

**ANALISIS ITEM SOAL DENGAN MENGGUNAKAN RASCH MODEL
SEBAGAI UKURAN KUALITAS MADRASAH IBTIDAIYAH
PADA MATA PELAJARAN IPA**

Oleh: Lilla Septiliana

Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga Yogyakarta

Email: lillaseptiliana21@gmail.com

Abstract

This study aims to analyze the item items using the Rasch model as a measure of the quality of Islamic elementary schools in science subjects. Comprehensive Assessment as an Effort to Improve Evaluation Quality to Measure the Level of Difficulty of Questions, Item Fit, and Possibility of Item Bias. This exploratory and descriptive research aims to retro-fit the test items and student answer data in the 2nd semester of the Natural Sciences Test for the 2022/2023 academic year. This research was conducted in 3 madrasas, namely MIN 2 OKU East, MIN 3 OKU East, and MIN 4 OKU East. The research findings include the following: 1. The difficulty level of the questions was identified in item S14 with a measurement value of +2.82, and only 25 students answered correctly. 2. Item fit meets the criteria except for item S14, which indicates an answer outside the Outfit Trust limits. However, the mean clothing values and spot size correlations are still in agreement. 3. Detected bias in question item S13, with a probability value of 0.02, which is below the 5% limit (0.05).

Keywords: Item Analysis, Tests: Rasch Model

Abstrak

Penelitian ini bertujuan untuk menganalisis item soal menggunakan rasch model sebagai ukuran kualitas madrasah ibtidaiyah pada mata pelajaran IPA. Penilaian Komprehensif sebagai Upaya Peningkatan Kualitas Evaluasi untuk Mengukur Tingkat Kesulitan Soal, Kesesuaian Butir (*Item Fit*), dan Kemungkinan Bias Butir Soal. Penelitian yang bersifat eksploratif dan deskriptif bertujuan untuk menganalisis secara *retro-fitting* butir soal dan data jawaban siswa pada Tes Ilmu Pengetahuan Alam semester 2 tahun ajaran 2022/2023. Penelitian ini dilakukan di di 3 madrasah, yaitu MIN 2 OKU Timur, MIN 3 OKU Timur, serta MIN 4 OKU Timur. Temuan penelitian meliputi hal-hal berikut: 1. Tingkat kesulitan soal teridentifikasi pada butir soal S14 dengan nilai Ukur +2,82, dan hanya 25 siswa yang menjawab benar. 2. Kesesuaian item pertanyaan (item fit) memenuhi kriteria kecuali pada item S14, yang menunjukkan adanya jawaban di luar batas *Outfit Trust*. Namun, nilai rata-rata pakaian dan korelasi ukuran titik masih sesuai. 3. Terdeteksi adanya bias pada item pertanyaan S13, dengan nilai probabilitas sebesar 0,02, yang berada di bawah batas 5% (0,05).

Kata Kunci: Analisis Item, Tes: Rasch Model

A. Pendahuluan

Tes ialah metode yang berarti dalam pengukuran ataupun metode yang bisa dipakai dan evaluasi di aspek pendidikan. Seseorang pengajar wajib memahami keahlian mengenai penilaian hasil belajar anak didik. Dalam hasil belajar serta evaluasi cara di institusi pendidikan, ada pandangan terpaut penentuan instrumen evaluasi, semacam analisa item pertanyaan serta kategorisasi pertanyaan. Perihal ini dicoba buat memperoleh mutu pertanyaan yang mencukupi serta menganalisa informasi dari hasil evaluasi.¹

Tujuan dari tes merupakan buat melaksanakan langkah filtrasi dini, remedi, ataupun pengetesan untuk mengenali tingkatan kesusahan sesuatu item pertanyaan, bagus itu susah, amat susah, ataupun gampang ataupun amat gampang buat digarap. Tes dicoba buat menilai kesesuaian item pertanyaan (*item fit*) ataupun mutu kesesuaian item dengan bentuk yang terdapat, dan mengetahui mungkin terdapatnya bias pada item pertanyaan itu. Data yang didapat dari hasil tes bisa dipakai buat menilai apakah sesuatu tes penuh standar ataupun belum, yang mana tes standar ini butuh direncanakan dengan bagus supaya bisa diketahui dengan cara besar. Dalam riset ini, tes standar yang didesain merupakan berbentuk item pertanyaan opsi dobel dengan opsi balasan yang betul, di mana anak didik diserahkan 4 alternatif balasan.

Keahlian seseorang guru dalam meningkatkan tes dengan tujuan tingkatkan kompetensi anak didik amat berarti. Oleh sebab itu, tes yang dipakai dalam tiap kompetensi dasar ataupun tes akhir yang dicoba oleh guru wajib penuh standar yang diresmikan. Dikala ini, tes di perguruan sedang berbentuk Berkas Item (tes yang karakteristiknya belum dikenal serta belum bisa dikategorikan selaku tes standar). Perihal ini berlainan dengan Item Bank (tes yang karakteristiknya telah dikenal serta bisa dikira selaku tes standar, tercantum tingkatan kesusahan, energi beda, serta keakuratan balasan). Pemakaian tes yang belum standar bisa menciptakan data yang bias serta tidak cermat terpaut dengan keahlian anak didik, alhasil bukti informasi serta data yang didapat jadi diragukan.

Untuk menilai tingkatan kesusahan biji pertanyaan (*Item Measure*), kesesuaian biji pertanyaan (*item fit*), serta penemuan terdapatnya bias pada biji pertanyaan, pengarang memakai pendekatan bentuk Rasch di Madrasah Ibtidaiyah OKU Timur Sumatera Selatan. Pengukuran Rasch Bentuk sendiri yakni sesuatu bentuk pengukuran yang tercipta hasil dari estimasi yang mengutip duga keabsahan serta reliabilitas tiap calon responden yang menanggapi item atau pertanyaan serta kepayahan item atau pertanyaan untuk tiap pertanyaan atau item.² Pendekatan ini diseleksi sebab guru-guru perguruan itu belum melakukan tes yang menyeluruh buat tingkatkan mutu evaluasi. Oleh sebab itu, mutu pertanyaan, tingkatan kesusahan pertanyaan, serta penemuan miskonsepsi dan bias pertanyaan belum dikenal. Riset lebih dahulu oleh Isnaini, dkk. pula membawa alamat kalau pemakaian bentuk Rasch

¹ Abdul Kadir, Menyusun Dan Menganalisis Tes Hasil Belajar, *Al-Ta'dib: Jurnal Kajian Ilmu Pendidikan*, Volume 8 Nomor 2 (2015), h. 70-81.

² Lisabeth M. Searing, *Family Functioning Scale Validation: A Rasch Analysis*. Chicago: MPH, University of Illinois, 2008.

dalam tes bisa menciptakan tes yang bagus serta bisa berspekulasi keahlian anak didik dengan cermat.³

Bersumber pada penjelasan itu hingga pengarang menyangka perlunya diadakan riset mengenai analisa item soal dengan menggunakan pendekatan rasch model sebagai ukuran kualitas Madrasah Ibtidaiyah pada mata pelajaran IPA.

B. Metode Penelitian

Penelitian ini ialah tipe riset deskriptif eksploratif yang dicoba kepada informasi item serta asumsi anak didik di Madrasah Ibtidaiyah OKU Timur, Sumatera Selatan, pada semester 2 mata pelajaran IPA Tahun Pelajaran 2022 atau 2023. Pendekatan yang dipakai dalam riset ini merupakan *post-hoc diagnostic* yang dihadirkan dalam wujud pendekatan *retro-fitting*. Pendekatan ini mengaitkan analisa biji pertanyaan untuk menguak data hal mutu pertanyaan yang dipakai dalam tes akhir semester 2 mata pelajaran IPA untuk anak didik kategori V pada tahun pelajaran 2022 atau 2023.

Partisipan dalam riset ini merupakan anak didik kategori V yang sudah menanggapi pertanyaan tes standarisasi kualitas Madrasah Ibtidaiyah pada semester 2 mata pelajaran IPA. Tes itu dibuktikan di 3 madrasah, ialah MIN 2 OKU Timur, MIN 3 OKU Timur, serta MIN 4 OKU Timur. Fokus analisa dalam riset ini merupakan pada 20 biji pertanyaan *opsi dobel* yang mempunyai 4 alternatif balasan, ialah A, B, C, serta D. Tata cara pengumpulan informasi dalam riset ini memakai akta, yang mengaitkan pengumpulan informasi dari lembar balasan anak didik kategori V di MIN 2 OKU Timur, MIN 3 OKU Timur, serta MIN 4 OKU Timur, Sumatera Selatan. Pemilihan yang dipakai dalam pengumpulan informasi meliputi:

- a. Lembar soal siswa untuk tes standarisasi mutu Madrasah Ibtidaiyah pada semester 2 mata pelajaran IPA kelas V tahun pelajaran 2022-2023.
- b. Lembar jawaban siswa untuk soal tes standarisasi mutu Madrasah Ibtidaiyah pada semester 2 mata pelajaran IPA kelas V tahun pelajaran 2022-2023.

Pemrosesan dan analisis data digunakan untuk mengolah data yang telah diatur dengan cara yang dapat dipahami. Data yang telah dikumpulkan kemudian dianalisis menggunakan software Rasch Model. Dalam penelitian ini, pemodelan Rasch dilakukan dengan menggunakan perangkat lunak Winsteps 3.92.1.⁴

C. Kajian Teori

Penganalisisan item soal dengan memanfaatkan Rasch Model merupakan salah satu metode yang digunakan untuk mengevaluasi kualitas soal dalam pendidikan. Rasch Model merupakan suatu kerangka pengukuran yang digunakan untuk memperoleh ukuran yang bersifat kualitatif atau kuantitatif dalam proses penilaian pendidikan. Saat diterapkan pada konteks Madrasah Ibtidaiyah (MI) dan mata pelajaran IPA, analisis item soal dengan Rasch Model

³ Isnani et al, "Estimation Of College Students' Ability On Real Analysis Course Using Rasch Model", *Reaseach and Evaluation In Education* No. 2, 2019: 95-102.

⁴ Napitupulu, D, Analisis Model Rasch Pada Instrumen Keberterimaan Marka Optik Penanda Jarak Aman Kendaraan. *Jurnal Penelitian Transportasi Darat*, 2018), 19(1), 37.

mampu memberikan informasi mengenai tingkat kesulitan soal, daya pembeda antara siswa dengan kemampuan yang berbeda, serta konsistensi internal dari instrumen penilaian yang digunakan.

1. Model Rasch

Pada dekade 1960, Georg Rasch mengembangkan sebuah model analisis yang dikenal dengan sebutan 1PL (satu parameter *logistic*), yang merupakan suatu bentuk dari teori respon butir (*Item Response Theory*, IRT).⁵ Selanjutnya, model matematika ini diperkenalkan secara luas oleh Ben Wright.⁶ Dengan menggunakan data yang berupa respons dikotomis (berupa benar atau salah) yang mengindikasikan kemampuan siswa, Rasch mengembangkan sebuah model yang menghubungkan siswa dengan item.⁷

Sebagai contoh, jika seorang siswa dapat menjawab dengan benar 80% dari total soal, tentunya menunjukkan bahwa kemampuan siswa tersebut lebih tinggi daripada siswa lain yang hanya mampu menjawab 65% soal. Data yang diperoleh (dalam bentuk persentase) menunjukkan bahwa data mentah tersebut sebenarnya merupakan jenis data ordinal yang mengindikasikan peringkat dan tidak bersifat linear.⁸ Karena data ordinal tidak memiliki interval yang konsisten, perlu dilakukan konversi data menjadi data rasio untuk keperluan analisis statistik. Dengan demikian, jika seseorang memperoleh skor 80%, maka rasio odds-nya adalah 80:20, yang sebenarnya merupakan data rasio yang lebih sesuai untuk tujuan pengukuran. Melalui penggunaan data rasio ini, Rasch mengembangkan sebuah model pengukuran yang menentukan hubungan antara tingkat kemampuan siswa (*person ability*) dan tingkat kesulitan aitem (*item difficulty*), dengan menggunakan fungsi logaritma untuk menghasilkan pengukuran dengan interval yang seragam. Dampaknya adalah terciptanya satuan baru yang dikenal sebagai logit (satuan *log odds*), yang menggambarkan kemampuan siswa dan tingkat kesulitan aitem. Dengan menggunakan nilai logit yang diperoleh, dapat disimpulkan bahwa tingkat keberhasilan siswa dalam menjawab soal sangat bergantung pada tingkat kemampuan mereka dan tingkat kesulitan soal yang dihadapi.⁹

2. Kualitas Madrasah Ibtidaiyah

Kualitas Madrasah Ibtidaiyah merujuk pada sejauh mana institusi tersebut memenuhi kriteria yang telah ditetapkan untuk memberikan pendidikan yang unggul kepada siswa. Evaluasi kualitas madrasah melibatkan aspek akademik

⁵ Olsen, L. W, *Essays on Georg Rasch and his contributions to statistics*, (Unpublished PhD thesis at Institute of economics university of copenhagen, 2003).

⁶ Linacre, J.M, *A User's Guide to Winsteps Ministep: Rasch Model Computer Program*, (Program Manual 3.73, 2011).

⁷ Sumintono, B dan Widhiarso, W, *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial*, (Cimahi: Trim Komunikata Publishing House, 2003).

⁸ Linacre, J.M, Investigating Rating Scale Category Utility, *Journal of Outcome Measurement*, 1999) 3(2), 103-122.

⁹ Linacre, J.M, *A User's Guide to Winsteps Ministep: Rasch Model Computer Program*, (Program Manual 3.73, 2011).

dan non-akademik, seperti pengelolaan, sarana prasarana, lingkungan belajar, dan pembentukan karakter siswa.¹⁰

Beberapa faktor yang berpengaruh terhadap kualitas Madrasah Ibtidaiyah antara lain:

- a. Rancangan Kurikulum: Kualitas madrasah dapat dinilai berdasarkan keberhasilan kurikulum dalam menghadirkan materi pembelajaran yang relevan, mutakhir, dan sesuai dengan kebutuhan siswa. Kurikulum harus mencakup standar kompetensi yang jelas, tujuan pembelajaran yang terukur, serta metode dan strategi pengajaran yang efektif.¹¹
- b. Tenaga Pendidik: Kualitas madrasah juga ditentukan oleh keahlian, pengalaman, dan kualitas tenaga pendidik. Guru-guru yang berkualitas memiliki pemahaman yang mendalam terhadap mata pelajaran yang diajarkan, mampu menyampaikan materi secara efektif, dan memiliki kemampuan untuk memotivasi serta mendukung perkembangan siswa.
- c. Manajemen dan Kepemimpinan: Kualitas madrasah juga tergantung pada kemampuan manajemen yang baik dan kepemimpinan yang efektif. Manajemen yang baik melibatkan perencanaan yang matang, pengorganisasian yang efisien, pengawasan yang tepat, dan evaluasi yang teratur. Kepemimpinan yang efektif mencakup pengambilan keputusan yang bijaksana, komunikasi yang efektif, dan pengembangan staf yang terarah.
- d. Fasilitas dan Sarana Pembelajaran: Fasilitas dan sarana pembelajaran yang memadai memiliki peran penting dalam menciptakan lingkungan belajar yang kondusif. Madrasah dengan fasilitas yang memadai, seperti ruang kelas yang nyaman, perpustakaan, laboratorium, dan fasilitas olahraga, mampu memberikan pengalaman pembelajaran yang lebih baik bagi siswa.
- e. Kegiatan Ekstrakurikuler: Kualitas madrasah juga dapat dinilai dari kegiatan ekstrakurikuler yang ditawarkan kepada siswa. Kegiatan ini membantu mengembangkan potensi siswa di luar pembelajaran akademik, seperti seni, olahraga, kegiatan sosial, dan kepemimpinan.
- f. Pengawasan dan Evaluasi: Pengawasan yang ketat dan evaluasi berkelanjutan terhadap proses pembelajaran dan prestasi siswa juga merupakan faktor penting dalam menentukan kualitas madrasah. Melalui pengawasan dan evaluasi yang baik, masalah dapat diidentifikasi dan solusi dapat diterapkan untuk meningkatkan kualitas pendidikan.

Secara keseluruhan, kualitas Madrasah Ibtidaiyah dapat diukur berdasarkan pencapaian akademik siswa, partisipasi siswa dalam kegiatan pembelajaran, kepuasan orang tua terhadap pelayanan pendidikan, dan reputasi madrasah di masyarakat.

3. Ilmu Pengetahuan Alam (IPA)

Ilmu Pengetahuan Alam (IPA) atau sains dalam pengertian yang lebih spesifik merujuk pada disiplin ilmu yang terdiri dari *physical sciences* dan *life*

¹⁰ Nur Hasanah, Dampak Kompetensi Profesional Guru Dalam Meningkatkan Mutu Pendidikan Madrasah Ibtidaiyah di Kota Salatiga, *Jurnal Penelitian Sosial Keagamaan*, Volume 9 Nomor 2 (2015), h. 457-458.

¹¹ Rahmat Hidayat, Abdillah, "*Ilmu Pendidikan (Konsep, Teori dan Aplikasinya)*", (Medan: Lembaga Peduli Pengembangan Pendidikan Indonesia, 2019), h. 65-66.

sciences. Physical sciences melibatkan bidang studi seperti kimia, astronomi, dan mineralogi. Sementara itu, *life sciences* mencakup bidang biologi, seperti anatomi, zoologi, sitologi, fisiologi, dan lain sebagainya. James Conant, dalam karyanya, memberikan definisi sains sebagai sebuah rangkaian konsep dan skema konseptual yang saling terkait, yang berkembang sebagai hasil dari eksperimen dan observasi, dan memiliki nilai penting untuk diamati dan diujicobakan lebih lanjut.¹²

Tujuan dari pembelajaran dalam bidang studi IPA di Madrasah adalah:

- a. Meningkatkan rasa ingin tahu dan kesadaran tentang berbagai jenis lingkungan alam dan lingkungan buatan serta hubungannya dengan manfaatnya dalam kehidupan sehari-hari manusia.
- b. Mengembangkan keterampilan proses siswa agar dapat memecahkan masalah melalui pendekatan "*doing science*".
- c. Mengembangkan kemampuan untuk menerapkan IPA, teknologi, dan keterampilan yang bermanfaat dalam kehidupan sehari-hari serta melanjutkan pendidikan ke jenjang yang lebih tinggi.
- d. Mengembangkan pemahaman, sikap, dan nilai-nilai yang relevan dengan kemajuan IPTEK, keberlanjutan lingkungan yang bermanfaat dalam kehidupan sehari-hari, dan pelestariannya.

Dengan mempertimbangkan berbagai fungsi pembelajaran IPA tersebut, penting bagi pembelajaran IPA untuk terintegrasi dengan pembelajaran lainnya. Hal ini disebabkan karena IPA pada dasarnya merupakan sarana untuk mengembangkan ilmu pengetahuan, dengan menggunakan pendekatan ilmiah sebagai salah satu pendukung dalam proses pembelajaran.

D. HASIL DAN PEMBAHASAN

Meskipun Rasch Model telah diterapkan secara luas dalam analisis item soal pada berbagai tingkat pendidikan, termasuk sekolah umum, penelitian yang khusus memusatkan pada penggunaan Rasch Model untuk mengevaluasi kualitas soal di Madrasah Ibtidaiyah pada mata pelajaran IPA masih terbatas. Madrasah Ibtidaiyah memiliki ciri khasnya sendiri, termasuk perbedaan kurikulum dan pendekatan pengajaran dengan sekolah umum. Karena itu, penting bagi kita untuk melaksanakan penelitian yang secara spesifik menjelajahi pemanfaatan Rasch Model untuk mengukur kualitas item soal di Madrasah Ibtidaiyah pada mata pelajaran IPA.

Penelitian semacam ini memiliki manfaat penting bagi Madrasah Ibtidaiyah dalam meningkatkan penilaian dan evaluasi pembelajaran pada mata pelajaran IPA. Dengan menggunakan Rasch Model, maka kita dapat mengidentifikasi dan memperbaiki item soal yang tidak memadai, mengukur tingkat kesulitan dan daya pembeda dari setiap item soal, serta mengevaluasi konsistensi internal dari instrumen penilaian yang digunakan di Madrasah Ibtidaiyah.

Melakukan penelitian tentang analisis item soal dengan menggunakan Rasch Model sebagai alat untuk menilai kualitas Madrasah Ibtidaiyah pada mata pelajaran IPA akan memberikan wawasan dan pemahaman yang lebih mendalam tentang metode penilaian dan evaluasi yang efektif dalam konteks Madrasah Ibtidaiyah. Selain itu, penelitian ini juga akan membantu dalam

¹² Usman Samawanto, *Pembelajaran IPA di Sekolah Dasar*, (Jakarta: PT Indeks, 2018), h. 1.

pengembangan kurikulum yang sesuai dan peningkatan pembelajaran pada mata pelajaran IPA di Madrasah Ibtidaiyah.

Oleh karena itu, penelitian semacam ini akan memberikan kontribusi yang signifikan dalam mengisi kesenjangan penelitian yang ada dan meningkatkan pemahaman kita tentang penggunaan Rasch Model sebagai alat untuk menganalisis kualitas item soal di Madrasah Ibtidaiyah pada mata pelajaran IPA. Temuan yang dilakukan dengan menggunakan aplikasi *winstep* memperoleh hasil sebagai berikut:

1. Tingkat Kesulitan, Item Soal (*Item Measure*)

Tingkatan kesusahan pada tiap item pertanyaan (*item measure* ataupun item logit) bisa didapat lewat pemakaian program *winstep*, yang terdapat pada bagan ke- 4 (*item measure*). Bagan 1 menunjukkan sebagian kolom yang membagikan informasi mengenai tingkatan kesusahan dari tiap item pertanyaan (*item measure*), semacam yang ditunjukkan dalam tabel di bawah ini:

Tabel 1.1
Item Measure

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL		INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		Item
				S. E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%		
14	25	81	2.82	.28	1.26	1.9	2.38	3.3	.31	.51	67.9	76.5	S14	
10	36	81	2.00	.27	1.08	.7	1.17	.8	.48	.53	74.4	72.4	S10	
19	38	81	1.86	.27	1.10	.9	1.28	1.4	.47	.53	67.9	72.4	S19	
13	41	81	1.65	.26	1.01	.1	.86	-.7	.55	.54	66.7	72.3	S13	
6	46	81	1.30	.27	.93	-.5	.98	.0	.56	.53	78.2	73.1	S6	
3	48	81	1.16	.27	1.21	1.7	1.29	1.5	.41	.53	70.5	73.6	S3	
12	49	81	1.08	.27	1.02	.2	.99	.0	.52	.53	74.4	73.8	S12	
9	55	81	.63	.28	.98	-.1	.84	-.6	.54	.52	74.4	76.6	S9	
2	59	81	.30	.29	1.14	1.0	1.01	.1	.44	.50	73.1	79.1	S2	
11	60	81	.21	.30	1.02	.2	1.07	.3	.47	.49	82.1	79.7	S11	
4	61	81	.12	.30	1.05	.4	1.29	1.0	.45	.49	78.2	80.3	S4	
20	66	81	-.38	.33	.80	-1.1	.67	-.8	.55	.45	87.2	83.7	S20	
7	67	81	-.49	.34	1.11	.6	1.09	.4	.39	.44	80.8	84.6	S7	
16	69	81	-.73	.36	.77	-1.1	.46	-1.2	.56	.42	91.0	86.4	S16	
17	72	81	-1.14	.39	.56	-2.0	.24	-1.7	.61	.38	94.9	88.9	S17	
18	73	81	-1.31	.41	.80	-.7	.33	-1.2	.50	.37	88.5	89.9	S18	
15	75	81	-1.69	.46	.83	-.5	.99	.3	.37	.33	92.3	92.3	S15	
5	76	81	-1.91	.50	.85	-.3	.27	-.8	.42	.30	93.6	93.5	S5	
1	78	81	-2.52	.62	.93	.0	.69	.0	.27	.24	96.2	96.1	S1	
8	79	81	-2.97	.74	.80	-.1	.15	-.8	.33	.20	97.4	97.4	S8	
MEAN	58.6	81.0	.00	.36	.96	.1	.90	.1			81.5	82.1		
S. D.	15.4	.0	1.56	.13	.17	.9	.49	1.1			10.1	8.3		

Dalam tabel itu, ada data hal tingkatan kesusahan biji pertanyaan (*item measure*) yang bisa diamati lewat angka Measure pada kolom yang cocok. Angka Measure yang besar membuktikan tingkatan kesusahan yang besar pada biji pertanyaan itu. Penemuan ini pula searah dengan Keseluruhan Score, yang membawa alamat jumlah balasan betul. Ilustrasinya, pada biji pertanyaan ke-14, nampak kalau angka Measure- nya merupakan+2. 82, serta cuma 25 anak didik yang menanggapi dengan betul(nampak pada kolom keseluruhan score).

Dari bagan Measure ini, ada data lain yang bisa menarangkan hubungannya dengan tingkatan kesusahan. Bila menyamakan biji pertanyaan (S10) dengan biji pertanyaan (S9), bisa disimpulkan kalau tingkatan kesusahan pada biji pertanyaan (S10) 3 kali lebih besar dari pada biji pertanyaan (S9). Tetapi, ada perbandingan sedikit kala memandang angka yang didapat, dimana ada 36 anak didik yang menanggapi dengan betul pada biji pertanyaan (S10), sebaliknya pada biji pertanyaan (S9) ada 55 anak didik yang menanggapi dengan betul.

Pada Item Measure pula bisa membagikan data yang bermanfaat buat membagi tingkatan kesusahan dari tiap biji pertanyaan (*Item Measure*), yang bisa diamati lewat angka Standar Digresi (1. 71). Kala angka ini digabungkan dengan pada umumnya Measure, ada 4 golongan pertanyaan bersumber pada tingkatan kesulitannya. Dalam perihal ini, tingkatan kesusahan biji pertanyaan dikelompokkan jadi 4 jenis, ialah amat susah, susah, amat gampang, serta gampang. Buat golongan biji pertanyaan susah, ada pertanyaan (S19, S13, S6, S3, serta S12) dengan determinasi (0, 0 logit+1SD). Sedangkan itu, golongan biji pertanyaan jenis tingkatan kesusahan amat susah ada pada pertanyaan (S14 serta S10) dengan determinasi(lebih besar dari+1SD). Buat pengelompokan biji pertanyaan dengan tingkatan kesusahan gampang, ada pertanyaan (S9, S2, S11, S4, S20, S7, S16, S17, S18, S15, serta S5) dengan determinasi (0, 0 Logit- 1SD). Sebaliknya, buat pengelompokan tingkatan kesusahan biji pertanyaan amat gampang, ada pada pertanyaan (S1 serta S8) dengan determinasi(lebih kecil dari- 1SD).

2. Tingkat Kesesuaian Item Soal (*Item Fit*)

Selanjutnya, uraian dicoba dengan memandang Item Bugat memakai pemodelan rasch. Item Bugat membagikan data mengenai apakah biji pertanyaan berperan dengan cara wajar serta cocok dengan tujuannya. Bila ada biji pertanyaan yang membuktikan ketidaksesuaian, perihal itu membawa alamat terdapatnya miskonsepsi pada biji pertanyaan itu. Informasi ini amat bernilai untuk guru buat tingkatkan mutu biji pertanyaan serta menghindari terbentuknya miskonsepsi kala mengonsep balik biji pertanyaan itu. Selanjutnya ini, bagan item bugat menyusunkan biji pertanyaan (*item measure*) secara berurutan:

Tabel 1.2
Item Fit

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S. E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT OBS%	MATCH EXP%	Item
14	25	81	2.82	.28	1.26	1.9	2.38	3.3	A .31	.51	67.9	76.5	S14
3	48	81	1.16	.27	1.21	1.7	1.29	1.5	B .41	.53	70.5	73.6	S3
4	61	81	.12	.30	1.05	.4	1.29	1.0	C .45	.49	78.2	80.3	S4
19	38	81	1.86	.27	1.10	.9	1.28	1.4	D .47	.53	67.9	72.4	S19
10	36	81	2.00	.27	1.08	.7	1.17	.8	E .48	.53	74.4	72.4	S10
2	59	81	.30	.29	1.14	1.0	1.01	.1	F .44	.50	73.1	79.1	S2
7	67	81	-.49	.34	1.11	.6	1.09	.4	G .39	.44	80.8	84.6	S7
11	60	81	.21	.30	1.02	2	1.07	.3	H .47	.49	82.1	79.7	S11
12	49	81	1.08	.27	1.02	.2	.99	.0	I .52	.53	74.4	73.8	S12
13	41	81	1.65	.26	1.01	.1	.86	-.7	J .55	.54	66.7	72.3	S13
15	75	81	-1.69	.46	.83	-.5	.99	-.3	J .37	.33	92.3	92.3	S15
9	55	81	.63	.28	.98	-.1	.84	-.6	I .54	.52	74.4	76.6	S9
6	46	81	1.30	.27	.93	-.5	.98	.0	H .56	.53	78.2	73.1	S6
1	78	81	-2.52	.62	.93	.0	.69	.0	G .27	.24	96.2	96.1	S1
5	76	81	-1.91	.50	.85	-.3	.27	-.8	F .42	.30	93.6	93.5	S5
8	79	81	-2.97	.74	.80	-.1	.15	-.8	E .33	.20	97.4	97.4	S8
20	66	81	-.38	.33	.80	-1.1	.67	-.8	D .55	.45	87.2	83.7	S20
18	73	81	-1.31	.41	.80	-.7	.33	-1.2	C .50	.37	88.5	89.9	S18
16	69	81	-.73	.36	.77	-1.1	.46	-1.2	B .56	.42	91.0	86.4	S16
17	72	81	-1.14	.39	.56	-2.0	.24	-1.7	A .61	.38	94.9	88.9	S17
MEAN	58.6	81.0	.00	.36	.96	.1	.90	.1			81.5	82.1	
S. D.	15.4	.0	1.56	.13	.17	.9	.49	1.1			10.1	8.3	

Menurut pernyataan Boone, Bond, serta Fox, dituturkan kalau angka outfit z- standard, outfit means square, serta angka poin measure correlation dikira bisa diperoleh bila tingkatan kesesuaian biji pertanyaan (*item bugat*) sudah penuh patokan yang diresmikan. Bila ketiga patokan itu belum terkabul pada sesuatu biji pertanyaan, hingga biji pertanyaan itu bisa ditentukan tidak bisa

diperoleh serta butuh dicoba koreksi. Aksi ini dicoba buat membenarkan terdapatnya kesesuaian serta mutu biji pertanyaan yang mencukupi di era kelak.

Menurut Boone¹³, untuk memverifikasi kesesuaian butir soal yang tidak diterima (*misfit* atau *outliers*) kriteria yang digunakan adalah:

- a. Nilai ZSTD diterima jika $-2,0 < ZSTD < +2,$
- b. Nilai MNSQ diterima jika $0,5 < MNSQ < 1,5$
- c. Nilai Pt Mean Corr diterima jika $0,4 < Pt \text{ Measure Corr} < 0,85$

Dari data yang tertera pada Tabel 2, terlihat bahwa terdapat kelemahan pada butir soal S14 yang memerlukan perbaikan. Hal ini disebabkan oleh ketidaksesuaian nilai ZSTD, MNSQ, dan Pt. *Mean Corr* dengan kriteria yang telah ditetapkan. Oleh karena itu, kesimpulan akhir adalah bahwa hanya butir soal S14 yang memerlukan perbaikan.

3. Deteksi adanya item soal yang bias

Dalam permodelan Rasch untuk mendeteksi bias ini disebut dengan deteksi DIF (*differential item functioning*) atau keberfungsian butir diferensial.¹⁴ Jika terdapat indikasi bahwa siswa dengan karakteristik tertentu mendapatkan keuntungan lebih dibandingkan siswa lainnya, hal tersebut mengindikasikan adanya bias pada butir soal. Misalnya, jika butir soal lebih mudah bagi siswa perempuan daripada siswa laki-laki, maka dapat dikatakan bahwa terdapat bias gender pada butir soal tersebut. Untuk mendeteksi bias pada butir soal, dapat menggunakan pendekatan *Differential Item Functioning* (DIF) atau fungsi butir diferensial. Informasi lebih lanjut dapat dilihat pada tabel berikut:

Tabel 1.3
Deteksi adanya item soal yang bias

DIF class specification is: DIF=\$S3W1

Person CLASS	DIF MEASURE	DIF S.E.	Person CLASS	DIF MEASURE	DIF S.E.	DIF CONTRAST	JOINT S.E.	Welch t	Welch d.f.	Welch Prob.	Mantel-Haenszel	Size	Item Number	Item Name
L	-1.95	.64	P	-3.99	1.84	2.04	1.95	1.05	62	.2992	.5000	.4795	1	S1
L	.52	.38	P	-.06	.49	-.59	.62	-.95	75	.3465	.3411	.5592	2	S2
L	1.60	.37	P	.57	.44	1.03	.57	1.80	75	.0753	1.4555	.2276	3	S3
L	.52	.38	P	-.61	.56	1.13	.67	1.68	73	.0963	1.3251	.2497	4	S4
L	-1.59	.57	P	-2.71	1.07	1.12	1.21	-.92	70	.3583	.5000	.4795	5	S5
L	1.74	.37	P	.76	.42	-.98	.56	1.75	75	.0848	4.6862	.0304	6	S6
L	-.42	.43	P	-.61	.56	-.19	.70	-.27	75	.7913	.4994	.4798	7	S7
L	-3.19	1.04	P	-2.71	1.07	-.49	1.49	-.33	75	.7442	.3333	.5637	8	S8
L	.38	.38	P	-.93	.41	-.55	.56	-.99	75	.3278	.8992	.3430	9	S9
L	1.60	.37	P	2.38	.37	-.78	.52	-1.49	75	.1392	1.4248	.2326	10	S10
L	.52	.38	P	-.32	.52	-.84	.64	-1.31	74	.1935	.9775	.3228	11	S11
L	1.20	.37	P	-.93	.41	-.27	.55	-.49	75	.6255	.0023	.9614	12	S12
L	1.07	.37	P	2.25	.37	-1.18	.52	-2.28	75	.0255	5.6259	.0177	13	S13
L	2.90	.41	P	3.07	.38	-.57	.56	-1.01	75	.3147	.0082	.9279	14	S14
L	-1.59	.57	P	-1.88	.79	-.29	.98	-.30	74	.7668	.0050	.9438	15	S15
L	-.81	.46	P	-.61	.56	-.21	.72	-.29	75	.7749	.0744	.7850	16	S16
L	-1.29	.52	P	-.94	.60	-.35	.80	-.44	75	.6603	.7500	.3865	17	S17
L	-1.95	.64	P	-.61	.56	-1.34	.85	-1.59	75	.1165	3.5854	.0583	18	S18
L	1.74	.37	P	1.98	.37	-.23	.53	-.45	75	.6570	.0086	.9262	19	S19
L	-.81	.46	P	-.17	.47	-.98	.66	-1.49	75	.1405	.9655	.3258	20	S20
P	-3.99	1.84	L	-1.95	.64	-2.04	1.95	-1.05	62	.2992	.5000	.4795	1	S1
P	-.06	.49	L	.52	.38	-.59	.62	-.95	75	.3465	.3411	.5592	2	S2
P	.57	.44	L	1.60	.37	-1.03	.57	-1.80	75	.0753	1.4555	.2276	3	S3
P	-.61	.56	L	-.52	.38	-1.13	.67	-1.68	73	.0963	1.3251	.2497	4	S4
P	-2.71	1.07	L	-1.59	.57	-1.12	1.21	-.92	70	.3583	.5000	.4795	5	S5
P	.76	.42	L	1.74	.37	-.98	.56	1.75	75	.0848	4.6862	.0304	6	S6
P	-.61	.56	L	-.42	.43	-.19	.70	-.27	75	.7913	.4994	.4798	7	S7
P	-2.71	1.07	L	-3.19	1.04	-.49	1.49	-.33	75	.7442	.3333	.5637	8	S8
P	.38	.38	L	-.93	.41	-.55	.56	-.99	75	.3278	.8992	.3430	9	S9
P	2.38	.37	L	1.60	.37	-.78	.52	-1.49	75	.1392	1.4248	.2326	10	S10
P	-.32	.52	L	1.52	.38	-.84	.64	-1.31	74	.1935	.9775	.3228	11	S11
P	-.93	.41	L	1.20	.37	-.27	.55	-.49	75	.6255	.0023	.9614	12	S12
P	2.25	.37	L	1.07	.37	-1.18	.52	-2.28	75	.0255	5.6259	.0177	13	S13
P	3.07	.38	L	2.90	.41	-.57	.56	-1.01	75	.3147	.0082	.9279	14	S14
P	-1.88	.79	L	-1.59	.57	-.29	.98	-.30	74	.7668	.0050	.9438	15	S15
P	-.61	.56	L	-.81	.46	-.21	.72	-.29	75	.7749	.0744	.7850	16	S16
P	-.94	.60	L	-1.29	.52	-.35	.80	-.44	75	.6603	.7500	.3865	17	S17
P	1.98	.37	L	1.74	.37	-.23	.53	-.45	75	.6570	.0086	.9262	18	S18
P	-.17	.47	L	-.81	.46	-.98	.66	-1.49	75	.1405	.9655	.3258	19	S19
P			L								1.06		20	S20

¹³ Boone, Staver, and Yale, *RASCH ANALYSIS IN THE HUMAN SCIENCES*

¹⁴ Untary, H, *Analisis Data Penelitian dengan Winstep dan Ministep* (H. Untary, ed.). Bogor: Halaman Moeka Publishing, 2020.

Dalam kondisi ini, data mentah yang dipakai melingkupi 2 variabel demografi, ialah jenis kelamin (laki-laki dan perempuan) serta asal madrasah (MIN 2 OKU Timur, MIN 3 OKU Timur, serta MIN 4 OKU Timur). Dengan begitu, ada 3 variabel demografi yang bisa dipakai buat mengetahui terdapatnya bias. Bila ditemui kalau probabilitas nilai butir pertanyaan terletak di bawah 5% (0,05), hingga bisa disimpulkan kalau ada bias pada soal itu. Perihal ini bisa diamati pada Tabel 3 di atas, dimana ada bias pada biji pertanyaan S13 dengan probabilitas nilai butir sebesar 0,02, yang terletak di bawah 5% (0,05).

PEMBAHASAN

Penggunaan rasch model dalam kaitannya dengan analisis butir bukanlah hal yang baru, tetapi penelitian yang khusus memusatkan pada penggunaan Rasch Model untuk mengevaluasi kualitas soal di Madrasah Ibtidaiyah pada mata pelajaran IPA masih sangat terbatas. Hal tersebut ditunjukkan dengan penelitian yang dilakukan oleh Kathryn Scantlebury yang mengatakan bahwasanya kekuatan model Rasch sebagai alat psikometri dan teknik analisis, mengacu pada peta item orang, penahan, fungsi item diferensial, dan kesesuaian item orang. Ia pun menjelaskan agar pendidik sains harus hati-hati memeriksa alat yang mereka gunakan untuk mengukur dan mendokumentasikan perubahan dalam sistem pendidikan. Sedangkan menurut Nuryanto peningkatan mutu Madrasah Ibtidaiyah (*school-based management*) dan peningkatan kompetensi guru dapat dilakukan dengan strategi, pemenuhan kebutuhan belajar dan manajemen berbasis sekolah.¹⁵ Oleh karena itu peneliti melakukan penelitian dengan cara menganalisis item soal dengan menggunakan rasch model sebagai ukuran kualitas Madrasah Ibtidaiyah pada mata pelajaran IPA.

Penelitian ini menemukan bahwa tingkat kesusahan pada bagan ke-4, terdapat item measure yang menunjukkan tingkatan kesusahan dari masing-masing item pertanyaan. Angka measure yang besar menunjukkan tingkat kesusahan yang besar pada item pertanyaan tersebut. Misalnya, pada biji pertanyaan ke-14, angka measure adalah +2.82, dan hanya 25 anak didik yang menjawab dengan benar (terlihat pada kolom keseluruhan score). Selain itu, bagian tersebut juga menjelaskan bahwa data pada item measure dapat memberikan informasi lain yang dapat menggambarkan hubungannya dengan tingkatan kesusahan. Misalnya, jika membandingkan biji pertanyaan (S10) dengan biji pertanyaan (S9), dapat disimpulkan bahwa tingkatan kesusahan pada biji pertanyaan (S10) adalah tiga kali lebih besar daripada biji pertanyaan (S9). Namun, perlu diperhatikan bahwa terdapat perbedaan dalam jumlah siswa yang menjawab dengan benar, yaitu 36 siswa pada biji pertanyaan (S10) dan 55 siswa pada biji pertanyaan (S9). Pada tingkat kesusahan dalam Measure, ada 4 golongan pertanyaan bersumber pada tingkatan kesulitannya. Dalam analisis ini sudah mencakup 4 jenis tingkatan kesusahan biji pertanyaan, ialah amat susah, susah, amat gampang, serta gampang hal ini sejalan dengan kajian

¹⁵ Nuryanto, Peningkatan Mutu Madrasah Ibtidaiyah: *School-Based Management* Dan Kompetensi Guru Madrasah, *Elementary*, Vol. 4 Edisi Januari-Juni 2018, h. 27-30.

Purba yang menyatakan bahwa semakin baik instrumen tersebut karena dapat mengidentifikasi kelompok (responden dan butir soal secara lebih rinci).¹⁶

Tingkat Kesesuaian Item Soal (*Item Fit*): Bagian ini membahas tentang penggunaan Item Fit dalam pemodelan Rasch untuk mengevaluasi kesesuaian dan kecocokan item pertanyaan dengan tujuannya. Item Fit memberikan data mengenai apakah item pertanyaan berperan secara wajar dan cocok dengan tujuan yang ditetapkan. Jika terdapat ketidaksesuaian, hal tersebut dapat mengindikasikan adanya miskonsepsi pada item pertanyaan tersebut. Hal tersebut seperti yang dibahas oleh Azizah dan wahyuningsih bahwa batas-batas butir soal dikatakan layak (*valid*) jika memenuhi salah satu atau kedua prasyarat *Outfit Mean Square*, *Outfit Z-Standard* dan *Point Measure Correlation*.¹⁷ Oleh karena itu disimpulkan bahwa pada Tabel 2, terlihat bahwa terdapat kelemahan pada butir soal S14 yang memerlukan perbaikan. Hal ini disebabkan oleh ketidaksesuaian nilai 5 ZSTD, MNSQ, dan Pt.4 *Mean Corr* dengan kriteria yang telah ditetapkan. Oleh karena itu, kesimpulan akhir adalah hanya butir soal S14 yang perlu diperbaiki.

Deteksi Adanya Item Soal yang Bias: Bagian ini membahas tentang deteksi adanya bias pada item pertanyaan berdasarkan data demografi yang digunakan dalam penelitian. Dalam hal ini, data demografi mencakup dua variabel, yaitu jenis kelamin (laki-laki dan perempuan) dan asal madrasah (MIN 2 OKU Timur, MIN 3 OKU Timur, dan MIN 4 OKU Timur). Untuk mengetahui adanya bias, jika probabilitas nilai butir pertanyaan berada di bawah 5% (0,05), maka dapat disimpulkan bahwa terdapat bias pada item pertanyaan tersebut. Dalam contoh yang diberikan pada Tabel 3, terlihat bahwa terdapat bias pada biji pertanyaan S13 dengan probabilitas nilai butir sebesar 0,02, yang berada di bawah 5% (0,05).

Dengan demikian, bagian-bagian tersebut memberikan informasi mengenai tingkatan kesusahan item pertanyaan, kesesuaian item soal, dan deteksi adanya bias pada item pertanyaan. Informasi tersebut dapat digunakan oleh guru atau peneliti untuk meningkatkan kualitas item pertanyaan dan mencegah terbentuknya miskonsepsi pada saat mempelajari kembali item pertanyaan tersebut.

SIMPULAN

Analisa kepada butir pertanyaan dalam tes standarisasi kualitas mata pelajaran IPA di Madrasah Ibtidaiyah OKU Timur Sumatera Selatan memakai pendekatan Rasch Bentuk menciptakan penemuan selaku selanjutnya. Pertama, ada tingkatan kesusahan pada butir pertanyaan (*item measure*) yang bisa diamati pada butir pertanyaan (S14), dengan nilai Measure sebesar +2.82 serta cuma 25 anak didik yang sanggup menanggapi dengan betul. Kedua, sebagian besar butir soal (*item fit*) penuhi patokan outfit z- standard, nilai outfit means-square, serta poin *measure correlation*, tetapi ada satu reaksi yang terletak di luar batasan ruang keyakinan *Outfit* pada butir pertanyaan (S14). Ketiga, ada

¹⁶ Purba, S.E.D, Analisis Instrumen Tes Prestasi Model Rasch Pada Mata Pelajaran Dasar dan Pengukuran Listrik, *Wiyata Dharma: Jurnal Penelitian Dan Evaluasi Pendidikan*, 6(2), (2018), h. 142-147.

¹⁷ Azizah & Wahuningsih, Penggunaan Model Rasch Untuk Menganalisis Instrumen Tes Pada Mata Kuliah Matematika Akuatria, *Jurnal Pendidikan Matematika (JUPITEK)*, 3(1), (2020), h. 45-50.

penemuan terdapatnya bias pada butir soal yang terjalin pada butir soal (S13), dengan probabilitas butir sebesar 0.02 yang terletak di bawah nilai signifikansi 5% (0.05).

DAFTAR PUSTAKA

- Boone, J. William, John R S, and Melissa S Yale. 2013. *"Rasch Analysis In The Human Sciences"*. Springer.
- Hasanah, Nur. 2015. *"Dampak Kompetensi Profesional Guru Dalam Meningkatkan Mutu Pendidikan Madrasah Ibtidaiyah di Kota Salatiga"*. Jurnal Penelitian Sosial Keagamaan. Vol. 9 (2).
- Hidayat, Rahmat, Abdillah. 2019. *"Ilmu Pendidikan (Konsep, Teori dan Aplikasinya)"*. Medan: Lembaga Peduli Pengembangan Pendidikan Indonesia.
- Isnani et al. 2019. *"Estimation Of College Students' Ability On Real Analysis Course Using Rasch Model"*. Reaseach and Evaluation In Education. No. 2.
- Kadir, Abdul. 2015. *"Menyusun Dan Menganalisis Tes Hasil Belajar"*. Al-Ta'dib: Jurnal Kajian Ilmu Pendidikan. Vol. 8. No. 2.
- Linacre, J.M. 1999. *"Investigating Rating Scale Category Utility"*. Journal Of Outcome Measurement. Vol. 3 (2).
- Linacre, J.M. 2011. *"A User's Guide To Winsteps Ministep: Rasch Model Computer Program, Program Manual 3.73."*
- Napitupulu, D. 2018. *"Analisis Model Rasch Pada Instrumen Keberterimaan Marka Optik Penanda Jarak Aman Kendaraan"*. Jurnal Penelitian Transportasi Darat. Vol. 19(1).
- Nuryanto. Peningkatan Mutu Madrasah Ibtidaiyah: *School-Based Management Dan Kompetensi Guru Madrasah. Elementary*. Vol. 4 Edisi Januari-Juni 2018.
- Olsen, L. W. 2003. *"Essays On Georg Rasch And His Contributions To Ststistics"*. Unpulished Phd Thesis At Institure Of Economics University Of Copemhagen, 2003.
- Purba, S.E.D. 2018. Analisis Instrumen Tes Prestasi Model Rasch Pada Mata Pelajaran Dasar dan Pengukuran Listrik. *Wiyata Dharma: Jurnal Penelitian Dan Evaluasi Pendidikan*. Vol. 6(2).
- Samawanto, Usman. 2018. *"Pembelajaran IPA di Sekolah Dasar"*. Jakarta: PT Indeks.
- Searing, M. Lisabeth M. Searing. 2008 *"Family Functioning Scale Validition: A Rasch Analysis"*. Chicago: MPH, University of Illinois.
- Sumintono, B Dan Widhiarso, W. 2003. *"Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial"*. Cimahi: Trim Komunikata Publishing House.
- Trevor G Bond and Christine M Fox. 2013. *"Applying The Rasch Model: Fundamental Measurement In The Human Sciences"*. Psychology Press.
- Untary, H. 2020. *"Analisis Data Penelitian dengan Winstep dan Ministep"* (H. Untary, ed.). Bogor: Halaman Moeka Publishing.
- Wahyuningsih, & Azizah. 2020. Penggunaan Model Rasch Untuk Menganalisis Instrumen Tes Pada Mata Kuliah Matematika Akuatria. *Jurnal Pendidikan Matematika (JUPITEK)*. Vol. 3(1).