



Penggunaan Rasch Model dalam Menganalisis Butir Soal Matematika Kelas 5 Sekolah Dasar

Debbi Silvia Aqida^{1*}

¹Universitas Islam Negeri Sunan Kalijaga Yogyakarta

*e-mail: 22204085003@student.uin-suka.ac.id

<i>Submitted: February 17, 2024</i>	<i>Revised: March 4, 2024</i>	<i>Approved: April 22, 2024</i>
-------------------------------------	-------------------------------	---------------------------------

Abstrak: Penelitian ini dilatarbelakangi oleh banyaknya perangkat tes yang digunakan di sekolah belum diketahui kualitas butir soalnya. Oleh karena itu, hasil penilaian dapat menjadi semu atau tidak mencerminkan kemampuan sebenarnya dari siswa. Penelitian ini dilakukan untuk menganalisa butir soal yang digunakan untuk mengukur kesesuaian butir soal dan tingkat kesukaran soal pada peserta didik kelas V mata pelajaran matematika di Sekolah Dasar Swasta Pekalongan. Metode penelitian yang digunakan bersifat evaluasi dengan pendekatan deskriptif kuantitatif. Model rasch digunakan untuk mendapatkan butir soal yang fit. Analisa ini dilakukan dengan bantuan *software Winsteps*. Dari output program Winsteps diperoleh hasil 15 butir soal sesuai dengan model Rasch dengan nilai Outfit MNSQ sebesar 1,03. Sedangkan nilai Outfit ZSTD sebesar 0,0. Serta *Pt-measure Corr* yang bernilai positif ada 11 butir soal, dan 4 butir soal bernilai negative. Artinya ada 4 butir soal tidak cocok terhadap model. Hasil penelitian menunjukkan bahwa soal assesmen matematika dapat memenuhi minimal satu dari tiga kriteria yang telah ditetapkan dan memiliki kesesuaian dengan karakteristik fit. Di sisi lain, satu soal dianggap mudah berdasarkan tingkat kesukaran soal.

Kata Kunci: Analisis butir soal, Soal matematika, Rasch model, Sekolah dasar

Abstract: This research is motivated by the large number of test equipment used in schools and the quality of the questions is unknown. Therefore, the assessment results may be false or not reflect the students' true abilities. This research was conducted to analyze the questions used to measure the suitability of the questions and the level of difficulty of the questions for class V students in mathematics at Pekalongan Private Elementary Schools. The research method used is evaluation with a quantitative descriptive approach. The Rasch model is used to obtain fit test items. This analysis was carried out with the help of Winsteps software. From the output of the Winsteps program, the results obtained were 15 questions according to the Rasch model with an Outfit MNSQ value of 1.03. Meanwhile, the Outfit ZSTD value is 0.0. As well as the PT-measure Corr which has a positive value, there are 11 questions and 4 questions have a negative value. This means that there are 4 questions that do not fit the model. The results of the research show that the mathematics assessment questions can meet at least one of the three predetermined criteria and are in accordance with the fit characteristics. On the other hand, one question is considered easy based on the difficulty level of the question.

Keywords: Item analysis, Mathematical test, Rasch model, Elementary school

PENDAHULUAN

Salah satu upaya yang dapat dilakukan agar Indonesia mampu bersaing di era globalisasi adalah dengan meningkatkan kualitas pendidikan. Proses pembelajaran adalah komponen penting yang menentukan kualitas pendidikan. Peran pendidik tidak hanya terbatas pada desain dan pelaksanaan pembelajaran, tetapi mereka juga bertanggung jawab untuk mengevaluasi proses dan hasil belajar. Evaluasi pembelajaran merupakan proses yang harus dilakukan secara teratur untuk memastikan efektivitas pengajaran yang diterapkan. Pendekatan evaluatif yang holistik dan inklusif dapat membantu pendidik memahami kebutuhan individual siswa dan menyesuaikan strategi pembelajaran secara lebih tepat. Selain itu, dengan memperhatikan hasil evaluasi, pendidik dapat mengidentifikasi area-area yang perlu ditingkatkan dalam proses pembelajaran sehingga dapat terus melakukan peningkatan demi mencapai standar pendidikan yang lebih baik.

Guru memainkan peran penting dalam evaluasi dan penilaian kemampuan siswa dalam mata pelajaran di sekolah dasar. Proses evaluasi harus dilakukan dengan benar agar dapat mengukur kemampuan siswa yang sebenarnya. Selain itu, evaluasi yang efektif membutuhkan alat tes yang baik. Hal ini karena alat tes yang tidak baik dapat menjadi salah satu kesulitan maupun kesalahan peserta didik dalam memahami soal terutama pada soal-soal HOTS (Anggarani & Hidayat, 2022). Guru dapat menilai hasil belajar siswa mereka, yang merupakan langkah penting dalam memperbaiki proses pembelajaran. Hal ini juga dapat berfungsi untuk meningkatkan kualitas belajar siswa. Proses evaluasi harus dilakukan secara sistematis, bertahap, dan terus menerus agar tujuan dapat dicapai (Syahputri, 2018).

Penilaian proses pembelajaran memiliki tiga tujuan: mengidentifikasi masalah belajar peserta didik, mengukur peningkatan secara bertahap, dan mengumpulkan informasi yang dapat digunakan oleh peserta didik untuk meningkatkan prestasi mereka. Sampai saat ini, penilaian pembelajaran yang dilakukan oleh guru terutama terkait dengan proses perolehan atau pengumpulan data tentang kemampuan belajar peserta didik. Pada praktiknya, tes dapat digunakan untuk mengukur kompetensi tertentu selama proses pembelajaran. Diharapkan bahwa metode tes yang tepat dapat dengan akurat mengukur kemampuan peserta didik (Dunn et al., 2003).

Tes adalah bagian terpenting dari pembelajaran. Suatu alat diperlukan untuk mengevaluasi kemampuan siswa untuk tes. Tes adalah alat yang digunakan untuk mengukur kemampuan peserta didik yang sebenarnya. Tes juga merupakan salah satu cara untuk mengukur kompetensi seseorang setelah mempelajari sesuatu (Utomo, 2022). Tiga kriteria harus dipenuhi oleh perangkat tes yang berkualitas tinggi. Teks harus sesuai dengan materi yang diujikan (*validitas isi*), memiliki konstruk yang baik (*validitas konstruk*), dan jelas

(*realibilitas*) (Sumarni, 2019). Tes dianggap reliabel jika dilakukan berulang kali pada peserta yang sama atau berbeda dan hasilnya konsisten (Retnawati, 2016).

Pada kenyataannya, banyak perangkat tes belum diketahui kualitas soalnya, yang mengakibatkan penilaian semu yang berdampak pada kemampuan siswa yang sebenarnya tidak dapat diukur. Hal ini dapat disebabkan oleh ketidakmampuan guru dalam merancang dan mengevaluasi instrumen ujian. Di sini, evaluasi dilakukan untuk mengetahui apakah tujuan pembelajaran tercapai atau tidak dan sejauh mana kualitas soal yang dibuat (Priowuntato, 2020).

Analisis butir soal dapat membantu meningkatkan kualitas soal melalui revisi atau penghapusan soal yang tidak efektif. Selain itu, analisis ini dapat digunakan sebagai alat diagnostik untuk menentukan apakah siswa sudah memahami materi yang telah diajarkan. Kegiatan menganalisis butir soal menjadi kegiatan yang wajib dilakukan guru dengan tujuan untuk meningkatkan kualitas soal yang telah disusun dan bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan yang buruk (Ida & Musyarofah, 2021). Dalam pendidikan, ada dua cara untuk melakukan analisis butir soal: klasik dan modern. Pendekatan klasik melibatkan penelaahan butir soal melalui informasi dari jawaban siswa untuk meningkatkan kualitas butir soal yang relevan dengan menggunakan teori tes klasik (Retnawati, 2016). Pendekatan modern menggunakan Winstep untuk melakukan analisis butir soal. Winsteps adalah perangkat lunak berbasis Windows yang membantu komputasi Rasch model, terutama di bidang evaluasi pendidikan, survei sikap, dan analisis skala. Analisis menggunakan model Rasch memberikan informasi yang lebih komprehensif terkait karakteristik instrumen yang dikembangkan (Asdar & Afriadi, 2022, p. 91; Sabekti & Khoirunnisa, 2018).

Rasch model digunakan untuk menganalisis hasil tes jawaban pilihan ganda dalam penelitian ini. Teori respons butir termasuk model Rasch, yang dapat mengklasifikasikan perhitungan item dan orang dalam peta distribusi (Rozeha et al., 2007). Menurut Sumintono dan Widhiarso (2015), tujuan utama model Rasch yaitu untuk menghasilkan suatu skala pengukuran dengan interval yang sama yang selanjutnya dapat memberikan informasi yang akurat tentang peserta tes maupun kualitas soal yang dikerjakan (Tarigan et al., 2022, p. 93). Terdapat dua prinsip menentukan model Rasch. Prinsip pertama adalah kemampuan subjek, yaitu kemampuan siswa untuk menjawab pertanyaan dengan menggunakan beberapa faktor yang dikenal sebagai *traits*. Traits adalah aspek kemampuan seseorang yang dapat mencakup keterampilan psikomotorik, verbal, dan kognitif. Prinsip kedua menjelaskan hubungan antara kemampuan siswa. Kurva karakteristik butir dapat digambarkan sebagai hubungan antara kemampuan siswa dalam suatu tugas atau pertanyaan dengan kemampuan mereka yang berbeda (Hambleton et al., 1991).

Menurut model rasch, siswa yang memiliki kemampuan yang tinggi seharusnya memiliki kemungkinan yang lebih besar untuk menjawab soal dengan benar daripada siswa lainnya. Sebaliknya, siswa dengan tingkat kesulitan yang lebih tinggi memiliki kemungkinan yang lebih rendah untuk menjawab soal dengan benar (Sumintono & Widhiarso, 2015). Metode model Rasch mempertimbangkan aspek respons dan korelasi selain item (Ardiyanti, 2016).

Analisis butir soal dalam pembelajaran matematika sangat penting bagi guru, terutama karena matematika sering dianggap sebagai mata pelajaran yang sulit bagi sebagian besar siswa sekolah dasar. Analisis dapat membantu guru menemukan masalah siswa dalam belajar matematika. Banyak guru pendidikan dasar belum melakukan analisis rincian soal. Hal ini disebabkan oleh fakta bahwa jika analisis dilakukan secara manual, terlalu banyak perhitungan yang harus dilakukan, dan beberapa guru khawatir bahwa soal yang mereka buat akan bocor. Inilah mengapa ujian yang diberikan kepada siswa masih berkualitas rendah (Mustafidah & Harjono, 2019).

Hasil observasi dan wawancara dengan guru mata pelajaran matematika di SD Salafiyah fityatul Huda menyatakan bahwa, dirinya dan hampir semua guru tidak melakukan analisis butir soal yang dibuatnya, karena hal tersebut memerlukan waktu luang yang cukup lama. Guru tersebut mengakui memang seharusnya perlu melakukan analisis butir soal guna mengetahui kualitas sebuah soal yang dibuatnya. (hasil observasi dengan guru matematika pada tanggal 1 mei 2023).

Penulis juga menemukan beberapa penelitian terdahulu yang berkaitan dengan analisis butir soal menggunakan rasch model. Pertama penelitian yang dilakukan oleh Anis Fauziana dan Andhita Dessy Wulansari tahun 2021 yang berjudul “Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch”. Penelitian tersebut melibatkan satu parameter untuk mengukur tingkat kesulitan soal pada materi lingkaran kelas VI di SDN Klakahrejo 1/578. Hasil penelitian menunjukkan bahwa semua soal yang diujikan pada peserta tes adalah valid dan reliabilitas. Hasil tingkat kesukaran dari 10 soal yang diuji adalah 4 soal tergolong mudah, 4 soal tergolong sedang dan 2 soal tergolong sangat sukar (Fauziana & Wulansari, 2021). Penelitian tersebut memiliki persamaan serta perbedaan dengan penelitian yang dilakukan oleh penulis. Persamaan penelitian tersebut dengan penelitian ini yaitu sama-sama menggunakan metode deskriptif kuantitatif. Sedangkan perbedaannya terletak pada jumlah soal yang dianalisis serta tingkatan kelas yang berbeda.

Kedua penelitian dari Ainun Nurul Syadiah dan Ghullam Hamdu tahun 2020 yang berjudul “Analisis Rasch untuk Soal tes berpikir kritis pada pembelajaran STEM di Sekolah Dasar. Penelitian tersebut bertujuan untuk menganalisis bagaimana kemampuan siswa kelas

5 sekolah dasar di Tasikmalaya, Jawa Barat yang diukur dengan menggunakan soal tes berpikir kritis dalam pembelajaran STEM (*Science, Technology, Engineering, Mathematic*). Hasil tes tersebut di analisis dengan model Rasch sebagai alat untuk mengukur kesesuaian butir soal dan abilitas siswa, sehingga seorang guru akan lebih mudah untuk mengidentifikasi kesalahan yang terjadi (Syadiah & Hamdu, 2020a). Penelitian tersebut memiliki persamaan serta perbedaan dengan penelitian yang dilakukan penulis. Persamaan penelitian tersebut dengan penelitian ini yaitu sama-sama menggunakan model rasch untuk menganalisis butir soal dengan bantuan aplikasi winstep. Perbedaannya yaitu penelitian tersebut lebih menekankan pada tingkatan jenis soal mulai dari soal tingkat sangat tinggi sampai soal tingkat rendah. Sehingga, hasil tersebut mampu menggambarkan bagaimana keterampilan yang ada pada soal untuk melihat sejauh mana soal itu dibuat dalam indikator berpikir kritis. Sedangkan pada penelitian yang dilakukan penulis hanya menganalisis kriteria soal itu fit atau tidak dengan melihat syarat fit, serta hanya menjelaskan kriteria sebuah soal sulit dan mudah. Oleh karena itu, menganalisis butir soal penting untuk dilakukan, maka tujuan dari penelitian ini adalah untuk mengetahui kualitas butir soal sumatif di Kelas V Matematika di SD Salafiyah Fityatul Huda Pekalongan Tahun Ajaran 2022/2023.

METODE

Penelitian ini menggunakan metodologi penelitian yang bersifat evaluasi dengan pendekatan deskriptif kuantitatif. Penelitian dilaksanakan secara kuantitatif namun tidak untuk menolak atau menerima hipotesis melainkan guna menganalisis butir soal, kesesuaian butir soal, dan tingkat kesukaran soal. (Rukajat, 2018). Penelitian ini dilaksanakan di SD Salafiyah Fityatul Huda. Data yang dianalisis adalah soal matematika kelas 5 dengan jumlah 15 butir soal. Soal berupa pilihan ganda dengan skor benar 1 dan skor salah 0. Sehingga data yang diperoleh merupakan data dikotomi. Soal tersebut diujikan kepada 22 peserta didik.

Suatu butir soal dapat dikatakan baik apabila memenuhi kriteria uji kecocokan butir dengan memperhatikan hasil analisis *Outfit MNSQ*, *Outfit ZSTD*, dan *Pt-measure Corr*. Selain itu, butir soal dapat dikatakan baik apabila memiliki tingkat kesukaran yang baik dalam uji karakteristiknya. Analisis data dilakukan dengan menggunakan software Winstep. Dari output software Winsteps diperoleh beberapa parameter butir soal yang fit dengan *model Rasch*. Selain itu diperoleh nilai *Outfit MNSQ*, *Outfit ZSTD* dan nilai korelasi butir dengan soal secara keseluruhan menunjukkan batas item yang dinyatakan fit dengan model. Syarat butir dikatakan fit terhadap model pada program winstep antara lain jika nilai *Outfit MNSQ* sebesar 0,5 sampai dengan 1,5 dan nilai *Outfit ZSTD* sebesar -2 sampai dengan 2, serta *Pt-measure Corr* bernilai positif maka dapat dikatakan butir tersebut fit atau cocok terhadap model (Sumintono & Widhiarso, 2015a).

HASIL DAN PEMBAHASAN

Hasil

Uji Kecocokan Butir (Model Fit)

Uji kecocokan butir pada penelitian ini menggunakan asumsi *rasch model*. Uji kecocokan butir dengan asumsi pendekatan *rasch model* dilakukan dengan melihat fit atau tidaknya butir terhadap model. Uji ini dianalisis dengan menggunakan Program Winstep. Syarat butir dikatakan fit terhadap model pada program winstep antara lain jika nilai *Outfit MNSQ* sebesar 0,5 sampai dengan 1,5 dan nilai *Outfit ZSTD* sebesar -2 sampai dengan 2, serta *Pt-measure Corr* bernilai positif maka dapat dikatakan butir tersebut fit atau cocok terhadap model (Sumintono & Widhiarso, 2015a). Adapun item akan dikatakan fit apabila terpenuhinya salah satu dari ketiga syarat tersebut. Berikut adalah hasil analisis kecocokan butir menggunakan program winstep disajikan pada Tabel 1.

Tabel 1. Kecocokan butir instrumen Penilaian Context

Outfit MNSQ	Outfit ZSTD	PT-Measure Corr	Item
1.01	.1	.36	Item_01
1.12	.5	.10	Item_02
.67	-.8	.60	Item_06
.85	-.3	.35	Item_07
1.12	-1.8	.27	Item_09
.65	.4	.61	Item_11
1.74	-.9	-.24	Item_04
.72	1.4	.44	Item_08
1.42	-.5	-.10	Item_05
.45	.8	.64	Item_13
2.32	-.6	-.25	Item_03
.36	1.9	.61	Item_10
.45	.3	.53	Item_12
1.47	-1.7	-.08	Item_14
Minimum Measure		.00	Item_15

Tabel 1. diatas merupakan tabel hasil analisis kecocokan butir atau model fit instrumen *Penilaian Context*. Keocokan model pada tahap ini menggunakan *Outfit MNSQ*, *Outfit ZSTD*, atau *Pt-Measure corr*. Berdasarkan hasil analisis dari 15 butir pertanyaan instrumen *Penilaian Context*, dapat disimpulkan bahwa keseluruhan butir cocok terhadap model atau model fit. Karena telah memenuhi syarat butir dikatakan fit terhadap model pada program winstep, antara lain jika nilai *Outfit MNSQ* sebesar 0,5 sampai dengan 1,5 atau nilai *Outfit ZSTD* sebesar -2 sampai dengan 2, atau *Pt-measure Corr* bernilai positif maka dapat dikatakan butir tersebut fit atau cocok terhadap model (Sumintono & Widhiarso, 2015a).

Karakteristik butir soal

Setelah melakukan analisis kecocokan model atau model fit maka tahap selanjutnya yaitu penilaian terhadap karakteristik butir dalam instrumen yang dikembangkan. Karakteristik butir soal pada tahap ini melihat tingkat kesulitan butir soal, hal ini dikarenakan pengembangan instrument yang digunakan dalam penelitian menggunakan Pemodelan Rasch. Adapun hal ini dikarenakan Pemodelan Rasch hanya mengukur tingkat kesulitan butir. Berdasarkan Rasch model, *Item* dikatakan baik jika indeks kesukaran lebih dari -2,0 atau kurang dari 2,0 (Hambleton et al., 1991). Analisis karakteristik butir soal pada tahap ini menggunakan program winstep, dimana tingkat kesukaran dilihat pada *measure*. Berikut adalah hasil analisis karakteristik butir soal pada masing-masing instrument.

Butir yang dianalisis pada instrumen *Penilaian Context* sebanyak 15 butir. Adapun output hasil analisis karakteristik butir dapat dilihat pada Tabel 2.

Tabel 2. Output hasil analisis karakteristik butir instrumen Penilaian Context

Entry Number	Total Score	Total Count	Measure	Model S.E
1	13	22	1.35	.46
2	16	22	.66	.51
6	17	22	.39	.54
7	17	22	.39	.54
9	17	22	.39	.54
11	17	22	.39	.54
4	18	22	.08	.58
8	18	22	.08	.58
5	19	22	-.29	.65
13	19	22	-.29	.65
3	20	22	-.78	.76
10	20	22	-.78	.76
12	20	22	-.78	.76
14	20	22	-.78	.76
15	22	22	-2.79	1.82

Output tingkat kesukaran butir pada Tabel 2 sudah berurutan, mulai dari tingkat kesukaran tertinggi sampai terendah. Karakteristik butir soal pada tahap ini melihat tingkat kesukaran butir. Tingkat kesukaran butir dilihat berdasarkan hasil pada kolom *measure*. Berdasarkan Tabel 2 diatas menunjukkan bahwa butir dengan tingkat kesukaran tertinggi adalah butir ke 1 yaitu dengan indeks tingkat kesukaran sebesar 1,35, sedangkan untuk tingkat kesulitan terendah adalah butir ke 15 yaitu dengan indeks tingkat kesukaran sebesar -2,79. Namun hasil tersebut menunjukkan bahwa hanya 14 butir instrumen *Penilaian Context* memenuhi kriteria tingkat kesukaran yang baik, dan 1 butir instrument *Penilaian Context*

nomor 15 tidak memenuhi kriteria tingkat kesukaran yang baik karena indeks kesukaran lebih dari -2,0. Walaupun hanya 14 butir soal yang memenuhi kriteria tingkat kesukaran yang baik, namun 15 soal tersebut sudah memiliki kualitas butir soal yang fit. Karena 15 butir soal tersebut telah memenuhi kriteria minimal satu dari ketiga kriteria yang telah ditentukan. Syarat butir dikatakan fit terhadap model pada program winstep antara lain jika nilai *Outfit MNSQ* sebesar 0,5 sampai dengan 1,5 dan nilai *Outfit ZSTD* sebesar -2 sampai dengan 2, serta *Pt-measure Corr* bernilai positif maka dapat dikatakan butir tersebut fit atau cocok terhadap model (Sumintono & Widhiarso, 2015a).

Pembahasan

Rasch Model memungkinkan analisis yang lebih mendalam terhadap setiap butir soal matematika. Model ini tidak hanya mempertimbangkan jumlah jawaban yang benar, tetapi juga kompleksitas butir soal dan pola jawaban siswa secara individual. Hal ini memungkinkan identifikasi butir soal yang efektif dan memberikan informasi tentang tingkat kesulitan soal. Dari hasil analisis yang sudah dilakukan, dari 15 butir soal yang sudah dianalisis terdapat 14 butir soal yang memenuhi kriteria tingkat kesukaran yang baik, dan 1 butir soal tidak memenuhi kriteria tingkat kesukaran yang baik. Salah satu keunggulan utama Rasch Model adalah kemampuannya untuk memisahkan antara kemampuan siswa dan tingkat kesulitan butir soal. Dengan demikian, model ini dapat memberikan gambaran yang lebih akurat tentang seberapa baik sebuah soal dapat mengukur kemampuan siswa, tanpa terpengaruh oleh tingkat kesulitan soal itu sendiri. Dengan menggunakan Rasch Model, dapat diidentifikasi dengan lebih tepat butir soal mana yang memiliki masalah, seperti tingkat kesulitan yang tidak sesuai atau pola jawaban yang ambigu. Hal ini memungkinkan guru untuk melakukan revisi atau pengembangan butir soal yang lebih baik.

Lima belas pertanyaan diberikan kepada siswa untuk menganalisis, membuktikan, dan memberikan deskripsi sejauh mana siswa dapat menjawab soal dengan tingkat kesukaran atau *higher order thinking skill*. Keterampilan utama yang harus dimiliki dalam konteks abad 21 adalah keterampilan belajar dan berinovasi. Keterampilan ini berkenaan dengan kemampuan berfikir kritis. HOTS akan memampukan siswa dalam mengontruksi argumen yang tepat dan efektif untuk membuat keputusan atau solusi yang rasional. Kemampuan berpikir tingkat tinggi diperlukan siswa untuk mengerjakan model penilaian di abad 21 (Syadiah & Hamdu, 2020).

Dalam konteks model Rasch, pola penskoran yang 'menetap' ini tidak lain adalah pengukuran yang hasilnya bergantung pada siapa yang diukur (*test dependent scoring*); sedangkan yang harus dilakukan dalam riset kuantitatif dalam penilaian pendidikan adalah pengukuran yang objektif (*objective measurement*) (Sumintono & Widhiarso, 2015b). Hasil dari

data ini dapat memberikan informasi kepada guru agar lebih mudah untuk melakukan koreksi pada soal yang telah di buat agar tidak membuat soal dengan kesalahan yang sama. Maka analisis soal dengan pemodelan Rasch memberikan banyak gambaran yang bisa di peroleh oleh siapapun. Sebagian guru yang melakukan kegiatan pembelajaran dengan standar kurikulum nasional, pemikiran tingkat tinggi didekati sebagai "ujung tombak" Taksonomi Bloom (atau lainnya): Menganalisis, Mengevaluasi, dan Menciptakan, atau, dalam bahasa yang lebih tua, Analisis, Sintesis, dan Evaluasi (Liu et al., 2015).

SIMPULAN

Dari hasil uraian diatas dapat disimpulkan bahwa soal assesmen sekolah dasar mata pelajaran Matematika memiliki kualitas butir soal yang fit karena telah memenuhi kriteria minimal satu dari ketiga kriteria yang telah ditentukan. Di sisi lain pada tingkat kesukaran soal, ada satu soal yang tergolong mudah. Karena didapati semua siswa menjawab benar dengan satu nomor soal.

Saran dan harapan yang penulis inginkan terhadap pendidik maupun secara umum yaitu: (1) Perbaiki butir soal yang memiliki tingkat kesukaran yang tidak proporsional: Jika terdapat butir soal yang terlalu mudah atau terlalu sulit, langkah-langkah perbaikan harus dilakukan. Butir soal yang terlalu mudah dapat menunjukkan kurangnya tantangan bagi siswa, sedangkan butir soal yang terlalu sulit dapat menyebabkan frustrasi. Dengan merevisi butir soal yang tidak proporsional, tes akan lebih akurat dalam mengukur kemampuan siswa secara proporsional. (2) Tingkatkan daya pembeda butir soal: Jika terdapat butir soal dengan daya pembeda yang rendah, langkah-langkah perbaikan harus diambil. Daya pembeda yang rendah menandakan bahwa butir soal tidak efektif dalam membedakan antara siswa yang memiliki tingkat kemampuan yang berbeda. Dalam pengembangan tes selanjutnya, perhatian khusus harus diberikan pada penyusunan butir soal yang memiliki daya pembeda yang lebih tinggi. (3) Sebaiknya pihak sekolah mengadakan workshop atau pelatihan penyusunan instrument penilaian dan analisis butir soal bagi para guru dengan menghadirkan ahli/pakar guna meningkatkan kualitas penilaian/evaluasi pembelajaran di masa yang akan datang.

DAFTAR PUSTAKA

Anggarani, N., & Hidayat, R. A. (2022). Analisis Kesalahan Siswa dalam Mengerjakan Soal pada Penjumlahan dan Pengurangan Bilangan Bulat Berbasis HOTS pada Siswa Kelas V. *Indonesian Journal of Islamic Elementary Education*, 2(1), 49–58. <https://doi.org/https://doi.org/10.28918/ijjee.v2i1.5154>

- Ardiyanti, D. (2016). Aplikasi model rasch pada pengembangan skala efikasi diri dalam pengambilan keputusan karir siswa. *Jurnal Psikologi*, 43(3), 248–263. <https://doi.org/10.22146/jpsi.17801>
- Asdar, A. K., & Afriadi, B. (2022). Standardisasi Instrumen Integritas Akademik melalui Pendekatan Confirmatory Factor Analysis dan Rasch Model. *Jurnal Manajemen Pendidikan*, 13(2), 88–108. <https://doi.org/https://doi.org/10.21009/jmp.v13i2.31446>
- Dunn, L., Morgan, C., O'Reilly, M., & Parry, S. (2003). *The student assessment handbook: New directions in traditional and online assessment*. Routledge.
- Fauziana, A., & Wulansari, A. D. (2021). Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch. *Jurnal Ibriez: Jurnal Kependidikan Dasar Islam Berbasis Sains*, 6(1), 10–19. <https://doi.org/10.21154/ibriez.v6i1.112>
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Vol. 2). Sage.
- Ida, F. F., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab*, 1(1), 34–44. <https://doi.org/https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Liu, X., Huang, G., Huang, H., Wang, S., Xiao, Y., & Chen, W. (2015). Safety climate, safety behavior, and worker injuries in the Chinese manufacturing industry. *Safety Science*, 78, 173–178. <https://doi.org/10.1016/j.ssci.2015.04.023>
- Mustafidah, H., & Harjono, H. (2019). Implementasi Program QUEST untuk Menganalisis Butir Soal bagi Guru-Guru SMP Muhammadiyah 2 Karanglewas. *JPPM (Jurnal Pengabdian dan Pemberdayaan Masyarakat)*, 3(2), 321–328. <http://dx.doi.org/10.30595/jppm.v3i2.3378>
- Prijowuntato, S. W. (2020). *Evaluasi pembelajaran*. Sanata Dharma University Press.
- Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian (panduan peneliti, mahasiswa, dan psikometrian)*. Parama Publishing.
- Rozeha, A. R., Azami, Z., & Mohd Saidfudin, M. (2007). Application of Rasch Measurement in Evaluation of Learning Outcomes: A case study in electrical engineering. *Regional Conference on Engineering Mathematics, Mechanics, Manufacturing & Architecture (EM3ARC)*.
- Rukajat, A. (2018). *Pendekatan penelitian kuantitatif: quantitative research approach*. Deepublish.
- Sabekti, A. W., & Khoirunnisa, F. (2018). Penggunaan Rasch model untuk mengembangkan instrumen pengukuran kemampuan berikir kritis siswa pada topik ikatan kimia. *Jurnal Zarah*, 6(2), 68–75. <https://garuda.kemdikbud.go.id/documents/detail/817565>
- 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>

- Sumintono, B., & Widhiarso, W. (2015a). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim Komunikata.
- Syadiah, A. N., & Hamdu, G. (2020a). Analisis rasch untuk soal tes berpikir kritis pada pembelajaran STEM di sekolah dasar. *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, 10(2), 138–148. <https://dx.doi.org/10.25273/pe.v10i2.6524>
- Syahputri, N. (2018). Rancang Bangun Media Pembelajaran Matematika Sekolah Dasar Kelas 1 Menggunakan Metode Demonstrasi. *Jurnal Sistem Informasi Kaputama (JSIK)*, 2(1). <https://doi.org/10.59697/jsik.v2i1.797>
- Tarigan, E. F., Nilmarito, S., Islamiyah, K., Darmana, A., & Suyanti, R. D. (2022). Analisis Instrumen Tes Menggunakan Rasch Model dan Software SPSS 22.0. *Jurnal Inovasi Pendidikan Kimia*, 16(2), 92–96. <https://doi.org/https://doi.org/10.15294/jipk.v16i2.30530>
- Utomo, B. (2022). Analisis validitas isi butir soal sebagai salah satu upaya peningkatan kualitas pembelajaran di madrasah berbasis nilai-nilai islam. *Jurnal Pendidikan Matematika (Kudus)*, 1(2). <https://doi.org/http://dx.doi.org/10.21043/jmtk.v1i2.4868>