

**[VALIDITY AND RELIABILITY OF THE STUDENT ATTITUDES TOWARDS
ARABIC LANGUAGE LEARNING INSTRUMENT (I-SPPBA) USING
THE RASCH MEASUREMENT MODEL]**

**KESAHAN DAN KEBOLEHPERCAYAAN INSTRUMEN SIKAP PELAJAR
TERHADAP PEMBELAJARAN BAHASA ARAB (I-SPPBA)
MENGGUNAKAN MODEL PENGUKURAN RASCH**

Nur Afiqah Athirah Mohd Rushdi

p116037@siswa.ukm.edu.my (Corresponding Author)

Research Center for Arabic Language and Islamic Civilization,

Faculty of Islamic Studies, National University of Malaysia

43600 UKM Bangi, Selangor, Malaysia

Lily Hanefarezan Asbulah

lilyhane@ukm.edu.my

Research Center for Arabic Language and Islamic Civilization,

Faculty of Islamic Studies, National University of Malaysia

43600 UKM Bangi, Selangor, Malaysia

Abstrak

Antara ciri terpenting dalam pembelajaran bahasa Arab ialah sikap. Oleh itu, instrumen yang jitu dari segi kesahan konstruk serta kebolehpercayaan untuk mengukur tahap sikap pelajar terhadap pembelajaran bahasa Arab adalah penting. Penyelidikan ini bertujuan untuk mengesahkan instrumen yang digunakan untuk menentukan tahap sikap pelajar bahasa Arab di universiti awam. Bagi mengumpul data, pengedaran soal selidik telah dilakukan kepada 40 pelajar tahun akhir Kolej Universiti Islam Antarabangsa Selangor (KUIS) dan Universiti Islam Antarabangsa Sultan Abdul Halim Mu'adzam Shah (UniSHAMS) yang mengikuti program pengajian bahasa Arab. Bagi mengukur kesahan dan kebolehpercayaan dalam kajian ini, Model Pengukuran Rasch berbantuan Winsteps versi 5.2. telah digunakan. Dapatkan menemukan tahap *noise* instrumen adalah pada 7.0% iaitu kurang daripada nilai terkawal maksimum pada 15%. Selain itu, hanya tiga item yang direkodkan terkeluar dari julat kesesuaian Mean Square namun telah dikekalkan dan dimurnikan. Sementara itu, indeks kebolehpercayaan item adalah pada 0.96 yang dianggap sebagai tahap kebolehpercayaan yang tinggi dengan indeks pemisahan responden pada 3.31. Dapatkan ini menunjukkan kesahan konstruk instrumen sikap pembelajaran bahasa Arab. Kesimpulannya, Model Pengukuran Rasch telah membuktikan bahawa instrumen sikap terhadap pembelajaran bahasa Arab dalam kajian ini mempunyai tahap kesahan dan kebolehpercayaan yang tinggi seterusnya wajar digunakan untuk mengukur tahap sikap pelajar di institut pengajian tinggi.

Kata kunci: sikap; pembelajaran bahasa Arab; penilaian instrumen; model pengukuran Rasch

Abstract

Among the most important feature in learning Arabic language is attitude. Due to its importance, it is essential to precisely measure the level of students' attitude towards learning Arabic language in terms of construct validity as well as reliability. This research aims to validate the instrument being used to determine Arabic language students' attitude level in public universities. To collect the data, distribution of questionnaires was done to 40 final year students at Selangor International Islamic University College (KUIS) and Sultan Abdul Halim Mu'adzam Shah International Islamic University (UniSHAMS), who are taking Arabic language course. To measure validity and reliability in this study, Rasch Measurement Model using Winsteps version 5.2. was used. The findings recorded noise level at 7.0% which is less than maximum controlled value at 15%. Apart from that, only three items were recorded out of the Mean Square suitability range but were retained and refined..

Meanwhile, item reliability index was at 0.96 which is considered as high degree of reliability with the person separation index at 3.31. These findings demonstrate the construct validity of Arabic language learning attitude instrument. In conclusion, Rasch Measurement Model has proven that the Arabic language learning attitude instrument in this study has high degree of authenticity as well as reliability, hence it is appropriate to measure students' attitude level at higher institute of education.

Keyword: attitude; Arabic language learning; instrument assessment; Rasch Measurement Model

Article Received:
14 October 2021

Article Reviewed:
7 November 2021

Article Published:
15 December 2022

PENGENALAN

Bahasa Arab merupakan salah satu bahasa utama dunia dengan menganggarkan mempunyai lebih 422 juta penutur di seluruh dunia (Ku Fatahiyah et al., 2014). Dalam penelitian sejarah, bahasa Arab telah lama meninggalkan kesan kepada tamadun Melayu sehingga menjadi sebahagian daripada identiti Islam Melayu. Walaupun menghadapi cabaran semasa yang terus menerus bagi melemahkan status bahasa Arab dalam globalisasi dan hegemoni budaya semasa, namun kadar dan minat orang Melayu yang berterusan dalam mempelajari bahasa Arab adalah bukti bahawa bahasa Arab adalah bahasa yang hidup (Majdi & Akmal Khuzaairy, 2018).

Di Malaysia, pendidikan bahasa Arab semakin mendapat perhatian kerajaan melalui institusi pendidikan negara. Ia bukan sahaja dipelajari oleh pelajar di peringkat sekolah rendah dan menengah sahaja, bahkan menjadi begitu terkenal di Institut Pengajian Tinggi (IPT). Pada peringkat sekolah yang beraliran agama, bahasa Arab menjadi salah satu mata pelajaran wajib dipelajari dengan menggunakan pendekatannya yang tersendiri. Manakala bagi sekolah bukan beraliran agama atau Institut Pengajian Tinggi Awam (IPTA) dan Institut Pengajian Tinggi Swasta (IPTS), pendekatan yang berlainan digunakan seperti bahasa Arab sebagai bahasa asing (Azman, 2013). Ini sejajar dengan Akta Pendidikan 1996 yang mengklasifikasikan bahasa Arab diwartakan sebagai bahasa asing setaraf dengan bahasa Jepun, Jerman atau Perancis yang mana bahasa-bahasa ini “boleh diadakan jika didapati munasabah dan praktik berbuat sedemikian” (Akta Pendidikan 1996). Melihat kepada kepayahan pengajaran dan pembelajaran bahasa Arab dewasa ini, kursus bahasa Arab menjadi lumrah di institusi pengajian tinggi. Semua usaha ini akan membawa hasil sekiranya pelajar bahasa Arab mempunyai sikap dan persepsi yang positif terhadap kepentingan bahasa ini dalam kehidupan masa hadapan.

Sikap adalah sesuatu yang berbentuk dalaman dan sukar untuk difahami (Azizi et al., 2007). Robbins (1986) menyatakan bahawa sikap terbentuk daripada persepsi serta pengalaman seseorang terhadap sesuatu perkara. Ini disokong oleh Gardner dan Lambert (1972) yang menyatakan bahawa kemampuan pelajar untuk menguasai bahasa kedua bukan hanya dipengaruhi oleh kecekapan mental atau kemahiran berbahasa sahaja, tetapi juga pada sikap dan persepsi pelajar terhadap bahasa sasaran. Chalak dan Kassaian (2010) juga menyatakan bahawa pelajar bahasa kedua yang berjaya adalah orang yang mempunyai sikap positif terhadap bahasa, budaya dan masyarakat.

Sikap positif biasanya dikaitkan dengan penutur jati bahasa sasaran serta budayanya. Sikap positif sebegini boleh meningkatkan prestasi pembelajaran kerana pelajar boleh dijangka mempunyai keinginan untuk berkomunikasi dengan penutur asli bahasa yang mereka pelajari. Dalam erti kata lain, sekiranya pelajar tersebut mempunyai persepsi yang positif dengan negara dan budaya bahasa yang dipelajari, maka mereka akan lebih tertarik untuk mempelajari bahasa tersebut (Noels et al., 2000). Sebaliknya,

sikap negatif boleh menghalang pembelajaran bahasa kerana biasanya pelajar mempunyai persepsi yang negatif apabila menghadapi kesukaran dalam bahasa tersebut seterusnya tidak berminat untuk mempelajarinya (Ellis, 1994). Namun, sikap pelajar boleh berubah. Mereka mungkin mempunyai sikap yang negatif pada awal pembelajaran bahasa tetapi setelah mereka menyedari kebaikan dan keseronokan mengetahui serta menguasai bahasa tersebut, maka sikap mereka akan berubah. Dalam kajian ini, sikap didefinisikan sebagai tingkah laku dan persepsi pelajar terhadap bahasa Arab berdasarkan sub konstruk kognitif, afektif dan konatif.

Aspek kognitif merupakan komponen sikap yang diwakili oleh kepercayaan individu yang merujuk kepada kebarangkalian tingkah laku untuk menghasilkan sesuatu yang diinginkan itu dapat dinilai sebagai sesuatu yang baik atau tidak baik (Hale et al., 2003). Menurut Mohamad Jafre et al. (2012), aspek sikap ini melibatkan kepercayaan pelajar bahasa mengenai pengetahuan yang mereka terima dan pemahaman mereka dalam proses pembelajaran bahasa. Sikap kognitif ini melibatkan empat langkah iaitu menghubungkan pengetahuan sebelumnya dengan yang baru, mencipta pengetahuan baru, memeriksa pengetahuan baru dan menerapkan pengetahuan baru dalam banyak situasi.

Seterusnya, aspek afektif adalah komponen yang melibatkan perasaan dan emosi seseorang terhadap sesuatu seterusnya membentuk sikap positif atau negatif terhadap perkara tersebut (Zuchdi 1995). Feng dan Chen (2009) menyatakan bahawa proses belajar adalah proses emosi. Ia dipengaruhi oleh faktor yang emosi yang berbeza. Sikap dapat membantu pelajar untuk menyatakan sama ada mereka suka atau tidak menyukai sesuatu. Choy dan Troudi (2006) bersetuju bahawa perasaan dan emosi dalam pelajar bahasa kedua mempengaruhi perspektif dan sikap mereka terhadap bahasa sasaran.

Aspek konatif pula merupakan komponen tingkah laku yang berkaitan dengan cara seseorang bertindak dan bertindak balas dalam situasi tertentu (Mohamad Jafre et al., 2012). Kara (2009) menyatakan bahawa sikap positif mengarah pada tingkah laku positif terhadap kursus pengajian. Pelajar akan melibatkan diri untuk mendapatkan sesuatu dan berusaha untuk belajar lebih banyak. Pelajar seperti itu juga dilihat lebih bersemangat untuk menyelesaikan masalah, memperoleh maklumat dan kemahiran yang berguna untuk kehidupan seharian dan melibatkan diri secara emosional.

PENYATAAN MASALAH

Kajian lepas mendapati bahawa tahap penguasaan bahasa Arab pelajar di peringkat pengajian tinggi berada pada tahap yang membimbangkan seperti kajian Nurain Syafina Husain dan Norhayuza Mohamad (2020), yang menyatakan bahawa pelajar bahasa Arab diperingkat universiti masih kekurangan bekalan kosa kata yang memadai untuk menjalani proses pengajian di institusi pengajian tinggi. Menurut Stroud dan Wee (2006) yang menekankan bahawa pembelajaran bahasa adalah sukar kerana pelajar sering tidak yakin untuk bercakap bahasa sasaran dengan alasan mereka takut dinilai oleh guru atau penutur asli bahasa tersebut. Ini menunjukkan bahawa sikap positif sememangnya merupakan aspek penting ke arah kejayaan pembelajaran bahasa.

Selain itu, kajian lepas telah membuktikan bahawa penyelidikan mengenai kesahan dan kebolehpercayaan instrumen soal selidik sikap pelajar IPT terhadap bahasa Arab masih kurang dijalankan. Terdapat beberapa kajian lepas mengenai kesahan instrumen sikap seperti Wan Emril Nizar (2018) yang telah membina serta mengesahkan instrumen soal selidik sikap berdasarkan amalan adab belajar Imam al-Shafi'i, namun sasaran bagi instrumen tersebut adalah pelajar Sekolah Berasrama Penuh. Begitu juga dengan Norkumalasari Othman et al. (2020) yang juga menjalankan penyelidikan tentang pengesahan instrumen sikap, namun skop penyelidikannya adalah sikap pelajar sekolah menengah terhadap Matematik.

Oleh itu, bagi mengukur tahap sikap pelajar IPT terhadap pembelajaran bahasa Arab memerlukan kepada suatu alat pengukuran yang jitu seperti yang dicadangkan oleh Silke Stegmann (2013) dalam kajian beliau agar pengkaji menggunakan soal selidik yang sesuai untuk mengukur aspek sikap pelajar. Hal ini adalah untuk memastikan bahawa instrumen yang digunakan mampu mengukur dan menguji apa yang hendak diukur (Azizi et al., 2017). Justeru, tujuan kajian ini dijalankan adalah untuk membangunkan instrumen bagi mengukur tahap sikap pelajar terhadap pembelajaran bahasa Arab (i-SPPBA) dalam kalangan mahasiswa Bahasa Arab di institusi pengajian tinggi dan seterusnya menentukan kesahan dan kebolehpercayaan instrumen tersebut.

METODOLOGI

Kajian ini merupakan kajian kuantitatif yang berasaskan kaedah tinjauan. Kaedah persampelan rawak mudah telah digunakan bagi sampel kajian yang terdiri daripada 40 orang pelajar tahun akhir di dua buah universiti swasta iaitu, Kolej Universiti Islam Antarabangsa Selangor (KUIS) dan Universiti Islam Antarabangsa Sultan Abdul Halim Mu'adzam Shah (UniSHAMS) yang sedang mengikuti program pengajian bahasa Arab.

Instrumen i-SPPBA ini terdiri daripada dua bahagian. Bahagian A soal selidik ini mengandungi item yang berkaitan latar belakang responden seperti umur, jantina, nama universiti, program pengajian, tempoh pengalaman belajar bahasa Arab, tahap penguasaan bahasa Arab, faktor pendorong mengikut program pengajian dan bidang kerjaya yang diceburi. Bahagian B pula mengandungi 12 soalan berkaitan sikap pelajar terhadap pembelajaran bahasa Arab seperti dalam Jadual 1. Item-item tersebut meliputi tiga sub konstruk, iaitu afektif, kognitif dan konatif yang diadaptasi daripada instrumen Mohamad Jafre et al. (2012). Aspek kognitif mengukur tahap kepercayaan pelajar mengenai pengetahuan dan pemahaman mereka dalam proses pembelajaran bahasa. Aspek afektif pula mengukur perasaan dan emosi pelajar terhadap bahasa manakala aspek konatif mengukur tingkah laku seseorang sama ada mengarah kepada tindak balas positif atau negatif. Seterusnya, dua item daripada soal selidik terpilih telah diubahsuai daripada konteks ayat negatif kepada positif.

Dari segi skala kajian, skala Likert 4 mata telah dipilih bagi menggambarkan sikap pelajar terhadap pembelajaran bahasa Arab, iaitu 1 (sangat tidak setuju), 2 (tidak setuju), 3 (setuju) dan 4 (sangat setuju). Instumen soal selidik i-SPPBA ini tidak menggunakan skala neutral bagi mengelakkan pelajar memilih jalan tengah ketika menggambarkan pilihan mereka.

Proses pengumpulan data dilaksanakan secara atas talian secara sepenuhnya. Kebenaran telah dipohon daripada Dekan Kulliyah Bahasa Arab, UniSHAMS dan Dekan Fakulti Pengajian Peradaban Islam, KUIS untuk menjalankan kajian ini. Seterusnya, borang soal selidik diedarkan dalam bentuk *Google Form* melalui wakil pelajar dan juga ketua jabatan kepada responden disebabkan oleh proses pengajaran dan pembelajaran pada semester tersebut adalah sepenuhnya secara atas talian di kediaman masing-masing. Setelah proses pengumpulan data selesai, data tersebut dianalisis menggunakan model pengukuran Rasch dengan berbantuan perisian Winsteps versi 5.2.

KESAHAAN INSTRUMEN KAJIAN

Dalam kajian kesahan, kesahan kandungan (*content validity*) dan kesahan muka (*face validity*) adalah penting dalam memastikan instrumen yang digunakan dalam kajian adalah tepat dan mampu mencerminkan hasil dapatan dengan konstruk yang diuji (Mohammad Rahim & Wan Shahrazad, 2019).

Bagi kajian ini, seramai lima orang panel pakar telah dilantik bagi menjalankan proses kesahan kandungan terhadap instrumen i-SPPBA. Pemilihan panel pakar bagi mengesahkan kandungan item soal selidik ini adalah berdasarkan bidang kepakaran mereka dalam psikometrik, sains sosial, pembinaan soal selidik dan linguistik bahasa Arab. Pakar-pakar tersebut merupakan timbalan dekan Fakulti Pendidikan (UIAM), pensyarah Fakulti Pengajian Bahasa dan Pembangunan Insan (UMK), Fakulti Pengajian Islam (UKM), Akademi Pengajian Bahasa (UiTM) dan juga Guru Cemerlang Sekolah Menengah Sains Pasir Puteh. Panel pakar telah menyemak serta memberi pandangan terhadap struktur gaya bahasa, format serta kandungan soal selidik tersebut.

Setelah proses penilaian pakar selesai, kesahan kandungan bagi instrumen tersebut diukur menggunakan Nisbah Kesahan Kandungan (*Content Validity Ratio-CVR*) dan Indeks Kesahan Kandungan (*Content Validity Index-CVI*) iaitu prosedur pengukuran kuantitatif kesahan kandungan yang diperkenalkan oleh (Lawshe, 1975). CVR bertujuan untuk mengesahkan item melalui pengukuran empirikal dan memfokuskan mengenai kepentingan sesuatu item dalam instrumen seterusnya membantu pengkaji untuk menentukan samada item pada instrumen perlu digugurkan atau dikekalkan (Mohd Effendi @ Ewan Mohd Matore et al., 2017). Julat yang diterima bagi nilai CVR adalah -1 sehingga +1, dan nilai yang menghampiri +1 membuktikan persetujuan panel pakar akan kepentingan item tersebut dalam kesahan kandungan. Lawshe (1975) mencadangkan bahawa nilai CVR yang melebihi nilai sifar ($CVR>0$) membuktikan bahawa majoriti panel pakar yang terlibat telah bersetuju untuk menilai item sebagai sangat penting dan telah memenuhi kesahan kandungan. Secara keseluruhannya, kesemua item i-SPPBA telah mencapai pekali persetujuan yang baik iaitu melebihi nilai sifar dan nilai tersebut menunjukkan bahawa semua item adalah mewakili domain konstruk dan telah disahkan penting.

Nilai CVI pula akan dikira secara keseluruhan setelah mengenalpasti nilai CVR bagi setiap item. Lynn (1986) mencadangkan tiga hingga sepuluh orang panel pakar untuk menjalankan proses kesahan. Menurut beliau, nilai CVI yang diterima adalah >0.80 dengan mengambil kira purata tahap persetujuan antara panel pakar. Secara keseluruhannya, nilai CVI bagi kesemua item dalam instrumen i-SPPBA mencapai pekali persetujuan yang sesuai iaitu melebihi nilai 0.80. Seterusnya, instrumen i-SPPBA telah diakui berupaya untuk menguji aspek kandungan yang hendak diuji oleh panel pakar. Namun begitu, beberapa saranan dan cadangan oleh panel pakar telah diberi perhatian, antaranya:

1. Terdapat item yang bersifat double-barreled dan perlu dipisahkan menjadi dua item.
2. Terdapat item yang bersifat terlalu umum dan perlu dijelaskan dengan lebih terperinci.
3. Beberapa item perlu dilihat semula agar selari dengan sub-konstruk instrumen.

Justeru, pengkaji telah memperbaiki semula instrumen i-SPPBA mengikut saranan dan cadangan pakar. Seterusnya, instrumen diedarkan kepada pelajar dan data dianalisis bagi mendapatkan kebolehpercayaan item instrumen ini.

ANALISIS DAN PERBINCANGAN

Data dianalisis menggunakan pendekatan model pengukuran Rasch berbantuan perisian Winsteps versi 5.2. Pemeriksaan fungsian item dilakukan dari enam aspek, iaitu (a) indeks kebolehpercayaan dan pengasingan item-responden; (b) mengesan polariti item berdasarkan nilai PTMEA CORR; (c) mengesan kesesuaian item; (d) menentukan item bersandar berdasarkan nilai korelasi residual terpiawai; (e) mengukur keseragaman

dimensi (*unidimensionality*); dan (f) mengesan tahap kesukaran item dan kebolehan responden berdasarkan peta item-responden.

Indeks Kebolehpercayaan dan Pengasingan Item-Responden

Kebolehpercayaan item dan responden boleh mengesahkan sejauh mana item mematuhi model pengukuran Rasch serta indeks pengasingan item dan responden. Indeks pengasingan pula menunjukkan bilangan strata kesukaran item dan keupayaan individu yang dikenal pasti dalam kumpulan yang dikaji. Nilai Alpha Cronbach bagi instrumen i-SPPBA pula adalah 0.93 yang mana dikategorikan sebagai sangat baik dan efektif dengan tahap konsistensi yang tinggi seperti yang dinyatakan oleh Bond dan Fox (2015) dalam Jadual 1.

JADUAL 1. Interpretasi Skor Alpha Cronbach

Skor Alpha Cronbach	Interpretasi Kebolehpercayaan
0.9-1.0	Sangat baik dan efektif dengan tahap konsistensi yang tinggi
0.7-0.8	Baik dan boleh diterima
0.6-0.7	Boleh diterima
<0.6	Item perlu dibaiaki
<0.5	Item perlu digugurkan

Setelah pemeriksaan dan kajian fungsian dijalankan, nilai kebolehpercayaan item bagi 12 item i-SPPBA adalah 0.96 dengan nilai pengasingan item 4.68 seperti yang ditunjukkan dalam Jadual 2. Menurut Bond dan Fox (2015), nilai kebolehpercayaan item yang lebih tinggi daripada 0.80 adalah sangat baik dan efektif dengan tahap konsistensi yang tinggi. Indeks pengasingan item bagi i-SPPBA adalah 4.68, yang mana dapat ditafsirkan sebagai tinggi dan menepati syarat yang digariskan oleh Linacre (2002) yang menekankan bahawa nilai pengasingan item yang lebih tinggi daripada 2.00 adalah dikategorikan sebagai baik.

Nilai kebolehpercayaan responden bagi i-SPPBA pula adalah 0.92 dengan nilai pengasingan responden 3.31 seperti yang dinyatakan dalam Jadual 3. Kebolehpercayaan responden yang melebihi nilai 0.80 membuktikan tindak balas yang baik dan konsisten (Bond & Fox, 2015). Bagi nilai pengasingan, Linacre (2002) telah berpendapat bahawa nilai pengasingan responden yang melebihi 2.00 menunjukkan nilai pengasingan yang baik bagi kesukaran item. Oleh itu, indeks pengasingan bagi i-SPPBA ini menunjukkan bahawa lebih 3 strata kesukaran item dan keupayaan responden dapat dikesan. Secara keseluruhannya, i-SPPBA ini sangat sesuai digunakan untuk penyelidikan sebenar.

JADUAL 2. Nilai Kebolehpercayaan dan Pengasingan Item terhadap Instrumen i-SPPBA

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	121.8	40.0	.00	.38	.97	-.2	1.05	.3
S.D.	13.5	.0	1.91	.01	.28	1.2	.40	1.0
MAX.	137.0	40.0	4.06	.40	1.50	1.9	1.76	2.2
MIN.	92.0	40.0	-2.23	.35	.53	-2.4	.38	-1.0
REAL RMSE	.40	TRUE SD	1.87	SEPARATION	4.68	ITEM	RELIABILITY	.96
MODEL RMSE	.38	TRUE SD	1.87	SEPARATION	4.89	ITEM	RELIABILITY	.96
S.E. OF ITEM MEAN	= .58							

JADUAL 3. Nilai Kebolehpercayaan dan Pengasingan Responden terhadap Instrumen i-SPPBA

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	36.6	12.0		2.92 .74				
S.D.	5.9	.0		2.95 .24				
MAX.	48.0	12.0		10.71 1.96				
MIN.	23.0	12.0		-2.94 .59	.12	-2.4	.10	-2.3
REAL RMSE	.85	TRUE SD	2.83	SEPARATION	3.31	PERSON RELIABILITY	.92	
MODEL RMSE	.78	TRUE SD	2.85	SEPARATION	3.65	PERSON RELIABILITY	.93	
S.E. OF PERSON MEAN = .47								

Polariti Item

Analisis polariti item (*Point Measure Correlation- PTMEA CORR*) merupakan pengesahan awal atau analisis asas kesahan konstruk yang menunjukkan item bagi suatu instrumen itu bergerak ke arah penilaian sub-konstruk tunggal (Bond & Fox, 2015). Menurut Linacre (2005), pengukuran yang menunjukkan indeks positif bermaksud item mengukur konstruk yang ingin diukur dan ia berfungsi secara selari untuk mengukur konstruk. Namun, item tersebut perlu diteliti semula sekiranya ia mempunyai indeks yang negatif, sama ada perlu dikelaskan dan dimurnikan, atau digugurkan. Selain itu, Nunnally dan Bernstein (1994) menyatakan bahawa nilai PTMEA CORR yang kurang daripada 0.30 menunjukkan item mengendur.

Jadual 4 menunjukkan nilai polariti item bagi instrumen i-SPPBA. Nilai PTMEA CORR minimum bagi item instrumen i-SPPBA yang direkodkan adalah (.62) manakala nilai PTMEA CORR maksimum adalah (.82). Hasil analisis menunjukkan kesemua item melebihi nilai 0.30 dan tiada item yang menunjukkan nilai negatif. Justeru, didapati bahawa item-item bagi instrumen i-SPPBA ini bergerak pada satu arah yang ditetapkan dan dipercayai berupaya untuk membezakan sikap terhadap pembelajaran bahasa arab yang dimiliki oleh responden.

JADUAL 4. Nilai Polariti Item

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	TOTAL MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	EXACT EXP.	MATCH OBS%	MATCH EXP%	ITEM
8	127	40	-.70	.39	1.40	1.5	1.41	.9	.62	.74	61.5	77.5	D1.8
3	128	40	-.85	.39	1.03	.2	.91	.0	.71	.73	82.1	77.3	D1.3
5	106	40	2.30	.36	1.50	1.9	1.69	2.2	.72	.79	69.2	74.6	D1.5
6	137	40	-2.23	.48	.67	-1.6	1.03	.3	.73	.69	92.3	77.7	D1.6
12	120	40	.34	.38	1.06	.3	1.76	1.9	.74	.76	76.9	76.3	D1.12
4	128	40	-.85	.39	1.12	.6	1.14	.5	.75	.73	76.9	77.3	D1.4
2	133	40	-1.61	.39	.91	-.3	.86	.0	.76	.71	74.4	76.3	D1.2
1	135	40	-1.92	.39	.53	-2.4	.38	-1.0	.76	.70	87.2	76.7	D1.1
7	130	40	-1.15	.39	.77	-1.0	.65	-.6	.79	.72	79.5	77.0	D1.7
9	92	40	4.06	.35	.95	-.2	.99	.1	.79	.80	76.9	71.4	D1.9
11	123	40	-.10	.39	.67	-1.4	.62	-1.0	.81	.75	84.6	77.3	D1.11
10	103	40	2.69	.36	1.00	.1	1.13	.6	.82	.79	71.8	73.6	D1.10
MEAN	121.8	40.0	.00	.38	.97	-.2	1.05	.3			77.8	76.1	
S.D.	13.5	.0	1.91	.01	.28	1.2	.40	1.0			7.9	1.8	

Kesesuaian Item (*Item Fit*)

Analisis menggunakan model pengukuran Rasch menganggarkan darjah kesesuaian item yang mengukur sesuatu pemboleh ubah terpendam. Semakan kesesuaian item bermula dengan memastikan nilai *Mean-Square* (MNSQ) item berada dalam julat 0.5 hingga 1.5. Item dengan nilai yang kurang daripada 0.5 menunjukkan pertindihan domain dengan item yang lain manakala item dengan nilai yang lebih tinggi daripada 1.5 adalah tidak homogen dengan item-item lain dalam satu skala pengukuran. Lazimnya, sesuatu MNSQ yang terkeluar daripada julat akan menunjukkan nilai z-Std yang tinggi dan melebihi julat yang diterima iaitu $-2.0 < ZSTD < +2.0$.

Menurut Bond et al. (2021), Statistik MNSQ *infit* yang menyimpang biasanya akan menyebabkan lebih keimbangan berbanding MNSQ *outfit* yang terkeluar dari julat kesesuaian. MNSQ *outfit* berfungsi untuk membantu mengesan respons kepada tahap kesukaran item. Justeru, MNSQ *outfit* perlu diberi keutamaan dalam menentukan kesesuaian item yang mengukur sesuatu konstruk.

Berdasarkan Jadual 5, terdapat 3 item yang terkeluar daripada julat MNSQ yang diterima. Item-item yang melebih nilai MNSQ 1.5 adalah D1.12 (1.76) dan D1.5 (1.69), manakala item yang kurang daripada nilai 0.5 adalah D1.1 (0.38). Oleh itu, item-item ini perlu dipertimbangkan sama ada perlu dimurnikan atau digugurkan. Walau bagaimanapun, Boone et. al (2014) telah mencadangkan untuk mengenal pasti item yang tidak sesuai berdasarkan tiga kriteria (MNSQ, ZSTD dan PTMEA CORR), maka item D1.12 dan D1.1 memenuhi dua daripada tiga kriteria tersebut. Manakala item D1.5 memenuhi satu kriteria iaitu PTMEA CORR (0.72). Sumintono dan Widhiarso (2015) berpendapat bahawa apabila sesuatu item tersebut melengkapi sekurang-kurangnya satu daripada tiga kriteria, maka ia perlu dikekalkan. Begitu juga dengan penyataan Abdul Aziz et. al (2014), iaitu suatu item tersebut hanya dikira tidak sesuai sekiranya ia tidak memenuhi julat kesesuaian bagi ketiga-tiga kriteria. Oleh itu, ketiga-tiga item ini telah dikekalkan dan dimurnikan untuk digunakan dalam kajian sebenar.

JADUAL 5. Senarai Ketidaksesuaian Item (*Item Misfit*)

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MODEL MEASURE	INFIT S.E.	MNSQ ZSTD	OUTFIT MNSQ ZSTD	PT-MEASURE CORR.	EXACT EXP.	MATCH OBS% EXP%	ITEM
12	120	40	.34	.38 1.06	.3 1.76	1.9 A	.74	.76	76.9	D1.12
5	106	40	2.30	.36 1.50	1.9 1.69	2.2 B	.72	.79	69.2	D1.5
8	127	40	-.70	.39 1.40	1.5 1.41	.9 C	.62	.74	61.5	D1.8
4	128	40	-.85	.39 1.12	.6 1.14	.5 D	.75	.73	76.9	D1.4
10	103	40	2.69	.36 1.00	.1 1.13	.6 E	.82	.79	71.8	D1.10
6	137	40	-2.23	.40 .67	-1.6 1.03	.3 F	.73	.69	92.3	D1.6
3	128	40	-.85	.39 1.03	.2 .91	.0 f	.71	.73	82.1	D1.3
9	92	40	4.06	.35 .95	-.2 .99	.1 e	.79	.80	76.9	D1.9
2	133	40	-1.61	.39 .91	-.3 .86	.0 d	.76	.71	74.4	D1.2
7	130	40	-1.15	.39 .77	-1.0 .65	-.6 c	.79	.72	79.5	D1.7
11	123	40	-.10	.39 .67	-1.4 .62	-1.0 b	.81	.75	84.6	D1.11
1	135	40	-1.92	.39 .53	-2.4 .38	-1.0 a	.76	.70	87.2	D1.1

Nilai Korelasi Residual Terpiawai

Ujian penentuan piawai residual korelasi dilakukan bertujuan untuk mengenal pasti sekiranya terdapat dua item yang bertindih dan mengelirukan. Apabila terdapat dua item dengan nilai korelasi residual yang tinggi, maka ia menunjukkan bahawa item-item itu tidak bebas, dengan alasan kerana kedua-dua item tersebut memiliki ciri yang sama atau item-item tersebut berkongsi beberapa dimensi lain. Linacre (2005) meletakkan nilai 0.7

sebagai nilai korelasi residual maksimum. Sekiranya terdapat pasangan item dengan nilai melebihi 0.7, maka salah satu daripada item tersebut perlu digugurkan kerana ia mempunyai nilai korelasi yang tinggi. Jadual 6 menunjukkan bahawa sepuluh padanan item dalam instrumen i-SPPBA ini mempunyai nilai korelasi residual terpiawai yang menepati syarat kebebasan setempat, iaitu antara -0.31 sehingga 0.35. Justeru, kesemua item dipercayai tidak berkait antara satu sama lain dan tiada pasangan item yang dilihat mengelirukan oleh responden.

JADUAL 6. Nilai Korelasi Terpiawai Terbesar pada Item

CORREL- ATION	ENTRY NUMBER	ITEM	ENTRY NUMBER	ITEM
.35	6	D1.6	8	D1.8
.35	3	D1.3	4	D1.4

-.54	2	D1.2	8	D1.8
-.49	7	D1.7	10	D1.10
-.43	5	D1.5	6	D1.6
-.40	4	D1.4	11	D1.11
-.37	4	D1.4	9	D1.9
-.33	3	D1.3	5	D1.5
-.32	6	D1.6	9	D1.9
-.31	4	D1.4	5	D1.5

Keseragaman Dimensi

Salah satu aspek yang kritikal bagi mencapai dan mengukuhkan objektif instrumen untuk mengukur sesuatu adalah keseragaman dimensi. Sesuatu instrumen boleh memberikan hasil yang mengelirukan apabila mempunyai ciri kesamaran dalam proses pengukuran (Azrilah et al., 2013). Teknik *Residual Principal Component Analysis* (PCA) telah digunakan dalam analisis Rasch bagi mengesan kebolehan sesuatu instrumen mengukur dalam satu dimensi yang seragam dengan tahap gangguan (*noise*) yang boleh diterima.

Sebaik-baik nilai varians yang digariskan oleh Linacre (2002) adalah >60%. Jadual 7 menunjukkan bahawa varians kasar yang diterangkan oleh PCA bagi instrumen i-SPPBA adalah melebihi nilai 60% iaitu sebanyak 66.9%. Manakala tahap gangguan item dalam kontras satu yang direkodkan oleh instrumen i-SPPBA adalah setinggi 7.0% di mana nilai tersebut adalah terkawal baik dan berada di bawah nilai siling yang digariskan oleh Fisher (2007) iaitu 15%.

JADUAL 7. Keseragaman Dimensi Instrumen i-SPPBA

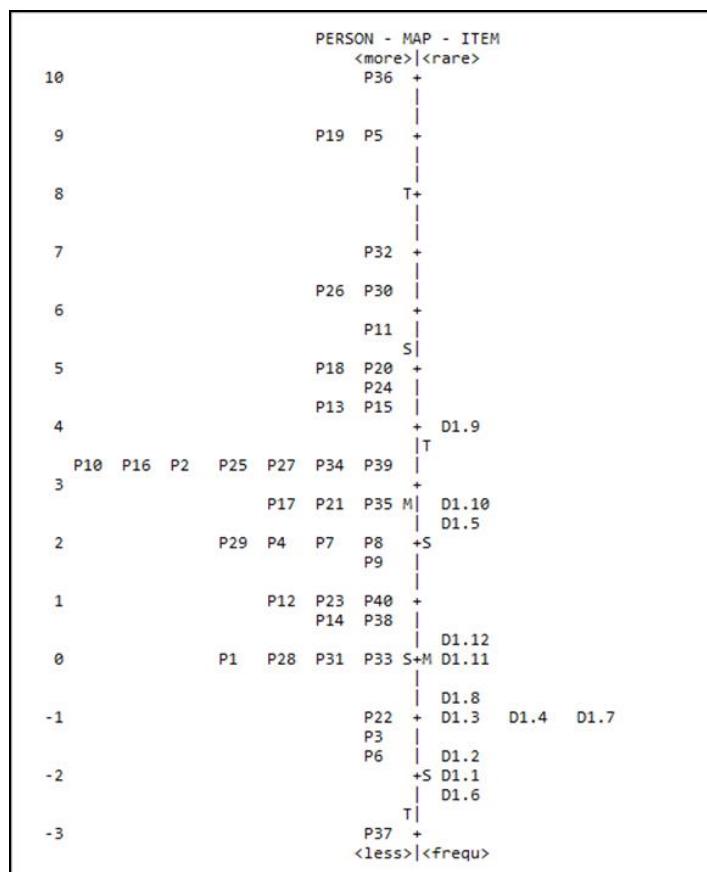
		-- Empirical --	Modeled
Total raw variance in observations	=	36.3 100.0%	100.0%
Raw variance explained by measures	=	24.3 66.9%	66.1%
Raw variance explained by persons	=	14.3 39.5%	39.0%
Raw Variance explained by items	=	10.0 27.5%	27.1%
Raw unexplained variance (total)	=	12.0 33.1% 100.0%	33.9%
Unexplnd variance in 1st contrast	=	2.5 7.0%	21.0%
Unexplnd variance in 2nd contrast	=	1.8 4.9%	14.9%
Unexplnd variance in 3rd contrast	=	1.6 4.5%	13.6%
Unexplnd variance in 4th contrast	=	1.4 3.8%	11.6%
Unexplnd variance in 5th contrast	=	1.1 3.1%	9.4%

Peta Item-Responden

Peta item-Responden menunjukkan taburan dalam skala ukur logit bagi responden yang terletak di sebelah kiri peta manakala di sebelah kanan peta menunjukkan taburan item yang (Bond & Fox, 2015). Tujuan peta ini adalah untuk memperlihatkan kaitan di antara aras kesukaran item dengan keupayaan responden (Rafidah & Mohd Effendi, 2019). Skala yang digunakan dalam peta ini ialah skala menurun, iaitu skala dengan item yang paling sukar dipersejui berada di atas diikuti dengan skala yang berada di bawah iaitu item yang paling mudah dipersejui.

Rajah 1 membentangkan peta pembolehubah yang menunjukkan taburan responden dan item dalam skala ukuran logit. Berdasarkan analisis daripada peta pembolehubah, nilai min keupayaan responden adalah lebih tinggi daripada nilai min kesukaran item bermakna instrumen i-SPPBA adalah sesuai dan boleh dijawab oleh responden. Secara purata, pelajar diukur mempunyai kebolehan dengan ukuran +2.61 logit iaitu P17, P21 dan P35. Selain itu, terdapat tiga orang pelajar (P19, P5 dan P36) melepas sempadan atas T (dua sisi piawai), yang menunjukkan bahawa pelajar-pelajar ini mempunyai kebolehan lebih tinggi yang berbeza berbanding dengan yang lain. Manakala hanya seorang sahaja pelajar melepas sempadan bawah T iaitu P37 dengan nilai ukuran -2.94, yang menunjukkan bahawa pelajar tersebut mempunyai kebolehan paling rendah dalam kalangan pelajar lain. Berdasarkan sebelah kanan peta pula, item yang paling sukar dipersejui oleh responden adalah item D1.9 (Saya sering bertutur dalam bahasa Arab di mana-mana sahaja) dengan ukuran +4.06 logit. Manakala item yang paling mudah dipersejui oleh responden adalah item D1.6 (Saya mendapat banyak pengetahuan ketika belajar bahasa Arab), iaitu dengan ukuran -2.23 logit.

RAJAH 1. Peta Item-Responden



KESIMPULAN

Secara keseluruhannya, dapatlah dirumuskan bahawa proses kesahan dan kebolehpercayaan pelu dijalankan terhadap setiap instrumen baru yang dibangunkan agar dapat membentuk satu instrumen yang sah dan tepat bagi mendapatkan satu pengukuran yang jitu. Justeru, kajian ini diharapkan dapat menyumbang serta dapat dijadikan panduan kepada tenaga pengajar atau para penyelidik yang berhasrat untuk membina instrumen penilaian dan pengukuran yang sah dan jitu.

Berdasarkan model pengukuran Rasch, kefungsian setiap item instrumen i-SPPBA boleh dianalisis secara mendalam untuk menentukan kesahan dan kebolehpercayaannya. Hasil pemeriksaan menunjukkan bahawa kesemua item i-SPPBA telah disahkan serta adil untuk mengukur sikap pelajar terhadap pembelajaran bahasa arab. Secara statistiknya, hasil penyelidikan ini telah berjaya membuktikan bahawa tahap kesahan dan kebolehpercayaan bagi item-item i-SPPBA ini adalah tinggi serta wajar untuk digunakan terhadap mahasiswa di peringkat institusi pengajian tinggi.

PENGHARGAAN

Penghargaan kepada Fakulti Pengajian Islam, Universiti Kebangsaan Malaysia di atas tajaan penyelidikan ini (GGPM-2021-016 Geran Galakan Penyelidik Muda (GGPM): Model Penguasaan Kosa Kata Bahasa Arab di Peringkat Universiti Awam Malaysia).

RUJUKAN

- Azizi Yahaya, Jamaluddin Ramli & Yusof Boon. (2007). Sumbangan Sikap terhadap Pencapaian Pelajar Mata Pelajaran Matematik: Sejauh Manakah Hubungan Ini Relevan? Universiti Teknologi Malaysia Institutional Repository Universiti Teknologi Malaysia Institutional Repository. <http://eprints.utm.my/id/eprint/2355/>
- Azizi Yahya, Peter Voo, Ismail Maakip & Mohd Dahlan A. Malek. (2017). *Kaedah Penyelidikan dalam Pendidikan*. Tanjung Malim: Universiti Pendidikan Sultan Idris.
- Azman Che Mat. (2013). Sikap Pelajar terhadap Pembelajaran Bahasa Arab di IPTA. *Academia Journal UiTMT*. 2(2), 61–68. <http://journals-academiauitmt.uitm.edu.my/>
- Azrilah Abdul Aziz, Mohd Saidfudin Masodi & Azami Zaharim. (2013). *Asas Model Pengukuran Rasch: Pembentukan Skala & Struktur Pengukuran*. Bangi: Penerbit Universiti Kebangsaan Malaysia.
- Bond, T. G. & Fox, C. M. (2015). *Applying the Rasch Model Fundamental Measurement in the Human Sciences*. London: Routledge.
- Bond, T. G., Yan, Z. & Heene, M. (2021). *Applying The Rasch Model Fundamental Measurement in the Human Sciences*. London: Routledge.
- Boone, W. J., Staver, J. R. & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. Netherlands: Springer. <https://doi.org/10.1007/978-94-007-6857-4>
- Chalak, A., & Kassaian, Z. (2010). Motivation and Attitudes of Iranian Undergraduate EFL Students towards Learning English. *GEMA Online Journal of Language Studies*. 10(2), 37–56.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Fisher, J. W. P. (2007). Rating Scale Instrument Quality Criteria. *Rasch Measurement Transactions*. 21(1), 1095.
- Gardner, R. & Lambert, W. (1972). *Attitudes and Motivation in Second Language Learning*. London: Newbury House.

- Kaseh Abu Bakar, Hakim Zainal, Nadzirah Norudin& Mohamad Azwan Kamarudin. (2021). Aplikasi Model Rasch bagi Kesahan Item Kosa Kata Kerja Quranic Vocabulary Test (QVT). *International Journal of West Asian Studies*. 13, 259–287.
- Ku Fatahiyah Ku Azizan, Awatif Abd Rahman & Hairun Najwah Jamali. (2014). Tahap Kemahiran Asas Bahsa Arab di IPT Malaysia Kaedah P&P Bahasa Arab di Malaysia. *GSE e-Journal of Education*. 1(2), 20–32. <https://doi.org/10.13140/RG.2.2.13788.44164>
- Lawshe, C. H. (1975). A Quantitative Approach to Content Validity. *Personnel Psychology*. 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Linacre, J. (2002). What Do Infit, Outfit, Mean-Square, and Standardization Mean? *Archives of Rasch Measurement*. 16, 871–882.
- Linacre, J. M. (2005). *A User's Guide to Winsteps/Ministeps Rasch Model Programs*. San Diego: MESA Press.
- Lynn, M. R. (1986). Determination and Quantification of Content Validity. *Nursing Research*. 35(6), 382–385.
- Majdi Ibrahim & Akmal Khuzaairy Abd Rahman. (2018). Teaching of Arabic in Malaysia. *Intellectual Discourse*, 26(1), 189–206.
- Mohamad Jafre Zainol Abidin, Majid Pour-Mohammadi & Hanan Alzwari. (2012). EFL Students' Attitudes towards Learning English Language: The Case of Libyan Secondary School Students. *Asian Social Science*. 8(2), 119–134. <https://doi.org/10.5539/ass.v8n2p119>
- Mohammad Rahim Kamaluddin & Wan Shahrazad Wan Sulaiman. (2019). *Teknik Kesahan dan Kebolehpercayaan Alat Ujian Psikologi*. Bangi: Penerbit Universiti Kebangsaan Malaysia.
- Mohd Effendi @ Ewan Mohd Matore, Hisyamsani Idris, Normawati Abdul Rahman & Ahmad Zamri Khairani. (2017). Kesahan Kandungan Pakar Instrumen IKBAR Bagi Pengukuran AQ Menggunakan Nisbah Kesahan Kandungan. Proceeding of International Conference on Global Education V (ICGE V), 10–11 April, Padang, Indonesia, 979–997.
- Noels, K. A., Pelletier, L. G., Clément, R. & Vallerand, R. J. (2000). Why are You Learning a Second Language? Motivational Orientations and Self-Determination Theory. *Language Learning*. 50(1), 57–85. <https://doi.org/10.1111/0023-8333.00111>
- Norkumalasari Othman, Nor Hasnida Che Md Ghazali, Mohd Nazir Md Zabit & Nazaruddin Abdul Hadi. (2020). Penerokaan dan Pengesahan Instrumen Sikap terhadap Matematik dalam Kalangan Guru Matematik Sekolah Menengah. *Lum & Pengajaran Asia Pasifik*. 8(4), 1-11.
- Nunnally, J. C. & Bernstein, I. H. (1994). The Assessment of Reliability. *Psychometric Theory*. 3, 248–292.
- Nurain Syafina Husain & Norhayuza Mohamad. (2020). Saiz Kosa Kata Bahasa Arab dalam Kalangan Pelajar Ijazah Sarjana Muda Bahasa Arab Komunikasi Professional, UiTM. *International Journal of Modern Languages and Applied Linguistics*. 4(2), 38.
- Rafidah Mohd Adnan & Mohd Effendi. (2019). Kesahan dan Kebolehpercayaan Instrumen I-CGPKM Menggunakan Model Rasch. *Journal of Quality Measurement and Analysis*. 15(1), 1–14.
- Robbins, P. (1986). Changing Attitudes of Teachers, Principals and Students: What We Didn't Measure. Reports - Research / Technical No. 143. <https://eric.ed.gov/?id=ED270415>
- Stegmann, Silke. (2013). Motivation and Attitudes towards Second Language Learning at Primary Schools; A Comparison of Teaching Programmes. Unpublished Master's Thesis, Utrecht University.
- Stroud, C. & Wee, L. (2006). Anxiety and Identity in the Language Classroom. *RELC Journal*. 37(3), 299–307. <https://doi.org/10.1177/0033688206071311>

- Wan Emril Nizar. (2018). Pembinaan, Kesahan dan Kebolehpercayaan Soal Selidik Amalan Adab Belajar Imam Syafie. *Journal of Sciences and Management Research.* 3.
https://www.widad.edu.my/download.php?f=Journal_Vol3_Oct2018_01.pdf

BIODATA PENGARANG

Nur Afiqah Athirah Mohd Rushdi, pelajar sarjana Pengajian Bahasa Arab dan Tamadun Islam, Fakulti Pengajian Islam, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia.

Lily Hanefarezan Asbulah, pensyarah kanan di Pusat Kajian Bahasa Arab dan Tamadun Islam, Fakulti Pengajian Islam, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia. Bidang penyelidikan beliau adalah termasuk Pengajaran Bahasa Arab sebagai bahasa Asing, Pengujian Bahasa, Linguistik Arab, Pengukuran dan Penilaian Sains Sosial.