

# **Analisis *Multi-Rater* Pelajaran Pendidikan Islam pada ujian Pentaksiran Tingkatan Tiga (PT3) di Malaysia<sup>1</sup>**

Nurul Nadia binti Abd Latib, Shahrir bin Jamaluddin & Bambang Sumintono<sup>2</sup>

Fakulti Pendidikan, Universiti Malaya

## **Abstrak**

Sejak tahun 2014, Malaysia menerapkan sistem ujian baru secara nasional di kelas 9 (*form 3*) yang bernama Pentaksiran Tingkatan Tiga (PT3) dengan menggunakan jenis soal essay saja. Studi ini meneliti tentang analisis ujian pada satu set soal pada pelajaran Pendidikan Islam yang menerapkan penilai berganda (*multi rater*) pada hasil kerja siswa. *Many facet rasch model* digunakan untuk analisis satu set soal yang terdiri dari 29 aitem, yang dikerjakan oleh 39 peserta tes dan dinilai oleh dua penilai. Info reliabilitas aitem dan separasinya menunjukkan tingkatan yang bagus; keragaman tingkat kesulitan soal dan proporsi di tiap kelompok kesulitan dan komponen ujian menunjukkan hal yang bagus. Properti psikometrik soal juga menunjukkan mayoritas aitem yang digunakan dalam ujian mempunyai tingkat kesesuaian yang bagus.

Keywords: Many Facet Rasch Model; Pentaksiran Tingkatan Tiga PT3; Pendidikan Islam Malaysia; Sekolah Menengah Malaysia

## **Pendahuluan**

Malaysia yang terdiri dari semenanjung Malaya dan bagian Utara Pulau Kalimantan (Borneo) mendapatkan kemerdekaan tahun 1957, telah mengalami kemajuan pesat sejak dekade 1990an. Dengan pendapatan perkapita telah mencapai US\$ 10.000 (sejak tahun 2012) serta angka kemiskinan 3% dan pertumbuhan ekonominya di atas 4% (Anonym, 2018) menunjukkan posisinya sebagai negara penghasil menengah (*middle income country*). Terdapat dua buah '*landmark*' modernnya yaitu menara kembar petronas dan ibukota administratif Putrajaya menjadi objek wisata dan kebanggaan negara, serta menjadi contoh modernitas negara muslim (60% penduduk Malaysia beragama Islam). Dalam bidang pendidikan, Malaysia telah sukses menempatkan satu universitasnya dalam ranking 100 universitas terbaik dunia menurut versi QS pada tahun 2018 (QS Top Universities, 2018), yaitu Universiti Malaya di peringkat 87; dan aktif bersaing merebut pasar mahasiswa asing (*international students*) dengan negara maju lainnya.

Sistem pendidikan Malaysia dilakukan secara terpusat, walaupun faktanya Malaysia adalah negara dengan sistem federal. Total jumlah populasi siswa sekolah adalah lima juta orang yang bersekolah pada 10 ribu lebih sekolah (lebih dari 95%-nya adalah sekolah

---

<sup>1</sup> Makalah dipresentasikan dalam 1<sup>st</sup> National Conference on Educational Assessment and Policy (NCEAP), Puspendik, Balitbang, Kemdikbud, Jakarta. Kamis, 28 Juni 2018, Hotel Le Meridien, Jakarta, Indonesia.

<sup>2</sup> Corresponding author, email: bambang@um.edu.my

negeri/publik). Populasi guru di Malaysia sebanyak 423 ribu orang dimana 70%-nya adalah guru perempuan (KPM, 2018). Kualifikasi minimal guru di Malaysia adalah sarjana (S1); pendidikan guru untuk sekolah dasar dilaksanakan oleh institut keguruan (disebut maktab) yang diawasi langsung oleh kementerian pendidikan; sedangkan untuk guru sekolah menengah pendidikannya dilaksanakan oleh 13 fakultas pendidikan di berbagai universitas. Semua pendidikan guru adalah lembaga publik/negeri yang dikontrol penuh oleh kementerian pendidikan dalam hal masukan dan keluarannya.

Sebagai negara bekas jajahan Inggris, Malaysia juga mengadopsi sistem pendidikan Inggris. Sistem sekolahnya terbagi dalam dua bagian besar, yaitu pendidikan dasar selama enam tahun yang dimulai pada usia tujuh tahun dengan diakhiri oleh ujian nasional pada tahun ke-6 (disebut Ujian Penilaian Sekolah Rendah atau UPSR). Pendidikan menengah dilaksanakan satu atap yang terbagi dua, yaitu sekolah menengah rendah selama tiga tahun yang diakhiri dengan satu ujian akhir (Pentaksiran Tingkatan Tiga atau PT3), kemudian pada tingkatan berikutnya selama dua tahun studi yang diselesaikan melalui ujian nasional lainnya (Sijil Pelajaran Malaysia atau SPM); ini tidak lain adalah O level (*ordinary level*) dalam sistem Inggris. Masa 11 tahun persekolahan ini adalah pendidikan wajib di Malaysia.

Makalah ini membahas tentang analisis ujian PT3 pada satu set soal pelajaran Pendidikan Agama Islam. Penilaian hasil ujian siswa dilakukan dengan penilai berganda (*multi rater*) berdasar panduan rubrik penskoran; data yang sama kemudian dianalisis dengan pemodelan rasch untuk mengetahui lebih detail tentang hasil ujian PT3 ini. Bagian berikutnya akan menjelaskan tentang konteks ujian PT3, dilanjutkan dengan metodologi, kemudian hasil dan pembahasan serta diakhiri dengan kesimpulan.

### **Ujian Pentaksiran Tingkatan Tiga (PT3)**

Pada tahun 2014 berdasar hasil penelitian mendalam oleh Kementerian Pendidikan Malaysia, ujian pada tahun 9 (*form 3*) yang bernama Penilaian Menengah Rendah (atau PMR) diakhiri. Sebagai gantinya Pemerintah Malaysia mengubah penilaian tersebut dengan nama Pentaksiran Berasaskan Sekolah (PBS), yaitu sekolah lah yang lebih berperan aktif menilai prestasi siswa, dimana prestasi kemajuan siswa dicatat dan dilaporkan berdasarkan empat komponen utama yaitu Pentaksiran sekolah (Ps), Pentaksiran pusat (Pp), Pentaksiran aktiviti jasmani, sukan dan kokurikulum (PAJSK) serta Pentaksiran psikometrik (PPSi).

Lembaga Peperiksaan Malaysia (LPM) yang merupakan lembaga yang berwenang dalam ujian nasional (*public examination*) di Malaysia, mendefinisikan ujian sekolah sebagai komponen utama dalam “menguatkan pembelajaran murid serta meningkatkan kualiti pengajaran guru kerana ianya mampu memberi maklum balas dan maklumat yang sah tentang keberkesanan proses pembelajaran”. Pentaksiran Pusat (ujian nasional) sebagai salah satu komponen dari PBS, merupakan ujian sumatif untuk menilai prestasi akademik dalam bentuk yang disebut Pentaksiran Tingkatan Tiga (PT3). Berbeda dengan PMR yang menggunakan jenis soal pilihan ganda (*multiple choices*), semua mata pelajaran dalam PT3 menggunakan

jenis soal *essay*, dimana peserta ujian harus menemukannya dengan menulis jawaban secara manual (*paper-pencil test*).

Ujian pada pelajaran Pendidikan Islam Menengah Rendah adalah khusus untuk siswa yang beragama Islam, sedangkan non-muslim mengambil pelajaran lainnya yaitu Pendidikan Moral (Sumintono, Lokman & Mohd Anuar, 2012). Ujian pelajaran Pendidikan Islam bertujuan untuk mengukur pemahaman dan kemampuan siswa terhadap topik-topik yang telah dipelajari dari tahun 7 sampai 9 (*form 1 to form 3*). Ujian ini dilaksanakan dengan pelajaran lainnya yang terlibat dalam Pentaksiran Pusat pada bulan Oktober setiap tahunnya, serta dilaksanakan oleh pihak sekolah masing-masing.

Instrumen ujian (soal-soal) telah disediakan oleh Lembaga Peperiksaan Malaysia dan pihak sekolah dikehendaki untuk mengunduh atau memuat turun (*download*) dokumen ujian mengikut jadwal yang telah disediakan, untuk kemudian dicetak sendiri. Aitem soal-soal ujian Pendidikan Islam terdiri dari tujuh komponen berikut:

Tabel 1. Komponen ujian pelajaran Pendidikan Islam PT3

No	Komponen	Persentase	Kode Soal
1	Tilawah al-Quran	10%	T
2	Hadis	10%	H
3	Akidah	20%	AK
4	Ibadah	20%	IB
5	Sirah	10%	S
6	Adab	20%	AD
7	Hafazan & Jawi	10%	HAF & Jawi
<b>Jumlah</b>		100%	

Dalam satu set ujian Pendidikan Islam PT3 terdiri dari 29 soal, dimana 40% soalnya di tiap komponen mengandung soal berjenis kemampuan berpikir tingkat tinggi atau *Higher Order Thinking Skills* (HOTS) yang dalam bahasa Malaysia disebut *Kemampuan Berpikir Aras Tinggi* (KBAT). Hasil ujian siswa kemudian diperiksa oleh dua orang guru di sekolah yang sama berdasar rubrik yang sudah disiapkan oleh Lembaga Peperiksaan Malaysia.

## Metodologi

### Data

Ujian PT3 Pendidikan Agama Islam yang dilakukan di Malaysia terdiri dari tiga set soal dan setiap sekolah mengunduh satu set soal. Dalam paper ini, soal yang dianalisis disebut sebagai Set-1 didapat dari satu sekolah menengah yang berada di daerah Wilayah Persekutuan Kuala Lumpur. Soal Set-1 terdiri dari 29 soal yang dikerjakan oleh 39 orang siswa, yang kemudian dinilai oleh dua orang guru, total data adalah  $29 \times 39 \times 2 = 2262$ , serta tidak terdapat data hilang dalam analisis.

### Penyiapan Analisis

Hasil penilaian oleh dua orang guru yang dilakukan dengan penskoran secara manual, kemudian dilakukan *data entry* ke dalam program Microsoft Excel. Saat yang sama disiapkan program analisis untuk perangkat lunak (*software*) Facets versi 3.7.2. yang melakukan analisis penilai ganda (*multi rater*) dengan pendekatan pemodelan rasch (Rasch Model). Program dan data kemudian disatukan untuk dianalisis.

### Many Facet Rasch Model (MFRM)

Pengukuran objektif (*objective measurement*) yang diinisiasi oleh Georg Rasch pada tahun 1960, telah banyak berkembang dan memberikan hasil pengukuran yang lebih untuk tes kognitif maupun non-kognitif (Olsen, 2003). Pemodelan Rasch (atau *Rasch Model*) tidak lain dari mengubah data mentah yang berjenis ordinal (*ordinal type*) dengan melihat pada *odd-probability*-nya untuk kemudian digunakan fungsi logaritma, sehingga menghasilkan data pengukuran yang bersifat interval sama yang disebut dengan logit (*log odds unit*) (Boone, Yale & Staver, 2014; Englehard, 2013; Sumintono & Widhiarso, 2015). Logit adalah data interval dengan kualitas yang lebih bagus hasil dari transformasi probabilitas pemodelan rasch (Bond & Fox, 2015).

Model analisis dikotomi dari Georg Rasch, kemudian berkembang ke bentuk lainnya, seperti *rating scale model* (RSM) oleh David Andrich (Andrich, 1988), *partial credit model* (PCM) dari Geoff Masters, maupun *many facets rasch measurement* (MFRM) yang dikembangkan oleh Linacre (Englehard, 2013; Boone et al., 2014).

Model analisis dari MFRM tidak lain analisis untuk penilai berganda/*multi rater* (lebih dari satu orang); yang melakukan analisis dari interaksi aitem soal, jawaban siswa dan penilaian guru. Tingkat kesulitan dari setiap aspek dikalibrasi berdasarkan ketajaman (C) penilai (j); ketajaman ini dianggap sebagai estimasi probabilitas dari siswa (n) yang memberikan jawaban pada aitem soal (i) untuk kategori ambang (k) untuk penilai (j), dengan persamaannya adalah:

$$P_{nikj} = \frac{e^{(\beta_n - \delta_i - F_k - C_j)}}{1 + e^{(\beta_n - \delta_i - F_k - C_j)}}$$

yaitu fungsi dari kemampuan siswa ( $\beta_n$ ), tingkat kesulitan aitem soal ( $\delta_i$ ), tingkat kesulitan ambang ( $F_k$ ), dan ketajaman penilai ( $C_j$ ) (Boone, 2014; Englehard, 2013).

## Hasil dan Pembahasan

### *Analisis secara manual*

Sebagai perbandingan untuk analisis dengan MFRM, dua orang guru yang menilai 39 peserta ujian pelajatan Pendidikan Islam PT3, hasilnya berdasar skor adalah:

Tabel 2. Hasil ujian PT3 Pendidikan Islam secara manual

No	Total Skor	Jumlah
1	80%-90%	12
2	70%-79%	13
3	60%-69%	10
4	50%-59%	3
5	40%-49%	1
<b>Jumlah</b>		39

Rata-rata skor dari seluruh peserta adalah 72,7, dimana 12 orang siswa (30%) berada pada total skor 80-90%; sedangkan 13 siswa (33%) prestasinya ada di rentang 70-79%; terdapat 25% (10 siswa) berada dalam rentang 60-69%; sisanya sebanyak 4 orang (10%) prestasinya dibawah 60%.

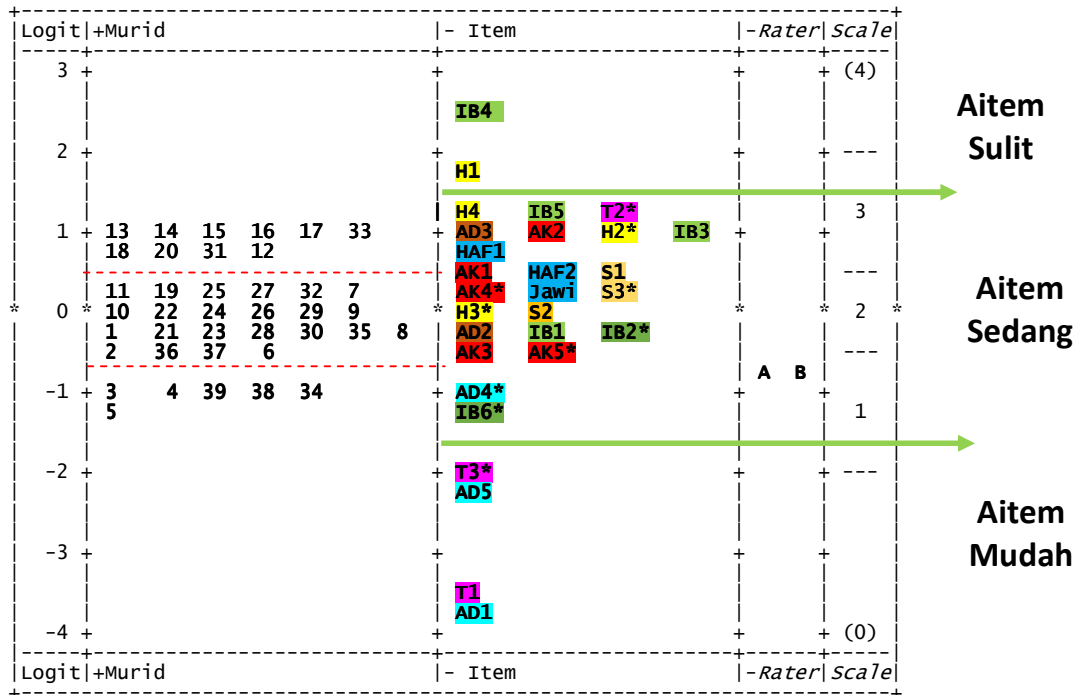
### *Analisis MFRM*

Hasil analisis reliabilitas dengan MFRM menunjukkan tingkat yang bagus dalam aspek aitem soal dan peserta tes seperti pada Tabel 3 di bawah ini. Tingkat separasi menunjukkan bahwa untuk siswa terbagi dalam tiga kelompok kemampuan secara kognitif, sedangkan tingkat kesulitan soal lebih dari tiga kelompok tingkat kesulitan yang menurut Fisher (2007), kedua info di atas menunjukkan tingkat reliabilitas yang baik.

Tabel 3. Reliabilitas ujian PT3 Pendidikan Islam dengan MFRM

Facet	Reliabilitas	Separasi
Aitem	0,98	7,47
Siswa	0,90	3,00

*Wright map* (peta aitem-siswa-penilai), dari Set-1 terdapat pada grafik di bawah ini. Kolom pertama menunjukkan skala logit, kolom kedua menunjukkan prestasi siswa dalam ujian Pendidikan Islam PT3 dimana posisi di atas menunjukkan prestasi lebih bagus; kolom ketiga menunjukkan tingkat kesulitan aitem, soal IB4 (ibadah) adalah soal tersulit, sedangkan AD1 (adab) adalah soal yang paling mudah; kolom keempat adalah penilai (rater), dimana ada dua orang (kode A dan B), dan keduanya termasuk dalam kategori *lenient rater* (logit negatif).



Grafik 1. Peta Aitem-Siswa-Penilai Pendidikan Islam PT3 Set-1

Keterangan: kode soal dengan tanda '\*' menunjukkan soal jenis HOTS (KBAT)

Pengelompokan prestasi siswa pada ujian Set-1, terbagi tiga dengan menggunakan nilai rata-rata logit siswa (0,0 logit) dan deviasi standarnya (0,50). Siswa yang dianggap lebih mampu mempunyai nilai di atas +0,50 logit sebanyak 10 orang atau 26% (bagian atas kolom ke dua di Grafik 1); kelompok kemampuan sedang berada diantara +0,5 logit sampai ke -0,50 logit jumlahnya 23 orang (59%); dan yang kemampuan rendah dengan nilai dibawah -0,50 logit ada enam orang (15%).

Berdasar data hasil analisis, pengelompokan tingkat kesulitan soal juga dapat dibagi tiga dengan menggunakan nilai rata-rata logit aitem (0,00 logit) dan deviasi standarnya (1,35). Sehingga soal yang kategori sulit dengan nilai lebih dari +1,35 logit terdapat dua aitem (7%) yaitu soal IB4 dan H1 yang secara relatif tidak bisa dikerjakan dengan benar oleh semua peserta ujian; soal kategori kesulitan sedang berada antara +1,35 logit sampai kepada -1,35 logit sebanyak 23 soal (79%); dan soal kategori mudah dengan nilai kurang dari -1,35 logit ada empat soal (14%) (lihat Grafik 1) yang menunjukkan semua peserta bisa mengerjakan dengan benar soal T3, AD5, T1 dan AD1.

Hasil analisis juga menunjukkan bahwa 10 soal (34%) dengan jenis HOTS (atau KBAT), terdapat satu aitem yang masuk kategori mudah (T3\*), dan sembilan soal lainnya termasuk kategori kesulitan sedang. Hal ini mengindikasikan bahwa soal yang menguji kemampuan berpikir tingkat tinggi sebagian besarnya bisa dikerjakan dengan benar oleh

banyak peserta ujian. Tidak terdapat soal jenis HOTs yang tidak bisa dikerjakan oleh peserta ujian.

Tabel 4 tentang kesesuaian statistik di bawah ini menunjukkan atribut psikometri secara keseluruhan dari soal Set-1 pada ujian Pendidikan Islam PT3. Berdasar kriteria fit statistik dari Boone et al., (2014), maka didapati empat item (14%) yang mempunyai kecenderungan tidak fit, yaitu soal T2\*, H2, IB1 dan AD2 dimana nilai *outfit mean-square* dan *outfit zstd*-nya berada di luar *range* kesesuaian statistik.

Tabel 4. Properti psikometrik soal-soal Set-1 Pendidikan Islam PT3

No.	Kode Aitem	Measure	Model S.E.	Infit		Outfit		Correlation	
				MnSq	ZStd	MnSq	ZStd	PtMea	PtExp
1	T1	-3.45	.41	1.23	.6	1.14	.4	.15	.14
<b>2</b>	<b>T2*</b>	<b>1.33</b>	<b>.12</b>	<b>.33</b>	<b>-6.4</b>	<b>.33</b>	<b>-6.4</b>	<b>.36</b>	<b>.43</b>
3	T3*	-1.56	.19	.72	-1.5	.66	-1.8	.62	.30
4	H1	1.35	.12	.58	-3.4	.58	-3.4	.32	.43
<b>5</b>	<b>H2*</b>	<b>.96</b>	<b>.12</b>	<b>.33</b>	<b>-6.2</b>	<b>.34</b>	<b>-6.2</b>	<b>.28</b>	<b>.43</b>
6	H3*	-.08	.13	1.16	1.0	1.16	1.0	.58	.41
7	H4	1.17	.12	.49	-4.3	.51	-4.1	.36	.43
8	AK1	.50	.12	.67	-2.4	.67	-2.4	.42	.43
9	AK2	1.10	.12	1.52	3.1	1.52	3.1	.33	.43
10	AK3	-.41	.14	1.76	4.0	1.79	4.1	.29	.39
11	AK4*	.20	.13	1.51	2.9	1.51	2.9	.57	.42
12	AK5*	-.47	.14	1.47	2.6	1.37	2.1	.64	.39
<b>13</b>	<b>IB1</b>	<b>-.22</b>	<b>.13</b>	<b>.47</b>	<b>-4.2</b>	<b>.43</b>	<b>-4.7</b>	<b>.22</b>	<b>.40</b>
14	IB2*	-.13	.13	1.11	.7	1.10	.7	.76	.40
15	IB3	1.10	.12	.99	.0	.98	.0	.62	.43
16	IB4	2.43	.15	.86	-.8	1.00	.0	.38	.36
17	IB5	1.17	.12	1.12	.8	1.13	.8	-.11	.43
18	IB6*	-1.27	.17	1.24	1.3	1.09	.5	.51	.33
19	S1	.55	.12	.97	-.1	.96	-.2	.38	.43
20	S2	.12	.13	.89	-.7	.90	-.6	.07	.42
21	S3*	.23	.13	2.00	5.2	1.91	4.8	.62	.42
22	AD1	-3.86	.50	1.89	1.5	.94	.0	.28	.12
<b>23</b>	<b>AD2</b>	<b>-.33</b>	<b>.14</b>	<b>.50</b>	<b>-3.9</b>	<b>.48</b>	<b>-4.1</b>	<b>.16</b>	<b>.39</b>
24	AD3	.89	.12	1.36	2.2	1.37	2.2	.20	.43
25	AD4*	-1.02	.16	1.51	2.7	1.49	2.5	.32	.35
26	AD5*	-1.63	.19	.66	-1.9	.63	-1.9	.49	.29
27	HAF1	.73	.12	1.31	1.9	1.29	1.8	.38	.43
28	HAG2	.44	.12	1.12	.8	1.07	.4	.41	.43
29	Jawi	.19	.13	.83	-1.1	.86	-.8	.38	.42

Keterangan: \* soal jenis HOTs

Soal cetak *tebal-miring*, tidak memenuhi kriteria kesesuaian statistik

## Kesimpulan

Hasil analisis ujian Set-1 Pendidikan Islam PT3 dengan MFRM menunjukkan hasil yang menarik dan lebih lengkap yang bisa memberikan informasi lebih tepat tentang hasil ujian dan kualitas soal. Penilaian *multirater* yang dimaksudkan untuk menghindari bias bisa dieksekusi dengan lebih tepat melalui MFRM dimana interaksi peserta tes, aitem soal dan penilai dikalibrasi dalam suatu skala *equal-interval*. Hasil analisis menunjukkan bahwa tingkat

kesulitan soal Set-1 sangat beragam (separasi 7,47) yang menunjukkan soal dapat memetakan tingkat kemampuan peserta tes pada berbagai komponen ujian PT3 pelajaran Pendidikan Islam. Perbandingan tingkat kesulitan soal sukar-sedang-mudah menunjukkan hal yang proporsional. Aitem soal yang berjenis kemampuan berpikir tingkat tinggi (HOTs atau KBAT), menunjukkan bisa diselesaikan oleh peserta tes dan bukan soal yang paling sukar untuk dikerjakan. Dari Aspek kesesuaian statistic didapati bahwa gterdapat empat soal (14%) yang mempunyai kecenderungan tidak fit pada atribut psikometriknya.

### **Daftar Pustaka**

- Andrich, D. (1988). *Rasch Model for Measurement*. (series: Quantitative Application in the Socials Sciences). Newburry Park, California: Sage Publication.
- Anonym. (2018). Malaysia. The Worldfact Book – CIA. Available at; <https://www.cia.gov/library/publications/the-world-factbook/geos/my.html>
- Bond, T.G. dan Fox, C.M. (2015). *Applying The Rasch Model, Fundamentals Measurement in the Human Sciences* (3rd edition). New York: Routledge.
- Boone, W.J., Staver, J.R., dan Yale, M.S. (2014). *Rasch Analysis in the Human Sciences*. Dordrecht: Springer.
- Engelhard, G., Jr. (2013). *Invariant Measurement, using rasch models in the social, behavioral and health sciences*. New York: Routledge.
- Fisher, W.P. Jr. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions*, 21(1) hal. 1095 <http://www.rasch.org/rmt/rmt211m.htm>
- KPM. (2018). Educational Management Information System. Available at; <http://emisportal.moe.gov.my/>
- Olsen, L. W. (2003). Essays on Georg Rasch and his contribution to statistics. Unpblished PhD thesis, Institute of Economics, University of Copenhagen. <http://www.rasch.org/olsen.htm>
- QS Top Universities, 2018. World Universitiy Rankings 2019. Online accessed 8 June 2018. URL <https://www.topuniversities.com/university-rankings/world-university-rankings/2019>
- Sumintono, B., Lokman Mohd Tahir, Mohd Anuar Abdul Rahman. (2012). Pendidikan Moral di Malaysia: Tantangan dan Implementasi Pendidikan Karakter di Sekolah. *Jurnal Pendidikan Karakter*, vol 2, no 1. pp 14-22. [available at: <http://eprints.utm.my/24588/>]
- Sumintono, B dan Widhiarso, W. (2014). Aplikasi Pemodelan Rasch pada Asesment Pendidikan. Cimahi: Trimkomunikata.



Lampiran: Contoh Soal Ujian Pendidikan Islam PT3

**1. Soal IB4: Aitem paling sulit**

(ii) Nyatakan **dua** cara beristinjak menurut syariat Islam.

1. ....

2. ....

[2 markah]

[Lihat halaman sebelah  
**TERHAD**

8

**2. Soal AD5 – aitem HOTs (KBAT) yang paling mudah**

(e) Terdapat anak yang menghantar ibu bapa ke pusat penjagaan harian orang tua dengan alasan memudahkan urusan.  
Kemukakan pendapat dan hujah anda.

.....

.....

.....

.....

.....

.....

[4 markah]

**3. Soal T2\* - aitem yang tidak fit**

(b) Berdasarkan ayat di atas, bagaimanakah makanan yang halal memberi kesan terhadap tubuh badan?

.....

.....

.....

[2 markah]