

Artikel

Kepimpinan Sebagai Penggerak Kajian Tindakan: Penilaian Instrumen bagi Menyemai Kualiti Guru Sekolah Rendah

(Leadership as a Catalyst for Action Research: Instrument Validation for Cultivating Teacher Quality in Primary Schools)

Azlin Moktar, Bity Salwana Alias* & Nurazidawati Mohamad Arsad

Fakulti Pendidikan, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

*Pengarang Koresponden: bity@ukm.edu.my

Diserah: 28 Julai 2025

Diterima: 11 November 2025

Abstrak: Peningkatan profesionalisme guru dalam pendidikan abad ke-21 menuntut pendekatan reflektif dan berasaskan data seperti kajian tindakan. Namun, pelaksanaan kajian tindakan dalam kalangan guru sekolah rendah masih bersifat tidak konsisten akibat kekangan masa, kurang sokongan kepimpinan, serta ketiadaan instrumen yang sah untuk menilai amalan ini secara sistematik. Kajian ini bertujuan membangunkan dan menilai kesahan serta kebolehpercayaan satu instrumen yang mengukur pembudayaan kajian tindakan oleh guru besar, komitmen guru terhadap kajian tindakan, dan kualiti guru sekolah rendah. Kajian rintis ini menggunakan reka bentuk tinjauan kuantitatif melibatkan 107 orang guru sebagai sampel rintis. Analisis kesahan meliputi penilaian kandungan oleh lima pakar bidang, kesahan muka oleh tiga pakar bahasa dan 30 guru, serta kesahan konstruk melalui Model Rasch menggunakan perisian *Winsteps* versi 3.73. Hasil analisis menunjukkan 65 item akhir berada dalam julat statistik kesesuaian yang diterima (*Infit/Outfit MNSQ* 0.5-1.5), unidimensionaliti disokong melalui *PCA residual*, dan instrumen mempunyai keupayaan diskriminatif yang tinggi dengan indeks pengasingan responden 6.28. Nilai kebolehpercayaan dalaman adalah sangat tinggi dengan *Cronbach's Alpha* (0.99) dan korelasi skor-logit (0.96). Pemetaan *Wright* menunjukkan penyebaran item yang seimbang, walaupun terdapat keperluan menambah item berkesukaran tinggi bagi menangani isu *mistargeting*. Secara keseluruhan, instrumen ini sah, konsisten, dan kontekstual untuk menilai budaya penyelidikan di sekolah, merancang intervensi latihan guru, serta menyokong pelaksanaan dasar pendidikan berasaskan data.

Kata kunci: Kajian tindakan, kesahan instrumen, kebolehpercayaan, kualiti guru, Model Rasch

Abstract: The enhancement of teacher professionalism in 21st-century education requires reflective and data-driven approaches such as action research. However, the implementation of action research among primary school teachers in Malaysia remains inconsistent due to time constraints, heavy workload, lack of leadership support, and the absence of valid instruments to systematically evaluate this practice. This study aimed to develop and validate a new instrument measuring the culture of action research led by principals, teachers' commitment to action research, and teacher quality. A quantitative survey design was employed with 107 primary school teachers as pilot respondents. Validity analyses included content validation by five subject-matter experts, face validation by three Malay language experts and 30 teachers, and construct validation through the Rasch Measurement Model using *Winsteps* version 3.73. Findings indicated that 65 final items fell within the acceptable fit range (*Infit/Outfit MNSQ* 0.5–1.5), with unidimensionality supported by *PCA residuals*. The instrument demonstrated high discriminative power (person separation index = 6.28) and

excellent internal consistency (Cronbach's Alpha = 0.99; person score-to-measure correlation = 0.96). The Wright Map confirmed balanced targeting between items and respondents, though additional high-difficulty items are recommended to address mistargeting issues. Overall, the instrument is valid, reliable, and contextually relevant for large-scale studies, enabling the assessment of research culture in schools, guiding teacher training interventions, and supporting evidence-based educational policies.

Keywords: action research, instrument validity, reliability, teacher quality, Rasch Model.

Pengenalan

Peningkatan profesionalisme guru merupakan asas kepada pemerkasaan sistem pendidikan, terutamanya dalam menghadapi cabaran Revolusi Industri Keempat (BPG, 2023). Salah satu pendekatan yang diiktiraf berkesan dalam membina kompetensi guru ialah kajian tindakan, yang berfungsi sebagai alat reflektif untuk memperkukuh amalan pedagogi, penguasaan pengetahuan profesional dan kepimpinan instruksional (Widayati et al., 2021; Tifanny & Roie, 2024). Tambahan pula, pendekatan pedagogi yang berteraskan refleksi dan adaptasi berasaskan bukti telah terbukti berkesan dalam meningkatkan ketahanan guru terhadap cabaran pendidikan semasa, termasuk perubahan dasar dan teknologi pendidikan (Chukwuma & Zondo, 2025). Namun begitu, dalam konteks sekolah rendah di Malaysia, pelaksanaan kajian tindakan masih terhad akibat kekangan masa, beban tugas yang tinggi, dan kurangnya keyakinan terhadap keberkesanan kaedah ini (Ag-Ahmad et al., 2023; Widayati et al., 2021, Adhikari, 2023).

Beberapa inisiatif telah dilaksanakan oleh agensi seperti IAB, BPPDP dan JPN untuk menyemarakkan amalan kajian tindakan, namun keberkesanan program ini masih kurang dinilai secara sistematik. Dalam hal ini, peranan guru besar sebagai peneraju budaya kajian tindakan menjadi semakin penting. Kepimpinan sekolah yang menyokong, membimbing dan menyediakan sumber yang mencukupi dapat meningkatkan komitmen guru dan mendorong amalan reflektif yang konsisten (Fullan, 2015; Aulls & Shore, 2023). Kajian oleh Youngs dan Printy (2022) menunjukkan bahawa kekurangan latihan dan bimbingan kepada pemimpin sekolah telah menyekat keupayaan mereka untuk memanfaatkan kajian tindakan sebagai mekanisme penambahbaikan pengajaran. Dapatan daripada Latip dan Abd Latif (2024) turut menegaskan keperluan pemimpin sekolah untuk memiliki kapasiti instruksional yang mantap serta kemahiran menyantuni guru secara kolaboratif agar inovasi pedagogi dapat diterapkan secara berkesan. Dapatan serupa turut dikemukakan oleh Moktar et al. (2024), yang menekankan bahawa sokongan pengurusan yang terhad merupakan cabaran utama pelaksanaan kajian tindakan di sekolah rendah.

Tambahan pula, dapatan kuantitatif oleh Talib et al. (2024) menunjukkan bahawa kompetensi kepimpinan memainkan peranan penting dalam memastikan pelaksanaan inisiatif pendidikan seperti *STEM* berjalan dengan berkesan. Mereka mendapati bahawa kejayaan transformasi sekolah sangat bergantung pada keupayaan pemimpin menyediakan hala tuju strategik, menyokong pembelajaran guru, serta memupuk komitmen melalui komunikasi dan sokongan berterusan. Penemuan ini selari dengan Ambon et al. (2024), yang menegaskan bahawa pembinaan budaya profesional dan pelaksanaan perubahan kurikulum bergantung rapat antara kepimpinan instruksional dan kesediaan guru. Dalam konteks kajian tindakan, hal ini memperlihatkan bahawa faktor kepimpinan dan komitmen merupakan asas utama kepada kejayaan pembudayaan penyelidikan dalam kalangan guru.

Meskipun dokumen seperti Standard Guru Malaysia 2.0 (SGM 2.0) dan PPPB telah digariskan untuk menyokong pembangunan profesional guru, masih tiada instrumen khusus yang sah dan boleh dipercayai untuk menilai pembudayaan kajian tindakan yang digerakkan oleh kepimpinan sekolah, tahap komitmen guru, serta hubungannya dengan kualiti guru. Instrumen-instrumen sedia ada yang mengukur aspek profesionalisme guru dan kepimpinan pendidikan kebanyakannya dibangunkan dalam konteks luar negara dan tidak mengambil kira kerangka dasar tempatan seperti SKPMg2, SGM 2.0 dan PPPB secara menyeluruh (Latip & Abd Latif, 2024; Talib, Alias & Matore, 2024). Tambahan pula, kebanyakan alat ukur terdahulu menilai komitmen, kepimpinan atau kualiti secara berasingan tanpa mempertimbangkan interaksi dinamik antara ketiga-tiga konstruk tersebut dalam persekitaran sekolah rendah di Malaysia. Justeru, ketiadaan alat ukur yang

menyepadukan pembudayaan kajian tindakan oleh pemimpin, komitmen guru dan kualiti guru dalam kerangka pengukuran yang seragam mencetuskan keperluan mendesak terhadap pembangunan instrumen yang kontekstual dan holistik (Mustapha & Rahman, 2024; Spencer & Spencer, 1993). Seiring dengan penekanan terhadap pendidikan yang berdaya saing dan berasaskan penyelidikan, kajian rintis ini juga bertujuan mengisi kelompongan terhadap amalan penilaian empirikal yang menyeluruh terhadap budaya penyelidikan guru di sekolah. Oleh yang demikian, satu kajian rintis telah dijalankan bagi menilai kesahan dan kebolehpercayaan instrumen yang dibangunkan untuk mengukur ketiga-tiga konstruk tersebut dalam konteks sekolah rendah di Malaysia.

Objektif kajian:

- i. Menilai kesahan kandungan, kesahan muka dan kesahan konstruk instrumen kajian.
- ii. Menentukan kebolehpercayaan dalaman instrumen dari segi konsistensi item.
- iii. Mengukuhkan asas penggunaan instrumen bagi kajian utama pembudayaan kajian tindakan, komitmen guru dan kualiti guru.

Metodologi

1. Reka Bentuk Kajian

Kajian ini menggunakan pendekatan kuantitatif dengan reka bentuk tinjauan, yang bertujuan untuk mengumpulkan data berkaitan pembudayaan kajian tindakan, komitmen guru, dan kualiti guru sekolah rendah di Malaysia. Reka bentuk ini membolehkan pengukuran persepsi dan pengalaman responden dikumpulkan bagi menjawab persoalan kajian. Kajian ini memberi tumpuan kepada penilaian kajian rintis bagi menguji kesahan dan kebolehpercayaan instrumen yang dibangunkan.

2. Pensampelan

Seramai 107 orang guru sekolah rendah dipilih sebagai responden kajian rintis ini. Kaedah pensampelan kebarangkalian secara rawak mudah digunakan bagi memastikan setiap guru dalam populasi sasaran mempunyai peluang yang saksama untuk dipilih. Pemilihan responden melibatkan guru-guru daripada pelbagai lokasi dan latar belakang bagi meningkatkan kebolehwakilan dapatan kajian rintis ini terhadap populasi guru sekolah rendah kebangsaan. Pendekatan ini bertujuan meminimumkan bias pemilihan dan meningkatkan kesahan luar instrumen yang diuji. Pemilihan dibuat berdasarkan senarai nama guru yang dibekalkan oleh pihak sekolah, dan hanya guru yang bersetuju secara sukarela untuk mengambil bahagian serta memberikan maklum balas melalui soal selidik telah dimasukkan dalam sampel kajian ini.

3. Pembangunan Instrumen

Instrumen kajian ini merupakan satu set soal selidik yang dibangunkan khusus bagi mengukur tiga konstruk utama, iaitu pembudayaan kajian tindakan oleh guru besar, komitmen guru terhadap kajian tindakan, dan kualiti guru sekolah rendah. Sebanyak 70 item telah dibina dan disusun berdasarkan kerangka teori yang kukuh dan sejajar dengan konteks pendidikan di Malaysia. Antara teori yang menjadi asas kepada pembangunan item ini termasuk Teori Pengurusan Perubahan John Kotter (1996), Teori Guest (1997), Model Komitmen untuk Perubahan Organisasi oleh Herscovitch dan Meyer (2002), Teori Kompetensi oleh Spencer dan Spencer (1993), Standard Kompetensi Pemimpin Sekolah 2.0 (KOMPAS 2.0) (IAB, 2020), Standard Kualiti Pendidikan Malaysia Gelombang 2 (SKPMg2)(KPM, 2018), dan Standard Guru Malaysia 2.0 (SGM 2.0)(KPM, 2021).

Proses penjaan item dilaksanakan secara sistematik dengan memetakan konstruk dan subkonstruk kepada indikator berdasarkan kerangka teori dan standard profesional. Item dibangunkan secara deduktif dengan rujukan kepada domain teori serta dapatan kajian lepas yang relevan dalam konteks Malaysia. Bagi kesahan kandungan, lima (5) orang pakar terlibat, iaitu tiga (3) pakar dalam bidang pengurusan pendidikan dan dua (2) pakar dalam bidang kajian tindakan. Penilaian ini menekankan aspek kerelevanan item, keselarasan dengan teori, dan kejelasan dalam mewakili konstruk yang diukur. Sumbangan pakar ini

membantu dalam menyaring item-item yang kurang sesuai dan memperkukuh kesepadanan antara item dan konstruk.

Kesahan muka pula dinilai dalam dua fasa: pertama, melalui semakan oleh tiga (3) pakar Bahasa Melayu yang memfokuskan kepada aspek linguistik dan kejelasan bahasa; dan kedua, melalui ujian persepsi awal oleh 30 orang guru sekolah rendah, bagi menilai tahap kefahaman terhadap setiap item. Maklum balas daripada kedua-dua kumpulan ini digunakan untuk menambah baik struktur ayat, mengganti istilah teknikal yang terlalu abstrak, serta memastikan kebolehfahaman yang tinggi dalam kalangan responden sebenar kajian utama. Kesahan konstruk instrumen dinilai melalui analisis Model Pengukuran Rasch menggunakan perisian *Winsteps* versi 3.73. Analisis ini membolehkan pemeriksaan terhadap fungsi item, kesesuaian responden, unidimensionaliti konstruk, serta kestabilan skala pengukuran. Ciri-ciri ini penting bagi memastikan instrumen berupaya menilai konstruk secara sah dan menyeluruh, selaras dengan prinsip pengukuran psikometrik yang mantap (Bond & Fox, 2015; Linacre, 2022).

Semua item disusun dalam bentuk skala Likert lima mata, iaitu daripada 1=Sangat Tidak Setuju hingga 5=Sangat Setuju. Skala ini dipilih kerana kemampuannya mengukur konstruk laten secara konsisten serta memudahkan pelaksanaan analisis statistik lanjutan (El-Den et al., 2020). Setiap item telah dirumus agar selaras dengan dimensi teori yang disandarkan, bertujuan untuk memastikan kesahan kandungan dan kesepaduan konsep. Struktur instrumen ini bukan sahaja memudahkan responden menjawab, malah direka bentuk bagi menyokong penilaian psikometrik yang kukuh melalui prosedur kesahan dan kebolehpercayaan.

4. Prosedur Pengumpulan Data

Proses pengumpulan data dimulakan dengan memperoleh kebenaran rasmi daripada eRAs KPM, Jabatan Pendidikan Negeri (JPN), dan pihak pengurusan sekolah yang terlibat. Setelah kelulusan diperolehi, senarai nama guru dikemukakan oleh pihak sekolah dan pemilihan responden dilakukan secara rawak mudah daripada senarai tersebut. Guru-guru yang terpilih diberikan pautan borang persetujuan penyertaan, yang mengandungi maklumat berkaitan tujuan kajian, jaminan kerahsiaan data, dan hak untuk menarik diri pada bila-bila masa.

Seterusnya, satu sesi taklimat secara dalam talian telah diadakan untuk memberikan penerangan lanjut kepada guru yang terpilih, termasuk tatacara menjawab soal selidik dan tempoh pelaksanaannya. Setelah guru memberikan persetujuan, pautan soal selidik dalam talian telah diedarkan melalui emel kepada semua responden yang bersetuju untuk menyertai kajian. Proses pengumpulan data dilaksanakan dalam tempoh 14 hari, bagi memberi ruang mencukupi kepada guru menyempurnakan soal selidik secara sendiri tanpa gangguan kepada komitmen tugas harian mereka.

5. Analisis Data

Data yang diperolehi dianalisis menggunakan beberapa teknik statistik bagi menilai sifat psikometrik instrumen. Analisis deskriptif melibatkan pengiraan min dan sisihan piawai bagi setiap item dan konstruk untuk mendapatkan gambaran awal mengenai taburan skor responden. Bagi menilai kesahan kandungan, Indeks Kesahan Kandungan (CVI) telah digunakan berdasarkan penilaian lima orang pakar dalam bidang kepimpinan dan kajian tindakan. Kesahan muka pula diperolehi melalui maklum balas daripada pakar Bahasa Melayu dan guru berkaitan kejelasan dan kesesuaian item yang dibangunkan.

Seterusnya, kesahan konstruk dinilai menggunakan Model Pengukuran Rasch melalui perisian *Winsteps* versi 3.73. Model *Rating Scale* digunakan memandangkan semua item menggunakan skala Likert lima mata yang seragam. Anggaran parameter dilakukan menggunakan pendekatan *Joint Maximum Likelihood Estimation* (JMLE), dan pemantauan kesesuaian item dilakukan melalui statistik *Infit* dan *Outfit Mean Square (MNSQ)* dalam julat penerimaan 0.5 hingga 1.5. Struktur kategori respons diperiksa melalui *Andrich thresholds*, dan keputusan menunjukkan urutan kategori tersusun secara logik, menandakan skala Likert berfungsi dengan baik dari segi psikometrik.

Pemetaan *Wright* turut dijalankan bagi menilai kesepadanan antara tahap kesukaran item dan kebolehan responden. Nilai indeks kebolehpercayaan item dan responden serta indeks pengasingan menunjukkan keupayaan diskriminatif instrumen. Rancangan *Differential Item Functioning* (DIF) tidak dilaksanakan kerana saiz sampel rintis ($n=47$) tidak mencukupi untuk analisis perbandingan antara

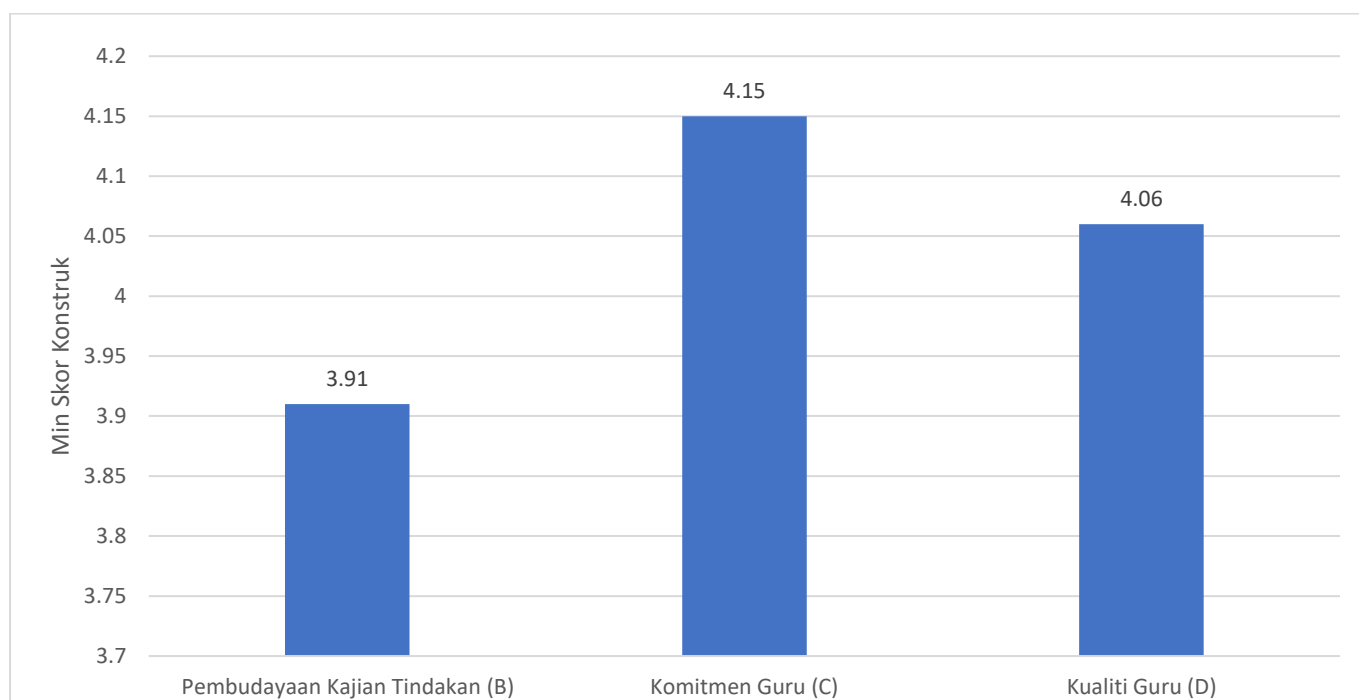
subkumpulan. Akhir sekali, kebolehpercayaan dalaman instrumen disahkan melalui nilai pekali *Cronbach's Alpha*, yang digunakan untuk menilai tahap konsistensi antara item-item dalam setiap konstruk.

Hasil Kajian dan Perbincangan

1. Statistik Deskriptif Item

Dapatan analisis deskriptif menunjukkan bahawa ketiga-tiga konstruk yang dikaji mencatatkan skor min yang berada dalam julat sederhana tinggi hingga tinggi. Konstruk Pembudayaan Kajian Tindakan (Item B1-B25) merekodkan skor purata 3.91, manakala komitmen guru (Item C1-C15) mencatatkan skor tertinggi iaitu 4.15, diikuti oleh kualiti guru (Item D1-D30) sebanyak 4.06 seperti dipaparkan di Rajah 1. Skor min item tertinggi dalam setiap konstruk menunjukkan kecenderungan positif terhadap elemen-elemen seperti inisiatif perubahan, kesediaan guru untuk terlibat dalam kajian tindakan, serta keyakinan terhadap kompetensi profesional.

Tambahan laporan taburan *ceiling/floor* menunjukkan bahawa beberapa item seperti C3 (komitmen afektif) mencatat *ceiling effect* > 70%, manakala item seperti D1 (pengetahuan guru) menunjukkan *floor effect* rendah (<10%). Dapatan ini menunjukkan wujudnya respons bersetuju yang dominan dalam kalangan responden terhadap item-item positif, konsisten dengan dapatan Talib et al., (2024) yang menegaskan bahawa kepimpinan sekolah yang efektif berperanan penting dalam meningkatkan komitmen guru terhadap pelaksanaan inisiatif berasaskan penyelidikan. Tambahan pula, skor tinggi dalam kualiti guru menunjukkan keselarasan dengan dapatan Ambon et al. (2024) yang mendapati bahawa penglibatan guru dalam komuniti pembelajaran profesional seperti kajian tindakan menyumbang kepada peningkatan profesionalisme dan refleksi sendiri guru.



Rajah 1. Carta Bar Min Skor bagi Konstruk

2. Analisis Kesahan

Kesahan kandungan instrumen telah disahkan melalui penilaian lima orang pakar (tiga dalam pengurusan pendidikan, dua dalam kajian tindakan). Nilai Indeks Kesahan Kandungan Purata (S-CVI/Ave) bagi hampir semua domain adalah pada tahap maksimum (1.00), manakala domain *Tanggungjawab Profesional* menunjukkan sedikit deviasi (S-CVI/Ave = 0.96; S-CVI/UA = 0.80), namun masih dalam julat diterima (Polit

& Beck, 2006). Penemuan ini menunjukkan bahawa item-item yang dibangunkan adalah relevan dan representatif terhadap konstruk yang diukur.

Penilaian kesahan muka menunjukkan kebolehfahaman dan kejelasan ayat yang tinggi berdasarkan semakan tiga pakar Bahasa Melayu dan 30 guru sekolah rendah. Walaupun terdapat beberapa item yang memerlukan penambahbaikan dari segi istilah teknikal seperti 'berwawasan' dan 'trait', dapatan ini secara keseluruhan menyokong kesesuaian linguistik instrumen. Kajian lepas oleh Makhsin, Teoh & Ismail (2022) turut menekankan kepentingan penyesuaian bahasa instrumen agar relevan dengan konteks pendidikan Malaysia.

3. Analisis Kesahan Konstruk

Kajian ini menggunakan pendekatan Model Pengukuran Rasch dengan perisian *Winstep* versi 3.73, untuk menganalisis kesahan konstruk secara terperinci. Daripada 107 respons awal yang diperoleh dalam kajian rintis ini, hanya 47 respons digunakan dalam analisis akhir setelah 60 respons digugurkan berikutan ketidakpatuhan terhadap julat statistik fit yang dibenarkan.

Kesesuaian Item dan Responden

Sebanyak 65 item akhir memenuhi kriteria *fit* berdasarkan nilai *Infit* dan *Outfit MNSQ* dalam julat 0.5 hingga 1.5 (Bond & Foz, 2015). Nilai purata *Infit MNSQ* ialah 0.97 dan *Outfit MNSQ* ialah 0.69 seperti dipaparkan di Jadual 1. Korelasi titik-ukuran pula mencatat nilai 0.72 (melebihi ambang 0.30), menunjukkan setiap item menyumbang secara positif terhadap konstruk yang diukur. Fungsi kategori diperiksa melalui analisis *Andrich thresholds*, yang menunjukkan urutan tersusun secara logik, menandakan skala Likert lima mata berfungsi secara stabil (Linacre, 2022).

Tambahan pula, instrumen menunjukkan kebolehpercayaan yang tinggi dengan nilai kebolehpercayaan item 0.84 dan indeks pengasingan item 2.28, kedua-duanya melepasi had minimum ≥ 0.80 dan ≥ 0.20 (Bond & Fox, 2015). Analisis semua parameter *fit* serta rasional pengguguran lima item (B5, B11, B15, C4, D1) yang menunjukkan statistik luar julat atau korelasi titik-ukuran rendah ditunjukkan dalam Jadual 2. Keputusan ini mengesahkan keupayaan diskriminatif instrumen untuk membezakan tahap keupayaan guru secara berkesan, dan kestabilan fungsi item dalam pelbagai keadaan pentaksiran.

Jadual 1. Kesesuaian Item dan Responden

Parameter	Nilai	Kriteria disyorkan
<i>Infit MNSQ</i> (Purata)	0.97	0.5 - 1.5
<i>Outfit MNSQ</i> (Purata)	0.69	0.5 - 1.5
Korelasi Titik-Ukur (Purata)	0.72	> 0.3
Kebolehpercayaan Item	0.84	≥ 0.80
Indeks Pengasingan Item	2.28	≥ 2.0

Sumber: Bond & Fox (2015)

Jadual 2. Jadual fit item luar julat

Item	<i>Infit MNSQ</i>	<i>Outfit MNSQ</i>	Korelasi titik ukur	Keputusan
B5	1.61	1.81	0.28	Digugurkan- <i>Infit</i> , <i>Outfit</i> dan korelasi luar julat
B11	1.85	0.48	0.22	Digugurkan- <i>Infit</i> , <i>Outfit</i> dan korelasi luar julat
B15	1.52	1.67	0.26	Digugurkan- <i>Infit</i> , <i>Outfit</i> dan korelasi luar julat
C4	0.42	1.48	0.21	Digugurkan- <i>Infit</i> dan korelasi rendah
D1	1.55	1.6	0.25	Digugurkan-semua indikator tidak memenuhi syarat

Unidimensionaliti

Unidimensionaliti merupakan syarat asas dalam Model Pengukuran Rasch yang memastikan setiap subkonstruk hanya mengukur satu ciri laten yang konsisten dan tidak diganggu oleh dimensi sekunder. Bagi tujuan ini, analisis *Principal Component Analysis of Residuals* (PCA), yang mengukur varians yang dijelaskan serta nilai eigen kontras pertama bagi setiap subkonstruk telah dijalankan secara berasingan terhadap empat belas (14) subkonstruk dalam instrumen ini.

Merujuk Jadual 3, semua subkonstruk mencatatkan nilai peratusan varians dijelaskan melebihi 40% dan nilai *eigen* kontras pertama di bawah 3.0, seperti disarankan oleh Linacre (2022), menandakan bahawa struktur setiap subkonstruk adalah bersifat unidimensi. Subkonstruk seperti berwawasan (84.3%, *eigen* = 1.8), pembelajaran dan pemudahcaraan (88.7%, *eigen* = 2.0), dan instruksional (76.0%, *eigen* = 1.6) menunjukkan struktur unidimensi yang kukuh. Walaupun subkonstruk komitmen berterusan (78.0%, *eigen* = 2.8) dan komitmen normatif (42.0%, *eigen* = 2.5) hampir menghampiri nilai ambang, ia masih berada dalam julat diterima dan menyokong dimensi tunggal secara minimum.

Dapatan ini sejajar dengan kajian oleh Bond & Fox (2015) yang menekankan bahawa model Rasch menuntut unidimensionaliti sebagai prinsip teras dalam pengukuran, serta kajian Linacre (2022) yang menunjukkan bahawa nilai *eigen* di bawah 3.0 menandakan tiada gangguan dimensi sekunder yang ketara. Penemuan ini turut disokong oleh Boone et al. (2014) yang menyatakan bahawa struktur konstruk yang tersusun rapi dengan satu dimensi meningkatkan ketepatan dan keabsahan keputusan pentaksiran psikometrik. Secara keseluruhan, dapatan ini mengesahkan kesahan struktur dalaman skala dan membuktikan bahawa semua subkonstruk dalam instrumen ini mampu mengukur konstruk sasaran secara tersendiri dan konsisten.

Jadual 3. Unidimensionaliti

Subkonstruk	Varians Dijelaskan (%)	<i>Eigenvalue</i> Kontras Pertama	Unidimensionaliti Disokong
Berwawasan	84.3	1.8	Ya
Instruksional	76.0	1.6	Ya
Sumber Operasi dan Iklim	70.9	1.7	Ya
Kerjasama Profesional	80.7	2.0	Ya
Mewujudkan inisiatif 'quick wins'	80.2	1.6	Ya
Komitmen Afektif	78.0	1.8	Ya
Komitmen Berterusan	78.0	2.8	Ya (hampir tinggi)
Komitmen Normatif	42.0	2.5	Ya (minimum diterima)
Pengetahuan	77.2	1.5	Ya
Kemahiran	65.7	1.3	Ya
Motif	65.7	1.7	Ya
Trait	63.3	1.9	Ya
Pembelajaran dan pemudahcaraan	88.7	2.0	Ya
Tanggungjawab terhadap profession	63.2	1.9	Ya

Kebolehpercayaan dan Indeks Pengasingan

Kebolehpercayaan dan indeks pengasingan memainkan peranan penting dalam menentukan ketekalan serta keupayaan instrumen untuk membezakan tahap keupayaan responden. Merujuk kepada Jadual 3, instrumen ini mencatatkan nilai kebolehpercayaan responden sebanyak 0.98 dan indeks pengasingan responden 6.28, jauh melebihi nilai ambang minimum ≥ 0.80 dan ≥ 1.5 seperti yang disarankan oleh Wright & Masters (1982). Ini membuktikan bahawa instrumen berkeupayaan tinggi untuk membezakan responden kepada sekurang-kurangnya enam kumpulan keupayaan berbeza secara konsisten, sejajar dengan penemuan oleh Bond & Fox (2015).

Bagi item, kebolehpercayaan dicatatkan pada 0.84, manakala indeks pengasingan item ialah 2.28, menandakan bahawa hierarki kesukaran item adalah stabil dan boleh diguna semula dalam konteks populasi guru yang serupa (Linacre, 2022). Selain itu, pekali Cronbach's Alpha ialah 0.99 yang melebihi tahap ≥ 0.90

yang lazimnya ditafsirkan sebagai sangat tinggi dan menunjukkan konsistensi dalaman yang stabil dalam kalangan item (Tavakol & Dennick, 2011; Pallant, 2020). Akhir sekali, nilai korelasi antara skor mentah responden dan ukuran logit (0.96) mengesahkan kewujudan hubungan yang sangat kukuh antara respon individu dan konstruk yang diukur. Kesemua dapatan ini membuktikan bahawa instrumen yang dibangunkan bukan sahaja sah dari sudut konstruk, malah stabil dan diskriminatif, serta sesuai digunakan dalam kajian berskala besar dan aplikasi profesionalisme guru di sekolah rendah.

Jadual 4: Indeks kebolehpercayaan dan pengasingan instrumen

Aspek	Nilai	Kriteria Piawai	Interpretasi
Kebolehpercayaan responden	0.98	≥ 0.80	Ketekalan tinggi antara responden
Indeks pengasingan responden	6.28	≥ 1.5 (min)	Keupayaan tinggi membezakan kumpulan
Kebolehpercayaan item	0.84	≥ 0.80	Stabil dan boleh diguna semula
Indeks pengasingan item	2.28	≥ 2.0 (baik)	Susunan kesukaran item adalah stabil
Cronbach's Alpha	0.99	≥ 0.90 (sangat baik)	Konsistensi dalaman sangat tinggi
Person score-to-measure correlation	0.96	≥ 0.70	Hubungan kuat antara skor dan konstruk

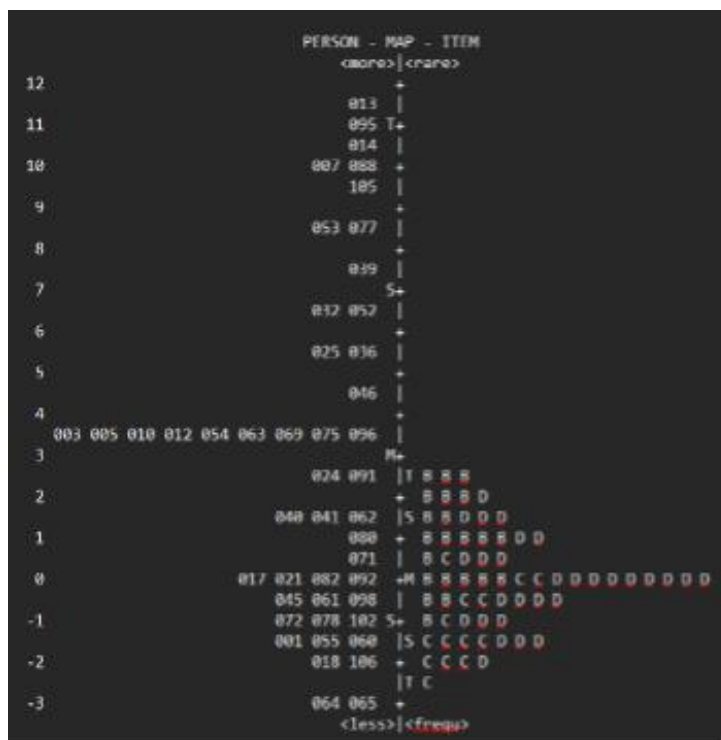
Sumber: Bond & Fox, 2015; Linacre, 2022; Wright & Masters (1982); Tavakol & Dennick (2011), Pallant (2020)

Penyebaran Ukuran dan Pemetaan Wright

Pemetaan Wright (Wright Map) memberikan gambaran visual yang penting mengenai sejauh mana instrumen ini berupaya menyasarkan tahap kebolehan responden dengan tepat melalui penyebaran tahap kesukaran item dalam skala logit yang sama. Dalam Rajah 1, bahagian kiri menunjukkan taburan keupayaan responden, manakala bahagian kanan menunjukkan kesukaran item. Analisis menunjukkan bahawa julat kesukaran item berada antara -2.64 hingga $+2.54$ logit, manakala kebolehan responden meliputi hingga $+12$ logit, mencerminkan jurang penyasaran (*mistargeting*) di mana sebilangan besar responden menunjukkan tahap keupayaan melebihi julat kesukaran item yang disediakan. Taburan item tertumpu di sekitar nilai logit 0, iaitu tahap kesukaran sederhana, manakala sebahagian besar responden berada dalam julat positif yang lebih tinggi. Ciri ini menunjukkan bahawa walaupun tiada pengelompokan ekstrem atau jurang besar dalam taburan item (yang menandakan keutuhan struktur instrumen), terdapat keperluan untuk penambahbaikan ke atas penyasaran. Menurut Bond dan Fox (2015), pengukuran Rasch yang efektif memerlukan pertindihan yang mencukupi antara kebolehan responden dan kesukaran item bagi membolehkan pengukuran dilakukan secara adil dan diskriminatif terhadap keseluruhan spektrum keupayaan.

Keupayaan responden yang tinggi berbanding kesukaran item juga mengesahkan *mistargeting*, iaitu instrumen tidak mencukupi dalam menilai guru-guru yang memiliki tahap kompetensi tinggi. Dalam konteks ini, cadangan penambahbaikan yang dicadangkan ialah pengayaan set item dengan menambah item-item yang lebih sukar, khususnya berkaitan kecekapan mendalam dalam pelaksanaan kajian tindakan, kepimpinan berimpak tinggi, atau inovasi reflektif dalam amalan pedagogi. Dapatan ini selari dengan dapatan kajian Talib et al. (2024) yang menekankan kepentingan instrumen yang dapat mengenal pasti guru pelbagai tahap kesiapsiagaan dan komitmen dalam pelaksanaan dasar seperti *STEM* dan kajian tindakan. Ambon et al. (2024) turut menyatakan bahawa penyasaran instrumen yang seimbang adalah asas penting dalam menilai impak sebenar intervensi profesionalisme guru.

Secara keseluruhan, walaupun penyebaran item adalah baik bagi mengukur kompetensi sederhana, penyesuaian tambahan diperlukan bagi meningkatkan keupayaan instrumen menilai guru cemerlang dan berpengalaman secara lebih berstrata. Penambahbaikan ini akan meningkatkan diskriminasi instrumen serta menjadikannya lebih relevan untuk pentaksiran diagnostik, pelaksanaan formatif, dan pembangunan profesional bersasar di peringkat sistem pendidikan rendah.



Rajah 1. Pemetaan Wright (Wright Map)

Secara keseluruhan, dapatan kajian rintis ini mengesahkan bahawa instrumen yang dibangunkan adalah sah dan boleh dipercayai dari aspek kandungan, muka, dan konstruk. Analisis Rasch menunjukkan bahawa skala berfungsi dengan baik, item berada dalam julat kesesuaian yang diterima, struktur instrumen adalah unidimensi, serta mempunyai daya diskriminatif yang tinggi. Nilai kebolehpercayaan yang dicatatkan juga amat konsisten, menandakan instrumen ini berpotensi digunakan dalam penilaian berskala besar. Walau bagaimanapun, nilai kebolehpercayaan yang sangat tinggi dan dapatan pemetaan Wright menunjukkan kemungkinan wujudnya *mistargeting*, iaitu taburan item yang lebih terkumpul pada tahap kesukaran sederhana sedangkan sebahagian besar responden berada pada tahap kebolehan tinggi. Hal ini membayangkan keperluan kepada pengayaan item yang lebih sukar bagi memastikan semua tahap keupayaan guru dapat diukur secara adil dan menyeluruh. Dapatan ini selari dengan saranan Youngs & Printy (2022), Makhsin et al. (2022), dan Rostini et al. (2022) yang menegaskan pentingnya instrumen reflektif berasaskan data dalam menyokong pembangunan profesional guru secara sistematik.

Di samping itu, batas kebolehgunaan instrumen dalam kajian rintis ini perlu diperakui, memandangkan saiz sampel analitik yang kecil ($n=47$) serta proses penapisan respons yang agak ketat, yang berpotensi mempengaruhi generalisasi dapatan. Justeru, pada fasa kajian utama, beberapa pelan penambahbaikan telah dirancang. Pertama, Analisis Pengesahan Faktor (CFA) akan dijalankan menggunakan perisian SmartPLS untuk mengesahkan struktur faktor konstruk aras pertama dan aras kedua (*higher order constructs*). Pelaporan CFA akan merangkumi indikator psikometrik penting termasuk muatan faktor, *Average Variance Extracted* (AVE), *Composite Reliability* (CR), *Heterotrait-Monotrait Ratio* (HTMT), serta indeks kesepadanan model seperti *Standardised Root Mean Square Residual* (SRMR). Analisis ini dijangka dapat mengukuhkan kesahan konvergen dan diskriminan, serta memberikan sokongan empirikal terhadap model konseptual yang dicadangkan.

Kedua, bagi mengatasi isu *mistargeting*, item baharu beraras sukar akan dibangunkan secara khusus berdasarkan hasil dapatan pemetaan Wright. Proses pembinaan item ini akan menggunakan pendekatan triangulasi bersama guru berpengalaman dan pakar bidang, bagi memastikan ia benar-benar mewakili cabaran sebenar dalam amalan kajian tindakan dan pembangunan profesional guru. Penambahan item ini diharap dapat meningkatkan keupayaan diskriminatif instrumen dalam menilai guru pada tahap kompetensi tinggi, sekali gus menjadikan instrumen lebih granular, stabil, relevan dalam konteks pendidikan rendah di Malaysia.

Kesimpulan

Kajian ini telah berjaya membangunkan dan menilai kesahan serta kebolehpercayaan satu instrumen yang mengukur pembudayaan kajian tindakan oleh guru besar, komitmen guru terhadap kajian tindakan, dan kualiti guru sekolah rendah di Malaysia. Instrumen yang dibangunkan menepati piawaian psikometrik yang tinggi, dengan bukti kesahan kandungan, kesahan muka dan kesahan konstruk yang kukuh. Analisis Model Rasch menunjukkan kesesuaian item yang memuaskan, struktur yang unidimensi, serta daya diskriminatif yang baik merentas tahap keupayaan guru. Kebolehpercayaan dalaman yang tinggi (*Cronbach's Alpha* = 0.99) dan nilai pengasingan yang stabil turut menyokong potensi instrumen ini untuk digunakan dalam kajian berskala besar atau penilaian impak dasar pendidikan. Sumbangan utama kajian ini adalah dalam menyediakan alat ukur yang sah, boleh dipercayai, serta berpaksikan teori serta konteks pendidikan Malaysia. Instrumen ini dapat dimanfaatkan oleh penyelidik, pembuat dasar, dan pentadbir sekolah untuk menilai tahap pembudayaan kajian tindakan, komitmen guru, serta kualiti profesional guru secara lebih sistematik dan berasaskan data. Secara tidak langsung, ia menyokong pelaksanaan inisiatif seperti PLC, PPPB dan program pembangunan kepimpinan instruksional yang berteraskan refleksi dan amalan berasaskan bukti.

Walau bagaimanapun, kajian ini tertakluk kepada beberapa batasan. Pertama, saiz sampel akhir yang kecil ($n = 47$) selepas proses pembuangan 60 respons tidak sesuai untuk generalisasi dapatan. Kedua, isu targeting dalam *Wright Map* menunjukkan bahawa sebahagian besar responden berada pada aras keupayaan lebih tinggi berbanding tahap kesukaran item, menandakan perlunya pengayaan instrumen dengan item yang lebih sukar. Faktor-faktor ini berpotensi mempengaruhi kebolehgunaan instrumen dalam konteks yang lebih luas dan pelbagai. Bagi kajian utama, beberapa langkah penambahbaikan digariskan. Pertama, rangka sampel akan diperluaskan melibatkan lebih banyak negeri dan jenis lokasi sekolah bagi memastikan representasi yang lebih menyeluruh. Kedua, saiz sampel yang lebih besar akan dipertimbangkan untuk memastikan kuasa statistik mencukupi bagi analisis pengesanan. Ketiga, Analisis Pengesanan Faktor (CFA) akan dijalankan bagi mengesahkan struktur faktor instrumen. Selain itu, instrumen akan diperkaya dengan penambahan item berkesukaran tinggi, dibangunkan berdasarkan dapatan *Wright Map* dan triangulasi bersama guru berpengalaman serta pakar bidang, bagi meningkatkan keupayaan instrumen dalam menilai guru pada aras kompetensi yang lebih tinggi. Secara keseluruhannya, instrumen ini merupakan satu sumbangan penting ke arah pemantapan amalan penyelidikan pendidikan dan pembangunan profesionalisme guru secara lestari dan berimpak tinggi.

Penghargaan: Kami merakamkan setinggi-tinggi penghargaan kepada Kementerian Pendidikan Malaysia atas sokongan kewangan yang diberikan (rujukan surat:KPM.BT.700-30/23/12(6)). Kami juga mengucapkan setinggi-tinggi terima kasih kepada semua pihak dan individu yang terlibat secara langsung atau tidak langsung dalam menjayakan kajian ini.

Penyataan Persetujuan Dimaklumkan: Persetujuan dimaklumkan telah diperoleh daripada peserta kajian yang terlibat dalam kajian ini.

Konflik Kepentingan: Tiada sebarang konflik kepentingan.

Rujukan

- Adhikari, N. P. (2023). Teachers' problems and concerns in teacher education. *Interdisciplinary Research in Education*, 8(1), 125-135. <https://doi.org/10.3126/ire.v8i1.56732>
- Ag-Ahmad, N., Mohamed, A. T. F. S., & Bakar, E. W. (2023). Systematic review on issues and challenges of pre-service English teachers in Malaysia. *Journal a Publication on the Teaching and Learning of English*, 34(1), 1-20. <https://doi.org/10.15639/teflinjournal.v34i1/1-20>
- Ambon, J., Alias, B. S., & Mansor, A. N. (2024). Hubungan amalan pembangunan profesionalisme berterusan guru dengan kualiti PdPc guru di Sarawak. *e-Bangi: Journal of Social Sciences & Humanities*, 21(4), 1-15. <https://doi.org/10.17576/ebangi.2024.2104.07>

- Aulls, M. W., & Shore, B. M. (2023). *Inquiry in education, Volume I: The conceptual foundations for research as a curricular imperative*. Routledge.
- Bahagian Profesionalisme Guru. (2023). *Dokumen Kompetensi Guru Bidang STEM*. Kementerian Pendidikan Malaysia.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.
- Chukwuma, N., & Zondo, R. W. D. (2025). Impact of Climate Change on Teaching and Learning: A Literature Review Perspective of Higher Education in Africa. *e-Bangi: Journal of Social Sciences & Humanities*, 22(2), 45-59. <https://doi.org/10.17576/ebangi.2025.2202.45>
- El-Den, S., Schneider, C. R., Mirzaei, A., & Carter, S. R. (2020). How to measure a latent construct: Psychometric principles for the development and validation of measurement instruments. *International Journal of Pharmacy Practice*, 28(4), 326–336. <https://doi.org/10.1111/IJPP.12600>
- Fullan, M. 2015. *The new meaning of educational change* (5th ed.). Teachers College Press.
- Latip, N. A., & Abd Latif, S. (2024). Strategi pemimpin sekolah dalam menangani isu ketidakupayaan guru melaksanakan inovasi pengajaran di sekolah menengah. *e-Bangi: Journal of Social Sciences & Humanities*, 21(4), 122–134.
- Linacre, J. M. (2022). *Winsteps® Rasch measurement computer program user's guide* (Version 5.2.5). Winsteps.com.
- Makhsin, M., Teoh, Y., & Ismail, N. (2022). Malaysian Teacher Standard 2.0 in becoming the soul of educators. *International Journal of Education, Psychology and Counseling*, 7(46), 26-39. <https://doi.org/10.35631/ijepc.746026>.
- Moktar, A., Alias, B. S., & Rasul, M. S. (2024). Exploring factors in teachers' action research: A case study analysis, *Educational Administration: Theory And Practice*, 30(4), 10106-10116. <https://doi.org/10.53555/kuey.v30i4.6177>
- Mustapha, N. A., & Rahman, A. F. A. (2024). Amalan Penyelidikan dalam Kalangan Guru: Cabaran dan Strategi Pengukuhan di Sekolah Rendah. *e-Bangi: Journal of Social Sciences & Humanities*, 21(4), 89–101.
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). Routledge.
- Rostini, D., Syam, R.Z.A., & Achmad, W. (2022). The significance of principal management on teacher performance and quality of learning. *Al-Ishlah: Jurnal Pendidikan* 14(2), 2513-2520. <https://doi.org/10.35445/alishlah.v14i2.1721>
- Talib, S., Alias, B. S., & Matore, M. E. M. (2024). Principal competency, teacher readiness, and teacher commitment in STEM education implementation: A qualitative study. *e-Bangi: Journal of Social Sciences & Humanities*, 21(4), Article 30. <https://doi.org/10.17576/ebangi.2024.2104.30>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tifanny, R. S. N., & Roie, M. U. (2024). Cultivating action research culture among Teachers. *International Journal of Multidisciplinary Research and Analysis*, 7(7), 3179-3194. <https://doi.org/10.47191/ijmra/v7-i07-15>
- Widayati, A., MacCallum, J., & Woods-McConney, A. (2021). Teachers' perceptions of continuing professional development: A study of vocational high school teachers in Indonesia. *Teacher Development*, 25(5), 604-621. <https://doi.org/10.1080/13664530.2021.1933159>
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.
- Youngs, P., & Printy, S. M. (2022). Principal leadership for instructional improvement: What does research say? *Journal of Educational Administration*, 60(2), 161-186. <https://doi.org/10.1108/JEA-03-2021-0042>