapat mengukur pelbagai tahap kebolehan secara menyeluruh.

Nilai purata logit individu (-0.05 logit) adalah sedikit lebih rendah daripada purata logit item (0.00 logit), namun perbezaan ini adalah kecil dan tidak signifikan dalam konteks Rasch. Ini menunjukkan bahawa aras kesukaran item secara umumnya sesuai dengan keupayaan responden, serta menyumbang kepada pengukuran yang adil dan seimbang.

Jadual 3: Tahap Kompetensi Guru dalam Melaksanakan PBD bagi Mata Pelajaran Bahasa Melayu Mengikut Demografi

Demografi	Kompetensi tinggi (Logit individu -0.05 hingga 2.15)	Kompetensi sederhana (Logit individu -0.05 hingga -0.93)	Kompetensi rendah (Logit individu <-0.93)
Jantina			
Lelaki	13 (18.6%)	24 (34.3%)	33 (47.1%)
Perempuan	16 (15.7%)	32 (31.4%)	54 (52.9%)
Demografi	Kompetensi tinggi (Logit individu -0.05 hingga 2.15)	Kompetensi sederhana (Logit individu -0.05 hingga -0.93)	Kompetensi rendah (Logit individu <-0.93)
Pengalaman			
Kurang 5 tahun	3 (12.0%)	10 (40.0%)	12 (48.0%)
6 hingga 10 tahun	11 (16.9%)	20 (30.8%)	34 (52.3%)
11 hingga 15 tahun	9 (18.0%)	15 (30.0%)	26 (52.0%)
16 hingga 20 tahun	3 (13.6%)	8 (36.4%)	11 (50.0%)
Melebihi 20 tahun	3 (30.0%)	3 (30.0%)	4 (40.0%)

Jadual 3 di atas menunjukkan pengelasan responden kepada tiga tahap kompetensi berdasarkan nilai logit individu. Seramai 87 responden dikenal pasti berada dalam kategori kompetensi rendah, 56 dalam kategori sederhana, dan 29 dalam kategori tinggi. Dari segi pengalaman mengajar, guru yang berada dalam fasa pertengahan kerjaya (6 hingga 10 tahun) mencatatkan bilangan tertinggi dalam kategori kompetensi tinggi. Manakala dari segi jantina, guru perempuan dilihat lebih dominan dalam kategori kompetensi tinggi dan sederhana berbanding guru lelaki. Kesimpulannya, ketiga-tiga analisis yang dijalankan menunjukkan bahawa item-item dalam IKGPBD BM memenuhi ciri pengukuran yang sah dan stabil berdasarkan Model Rasch.

Persoalan Kajian 4

Adakah terdapat perbezaan signifikan nilai min logit tahap kompetensi guru dalam pelaksanaan pentaksiran bilik darjah bagi mata pelajaran Bahasa Melayu berdasarkan tempoh pengalaman mengajar?

Kajian ini turut meneliti sama ada wujud perbezaan yang signifikan terhadap tahap kompetensi guru dalam melaksanakan Pentaksiran Bilik Darjah (PBD) bagi mata pelajaran Bahasa Melayu berdasarkan tempoh pengalaman mengajar. Analisis ini dirangka berlandaskan hipotesis kajian yang menyatakan tidak terdapat perbezaan signifikan antara kumpulan guru berdasarkan pengalaman mengajar, berbanding hipotesis alternatif yang menjangka wujud perbezaan.

Bagi tujuan ini, pengalaman mengajar guru-guru yang terlibat dalam kajian telah dibahagikan kepada lima kategori, iaitu: kurang daripada 5 tahun, 6 hingga 10 tahun, 11 hingga 15 tahun, 16 hingga 20 tahun, dan melebihi 20 tahun. Nilai logit yang diperoleh daripada Model Pengukuran Rasch digunakan sebagai asas bagi menentukan tahap kompetensi setiap kumpulan guru.

Analisis deskriptif menunjukkan terdapat sedikit variasi min logit antara kumpulan. Guru dengan pengalaman kurang lima tahun mencatatkan skor purata tertinggi ($M = 0.01$), manakala kumpulan paling berpengalaman melebihi dua puluh tahun memperoleh skor purata terendah ($M = -0.26$). Nilai sisihan piawai kumpulan ini juga paling besar, menandakan variasi kompetensi yang lebih luas berbanding kumpulan lain.

Jadual 4: Hasil Dapatan Analisis Anova

	Jumlah kuasa dua	Darjah kebebasan	Purata kuasa dua	F	Sig.
Antara kumpulan	0.662	4	0.166	0.207	0.934
Dalam kumpulan	133.75	167	0.801		
Jumlah	134.41	171			

Bagi mengenal pasti sama ada perbezaan nilai min logit antara kumpulan tersebut adalah signifikan secara statistik, satu analisis ANOVA sehalu telah dijalankan. Hasil analisis dalam jadual 4 di atas menunjukkan bahawa tiada perbezaan yang signifikan dalam tahap kompetensi guru berdasarkan pengalaman mengajar, seperti yang ditunjukkan melalui nilai $F(4, 167) = 0.207$ dan nilai signifikan $p = 0.934$. Nilai p ini jauh melebihi aras signifikan yang lazim digunakan iaitu 0.05, maka hipotesis nol diterima dan hipotesis alternatif ditolak.

Kesimpulannya, walaupun terdapat sedikit variasi purata nilai logit antara kumpulan guru berdasarkan pengalaman, analisis statistik membuktikan bahawa perbezaan tersebut tidak signifikan. Ini membawa maksud bahawa tahap kompetensi pelaksanaan PBD dalam kalangan guru Bahasa Melayu adalah relatif seragam merentas pelbagai peringkat pengalaman mengajar, dan sebarang perbezaan yang wujud tidak cukup signifikan untuk membuat inferens umum mengenainya.

Perbincangan

Kajian ini telah berjaya membina dan mengesahkan instrumen bagi mengukur tahap kompetensi guru dalam pelaksanaan Pentaksiran Bilik Darjah (PBD) bagi mata pelajaran Bahasa Melayu di sekolah rendah. Perbincangan ini menghuraikan dapatan berdasarkan empat objektif utama serta mengaitkannya dengan teori, literatur terdahulu, faktor kontekstual dalam pendidikan Malaysia, dan implikasi praktikal.

Secara keseluruhan, dapatan kajian ini menunjukkan instrumen yang dibangunkan mempunyai kesahan kandungan yang kukuh, hasil daripada penilaian pakar dan analisis Model Rasch pelbagai faset. Kesepakatan tinggi antara pakar menunjukkan bahawa item instrumen benar-

benar mencerminkan konstruk yang diukur. Dapatan ini menyokong hujah Sidek (2020) yang menegaskan kepentingan semakan pakar dalam memastikan liputan konstruk yang menyeluruh dan ketepatan konteks. Ia membuktikan bahawa proses pengesahan kandungan merupakan asas penting dalam pembangunan instrumen, terutama dalam konteks pentaksiran pendidikan.

Kesahan konstruk dan kebolehpercayaan instrumen turut disokong melalui analisis Rasch. Instrumen mampu membezakan tahap kompetensi guru secara konsisten, dengan unidimensi konstruk yang jelas seperti dicadangkan oleh Bond dan Fox (2015). Hal ini membuktikan bahawa pengukuran benar-benar menilai kompetensi yang sama, bukan faktor luar. Di samping itu, dapatan ini menambah keyakinan bahawa Model Rasch lebih berkesan berbanding analisis klasik seperti EFA atau CFA, kerana ia menilai kesepadanan item secara individu.

Analisis terhadap kualiti item menunjukkan bahawa proses saringan yang teliti melalui kajian rintis dan lapangan telah menghasilkan set item yang mantap. Item yang tidak menepati kriteria statistik digugurkan bagi memastikan integriti pengukuran, selaras dengan pandangan Linacre (2021). Selain itu, analisis skala penarafan membuktikan skala empat mata berfungsi dengan baik dan membantu mengelakkan jawapan neutral. Hal ini memastikan responden membuat pilihan yang jelas, sama ada positif atau negatif.

Dapatan mengenai kompetensi guru berdasarkan pengalaman mengajar menunjukkan tiada perbezaan signifikan tahap kompetensi guru berdasarkan pengalaman mengajar. Hal ini membuktikan bahawa pengalaman semata-mata tidak menjamin kompetensi yang lebih tinggi. Faktor lain seperti keberkesanan latihan, sokongan profesional, dan motivasi individu lebih mempengaruhi, selaras dengan dapatan Khairuddin dan Yaakub (2022). Dalam konteks pendidikan di Malaysia, beban tugas guru yang tinggi serta pelaksanaan dasar baharu seperti PBD menuntut penyesuaian berterusan. Guru baharu lebih terbiasa dengan pendekatan PBD kerana terlatih dengan dasar semasa, manakala guru berpengalaman berdepan cabaran untuk menyesuaikan diri dengan perubahan.

Dapatan kajian ini mempunyai implikasi penting terhadap dasar pendidikan, khususnya berkaitan pelaksanaan Pentaksiran Bilik Darjah (PBD). Instrumen IKGPBD BM yang dibangunkan boleh dijadikan asas oleh Kementerian Pendidikan Malaysia (KPM) untuk menilai tahap kompetensi guru secara lebih sistematik dan berasaskan bukti. Dengan adanya data empirikal, KPM dapat merangka program pembangunan profesional yang lebih tersasar, misalnya dengan menekankan latihan berfokus kepada aspek kelemahan tertentu guru. Hal ini juga dapat membantu mengurangkan beban tugas guru kerana latihan yang dirancang lebih relevan dengan keperluan sebenar, bukan sekadar bersifat umum.

Selain itu, Institut Pendidikan Guru (IPG) dan fakulti pendidikan boleh memanfaatkan instrumen ini untuk menambah baik kurikulum latihan perguruan agar lebih selaras dengan dasar semasa PBD. Implikasi ini menunjukkan bahawa hasil kajian bukan sahaja menyumbang kepada peningkatan amalan bilik darjah, tetapi juga menyokong usaha reformasi dasar pendidikan negara ke arah pelaksanaan pentaksiran yang lebih autentik, formatif, dan berkesan.

Dari sudut teori, kajian ini memperlihatkan sumbangan penting Model Rasch dalam pengukuran instrumen pendidikan di Malaysia. Model ini bukan sahaja menyediakan penilaian yang objektif dan bebas sampel, tetapi juga memperkenalkan pendekatan pemetaan yang lebih bermakna dalam menghubungkan kebolehan responden dengan tahap kesukaran item. Hal ini

menunjukkan anjakan ke arah penggunaan analisis psikometrik yang lebih saintifik dan relevan dalam penyelidikan pendidikan.

Akhir sekali, walaupun kajian ini telah menyumbang secara signifikan, ruang untuk penambahbaikan masih wujud. Kajian lanjutan boleh memperluas skop kepada guru sekolah menengah atau guru dari negeri lain untuk menilai kesahan instrumen dalam konteks yang lebih luas. Analisis tambahan seperti Differential Item Functioning (DIF) atau Confirmatory Factor Analysis (CFA) juga boleh digunakan bagi menguji bias item dan mengesahkan struktur konstruk. Kajian kualitatif juga boleh dijalankan untuk meneroka cabaran sebenar guru dalam melaksanakan PBD, sekali gus melengkapkan dapatan kuantitatif yang diperoleh.

Kesimpulan

Secara keseluruhannya, instrumen yang dibangunkan dan disahkan dalam kajian ini telah memenuhi syarat psikometrik dari aspek kesahan kandungan, kesahan konstruk, kebolehpercayaan, dan kualiti item. Kajian ini bukan sahaja membuktikan keberkesanan instrumen ini dalam konteks pendidikan Bahasa Melayu, malah turut menyumbang kepada pembangunan instrumen pendidikan yang berkualiti, sahih dan sesuai digunakan oleh pelbagai pihak untuk mempertingkatkan amalan pentaksiran bilik darjah di Malaysia.

Penghargaan

Setinggi-tinggi penghargaan dan terima kasih kepada Assoc. Prof. Dr. Mohd Nazri bin Abdul Rahman atas segala tunjuk ajar serta bimbingan yang diberikan. Ucapan terima kasih juga ditujukan kepada Bahagian Tajaan Pendidikan, Kementerian Pendidikan Malaysia, pihak pentadbir serta guru besar sekolah. Tidak dilupakan para responden yang terlibat secara langsung dalam memberikan maklumat dan pengalaman berharga dalam menjayakan kajian ini.

Rujukan

- Andrich, D. (2004). Controversy and the Rasch model: A characteristic of incompatible paradigms. *Medical Care*, 42(1), 1–7.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.
- Boone, W. J. (2016). *Rasch analysis for instrument development: Why, when, and how?* CBE—Life Sciences Education.
- Cai, Q. Y. (2025). Formative quizzes in L2 Chinese classrooms: Effects on summative assessment outcomes. *International Journal for Multilingual Education*.
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE Publications.
- El Zerk, M. (2025). Classroom-based assessment techniques for EFL skills and components in Moroccan high schools. *International Jordanian Journal Aryam for Humanities*.
- Fine, C. G. M. C., & Furtak, E. M. (2025). A call to reimagine science formative assessment systems through translanguaging. *Science Education*. <https://doi.org/10.1002/sce.70008>
- Han, C., Lu, X., & Fan, Q. (2025). Taming generative AI for interpreter education: Using large language models in classroom-based assessment of English-Chinese consecutive interpreting. *The Interpreter and Translator Trainer*.
- Jain, A., & Narayanan, P. (2021). Competency mapping in teacher education: Designing holistic assessment tools. *Journal of Educational Research and Practice*, 11(2), 98–112.

- Jamil, H., & Said, M. N. H. (2019). Cabaran pelaksanaan pentaksiran bilik darjah dalam kalangan guru Bahasa Melayu. *Jurnal Pendidikan Bahasa Melayu*, 9(1), 24–35.
- Kay, S. (2025). Beyond memorization: Developing three-dimensional assessments for Earth science through the carbon cycle unit (Master's thesis, Grand Valley State University).
- Khairuddin, M. Z., & Yaakub, N. F. (2022). Assessing the validity of teacher competency instruments: A review of current practices and gaps. *Malaysian Journal of Learning and Instruction*, 19(1), 101–120.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575.
- Li, M., Zan, Z., Liu, L., Gu, H., Zhang, Y., & You, L. (2025). Development and testing of physical literacy scales for Chinese elementary school students
- Linacre, J. M. (2021). A user's guide to WINSTEPS: Rasch-model computer programs.