

KESAHAN DAN KEBOLEHPERCAYAAN INSTRUMEN LITERASI PENTAKSIRAN BILIK DARJAH MELALUI MODEL PENGUKURAN RASCH

Mazliana Md Said¹, Zolkepli Haron², Shahlan Surat²

¹Fakulti Pendidikan
Universiti Kebangsaan Malaysia (UKM)
43600 Bangi, Selangor, Malaysia.

²Fakulti Pendidikan
Universiti Kebangsaan Malaysia
43600 Bangi, Selangor, Malaysia.

p88421@siswa.ukm.edu.my

Abstrak

Kajian rintis ini bertujuan untuk menghasilkan bukti empirikal kesahan dan kebolehpercayaan item Instrumen Literasi Pentaksiran Bilik Darjah yang diadaptasi daripada Approaches to Classroom Assessment Inventory (DeLuca et al. 2016). Kesahan dan kebolehpercayaan instrumen ini dianalisis menggunakan Model Pengukuran Rasch dibantu oleh perisian Winstep 3.71.0.1. Instrumen kajian rintis mengandungi 32 item dan telah diedarkan kepada 50 orang guru Matematik dalam daerah Muar. Analisis kesahan instrumen dibuat melalui pemeriksaan polariti item, kesesuaian item, unidimensionaliti dan kebebasan setempat. Analisis kebolehpercayaan item dibuat menerusi pemeriksaan kefungsi item dari aspek kebolehpercayaan dan indeks pengasingan item-individu. Hasil analisis polariti item mendapati semua item mempunyai nilai PMC yang positif di antara 0.33 hingga 0.77. Analisis kesepadanan item mendapati terdapat 31 item dalam instrumen yang mempunyai nilai min kuasa dua infit di antara 0.68 hingga 1.33 dan nilai min kuasa dua outfit di antara 0.68 hingga 1.25. Terdapat satu item yang terkeluar daripada julat yang dibenarkan. Analisis unidimensionaliti mencatatkan varians reja terpiawai sebanyak 43.3% dengan dimensi sekunder terbesar dalam kontras pertama mencatatkan varians sebanyak 4.9% dan 31 item menepati syarat kebebasan setempat. Dari sudut kebolehpercayaan pula, analisis Rasch menunjukkan nilai kebolehpercayaan individu adalah 0.92, nilai kebolehpercayaan item adalah 0.89, nilai pengasingan individu adalah 3.51 dan indeks pengasingan item adalah 2.82. Dapatan ini telah memberikan membuktikan bahawa Instrumen Literasi Pentaksiran Guru mempunyai tahap kesahan dan kebolehpercayaan yang tinggi untuk digunakan dalam pengukuran literasi pentaksiran bilik darjah. **Kata kunci:** Literasi pentaksiran bilik darjah; Model Pengukuran Rasch; analisis item; kesahan; kebolehpercayaan.

Abstract

This pilot study aimed to produce empirical validity and reliability evidence of the Classroom Assessment Literacy Instrument items that were adapted from the Approaches to Classroom Assessment Inventory (DeLuca et al. 2016). The validity and reliability of this instrument was analyzed using the Rasch Measurement Model assisted by the Winstep 3.71.0.1 software. Pilot study instrument contains 32 items and has been distributed to 50 Mathematics teachers in Muar district. The validity analysis of the instrument was made by checking the items polarity, item fit, unidimensionality and local independence. Item reliability analysis is done through item functionality checks on item-person reliability and separation index. Results of item polarity analysis found that all items had positive PMC values ranging from 0.33 to 0.77. The item fitness analysis found 31 items in the instruments that have infit square mean values between 0.68 to 1.33 and the outfit mean square values between 0.68 to 1.25. There is one item that fall out of the allowed range. The unidimensionality analysis recorded a standardized residual variance of 43.3% with the largest secondary dimension in the first contrast which recorded a variance of 4.9 % and 31 items meet local independence requirements. For reliability, Rasch's analysis showed that individual reliability was 0.92, item reliability was 0.89, the person separation value was 3.51 and the item separation value was 2.82. This finding has

proven that the Teacher Assessment Literacy Instrument has a high degree of validity and reliability to be used in measuring the classroom assessments literacy.

Keywords: Classroom assessment literacy; Rasch Measurement Model; item analysis; validity; reliability.

1.0 PENGENALAN

Pentaksiran bilik darjah menjadi suatu aspek yang paling kritikal dalam tugas guru pada hari ini. Banyak penyelidikan menyatakan bahawa tugas pentaksiran bilik darjah sebagai salah satu tugas guru yang paling mencabar, rumit dan penting (Metin, 2013). Guru yang menguasai literasi pentaksiran bilik darjah bererti mempunyai pengetahuan dan kemahiran yang mencukupi tentang pentaksiran bilik darjah dan berupaya mengintegrasikan pentaksiran bilik darjah secara berterusan dalam proses pengajarannya untuk menyokong pembelajaran pelajar (Birenbaum et. al., 2015; DeLuca, 2012; Popham, 2013).

Banyak kajian-kajian lepas yang menyatakan kepentingan pentaksiran bilik darjah dan keberkesannya mengukuhkan proses pengajaran dan pembelajaran (Birenbaum et. al., 2015; Brookhart, 2011; DeLuca, 2012, Popham, 2013, Ragonis, 2012). Selain itu, kajian empirikal juga menunjukkan peningkatan yang signifikan dalam pencapaian pelajar, fungsi metakognitif dan motivasi untuk belajar apabila guru mengintegrasikan pentaksiran dalam pengajaran mereka (Black & Wiliam 1998; Earl, 2003; Gardner, 2006; Willis, 2010). KPM juga menerusi surat Ruj:KPM.600-5/115 Jld.3(6) bertarikh 23 November 2017 telah mengarahkan semua sekolah perlu melaksanakan Pentaksiran Bilik Darjah (PBD) dalam sesi persekolahan 2018.

2.0 LATARBELAKANG / PERNYATAAN MASALAH

Terdapat banyak usaha membangunkan instrumen untuk mengukur literasi pentaksiran bilik darjah (Campbell et al. 2002; Plake et al. 1993; Mertler, 2005; Volante & Fazio, 2007; Zheng, 2009) namun penyelidikan mendapati bahawa instrumen-instrumen tersebut tidak lagi relevan untuk mewakili perubahan yang berlaku dalam landskap pentaksiran hari ini (Brookhart, 2011; Gotch & Perancis, 2014). Penemuan ini tidaklah menghairankan memandangkan kebanyakan instrumen pengukuran literasi pentaksiran bilik darjah yang dibina adalah berasaskan standard pentaksiran yang dibangunkan pada awal tahun 1990an iaitu Standards for Teacher Competency in Educational Assessment of Students (STCEAS), American Federation of Teachers (EFT), National Council on Measurement in Education (NCME) dan National Education Association (NEA) (DeLuca, LaPointe-McEwan, dan Luhanga, 2016; Gotch & French, 2014). Brookhart (2011) menyatakan bahawa standard pentaksiran tahun 1990an telah ketinggalan zaman dalam dua aspek iaitu ia tidak mengambil kira konsep pentaksiran formatif semasa dan tidak mengambil kira isu pembaharuan dalam pendidikan yang berlaku dalam abad ke-21 ini.

Mengambil kira pandangan kajian-kajian lepas, DeLuca et al. (2016) telah membangunkan instrumen Approaches to Classroom Assessment Inventory (ACAI), instrumen yang boleh dipercayai untuk menggambarkan literasi pentaksiran bilik darjah guru-guru dalam konteks masa kini. ACAI dibangunkan berasaskan standard pentaksiran bilik darjah baharu yang diterbitkan oleh Joint Committee for Standards on Evaluation (JCSEE, 2015).

Pengkaji telah mengadaptasi instrumen ACAI bertujuan untuk mengukur tahap literasi pentaksiran bilik darjah dalam kalangan guru-guru Matematik di daerah Muar. Maka fokus utama kajian ini dijalankan ialah untuk menghasilkan bukti empirikal kesahan dan kebolehpercayaan instrumen ACAI yang diadaptasi ke dalam Bahasa Melayu menggunakan Model Pengukuran Rasch. Kajian kesahan dan kebolehpercayaan instrumen adalah sangat penting untuk mengekalkan ketepatan instrumen dalam memberikan maklumat yang konsisten. Menurut Howard dan Henry (1988) konsisten bermakna apabila item yang sama diuji beberapa kali mengenai perkara yang sama pada selang masa yang berbeza, hasil keputusan skor atau jawapan yang diberikan adalah lebih kurang sama.

3.0 SOROTAN KAJIAN / KAJIAN LITERATUR

Literasi pentaksiran ialah istilah yang dicipta untuk kali pertama oleh Stiggins (1991) yang menyatakan bahawa guru dengan literasi pentaksiran tahu apakah yang perlu ditaksir, tujuan mentaksir, kaedah mentaksir, kemungkinan masalah-masalah yang timbul ketika mentaksir dan tindakan mencegahnya daripada berlaku. Seterusnya definisi literasi pentaksiran bilik darjah diperluaskan lagi sebagai pengetahuan dan kemahiran guru untuk menyusun data tentang pencapaian pelajar dan menggunakan hasil tersebut

untuk meningkatkan pencapaian pelajar (Chappuis et al. 2012). Proses ini dilakukan secara sistematik dan menjadi asas kepada guru untuk membuat keputusan dan perancangan berkaitan pengajaran dan pembelajaran (Yamtim & Wongwanich, 2014).

3.1 Instrumen-instrumen Mengukur Literasi Pentaksiran Bilik Darjah

Para penyelidik telah membangunkan dan menggunakan pelbagai instrumen untuk mengukur literasi pentaksiran bilik darjah (atau literasi pentaksiran guru) selama beberapa dekad yang lalu sebagai usaha menyokong amalan pentaksiran bilik darjah. Kebanyakan kajian ini telah menggunakan langkah-langkah kuantitatif yang merujuk Standard Kompetensi Guru Dalam Pentaksiran Pendidikan Pelajar (STCEAS, 1990).

Jadual 1.0: Analisis Instrumen Pentaksiran Dalam dan Luar Negara

Instrumen	Tujuan	Rujukan
Teacher Assessment Literacy Questionnaire (TALQ) (Plake et. al., 1993)	Mengkaji literasi pentaksiran guru dalam perkhidmatan	STCEAS 1990
Assessment Literacy Inventory (ALI) (Campbell et. al., 2002)	Mengkaji literasi pentaksiran guru pra perkhidmatan	STCEAS 1990
Classroom Assessment Literacy Inventory (CALI) (Mertler, 2003)	Mengkaji literasi pentaksiran guru pra perkhidmatan dan dalam perkhidmatan	STCEAS 1990
Revised Assessment Literacy Inventory (ALI) (Campbell & Mertler, 2005)	Mengkaji literasi pentaksiran guru pra perkhidmatan	STCEAS 1990
Ujian Literasi Pentaksiran (ULP) (Rohaya, 2009)	Mengukur tahap literasi pentaksiran guru-guru sekolah menengah	STCEAS 1990
Inventori Literasi Bilik Kelas (ILPBK) (Abd Wahab, 2013)	Mengukur tahap literasi pentaksiran guru-guru sekolah rendah	STCEAS 1990
Instrumen Literasi Pentaksiran Institut Pendidikan Guru (ILP IPG) (Zakiah, 2016)	Mengukur literasi pentaksiran pelajar ijazah sarjana muda perguruan	STCEAS 1990

3.2 Standard Kompetensi Guru Dalam Pentaksiran Pendidikan Pelajar (1990)

Pada tahun 1987, The National Council on Measurement in Education (NCME) bekerjasama dengan tiga pertubuhan profesional dalam bidang pendidikan telah membangunkan standard untuk pentaksiran bilik darjah yang menumpukan kepada pentaksiran pelajar. Tiga pertubuhan profesional yang terlibat secara langsung itu ialah American Association of Colleges for Teacher Education (AACTE), American Federation of Teachers (AFT) dan National Education Association (NEA). Laporan projek ini telah diterbitkan pada tahun 1990 dan dinamakan sebagai Standard Kompetensi Guru Dalam Pentaksiran Pendidikan Pelajar (STCEAS). STCEAS mengandungi tujuh standard berhubung dengan kompetensi pentaksiran pelajar iaitu;

- a) Memilih kaedah pentaksiran yang bersesuaian dengan pengajaran.
- b) Membina kaedah pentaksiran yang bersesuaian dengan pengajaran.
- c) Mentadbir, menskor dan mentafsir hasil pentaksiran.
- d) Menggunakan hasil pentaksiran untuk membuat keputusan berkaitan pelajar, pengajaran, kurikulum dan sekolah.
- e) Membina prosedur penggredan yang sah untuk pelajar.

- f) Menyampaikan hasil pentaksiran kepada pelajar, ibu bapa dan pihak-pihak berkepentingan.
- g) Mengenal pasti amalan pentaksiran yang tidak beretika dan melanggar peraturan.

3.3 Standard Pentaksiran Bilik Darjah (2015)

The Joint Committee on Standards for Educational Evaluation (JCSEE) telah memperkenalkan Standard Pentaksiran Bilik Darjah (Classroom Assessment Standard, 2015). Standard ini adalah hasil usaha kolaboratif 16 badan profesional dalam bidang akademik yang menjadi ahli JCSEE seperti American Educational Research Association (AERA), National Council on Measurement in Education (NCME), American Psychological Association (APA), Canadian Evaluation Society (CES) dan Canadian Society for the Study of Education (CSSE). Standard ini merupakan garis panduan yang diterima umum oleh organisasi profesional dan antarabangsa sebagai petunjuk amalan pentaksiran murid yang adil. Standard yang dibina ini berdasarkan kepada pernyataan bahawa pentaksiran murid di dalam bilik darjah merupakan pusat pembelajaran murid. Standard Pentaksiran Bilik Darjah mengandungi 17 standard berhubung dengan kompetensi pentaksiran pelajar iaitu;

- a) Tujuan pentaksiran: Amalan pentaksiran bilik darjah perlu mempunyai tujuan yang jelas yang menyokong pengajaran dan pembelajaran.
- b) Objektif pentaksiran: Amalan pentaksiran bilik darjah perlu selari dengan objektif pembelajaran dan sesuai untuk setiap pelajar.
- c) Reka bentuk pentaksiran: Jenis dan kaedah pentaksiran bilik darjah yang digunakan hendaklah dengan jelas membenarkan/membolehkan pelajar menunjukkan pembelajaran mereka.
- d) Penglibatan pelajar dalam pentaksiran: Pelajar harus terlibat dalam proses pentaksiran dan menggunakan bukti pentaksiran untuk meningkatkan pembelajaran mereka.
- e) Penyediaan pentaksiran: Penyediaan guru dan pelajar yang mencukupi dari segi sumber, masa dan peluang belajar harus menjadi sebahagian daripada amalan pentaksiran kelas.
- f) Makluman kepada pelajar dan ibu bapa atau penjaga: Tujuan dan kegunaan pentaksiran bilik darjah perlu disampaikan kepada pelajar dan, apabila sesuai, ibu bapa / penjaga.
- g) Analisis prestasi pelajar: Kaedah untuk menganalisis bukti pembelajaran pelajar sepatutnya sesuai untuk tujuan pentaksiran dan amalan.
- h) Maklum balas berkesan: Amalan pentaksiran bilik darjah perlu memberi maklum balas yang tepat pada masanya dan berguna untuk meningkatkan pembelajaran pelajar.
- i) Tindakan susulan selepas pengajaran: Analisis prestasi pelajar hendaklah memaklumkan perancangan pengajaran dan langkah seterusnya untuk menyokong pembelajaran pelajar yang berterusan.
- j) Gred dan ulasan: Gred dan ulasan pentaksiran haruslah tepat menunjukkan tahap penguasaan murid dan sasaran pembelajaran kurikulum tertentu.
- k) Pelaporan: Laporan pentaksiran murid hendaklah berdasarkan bukti-bukti yang mencukupi dan memberikan ringkasan pembelajaran murid dengan cara yang jelas, tepat pada masanya dan berguna.
- l) Kepelbagaian budaya dan linguistik: Amalan pentaksiran bilik darjah perlu menghormati kepelbagaian budaya dan bahasa pelajar dan komuniti mereka.
- m) Pendidikan khas dan pembelajaran terbeza: Amalan pentaksiran bilik darjah perlu dibezakan dengan tepat untuk memenuhi keperluan pendidikan khusus semua pelajar.
- n) Bebas dari bias: Amalan dan keputusan pentaksiran bilik darjah tidak boleh dipengaruhi oleh faktor-faktor yang tidak berkaitan dengan tujuan pentaksiran.
- o) Kesahan: Amalan pentaksiran bilik darjah harus menyediakan maklumat yang mencukupi dan sesuai bagi menyokong keputusan yang baik mengenai pengetahuan dan kemahiran setiap pelajar.
- p) Kebolehpercayaan: Amalan pentaksiran bilik darjah perlu memberikan maklumat yang konsisten dan boleh dipercayai bagi menyokong keputusan yang baik mengenai pengetahuan dan kemahiran setiap pelajar.

- q) Refleksi: Amalan pentaksiran bilik darjah perlu dipantau dan disemak untuk meningkatkan kualiti keseluruhan pengajaran dan pembelajaran.

3.4 Pendekatan Pentaksiran Bilik Darjah

a) Pentaksiran Untuk Pembelajaran (Assessment for Learning)

Pentaksiran untuk pembelajaran merupakan proses pembelajaran berterusan, sepanjang unit pembelajaran (McMillan, 2007) dan berlaku dalam bilik darjah (Mohamad Azhar & Shahrir, 2007). Melalui pendekatan pentaksiran ini, pelajar akan menerima maklum balas bagi mempertingkatkan lagi pembelajaran (Shepard et al. 2005). Pentaksiran untuk pembelajaran membolehkan guru menentukan cara dan kaedah yang sesuai untuk membantu pelajar bagi meningkatkan prestasi mereka (Assessment Reform Group, 2002).

Kriteria 'kejayaan' digunakan untuk mengetahui sama ada pelajar sudah mencapai matlamat atau tidak. Hal ini penting agar guru mendapat maklum balas tentang perkembangan pelajar. Dapatlah dirumuskan bahawa pentaksiran untuk pembelajaran ditakrifkan sebagai proses mendapatkan data dan menginterpretasikan data itu bagi tujuan membuat keputusan mengenai kedudukan pelajar dalam pembelajarannya, hala tuju mereka serta cara terbaik untuk mempertingkatkan pembelajaran mereka (Noraini, 1995).

b) Pentaksiran Sebagai Pembelajaran (Assessment as Learning)

Pentaksiran sebagai pembelajaran menggunakan aktiviti atau latihan untuk mengukur keupayaan pelajar. Namun begitu, pentaksiran sebagai pembelajaran ini dilakukan sendiri oleh pelajar atau melalui rakan sebaya, untuk membolehkan pelajar mengenal pasti kelemahan dan kekuatan yang ada di dalam diri (Azizi, 2010). Tujuan pentaksiran sebagai pembelajaran ialah memberikan pelajar ruang dan meningkatkan autonomi pelajar untuk memperbaiki diri secara sendiri kerana mereka perlu menetapkan matlamat sendiri atau sasaran yang ingin dicapai. Pentaksiran ini boleh dikatakan formatif kerana mengukur sesuatu itu secara terus dan spesifik dalam sesuatu perkara bukannya menilai secara keseluruhan topik.

c) Pentaksiran Terhadap Pembelajaran (Assessment of Learning)

Pentaksiran terhadap pembelajaran adalah penggunaan tugas atau kegiatan untuk mengukur dan melaporkan tahap pencapaian keseluruhan pelajar dalam bidang pembelajaran tertentu pada masa tertentu (Harlen, 2004). Pentaksiran ini juga dikenali sebagai pentaksiran sumatif kerana pentaksiran kepada pembelajaran diukur secara keseluruhan subjek. Biasanya guru akan mentaksir pelajar dengan cara peperiksaan iaitu menyediakan soalan-soalan tentang semua topik yang diajar bagi sesuatu subjek dan kemudian meringkaskan pencapaian pelajar (Crooks 2001). Ringkasan pencapaian pelajar biasanya diberi dalam bentuk gred atau pun markah yang diperolehi oleh pelajar selepas peperiksaan. Rubrik permarkahan juga digunakan untuk melakukan pentaksiran kepada pembelajaran.

Pentaksiran terhadap pembelajaran membolehkan guru mengasingkan pelajar-pelajar melalui markah yang mereka perolehi. Kesimpulannya, pentaksiran terhadap pembelajaran bertujuan untuk menentukan sejauh mana objektif-objektif bagi sesuatu bidang mata pelajaran yang diajar itu berjaya dicapai oleh pelajar-pelajar. Selain itu dengan adanya gred pencapaian, dapatlah dibuat perbandingan antara individu atau kumpulan (Abu Bakar & Bhasah, 2008).

3.5 Amalan Pentaksiran Bilik Darjah

Sebelum pertengahan tahun 1980-an, pentaksiran pendidikan kebanyakannya tertumpu kepada ujian berskala besar (ujian standard) dan memberi tumpuan kepada ujian kertas dan pensel (Stiggins & Conklin, 1992). Selain itu, kebanyakan ujian kertas dan pensil yang digunakan untuk pentaksiran pula tidak konsisten dan sering tidak berkaitan dengan realiti pengajaran (Shulman, 1980). Akhirnya, pada tahun 1980an, barulah timbul isu tentang keputusan dan amalan tingkah laku guru, pencapaian pelajar dan mengenai pentaksiran bilik darjah berkaitan pengajaran dan pembelajaran (McMillan, 1998).

Kajian-kajian berkaitan amalan pentaksiran guru menunjukkan amalan pentaksiran guru bergantung kepada mata pelajaran yang diajar serta peringkat persekolahan yang diajar (Bol, et al. 1998; Marso & Pigge, 1987; Zhang & Burry-Stock, 2003). Misalnya, guru-guru Matematik didapati lebih kerap menggunakan item penyelesaian masalah (Marso & Pigge, 1987), item pengiraan (McMorris & Boothroyd, 1993) dan kurang menggunakan kaedah pentaksiran tradisional seperti ujian objektif berbanding dengan guru-guru mata pelajaran lain (Bol et. al., 1998). Guru Bahasa Inggeris pula lebih kerap menggunakan item esei seperti menghasilkan karangan, laporan dan dokumen untuk mentaksir pembelajaran pelajar. Manakala guru Sains lebih suka mengguna item aneka-pilihan (McMorris & Boothroyd, 1993).

3.6 Pembelajaran Profesionalisme Guru

Pembelajaran (Perkembangan) professional guru merujuk kepada apa sahaja pengalaman yang memperkayakan ilmu, penghayatan, kemahiran dan kefahaman guru-guru berhubung kerjayanya (Glickman et. al., 2001). Pengetahuan dan kemahiran guru boleh ditingkatkan menerusi pembelajaran professional guru (Mahaliza, 2014) dan pembelajaran profesional yang berkesan mempunyai elemen-elemen yang berikut iaitu berterusan, berasaskan sekolah, dilaksanakan secara job-embedded dan melibatkan guru dalam identifikasi dan mereka bentuk pembelajaran profesional untuk memenuhi keperluan mereka (Bredeson, 2013). Dalam persekitaran yang sentiasa berubah dan penuh ketidakpastian, guru-guru perlu sentiasa meningkatkan pengetahuan, kemahiran dan kompetensi sebagai asas pengajaran agar sentiasa relevan dengan peredaran masa (Fullan & Mascal, 2000). Latihan diperlukan apabila terdapat jurang di antara kompetensi, kemahiran, pengetahuan dan sikap yang dimiliki oleh guru dengan standard tugas yang ditetapkan. Hal ini kerana latihan dapat menambah baik pengetahuan, kemahiran atau tingkah laku guru berkaitan tugas mereka (Sparks & Horsely, 1989; Pekeliling Perkhidmatan Bilangan 6, Tahun 2005).

4.0 METODOLOGI

Perbincangan bahagian ini akan menerangkan sampel kajian, instrumen kajian dan tatacara penganalisan data.

4.1 Sampel Kajian dan Pengumpulan Data

Kajian rintis ini telah dilakukan dengan menggunakan kaedah kuantitatif kerana kemampuannya untuk mengumpul dan menganalisis data berbentuk numerikal bagi tujuan untuk menjelaskan fenomena yang dikaji (Gay et al., 2012). Sampel kajian yang terlibat ialah 50 orang guru Matematik sekolah menengah dalam daerah Muar. Bilangan sampel ini adalah mencukupi untuk penentuukuran yang stabil (Linacre, 1994).

Jadual 2.0: Ukuran Saiz Sampel dalam Model Pengukuran Rasch

Kalibrasi item stabil dalaman	Aras keyakinan	Julat saiz sampel minimum	Saiz sampel yang sesuai
± 1 logits	95 %	16 – 36	30
± 1 logits	99 %	27 – 61	50
± 0.5 logits	95 %	64 – 144	100
± 0.5 logits	99 %	108 - 243	150

Linacre 1994

4.2 Instrumen Kajian

Dalam kajian ini, pengkaji telah mengubah suai instrumen yang dibina oleh Christopher DeLuca et al. (2016) iaitu Approaches to Classroom Inventory (ACAI) untuk mengukur literasi pentaksiran bilik darjah guru. Pemilihan instrumen ini dibuat kerana ia merupakan instrumen pentaksiran bilik darjah yang menggunakan standard pentaksiran masa kini iaitu Standard Pentaksiran Bilik Darjah 2015. Bahagian pertama instrumen mengandungi item yang berkaitan dengan latar belakang sampel seperti kelulusan akademik, jantina, bangsa, umur, pengalaman mengajar, gred jawatan dan kategori sekolah. Bahagian kedua adalah berkaitan dengan literasi pentaksiran bilik darjah yang terdiri daripada tiga konstruk utama iaitu pendekatan pentaksiran bilik darjah (12 item), amalan pentaksiran bilik darjah (8 item) dan keperluan pembelajaran profesional (12 item).

Jadual 3.0: Kandungan Skala Pengukuran Literasi Pentaksiran Bilik Darjah

Bil	Konstruk	ID Item	Skala Pengukuran
1	Pendekatan pentaksiran bilik darjah (PPBD)	B1 – B12	Skala Likert 6 mata
2	Amalan pentaksiran bilik darjah (APBD)	C13 – C21	Skala Likert 5 mata
3	Keperluan pembelajaran profesionalisme (KPP)	D22 – D32	Skala Likert 5 mata

4.3 Analisis Data

Semua data yang dikumpul daripada instrumen dianalisis menggunakan perisian Winsteps Version 3.71.0.1 dengan pendekatan Model Pengukuran Rasch. Pendekatan Model Pengukuran Rasch digunakan untuk melihat kesahan dan kebolehpercayaan konstruk instrumen secara lebih mendalam melalui diagnosis seperti yang dinyatakan dalam jadual 4.0. Daripada 72 item yang dirintiskan dalam kajian rintis pertama, sebanyak 40 item telah digugurkan kerana tidak memenuhi keperluan Rasch. 32 item masih dikekalkan untuk digunakan dalam kajian sebenar.

Jadual 4.0: Prosedur Penganalisan Data

Aspek	Tujuan Analisis	Prosedur Analisis
Penentuan kesahan instrumen	Kesahan konstruk	<ol style="list-style-type: none"> 1. Unidimensionaliti 2. Kebebasan setempat 3. Polariti item 4. Keserasian item
Penentuan kebolehpercayaan instrumen	Kebolehpercayaan	<ol style="list-style-type: none"> 1. Kebolehpercayaan item 2. Kebolehpercayaan individu 3. Indeks pemisahan item 4. Indeks pemisahan individu

Unidimensionaliti ialah andaian dalam model Rasch yang perlu dibuktikan terlebih dahulu sebelum proses analisis yang lain dijalankan. Syarat asas dalam IRT ini bertujuan untuk memastikan bahawa data yang dikumpul bersifat selari atau searah (Wright & Masters, 1982) dalam mengukur satu konstruk yang sama. Dalam menilai unidimensionaliti, kriteria pertama yang diperhatikan adalah varians kasar yang diterangkan oleh pengukuran sekurang-kurangnya 40% (Linacre, 2006). Kriteria kedua adalah varians yang tidak dijelaskan dalam kontras pertama tidak lebih dari 15% (Azrilah et al., 2015). Kriteria ketiga adalah nisbah minimum 3: 1 untuk varians kasar yang diterangkan oleh pengukuran berbanding dengan varians yang tidak dijelaskan dalam kontras pertama (Conrad et al., 2012; Embretson & Reise, 2000). Akhir sekali kriteria keempat adalah nilai eigen mesti kurang dari lima, menunjukkan tidak wujud dengan jelas dimensi kedua (Linacre, 2005).

Ciri kebebasan setempat diperiksa untuk menentukan sejauh mana respon kebolehan individu kepada mana-mana item adalah tidak berkait dengan respon item lain dalam konstruk yang sama. Julat yang menepati syarat kebebasan setempat adalah nilai piawai residual korelasi yang kurang daripada 0.70 (Linacre, 2012). Item yang mempunyai nilai piawai residual korelasi yang bersamaan atau melebihi nilai 0.70 menunjukkan responden melihat pasangan item yang berkaitan sebagai sama dan mengelirukan.

Analisis polariti item (PMC) bertujuan menentukan sejauh mana item-item mengukur konstruk yang sama dan semua item mengukur sub konstruk tunggal Bond & Fox, 2007). Jika nilai PMC adalah tinggi, bererti item tersebut berupaya untuk membezakan kemampuan antara individu dan mengukur konstruk yang ingin diukur. Sebaliknya jika nilai adalah negatif, item yang dibangunkan tidak mengukur konstruk yang ingin diukur (Linacre, 2003).

Kesesuaian (fit) item dalam mengukur konstruk dapat dilihat melalui nilai infit MNSQ dan outfit MNSQ. Menurut Bond dan Fox (2007), nilai outfit MNSQ dan infit MNSQ perlu berada pada julat di antara 0.5 logits hingga 1.5 logits bagi memastikan item yang dibangunkan adalah sesuai untuk mengukur konstruk. Namun indeks outfit MNSQ perlu diberi perhatian terlebih dahulu berbanding infit bagi menentukan kesesuaian item yang mengukur sesuatu konstruk atau latent variable (Azrilah et al. 2015). Sekiranya nilai MNSQ lebih daripada 1.5 logits, maka ia memberi maksud item mengelirukan. Jika nilai MNSQ kurang daripada 0.5 logits pula ia menunjukkan item terlalu mudah dijangka oleh individu (Linacre, 2007). Selain itu, nilai outfit ZSTD dan infit ZSTD perlu juga berada dalam lingkungan -2 hingga +2 (Bond & Fox, 2007), namun jika nilai outfit dan infit MNSQ diterima, indeks ZSTD boleh diabaikan (Linacre, 2005). Jika syarat ini tidak ditepati, maka item boleh dipertimbangkan untuk disingkirkan atau dimurnikan.

Nilai kebolehpercayaan individu bermaksud andaian kebolehan individu dalam sampel adalah tekal walaupun diberi set item yang berlainan tetapi masih mengukur konstruk yang sama (Linacre, 2012). Nilai kebolehpercayaan item pula bermaksud item adalah mencukupi untuk mengukur perkara yang ingin diukur (Azrilah et al. 2013). Indeks pemisahan merupakan indeks yang mengelaskan individu atau item kepada

beberapa kumpulan. Nilai indeks pemisah individu dan item yang lebih daripada nilai 2 dianggap sebagai baik (Linacre, 2005).

5.0 DAPATAN KAJIAN

Bahagian ini akan melaporkan dapatan analisis kajian rintis berkaitan kesahan dan kebolehpercayaan instrumen Literasi Pentaksiran Bilik Darjah.

5.1 Kesahan Instrumen Literasi Pentaksiran Bilik Darjah

Berdasarkan Jadual 5.0, hasil analisis unidimensionaliti mendapati instrumen kajian ini mempunyai nilai varians reja terpiawai 43.3% dan dimensi sekunder terbesar dalam kontras pertama 8.6% adalah tidak melebihi had kawalan seperti yang ditetapkan oleh Model Pengukuran Rasch. Nisbah untuk varians kasar yang diterangkan oleh pengukuran berbanding dengan varians yang tidak dijelaskan dalam kontras pertama adalah 5:1 dan nilai eigen ialah 4.9. Ini menunjukkan instrumen ini bebas daripada sebarang kekeliruan dari segi tujuan kajian ini dilakukan.

Jadual 6.0 pula memaparkan hasil analisis kebebasan setempat. Terdapat dua item yang mempunyai nilai korelasi yang melebihi 0.7. Ini menunjukkan responden melihat pasangan item yang berkaitan sebagai satu perkara yang sama dan mengelirukan. Terdapat dua pendekatan yang boleh diambil iaitu sama ada dimurnikan susun atur ayatnya supaya menjadi lebih jelas atau digugurkan. Namun proses menggugurkan item perlu dilakukan dengan berhati-hati bagi memastikan kesahan kandungan dan konstruk tidak terancam.

Jadual 7.0 menunjukkan keseluruhan 32 item Instrumen Literasi Pentaksiran Guru mempunyai nilai PMC positif antara 0.33 hingga 0.77. Ini mengesahkan item yang diukur adalah dalam satu arah yang sama sahaja dan tidak bercanggah dengan konstruk yang hendak diukur.

Analisis kesesuaian item pula mendapati 31 item mempunyai nilai misfit antara 0.68 hingga 1.25. Nilai ini menepati julat pengukuran Bond & Fox (2007). Terdapat satu item (D23) terkeluar daripada julat keserasian yang ditetapkan. Namun item D23 perlu dipertimbangkan untuk dimurnikan atau digugurkan dengan melihat kepada keperluan pengkaji dan pandangan pakar.

Jadual 5.0: Analisis Komponen Prinsipal

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)					
			-- Empirical --		Modeled
Total raw variance in observations	=	56.4	100.0%		100.0%
Raw variance explained by measures	=	24.4	43.3%		44.2%
Raw variance explained by persons	=	12.7	22.4%		22.9%
Raw Variance explained by items	=	11.8	20.9%		21.3%
Raw unexplained variance (total)	=	32.0	56.7%	100.0%	55.8%
Unexplnd variance in 1st contrast	=	4.9	8.6%	15.2%	
Unexplnd variance in 2nd contrast	=	3.4	6.0%	10.6%	
Unexplnd variance in 3rd contrast	=	3.0	5.3%	9.3%	
Unexplnd variance in 4th contrast	=	2.5	4.5%	7.9%	
Unexplnd variance in 5th contrast	=	2.2	3.9%	7.0%	

Jadual 6.0: Analisis Kebebasan Setempat

CORREL- ATION	ENTRY NUMBER	ITE	ENTRY NUMBER	ITE
.75	29	D29	30	D30
.61	18	C18	24	D24
.60	2	B2	3	B3
.58	27	D27	28	D28
.51	15	C15	17	C17
.48	22	D22	23	D23
.47	21	D21	22	D22
.44	16	C16	20	C20
-.53	4	B4	11	B11
-.53	9	B9	13	C13

Jadual 7.0: Nilai Point Measure Correlation

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	ITEM	G
23	182	50	-2.58	.32	1.21	1.0	1.65	1.8	.33	.50	72.0	73.7	D23	0
4	221	50	-.31	.24	1.33	1.5	1.25	1.0	.45	.58	62.0	63.7	B4	0
3	215	50	.38	.29	1.15	.8	1.12	.6	.45	.52	64.0	69.7	B3	0
24	193	50	-1.11	.26	1.14	.7	1.21	1.0	.48	.57	70.0	65.6	D24	0
5	217	50	.32	.25	1.20	1.0	1.19	1.0	.48	.58	52.0	62.5	B5	0
25	180	50	-.17	.30	1.07	.4	1.10	.5	.49	.53	70.0	70.0	D25	0
2	218	50	-.15	.29	1.08	.4	1.01	.1	.50	.53	64.0	69.4	B2	0
30	160	50	.61	.22	1.25	1.3	1.24	1.2	.51	.64	36.0	56.4	D30	0
8	236	50	.70	.24	1.25	1.3	1.16	.7	.51	.60	62.0	59.1	B8	0
1	213	50	-.97	.20	1.31	1.2	1.53	2.0	.51	.62	48.0	56.4	B1	0
16	165	50	.79	.24	1.18	1.0	1.14	.8	.53	.60	66.0	59.1	C16	0
10	237	50	.74	.25	1.12	.7	1.12	.6	.53	.59	70.0	62.6	B10	0
32	186	50	-.63	.23	1.20	1.1	1.20	1.0	.54	.63	64.0	57.2	D32	0
11	226	50	-.73	.24	1.15	.7	1.23	1.0	.56	.60	74.0	63.0	B11	0
27	185	50	-.52	.26	1.02	.2	1.06	.4	.57	.58	74.0	65.6	D27	0
21	187	50	1.58	.34	.84	-.8	.78	-.8	.57	.49	82.0	75.9	D21	0
13	169	50	.21	.24	1.06	.4	1.05	.3	.58	.61	54.0	59.1	C13	0
9	229	50	-.47	.24	.98	-.1	.94	-.2	.60	.60	74.0	61.5	B9	0
6	221	50	.14	.26	.95	-.2	.93	-.3	.60	.57	62.0	63.3	B6	0
12	232	50	1.19	.27	.94	-.2	.87	-.6	.61	.56	74.0	65.9	B12	0
22	188	50	-.36	.29	.75	-1.0	.92	-.2	.61	.53	80.0	74.3	D22	0
29	166	50	.47	.24	.99	.0	.97	-.1	.61	.60	62.0	59.0	D29	0
18	184	50	-.37	.25	.89	-.5	.94	-.2	.64	.59	66.0	63.7	C18	0
7	230	50	-.78	.25	.87	-.6	.80	-.9	.65	.59	62.0	61.9	B7	0
26	188	50	1.19	.32	.78	-1.1	.77	-1.0	.66	.51	80.0	74.0	D26	0
15	177	50	-.12	.27	.82	-.9	.81	-.9	.66	.57	68.0	65.5	C15	0
20	167	50	.45	.24	.85	-.8	.84	-.8	.68	.61	66.0	58.8	C20	0
28	185	50	-.66	.28	.79	-1.0	.79	-1.0	.69	.55	80.0	68.3	D28	0
31	181	50	1.13	.26	.74	-1.4	.67	-1.6	.71	.57	70.0	64.7	D31	0
19	169	50	.21	.24	.73	-1.5	.73	-1.5	.74	.61	66.0	59.1	C19	0
14	174	50	-.06	.25	.68	-1.8	.68	-1.8	.76	.59	76.0	62.3	C14	0
17	178	50	-.14	.25	.68	-1.8	.68	-1.7	.77	.59	68.0	62.9	C17	0
MEAN	195.6	50.0	.00	.26	1.00	.0	1.01	.0			66.8	64.2		
S.D.	24.1	.0	.81	.03	.19	1.0	.23	1.0			9.6	5.3		

5.2 Kebolehpercayaan Instrumen Literasi Pentaksiran Bilik Darjah

Jadual 8.0 menunjukkan nilai kebolehpercayaan individu adalah 0.92 dan indeks pemisahan individu adalah 3.51. Ini menunjukkan kebolehan individu bagi kesemua konstruk adalah tekal walaupun diberi set item yang berlainan tetapi masih mengukur konstruk yang sama dan item-item instrumen berupaya membezakan individu kepada empat kumpulan tahap kebolehan.

Jadual 9.0 pula menunjukkan nilai kebolehpercayaan item dan indeks pemisah item masing-masing adalah 0.89 dan 2.82. Ini bererti item-item instrumen adalah mencukupi untuk mengukur apa yang seharusnya diukur dan bilangan sampel individu berupaya mengesahkan terdapat tiga hierarki kesukaran item instrumen.

Jadual 8.0: Analisis Indeks Kebolehpercayaan Individu dan Indeks Pemisahan Individu

SUMMARY OF 50 MEASURED PERSON									
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	
MEAN	125.2	32.0	-.08	.32	1.01	-.2	1.01	-.3	
S.D.	12.6	.0	1.29	.01	.60	2.2	.64	2.2	
MAX.	157.0	32.0	3.31	.36	2.90	5.3	3.15	5.6	
MIN.	100.0	32.0	-2.66	.32	.17	-5.1	.16	-4.5	
REAL RMSE	.35	TRUE SD	1.24	SEPARATION	3.51	PERSON RELIABILITY	.92		
MODEL RMSE	.32	TRUE SD	1.25	SEPARATION	3.91	PERSON RELIABILITY	.94		
S.E. OF PERSON MEAN	= .18								

Jadual 9.0: Analisis Indeks Kebolehpercayaan Item dan Indeks Pemisahan Item

SUMMARY OF 32 MEASURED ITEM								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	195.6	50.0	.00	.26	1.00	.0	1.01	.0
S.D.	24.1	.0	.81	.03	.19	1.0	.23	1.0
MAX.	237.0	50.0	1.58	.34	1.33	1.5	1.65	2.0
MIN.	160.0	50.0	-2.58	.20	.68	-1.8	.67	-1.8
REAL RMSE	.27	TRUE SD	.77	SEPARATION	2.82	ITEM	RELIABILITY	.89
MODEL RMSE	.26	TRUE SD	.77	SEPARATION	2.94	ITEM	RELIABILITY	.90
S.E. OF ITEM MEAN = .15								

6.0 PERBINCANGAN DAN RUMUSAN

Setelah data dianalisis, semakan kembali setiap item dilakukan dengan berpandukan indeks standard serta syarat-syarat yang perlu diikuti untuk mencapai standard kesahan dan kebolehpercayaan instrumen berdasarkan Model Pengukuran Rasch. Berdasarkan hasil dapatan kajian rintis kedua yang telah dijalankan, terdapat empat item yang tidak menepati standard analisis yang telah ditetapkan dan perlu disingkirkan atau dimurnikan dengan mengambil kira pandangan dan pentaksiran pakar. Manakala 28 item lagi mempunyai kesahan dan kebolehpercayaan yang tinggi untuk mengukur literasi pentaksiran bilik darjah. Hasil perbincangan bersama pakar, item-item tersebut digugurkan. Berdasarkan kajian rintis ini, dapat disimpulkan bahawa pemeriksaan kesahan dan kebolehpercayaan instrumen menunjukkan Instrumen Literasi Pentaksiran Guru mempunyai kualiti yang boleh digunakan untuk mengukur literasi pentaksiran bilik darjah dalam konteks Malaysia. Implikasi dari analisis ini dapat membantu pengkaji dalam membangunkan instrumen pentaksiran bilik darjah mengikut standard pentaksiran terkini dalam konteks negara Malaysia. Ini merupakan langkah awal dalam membantu semua pihak yang terlibat dalam melahirkan guru-guru yang mencapai standard pentaksiran yang ditetapkan oleh Kementerian Pendidikan Malaysia.

7.0 RUJUKAN

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). Standards for competence in the educational assessment of students. Retrieved from <http://www.unl.edu/buros/article3.html>

Azrilah, A., Azlinah, M., NoorHabibah, A., Sohaimi, Z., Azami, Z., Hamza, A.G. & Saidfudin, M. (2008). Application of rasch model in validating the construct of measurement instrument. *International Journal of Education and Information Technologies* Vol 2, 2008, 105-112. doi.org/10.1016/j.sbspro.2010.12.225

Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: implications for policy and practice. *Policy Futures in Education*, 13, 117-140. [doi/pdf/10.1177/1478210314566733](https://doi.org/10.1177/1478210314566733)

Bond, TG; Fox, C.M. (2001). *Applying the rasch model fundamental measurement in the human sciences*. ERL Lawrence Baum Associates Publishers, London. Retrieved from <https://www.rasch.org/rmt/rmt151c.htm>

Bredeson, P. V. (2003). *Designs for learning*. Thousand Oaks: Corwin Press.

Brookhart, S. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30, 3-12. [doi/full/10.1111/j.1745-3992.2010.00195.x](https://doi.org/10.1111/j.1745-3992.2010.00195.x)

- Chappuis, J., Stiggins, R., Chappuis, S., Arter, J. (2012). Classroom assessment for student learning: doing it right-using it well. Pearson Assessment Training Institute.
- DeLuca, C. (2012). Preparing teachers for the age of accountability: toward a framework for education. *Teacher Education Yearbook XXI: A Special Issue of Action in Teacher Education*, 34, 576–591. Retrieved from <https://doi.org/10.1080/01626620.2012.730347>
- DeLuca, C., & Bellara, A. (2013). The current state of assessment education: aligning policy, standards and teacher education curriculum. *Journal of Teacher Education*, 64, 356–372. doi.org/10.1177/0022487113488144
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17, 419–438. doi.org/10.1080/0969594X.2010.516643
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to classroom assessment inventory: a new instrument to support teacher assessment literacy. *Educational Assessment*, 21(4), 251–272. doi.org/10.1080/10627197.2016.1236677
- Fisher Jr., W. P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transaction*, 21, 1095. Retrieved from <http://www.rasch.org/rmt/rmt211a.htm>
- Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice*, 33, 14–18. doi.org/10.1111/emip.12030
- Howard W. & Henry I.B. (1988). Test validity. New Jersey: Lawrence Erlbaum Associates Publishers.
- Joint Committee for Standards on Educational Evaluation [JCSEE]. (2015). Classroom assessment standards: practices for PK-12 teachers. Retrieved from <http://www.jcsee.org/the-classroomassessment-standards-new-standards>
- Kashfi Mohd Jailani. (2011). Manual pengenalan pengukuran rasch & winstep. Pengukuran dan Penilaian dalam Pendidikan. Fakulti Pendidikan Universiti Kebangsaan Malaysia.
- Linacre J.M. (2003). Winsteps computer program version 3.48. Chicago: www.winsteps.com
- Linacre, J.M. (2004). Test validity and rasch measurement: construct, content, etc. *Rasch Measurement Transactions*. Retrieved from <https://www.rasch.org/rmt/rmt181h.htm>
- Linacre, J.M. (2010). User's guide to winsteps ministep rasch-model computer programs. Retrieved from <http://www.winsteps.com/winman/>
- Mahaliza Mansor. (2013). Hubungan antara model-model pembelajaran profesional terhadap amalan guru sekolah menengah harian di Malaysia. https://www.researchgate.net/publication/279205906_HUBUNGAN_ANTARA_MODEL-MODEL_PEMBELAJARAN_PROFESIONAL_TERHADAP_AMALAN_GURU_SEKOLAH_MENENGAH_HARIAN_DI_MALAYSIA
- Mertler, C. A. (2003). Preservice and Inservice Teachers' Assessment Literacy: Does Classroom Experience Make A Difference?. Annual meeting of the Mid-Western Educational Research Association, Columbus, OH. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.556.4217&rep=rep1&type=pdf>
- Mertler, C. A. (n.d). Classroom Assessment Literacy Inventory. Retrieved from <http://pareonline.net/htm/v8n22/cali.htm>

- Mertler, C. A., & Campbell, C. (2005). Measuring Teachers' Knowledge & Application of Classroom Assessment Concepts: Development of the Assessment Literacy Inventory. Annual meeting of the American Educational Research Association, Montreal, QC. Retrived from <https://files.eric.ed.gov/fulltext/ED490355.pdf>
- Metin, M. (2013). Teachers' difficulties in preparation and implementation of performance task. *Energy Education Science and Technology: Theory & Practice* - 13(3). 1664-1673. Retrived from <https://files.eric.ed.gov/fulltext/EJ1017658.pdf>
- Plake, B., Impara, J., & Fager, J. (1993). Assessment competencies of teachers: a national survey. *Educational Measurement: Issues and Practice*, 12, 10–39. doi:10.1111/j.1745-3992.1993.tb00548.x
- Popham, W. J. (2013). *Classroom assessment: what teachers need to know* (7th ed.). Boston, MA: Pearson.
- Rossen, D., Mazalah, A., M.Faisal, K., Norhaslinda, M.S., Aidah, A.K., Nur Ayu, J., Kamaruzaman, J., Mohamad Shanudin, Z., Khairul Anwar, M., & Siti Rahayah, A. (2009). Kesahan dan kebolehppercayaan soal selidik gaya e-pembelajaran (else) versi 8.1 menggunakan model pengukur rasch. *Jurnal Pengukuran Kualiti dan Analisis* 5(2), 15-27. Dimuat turun dari <http://journalarticle.ukm.my/1930/>
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29, 4–14. doi.org/10.3102/0013189X029007004
- Stiggins, R. J. (1991). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practices*, 10(1), 7–12. doi.org/10.1111/j.1745-3992.1991.tb00171.x
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: implications for teacher education reform and professional development. *Canadian Journal of Education*, 30, 749–770. Retrived from <https://files.eric.ed.gov/fulltext/EJ780818.pdf>
- Wright B. & Linacre J. (1992). Combining and splitting categories. Institute for Objective Measurement Inc. Retrived from <https://www.rasch.org/rmt/rmt63f.htm>
- Yamtim, V. & Wonwanich S. (2013). A study of classroom assessment literacy of primary school teachers. 5th World Conference on Educational Sciences, WCES. Retrived from <https://www.sciencedirect.com/science/article/pii/S1877042814007137>