



ANALISIS KUALITAS BUTIR SOAL DAN ABILITAS SISWA PADA ASESMEN MADRASAH IBTIDAIYAH DENGAN RASCH MODEL

Raudhatul Jannah

Magister Pendidikan Guru Madrasah Ibtidaiyah Universitas Islam Negeri Sunan
Kalijaga Yogyakarta, Yogyakarta, Indonesia

Contributor Email: 22204085018@student.uin-suka.ac.id

Received: June 18, 2023

Accepted: November 7, 2023

Published: March 30, 2024

Article Url: <https://ojsdikdas.kemdikbud.go.id/index.php/didaktika/article/view/1234>

Abstract

Item analysis is very important because it can help teachers to find out whether the measurement tools given to students can measure students' abilities. Some things that need to be analyzed from the questions and student learning outcomes are the suitability of the items, the level of difficulty, and also the abilities of students. So this article will discuss it. This research is a quantitative descriptive evaluative research. The research subjects were students' answers to questions on Madrasah Assessment in Social Sciences. Data analysis of student learning outcomes was carried out using the Winstep 3.73 application. The results of this study note that the IPS Assessment questions have compatibility with the fit characteristics because they can meet at least one of the three criteria that have been set. Then the questions have varying levels of difficulty because three questions fall into the category of too difficult and too easy. Furthermore, if seen from the student's ability to answer the questions, abnormal patterns were found, so the teacher needs to review the student's learning outcomes.

Keywords: *Item Analysis; Rasch Model; Madrasah Ibtidaiyah.*

Abstrak

Analisis soal merupakan hal yang sangat penting karena membantu guru untuk mengetahui apakah alat pengukuran yang diberikan kepada siswa benar mampu mengukur kemampuan siswa. Beberapa hal yang perlu dianalisis dari soal dan hasil belajar siswa adalah kesesuaian butir soal, tingkat kesukaran, dan abilitas siswa. Penelitian ini adalah penelitian evaluatif deskriptif kuantitatif. Subjek penelitian adalah jawaban siswa terhadap soal Asesmen Madrasah mata pelajaran Ilmu Pengetahuan Sosial. Analisis data hasil belajar siswa tersebut dilakukan dengan menggunakan aplikasi Winstep 3.73. Hasil penelitian ini menunjukkan bahwa soal asesmen Ilmu Pengetahuan Sosial memiliki kesesuaian dengan karakteristik fit karena mampu memenuhi minimal satu dari tiga kriteria yang sudah ditetapkan. Soal memiliki tingkat kesulitan yang bervariasi yaitu tiga soal masuk pada kategori terlalu sulit, dan tiga soal masuk kategori terlalu mudah. Dilihat dari abilitas siswa dalam menjawab soal ditemukan adanya pola-pola yang tidak normal sehingga guru perlu melakukan peninjauan kembali terhadap hasil belajar siswa.

Kata Kunci: *Analisis Butir Soal; Rasch Model; Madrasah Ibtidaiyah.*

A. Pendahuluan

Melakukan evaluasi adalah salah satu tugas guru sebagai pendidik dalam menyelenggarakan pembelajaran yang utuh (Haryanto & Amalia, 2020). Evaluasi pembelajaran bertujuan untuk mengetahui keefektifan dan efisiensi sistem pembelajaran serta membantu untuk mengidentifikasi kekuatan dan kelemahan yang dimiliki oleh peserta didik (Kusmiyati, 2022). Penyiapan instrumen dalam sebuah evaluasi tidak dapat dilakukan asal-asalan dan sembarangan. Jika instrumen evaluasi salah, secara otomatis hasil evaluasi tersebut juga salah. Salah satu jenis evaluasi adalah asesmen, dalam bukunya Russel dan Airasian mengungkapkan bahwa komponen yang penting dalam pembelajaran adalah asesmen (Sumarnia, 2019; Wardani et al., 2023). Oleh karena itu, dalam menyelenggarakan asesmen, guru membutuhkan alat. Salah satunya adalah tes. Agar dapat menghasilkan hasil asesmen yang baik, dibutuhkan tes yang baik yaitu tes yang memenuhi tiga syarat yaitu validitas isi, validitas konstruk dan reliabilitas yang baik.

Pemenuhan syarat instrumen tes ini terbentur dengan keadaan di madrasah yaitu belum semua guru mampu melakukan penyusunan

instrumen dengan baik (Nurhafidhah et al., 2021). Banyak instrumen tes yang belum diketahui kualitasnya sehingga tidak diketahui efektivitasnya dalam menilai kemampuan peserta didik atau istilahnya adalah terjadi penilaian semu ketika melakukan evaluasi. Untuk menginisiasi hal tersebut, melakukan analisis butir soal adalah upaya yang strategis bagi guru untuk mengetahui kualitas tesnya.

Salah satu alat analisis soal adalah model rasch. Model rasch adalah model analisis tes yang hadir sebagai solusi dari model sebelumnya yaitu teori tes klasik, Model ini mampu menghadirkan peta butir orang yang menunjukkan apakah tes yang dibuat mampu mengakomodasi berbagai karakteristik siswa di kelas (Widhiarso, 2016).

Berdasarkan pencarian yang dilakukan penulis, di *Publish or Perish*, diketahui bahwa publikasi penelitian mengenai analisis Rasch Model hanya ada enam penelitian. Dari keenam penelitian tersebut ada dua penelitian yang meneliti pada ranah sekolah dasar, sedangkan pada ranah madrasah ibtidaiyah belum ada sama sekali. Dengan demikian, jelas bahwa penelitian ini masih jarang dilakukan dan memiliki unsur kebaruan. Dilihat dari kontennya, penelitian sebelumnya rata-rata hanya menganalisis butir soal tanpa menyajikan abilitas siswa dalam mengerjakan soal tersebut. Dengan demikian, penelitian ini sangat penting agar guru mampu mengetahui kualitas soal yang disusunnya dan juga mampu mengukur kelebihan dan kelemahan peserta didiknya.

Berdasarkan paparan di atas, ditetapkan tujuan penelitian ini adalah menganalisis kualitas butir dan tingkat kesulitan soal yang diujikan pada asesmen madrasah dan meneliti abilitas siswa dalam mengerjakan soal tersebut.

B. Metode

Penelitian ini menggunakan pendekatan deskriptif kuantitatif dan bersifat evaluatif. Karena bersifat evaluatif, penelitian ini bukan bertujuan untuk membuktikan sebuah hipotesis, tetapi untuk menguraikan keadaan dalam lingkup penelitian dengan bantuan alat analisis data kuantitatif (Walidin et al., 2015).

Data yang dianalisis adalah soal mata pelajaran Ilmu Pengetahuan Sosial yang diujikan pada Asesmen Madrasah di Madrasah Ibtidaiyah. Soal ini diujikan kepada 32 siswa. Teknik analisis data yang digunakan adalah Rasch Model. Menurut Sumintono dan Widhiarso (dalam Ernawati, 2023). Rasch Model adalah alat analisis yang sangat berguna untuk menguji validitas, realibilitas instrumen, serta person dan item secara sekaligus. Rasch model telah memenuhi lima prinsip model pengukuran yaitu (a) mampu memberikan ukuran yang linier dengan interval yang sama, (b) dapat mengatasi data yang hilang, (c) bisa memberikan estimasi yang lebih tepat, (d) mampu mendeteksi ketidaktepatan model, dan (e) memberikan instrument pengukuran yang independen dari parameter yang diteliti.

C. Hasil dan Pembahasan

1. Analisis Kesesuaian Butir Soal dan Tingkat Kesulitan Soal

Analisis butir soal ini dilakukan pada soal Ilmu Pengetahuan Sosial dalam Asesmen Madrasah Ibtidaiyah. Hasil analisis data dapat diketahui bahwa butirnya tergolong *outliers atau misfit* dapat dilihat dari nilai *statistic outfit MNSQ (outliner-sensitive or information weighted fit Mean Square)*, *outfit ZSTD (z-standard)*, dan *PT-Measure Corr (point measure correlation)*. Menurut Sumintono dan Widhiarso (dalam Ernawati, 2023) syaratnya adalah sebagai berikut.

Nilai Outfit MNSQ: 0,5 - 1,5

Nilai Outfit ZSTD: -2 - 2

Pt- measure corr: harus bernilai positif

Ketiga kriteria tersebut jika terpenuhi salah satu saja, maka soal masih tergolong fit. Jika ketiganya tidak terpenuhi, tergolong misfit. Untuk mengetahui apakah butir-butir soal mata pelajaran IPS ini sudah memenuhi kriteria tersebut dapat dilihat dari tabel di bawah ini.

Tabel.1 Ringkasan Statistik Outfit MNZQ, Outfit ZSTD & PT-Measure Corr

Outfit MNSQ	Outfit ZSTD	PT-Measure Corr	Item
.72	-.3	.54	it28
.79	-.5	.52	it02
1.38	1.2	.27	it05
.80	-.6	.52	it12
1.74	2.4	.19	it33
.95	-.1	.44	it24
.88	-.4	.47	it23
1.65	2.3	.10	it04
1.28	1.1	.24	it16
1.21	.8	.28	it29
1.02	.2	.37	it03
.94	-.1	.41	it17
.75	-.8	.53	it31
1.15	.5	.29	it01
.95	-.1	.37	it06
.95	.0	.41	it40
.80	-.4	.46	it07
.64	-.9	.55	it08
.57	-1.1	.59	it39
.96	.1	.41	it11
.76	-.4	.45	it27
1.20	.5	.26	it09
.78	-.3	.43	it10
.79	-.3	.40	it13
1.31	.7	.13	it26
1.17	.5	.32	it15
.56	-.7	.50	it19
.97	.1	.31	it21
.94	.1	.31	it18
.45	-.8	.52	it34
.79	.0	.31	it14
.40	-.7	.49	it20
.75	-.1	.37	it22
.52	-.5	.46	it30

Outfit MNSQ	Outfit ZSTD	PT-Measure Corr	Item
1.27	.6	.19	it36
.32	-.9	.56	it37
.83	.1	.26	it25
.85	.1	.21	it35
1.84	1.1	.13	it38
.72	.1	.29	it32

Berdasarkan hasil analisis di atas diketahui bahwa dilihat dari nilai *Outfit MNSQ* dapat diketahui bahwa terdapat enam soal yang keluar dari ketentuan. Berdasar pada *Outfit ZSTD*, terdapat dua soal yang keluar dari ketentuan dan nilai *Pt-Measure corr* terlihat sesuai seluruhnya. Karena untuk mengetahui kesesuaian butir soal diperlukan minimal satu nilai terpenuhi, disimpulkan seluruh soal memenuhi kriteria atau seluruh item fit.

Selanjutnya adalah melihat bagaimana tiap butir soal memiliki karakteristik yang baik bagi peserta didik dalam memetakan peserta didik. Karakteristik butir soal dalam model Rasch ini adalah bagaimana tingkat kesulitan butir soal bagi peserta didik. Syarat item dikatakan memiliki karakter yang baik adalah jika indeks kesulitan (nilai measure) di antara -2,0 dan 2 (Hambleton & Swaminathan, 1985). Untuk mengetahuinya dapat dicermati dari tabel di bawah ini.

Tabel. 2 Ringkasan Nilai Measure

Entry Number	Total Score	Total Count	Measure	Model S.E
28	6	32	3.08	.53
2	10	32	2.18	.43
5	11	32	2.00	.42
12	12	32	1.82	.41
33	14	32	1.50	.40
24	15	32	1.34	.40
23	16	32	1.18	.39
4	17	32	1.03	.39
16	18	32	.87	.40
29	18	32	.87	.40
3	19	32	.71	.40

Entry Number	Total Score	Total Count	Measure	Model S.E
17	19	32	.71	.40
31	20	32	.55	.41
1	21	32	.38	.41
6	21	32	.38	.41
40	21	32	.38	.41
7	23	32	.03	.43
8	23	32	.03	.43
39	23	32	.03	.43
11	24	32	-.17	.45
27	24	32	-.17	.45
9	25	32	-.38	.47
10	25	32	-.38	.47
13	25	32	-.38	.47
26	25	32	-.38	.47
15	26	32	-.61	.49
19	26	32	-.61	.49
21	26	32	-.61	.49
18	27	32	-.87	.53
34	27	32	-.87	.53
14	28	32	-1.17	.57
20	28	32	-1.17	.57
22	28	32	-1.17	.57
30	28	32	-1.17	.57
36	28	32	-1.17	.57
37	28	32	-1.17	.57
25	29	32	-1.54	.64
35	29	32	-1.54	.64
38	29	32	-1.54	.64
32	30	32	-2.03	.76

Berdasarkan nilai pada tabel di atas diketahui bahwa pada kolom entri number 28 dan 2 bernilai measure lebih dari 2, entri number 32 bernilai measure kurang dari -2. Hal ini menunjukkan nilai yang tidak sesuai dengan syarat tingkat kesulitan soal. Jika ditelaah lebih dalam, diketahui bahwa soal nomor 28 terlalu sulit karena hanya ada 6 anak yang dapat menjawab benar. Soal nomor 2 hanya 10 anak yang dapat

Guttman Scalogram of Responses:

21	+111111111101110111111111100011010011010	Sis21
27	+1111111111101101111111111100011000100	Sis27
29	+1111111111110111010110101111001001011	Sis29
1	+11111111111111111111111110001110010000100	Sis01
26	+11111111111111111001111011011010010000	Sis26
31	+111101111111111111011110100110101010010	Sis31
7	+110111101111111111101010010110011110100	Sis07
11	+1011111111110101111110000111110100100	Sis11
32	+1111111111110111111110011100001001000	Sis32
4	+111111111110110111100111011001011100000	Sis04
6	+011011111111111101101011111000100011010	Sis06
16	+11111101111111111011111011000100001000	Sis16
10	+101101111101111101111101101001000010000	Sis10
28	+111110111111111010100010010011000000100	Sis28
23	+1101111111000011001010001000011011110100	Sis23
24	+1111001011101001100011100100110110100000	Sis24
25	+1101101110111011001011000000100110010100	Sis25
3	+1111100110011001011100000100101100000000	Sis03
22	+1010011010000100101000001100001110000000	Sis22
9	+0111100000100010011000001010000000000000	Sis09

Pada data skalogram jika menunjukkan pola respon secara sistematis antara siswa (diurutkan dari yang abilitas tinggi ke yang rendah secara vertikal, atas ke bawah) dan butir soal (diurutkan dari yang mudah ke yang susah secara horizontal, kiri ke kanan) (Kurniawan *et al.*, 2018).

Dari tabel data ini dapat diketahui bahwa mayoritas siswa menunjukkan pola scalogram yang acak. Misalnya pada "Sis13" menunjukkan pola yang kurang normal sebab sebelumnya diketahui bahwa semakin ke kanan maka soal akan semakin sulit. Data Sis13 menunjukkan yang terakhir tidak bisa ia jawab, sedangkan soal terakhir yang paling sulit dapat ia jawab. Soal-soal yang lebih mudah pada nomor urut sebelumnya juga ada beberapa yang tidak bisa dijawab olehnya. Hal

yang serupa terjadi pada siswa lainnya. Dengan mendasarkan pada definisi Rasch model, yakni siswa dengan kemampuan lebih rendah tidak akan mempunyai peluang menyelesaikan taraf soal yang lebih sukar. Oleh karena itu, dapat disimpulkan bahwa jawaban yang diberikan siswa-siswa di atas mungkin saja adalah tebakan yang kebetulan benar (*lucky guessing*) atau bahkan hasil dari *cheating*. Menurut Karabatsos (dalam Amelia, 2021), ada lima hal penyebab *person misfit* yaitu *cheating*, *careless responding*, *lucky guessing*, *creative responding*, dan *random responding*. *Cheating*, misalnya menyalin jawaban dari *testee* lain, mengacu pada perilaku yang tidak adil atas jawaban benar pada item soal yang sebenarnya tidak dapat dia jawab dengan benar. *Careless responding* terjadi saat *testee* menjawab dengan benar item sukar namun dengan cara yang tidak jelas justru menjawab salah pada item mudah. *Lucky guessing* terjadi jika *testee* menebak dengan benar item yang sebenarnya dia tidak tau jawaban mana yang benar (Ninoersy et al., 2019; Thamrin et al., 2023; Sidharta et al., 2023). *Creative responding* hanya terjadi pada *testee* yang memiliki kemampuan tinggi saat merespon dengan salah pada item yang sebenarnya mudah karena mereka mengartikan item tersebut dengan cara yang unik dan kreatif. *Random responding* mengacu pada situasi di mana *testee* memilih opsi pilihan ganda secara acak dalam merespons item soal.

D. Penutup

Soal Asesmen Madrasah di Madrasah Ibtidaiyah mata pelajaran Ilmu Pengetahuan Sosial ini memiliki kualitas butir soal yang fit karena telah memenuhi kriteria minimal satu dari ketiga kriteria yang ditentukan. Dalam kaitannya dengan tingkat kesukaran soal masih ada empat soal yang perlu mendapat perhatian lebih oleh guru karena tergolong terlalu mudah dan terlalu sulit. Kemudian jika dilihat dari person terdapat banyak anomaly dari pola jawaban siswa. Didapati beberapa siswa mampu menjawab soal dengan tingkat kesukaran tinggi sedangkan soal dengan tingkat kesukaran rendah tidak dapat dijawab. Dapat disimpulkan bahwa soal sudah termasuk fit. Hanya saja beberapa materi perlu ada peninjauan kembali terhadap abilitas siswa.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada guru-guru di Madrasah Ibtidaiyah daerah Tangerang Selatan yang telah membantu dan kooperatif dalam pemberian data penelitian. Juga kepada dosen yang membimbing penelitian ini.

Daftar Referensi

- Amelia, Noor R. (2021). *Identifikasi Item Fit dan Person Fit dalam Pengukuran Hasil Belajar Kimia*.
- Ernawati, E. (2023). Analisis Soal Ujian Matematika Berkategori Higher Order Thinking Skills (HOTS) Siswa SMP dengan Analisis Rasch Model. *Jurnal Didaktika Pendidikan Dasar*, 7(1), 277-294. <https://doi.org/10.26811/didaktika.v7i1.1027>.
- Hartono, M. S. (2018). *Bimbingan Karier*. Prenada Media. <https://books.google.co.id/books?id=eDa2DwAAQBAJ>.
- Haryanto, M. P., & Amalia, S. (2020). *Evaluasi Pembelajaran (Konsep dan Manajemen)*. UNY Press. <https://books.google.co.id/books?id=YIVBEAAAQBAJ>.
- Joko Subando, S. S. M. P., & HMJ. Nashir Wahid, S. A. M. A. (2022). *Evaluasi Hasil Belajar Pendidikan Agama Islam*. Penerbit Lakeisha.
- Kurniawan, U., & Andriyani, K. D. K. (2018). Analisis soal pilihan ganda dengan RASCH Model. *Jurnal Statistika Universitas Muhammadiyah Semarang*, 6(1), 34-39. <https://doi.org/10.26714/jsunimus.6.1.2018.%25p>.
- Kusmiyati. (2022). *Konsep Dasar Evaluasi Pembelajaran*. Penerbit P4I. <https://books.google.co.id/books?id=iYV-EAAAQBAJ>.
- Ninoersy, T., Tabrani, Z. A., & Wathan, N. (2019). Manajemen Perencanaan Pembelajaran Bahasa Arab Berbasis Kurikulum 2013 pada SMAN 1 Aceh Barat. *FITRAH: Jurnal Kajian Ilmu-ilmu Keislaman*, 5(1), 83-102. doi: <https://doi.org/10.24952/fitrah.v5i1.1759>
- Nurhafidhah, N., Hasby, H., & Alvina, S. (2021). The Analysis of Student Character Values in the Use of Secondary Metabolic Utilization Lab Module. *Jurnal Ilmiah Peuradeun*, 9(1), 179-188. <https://doi.org/10.26811/peuradeun.v9i1.484>

- Sidharta, I., Riadi, F., Rahman, R., Ramadhani, G., & Hidayatullah, A. (2023). Determinant Aspects of Students to Become Entrepreneurs During the Second Wave of Covid-19. *Jurnal Ilmiah Peuradeun*, 11(3), 887-904. <https://doi.org/10.26811/peuradeun.v11i3.900>
- Sumarni, S. (2019). Designing ICT Competences-Integrated Assessment Instruments of Practical Key Teaching Competences for English Language Education Study Program. *IJLECR International Journal of Language Education and Cultural Review*, 5(1), 47-55. <https://doi.org/10.21009/IJLECR.051.05>.
- Thamrin, L., Ghasya, D., & Pranata, R. (2023). The Practicality of Developing a Multi-Split Model Based on the Hands-on-Mind Approach as a Reconstruction of Basic Literacy and Character Values. *Jurnal Ilmiah Peuradeun*, 11(3), 949-964. <https://doi.org/10.26811/peuradeun.v11i3.905>
- Walidin, W., Idris, S., & Tabrani ZA. (2015). *Metodologi Penelitian Kualitatif & Grounded Theory*. Banda Aceh: FTK Ar-Raniry Press.
- Wardani, H., Sujarwo, S., Rakhmawati, Y., & Cahyandaru, P. (2023). Analysis of the Impact of the Merdeka Curriculum Policy on Stakeholders at Primary School. *Jurnal Ilmiah Peuradeun*, 11(2), 513-530. <https://doi.org/10.26811/peuradeun.v11i2.801>
- Widhiarso, W. (2016). Penerapan Model Rasch untuk Mengevaluasi Tes UKKS Dan UKPS. *Tenaga Kependidikan*, 1(1).