

PENGGUNAAN MODEL RASCH UNTUK MENDETEKSI KUALITAS SOAL KIMIA SMA PADA UJIAN AKHIR SEMESTER (UAS)

Lisa Ariyanti Pohan^{1*)}, Kiki Ledy²⁾, Vera Roni Setiawan²⁾, Ayi Darmana²⁾, Retno Dwi Suyanti²⁾

¹⁾Prodi Pendidikan Kimia, Fakultas Keguruan dan Ilmu Pendidikan Universitas Islam Sumatera Utara, Medan, Indonesia

²⁾Prodi Pendidikan Kimia, Universitas Negeri Medan, Sumatera Utara, Indonesia

*e-mail: lisa.ariyantipohan@fkip.uisu.ac.id.

(Received 24 Mei 2022, Accepted 10 Juni 2022)

Abstract

The investigation of the test quality for the Final Semester Examination was carried out to explore the conformity of the students' grade of ability with the test presented. The tests were used to assess student's competence in chemistry subjects in the science of class XII of SMA Negeri 2 Lhokseumawe. The tests set are in the form of multiple choice questions totaling 25 related to integrated calculation and theory material in chemistry subject matter from class X, XI and XII and given to 75 students. The Model Rasch is used to acquire fit items. From the output of the Winsteps program, the results of the final semester exam (UAS) in chemistry subjects were tested using the Model Rasch to produce data, namely the item score (item reliability) of 0.77 with consistent information that the answers were in the sufficient category and the reliability of the person (person reliability) was 0.95 with information on the quality of the items in the special category and Cronbach's alpha value of 0.85. The Outfit MNSQ value for the suitability of the questions and the level of difficulty of the questions is $(1.54 = 1.5)$ in the person column. The Outfit ZSTD value for the suitability of the tests and the grade of difficulty of the tests in the item column is 0.55. The Outfit MNSQ value for students' abilities is 1.36 in the person column. The Outfit ZSTD value for students' abilities in the item column is 0.19. The test questions used for chemistry subjects are 25 questions, after being tested with a fit Model Rasch, 18 questions (72%) can be used and as many as 7 questions (28%) are not fit or cannot be used.

Keyword: *model rasch, kualitas soal, reliabilitas, fit.*

Abstrak

Penyelidikan kualitas soal Ujian Akhir Semester dilakukan untuk mengetahui kesesuaian tingkat kemampuan siswa dengan soal yang disajikan. Kemampuan siswa dalam memahami mata pelajaran kimia bagi siswa di jurusan IPA kelas XII diukur dengan menggunakan Soal Ujian Akhir Semester (UAS) yang dilaksanakan di SMA Negeri 2 Lhokseumawe. Soal diujikan kepada 75 orang siswa di SMA Negeri 2 Lhokseumawe, menggunakan soal pilihan ganda berjumlah 25 butir soal yang memuat materi perhitungan dan teori dalam mata pelajaran kimia selama mereka belajar di SMA. Kualitas butir soal diukur menggunakan Rasch Model untuk mengetahui butir soal yang fit. Diperoleh hasil output dengan bantuan winsteps berupa reliabilitas item (item reliability) sebesar 0,77 dengan konsistensi jawaban siswa berada pada kategori cukup dan reliabilitas person sebesar 0,95, dengan keterangan kualitas butir soal berada pada kategori istimewa serta nilai alpha Cronbach sebesar 0,85. Nilai Outfit MNSQ menunjukkan kesesuaian soal dan tingkat kesukaran soal sebesar $(1,54 = 1,5)$ pada kolom person. Sedangkan nilai Outfit ZSTD menunjukkan kesesuaian soal dan tingkat kesukaran soal pada kolom item sebesar 0,55. Nilai Outfit MNSQ untuk kemampuan siswa sebesar 1,36 pada kolom person. Pada kemampuan siswa kolom item diperoleh nilai Outfit sebesar 0,19. Soal ujian yang digunakan untuk mata pelajaran kimia berjumlah 25 soal, setelah diuji dengan Model Rasch yang fit dapat digunakan 18 butir soal (72 %) dan sebanyak 7 butir soal (28 %) tidak fit atau tidak dapat digunakan.

Kata Kunci: *rasch model, reliabilitas, kualitas soal*

PENDAHULUAN

Upaya guru dalam memastikan tingkat pencapaian kemampuan siswa setelah proses belajar mengajar dilakukan menggunakan sistem penilaian sangat penting dilaksanakan terhadap materi ajar yang diampu (Riadi, 2017). Salah satu indikator evaluasi yang baik dapat dilaksanakan dengan mengukur kemampuan siswa adalah menggunakan soal-soal tes. Proses tes evaluasi ini sangat penting dilaksanakan untuk mengetahui sejauh mana daya serap siswa dalam memahami materi pelajaran yang telah dilakukan di kelas. Evaluasi terhadap matapelajaran yang telah diajarkan kepada siswa sangat penting dilakukan untuk mendapatkan informasi tentang kualitas pembelajaran yang terlaksana sehingga guru mendapatkan gambaran tentang rencana selanjutnya yang perlu dilakukan terkait peningkatan kualitas belajar siswa. Adapun pengukuran kemampuan siswa untuk menilai kemampuan belajarnya dilakukan dengan metode tes.

Soal tes ini terdiri dari campuran soal latihan yang digunakan pada mata pelajaran kimia yang dikembangkan dan digabungkan menjadi soal UAS pada pelajaran kimia disusun berasal dari materi yang dipelajari di mata pelajaran kimia kelas XII. Soal dikembangkan secara khusus dengan memadukan materi pelajaran kimia secara heterogen untuk menilai kompetensi siswa dalam mempelajari kimia SMA. Soal tes kimia dibuat dengan memperhatikan kriteria kontruk dan deskriptif, dalam hal ini juga memperhatikan kriteria kesesuaian soal, tingkat kesukaran soal serta kemampuan siswa dalam menjawab soal ujian akhir sekolah ini. Soal yang mempunyai kriteria konstruk yang baik akan memberikan indikasi terhadap jawaban, mulai dari tahapan rendah ke tahapan yang lebih tinggi dari jawaban yang akan didapatkan (Dunn et al, 2003). Adapun aspek penekanan dalam penelitian ini adalah bagaimana mendeteksi kualitas soal serta kemampuan siswa dalam menjawab soal ujian, dimana soal ujian ini memuat materi kimia baik itu teori maupun perhitungan yang pernah diajarkan selama siswa menempuh jenjang sekolah menengah atas. Soal diberikan kepada siswa kelas XII SMA Negeri 2 Lhokseumawe, kemudian soal tersebut dianalisis menggunakan Model Rasch untuk mengetahui kualitas butir soal yang telah di ujikan kepada siswa tersebut.

Teori penilaian modern yang terkenal saat ini adalah Model Rasch, dalam model ini dapat mengklasifikasikan perhitungan antara item soal dan orang yang menjawab (person) dalam hal ini siswa pada suatu gambar peta terdistribusi (Rozehaet al, 2007). Thissenet al, (2001) menyatakan bahwa model Rosch adalah bagian dari teori respon butir soal yang di berikan kepada peserta tes. Model Rasch ini menitik beratkan pengukuran pada hubungan antara kemampuan orang yang menjawab soal (*person ability*) terhadap tingkat kesukaran soal yang diberikan (*item difficulty*) menggunakan fungsi logaritma sehingga menghasilkan pengukuran dengan inteval yang sama (Olsen, 2003). Model Rasch dapat mendeteksi kemampuan (*ability*) yang tinggi pada peserta tes melalui kemungkinan yang lebih besar untuk menjawab soal dengan benar dbandingkan dengan peserta tes lainnya. Pada soal yang mudah dengan tingkat kesukaran yang rendah memiliki peluang besar siswa menjawab dengan benar sedangkan soal dengan tingkat kesulitan tinggi akan terjadi hal sebaliknya. (Sumintono & Widhiarso, 2014; Azizah & Wahyuningsih, 2020). Dalam analisis soal menggunakan model rasch aspek respond dan korelasi menjadi dua hal penting yang diperiksa (Ardiyanti, 2016). Kelebihan model Rasch dibandingkan m=dengan model klasik enurut Aziz (2015) adalah kemampuan model tersebut dalam menganalisis data yang hilang atau missing data.

Teori klasik dalam penelitian analisis butir soal mempunyai keterbatasan yaitu karakteristik butir soal sangat tergantung pada kelompok peserta tes. Lain halnya dengan analisis butir soal menggunakan teori modern/teori respon butir soal. Kemampuan individu

peserta tes menjadi dasar penilaian bukan kemampuan kelompok peserta tes (Ratnaningsih, & Isfarudi, 2013). Penelitian yang dilakukan oleh Hayati dan Lailatussaadah (2016) menyatakan bahwa instrumen soal akan baik jika diperiksa kualitas soal, reabilitas dan validitas soal secara berkala. Instrumen dikatakan sah dan dapat digunakan dalam penelitian adalah instrumen yang memperoleh indeks reabilitas soal dan indeks reabilitas respon berada pada kategori tinggi.

Penelitian yang dilakukan oleh Sabekti & Khoirunnisa (2018) dalam menganalisis kompetensi dasar dan kemampuan berfikir kritis menunjukkan hanya satu soal yang tidak memenuhi kriteria soal baik dari 13 soal yang di ujikan, sedangkan 12 item soal sudah baik berdasarkan analisis Model Rasch dengan reabilitas instrumen 0,72 berada pada kategori cukup. Hasil penelitian Purba (2008) diperoleh nilai outfit ZSTD sebesar -0,7; outfit MNSQ sebesar 1,10; nilai point Measure Correlation 0,4 sampai 0,85; reabilitas sebesar 0,99 dan nilai Alpha Cronbach sebesar 0,99 dianalisis pada mata pelajaran dasar dan pengukuran listrik, dari hasil tersebut diperoleh kategori soal sangat baik.

Pada penelitian lainnya juga menampilkan data instrumen pengukuran atribut tes harus valid dan reliabel. Penelitian ini dilakukan sejak tahunuji validitas dan reliabilitas item kimia yang digunakan oleh testee diperlukan. Kajian ini bertujuan untukmemperkirakan validitas dan menentukan reliabilitas instrumen uji kimia berorientasi Next Generation Standar Sains (NGSS). Penelitian ini dilakukan melalui pendekatan deskriptif kuantitatif dalamdua sekolah kejuruan program studi teknik yang memiliki 130 testee. Instrumen yang digunakan adalahInstrumen tes kimia berorientasi NGSS berisi 35 butir soal dan angket validasi ahli. Itudiperoleh tanggapan peserta tes dari instrumen tes dikumpulkan melalui dokumentasimetode. Butir soal tes NGSS dipresentasikan kepada tiga ahli materi. Validitas yang digunakan untuk menguji sis dan kontruksi dari soal yang digunakan. Keandalan diuji untuk melihat konsistensi internal dan pendekatan konsistensi antar penilai. Hasil penelitian menunjukkan bahwa validitas isi (Aiken's V) berada pada kisaran 0,50ke 1.00. Nilai unexplained variance kurang dari 10% yang berarti berkategori baik.Analisis ini diperkuat dengan CFA yang memiliki goodness of fit dan model pengukuran yang baik. ItuParameter yang digunakan untuk menguji model fit adalah CFI, NFI, RMSEA dan nilai loading factor. Beberapa hasilnilai lebih dari 0,90 dan RMSEA adalah 0,00 dan lebih dari 0,3 nilai faktor pemuatan pada setiap item. Semua skala memiliki reliabilitas alpha lebih dari kriteria 0,70. Dengan demikian, soal tes kimia yang dikembangkan terbukti sebagaiinstrumen yang valid dan reliabel (Lia et al, 2020).

Pada penelitian yang dilakukan oleh Ramadhani & Fitri (2020) diperoleh hasil uji validitas instrumen tes dan reabilitas hanya 24%, maka perlu direvisi 19 intrumen tes untuk memenuhi kriteria instrumen tes yang valid dan reliabel dan dites pada tahap berikutnya.. Hasil penelitian lainnya menunjukkan ada sembilan item soal merupakan kategori berkualitas baik pada materi dimensi *compulsive internet use* dan bersifat unidimensi, serta nilai reliabilitas item lebih baik dibandingkan dengan penelitian sebelumnya yang menggunakan *confirmatory factor analysis*. Instrumen yang cocok mengukur *problematic internet use* pada mahasiswa di Indonesia adalah menggunakan IPIUS (Natanael, 2021).

Pada penelitian sebelumnya juga dapat dilihat, untuk menguji kemampuan berfikir tingkat tinggi/HOTS kimia larutan berbasis tauhid, instrumen uji HOTS menggunakan MFRM menunjukkan hasil yang menarik dan bagus yang dapat memberikan informasi yang lengkap tentang kualitas pertanyaan. Bias dapat dihindari dengan penilaian multirater dan dapat dianalisis lebih tepat melalui MFRM di mana interaksi item-item, item dan penilai adalah dikalibrasi pada skala interval yang sama. Hasil analisis menunjukkan bahwa (1) terdapat 17 dari40 item dianggap kurang baik oleh ahli (42,5%), sedangkan 23 dari 40 item (57,5%) dianggapkualitas baik; (2) menurut kelima validator, butir-butir instrumen yang paling sulit/yang sulit dipenuhi dari soal adalah B4, K4, dan K8 (harga logitnya paling

besar/lebih besardari +1), sedangkan K10 (harga logit paling rendah/kurang dari -4) merupakan item pernyataan yang paling mudah untuk memenuhi dari item; (3) Dari aspek kecocokan statistik dapat diketahui bahwa Val.A dan Pakar Val.F adalah yang paling konsisten (nilai Pakaian MNSQ dan Pakaian ZSTD sama-sama cocokkisan barang); (4) reliabilitas antar penilai diperoleh nilai sangat baik (0,97) dan data dalam sesuai dengan model pengukuran yang terbukti signifikan dari uji Chi-square nilai ($p < 0,01$); (5) kesamaan penilaian oleh lima ahli (kesepakatan yang tepat) yang mencapai 93,0% yang menunjukkan bahwa tidak banyak perbedaan antara kelima ahli dalam menilai semua item (Darmana et al, 2021).

Penelitian yang dilakukan oleh Fauziana dan Wulansari (2021) menunjukkan kemampuan siswa berda pada level sangat baik, diperoleh melalui tes evaluatif dari 10 soal yang diujikan dimana 4 soal digolongkan kedalam kategori mudah, 4 soal dalam kategori sedang dan 2 soal dalam kategori sangat sulit. Dari hasil analisis Model Rasch juga diperoleh estimasi nilai kemampuan 31 orang siswa yang mengikuti tes bervariasi berada pada range - 0,69 sampai dengan 3,05. Purba et al (2014) meneliti mengenai kemampuan calon guru di lembaga penyelenggara pendidikan terhadap soal yang diujikan memberi gambaran akan kualitas soal yang digunakan meliputi: tingkat kesulitan butir soal, kemampuan dan tingkat pengetahuan peserta tes serta interaksi antara butir soal dan peserta tes.

Pengolahan data Model Rasch dilakukan melalui program aplikasi *Winsteps*, kesesuaian data dengan model yang digunakan dapat dilihat berdasarkan nilai outfit MNSQ dimana nilai mean square sama dengan satu (jika kurang dari satu disimpulkan variasi soal yang digunakan lebih sedikit dibandingkan prediksi yang akan dilakukan). Dalam penelitian ini yang menjadi titik berat penelitian adalah bagaimana mendeteksi kualitas soal meliputi reliabilitas butir soal dan butir soal yang fit yaitu butir soal yang baik digunakan sebagai instrumen penilaian. Dimana soal tes dibuat dengan memadukan kumpulan soal-soal latihan kimia yang kemudian dikembangkan menjadi Soal ujian akhir sekolah untuk kelas XII jurusan IPA. selanjutnya akan dianalisis butir soal yang layak digunakan (soal fit) dan butir soal yang tidak layak digunakan (soal tidak fit) sebagai instrumen analisis menggunakan model Rasch. Untuk reliabilitas butir soal di lihat dari nilai alfa cronbach yang dihasilkan dengan bantuan aplikasi software *Winstep*.

METODE

Penelitian ini menggunakan metode penelitian deskriptif. Tujuan utama penelitian untuk mendeteksi kualitas butir soal menggunakan Model Rasch. Sampel merupakan siswa kelas XII jurusan IPA di SMA Negeri 2 Lhokseumawe tahun ajaran 2021-2022 yang mengikuti mata pelajaran kimia sebanyak 75 siswa. Adapun jumlah soal yang diujikan berjumlah 25 soal, dimana soal tersebut telah dirancang dengan level kesulitan tertentu, mulai dari level rendah, sedang hingga tinggi. Pengujian soal terhadap siswa dilakukan pada Ujian Akhir Semester (UAS). Kriteria soal yang diberikan adalah soal pilihan ganda, dengan skor satu jika menjawab benar dan skor nol jika menjawab salah, dimana data yang diperoleh bersifat dikotomi. Untuk mendeteksi kualitas butir soal menggunakan Model Rasch dibantu menggunakan software *Winsteps*. Hasil tes berupa skor dianalisis dengan bantuan software tersebut. Adapun kualitas butir soal yang diteliti meliputi: analisis butir soal fit dan tidak fit melalui model Rasch serta reliabilitas butir soal menggunakan outfit nilai cronbach. Kesesuaian soal, kesukaran soal dan kemampuan siswa dilihat dari nilai outfit MNSQ dan nilai outfit ZSTD yang dihasilkan. Soal dikatakan fit apabila memenuhi kriteria nilai Outfit ZSTD, outfit MNSQ dan outfit korelasi butir soal. Kriteria soal fit dapat dilihat dari hasil analisis berdasarkan software yaitu untuk outfit ZSTD berada pada kisaran -0,2 sampai 2,0; outfit MNSQ berada pada kisaran 0,5 sampai 1,5; sedangkan untuk korelasi butir soal berada

pada kisaran 0,4 sampai 0,85 (Sumintono & Widhiarso, 2015).

HASIL DAN PEMBAHASAN

Hasil analisis dari 25 butir soal yang diujikan kepada 75 siswa SMA Negeri 2 Lhokseumawe menggunakan *software Winsteps*, terdapat 18 Item soal yang dinilai fit dengan model Rasch sedangkan sisanya 7 item soal dinyatakan tidak fit. Berikut hasil analisis secara lengkap yang dirangkum dalam Tabel 1.

Tabel 1. Ringkasaan Statistik

| Kriteria Penilaian | Outfit yang dihasilkan | Nilai |
|---------------------------|-------------------------------|--------------|
| Reliabilitas | Person Realiability | 0.77 |
| | Item Reliability | 0.95 |
| | Alpha Cronbach | 0.85 |
| Kesesuaian Soal | Outfit MNSQ | 1.54=1,5 |
| | Outfit ZSTD | 0.55 |
| Kesukaran Soal | Outfit MNSQ | 1.54=1,5 |
| | Outfit ZSTD | 0.55 |
| Kemampuan Siswa | Outfit MNSQ | 1.36 |
| | Outfit ZSTD | 0.19 |

Tabel 1 menunjukkan interaksi antar person (siswa) dengan item (soal) secara menyeluruh dikategorikan bernilai cukup baik dengan nilai Alpha Cronbach sebesar 0,85. Kemudian nilai reabilitas person yang dihasilkan sebesar 0,77 yang dapat disimpulkan bahwa konsistensi jawaban siswa sudah cukup tinggi dengan demikian kualitas butir soal yang digunakan pada instrumen tes cukup baik.

Untuk melihat kriteria soal fit dan tidak fit didasarkan pada hasil outfit MNSQ dan ZSTTD. Dari Tabel 1 dapat dilihat nilai Outfit Mean Squared (Outfit MNSQ) sebesar 1,54. Nilai 1,54 =1,5 berada dalam kategori soal fit yang terletak pada kisaran angka $0,5 < \text{MNSQ} < 1,5$, hal ini berarti bahwa instrumen tes berupa soal kimia yang dibuat sudah sesuai dengan model rasch untuk mengukur kompetensi siswa dalam mempelajari mata pelajaran kimia SMA. Sedangkan untuk nilai Outfit Z Standardized (Outfit ZSTD) dihasilkan sebesar 0.55. Nilai ZSTD yang dihasilkan berada pada kisaran angka $-2,0 < \text{ZSTD} < 2,0$ yang termasuk dalam kriteria soal fit. Dari nilai ZSTD juga dapat disimpulkan bahwa keseluruhan item (butir soal) sesuai dengan model Rasch dan dapat dijadikan sebagai instrumen tes yang baik pada mata pelajaran kimia secara keseluruhan.

Kategori soal fit dan tidak fit dapat dilihat dari distribusi item (soal) pada Tabel 2 dibawah. Adapun kriteria soal fit jika memenuhi dua karakteristik berikut. Yang pertama batas item (soal) dikatakan fit jika nilai outfit MNSQ terletak pada kisaran angka 1,5 sampai 2,0. Syarat kedua yaitu memiliki nilai outfit ZSTD pada kisaran -2,0 sampai 2,0; dan untuk nilai *point measure correlation*/korelasi butir dengan skor total berada pada kisaran 0,4 sampai 0,85 (Sumintono & Widhiarso, 2014, p.71-72). Berikut adalah sebaran data soal yang tidak fit berdasarkan model Rasch menggunakan bantuan *software Winsteps*.

Dari hasil analisis software Winstep pada Tabel 2 diperoleh ada 7 butir soal yang tidak fit. Ketujuh butir soal tersebut yaitu nomor 7, 10, 6, 8, 24, 4 dan 1 termasuk dalam kriteria tidak memenuhi syarat soal fit sehingga item soal harus diganti karena tidak memenuhi kriteria. Untuk butir soal nomor 11 nilai outfit ZSTD sebesar 1,35 masih memenuhi syarat memenuhi kriteria Outfit ZSTD sehingga soal tersebut masih layak untuk digunakan. Jumlah butir soal fit yang memenuhi kriteria ada 18 butir soal dari 25 butir soal

mata pelajaran kimia yang diujikan pada siswa kelas XII jurusan IPA di SMA Negeri 2 Lhokseumawe.

Tabel 2. Sebaran data soal tidak fit berdasarkan Model Rasch

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | JMLE MEASURE | MODEL S.E. | INFIT | | OUTFIT | | PTMEASUR-AL | | EXACT MATCH | | Item |
|--------------|-------------|-------------|--------------|------------|-------|-------|--------|-------|-------------|------|-------------|------|------|
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD | CORR. | EXP. | OBS% | EXP% | |
| 7 | 1 | 75 | 4.97 | 1.01 | 1.05 | .37 | 6.93 | 2.54 | A-.11 | .11 | 98.7 | 98.7 | S7 |
| 10 | 10 | 75 | 2.38 | .36 | 1.04 | .23 | 3.46 | 3.12 | B .13 | .32 | 86.7 | 86.6 | S10 |
| 6 | 58 | 75 | -1.60 | .31 | 1.26 | 1.49 | 3.17 | 4.28 | C .17 | .46 | 74.7 | 80.7 | S6 |
| 8 | 7 | 75 | 2.83 | .42 | .94 | -.10 | 2.99 | 2.25 | D .17 | .28 | 90.7 | 90.6 | S8 |
| 24 | 55 | 75 | -1.32 | .30 | 1.40 | 2.47 | 2.88 | 4.45 | E .10 | .47 | 72.0 | 78.4 | S24 |
| 4 | 55 | 75 | -1.32 | .30 | 1.05 | .40 | 2.60 | 3.97 | F .31 | .47 | 82.7 | 78.4 | S4 |
| 1 | 46 | 75 | -.60 | .27 | 1.53 | 3.83 | 1.80 | 3.11 | G .12 | .49 | 58.7 | 72.9 | S1 |
| 11 | 13 | 75 | 2.02 | .33 | 1.09 | .53 | 1.63 | 1.35 | H .23 | .36 | 88.0 | 83.1 | S11 |
| 5 | 72 | 75 | -4.02 | .64 | 1.14 | .44 | .34 | -.47 | I .33 | .30 | 94.7 | 96.3 | S5 |
| 3 | 62 | 75 | -2.03 | .34 | .92 | -.34 | 1.12 | .40 | J .45 | .43 | 89.3 | 84.8 | S3 |
| 19 | 29 | 75 | .63 | .27 | .68 | -2.97 | 1.03 | .18 | K .64 | .47 | 89.3 | 72.9 | S19 |
| 22 | 17 | 75 | 1.61 | .31 | 1.03 | .27 | .95 | -.02 | L .39 | .40 | 77.3 | 79.4 | S22 |
| 21 | 52 | 75 | -1.07 | .29 | .99 | -.05 | .99 | .05 | l .49 | .48 | 74.7 | 76.3 | S21 |
| 12 | 30 | 75 | .56 | .27 | .89 | -.96 | .92 | -.26 | k .54 | .48 | 77.3 | 72.9 | S12 |
| 18 | 22 | 75 | 1.18 | .29 | .91 | -.64 | .87 | -.27 | j .49 | .44 | 77.3 | 75.6 | S18 |
| 13 | 35 | 75 | .20 | .27 | .86 | -1.18 | .88 | -.51 | i .57 | .49 | 82.7 | 72.8 | S13 |
| 25 | 60 | 75 | -1.81 | .33 | .83 | -.93 | .71 | -.72 | h .56 | .45 | 82.7 | 82.8 | S25 |
| 20 | 56 | 75 | -1.41 | .30 | .81 | -1.27 | .57 | -1.56 | g .61 | .47 | 84.0 | 79.2 | S20 |
| 9 | 62 | 75 | -2.03 | .34 | .76 | -1.25 | .48 | -1.39 | f .60 | .43 | 86.7 | 84.8 | S9 |
| 14 | 4 | 75 | 3.49 | .53 | .76 | -.49 | .24 | -1.04 | e .40 | .21 | 94.7 | 94.6 | S14 |
| 15 | 46 | 75 | -.60 | .27 | .76 | -2.09 | .64 | -1.89 | d .66 | .49 | 80.0 | 72.9 | S15 |
| 2 | 67 | 75 | -2.74 | .42 | .75 | -.87 | .38 | -1.12 | c .56 | .39 | 90.7 | 90.3 | S2 |
| 16 | 40 | 75 | -.16 | .27 | .68 | -3.03 | .64 | -1.95 | b .69 | .50 | 88.0 | 72.2 | S16 |
| 17 | 26 | 75 | .86 | .28 | .68 | -2.84 | .65 | -1.27 | a .66 | .46 | 85.3 | 73.6 | S17 |
| MEAN | 37.0 | 75.0 | .25 | .42 | .95 | -.37 | 1.54 | .55 | | | 83.6 | 81.3 | |
| P.SD | 22.4 | .0 | 2.38 | .33 | .22 | 1.57 | 1.48 | 2.00 | | | 8.5 | 7.9 | |

Dari hasil analisis butir soal menggunakan Model Rasch sebanyak 25 butir soal yang diujikan pada siswa kelas XII mata pelajaran Kimia dalam ujian akhir semester diperoleh dalam Tabel 3. Tingkat kategori kesulitan soal dibedakan atas, kelompok soal sangat sulit, sulit, mudah dan sangat mudah. Berdasarkan hasil analisis soal dengan software Winsteps dari Tabel 3 dapat dibaca data sebaran kategori tingkat kesulitan soal yaitu pada kelompok soal sangat sulit terdapat empat soal yaitu pada butir soal nomor 23, 7, 14 dan 8. Kelompok soal sulit terdapat delapan soal yaitu pada butir soal nomor 10, 11, 22, 18, 17, 19, 12 dan 13. Kelompok soal yang mudah terdapat sebelas soal yaitu pada butir soal nomor 16, 1, 15, 21, 4,

24, 20, 6, 25, 3 dan 9. Kelompok soal sangat mudah terdapat dua soal yaitu pada butir soal nomor 2 dan 5. Dengan persentase soal sangat sulit 16%, soal sulit 32%, soal mudah 44% dan soal sangat mudah sebesar 8%.

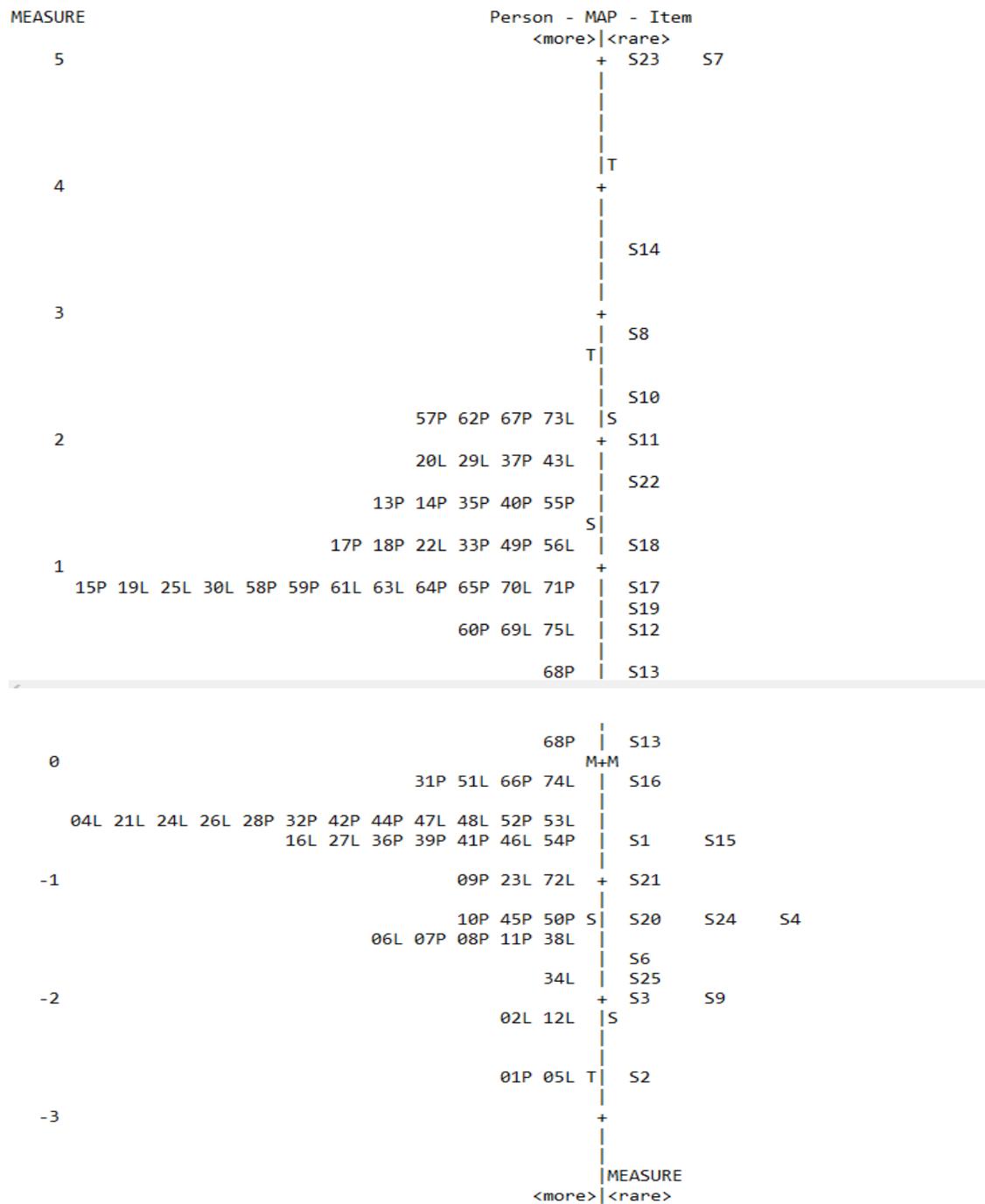
Tabel 3. Hasil Sebaran Tingkat Kesulitan Soal

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | JMLE MEASURE | MODEL S.E. | INFIT MNSQ | ZSTD | OUTFIT MNSQ | ZSTD | PTMEASUR-CORR. | AL-EXP. | EXACT OBS% | MATCH EXP% | Item |
|--------------|-------------|-------------|--------------|------------|-----------------|-------|-------------|-------|----------------|---------|------------|------------|------|
| 23 | 0 | 75 | 6.20 | 1.83 | MAXIMUM MEASURE | | | | .00 | .00 | 100.0 | 100.0 | S23 |
| 7 | 1 | 75 | 4.97 | 1.01 | 1.05 | .37 | 6.93 | 2.54 | -.11 | .11 | 98.7 | 98.7 | S7 |
| 14 | 4 | 75 | 3.49 | .53 | .76 | -.49 | .24 | -1.04 | .40 | .21 | 94.7 | 94.6 | S14 |
| 8 | 7 | 75 | 2.83 | .42 | .94 | -.10 | 2.99 | 2.25 | .17 | .28 | 90.7 | 90.6 | S8 |
| 10 | 10 | 75 | 2.38 | .36 | 1.04 | .23 | 3.46 | 3.12 | .13 | .32 | 86.7 | 86.6 | S10 |
| 11 | 13 | 75 | 2.02 | .33 | 1.09 | .53 | 1.63 | 1.35 | .23 | .36 | 88.0 | 83.1 | S11 |
| 22 | 17 | 75 | 1.61 | .31 | 1.03 | .27 | .95 | -.02 | .39 | .40 | 77.3 | 79.4 | S22 |
| 18 | 22 | 75 | 1.18 | .29 | .91 | -.64 | .87 | -.27 | .49 | .44 | 77.3 | 75.6 | S18 |
| 17 | 26 | 75 | .86 | .28 | .68 | -2.84 | .65 | -1.27 | .66 | .46 | 85.3 | 73.6 | S17 |
| 19 | 29 | 75 | .63 | .27 | .68 | -2.97 | 1.03 | .18 | .64 | .47 | 89.3 | 72.9 | S19 |
| 12 | 30 | 75 | .56 | .27 | .89 | -.96 | .92 | -.26 | .54 | .48 | 77.3 | 72.9 | S12 |
| 13 | 35 | 75 | .20 | .27 | .86 | -1.18 | .88 | -.51 | .57 | .49 | 82.7 | 72.8 | S13 |
| 16 | 40 | 75 | -.16 | .27 | .68 | -3.03 | .64 | -1.95 | .69 | .50 | 88.0 | 72.2 | S16 |
| 1 | 46 | 75 | -.60 | .27 | 1.53 | 3.83 | 1.80 | 3.11 | .12 | .49 | 58.7 | 72.9 | S1 |
| 15 | 46 | 75 | -.60 | .27 | .76 | -2.09 | .64 | -1.89 | .66 | .49 | 80.0 | 72.9 | S15 |
| 21 | 52 | 75 | -1.07 | .29 | .99 | -.05 | .99 | .05 | .49 | .48 | 74.7 | 76.3 | S21 |
| 4 | 55 | 75 | -1.32 | .30 | 1.05 | .40 | 2.60 | 3.97 | .31 | .47 | 82.7 | 78.4 | S4 |
| 24 | 55 | 75 | -1.32 | .30 | 1.40 | 2.47 | 2.88 | 4.45 | .10 | .47 | 72.0 | 78.4 | S24 |
| 20 | 56 | 75 | -1.41 | .30 | .81 | -1.27 | .57 | -1.56 | .61 | .47 | 84.0 | 79.2 | S20 |
| 6 | 58 | 75 | -1.60 | .31 | 1.26 | 1.49 | 3.17 | 4.28 | .17 | .46 | 74.7 | 80.7 | S6 |
| 25 | 60 | 75 | -1.81 | .33 | .83 | -.93 | .71 | -.72 | .56 | .45 | 82.7 | 82.8 | S25 |
| 3 | 62 | 75 | -2.03 | .34 | .92 | -.34 | 1.12 | .40 | .45 | .43 | 89.3 | 84.8 | S3 |
| 9 | 62 | 75 | -2.03 | .34 | .76 | -1.25 | .48 | -1.39 | .60 | .43 | 86.7 | 84.8 | S9 |
| 2 | 67 | 75 | -2.74 | .42 | .75 | -.87 | .38 | -1.12 | .56 | .39 | 90.7 | 90.3 | S2 |
| 5 | 72 | 75 | -4.02 | .64 | 1.14 | .44 | .34 | -.47 | .33 | .30 | 94.7 | 96.3 | S5 |
| MEAN | 37.0 | 75.0 | .25 | .42 | .95 | -.37 | 1.54 | .55 | | | 83.6 | 81.3 | |
| P.SD | 22.4 | .0 | 2.38 | .33 | .22 | 1.57 | 1.48 | 2.00 | | | 8.5 | 7.9 | |

Pada gambar 1 di perlihatkan gambar analisis data tingkat kemampuan siswa dalam menjawab soal ujian sekolah berikut. Bagian kanan peta wright dibawah ini merupakan sebaran tingkat kesulitan butir soal sedangkan bagian kiri merupakan sebaran abilitas (kemampuan) siswa.

Kemampuan siswa dalam mengerjakan soal yang dianalisis dengan bantuan software Winstep dalam bentuk peta wright dapat dilihat pada Gambar 1. Dari peta wright diperoleh data siswa dengan kode (P=Perempuan) dan (L=Laki-laki) dan nomor yang diberikan adalah nomor urut data siswa tersebut saat pengujian dalam tabel, sehingga diperoleh data 57P, 62P, 67P dan 73L memiliki kemampuan atau abilitas paling tinggi sedangkan siswa 01P dan 05L memiliki kemampuan paling rendah. Keterangan data Person reliability = 0,77 (Kategori cukup) yaitu menyatakan keterangan konsistensi jawaban berada pada kriteria cukup dan

Item reability = 0,95 (Kategori Istimewa) yaitu menyatakan keterangan kualitas butir soal berada pada kategori istimewa.



Gambar 1. Person Measure

KESIMPULAN

Berdasarkan analisis Model Rasch dari 25 butir soal yang diujikan menunjukkan 18 butir soal (72%) fit dan layak digunakan sedangkan 7 butir soal (28%) tidak fit dan tidak layak digunakan serta harus diganti. Hasil analisis butir soal mata pelajaran Kimia menggunakan Rasch Model pada Ujian Akhir Semester (UAS) diperoleh konsistensi jawaban berada pada kategori cukup dengan nilai item (*item reliability*) sebesar 0,77 dan reliabilitas person (*person reliability*) sebesar 0,95. Kualitas butir soal berada pada kategori istimewa dengan nilai alpha Cronbach sebesar 0,85. Tingkat kesukaran soal dan kesesuaian soal diperoleh nilai outfit Mean Square Statistic (Outfit MNSQ) sebesar (1,54 = 1,5) dan nilai Outfit Z Standard (ZSTD) sebesar 0,55. Untuk kemampuan siswa diperoleh nilai Outfit Mean Square Statistic (MNSQ) sebesar 1.36 pada kolom person dan nilai Outfit Z Standard (ZSTD) pada kolom item sebesar 0,19.

DAFTAR PUSTAKA

- Ardiyanti, D. (2016). Aplikasi Model Rasch pada Pengembangan Skala Efikasi Diri dalam Pengambilan Keputusan Karir Siswa. *Jurnal Psikologi*, 43(3), 248-263.
- Azizah, & Wahyuningsih, S. (2020). Penggunaan Model Rasch Untuk Analisis Instrumen Tes Pada Mata Kuliah Matematika Aktuaria. *Jurnal Pendidikan Matematika*, 3 (1), 45-50.
- Aziz, R. (2015). Aplikasi model Rasch dalam pengujian alat ukur kesehatan mental di tempat kerja. *Psiko islamika: Jurnal Psikologi dan Psikologi Islam*, 12(2), 29-39.
- Dunn, L., Morgan, C., O'Reilly, M., & Parry, S. (2003). *The student assessment handbook: New directions in traditional and online assessment*.
- Darmana, A., Sutiani, A., Nasution, H. A., Sylvia, N. A., Aminah, N., & Utami, T. (2021, March). Analisis of Multi Rater with Facetson Instruments HOTS of Solution Chemistry based on Tawheed. In *Journal of Physics: Conference Series* (Vol. 1819, No. 1, p. 012038). IOP Publishing.
- Fauziana, A., & Wulansari, A. D. (2021). Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch. *Ibriez: Jurnal Kependidikan Dasar Islam Berbasis Sains*, 6(1), 10-19.
- Hambleton R, Swaminathan H, Rogers J (1991). Background. In: Foster D, editor. *Fundamentals of Item Response Theory*. California: Sage Publications. pp. 1-6.
- Hayati, S., & Lailatussaadah, L. (2016). Validitas dan reliabilitas instrumen pengetahuan pembelajaran aktif, kreatif dan menyenangkan (pakem) menggunakan model rasch. *Jurnal Ilmiah Didaktika: Media Ilmiah Pendidikan dan Pengajaran*, 16(2), 169-179.
- Kumpulan Soal-soal Ujian Nasional (UN) Kimia SMA/MA/Sederajatnya. (2018). Erlangga-Jakarta.
- Kustriyono. (2004). Penyusunan Perangkat Soal Ujian Akhir Mata Pelajaran Sains Biologi Dalam Rangka Pengembangan Bank Soal. *Jurnal Penelitian dan Evaluasi Pendidikan* 2 (VI): 175-198.
- Lia, R. M., Rusilowati, A., & Isnaeni, W. (2020). NGSS-Oriented chemistry test instruments: validity and reliability analysis with the rasch model. *REiD (Research and Evaluation in Education)*, 6(1), 41-50.
- Natanael, Y. (2021). Analisis Rasch model Indonesia Problematic Internet Use Scale (IPIUS). *Persona: Jurnal Psikologi Indonesia*, 10(1), 167-186.
- Olsen, L. W. (2003). *Essay son Georg Rasch and his contribution stostatistics*. Un published PhD the sisat Institute Of Economics University of Copenhagen.

- Purba, S. E. D. (2018). Analisis model Rasch instrumen tes prestasi pada mata pelajaran dasar dan pengukuran listrik. *Wiyata Dharma: Jurnal Penelitian Dan Evaluasi Pendidikan*, 6(2), 142-147.
- Putra, Z. H., Hermita, N., & Alim, J. A. (2021). Analisis Pengetahuan Matematika, Didaktika, dan Teknologi Calon Guru Sekolah Dasar Menggunakan Rasch Model. *Mosharafa: Jurnal Pendidikan Matematika*, 10(3), 345-356.
- Ramadhani, R., & Fitri, Y. (2020). Validitas e-modul matematika berbasis EPUB3 menggunakan analisis rasch model. *Jurnal Gantang*, 5(2), 95-111.
- Ratnaningsih, D. J., & Isfarudi, I. (2013). Analisis butir tes objektif ujian akhir semester mahasiswa Universitas Terbuka berdasarkan teori tes modern. *Jurnal Pendidikan Terbuka dan Jarak Jauh*, 14(2), 98-109.
- Ratnaningsih, D. J., & Isfarudi, I. (2013). Analisis butir tes objektif ujian akhir semester mahasiswa universitas terbuka berdasarkan teori tes modern. *Jurnal Pendidikan Terbuka Dan Jarak Jauh*, 14 (2), 98-109.
- Riadi, A. (2017). Kompetensi Guru Dalam Pelaksanaan Evaluasi Pembelajaran. *Ittihad Jurnal Kopertais Wilayah XI Kalimantan*. 15 (28), 52-67.
- Rozeha, A. R., Azami, Z., & 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 2007 (EM3ARC).
- Routledge. Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of Item Response Theory. London: SAGE Publications. dalam Rangka Pengembangan Bank Soal. *Jurnal Penelitian dan Evaluasi Pendidikan*, 6(2).
- Sabekti, A. W., & Khoirunnisa, F. (2018). Penggunaan Rasch Model Untuk Mengembangkan Instrumen Pengukuran Kemampuan Berpikir Kritis Siswa Pada Topik Ikatan Kimia. *Jurnal Zarah*, 6(2), 68-75.
- Sumintono, B., & Widhiarso, W. (2014). Aplikasi Model Rasch Untuk Penelitian Ilmu-ilmu Sosial (edisi revisi). Bandung: Trim Komunikata Publishing House.
- Sumintono, B., & Widhiarso, W. (2015). Aplikasi Pemodelan Rasch Pada Assessment Pendidikan. Trim Komunikata.
- Swaminathan, H., Hambleton, R. K., & Rogers, H. J. (2007). Assessing the fit of item response models. *Handbook of Statistics*, 26, 683-718.
- Thissen, D., Nelson, L., Rosa, K., (2001). Item Response Theory for Items Scored in More than Two Categories dalam D. Thissen & H. Wainer, *Test Scoring* (pp. 141-184). New Jersey: Lawrence Erlbaum Associates Publishers