



INTERNATIONAL JOURNAL OF  
EDUCATION, PSYCHOLOGY  
AND COUNSELLING  
(IJEPC)

<https://gaexcellence.com/ijepe>



**ANALISIS DIFFERENTIAL ITEM FUNCTIONING (DIF)  
INSTRUMEN KEPERIBADIAN ISLAM REMAJA (IKIR)  
MERENTAS JANTINA**

*DIFFERENTIAL ITEM FUNCTIONING (DIF) ANALYSIS OF THE ISLAMIC  
ADOLESCENT PERSONALITY INSTRUMENT (IKIR) ACROSS GENDER*

Norhayati Ahmad<sup>1\*</sup>, Zaidi Fadzil<sup>2</sup>, Nurul Huda Ishak<sup>3</sup>

<sup>1</sup> Kuliyah Kaunseling, Pendidikan dan Sains Sosial (KAUSAS), UniSHAMS, Malaysia

 [hayyun66@gmail.com](mailto:hayyun66@gmail.com)

 <https://orcid.org/0009-0004-6064-2811>

<sup>2</sup> Kuliyah Kaunseling, Pendidikan dan Sains Sosial (KAUSAS), UniSHAMS, Malaysia

 [zaidifadzil@unishams.edu.my](mailto:zaidifadzil@unishams.edu.my)

 <https://orcid.org/0009-0003-4620-691X>

<sup>3</sup> Kuliyah Kaunseling, Pendidikan dan Sains Sosial (KAUSAS), UniSHAMS, Malaysia

 [nurulhudaishak@unishams.edu.my](mailto:nurulhudaishak@unishams.edu.my)

 <https://orcid.org/0009-0007-2253-3931>

\*Corresponding Author

**Article Info:**

**Article history:**

Received date: 15.01.2026

Revised date: 09.02.2026

Accepted date: 12.03.2026

Published date: 26.03.2026

**To cite this document:**

Ahmad, N., Fadzil, Z., & Ishak, N. H. (2026). Analisis Differential Item Functioning (DIF) Instrumen Kepribadian Islam Remaja (Ikir) Merentas Jantina. *International Journal of Education, Psychology and Counselling*, 11(62), 1060-1071.

**Abstrak:**

Instrumen Kepribadian Islam Remaja (IKIR) dibangunkan berasaskan kerangka psikospiritual Islam yang mengintegrasikan dimensi iman, tanggungjawab khalifah dan akhlak sebagai manifestasi ihsan dalam pembentukan sahsiah remaja Muslim. Pengukuran kepribadian Islam memerlukan instrumen yang sah, boleh dipercayai dan bebas daripada bias pengukuran agar interpretasi skor dapat dibuat secara adil merentas kumpulan responden. Sehubungan itu, kajian ini bertujuan menilai kesetaraan pengukuran Instrumen Kepribadian Islam Remaja merentas jantina menggunakan Model Rasch melalui analisis *Differential Item Functioning* (DIF). Kajian ini menggunakan pendekatan kuantitatif dengan reka bentuk tinjauan. Seramai 130 orang pelajar sekolah menengah terlibat sebagai responden kajian yang terdiri daripada 52 pelajar lelaki dan 78 pelajar perempuan daripada pelbagai latar institusi pendidikan. Data dianalisis menggunakan pemodelan Rasch bagi menilai fungsi item serta mengesan kewujudan DIF jantina. Dapatan analisis menunjukkan bahawa majoriti item dalam IKIR berfungsi secara setara bagi responden lelaki dan perempuan, dengan hanya sebilangan kecil item menunjukkan perbezaan sederhana yang tidak menjejaskan kestabilan konstruk keseluruhan. Secara keseluruhan, keputusan kajian mengesahkan bahawa IKIR mempunyai tahap

kesetaraan pengukuran yang memadai dan boleh digunakan untuk menilai keperibadian Islam remaja tanpa memerlukan pelarasan tafsiran berdasarkan jantina. Instrumen ini berpotensi diaplikasikan dalam konteks pendidikan dan kaunseling Islam sebagai alat penilaian sahsiah remaja yang sistematik, adil dan berasaskan bukti empirik.

DOI: 10.35631/IJEPC.1162062

### Kata Kunci:

Differential Item Functioning, Kesetaraan Pengukuran, Keperibadian Islam Remaja, Model Rasch, Psikometrik Islam

### Abstract:

The Islamic Personality Instrument for Adolescents (IKIR) was developed based on an Islamic psycho-spiritual framework that integrates the dimensions of faith (*iman*), the responsibility of humans as *khalifah*, and moral conduct (*akhlaq*) as manifestations of *ihsan* in the development of Muslim adolescents' character. Measuring Islamic personality requires an instrument that is valid, reliable, and free from measurement bias so that score interpretations can be made fairly across different respondent groups. Therefore, this study aims to examine the measurement equivalence of the Islamic Personality Instrument for Adolescents across gender using the Rasch Model through Differential Item Functioning (DIF) analysis. This study employed a quantitative survey design. A total of 130 secondary school students participated as respondents, consisting of 52 male students and 78 female students from various educational institutions. The data were analysed using Rasch modelling to evaluate item functioning and to detect the presence of gender-related DIF. The findings indicate that the majority of items in IKIR function equivalently across male and female respondents, with only a small number of items showing moderate differences that do not affect the stability of the overall constructs. Overall, the results confirm that IKIR demonstrates adequate measurement equivalence and can be used to assess the Islamic personality of adolescents without requiring gender-based adjustments in score interpretation. The instrument has the potential to be applied in Islamic education and counselling contexts as a systematic, fair, and evidence-based tool for assessing adolescent character development.

### Keyword:

Differential Item Functioning, Islamic Personality, Islamic Psychometrics, Rasch Model, Measurement Equivalence



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [ijepec@gaexcellence.com](mailto:ijepec@gaexcellence.com)

## Pengenalan

Pembinaan dan pengukuran keperibadian Islam dalam kalangan remaja merupakan agenda penting dalam pendidikan Islam kontemporari, khususnya bagi melahirkan generasi Muslim yang seimbang dari segi rohani, emosi, intelektual dan akhlak. Golongan remaja berada pada

fasa perkembangan kritikal yang melibatkan pembentukan identiti, sistem nilai serta hala tuju kehidupan, sekali gus menjadikan tempoh ini sangat signifikan dalam pembentukan sahsiah dan orientasi hidup beragama (Erikson, 1968; Santrock, 2019). Dalam konteks Islam, penilaian keperibadian remaja bukan sekadar berasaskan tingkah laku luaran, malah merangkumi pembentukan dalaman yang berpaksikan keimanan, penyucian jiwa, penghayatan ilmu dan kesedaran terhadap peranan manusia sebagai hamba serta khalifah Allah SWT (al-Ghazali, 2004; Rafii, 2021).

Keperibadian Islam dalam kajian ini dioperasionalkan sebagai tiga konstruk laten yang saling berkaitan tetapi berbeza secara konseptual, iaitu konstruk Hamba Cemerlang (orientasi tauhid dan keyakinan dalaman), Khalifah Cemerlang (kesedaran peranan dan tanggungjawab sosial) serta Akhlak Cemerlang (manifestasi tingkah laku berasaskan nilai Islam). Ketiga-tiga dimensi ini dimodelkan secara reflektif dalam kerangka psikometrik, iaitu item-item instrumen dianggap sebagai indikator kepada trait laten yang mendasari respon remaja. Pendekatan ini membolehkan konsep normatif tazkiyah al-nafs diterjemahkan kepada struktur pengukuran empirikal yang boleh diuji dari sudut kesahan dan kebolehpercayaan.

Secara teoretikal, prinsip iman, khalifah yang bertanggungjawab dan akhlak adalah tuntutan syariat yang bersifat universal dan merentas jantina. Namun demikian, literatur religiositi menunjukkan bahawa ekspresi dan tahap penghayatan agama boleh berbeza antara lelaki dan perempuan bergantung kepada konteks sosial dan budaya (Miller & Hoffmann, 1995; Francis, 1997). Oleh itu, walaupun konstruk keperibadian Islam diasaskan kepada prinsip sejagat, pengujian kesetaraan pengukuran merentas jantina adalah perlu bagi memastikan struktur instrumen benar-benar beroperasi secara setara bagi kedua-dua kumpulan.

Pengukuran keperibadian Islam dalam kalangan remaja turut menimbulkan cabaran metodologi yang signifikan, khususnya berkaitan aspek kesahan dan keadilan instrumen. Instrumen yang tidak diuji daripada sudut kesetaraan pengukuran berisiko menghasilkan inferens yang tidak adil apabila digunakan merentas kumpulan demografi yang berbeza (Zumbo, 2007; Millsap, 2011). Dalam konteks pendidikan Islam, isu ini menjadi lebih kritikal kerana penilaian keperibadian sering dijadikan asas kepada keputusan pendidikan, intervensi kaunseling dan pembangunan sahsiah remaja.

Isu bias jantina dalam pengukuran tidak bertujuan membandingkan kelebihan antara lelaki dan perempuan, sebaliknya menilai sama ada item dalam instrumen berfungsi secara adil selepas mengambil kira tahap trait yang setara. Prinsip ini selari dengan konsep keadilan (al-'adl) dalam Islam yang menuntut penilaian dilaksanakan secara objektif dan bebas daripada prasangka (al-Qaradawi, 2001). Oleh itu, pengukuran yang adil memerlukan pendekatan yang mampu mengesan potensi bias secara empirik.

Model Rasch menawarkan kerangka pengukuran yang menekankan prinsip *specific objectivity*, iaitu pemisahan parameter item daripada kebolehan responden (Rasch, 1960; Bond & Fox, 2015). Salah satu kekuatan utama model ini ialah analisis Differential Item Functioning (DIF), yang membolehkan pengesanan item yang menunjukkan perbezaan tahap kesukaran antara kumpulan walaupun memiliki tahap trait yang sama (Tennant & Conaghan, 2007). Pengujian DIF adalah kritikal kerana kewujudan item yang bias boleh menjejaskan kesahan konstruk serta menghasilkan keputusan yang tidak adil dalam perbandingan jantina.

Sehubungan itu, artikel ini bertujuan menilai kesetaraan fungsi item Instrumen Keperibadian Islam Remaja (IKIR) merentas jantina melalui analisis DIF menggunakan Model Rasch. Kajian ini bukan sahaja menyumbang kepada pembangunan instrumen psikometrik Islam yang sah dan boleh dipercayai, malah mengukuhkan prinsip keadilan pengukuran selaras dengan tuntutan metodologi moden dan nilai epistemologi Islam.

Kajian ini mempunyai beberapa keunikan berbanding kajian terdahulu. Pertama, kajian ini mengintegrasikan kerangka psikospiritual Islam dengan pendekatan psikometrik moden melalui penggunaan Model Rasch bagi menilai kesetaraan fungsi item instrumen keperibadian Islam. Kedua, kajian ini memberikan bukti empirik mengenai kesetaraan pengukuran konstruk berasaskan nilai Islam merentas jantina, yang jarang dibincangkan dalam kajian pembangunan instrumen religiositi. Ketiga, kajian ini menyumbang kepada pembangunan instrumen keperibadian Islam yang bukan sahaja sah dan reliabel, tetapi juga adil dari sudut pengukuran, selaras dengan prinsip keadilan (*al-'adl*) dalam epistemologi Islam.

### Sorotan Literatur

Isu bias jantina dalam pengukuran psikometrik telah lama menjadi fokus dalam penyelidikan pendidikan dan sains sosial, khususnya bagi konstruk yang melibatkan nilai, sikap dan keperibadian (Zumbo, 2007; Millsap, 2011). Instrumen yang tidak memenuhi prinsip kesetaraan pengukuran berpotensi menghasilkan perbezaan skor yang disalah tafsir sebagai perbezaan trait sebenar, sedangkan ia sebenarnya berpunca daripada fungsi item yang tidak setara.

Dalam konteks religiositi dan keperibadian berasaskan nilai, beberapa kajian menunjukkan kemungkinan perbezaan pola respons antara lelaki dan perempuan, khususnya dalam dimensi penghayatan dalaman dan ekspresi tingkah laku agama (Miller & Hoffmann, 1995; Francis, 1997). Walau bagaimanapun, perbezaan tersebut tidak semestinya mencerminkan perbezaan konstruk, sebaliknya mungkin berpunca daripada variasi interpretasi item atau pengaruh norma sosial. Oleh itu, analisis kesetaraan pengukuran penting bagi membezakan antara perbezaan trait sebenar dan bias instrumen.

Pendekatan berasaskan Teori Ujian Klasik (Classical Test Theory, CTT) mempunyai keterbatasan dalam menilai kesetaraan item kerana parameter item bergantung kepada sampel kajian (DeVellis, 2017). Sebaliknya, Model Rasch menyediakan estimasi parameter yang lebih stabil serta membolehkan pengujian Differential Item Functioning (DIF) dilakukan secara lebih objektif melalui perbandingan logit antara kumpulan responden (Bond & Fox, 2015; Boone et al., 2014).

Kesetaraan pengukuran merupakan syarat penting dalam memastikan interpretasi skor yang sah merentas kumpulan responden (Engelhard & Wind, 2018; Millsap, 2011). Model Rasch semakin diterima secara meluas dalam pembangunan instrumen psikometrik kerana keupayaannya menghasilkan ukuran yang bersifat invarian serta relatif bebas daripada pengaruh sampel kajian (Bond & Fox, 2015; Boone, 2020). Selain itu, pendekatan ini membolehkan penilaian kesetaraan pengukuran merentas kumpulan responden dilakukan secara lebih sistematik melalui analisis Differential Item Functioning (DIF) (Engelhard & Wind, 2018; Millsap, 2011).

Walaupun Model Rasch semakin meluas digunakan dalam kajian pendidikan dan kesihatan, penggunaannya dalam konteks pembangunan instrumen keperibadian Islam masih terhad. Kebanyakan kajian terdahulu lebih menumpukan kepada pembangunan konstruk dan kebolehpercayaan dalaman tanpa menilai kesetaraan fungsi item merentas jantina (Rafii, 2021). Kekurangan ini menunjukkan wujud jurang literatur dalam memastikan instrumen keperibadian Islam bukan sahaja sah dan reliabel, tetapi juga adil merentas kumpulan demografi utama.

Dari perspektif Islam, prinsip al-'adl (keadilan) menuntut agar penilaian terhadap manusia dilakukan secara objektif tanpa dipengaruhi ciri luaran seperti jantina. Oleh itu, pengujian DIF dalam kajian ini bukan sekadar prosedur statistik, tetapi juga satu tuntutan metodologi yang selaras dengan nilai keadilan dalam Islam.

Walaupun beberapa kajian terdahulu telah membincangkan pembangunan konstruk keperibadian Islam dalam kalangan remaja, kebanyakannya menumpukan kepada aspek kebolehpercayaan dalaman dan kesahan konstruk tanpa menilai kesetaraan fungsi item merentas kumpulan demografi seperti jantina. Kekurangan ini menyebabkan interpretasi perbandingan skor antara kumpulan berpotensi terdedah kepada bias pengukuran. Oleh itu, kajian ini mengisi jurang literatur dengan menilai kesetaraan fungsi item Instrumen Keperibadian Islam Remaja (IKIR) merentas jantina menggunakan Model Rasch melalui analisis Differential Item Functioning (DIF).

### ***Objektif Kajian***

Kajian ini bertujuan untuk menilai kesetaraan fungsi item Instrumen Keperibadian Islam Remaja (IKIR) merentas jantina menggunakan analisis *Differential Item Functioning* (DIF) berasaskan Model Rasch. Secara khusus, kajian ini mempunyai dua objektif utama:

1. Menilai kualiti pengukuran Instrumen Keperibadian Islam Remaja (IKIR) berdasarkan analisis kebolehpercayaan, indeks pemisahan dan padanan model menggunakan Model Rasch.
2. Mengenal pasti kewujudan *Differential Item Functioning* (DIF) bagi item IKIR merentas jantina bagi memastikan kesetaraan fungsi item dalam kalangan responden lelaki dan perempuan.

### ***Metodologi Kajian***

#### ***Reka Bentuk Kajian***

Kajian ini menggunakan reka bentuk kuantitatif dengan pendekatan pemodelan Rasch bagi menilai kualiti pengukuran serta kesetaraan fungsi item Instrumen Keperibadian Islam Remaja (IKIR) merentas jantina. Pendekatan ini dipilih kerana Model Rasch membolehkan penilaian empirikal yang sistematik terhadap kebolehpercayaan, kestabilan parameter item, padanan model serta keadilan pengukuran melalui prinsip *specific objectivity* (Rasch, 1960; Bond & Fox, 2015). Pendekatan ini semakin digunakan secara meluas dalam pembangunan instrumen pendidikan dan psikologi kerana keupayaannya menghasilkan ukuran yang bersifat invarian serta membolehkan penilaian kesetaraan item merentas kumpulan responden secara lebih objektif (Bond & Fox, 2015; Boone et al., 2014).

Secara khusus, analisis ini bertujuan menguji kewujudan *Differential Item Functioning* (DIF) jantina bagi memastikan item-item IKIR berfungsi secara setara bagi responden lelaki dan perempuan selepas mengambil kira tahap trait laten yang sama.

### ***Sampel Kajian***

Seramai 130 orang pelajar sekolah menengah terlibat dalam kajian ini. Daripada jumlah tersebut, 52 orang ialah responden lelaki (40.0%) dan 78 orang responden perempuan (60.0%). Responden dipilih daripada tiga jenis institusi pendidikan, iaitu sekolah menengah kebangsaan, sekolah menengah agama dan institusi tahfiz, bagi memastikan kepelbagaian latar pendidikan serta pengalaman pembelajaran Islam.

Walaupun saiz sampel ini bersifat sederhana, ia memenuhi keperluan minimum bagi analisis Model Rasch pada peringkat pembangunan dan pengesahan awal instrumen. Linacre (1994) menyatakan bahawa sampel antara 100 hingga 150 responden mampu menghasilkan anggaran parameter item yang stabil dalam lingkungan  $\pm 0.5$  logit pada tahap keyakinan 95%. Selain itu, setiap kumpulan jantina melepasi ambang minimum 30 responden yang lazim disarankan untuk analisis perbandingan kumpulan dalam konteks DIF (Boone et al., 2014).

Walau bagaimanapun, kajian lanjutan dengan sampel yang lebih besar dan lebih seimbang dicadangkan bagi meningkatkan kuasa statistik dan kebolehumuman dapatan.

### ***Instrumen Kajian***

Instrumen Keperibadian Islam Remaja (IKIR) dibangunkan berasaskan kerangka psikospiritual Islam yang menekankan pembangunan insan secara holistik merangkumi dimensi dalaman dan luaran. Pembinaan instrumen berpandukan sumber wahyu (al-Quran dan sunah), karya ulama klasik seperti Imam al-Ghazali serta literatur kontemporari dalam psikologi Islam.

IKIR mengandungi 66 item yang merangkumi enam subkonstruk dari tiga konstruk utama:

1. Konstruk Hamba Cemerlang
2. subkonstruk Pemantapan Iman
3. subkonstruk Penyucian Jiwa
4. Konstruk Khalifah Cemerlang
5. subkonstruk Penghayatan Ilmu
6. subkonstruk Insan Kamil
7. Konstruk Akhlak Cemerlang
8. subkonstruk Hubungan dengan Manusia
9. subkonstruk Hubungan dengan Alam

Item disusun dalam bentuk pernyataan berskala Likert lima mata bagi mengukur kecenderungan sikap, tingkah laku dan penghayatan nilai Islam dalam kalangan remaja. Sebelum analisis Rasch dijalankan, instrumen melalui proses pembangunan sistematik melibatkan penentuan konstruk, pembinaan item dan semakan kesahan kandungan oleh pakar (DeVellis, 2017).

### ***Prosedur Analisis Data***

Data dianalisis menggunakan perisian pemodelan Rasch. Analisis melibatkan penilaian:

- Kebolehpercayaan item dan responden
- Tahap pemisahan (separation index)
- Padanan model (infit dan outfit mean-square)
- Analisis Differential Item Functioning (DIF) merentas jantina

Bagi analisis DIF, dua kumpulan dibandingkan (lelaki dan perempuan) menggunakan ujian Rasch-Welch dan Mantel-Haenszel. Item ditafsir menunjukkan DIF ketara sekiranya memenuhi dua kriteria serentak:

1. Perbezaan logit (DIF contrast)  $\geq |0.50|$  sebagai ambang praktikal sederhana
2. Signifikan secara statistik pada aras  $p < 0.05$

Pendekatan ini membolehkan penilaian kesetaraan item secara objektif selepas mengawal tahap kebolehan responden.

### ***Dapatan Kajian***

Dapatan kajian dibentangkan berdasarkan analisis Model Rasch yang merangkumi penilaian kebolehpercayaan, indeks pemisahan dan padanan model bagi setiap subkonstruk IKIR. Seterusnya, analisis *Differential Item Functioning* (DIF) dijalankan untuk menilai kesetaraan fungsi item merentas jantina. Pendekatan ini bertujuan memastikan bahawa interpretasi skor adalah sah dan bebas daripada bias pengukuran.

**Jadual 1 Statistik Kebolehpercayaan dan Pemisahan Rasch Mengikut Subkonstruk IKIR**

| Subkonstruk             | Person Separation | Person Reliability | Person Infit MNSQ | Item Separation | Item Reliability | Item Infit MNSQ |
|-------------------------|-------------------|--------------------|-------------------|-----------------|------------------|-----------------|
| Pemantapan Iman         | 1.84              | .77                | 1.25              | 6.21            | .97              | .88             |
| Penyucian Jiwa          | 2.02              | .80                | 1.07              | 5.16            | .96              | .91             |
| Penghayatan Ilmu        | 1.97              | .80                | .96               | 6.35            | .98              | .98             |
| Insan Kamil             | 1.68              | .74                | .99               | 5.03            | .96              | 1.00            |
| Hubungan dengan Manusia | 2.20              | .83                | .99               | 3.82            | .94              | .97             |
| Hubungan dengan Alam    | 2.45              | .86                | .99               | 3.42            | .92              | .97             |

*Nota.* Nilai Infit MNSQ dalam julat 0.5–1.5 dianggap memadai (Bond & Fox, 2015).

Dari Jadual 1, analisis Model Rasch menunjukkan bahawa semua subkonstruk IKIR mempunyai tahap kebolehppercayaan yang memadai hingga tinggi.

Nilai *person reliability* berada dalam julat .74 hingga .86, manakala *item reliability* adalah tinggi bagi semua subkonstruk (.92 hingga .98), menunjukkan kestabilan hierarki kesukaran item.

Indeks pemisahan responden melebihi 1.5 bagi semua subkonstruk, menandakan keupayaan instrumen membezakan sekurang-kurangnya dua tahap kebolehan responden.

Nilai Infit MNSQ bagi item dan responden berada dalam julat boleh diterima (0.5–1.5), menunjukkan padanan model yang baik tanpa kewujudan item misfit yang serius.

Secara keseluruhan, dapatan ini mengesahkan bahawa struktur pengukuran IKIR adalah stabil dan sesuai untuk analisis DIF merentas jantina.

**Jadual 2 Ringkasan Analisis DIF Jantina Mengikut Subkonstruk IKIR**

| Konstruk                | Bil. Item | Item DIF Signifikan | Julat DIF Contrast | Kesimpulan       |
|-------------------------|-----------|---------------------|--------------------|------------------|
| Pemantapan Iman         | 11        | 0                   | -0.40 hingga 0.17  | Bebas DIF        |
| Penyucian Jiwa          | 11        | 5                   | -0.87 hingga 0.56  | DIF Sederhana    |
| Penghayatan Ilmu        | 11        | 0                   | -0.25 hingga 0.46  | Bebas DIF        |
| Insan Kamil             | 11        | 2                   | -0.87 hingga 0.36  | DIF Sederhana    |
| Hubungan dengan Manusia | 11        | 1*                  | -0.35 hingga 0.48  | Tiada DIF Ketara |
| Hubungan dengan Alam    | 11        | 1                   | -0.59 hingga 0.16  | DIF Sederhana    |

*Nota.* Item signifikan ditentukan berdasarkan  $|DIF\ contrast| \geq 0.50$  dan  $p < .05$ .

Seperti yang ditunjukkan dalam Jadual 2, tiga subkonstruk iaitu Pemantapan Iman, Penghayatan Ilmu dan Hubungan dengan Manusia tidak menunjukkan perbezaan signifikan dan bermakna DIF ketara merentas jantina.

Walau bagaimanapun, beberapa item dalam subkonstruk Penyucian Jiwa, Insan Kamil dan Hubungan dengan Alam menunjukkan perbezaan yang signifikan dan bermakna tahap kesukaran sederhana antara responden lelaki dan perempuan.

Item-item yang memenuhi kriteria DIF ketara dipaparkan dalam Jadual 3.

**Jadual 3 Item IKIR dengan DIF Jantina Signifikan**

| Konstruk             | Item | DIF Contrast | p    | Arah DIF  |
|----------------------|------|--------------|------|-----------|
| Penyucian Jiwa       | 12   | -0.87        | .000 | Perempuan |
| Penyucian Jiwa       | 13   | 0.56         | .014 | Lelaki    |
| Penyucian Jiwa       | 15   | -0.63        | .031 | Perempuan |
| Penyucian Jiwa       | 16   | -0.59        | .005 | Perempuan |
| Penyucian Jiwa       | 21   | -0.67        | .016 | Perempuan |
| Insan Kamil          | 41   | -0.87        | .001 | Perempuan |
| Insan Kamil          | 42   | -0.78        | .002 | Perempuan |
| Hubungan dengan Alam | 57   | -0.59        | .009 | Perempuan |

*Nota.* DIF dianggap signifikan apabila  $|DIF\ contrast| \geq 0.50$  dan  $p < .05$ . Arah DIF menunjukkan kumpulan yang itemnya lebih sukar.

Seperti yang dipaparkan dalam Jadual 3, hanya lapan daripada 66 item menunjukkan perbezaan DIF yang signifikan.

Walau bagaimanapun, tiada pola bias yang konsisten memihak kepada satu jantina secara sistematik pada tahap konstruk keseluruhan.

Dapatan ini menunjukkan bahawa majoriti item IKIR berfungsi secara setara merentas jantina, sekali gus menyokong kesetaraan pengukuran instrumen dalam kalangan remaja lelaki dan perempuan.

### Perbincangan

Kajian ini menilai kesetaraan fungsi item IKIR merentas jantina menggunakan analisis DIF berdasarkan Model Rasch. Secara keseluruhan, majoriti item didapati berfungsi secara setara bagi responden lelaki dan perempuan, dengan hanya sebilangan kecil item menunjukkan DIF sederhana. Dapatan ini menunjukkan bahawa IKIR mempunyai tahap kesetaraan pengukuran yang memadai untuk perbandingan skor merentas jantina.

Dari sudut psikometrik, fungsi item yang stabil menyokong prinsip *specific objectivity* dalam Model Rasch, yang menegaskan bahawa parameter item adalah bebas daripada ciri demografi responden apabila andaian model dipenuhi (Rasch, 1960; Bond & Fox, 2015). Walaupun beberapa item menunjukkan perbezaan logit sederhana, magnitudnya tidak membentuk pola bias yang sistematik terhadap satu jantina. Oleh itu, perbezaan skor keseluruhan lebih berkemungkinan mencerminkan variasi tahap keperibadian Islam sebenar berbanding kesan bias pengukuran.

Dapatan ini turut menyokong konsep *measurement invariance* yang menjadi prasyarat penting dalam perbandingan kumpulan secara sah (Millsap, 2011). Dalam konteks pengukuran keperibadian Islam, kesetaraan item merentas jantina adalah signifikan kerana konstruk yang diukur seperti iman, penyucian jiwa dan akhlak yang berasaskan prinsip spiritual dan tanggungjawab insan yang bersifat universal. Namun demikian, kewujudan beberapa item dengan DIF sederhana menunjukkan bahawa variasi pengalaman sosial atau interpretasi item mungkin mempengaruhi respons bagi item tertentu, dan hal ini wajar diteliti dalam kajian lanjutan.

Dari perspektif Islam, dapatan ini selari dengan prinsip keadilan (*al-'adl*) dalam penilaian insan. Islam menegaskan bahawa penilaian terhadap manusia hendaklah berasaskan ketakwaan dan kualiti amal, bukan ciri luaran seperti jantina (al-Hujurat 49:13). Oleh itu, pengujian dan pemastian kesetaraan item dalam IKIR bukan sahaja memenuhi tuntutan metodologi moden, malah mencerminkan keperluan etika dalam penyelidikan berteraskan Islam.

Selain itu, penggunaan Model Rasch membolehkan penilaian instrumen dilaksanakan secara lebih telus dan berasaskan bukti empirik berbanding pendekatan pengukuran tradisional (Boone et al., 2014; Tennant & Conaghan, 2007). Keupayaan model ini mengesan potensi bias item meningkatkan keyakinan terhadap kesahan dan kebolehgunaan IKIR dalam konteks pendidikan dan kaunseling remaja.

Walau bagaimanapun, kajian ini dijalankan dengan saiz sampel yang sederhana dan taburan jantina yang tidak sepenuhnya seimbang. Oleh itu, kajian lanjutan dengan sampel lebih besar dan pelbagai latar belakang dicadangkan bagi mengesahkan kestabilan dapatan DIF serta memperkukuh kebolehumuman instrumen.

## Kesimpulan

Kajian ini menunjukkan bahawa Instrumen Keperibadian Islam Remaja (IKIR) secara keseluruhan mempunyai tahap kesetaraan pengukuran yang memadai merentas jantina berdasarkan analisis *Differential Item Functioning* (DIF) menggunakan Model Rasch. Majoriti item didapati berfungsi secara setara bagi responden lelaki dan perempuan, manakala hanya sebilangan kecil item menunjukkan DIF sederhana. Walau bagaimanapun, magnitud perbezaan tersebut tidak menunjukkan pola bias yang sistematik terhadap mana-mana kumpulan. Oleh itu, dapatan ini menunjukkan bahawa IKIR mempunyai tahap keadilan pengukuran yang memadai dan boleh digunakan untuk perbandingan skor merentas jantina.

Item-item yang menunjukkan DIF tidak digugurkan secara automatik, sebaliknya dicadangkan untuk dinilai semula melalui semakan kandungan serta kajian lanjutan dengan sampel yang lebih besar bagi mengesahkan kestabilan dapatan. Pendekatan ini selaras dengan amalan pembangunan instrumen psikometrik yang menekankan keseimbangan antara pertimbangan statistik dan kesahan kandungan.

Dari sudut teori, kajian ini menyumbang kepada literatur psikometrik Islam dengan menyediakan bukti empirik bahawa konstruk keperibadian Islam dapat diukur secara setara merentas jantina melalui pendekatan pengukuran moden. Kajian ini turut mengisi jurang penyelidikan berkaitan penilaian bias jantina dalam instrumen berasaskan nilai Islam.

Dari sudut praktikal, IKIR berpotensi digunakan sebagai alat penilaian sahsiah remaja dalam konteks pendidikan dan kaunseling tanpa memerlukan pelarasan tafsiran berdasarkan jantina. Penggunaan instrumen yang telah diuji kesetaraan itemnya menyokong amalan penilaian yang lebih adil, telus dan beretika selaras dengan prinsip keadilan dalam Islam.

Secara keseluruhannya, dapatan kajian ini memperkukuh kedudukan IKIR sebagai instrumen pengukuran keperibadian Islam yang sah, stabil dan signifikan untuk menilai pembangunan sahsiah remaja Muslim secara empirikal dan holistik.

---

|                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Penghargaan:</b>                        | Penulis merakamkan setinggi-tinggi penghargaan kepada Universiti Islam Antarabangsa Sultan Abdul Halim Muad'zam Shah (UniSHAMS) atas sokongan dan kemudahan yang diberikan sepanjang kajian ini dijalankan.                                                                                                                                                                                                                                                       |
| <b>Penyataan<br/>Pembiayaan:</b>           | Kajian ini tidak menerima sebarang pembiayaan.                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Pernyataan Konflik<br/>Kepentingan:</b> | Penulis mengisytiharkan bahawa tiada konflik kepentingan berkaitan penerbitan kertas kerja ini. Semua penulis telah menyumbang kepada kajian ini dan telah meluluskan versi akhir manuskrip untuk penyerahan kepada International Journal of Education, Psychology and Counseling (IJEPC).                                                                                                                                                                        |
| <b>Pernyataan Etika:</b>                   | Kajian ini telah dijalankan selaras dengan piawaian etika penyelidikan. Persetujuan bermaklumat telah diperolehi daripada Kementerian Pelajaran Malaysia dan Pejabat Pendidikan Daerah sebelum pengumpulan data dijalankan. Penyertaan adalah secara sukarela, dan responden telah dijamin kerahsiaan serta anonimitas. Data yang dikumpulkan digunakan semata-mata bagi tujuan akademik.                                                                         |
| <b>Pernyataan<br/>Sumbangan Penulis:</b>   | Semua penulis telah menyumbang secara signifikan terhadap pembangunan manuskrip ini. [Penulis A] mengendalikan pengumpulan data, analisis serta tafsiran dapatan kajian. [Penulis B] bertanggungjawab terhadap pengkonsepan, metodologi dan penyeliaan keseluruhan kajian. [Penulis C] menyumbang kepada sorotan literatur, penyediaan draf, dan semakan kritikal manuskrip. Semua penulis telah membaca dan meluluskan versi akhir manuskrip sebelum penyerahan. |

---

## Rujukan

- Al-Ghazali. (2004). *Ihya' 'ulum al-din*. Dar al-Ma'rifah.
- Al-Munawi, A. R. (1994). *Fayd al-Qadir sharh al-Jami' al-Saghir*. Dar al-Kutub al-'Ilmiyyah.
- Al-Qaradawi, Y. (2001). *Fiqh al-awlawiyyat: Dirasah jadidah fi daw' al-Qur'an wa al-Sunnah*. Maktabah Wahbah.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.
- Boone, W. J. (2020). Rasch analysis for instrument development: Why, when, and how? *CBE—Life Sciences Education*, 19(3), 1–7. <https://doi.org/10.1187/cbe.19-12-0253>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage Publications.
- Erikson, E. H. (1968). *Identity: Youth and crisis*. Norton.
- Engelhard, G., & Wind, S. A. (2018). *Invariant measurement with raters and rating scales: Rasch models for rater-mediated assessments*. Routledge.
- Fisher, W. P., Jr. (2019). Rasch measurement applications in social science research. *Measurement: Interdisciplinary Research and Perspectives*, 17(2), 1–14. <https://doi.org/10.1080/15366367.2019.1588646>
- Francis, L. J. (1997). The psychology of gender differences in religion: A review of empirical research. *Religion*, 27(1), 81–96. <https://doi.org/10.1006/reli.1996.0066>
- Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, 7(4), 328.
- Miller, A. S., & Hoffmann, J. P. (1995). Risk and religion: An explanation of gender differences in religiosity. *Journal for the Scientific Study of Religion*, 34(1), 63–75. <https://doi.org/10.2307/1386523>
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. Routledge.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research.
- Santrock, J. W. (2019). *Adolescence* (17th ed.). McGraw-Hill Education.
- Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? *Arthritis & Rheumatism*, 57(8), 1358–1362. <https://doi.org/10.1002/art.23108>
- Wright, B. D., & Linacre, J. M. (2020). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 34(1), 1–2.
- Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4(2), 223–233. <https://doi.org/10.1080/15434300701375832>