

# Kesahan dan Kebolehpercayaan Instrumen Kepimpinan Beretika dan Amalan Pengurusan Konflik terhadap Budaya Kerja Guru Menggunakan Model Pengukuran Rasch

Zubir Abdullah, Bity Salwana Alias and Aida Hanim A.Hamid

**Abstract** – Budaya kerja guru yang positif merupakan pemacu utama keberkesanan sekolah dan pencapaian murid, namun isu konflik dalaman, beban tugas, dan motivasi guru terus dilaporkan dalam konteks pendidikan di Malaysia. Kajian ini dijalankan untuk mengesahkan instrumen yang diadaptasi dan diubahsuai mengikut konteks kajian bagi mengukur konstruk kepimpinan beretika, amalan pengurusan konflik, dan budaya kerja guru sebagai usaha memperkukuh asas penyelidikan berasaskan bukti dalam bidang kepimpinan sekolah. Menggunakan *Rasch Measurement Model* melalui perisian Winsteps versi 3.73, sifat psikometrik instrumen dianalisis secara komprehensif melalui kajian rintis melibatkan 133 guru dari tiga buah sekolah menengah kebangsaan di Melaka. Dapatan menunjukkan kebolehpercayaan individu (person reliability) yang sangat tinggi (0.96) dan indeks pemisahan individu (person separation index) (5.23), sekali gus membuktikan keupayaan instrumen membezakan tahap responden dengan ketepatan yang baik. Kebolehpercayaan Item (Item Reliability) (0.84) dan indeks pemisahan item (item separation index) (2.27) turut menyokong kestabilan hierarki item. Statistik Infit dan Outfit MNSQ mengesahkan kesesuaian item terhadap model pengukuran, manakala korelasi titik–ukuran (Point-Measure Correlations, PTMEA Corr) dan analisis komponen utama terhadap sisa (Principal Component Analysis, PCA of Residuals) mengesahkan kesahan konstruk serta unidimensionaliti instrumen. Keseluruhannya, dapatan kajian mengesahkan bahawa instrumen ini sah, boleh dipercayai dan bersedia untuk digunakan dalam kajian berskala besar. Dapatan ini memberi implikasi penting kepada pembangunan instrumen tempatan yang mantap bagi menilai kepimpinan beretika dan amalan pengurusan konflik, seiring menyokong agenda pendidikan transformasi dan peningkatan kualiti budaya kerja guru di sekolah.

**Keywords** – kepimpinan beretika, amalan pengurusan konflik, budaya kerja guru, kesahan, kebolehpercayaan, *Rasch Measurement Model*

## I. PENGENALAN

Budaya kerja guru memainkan peranan penting dalam membentuk kualiti pendidikan di sekolah, khususnya dalam konteks sekolah menengah di Malaysia. Ia mencerminkan nilai bersama, norma profesional dan jangkaan tingkah laku yang mempengaruhi amalan pengajaran, kolaborasi rakan sejawat serta iklim organisasi secara keseluruhan (Fullan, 2007; Schein, 2010). Selaras dengan *Malaysia Education Blueprint (MEB) 2013–2025*, pembudayaan kerja yang positif dalam kalangan guru

merupakan prasyarat peningkatan hasil pembelajaran murid dan keberkesanan kepimpinan sekolah (Kementerian Pendidikan Malaysia, 2013). Deal dan Peterson (2009) menegaskan bahawa budaya sekolah berperanan secara langsung dalam membentuk motivasi, komitmen serta kecekapan kolektif guru yang menjadi asas kepada pemantapan amalan pendidikan berkualiti. Budaya sekolah yang positif, menyokong dan beretika bukan sahaja menggalakkan sikap profesional yang proaktif dan disiplin dalam kalangan guru, malah turut memperkukuh kualiti kerja serta keberkesanan komunikasi sesama warga sekolah. Dalam jangka panjang, persekitaran budaya sekolah yang kondusif ini menyumbang kepada kesejahteraan emosi, kepuasan profesional, dan pembangunan berterusan guru sebagai tenaga penggerak utama pendidikan (Tschannen-Moran & Hoy, 2001; Bryk & Schneider, 2002). Walaupun kepentingan budaya kerja guru telah diiktiraf, penyelidikan empirikal mengenai instrumen pengukuran yang sah secara kontekstual masih terhad dalam landskap pendidikan Malaysia. Banyak kajian terdahulu kurang menimbang pengaruh kepimpinan beretika dan amalan pengurusan konflik di mana dua unsur kritikal yang membentuk kelakuan profesional dalam persekitaran sekolah yang pelbagai budaya (Salleh et al., 2019; Khalid et al., 2016). Kepimpinan beretika yang menekankan integriti, keadilan dan akauntabiliti terbukti mempengaruhi tingkah laku guru, kepercayaan dan iklim sekolah (Brown & Treviño, 2006; Langlois et al., 2014), manakala pengurusan konflik yang berkesan mengukuhkan kolaborasi profesional dan mengurangkan tekanan organisasi (Rahim, 2011; Ting-Toomey, 2018).

Bagi menangani jurang tersebut, kajian ini menggunakan pendekatan kuantitatif untuk mengesahkan instrumen yang diadaptasi mengikut konteks kajian merangkumi konstruk kepimpinan beretika, amalan pengurusan konflik dan budaya kerja guru. Instrumen ini telah diuji rintis melibatkan guru di tiga buah sekolah menengah harian di negeri Melaka selepas memperoleh kelulusan rasmi penyelidikan daripada Bahagian Perancangan dan Penyelidikan Dasar Pendidikan (EPRD), Jabatan Pendidikan Negeri dan pihak pentadbir sekolah. Bagi memastikan kekukuhan psikometrik, Model Pengukuran Rasch digunakan melalui perisian Winsteps versi 3.73. Model Rasch menyediakan kerangka mantap untuk menilai kebolehpercayaan dan kesahan pengukuran melalui parameter seperti kebolehpercayaan individu dan item, indeks pengasingan item, serta statistik kesepadanan (fit statistics) (Bond & Fox, 2015; Boone et al., 2014). Ia membolehkan transformasi data ordinal kepada ukuran aras sela dan menawarkan diagnostik seperti *Infit* dan *Outfit Mean Square* (MNSQ), Korelasi Titik-Ukuran

Zubir Abdullah, Universiti Kebangsaan Malaysia, Malaysia (Email address: P133503@siswa.ukm.edu.my).  
Bity Salwana Alias, Universiti Kebangsaan Malaysia, Malaysia (Email address: bity@ukm.edu.my).  
Aida Hanim A. Hamid, Universiti Kebangsaan Malaysia, Malaysia (Email address: pgh-ikom@ukm.edu.my).

(PTMEA Corr) dan Analisis Komponen Prinsipal (PCA) sisa residual bagi menilai kesahan kandungan dan konstruk (Wright & Linacre, 1994; Linacre, 2021).

Dengan aplikasi model ini, kajian memastikan kecukupan psikometrik instrumen untuk aplikasi yang lebih meluas dalam menilai hubungkait antara kepimpinan beretika, pengurusan konflik dan budaya kerja guru dalam konteks sekolah menengah Malaysia. Dapatan kajian rintis menyediakan bukti awal yang penting untuk menambahbaik seterusnya sebelum proses pengesahan dan kajian berskala besar.

## II. PERNYATAAN MASALAH

Budaya kerja guru yang positif ialah pemacu kualiti pengajaran, kolaborasi profesional dan kesejahteraan emosi warga sekolah. Dalam konteks sekolah menengah Malaysia, isyarat dasar semasa menunjukkan keperluan mendesak memperkukuh budaya organisasi dan kesejahteraan warga sebagai asas peningkatan hasil pembelajaran murid serta keberkesanan kepimpinan sekolah (Kementerian Pendidikan Malaysia, 2024). Kedua-dua dokumen ini menekankan keperluan ukuran yang sah dan boleh dipercayai bagi konstruk berkaitan budaya kerja guru untuk memaklumkan intervensi berasaskan bukti di peringkat sekolah, PPD dan JPN.

Seiring itu, penyelidikan semasa mendapati beban kerja dan tekanan kerjaya guru semakin signifikan, dengan implikasi terhadap komitmen, kerjasama profesional dan iklim organisasi sekolah iaitu faktor yang melemahkan budaya kerja kolaboratif jika tidak ditangani secara sistematik (Arba'in et al., 2024; Saad et al., 2025). Dapatan ini mengisyaratkan keperluan pengukuran yang tepat konteks untuk mengenal pasti domain budaya kerja yang perlu diintervensi dan ditambah baik melalui program kepimpinan serta sokongan organisasi.

Dari sudut pemboleh ubah penentu, dua komponen kritikal yang konsisten ditonjolkan dalam literatur terkini ialah kepimpinan beretika pengetua dan amalan pengurusan konflik. Meta-analisis terkini merumuskan bahawa kepimpinan beretika pengetua berkaitan positif dengan komitmen, kepercayaan, dan kepuasan kerja guru di mana menyokong hujah bahawa isyarat etika pemimpin membentuk norma dan budaya kerja (Özdoğan & Sarier, 2024). Selari dengan itu, kajian empirikal menunjukkan gaya pengurusan konflik yang integratif dan kooperatif berkait positif dengan prestasi kerja guru, justeru memperkukuh iklim profesional yang kondusif (Farooqi et al., 2023). Walau bagaimanapun, sebahagian besar kajian mengguna instrumen adaptasi tanpa pengesahan psikometrik yang memadai bagi kesesuaian budaya dan bahasa tempatan, membataskan kebolehbandingan dan kebolegunaan dapatan untuk amalan di sekolah menengah Malaysia.

Jurang utama yang perlu ditangani ialah kekurangan instrumen yang disahkan secara kontekstual untuk mengukur kepimpinan beretika, amalan pengurusan konflik dan budaya kerja guru. Bagi menutup jurang ini, kajian rintis ini memfokuskan pengesahan instrumen yang diadaptasi menggunakan Model Pengukuran Rasch dengan perisian Winsteps versi 3.73 untuk menilai

kesahan dan kebolehpercayaan pada aras item dan responden (kebolehpercayaan individu/item, indeks pengasingan, statistik kesepadanan Infit/Outfit, PTMEA Corr, dan PCA sisa). Pendekatan Rasch diperlukan untuk memastikan unidimensionaliti dan kestabilan ukuran sebelum aplikasi pengesahan skala penuh, seterusnya menyumbang alat ukur tempatan yang mantap untuk penyelidikan dan penambahbaikan kepimpinan pendidikan, selaras dengan hala tuju dasar semasa (Bond & Fox, 2015; Linacre, 2021; KPM, 2024).

## III. SOROTAN LITERATUR

Budaya kerja guru menjadi asas utama kepada keberkesanan sekolah kerana ia mempengaruhi kualiti pengajaran, kolaborasi profesional dan kesejahteraan emosi warga pendidik. Menurut Schein (2010), budaya organisasi merangkumi nilai bersama, norma dan andaian asas yang membentuk makna serta tindakan warga sekolah. Dalam konteks pendidikan, budaya kerja yang positif dikaitkan dengan peningkatan motivasi dan kecekapan kolektif guru (Deal & Peterson, 2009; Bryk & Schneider, 2004). Di Malaysia, Pelan Strategik Kementerian Pendidikan Malaysia 2024–2030 menegaskan keperluan memperkukuh budaya organisasi berasaskan integriti, profesionalisme dan kesejahteraan guru sebagai tonggak transformasi pendidikan (KPM, 2024). Namun, kajian semasa menunjukkan bahawa tekanan kerja dan beban tugas yang tinggi masih menjadi cabaran kepada pengukuhan budaya kerja kolaboratif dalam kalangan guru sekolah menengah (Arba'in et al., 2024; Saad et al., 2025). Hal ini menuntut penyelidikan empirikal yang menilai konstruk budaya kerja secara sah dan kontekstual bagi menyokong dasar pendidikan negara.

Kepimpinan beretika dikenal pasti sebagai faktor utama yang membentuk budaya kerja dan tingkah laku profesional warga sekolah. Brown dan Treviño (2006) mendefinisikan kepimpinan beretika sebagai demonstrasi tingkah laku normatif seperti keadilan, kejujuran dan tanggungjawab, serta pengaruh pemimpin terhadap pengikut melalui pemodelan nilai moral. Model *Ethical Leadership at Work (ELW)* oleh Kalshoven et al. (2011) memperincikan tujuh dimensi utama termasuk keadilan, integriti, orientasi manusia, keperihatinan terhadap kemampuan, etika bimbingan dan perkongsian kuasa yang terbukti meningkatkan kepercayaan serta kesetiaan guru (Özdoğan & Sarier, 2024). Dalam konteks sekolah, pemimpin beretika bukan sahaja menjadi teladan moral tetapi juga pemangkin amalan pengurusan konflik yang adil dan berasaskan komunikasi terbuka. Kajian oleh Farooqi et al. (2023) menunjukkan bahawa gaya pengurusan konflik pengetua yang kolaboratif meningkatkan kepuasan kerja dan kecekapan profesional guru, justeru mengukuhkan iklim sekolah yang positif dan saling mempercayai.

Walaupun bagaimanapun, banyak kajian tempatan menggunakan instrumen adaptasi tanpa pengesahan psikometrik yang menyeluruh terhadap kesesuaian budaya dan bahasa tempatan, menyebabkan dapatan sukar digeneralisasikan kepada konteks sekolah menengah Malaysia (Boone et al., 2014; Awang, 2018). Oleh itu,

Model Pengukuran Rasch dipilih dalam kajian ini bagi memastikan kesahan dan kebolehpercayaan instrumen diukur secara saintifik dan tepat. Model ini menilai kebolehpercayaan responden dan kebolehpercayaan item, indeks pengasingan serta kesepadanan item melalui statistik Infit dan Outfit, selain menguji unidimensionaliti menggunakan analisis komponen utama terhadap sisa (Bond & Fox, 2015; Linacre, 2021). Pendekatan ini bukan sahaja memperkukuh ketepatan konstruk instrumen yang mengukur kepimpinan beretika, amalan pengurusan konflik dan budaya kerja guru, tetapi turut menyumbang kepada pembangunan alat ukur tempatan yang sah dan sejajar dengan realiti sistem pendidikan Malaysia

#### IV. METODOLOGI

Kajian rintis ini menggunakan reka bentuk tinjauan yang memfokuskan kepada pengesahan sifat psikometrik bagi instrumen adaptasi yang mengukur kepimpinan beretika, amalan pengurusan konflik dan budaya kerja guru. Teknik persampelan bertujuan digunakan selaras dengan matlamat kajian untuk memastikan kepercayaan dan kesesuaian data yang diperoleh daripada responden sasaran (Palinkas et al., 2015). Pendekatan ini diiktiraf sebagai sesuai untuk kajian rintis, khususnya apabila kriteria pemilihan responden sejajar dengan rangka konseptual kajian (Etikan, Musa, & Alkassim, 2016). Peserta kajian yang telah terlibat terdiri daripada 133 orang guru sekolah menengah dari tiga buah Sekolah Menengah Kebangsaan (SMK) harian di negeri Melaka, Malaysia. Pemilihan sekolah dibuat berdasarkan pertimbangan praktikal seperti kebolehcapaian dan kebenaran institusi. Kriteria pemilihan responden ialah guru yang sedang berkhidmat sepenuh masa dan mempunyai sekurang-kurangnya tiga tahun pengalaman mengajar, sejajar dengan pandangan penyelidikan lepas bahawa kematangan dalam perkhidmatan penting bagi memastikan persepsi yang stabil terhadap kepimpinan dan budaya organisasi sekolah (Yuzainee et al., 2019).

Sebelum proses pengumpulan data dijalankan, kebenaran rasmi penyelidikan diperoleh melalui Sistem Permohonan Penyelidikan Pendidikan (eRas) di bawah Kementerian Pendidikan Malaysia (KPM). Kebenaran tambahan turut diperoleh daripada Jabatan Pendidikan Negeri Melaka (JPN Melaka), pengetua sekolah terlibat serta pihak pentadbiran sekolah masing-masing, sejajar dengan prosedur etika penyelidikan KPM. Soal selidik diedarkan secara bersemuka ke sekolah mengikut jadual yang ditetapkan, dengan kerjasama pentadbir sekolah bagi memastikan proses pengumpulan data dijalankan secara teratur dan menepati masa. Kaedah pengedaran bersemuka ini dipilih bagi memaksimumkan kadar respons dan meningkatkan ketepatan data, sebagaimana disarankan oleh Groves et al. (2009), khususnya dalam konteks penyelidikan berasaskan sekolah.

Bagi menentukan saiz sampel yang mencukupi, kajian ini merujuk kepada garis panduan kecukupan sampel dan kestabilan penentukuran item menggunakan Model Pengukuran Rasch (Linacre, 1994). Keperluan minimum bagi kestabilan penentukuran item pada tahap keyakinan 95% ialah antara 30 hingga 36 responden (Azrilah et al.,

2017). Kajian ini melangkaui keperluan tersebut dengan memperoleh 133 respons yang sah, sekali gus menjamin kebolehpercayaan data dan kesesuaian bagi analisis psikometrik seterusnya. Instrumen Struktur instrumen yang dibangunkan dan pembahagian item bagi setiap bahagian ditunjukkan dalam Jadual I.

**JADUAL I: PEMBAHAGIAN ITEM BAGI SETIAP BAHAGIAN INSTRUMEN PENYELIDIKAN**

Butiran	Kod	Bahagian	Bil.Item
A	A1 – A4	Profil Responden	-
B		Berorientasikan Manusia	5
	B1 - B5	Keadilan	5
	B6 - B10	Integriti	6
		Penjelasan peranan	5
	B11 - B16	Bimbingan etika	7
	B17 - B21	Perkongsian kuasa	4
		Keprihatinan terhadap kemampanan	6
	B22 - B28		
	B29 - B32		
	B33 - B38		
C	C1 – C5	Kolaborasi	5
	C6 – C10	Kompromi	5
	C11 – C15	Pengelakan	5
	C16 – C21	Menyesuai	6
	C22 – C26	Dominasi	5
D	D1 - D7	Sikap terhadap kerja	7
	D8 - D14	kerja	6
	D15 - D18	Disiplin kerja	4
	D19 - D23	Kualiti kerja	5
		Komunikasi kerja	
	Total		87

Kajian ini menggunakan borang soal selidik berstruktur yang mengandungi empat bahagian utama (Bahagian A hingga D), yang masing-masing dibangunkan untuk mengukur konstruk yang berbeza selaras dengan objektif kajian. Bahagian A mengumpul maklumat demografi responden seperti jantina, umur, tempoh pengalaman mengajar dan kelulusan akademik tertinggi. Maklumat ini bertujuan memberikan gambaran yang lebih jelas tentang kepelbagaian responden serta mengenal pasti kemungkinan pengaruh faktor pengantaraan. Bahagian B direka bentuk untuk menilai persepsi guru terhadap amalan kepimpinan beretika pengetua. Instrumen ini mengandungi 38 item yang merangkumi tujuh dimensi utama iaitu Berorientasikan manusia, Keadilan, Integriti, Penjelasan peranan, Bimbingan etika, Perkongsian kuasa, dan Keprihatinan terhadap kemampanan. Bahagian C menilai amalan pengurusan konflik pengetua dengan menggunakan 26 item yang mewakili lima dimensi utama: Kolaborasi, Kompromi, Pengelakan, Menyesuai, dan Dominasi. Bahagian D pula menumpukan kepada konstruk budaya kerja guru, terdiri daripada 23 item yang diedarkan dalam empat dimensi iaitu Sikap terhadap kerja, Disiplin kerja, Kualiti kerja, dan Komunikasi kerja.

Kesemua item dalam soal selidik ini dinilai menggunakan skala Likert lima mata bagi mengukur tahap persetujuan responden terhadap setiap pernyataan. Skala ini diukur seperti berikut:

- 1 = Sangat Tidak Setuju,
- 2 = Tidak Setuju,
- 3 = Kurang Setuju,

4 = Setuju, dan  
5 = Sangat Setuju.

Reka bentuk skala ini membolehkan pengumpulan data sikap yang lebih terperinci dan sesuai untuk penilaian psikometrik menggunakan Model Pengukuran Rasch dalam konteks penyelidikan pendidikan yang disesuaikan dengan budaya tempatan. Menurut Tsang (2012), kewujudan atau ketiadaan titik tengah dalam skala Likert tidak memberi kesan signifikan terhadap kebolehpercayaan atau kesahan instrumen, asalkan asas teorinya kukuh. Tambahan pula, Siti Hamidah et al. (2020) menegaskan bahawa penggunaan skala empat atau lima mata lebih sesuai secara budaya bagi konteks Malaysia dan Asia kerana dapat mengurangkan kecenderungan jawapan pertengahan serta meningkatkan kejelasan respons. Dalam konteks kajian berasaskan Rasch, Bond dan Fox (2015) serta Boone et al. (2014) mencadangkan penggunaan skala Likert empat hingga lima kategori kerana julat ini memberikan prestasi terbaik dari segi taburan respons, kesepadanan model dan penentuan item.

Analisis data kajian ini dijalankan menggunakan Model Pengukuran Rasch melalui perisian Winsteps versi 3.73 kerana keupayaannya yang mantap dalam menilai sifat psikometrik instrumen pengukuran pendidikan, khususnya dari aspek kesahan dan kebolehpercayaan. Bagi menilai kesahan instrumen, tiga petunjuk utama berasaskan Rasch digunakan. Pertama, statistik *Infit* dan *Outfit Mean Square* (MNSQ) yang menilai kesepadanan item dengan model pengukuran. Nilai dalam julat 0.5 hingga 1.5 logit dianggap memadai untuk menunjukkan bahawa item mempunyai kesepadanan yang baik serta menyokong kesahan kandungan instrumen (Wright & Linacre, 1994). Kedua, Korelasi Titik-Ukuran (Point Measure Correlation, PTMEA Corr) digunakan untuk menilai polariti item, di mana nilai positif menandakan setiap item sejajar dengan konstruk yang diukur. Ketiga, Analisis Komponen Utama terhadap sisa (Principal Component Analysis of Residuals, PCA) digunakan untuk menentukan unidimensionaliti, iaitu sejauh mana item dalam konstruk mengukur satu dimensi laten yang sama. Menurut Linacre (2005), unidimensionaliti disahkan apabila dimensi Rasch menerangkan sekurang-kurangnya 40% daripada varians keseluruhan, manakala nilai eigen bagi kontras pertama berada di bawah 2.0.

Penilaian kebolehpercayaan instrumen pula dilakukan melalui beberapa indeks Rasch utama. Antaranya ialah *Cronbach's Alpha* (KR-20) yang mengukur ketekalan dalaman serta kebolehpercayaan keseluruhan skala. Selain itu, indeks kebolehpercayaan individu dan item (Person and Item Reliability) digunakan untuk menilai kestabilan corak respons responden serta keupayaan instrumen menghasilkan ukuran yang boleh diulang (replicable). Dalam masa yang sama, indeks pengasingan individu dan item (Person and Item Separation Index) digunakan untuk menilai tahap daya diskriminasi instrumen terhadap perbezaan keupayaan responden dan kesukaran item. Nilai indeks pengasingan melebihi 2.0 disyorkan bagi memastikan pengelasan aras keupayaan dan kesukaran yang bermakna (Linacre, 1994; Fisher, 2007). Secara keseluruhan, nilai-nilai indeks ini

digunakan untuk menentukan kekuatan model dan kesepadanan data terhadap teori pengukuran Rasch.

Dalam kajian ini, kriteria dan julat nilai bagi penafsiran keputusan analisis Rasch diringkaskan dalam Jadual 2, yang merangkumi ambang penerimaan bagi setiap indeks. Di samping itu, kebolehpercayaan individu digunakan untuk menentukan sama ada responden disasarkan dengan tepat oleh item-item dalam instrumen, seterusnya memastikan konsistensi respons yang tinggi. Kebolehpercayaan item pula menunjukkan tahap ketekalan setiap item dalam mengukur konstruk laten yang dimaksudkan. Menurut Wright dan Masters (1982), nilai kebolehpercayaan item yang tinggi menandakan bahawa setiap item menyumbang secara konsisten terhadap pengukuran konstruk yang sama, sekali gus memperkukuh kesepaduan konstruk dan kestabilan instrumen bagi kegunaan kajian berskala besar pada peringkat seterusnya.

**JADUAL II: JULAT NILAI BAGI UJIAN KESAHAN BERDASARKAN ANALISIS RASCH**

Ujian / Indeks	Kriteria	Julat Nilai	Rujukan
Kebolehpercayaan	Kebolehpercayaan individu dan item	>5.0: Cemerlang	Tavakol & Dennick (2011)
		4.0 – 5.0: Sangat baik	
Indeks Pengasingan	Indeks pengasingan individu dan item	3.0 – 4.0: Baik	Linacre (1994); Fisher (2007)
		2.0 – 3.0: Sederhana	
Kesahan Kandungan	Statistik Infit/Outfit Mean Square (MNSQ) Z-Standard (Zstd)	<2.0: Lemah	Wright & Linacre (1994)
		Nilai >2.0 menunjukkan daya diskriminasi yang baik	
Kesahan Konstruk	Korelasi Titik-Ukuran (PTMEA Corr / Polariti) Analisis Komponen Prinsipal (PCA)	0.5 < MNSQ < 1.5	Bond & Fox (2015); Linacre (2005)
		-2.0 < Zstd < 2.0	
		0.20 < PTMEA Corr < 0.79	
		≥40% varians diterangkan; Kontras pertama < 2.0	

Jadual II menunjukkan julat nilai dan kriteria yang digunakan bagi menilai kesahan dan kebolehpercayaan instrumen berdasarkan Analisis Rasch. Dari segi kebolehpercayaan, indeks kebolehpercayaan person dan item dikategorikan kepada lima tahap: cemerlang (>5.0), sangat baik (4.0–5.0), baik (3.0–4.0), sederhana (2.0–3.0), dan lemah (<2.0) seperti yang disarankan oleh Tavakol dan Dennick (2011). Nilai yang tinggi mencerminkan kestabilan dan ketekalan responden serta item dalam mengukur konstruk yang dimaksudkan.

Indeks pengasingan individu dan item pula digunakan untuk menilai sejauh mana instrumen berupaya membezakan tahap keupayaan responden dan kesukaran item. Linacre (1994) dan Fisher (2007) menyatakan bahawa nilai indeks pengasingan melebihi 2.0 menunjukkan daya diskriminasi yang baik dan

memastikan klasifikasi keupayaan yang bermakna dalam kalangan responden.

Dari aspek kesahan kandungan, statistik *Infit* dan *Outfit Mean Square (MNSQ)* digunakan bagi menguji kesepadanan item terhadap model Rasch. Nilai dalam julat 0.5 hingga 1.5 logit dianggap sesuai dan menandakan kesepadanan item yang baik, manakala nilai *Z-Standard (Zstd)* antara -2.0 hingga +2.0 menggambarkan tahap keseragaman respons yang diterima (Wright & Linacre, 1994). Seterusnya, kesahan konstruk ditentukan melalui dua analisis utama: (i) Korelasi Titik-Ukuran (PTMEA Corr) yang menilai polariti item, di mana nilai antara 0.20 hingga 0.79 menunjukkan penjajaran yang baik antara item dengan konstruk yang diukur (Bond & Fox, 2015); dan (ii) Analisis Komponen Prinsipal terhadap sisa residual (PCA) yang mengesahkan unidimensionaliti, dengan sekurang-kurangnya 40% varians diterangkan oleh ukuran Rasch dan nilai kontras pertama kurang daripada 2.0, sebagaimana yang disarankan oleh Linacre (2005).

Secara keseluruhannya, kriteria dan julat nilai ini digunakan untuk memastikan bahawa instrumen yang dibangunkan memenuhi syarat kesahan dan kebolehpercayaan yang tinggi sebelum digunakan dalam kajian berskala penuh, sekali gus memperkukuh keabsahan ukuran konstruk kepimpinan beretika, amalan pengurusan konflik dan budaya kerja guru dalam konteks pendidikan Malaysia.

## V. DAPATAN

Bahagian ini membentangkan hasil analisis Rasch yang memfokuskan kepada indeks kebolehpercayaan dan pengasingan, serta penilaian kesahan kandungan dan kesahan konstruk bagi instrumen penyelidikan yang digunakan dalam kajian ini.

### *Kebolehpercayaan dan Indeks Pengasingan*

Berdasarkan Jadual III output analisis bagi 133 orang responden, indeks kebolehpercayaan person didapati bernilai 0.96 (real) dan 0.97 (model), menunjukkan tahap ketekalan dalaman yang sangat tinggi dalam mengukur persepsi individu terhadap konstruk yang dikaji. Menurut Tavakol dan Dennick (2011), nilai kebolehpercayaan melebihi 0.90 mencerminkan kestabilan ukuran yang cemerlang dan mengesahkan bahawa instrumen mampu menghasilkan keputusan yang konsisten merentas kumpulan responden yang berbeza. Tahap kebolehpercayaan person yang tinggi ini menandakan bahawa instrumen berupaya menangkap variasi keupayaan atau tahap persetujuan responden dengan ralat pengukuran yang minimum, sekali gus meningkatkan ketepatan interpretasi terhadap konstruk yang diukur.

Dari segi indeks pengasingan person, analisis menunjukkan nilai 5.23 (real) dan 5.54 (model), yang menandakan tahap pengasingan cemerlang berdasarkan kriteria yang ditetapkan oleh Fisher (2007). Nilai indeks pengasingan melebihi 5.0 menunjukkan bahawa instrumen ini berupaya membezakan sekurang-kurangnya lima strata keupayaan responden secara statistik. Dapatan ini memperlihatkan keupayaan diskriminatif yang kukuh

dalam membezakan responden dengan tahap persetujuan atau kecekapan yang berbeza terhadap konstruk yang diukur. Ciri ini amat penting dalam memastikan penilaian psikometrik yang mantap, khususnya dalam konteks penyelidikan pendidikan yang melibatkan konstruk kognitif dan afektif yang kompleks.

Selain itu, nilai Cronbach's Alpha (KR-20) bagi skor mentah person ialah 0.98, yang menunjukkan kebolehpercayaan dalaman yang luar biasa tinggi. Nilai ini mengesahkan bahawa item-item dalam soal selidik bersifat homogen dan mengukur konstruk asas yang sama secara konsisten, selaras dengan andaian unidimensionaliti dalam kerangka Rasch. Ketekalan dalaman yang tinggi ini turut mengurangkan kemungkinan berlakunya ralat rawak dan memperkukuh kebolehpercayaan inferens yang dibuat daripada data. Secara keseluruhannya, kesemua petunjuk ini memberikan bukti kukuh bahawa instrumen kajian ini memiliki kualiti psikometrik yang sangat baik dan sesuai digunakan bagi pengumpulan data berskala penuh dalam kajian sebenar yang akan datang.

**JADUAL III: PERSON RELIABILITY AND SEPARATION INDEX**

SUMMARY OF 133 MEASURED (EXTREME AND NON-EXTREME) PERSON									
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	368.6	87.0	2.60	.29					
S.D.	30.9	.0	1.97	.20					
MAX.	435.0	87.0	9.72	1.83					
MIN.	245.0	87.0	-1.08	.10	.05	-5.1	.03	-5.8	
REAL RMSE	.37	TRUE SD	1.93	SEPARATION	5.23	PERSON RELIABILITY	.96		
MODEL RMSE	.35	TRUE SD	1.94	SEPARATION	5.54	PERSON RELIABILITY	.97		
S.E. OF PERSON MEAN = .17									
PERSON RAW SCORE-TO-MEASURE CORRELATION = .96									
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .98									

Berdasarkan analisis Rasch bagi 87 item bukan ekstrem dalam instrumen, kebolehpercayaan item direkodkan pada nilai 0.84 (real) dan 0.86 (model), yang menunjukkan tahap ketekalan yang tinggi dalam pengukuran kesukaran item merentasi keseluruhan sampel. Menurut Bond dan Fox (2015), nilai kebolehpercayaan item yang melebihi 0.80 dianggap memuaskan dan mencerminkan kestabilan anggaran parameter item. Dapatan ini menunjukkan bahawa item-item dalam instrumen berfungsi secara konsisten merentasi responden dan sejajar dengan konstruk yang hendak diukur, sekali gus menyokong keteguhan serta keesahan skala yang dibangunkan.

Indeks pengasingan item pula direkodkan pada nilai 2.27 (real) dan 2.43 (model). Nilai ini menunjukkan bahawa item-item dalam instrumen dapat dikategorikan kepada kira-kira dua hingga tiga tahap kesukaran yang berbeza secara statistik. Menurut Linacre (1994), nilai indeks pengasingan melebihi 2.0 sudah memadai untuk menunjukkan perbezaan item yang bermakna dalam sesuatu konstruk. Ini menunjukkan bahawa instrumen kajian berupaya mendapatkan pelbagai tahap cabaran item dalam konstruk yang diukur, justeru memperkukuh ketepatan pengukuran dan perwakilan kandungan.

Secara keseluruhannya, nilai kebolehpercayaan dan pengasingan item ini mengesahkan bahawa item-item dalam soal selidik adalah kukuh dari segi psikometrik serta

menghasilkan keputusan yang konsisten dan boleh ditafsir dengan jelas merentasi pelbagai kumpulan responden. Dapatan ini memperkukuh kesahan keseluruhan instrumen dan membuktikan kesesuaiannya untuk digunakan bagi menilai konstruk psikologi atau tingkah laku yang dihasratkan dalam populasi guru yang lebih luas.

**Kesepadanan Item bagi Kesahan Kandungan**

**JADUAL IV: ITEM RELIABILITY AND SEPARATION INDEX**

SUMMARY OF 87 MEASURED (NON-EXTREME) ITEM									
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	563.5	133.0	.00	.21	1.01	-.1	.93	-.5	
S.D.	13.5	.0	.55	.02	.34	1.7	.56	1.9	
MAX.	597.0	133.0	1.73	.22	2.13	5.9	5.06	9.9	
MIN.	503.0	133.0	-1.55	.12	.48	-3.3	.34	-3.8	
REAL RMSE	.22	TRUE SD	.50	SEPARATION	2.27	ITEM	RELIABILITY	.84	
MODEL RMSE	.21	TRUE SD	.51	SEPARATION	2.43	ITEM	RELIABILITY	.86	
S.E. OF ITEM MEAN	= .06								

UMEAN=-.0000 USCALE=1.0000  
ITEM RAW SCORE-TO-MEASURE CORRELATION = -.99  
11397 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 11926.49 with 11177 d.f. p=.0000  
Global Root-Mean-Square Residual (excluding extreme scores): .4328

Berdasarkan analisis yang ditunjukkan dalam Jadual IV, statistik kesepadanan item telah diperiksa menggunakan nilai Infit dan Outfit Mean Square (MNSQ) serta Z-Standard (Zstd) bagi menilai sejauh mana setiap item mematuhi jangkakan Model Pengukuran Rasch. Secara umum, nilai MNSQ yang boleh diterima berada dalam julat 0.5 hingga 1.5 logit, yang menunjukkan prestasi item yang produktif dan menyumbang secara bermakna kepada konstruk laten yang diukur (Wright & Linacre, 1994). Dalam kajian ini, Jadual 5 *Item Fit* menunjukkan majoriti item memenuhi julat tersebut; namun, terdapat beberapa item yang melebihi had atas dan dikenal pasti sebagai item misfit berpotensi. Sebagai contoh, Item C22 mencatatkan nilai Infit MNSQ = 2.11, Outfit MNSQ = 5.06, serta Zstd = 9.9, iaitu jauh melebihi julat yang dibenarkan, menunjukkan pola respons yang tidak stabil atau tidak menentu.

Begitu juga, item-item lain seperti B9, C11, B17, dan B4 turut menunjukkan nilai Infit MNSQ melebihi 1.5, dengan Zstd antara 2.1 hingga 5.9, menunjukkan kemungkinan berlakunya ketidakselarasan tempatan (local misfit). Nilai-nilai ini mencadangkan bahawa item tersebut mungkin tidak berfungsi secara optimum dalam membezakan responden mengikut tahap keupayaan atau konstruk yang diukur. Fenomena ini juga berkemungkinan dipengaruhi oleh faktor kontekstual, linguistik, atau tafsiran responden terhadap pernyataan item. Walau bagaimanapun, perlu ditekankan bahawa statistik MNSQ bersifat sensitif terhadap saiz sampel, dan oleh itu, nilai overfit atau underfit tidak seharusnya dijadikan asas tunggal bagi penyingkiran item, khususnya dalam fasa kajian rintis (Bond & Fox, 2015).

Meskipun terdapat beberapa item yang menunjukkan nilai misfit secara statistik, tiada item yang digugurkan daripada instrumen dalam kajian ini. Keputusan tersebut dibuat berdasarkan beberapa pertimbangan utama. Pertama, semua item yang dikenal pasti sebagai misfit menunjukkan nilai PTMEA Corr yang positif (contohnya, 0.01 bagi C22 dan melebihi 0.30 bagi item lain), yang

menunjukkan bahawa item-item tersebut masih sejajar dengan arah konstruk laten yang diukur. Kedua, secara keseluruhan instrumen ini mempamerkan ciri psikometrik yang kukuh, termasuk kebolehpercayaan person yang tinggi (0.96), kebolehpercayaan item yang baik (0.84), dan indeks pengasingan item yang memuaskan (2.27), sekali gus mengesahkan kestabilan dan kepekaan instrumen terhadap variasi kesukaran item.

Selain itu, keputusan untuk mengekalkan semua item adalah selaras dengan amalan Rasch yang mapan dalam fasa awal pengesanan instrumen, di mana data misfit ditafsir secara berhati-hati dan lazimnya diperhalusi dalam kajian sebenar berskala besar (Boone, 2016; Linacre, 2021). Dengan mengekalkan keseluruhan set item, kajian ini memastikan tiada item bermakna digugurkan secara pramatang, sekali gus membolehkan penilaian psikometrik yang lebih menyeluruh dilaksanakan dalam konteks kajian sebenar dengan sampel yang lebih besar dan lebih representatif. Pada peringkat seterusnya, analisis dalam kajian penuh akan mempertimbangkan aspek tambahan seperti fungsi pembezaan item (Differential Item Functioning, DIF), pola taburan respons, dan keterkaitan kandungan item, bagi menilai prestasi setiap item secara lebih mendalam dan menyeluruh.

**JADUAL V: ITEM FIT**

ITEM STATISTICS: MISFIT ORDER												
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT [MNSQ]	OUTFIT [MNSQ]	[PT-MEASURE]	[EXACT MATCH]	ITEM			
68	583	133	1.73	.12	2.11	4.75	5.06	9.9	C22			
9	589	133	1.18	.22	1.13	5.91	1.64	2.5	B9			
89	561	133	1.15	.21	1.72	3.81	1.64	2.3	C11			
17	573	133	1.01	.22	1.02	3.11	1.51	2.1	B17			
68	582	133	1.17	.16	1.35	1.81	1.66	2.5	B4			
4	537	133	1.15	.21	1.64	4.31	1.51	2.4	B9			
62	581	133	.92	.18	1.94	1.61	1.54	2.8	C22			
12	551	133	.69	.28	1.53	1.91	1.64	2.3	B9			
61	587	133	.71	.19	1.58	1.91	1.98	1.5	C11			
13	589	133	1.18	.22	1.45	2.81	1.17	4.3	B9			
28	556	133	.86	.28	1.43	1.61	1.64	2.7	C11			
75	588	133	.82	.18	1.43	1.61	1.62	1.6	B17			
6	562	133	1.11	.21	1.27	1.31	1.43	1.7	B9			
15	576	133	1.15	.22	1.42	2.11	1.27	1.2	B9			
63	589	133	.68	.19	1.41	1.61	1.36	1.4	B9			
3	575	133	1.18	.22	1.48	2.11	1.13	1.6	B9			
18	568	133	1.19	.21	1.38	1.71	1.98	1.5	B9			
3	578	133	1.05	.22	1.17	1.61	1.16	1.7	B9			
7	561	133	1.15	.21	1.22	1.11	1.38	1.3	B9			
14	599	133	1.22	.22	1.38	1.51	1.85	1.7	B9			
85	582	133	.88	.18	1.28	1.81	1.38	1.1	B9			
16	576	133	.69	.28	1.24	1.31	1.98	1.8	B9			
11	587	133	1.11	.21	1.24	1.21	1.54	1.3	B9			
64	585	133	.78	.19	1.22	.91	1.16	.7	B9			
11	569	133	1.22	.22	1.18	1.81	1.65	1.6	B9			
76	585	133	.61	.28	1.17	1.81	1.68	1.2	B9			
39	582	133	.53	.28	1.41	1.11	1.48	1.1	B9			
15	575	133	1.18	.22	1.48	1.81	1.39	1.4	B9			
BETTER FITTING												
58	563	133	.86	.21	1.95	1.21	1.74	1.1	B9			
22	572	133	.36	.22	1.95	1.21	1.58	1.1	B9			
15	563	133	.86	.21	1.93	1.21	1.73	1.1	B9			
66	572	133	.36	.22	1.89	1.01	1.72	1.1	B9			
67	566	133	1.08	.21	1.88	1.01	1.71	1.1	B9			
19	568	133	1.19	.21	1.86	1.01	1.77	1.1	B9			
79	558	133	.86	.28	1.84	1.01	1.72	1.1	B9			
54	568	133	.38	.21	1.84	1.01	1.74	1.1	B9			
73	575	133	1.18	.22	1.83	1.01	1.88	1.1	B9			
16	568	133	.86	.28	1.81	1.01	1.77	1.1	B9			
23	578	133	1.05	.22	1.82	1.01	1.87	1.1	B9			
41	567	133	1.12	.22	1.82	1.01	1.72	1.1	B9			
34	565	133	1.02	.21	1.81	1.01	1.84	1.1	B9			
38	583	133	.84	.28	1.81	1.01	1.81	1.1	B9			
78	581	133	1.19	.22	1.81	1.01	1.76	1.1	B9			
86	579	133	1.19	.22	1.81	1.01	1.77	1.1	B9			
11	578	133	1.05	.22	1.88	1.01	1.78	1.1	B9			
38	575	133	1.08	.22	1.79	1.01	1.85	1.1	B9			
28	564	133	1.02	.21	1.78	1.01	1.91	1.1	B9			
33	557	133	1.12	.21	1.78	1.01	1.85	1.1	B9			
43	569	133	.94	.21	1.78	1.01	1.68	1.1	B9			
18	563	133	.86	.21	1.77	1.01	1.57	1.1	B9			
21	568	133	1.02	.21	1.77	1.01	1.54	1.1	B9			
27	563	133	.86	.21	1.77	1.01	1.55	1.1	B9			
42	568	133	1.19	.21	1.77	1.01	1.51	1.1	B9			
25	565	133	1.02	.21	1.75	1.01	1.69	1.1	B9			
16	566	133	.86	.21	1.74	1.01	1.71	1.1	B9			
38	566	133	.36	.28	1.71	1.01	1.68	1.1	B9			
36	567	133	1.12	.22	1.69	1.01	1.43	1.1	B9			
24	569	133	.94	.21	1.68	1.01	1.48	1.1	B9			
46	562	133	1.11	.21	1.65	1.01	1.54	1.1	B9			
96	568	133	1.17	.22	1.66	1.01	1.51	1.1	B9			
59	563	133	.86	.21	1.63	1.01	1.45	1.1	B9			
32	562	133	.73	.28	1.62	1.01	1.51	1.1	B9			
11	563	133	.86	.21	1.59	1.01	1.38	1.1	B9			
11	557	133	1.12	.21	1.58	1.01	1.49	1.1	B9			
47	569	133	1.22	.22	1.57	1.01	1.47	1.1	B9			
72	575	133	1.08	.22	1.57	1.01	1.58	1.1	B9			
45	578	133	1.05	.22	1.55	1.01	1.44	1.1	B9			
68	572	133	1.16	.22	1.54	1.01	1.42	1.1	B9			
58	567	133	1.12	.22	1.48	1.01	1.34	1.1	B9			
MEAN	563.5	133.0	.00	.21	1.81	1.93	1.93	1.93				
S.D.	13.5	.0	.55	.02	.34	1.7	.56	1.9				

**Polariti Item bagi Kesahan Kandungan**

Kesahan konstruk dalam kajian rintis ini dinilai melalui analisis polariti item menggunakan nilai Korelasi

Titik-Ukuran (Point Measure Correlation, PTMEA Corr) yang dijana daripada analisis Rasch. PTMEA Corr menunjukkan sejauh mana sesuatu item sejajar dengan sifat laten (latent trait) yang hendak diukur. Secara umumnya, nilai korelasi positif menandakan bahawa item berfungsi dalam arah yang betul serta menyumbang secara bermakna kepada pengukuran konstruk. Menurut Bond dan Fox (2015) serta Linacre (2012), nilai PTMEA Corr yang boleh diterima biasanya melebihi 0.30, yang mencerminkan sumbangan yang mencukupi terhadap pembentukan konstruk yang sah.

Dapatan kajian dalam Jadual VI di bawah menunjukkan ini menunjukkan bahawa majoriti daripada 87 item yang dianalisis mencatatkan nilai PTMEA Corr antara 0.29 hingga 0.80, menandakan penjajaran yang baik antara item dan konstruk masing-masing. Walaupun item C22 menunjukkan nilai korelasi yang agak rendah (0.01), item tersebut masih berada dalam julat yang boleh diterima bagi statistik Infit dan Outfit MNSQ, justeru tidak menjejaskan kesahan keseluruhan instrumen. Memandangkan kajian ini merupakan fasa rintis yang bersifat pemerolehan, dan struktur item perlu dikekalkan bagi tujuan pengesahan lanjut, pengkaji mengambil keputusan untuk mengekalkan kesemua item untuk digunakan dalam kajian sebenar. Pendekatan ini disokong oleh Teori Pengukuran Reflektif (Reflective Measurement Theory) yang menegaskan bahawa kekuatan hubungan antara item dan konstruk merupakan aspek penting dalam menetapkan kesahan konstruk (Hair et al., 2014). Oleh itu, pengesanan semua 87 item pada peringkat ini dianggap wajar, dengan penilaian lanjutan akan dijalankan dalam analisis faktor pengesahan (confirmatory factor analysis) pada fasa kajian seterusnya bagi menilai semula kesahan dan kesepadanan setiap item dalam struktur model yang lengkap.

**JADUAL VI: POINT MEASURE CORRELATION (PTMEA CORR) FOR POLARITY ITEM IN CONSTRUCT VALIDITY**

ITEM STATISTICS: CORRELATION ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL		INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		ITEM
				S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%		
60	503	133	1.73	.12	2.11	4.7	5.06	9.9	.01	.54	41.2	61.9	C22	
68	532	133	1.17	.16	1.35	1.4	1.66	2.5	.29	.56	80.2	75.5	D4	
75	544	133	.82	.18	1.43	1.6	1.42	1.6	.31	.59	71.0	77.8	D11	
69	535	133	1.09	.17	.97	-1.1	1.16	-7	.32	.57	80.2	76.7	D5	
85	542	133	.88	.18	1.20	.8	1.28	2.1	.38	.58	75.6	77.4	D21	
62	541	133	.92	.18	1.44	1.6	1.54	2.0	.42	.58	74.8	77.2	C24	
61	547	133	.71	.19	1.50	1.9	1.40	1.5	.42	.59	84.0	78.6	C23	
65	550	133	.60	.19	.99	.0	1.01	.1	.42	.60	79.4	79.3	D1	
64	545	133	.78	.19	1.22	.9	1.16	-7	.44	.59	83.2	77.9	C26	
63	549	133	.64	.19	1.41	1.6	1.36	1.4	.46	.60	85.5	79.2	C25	
76	555	133	.41	.20	1.17	.8	1.04	-2	.48	.61	76.3	79.9	D12	
9	589	133	-1.18	.22	2.13	5.9	1.64	2.5	.48	.61	74.0	79.1	B9	
82	561	133	.15	.21	1.17	.9	1.07	-4	.49	.63	78.6	80.3	D18	
80	558	133	.28	.21	1.03	.2	.92	-3	.51	.62	80.2	80.1	D16	
4	597	133	-1.55	.21	1.64	4.3	1.65	2.4	.52	.58	74.8	77.6	B4	
81	563	133	.06	.21	.97	-1.1	.89	-4	.54	.63	81.7	80.4	D17	
79	554	133	.45	.20	.84	-7	.72	-1.2	.57	.61	80.9	79.8	D15	
49	561	133	.15	.21	1.72	3.0	1.64	2.3	.57	.63	73.3	80.3	C11	
78	573	133	-.41	.22	1.01	.1	.83	-7	.58	.64	80.9	80.5	D14	
83	569	133	-.22	.22	.90	-5	.86	-6	.58	.63	79.4	80.6	D19	
84	569	133	-.22	.22	.98	-1	.98	.0	.59	.63	81.7	80.6	D20	
13	589	133	-1.18	.22	1.46	2.8	1.17	.8	.59	.61	77.1	79.1	B13	
71	575	133	-.50	.22	.97	-1	.83	-7	.59	.63	76.3	80.4	D7	
77	587	133	-1.08	.22	.94	-3	.93	-3	.60	.61	80.2	79.4	D13	
67	566	133	-.08	.21	.88	-6	.75	-1.1	.60	.63	80.2	80.5	D3	
87	576	133	-.55	.22	.92	-4	.98	.0	.60	.63	79.4	80.6	D23	
29	556	133	.36	.20	1.43	1.8	1.46	1.7	.61	.62	75.6	80.0	B29	
39	552	133	.53	.20	1.01	.1	.76	-1.0	.61	.61	85.5	79.6	C1	
66	572	133	-.36	.22	.89	-6	.72	-1.3	.62	.64	80.2	80.6	D2	
17	573	133	-.41	.22	1.67	3.1	1.51	2.1	.62	.64	70.2	80.5	B17	
1	576	133	-.55	.22	1.42	2.1	1.27	1.2	.62	.63	77.9	80.6	B1	
12	553	133	.49	.20	1.48	1.9	1.53	1.9	.62	.61	71.8	79.7	B12	
70	581	133	-.79	.22	.87	-7	.92	-3	.62	.63	82.4	80.2	D6	
14	569	133	-.22	.22	1.30	1.5	1.05	.3	.63	.63	76.3	80.6	B14	
37	544	133	.82	.18	.87	-5	.96	-1.1	.63	.59	77.1	77.8	B37	
73	575	133	-.50	.22	.83	-9	.80	-9	.64	.63	80.2	80.4	D9	
6	562	133	.11	.21	1.27	1.3	1.43	1.7	.64	.63	74.0	80.3	B6	
38	543	133	.85	.18	.75	-1.0	.81	-7	.64	.58	80.2	77.6	B38	
7	561	133	.15	.21	1.22	1.1	1.34	1.3	.65	.62	76.3	80.2	B7	

**Analisis Komponen Prinsipal (PCA) bagi Kesahan Konstruk**

Analisis Komponen Prinsipal (Principal Component Analysis, PCA) terhadap sisa piawai dalam analisis Rasch merupakan kaedah penting bagi menilai kesahan konstruk, khususnya dalam mengesahkan andaian unidimensionaliti, iaitu sama ada sekumpulan item benar-benar mengukur satu sifat laten (latent trait) yang sama. Berdasarkan Jadual VII, jumlah varian mentah keseluruhan yang diperhatikan dalam data ialah 145.3 unit, mewakili 100% daripada keseluruhan varian. Daripada jumlah tersebut, sebanyak 58.3-unit eigenvalue atau 40.1% varian dijelaskan oleh ukuran Rasch (iaitu dimensi utama model), menandakan struktur unidimensionaliti yang sederhana kukuh. Menurut Linacre (2005), peratusan varian yang dijelaskan oleh ukuran melebihi 40% dianggap memadai untuk menyokong kewujudan unidimensionaliti. Oleh itu, dapatan ini memenuhi kriteria tersebut dan menunjukkan bahawa majoriti varians dijelaskan oleh satu konstruk utama yang diukur oleh instrumen.

Sementara itu, varian tidak dijelaskan (unexplained variance) dalam kontras pertama dicatatkan sebanyak 11.7 unit eigenvalue atau 8.1%, iaitu jauh di bawah ambang maksimum 10% yang disarankan oleh Linacre (2005) dan Aziz et al. (2013). Nilai ini menunjukkan ketiadaan dimensi sekunder yang dominan yang boleh mengancam kesahan konstruk instrumen. Begitu juga, varian tidak dijelaskan dalam kontras ke-2 hingga ke-5 berada dalam julat 2.3% hingga 6.5%, mengesahkan bahawa tiada faktor sekunder yang cukup besar untuk mencabar andaian unidimensionaliti.

Secara keseluruhannya, hasil analisis PCA ini memberikan bukti empirikal yang kukuh terhadap kesahan konstruk instrumen, mengesahkan bahawa semua item berfungsi secara kolektif untuk mengukur satu konstruk laten yang sama seperti yang ditetapkan dalam andaian model Rasch. Dapatan ini juga sejajar dengan amalan terbaik dalam pengukuran Rasch bagi penilaian psikometrik instrumen dalam penyelidikan pendidikan, sebagaimana disarankan oleh Boone (2016) dan Bond dan Fox (2015).

**JADUAL VII: STANDARDISED RESIDUAL VARIANCE FOR PCA ANALYSIS**

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

	-- Empirical --	Modeled
Total raw variance in observations =	145.3 100.0%	100.0%
Raw variance explained by measures =	58.3 40.1%	42.5%
Raw variance explained by persons =	41.7 28.7%	30.3%
Raw Variance explained by items =	16.7 11.5%	12.1%
Raw unexplained variance (total) =	87.0 59.9%	57.5%
Unexplned variance in 1st contrast =	11.7 8.1%	13.5%
Unexplned variance in 2nd contrast =	9.5 6.5%	10.9%
Unexplned variance in 3rd contrast =	5.7 3.9%	6.6%
Unexplned variance in 4th contrast =	3.7 2.6%	4.3%
Unexplned variance in 5th contrast =	3.3 2.3%	3.8%

## VI. PERBINCANGAN

Dapatan kajian rintis ini mengesahkan bahawa instrumen adaptasi yang mengukur kepimpinan beretika, amalan pengurusan konflik dan budaya kerja guru mempunyai sifat psikometrik yang kukuh apabila dinilai menggunakan Model Pengukuran Rasch. Nilai indeks kebolehppercayaan person yang tinggi (0.96) serta indeks pengasingan person (5.23) menunjukkan bahawa responden dalam kajian ini mempunyai tahap kepelbagaian yang mencukupi untuk membolehkan instrumen membezakan tahap keupayaan dan persepsi secara berkesan antara individu (Tavakol & Dennick, 2011). Seterusnya, nilai kebolehppercayaan item (0.84) dan indeks pengasingan item (2.27) menunjukkan taburan item yang baik dan konsisten, sekali gus mengesahkan kestabilan struktur instrumen merentasi konstruk yang diukur (Fisher, 2007; Linacre, 2021).

Dari segi kesahan kandungan, majoriti item menunjukkan nilai Infit dan Outfit Mean Square (MNSQ) yang berada dalam julat 0.5 hingga 1.5 logit (Wright & Linacre, 1994), menandakan kesepadanan item yang baik dengan model Rasch serta menunjukkan bahawa setiap item menyumbang secara bermakna kepada pengukuran konstruk laten yang disasarkan. Walaupun terdapat satu item (C22) yang dikenal pasti mempunyai nilai misfit dan korelasi PTMEA yang rendah (0.01), item tersebut dikekalkan untuk penilaian lanjut dalam kajian sebenar kerana relevansinya dari sudut teori. Secara keseluruhan, nilai PTMEA Corr berada dalam julat 0.01 hingga 0.80, dengan majoriti item menunjukkan korelasi positif, sekali gus mengesahkan polariti item yang sejajar dengan konstruk masing-masing (Bond & Fox, 2015).

Selain itu, kesahan konstruk turut disokong oleh hasil Analisis Komponen Prinsipal (PCA) terhadap sisa residual. Dapatan menunjukkan bahawa dimensi Rasch menerangkan 40.1% daripada jumlah varians mentah, melebihi ambang 40% yang disyorkan untuk menyokong andaian unidimensionaliti (Linacre, 2005). Sementara itu, varians tidak dijelaskan dalam kontras pertama hanyalah

8.1%, iaitu di bawah had maksimum 10% yang diterima. Keputusan ini menunjukkan bahawa instrumen mempunyai struktur pengukuran tunggal yang kukuh serta gangguan multidimensi yang minimum.

Secara keseluruhannya, keputusan ini mengesahkan bahawa instrumen kajian ini adalah sah dan boleh dipercayai untuk menilai konstruk yang disasarkan, iaitu kepimpinan beretika, amalan pengurusan konflik dan budaya kerja guru dalam konteks sekolah menengah Malaysia. Keputusan ini juga memperkukuh keyakinan bahawa instrumen ini bersedia untuk digunakan dalam kajian sebenar berskala penuh, dengan potensi besar untuk menyumbang kepada penyelidikan empirikal berasaskan bukti dalam bidang kepimpinan pendidikan dan budaya organisasi sekolah.

## VII. KESIMPULAN, LIMITASI DAN CADANGAN KAJIAN LANJUTAN

Kajian rintis ini telah berjaya menegaskan kesahan dan kebolehppercayaan instrumen adaptasi baharu yang dibangunkan untuk mengukur kepimpinan beretika, amalan pengurusan konflik, dan budaya kerja guru di sekolah menengah. Melalui penggunaan Model Pengukuran Rasch, instrumen ini memperlihatkan ciri psikometrik yang mantap, termasuk ketekalan dalaman yang tinggi, kebolehppercayaan person dan item yang baik, kesepadanan item dalam julat diterima, serta struktur unidimensional yang kukuh. Dapatan ini memberi asas empirikal yang meyakinkan untuk meneruskan penggunaannya dalam kajian utama berskala lebih besar yang melibatkan populasi guru dari pelbagai negeri dan konteks sekolah.

Instrumen ini juga memberi sumbangan signifikan kepada pembangunan alat ukur berpaksikan konteks tempatan, sejajar dengan keperluan sistem pendidikan Malaysia yang menekankan nilai etika, kepimpinan berasaskan integriti, dan budaya kerja profesional. Dengan kesahan yang telah dibuktikan pada peringkat rintis, instrumen ini dapat digunakan untuk menilai secara lebih menyeluruh hubungan antara etika kepimpinan, strategi pengurusan konflik, dan pembentukan budaya kerja positif dalam kalangan warga pendidik. Hasilnya diharap dapat memberi implikasi praktikal kepada pembuat dasar, pemimpin sekolah dan agensi latihan kepimpinan dalam memperkukuh dasar pembangunan profesional guru dan amalan kepimpinan beretika di sekolah menengah.

Namun begitu, kajian ini mempunyai beberapa limitasi yang perlu diambil kira. Pertama, saiz sampel kajian rintis yang terhad kepada 133 responden dari tiga buah sekolah di Melaka mungkin tidak menggambarkan sepenuhnya kepelbagaian populasi guru di seluruh Malaysia. Oleh itu, generalisasi dapatan perlu dilakukan dengan berhati-hati. Kedua, analisis Rasch dijalankan dalam konteks instrumen adaptasi yang baru disahkan secara awal, justeru perlu diuji semula menggunakan kaedah Analisis Faktor Pengesahan (Confirmatory Factor Analysis, CFA) dan Pemodelan Persamaan Berstruktur (Structural Equation Modeling, SEM) untuk mengesahkan struktur konstruk secara lebih mendalam. Ketiga, faktor perbezaan budaya organisasi, lokasi geografi dan jenis

sekolah juga berpotensi mempengaruhi persepsi responden terhadap konstruk yang dikaji.

Sehubungan itu, kajian lanjutan disarankan agar dilaksanakan dengan saiz sampel yang lebih besar dan melibatkan pelbagai negeri, termasuk kawasan bandar, luar bandar dan pedalaman bagi memastikan kebolehlulangan (replicability) dan kebolegunaan umum instrumen. Kajian masa depan juga digalakkan untuk menilai fungsi pembezaan item (Differential Item Functioning, DIF) bagi memastikan kesetaraan ukuran merentasi jantina, pengalaman dan tahap akademik guru. Selain itu, penyelidik juga boleh menggabungkan pendekatan kaedah campuran (mixed-method) untuk menilai interpretasi kualitatif terhadap amalan kepimpinan beretika dan dinamika budaya kerja di sekolah.

Secara keseluruhannya, instrumen ini bukan sahaja menambah nilai dalam wacana penyelidikan kepimpinan pendidikan tetapi juga menyumbang alat ukur yang sah dan boleh dipercayai bagi memperkukuh pelaksanaan dasar pendidikan berasaskan etika dan integriti. Justeru, hasil kajian ini seiring dengan agenda pembaharuan pendidikan nasional, khususnya dalam memperkasa kualiti kepimpinan sekolah dan profesionalisme guru ke arah kecemerlangan sistem pendidikan Malaysia.

#### PENGHARGAAN

Penulis ingin merakamkan setinggi-tinggi penghargaan dan ucapan terima kasih kepada Bahagian Tajaan, Kementerian Pendidikan Malaysia atas sokongan yang diberikan terhadap pelaksanaan inisiatif penyelidikan ini. Penghargaan khas turut ditujukan kepada pengetua serta para guru di sekolah yang terlibat atas kerjasama dan komitmen mereka sepanjang fasa kajian rintis dijalankan. Seterusnya, penulis merakamkan setinggi-tinggi penghargaan kepada penyelia utama, Profesor Madya Dr. Bity Salwana Alias, dan penyelia bersama, Dr. Aida Hanim A. Hamid, atas bimbingan ilmiah yang bernas, nasihat akademik yang berterusan, serta sokongan padu sepanjang tempoh kajian ini. Ucapan terima kasih juga ditujukan kepada rakan penyelidik dan rakan seperjuangan akademik atas maklum balas yang membina serta dorongan moral yang telah banyak membantu dalam menyempurnakan kajian ini.

#### RUJUKAN

- Abushihab, I. (2016). Foreign words in Jordanian Arabic among Jordanians living in Irbid city: The impact of foreign languages on Jordanian Arabic. *Journal of Language Teaching and Research*, 7(2), 284–292.
- Al Hosni, S. (2014). Speaking difficulties encountered by young EFL learners. *International Journal on Studies in English Language and Literature (IJSELL)*, 2(6), 22–30.
- Arba'in, N., Mohd Noor, A., & Adnan, N. (2024). School resources, workload, and teacher job stress. *Attarbawiy: Malaysian Online Journal of Education*, 8(1), 1–17.
- Awang, Z. (2018). *A handbook on SEM: Structural equation modelling using AMOS*. MPWS Rich Resources.
- 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*, 15(4), rm4.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.
- Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: A review and future directions. *The Leadership Quarterly*, 17(6), 595–616.
- Bryk, A. S., & Schneider, B. (2004). *Trust in schools: A core resource for improvement*. Russell Sage Foundation.
- Deal, T. E., & Peterson, K. D. (2009). *Shaping school culture: Pitfalls, paradoxes, and promises* (2nd ed.). Jossey-Bass.
- Farooqi, M. T. K., Ali, S., & Ahmed, S. (2023). Effects of school principals' conflict management styles on teachers' job performance. *Bulletin of Education and Research*, 45(3), 19–34.
- Ghavifekr, S., Nair, A., & Ibrahim, M. S. (2019). Practising Conflict Management Strategies in Primary Schools: A Malaysian Scenario. *Global Business & Management Research*, 11(1).
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Pearson Education Limited.
- Kalshoven, K., Den Hartog, D. N., & De Hoogh, A. H. B. (2011). Ethical leadership at work (ELW): Development and validation of a multidimensional measure. *The Leadership Quarterly*, 22(1), 51–69.
- Kementerian Pendidikan Malaysia. (2024). *Pelan Strategik KPM 2024–2030*.
- Langlois, L., Lapointe, C., Valois, P., & de Leeuw, A. (2014). Development and validity of the ethical leadership questionnaire. *Journal of Educational Administration*, 52(3), 310–331.
- Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, 7(4), 328.
- Linacre, J. M. (2005). *A user's guide to Winsteps Rasch-model computer programs*. Winsteps.com.
- Linacre, J. M. (2012). *Winsteps® Rasch measurement computer program user's guide*. Winsteps.com.
- Linacre, J. M. (2021). *Winsteps® Rasch measurement computer program user's guide*. Winsteps.com.
- Ministry of Education Malaysia. (2023). *Malaysia Education Blueprint (MEB) 2013–2025: Annual report 2023*.
- Özdoğan, M., & Sarier, Y. (2024). The relationship of ethical leadership with teachers' organizational behavior, attitudes, and perceptions: A meta-analysis study. *Humanities and Social Sciences Communications*, 11, Article 640.
- Patimah, S. (2019). Teachers' Work Culture in an Islamic Junior High School in Lampung.

- Indonesia. *Pertanika Journal of Social Sciences & Humanities*, 27(1).
- Saad, S. C., Aziz, N. A., & Rahim, A. (2025). Teaching on the edge: How workload drives burnout among Malaysian educators. *International Journal of Academic Research in Business and Social Sciences*, 15(8), 1–13.
- Schein, E. H. (2010). *Organizational culture and leadership* (4th ed.). Jossey-Bass.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55.
- Ting-Toomey, S., & Dorjee, T. (2018). *Communicating across cultures*. Guilford Publications.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370–371.