



Tersedia online di EDUSAINS
Website: <http://journal.uinjkt.ac.id/index.php/edusains>
EDUSAINS,13(2), 2021, 106-118



Research Artikel

APLIKASI MODEL RASCH PADA INSTRUMEN TES UNTUK MENGANALISIS KEMAMPUAN PEMECAHAN MASALAH SISWA MATERI LARUTAN ASAM BASA

APPLICATION OF THE RASCH MODEL ON THE TEST INSTRUMENT TO ANALYZE THE PROBLEM-SOLVING ABILITY OF STUDENTS IN ACID-BASE SOLUTIONS MATERIALS

Nurkintan Aprilia^{1*}, Endang Susilaningsih¹, Sudarmin¹, Woro Sumarni², F. Widhi Mahatmanti²,
Naftalina Ulik Adhelia²

¹Magister Pendidikan Kimia, Pascasarjana, Universitas Negeri Semarang, Indonesia

²Pendidikan Kimia, FMIPA, Universitas Negeri Semarang, Indonesia
nurkintanapriliana@students.unnes.ac.id

Abstract

In 21st century education, students are required to have high problem-solving skills. This study aims to reveal the quality of the test instrument with the Rasch model. Descriptive quantitative research is used with total of 38 students of class XI SMAN 2 Semarang became the subjects in this study. The data obtained in this study is from a combination of student response patterns on the product development diagnostic test instrument. The data analysis used was IRT Rasch. Rasch model analysis for summary statistics, item-fit, wright-map, item-measure, person-measure, person-fit dan item-DIF. The value of person reliability is 0.78(sufficient), and item reliability value is 0.8(good). Cronbach's alpha value is 0.82 (very good). Based on the analysis Rasch model, it was found that there were two students (12P and 37P) whose response patterns were not ideal. The results showed problem-solving ability profile of students in the very high category of 26%, high of 29%, average of 16%, low of 13%, and very low of 16%. The results of this study indicate that the test instrument can be used for further research in measuring student problem solving and the Rasch model contributes to teachers and schools in identifying students' abilities in more detail.

Keywords: *problem-solving; rasch model analysis; test instruments.*

Abstrak

Pada pendidikan abad-21, siswa dituntut harus memiliki keterampilan *problem-solving* yang tinggi. Penilaian keterampilan *problem-solving* dapat dilakukan secara online yang terintegrasi pada soal tes pengetahuan. Penelitian ini bertujuan untuk mengungkap kualitas instrumen tes dengan model Rasch. Penelitian kuantitatif deskriptif digunakan dalam penelitian ini. Total 38 siswa kelas XI SMAN 2 Semarang menjadi subjek dalam penelitian ini. Data yang diperoleh dalam penelitian ini adalah dari kombinasi pola respon jawaban siswa pada instrumen tes diagnostik produk pengembangan. Analisis data yang digunakan menggunakan *Item Response Theory* (IRT) model Rasch. Analisis model Rasch untuk *summary statistics, item-fit, wright-map, item-measure, person-measure, person-fit dan item-DIF*. Nilai *person reliability* didapatkan sebesar 0,78, kriteria cukup, dan nilai *item reliability* sebesar 0,85, kriteria bagus. Nilai alpha Cronbach sebesar 0,82, kriteria bagus sekali. Berdasarkan hasil analisis Rasch model ditemukan ada dua siswa yang pola responnya tidak ideal yaitu siswa 12P dan 37P. Hasil penelitian menunjukkan profil kemampuan *problem-solving* siswa dengan kategori sangat tinggi sebesar 26%, tinggi sebesar 29%, rata-rata sebesar 16%, rendah sebesar 13%, dan sangat rendah sebesar 16%. Hasil penelitian ini menunjukkan bahwa instrumen tes dapat digunakan untuk penelitian lanjutan dalam mengukur pemecahan masalah siswa dan model Rasch memberikan kontribusi kepada guru dan sekolah dalam mengidentifikasi kemampuan siswa lebih rinci.

Kata Kunci: analisis model rasch; instrumen tes; pemecahan masalah.

Permalink/DOI: <http://doi.org/10.15408/es.v13i2.21375>

*Corresponding author

EDUSAINS, p-ISSN 1979-7281 e-ISSN 2443-1281

This is an open access article under CC-BY-SA license (<https://creativecommons.org/licenses/by-sa/4.0/>)

PENDAHULUAN

Pada pendidikan abad-21, siswa dituntut memiliki pondasi daya saing yang tinggi (Putri *et al.*, 2016). Salah satu pondasi yang dimaksud adalah keterampilan pemecahan masalah (Stehle & Peters-Burton, 2019). Keterampilan pemecahan masalah yang tinggi harus dimiliki oleh siswa di Indonesia agar dapat bersaing di dunia kerja. Keterampilan pemecahan masalah erat kaitannya dengan keterampilan berpikir. Siswa yang memiliki keterampilan pemecahan masalah yang tinggi akan memiliki keterampilan berpikir yang tinggi pula (Astuti *et al.*, 2020). Solusi yang digunakan dalam mewujudkan keterampilan abad ke-21 adalah dengan mengembangkan penilaian kemampuan pemecahan masalah. Tujuan pengembangan penilaian kemampuan pemecahan masalah adalah untuk memudahkan guru dalam proses pembelajaran yaitu dalam mengukur seluruh kompetensi siswa (Sari *et al.*, 2020).

Berdasarkan hasil wawancara dengan beberapa guru kimia menunjukkan bahwa belum banyak guru yang mengembangkan dan menggunakan instrumen penilaian kemampuan pemecahan masalah. Banyak guru yang masih mengalami kesulitan mencari kata kerja operasional yang sesuai dengan domain/indikator pada taksonomi bloom dan tingkat kognitif yang akan diukur. Hal inilah yang menyebabkan guru mengalami kesulitan menyusun instrumen yang baik sesuai dengan kaidah penyusunan instrumen penilaian yang baik. Guru khawatir jika instrumen yang disusun tidak valid dalam mengukur hasil belajar siswa terutama kemampuan pemecahan masalah siswa.

Selain itu, pengembangan penilaian kemampuan *problem-solving* merupakan keterampilan metakognitif yang sangat penting (Lestari *et al.*, 2019). Oleh karena itu, dalam pengembangan instrumen tes berbasis *problem-solving* memperhatikan setiap indikator *problem-solving* yang meliputi: (1) menemukan dan memahami masalah; (2) mengembangkan strategi *problem-solving* yang baik; (3) mencari solusi; (4) memikirkan dan mendefinisikan kembali masalah dan solusi dari waktu ke waktu (Widiasih *et al.*, 2018).

Instrumen tes produk pengembangan yang digunakan dalam penelitian ini merupakan instrumen pilihan ganda dua tingkat. Dimana tingkat kedua dimaksudkan untuk menganalisis keterampilan pemecahan masalah siswa (Afnia & Istiyono, 2020). Berdasarkan hasil penelitian Singaravelu (2017), yang menggunakan instrumen tes *problem-solving* yang dikembangkan oleh Darchhingpuii untuk menganalisis kemampuan pemecahan masalah kimia di India menyarankan berdasarkan hasil penelitiannya yang sudah dilakukan, bahwa guru harus memberikan latihan kasus atau masalah pada soal yang lebih bervariasi untuk siswa, demi mengembangkan daya pikir kreatif dan kemampuan pemecahan masalah siswa. Berdasarkan uraian tersebut maka instrumen tes produk pengembangan yang difokuskan pada pengukuran kemampuan *problem-solving* siswa menyajikan butir soal yang bervariasi yang terdiri dari 25 butir soal yang diintegrasikan dengan indikator pemecahan masalah dan indikator kompetensi dasar materi kimia yaitu larutan asam basa.

Instrumen tes yang dikembangkan dianalisis kemampuan pemecahan masalah siswa berdasarkan pola jawaban siswa menggunakan *Item Response Theory* (IRT) model Rasch. Analisis model Rasch membantu untuk menganalisis lebih jauh tentang kualitas item sampai dengan mengidentifikasi kemampuan konsep siswa (Bohori & Winny, 2019). Dimana kemampuan konsep ini erat kaitannya dengan kemampuan pemecahan masalah siswa. Siswa yang tinggi pemahaman konsepnya akan lebih mudah memecahkan masalah dengan baik. Rusmansyah & Almubarak (2020) pernah menganalisis pengetahuan kognitif siswa menggunakan analisis model Rasch dan menyimpulkan bahwa analisis Rasch efektif digunakan untuk menganalisis pengetahuan kognitif siswa secara rinci.

Yasin *et al.* (2018) menjelaskan bahwa model pengukuran Rasch merupakan model yang dibentuk sebagai hasil pertimbangan meliputi kemampuan masing-masing responden yang menjawab tes dan kesulitan setiap butir item tes. Winarti & Mubarak (2020) juga menegaskan bahwa model Rasch memberikan deskripsi yang

lebih komprehensif dan konkrit dalam pengukuran instrumen tes, hal ini karena model Rasch melibatkan dua aspek parameter yaitu (a) abilitas siswa, dan (b) tingkat kesulitan soal (Mariam *et al.*, 2018; Suryani, 2018).

Prinsip model Rasch mendeskripsikan bahwa tingkat kemampuan siswa yang lebih besar dibanding siswa lain dapat memungkinkan siswa untuk mendapatkan peluang yang lebih besar juga dalam menjawab satu butir soal dengan benar, yang artinya bahwa siswa yang memiliki kemampuan pemecahan masalah yang lebih tinggi dibanding siswa lain memungkinkan untuk dapat menjawab benar semua butir soal yang tersedia dengan mudah (Sumintono & Widhiarso, 2013; Wigati & Aini, 2020).

Oleh karena itu, tujuan dari penelitian ini adalah untuk (1) mengetahui kualitas instrumen tes berdasarkan hasil analisis model Rasch, (2) mendapatkan informasi kemampuan dan keterampilan pemecahan masalah siswa dari hasil analisis model Rasch.

METODE

Penelitian kuantitatif deskriptif digunakan dalam penelitian ini, tujuannya untuk mendapatkan informasi tentang kualitas butir soal dengan analisis model Rasch dan keterampilan pemecahan masalah siswa (Misbah *et al.*, 2018). Data yang diperoleh dari penelitian ini berupa pola respon jawaban siswa berdasarkan hasil pengerjakan tes berbasis pemecahan masalah, yang kemudian data tersebut dianalisis menggunakan analisis model Rasch. Model Rasch menekankan bahwa setiap siswa memiliki kesempatan yang sama dalam menjawab soal benar dan soal yang memiliki tingkat kesulitan yang berbeda yang disebut Rasch sebagai *person logit* dan *item logit* (Sihombing *et al.*, 2019).

Subjek dalam penelitian ini adalah siswa kelas XI SMAN 2 Semarang yang berjumlah 38 siswa. Pengumpulan data dalam penelitian ini menggunakan metode tes. Instrumen tes materi larutan asam basa yang dikembangkan mengadopsi 3 tahap dari Treagust (1988, 1995) yaitu tahap mendefinisikan konten, mengidentifikasi kelemahan konsep siswa, dan pengembangan instrumen tes.

Penyusunan kisi-kisi soal dimulai dengan menganalisis silabus pokok bahasan larutan asam basa, kemudian membedah kompetensi dasar menjadi indikator kompetensi dasar larutan asam basa yaitu sebanyak tujuh indikator meliputi (1) sifat larutan asam basa; (2) hubungan daya hantar listrik dan kekuatan asam-basa; (3) reaksi autoionisasi air dan Kw; (4) kekuatan asam basa; (5) derajat keasaman (pH); (6) indikator asam basa; dan (7) aplikasi pH dalam kehidupan sehari-hari. Ke-tujuh Indikator Kompetensi Dasar (IKD) tersebut kemudian dikaitkan dengan indikator pemecahan masalah siswa sebanyak empat indikator (Widiasih *et al.*, 2018).

Analisis IRT model Rasch menganalisis estimasi *separation item reability*, *item fit*, *wright map*, *item measure*, *person measure*, *person fit* dan deteksi butir soal bias (*item DIF*) (Aprilia *et al.*, 2020; Nurwahidah *et al.*, 2020). Analisis model Rasch dapat memberitahu bahwa ada individu yang pola responnya tidak sesuai dan butir soal yang tidak valid (*outliers* atau *misfits*). Menurut Boone *et al.* (2014), terdapat tiga kriteria yang dipakai untuk memeriksa kesesuaian butir soal yang tidak sesuai (*outliers* atau *misfits*) serta individu yang pola responnya tidak sesuai (tidak fit), yaitu nilai MNSQ: $0,5 < MNSQ < 1,5$; *Outfit ZSTD*: $-2,0 < ZSTD < +2,0$; dan *Pt Measure Corr*: $0,4 < Pt Measure Corr < 0,85$. Jika ada butir soal yang tidak memenuhi ketiga kriteria maka butir soalnya harus diubah, karena pemahaman siswa diuji melalui butir soal yang berkualitas.

Tabel 1 menunjukkan kategori kelompok soal berdasarkan tingkat kesulitannya. Tingkat kesulitan butir soal dikategorikan berdasarkan nilai rata-rata *logit* dan nilai standar deviasi pada *item measure* yaitu sebesar 1,09.

Tabel 1. Kategori Kelompok Soal Berdasarkan Tingkat Kesulitannya

Nilai Logit	Kategori
Lebih besar dari +1,09 SD	Sangat Sulit
0,0 logit +1,09 SD	Sulit
0,0 logit -1,09 SD	Sedang
Lebih kecil dari -1,09 SD	Mudah

(Sumintono & Widhiarso, 2015)

Modifikasi dilakukan karena berdasarkan nilai indeks pemisahan item hanya dapat dikelompokkan ke dalam dua sampai tiga tingkat kesulitan, dimana pengelompokkan ini tidak

menunjukkan proporsi yang sesuai sehingga dalam modifikasi tingkat kesulitan butir ini dikelompokkan menjadi empat kategori.

Sedangkan pada *person measure* nilai Standar Deviasi (SD) yang diperoleh sebesar 1,11 dan nilai rata-rata *logit person* sebesar 0,28. Tabel 2 berikut menjelaskan kriteria pengelompokan abilitas siswa.

Tabel 2. Kriteria Pengelompokan Abilitas Siswa

Nilai <i>Logit</i> Abilitas Siswa	Kriteria
Lebih besar dari +1,11	Tinggi
Lebih kecil dari +1,11	Sedang
Lebih kecil dari -0,28	Rendah

(Sumintono & Widhiarso, 2015)

Modifikasi dilakukan karena berdasarkan nilai indeks pemisahan orang hanya dapat dibedakan menjadi dua kelompok dalam data, dimana pengelompokan ini akan berpengaruh pada pembentukan grafik fungsi informasi pada kurva normal sehingga dalam modifikasi ini dikelompokkan menjadi tiga kriteria.

HASIL DAN PEMBAHASAN

Dalam analisis data model Rasch ini menggunakan *software* yang didesain khusus untuk Rasch yaitu *software* Ministep. Kemampuan maksimal olah data menggunakan Ministep sebanyak 25 butir dan 75 siswa (Sumintono & Widhiarso, 2015). Pada penelitian ini, instrumen tes yang dikembangkan adalah sebanyak 25 butir soal dengan 38 responden.

Ringkasan Statistik

Dalam pemodelan Rasch reliabilitas ditunjukkan dengan nilai reliabilitas individu dan butir serta nilai alpha Cronbach sebagai reliabilitas keseluruhan (Pratama & Husnayaini, 2020; Shuib *et al.*, 2020). Koefisien pemisahan dan perhitungan keandalan dihitung dengan dan tanpa elemen apa pun dengan ukuran ekstrem. Karena ukuran untuk skor ekstrim tidak tepat, statistik reliabilitas yang mencakup skor ekstrim seringkali lebih rendah daripada pada non-ekstrimnya. Winsteps menghitung nilai batas atas dan bawah untuk *True Reliability*. Batas bawah adalah *Real Reliability*. Batas atas adalah *Model Reliability*. *True Reliability* yang tidak dapat diketahui terletak di

antara keduanya. Karena sumber kebisingan yang kontradiktif dihapus dari data, *True Reliability* mendekati *Model Reliability* (Sumintono & Widhiarso, 2015).

Data reliabilitas dianalisis menggunakan model Rasch melalui *Summary Statistics*. Hasil *output* reliabilitas model Rasch disajikan pada Tabel 3.

Tabel 3. Ringkasan Statistik Instrumen

Informasi	Nilai	Kriteria
<i>Person Reliability</i>	0,78	Cukup
<i>Item Reliability</i>	0,85	Bagus
<i>Cronbach Alpha</i>	0,82	Bagus Sekali
<i>Separation Person</i>	1,91	Baik
<i>Separation Item</i>	2,42	Analisis Tingkat Individu

Tabel 3 menunjukkan indeks reliabilitas dan indeks pemisahan orang-butir. Tabel menunjukkan nilai reliabilitas orang sebesar 0,78 hal ini menunjukkan bahwa sampel yang terlibat dalam penelitian ini cukup dapat diandalkan. Nilai reliabilitas butir sebesar 0,85 yang menunjukkan bahwa kualitas butir instrumen tes yang dikembangkan memiliki keandalan yang bagus. Sedangkan untuk nilai Cronbach Alpha yang didapatkan sebesar 0,82. Hal ini menunjukkan bahwa seluruh item dalam tes sangat reliabel (Sumintono & Widhiarso, 2015).

Tabel 3 juga menunjukkan indeks pemisahan orang dan butir. Indeks pemisahan orang didapatkan sebesar 1,91, menunjukkan bahwa sekitar dua kelompok berbeda dapat diidentifikasi dalam data. Indeks pemisahan butir yang didapatkan sebesar 2,42, menunjukkan bahwa item dapat dikelompokkan ke dalam dua sampai tiga tingkat kesulitan. Informasi indeks pemisahan ini merupakan informasi tambahan yang sangat penting untuk evaluasi fungsi instrumen yang dikembangkan (Boone *et al.*, 2014).

Berdasarkan kriteria yang dikemukakan oleh para ahli, dapat disimpulkan bahwa nilai pemisahan orang dalam penelitian ini baik dan nilai pemisahan item menunjukkan kesesuaian item untuk dianalisis pada tingkat individu (Lian & Yew, 2020).

Item Fit

Tingkat kesesuaian butir menentukan validitas masing-masing butir soal (Rachman & Napitupulu, 2017). *Item fit* memberikan informasi mengenai butir soal yang berfungsi normal melakukan pengukuran atau tidak (Muller, 2020). Jika terdapat soal yang tidak *fit*, berarti terdapat indikasi terjadinya miskonsepsi pada butir soal itu (Susilaningsih *et al.*, 2021). Tentunya ini sangat membantu guru untuk memperbaiki sistem pengajarannya agar miskonsepsi bisa diminimalisir. Berdasarkan hasil analisis item fit ditemukan butir soal yang tidak valid (*unfit*) dengan kriteria seperti pada Tabel 4.

Tabel 4. Butir Soal Tidak Valid (*Unfit*)

Nomor Butir	Logit Measure	Outfit			Pt Measure Corr
		MNSQ	ZSTD		
13	2.81	3,71	2,37	0,04	
5	1.91	2,59	2,53	0,04	

Berdasarkan Tabel 4 pada hasil analisis *item fit order* terlihat bahwa butir soal yang tidak *fit* atau *outliers* adalah butir soal nomor 13 dan 5. Dimana butir tersebut tidak memenuhi tiga kriteria yang ditentukan. Oleh karena itu butir tersebut harus direvisi sebelum dilakukan uji coba selanjutnya. Gambar 1 menunjukkan butir soal nomor 13 yang tidak *fit* dan harus direvisi.

13. Suatu larutan asam klorida encer didalam label tertulis $1,0 \times 10^{-8}$ M. Sifat dari larutan tersebut adalah ... dan nilai pH yang dihasilkan sebesar...
- A. asam, pH= 8
 - B. basa, pH=8
 - C. netral, pH=7
 - D. asam, pH= Mendekati 7
 - E. asam, pH= 6

Alasan:

- A. larutan bersifat netral karena $[H^+]$ sangat kecil
- B. larutan asam klorida memiliki konsentrasi besar yaitu 10^{-8} oleh karena itu pH yang dihasilkan diatas 7
- C. $HCl \rightarrow H^+ + Cl^-$
 $[H^+] = Ma. valensi$
 $= 10^{-8} \cdot 1$
 $= 10^{-8}$
- D. $HCl \rightarrow H^+ + Cl^-$
 $[H^+] = Ma. valensi^+$
 $= 10^{-8} \cdot 1$
 $= 10^{-8}$

pH = 14-8 = 6
- E. Karena $[H^+]$ dari asam $< 10^{-7}$ maka
 $[H^+] = [H^+]$ dari asam + $[H^+]$ dari air
 $= 10^{-8} + 10^{-7}$
 $= 1,1 \times 10^{-7}$
[H^+]
 $= -\log 1,1 \times 10^{-7}$
 $= 7-\log 1,1$

Gambar 1. Butir Soal Nomor 13

Gambar 2 menunjukkan butir soal nomor 5 yang tidak *fit* dan harus direvisi.

5. Perhatikan Tabel nilai harga tetapan kesetimbangan (K_a) dari beberapa asam pada suhu dan tekanan yang sama (P, T) berikut!

No	1	2	3	4	5
Asam	HNO ₂	HClO	HCN	HCOOH	HF
K _a	$4,5 \times 10^{-4}$	$3,5 \times 10^{-8}$	$4,9 \times 10^{-10}$	$1,7 \times 10^{-4}$	$6,8 \times 10^{-4}$

Berdasarkan tabel tersebut, dapat ditarik kesimpulan bahwa urutan tingkat kekuatan asamnya adalah...

- A. HF>HNO₂>HCOOH
- B. HCN>HClO>HCOOH
- C. HF<HNO₂<HCOOH
- D. HClO>HNO₂>HF
- E. HNO₂<HCl<HCN

Alasan:

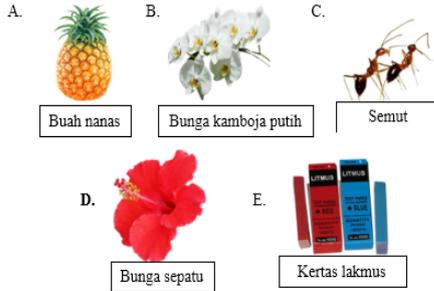
- A. semakin besar nilai K_a berarti semakin banyak ion OH⁻
- B. semakin kecil nilai K_a berarti semakin banyak ion H⁺
- C. semakin besar nilai K_a berarti semakin banyak ion H⁺
- D. larutan asam encer, sehingga derajat ionisasinya besar
- E. semakin besar nilai K_a , maka makin lemah kekuatan asam

Gambar 2. Butir Soal Nomor 5

Butir nomor 13 dan 5 harus direvisi, karena kedua butir tersebut tidak berfungsi normal melakukan pengukuran (Muller, 2020), sehingga menyebabkan miskonsepsi tinggi. Butir nomor 13 (IKD-5) meminta siswa untuk menganalisis pH asam. Sedangkan butir nomor 5 (IKD-4) meminta siswa untuk menyimpulkan kekuatan asam.

Sedangkan pada butir soal nomor 9 masih masuk ke dalam kategori butir soal yang *fit*, walaupun butir tersebut tidak memenuhi nilai *PT Measure Corr* dan *MNSQ* akan tetapi masih memenuhi nilai *ZSTD*, maka butir tersebut masih dianggap *fit*, artinya butir tetap dipertahankan dan tidak perlu diubah (revisi) (Sumintono & Widhiarso, 2015). Butir soal nomor 9 ditunjukkan pada Gambar 3.

9. Berbagai jenis zat warna yang dipisahkan dari asam atau basa organik lemah dapat digunakan sebagai indikator. Berikut ini yang termasuk ke dalam indikator alami asam basa adalah....



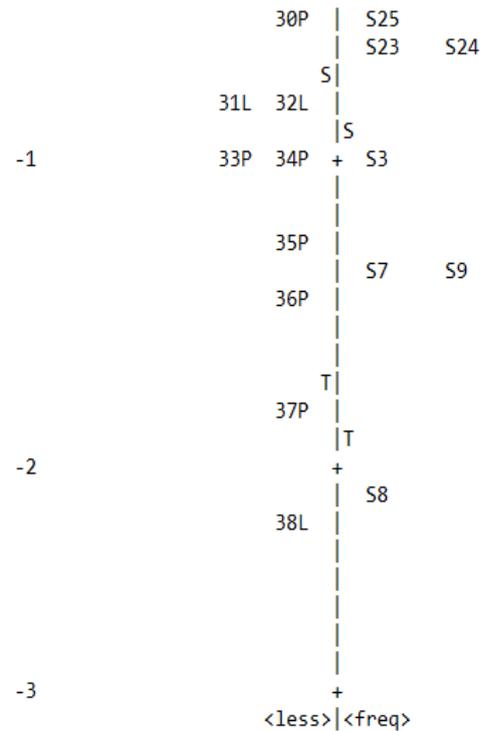
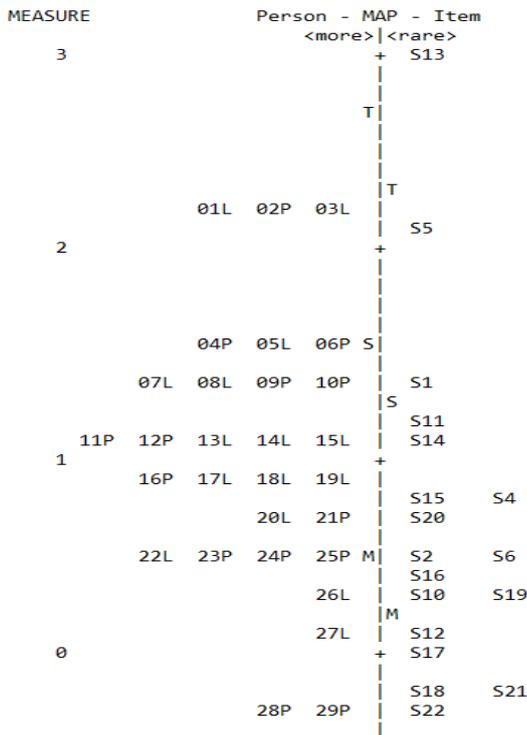
Alasan:

- A. indikator alami adalah indikator yang berasal dari alam dan harus mengandung sifat asam/basa.
- B. indikator alami adalah indikator yang berasal dari alam dan memiliki warna yang mencolok.
- C. indikator alami adalah indikator yang berasal dari hewan/tumbuhan berwarna tajam dan memiliki sifat asam atau basa.
- D. indikator alami adalah indikator yang berasal dari alam, memiliki warna mencolok dan menghasilkan warna yang berbeda ketika suasana asam atau basa.
- E. indikator alami adalah campuran dari beberapa indikator yang dapat memberikan warna berbeda pada pH berbeda.

Gambar 3. Butir Soal Nomor 9

Person-Item Map (Wright Map)

Wright Map menggambarkan sebaran kemampuan siswa dan tingkat kesulitan soal dengan skala yang sama. Artinya, analisis ini memberikan info yang sangat berharga bagi guru dalam mengidentifikasi abilitas siswa dan menganalisis kualitas butir soal yang diuji (Hilaliyah *et al.*, 2019; Ringo *et al.*, 2020). Hasil analisis peta *wright* yang terlihat pada Gambar 4.



Gambar 4. Peta *Wright*

Bagian kanan peta merupakan sebaran tingkat kesulitan butir soal sedangkan bagian kiri sebaran kemampuan siswa (Noben *et al.*, 2021). Berdasarkan peta di atas terlihat bahwa butir soal dengan tingkat kesulitan paling tinggi adalah butir soal nomor 13 sedangkan butir dengan tingkat kesulitan paling rendah atau paling mudah adalah butir soal nomor 8. Butir soal nomor 8 berada di luar batas dua deviasi standar (T). Artinya kemungkinan seluruh siswa untuk menjawab benar nomor 8 sangat besar.

Pada *wright map* terlihat bahwa siswa dengan abilitas yang paling tinggi yaitu 01L, 02P dan 03L, sedangkan siswa dengan abilitas paling rendah adalah siswa 38L. Abilitas siswa tertinggi masih masuk ke dalam kategori normal karena ada di dalam batas dua deviasi standar (T) (Ringo *et al.*, 2020). Sedangkan untuk siswa dengan abilitas terendah dimana nilai logitnya lebih kecil dari -2.0 *logit* masuk dalam kategori *outlier* karena siswa 38L berada di luar batas dua deviasi standar (T).

Selain itu, terdapat pula hubungan antara tingkat kemampuan siswa dengan tingkat kesulitan butir soal dalam *wright map* ini, yaitu berdasarkan Gambar 4 terlihat jelas bahwa siswa 01L, 02P dan 03L yang memiliki kemampuan tertinggi

dibandingkan siswa lain ternyata belum dapat menjawab dengan benar butir soal tersulit yaitu nomor 13. Hal ini karena nilai *logit* dari ketiga siswa tersebut hanya sebesar +2 *logit*, dimana nilai ini lebih rendah dibanding *logit* butir soal tersulit yaitu nomor 13 dengan nilai *logit* mencapai +3 *logit*. Sehingga wajar apabila ketiga siswa terpintar belum mampu menjawab butir soal tersulit tersebut dengan benar (Hilaliyah *et al.*, 2019).

Item Measure

Kemudian untuk mendapatkan informasi mengenai tingkat kesulitan butir soal dapat dilihat dari nilai *logit* tiap butir soal pada kolom *measure* (Hamdu *et al.*, 2020). Nilai *logit* tinggi menunjukkan tingkat kesulitan soal paling tinggi. Hasil analisis *item measure* disajikan pada Tabel 5.

Tabel 5. *Item Measure*

Nomor Butir	Total Score	Total Count	Measure	Tipe Soal
13	4	38	2,81	Sangat Sulit
5	8	38	1,91	Sangat Sulit
1	13	38	1,13	Sangat Sulit
11	14	38	0,99	Sulit
14	15	38	0,86	Sulit
4	17	38	0,59	Sulit
15	17	38	0,59	Sulit
20	18	38	0,46	Sulit
2	19	38	0,33	Sulit
6	19	38	0,33	Sulit
16	20	38	0,20	Sulit
10	21	38	0,07	Sulit
19	21	38	0,07	Sulit
12	22	38	-0,07	Sedang
17	23	38	-0,21	Sedang
18	24	38	-0,35	Sedang
21	24	38	-0,35	Sedang
22	25	38	-0,50	Sedang
25	26	38	-0,65	Sedang
23	27	38	-0,81	Sedang
24	27	38	-0,81	Sedang
3	29	38	-1,15	Mudah
7	31	38	-1,55	Mudah
9	31	38	-1,55	Mudah
8	34	38	-2,33	Mudah

Pada Tabel 5 terlihat bahwa butir dengan tingkat kesulitan tinggi adalah butir soal nomor 13 dengan nilai *logit* sebesar +2,81 *logit*. Responden yang dapat menjawab butir 13 dengan benar hanya sebanyak 4 dari total 38 responden. Sedangkan

butir dengan tingkat kesulitan yang rendah adalah butir soal nomor 8 dengan nilai *logit* sebesar -2,33. Butir soal nomor 8 merupakan butir soal yang paling mudah dikerjakan oleh responden, yaitu dibuktikan dengan 34 dari total 38 responden menjawab benar butir soal nomor 8. Sedangkan butir yang mempunyai nilai *logit* yang sama berarti kesulitan butirnya sama, seperti pada nomor 2 dan 6 dengan nilai *logit* sama yaitu +0,33

Person Measure

Pada tingkat abilitas individu digunakan untuk mengidentifikasi tingkat kemampuan siswa dalam menjawab butir soal, yang diurutkan mulai dari tertinggi sampai terendah berdasarkan nilai *logit* abilitasnya (Sudihartinih & Prabawanto, 2020). Pengelompokan abilitas siswa berdasarkan dari nilai rata-rata *logit person* dan nilai deviasi standar (SD) *person measure* yang diketahui (Sumintono & Widhiarso, 2015).

Berdasarkan Tabel 2 maka tingkat abilitas siswa terbagi menjadi tiga kategori kemampuan. Hasil pengelompokan abilitas siswa pada uji ini disajikan pada Gambar 5.



Gambar 5. Tingkat Abilitas Individu

Berdasarkan Gambar 5 didapatkan siswa dengan kategori kemampuan tinggi sebanyak 10 dari 38 siswa, siswa dengan kemampuan sedang sebanyak 17 dari 38 siswa, dan siswa dengan kemampuan rendah sebanyak 11 dari total 38 siswa. *Person measure* juga memberikan informasi tentang jumlah soal yang mampu dijawab dengan benar oleh masing-masing siswa.

Informasi *person measure* dapat dikaitkan dengan *item measure* pada Tabel 5, yaitu kemampuan siswa dalam menjawab butir soal sesuai dengan nilai *logit* abilitasnya. Akan tetapi informasi ini hanya sebagai perkiraan berdasarkan

nilai *logit* yang didapatkan. Informasi lebih jelasnya dapat diketahui melalui skalogram pada *person fit*. *Person measure* dan *item measure* juga menampilkan nilai *logit* yang lebih rinci daripada yang disajikan pada *wright map* pada Gambar 4.

Person Fit

Person fit ini bisa dideteksi apabila ada individu yang pola jawabannya berbeda. Pola jawaban berbeda yang dimaksud ini adalah terdapat ketidaksesuaian jawaban siswa berdasarkan nilai *logit* abilitasnya apabila dibandingkan dengan model ideal (Yazdani *et al.*, 2015), misalnya terlihat siswa yang tidak hati-hati dalam menjawab, siswa yang main tebakkan maupun siswa yang teridentifikasi menyontek (Holster & Lake, 2016). *Person Fit Order* ini akan memberikan informasi secara berurutan yaitu butir soal yang mempunyai kriteria tidak *fit* akan ditampilkan dibagian atas (Sumintono & Widhiarso, 2015).

Tabel 6 dibawah ini menyajikan cuplikan hasil analisis *person fit* yang memberikan informasi siswa yang memiliki pola respon yang tidak ideal berdasarkan kriteria menurut Boone *et al.* (2014).

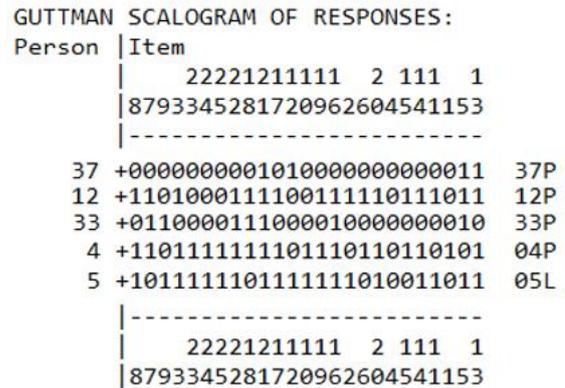
Tabel 6. Hasil Analisis *Person Fit Order*

Nomor Siswa	Measure	Outfit		
		MNSQ	ZSTD	Pt. Measure Corr
37P	-1,99	7,55	4,31	-0,43
12P	0,90	2,04	2,41	-0,17
05L	1,39	1,73	1,44	0,06
04P	1,39	1,56	1,19	0,19
33P	-1,16	1,68	1,44	0,18

Berdasarkan Tabel 6 terlihat bahwa siswa 37P dan 12P pola responnya tidak *fit*. Kedua siswa tersebut tidak memenuhi ketiga kriteria menurut Boone *et al.* (2014), sedangkan tiga siswa lainnya yaitu siswa 05L, 33P, dan 04P tidak memenuhi 2 kriteria, hanya nilai ZSTD yang terpenuhi. Akan tetapi ketiga siswa tersebut tetap harus diperiksa pola responnya melalui skalogram. Dan beberapa siswa lainnya hanya tidak memenuhi satu kriteria saja, sehingga pola responnya masih ideal.

Informasi pola respon secara rinci dapat diketahui melalui *scalograms* (Wigati & Aini, 2020). Melalui *scalograms* inilah dapat diketahui

penyebab mengapa pola respon siswa tidak *fit* atau tidak ideal. Gambar 6 dibawah ini menyajikan skalogram yang menampilkan pola respon kelima siswa pada Tabel 6.



Gambar 6. Informasi Skalogram

Siswa 37P dan 12P memiliki nilai *logit* abilitas sebesar -1,99 dan +0,90 *logit* yang termasuk ke dalam kategori siswa dengan abilitas rendah untuk 37P dan sedang untuk 12P. Terlihat pada *scalogram* (Gambar 6) kedua siswa ini dapat menjawab dengan benar soal nomor 13 dan 5. Dimana soal nomor 13 dan 5 merupakan butir dengan nilai *logit* tertinggi atau butir yang masuk dalam kategori sangat sulit (Tabel 5) yang memiliki nilai *logit* berturut-turut +2,81 dan +1,91 *logit*. Nilai *logit* abilitas kedua siswa 37P dan 12P jauh lebih rendah dibanding nilai *logit* kedua soal tersebut.

Oleh karena itu, seharusnya jika berdasarkan pola respon ideal, siswa 37P dan 12P tidak dapat menjawab butir 13 dan 5 dengan benar. Seharusnya hanya soal nomor 8 yang dapat dijawab benar oleh siswa 37P, akan tetapi ternyata soal nomor 8 sebagai soal termudah tidak dapat dijawab benar oleh siswa 37P. Begitu pula dengan siswa 12P, seharusnya hanya soal nomor 13, 5, 1 dan 11 yang tidak dapat dijawab benar oleh siswa 12P, tetapi kenyataannya siswa 12P dapat menjawab soal nomor 13, 5 dan 11 dengan benar, sedangkan beberapa soal lainnya yang memiliki nilai *logit* lebih kecil dibanding nilai *logit* abilitasnya justru ada yang tidak dapat dijawab dengan benar oleh siswa 12P ini, seperti contohnya soal nomor 4 dan 9, siswa 12P salah dalam menjawab soal tersebut.

Hal ini mengindikasikan bahwa dalam menjawab soal nomor 13, 5 serta 11 terdapat tebakan (*lucky guess*), sedangkan dalam menjawab soal nomor 8, 4 dan 9 terdapat ketidak cermatan (*careless*). Inilah yang menyebabkan pola respon siswa 37P dan siswa 12P menjadi tidak *fit* atau tidak ideal.

Berdasarkan Gambar 6 juga didapatkan informasi bahwa siswa 04P terindikasi menebak soal nomor 13 dan siswa 05L juga terindikasi menebak soal nomor 13 dan 5. Berdasarkan nilai *logit* abilitas kedua siswa ini yaitu +1,39 *logit* seharusnya tidak dapat menjawab kedua butir tersebut dengan benar. Karena nilai *logit* kedua butir tersebut lebih tinggi dari nilai *logit* abilitas kedua siswa tersebut (dapat dilihat pada Tabel 5).

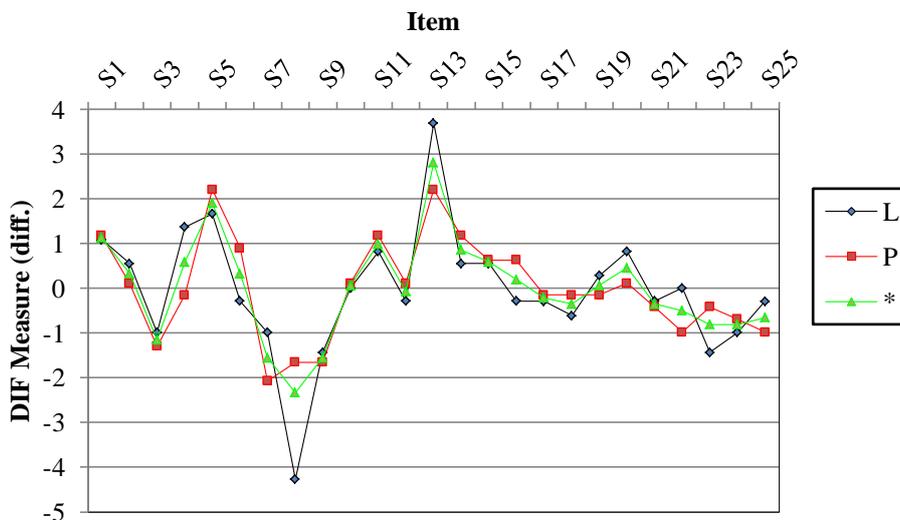
Sedangkan pada siswa 33P yang memiliki nilai *logit* -1,16 yang seharusnya hanya dapat menjawab dengan benar soal nomor 9, 7 dan 8 justru tidak dapat menjawab benar soal nomor 8, padahal soal tersebut memiliki nilai *logit* yang lebih kecil dibanding *logit* abilitas siswa yaitu -2,33 *logit* dan soal lainnya yang seharusnya tidak dapat dijawab benar oleh siswa 33P justru terdapat soal yang dapat dijawab benar oleh siswa tersebut seperti soal nomor 5, 16, 18, 21, dan 22. Hal ini mengindikasikan siswa tidak cermat dalam

menjawab soal nomor 8 dan menebak dalam menjawab soal nomor 5, 16, 18, 21, dan 22. Inilah yang menyebabkan pola respon siswa 33P menjadi tidak *fit* atau tidak ideal.

Tentu saja hasil *output* ini sangat berguna bagi guru untuk mengetahui seberapa besar kemampuan masing-masing siswa berdasarkan pola respon jawaban dan nilai *logit* mereka atau bahkan dapat mengadakan tes ulang pada siswa yang terindikasi curang (mencontek).

Butir Soal Bias (Item DIF)

Butir soal dikatakan mengandung bias apabila nilai probabilitas (PROB) butir soal berada dibawah 0,05 (5%) (Noben *et al.*, 2021). Nilai PROB dapat dilihat pada kolom PROB yang ada pada *Output Table 30. Item DIF* (Martinkova *et al.*, 2017). Hasilnya menunjukkan bahwa butir soal nomor 4 terdeteksi DIF karena memiliki nilai probabilitas kurang dari 5%. Nilai probabilitas butir soal nomor 4 adalah sebesar 0,0497. Sedangkan butir soal yang lain tidak terdeteksi DIF. Hal ini menunjukkan butir soal nomor 4 perlu diperbaiki terlebih dahulu sebelum dilakukan uji lebih lanjut supaya tidak merugikan kelompok gender tertentu. Gambar 7 berikut menampilkan Grafik butir soal DIF.



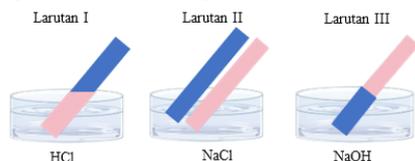
Gambar 7. Grafik Item DIF

Berdasarkan analisis DIF pada Gambar 7 terlihat bahwa, soal nomor 4 lebih mudah dijawab oleh siswa perempuan dan lebih sulit dijawab oleh siswa laki-laki. Oleh karena itu

soal nomor 4 harus direvisi. Butir soal nomor 4 (IKD-1) meminta siswa untuk mengklasifikasikan sifat asam, basa, dan netral dalam suatu larutan. Soal ini menampilkan

ilustrasi larutan I, II, III yaitu larutan HCl, NaCl, dan NaOH dengan masing-masing sampel diberikan data pada tabel. Pertanyaan dalam soal ini perihal penyebab larutan bersifat netral jika dihubungkan antara konsentrasi dan pH. Gambaran lebih jelas terkait soal nomor 4 yang mengandung DIF ini disajikan pada Gambar 8 dibawah.

4. Perhatikan gambar larutan asam, basa dan garam berikut!



Diperoleh hubungan antara $[H^+]$, $[OH^-]$, pH dan pOH pada larutan I, II, dan III sebagai berikut:

Sampel	$[H^+]$	$[OH^-]$	pH	pOH
Larutan I	10^{-1}	10^{-13}	1	13
Larutan II	10^{-7}	10^{-7}	7	7
Larutan III	10^{-13}	10^{-1}	13	1

Berdasarkan data yang diperoleh bahwa larutan bersifat basa apabila dapat menghasilkan ion hidroksida lebih banyak dibandingkan ion hidrogen dalam air, bersifat asam jika terjadi hal sebaliknya dan larutan bersifat netral jika....

- larutan memiliki pH =7 pada suhu ruang
- larutan memiliki $K_w = 10^{-14}$ pada suhu ruang
- larutan tidak memiliki rasa dan tidak berwarna
- larutan menghasilkan ion hidrogen dan ion hidroksida dengan jumlah berbeda
- larutan menghasilkan ion hidrogen dan ion hidroksida dengan jumlah yang sama

Alasan:

- pada keadaan larutan yang bersifat netral, $[H^+]$ pasti 10^{-8}
- pada keadaan larutan yang bersifat netral, $[OH^-]$ pasti 10^{-2}
- pada keadaan larutan yang bersifat netral, $[H^+]$ seimbang dengan $[OH^-]$
- pada keadaan larutan yang bersifat netral, $[H^+]$ lebih besar daripada $[OH^-]$
- pada keadaan larutan yang bersifat netral, $[H^+]$ lebih kecil daripada $[OH^-]$

Gambar 8. Butir Soal Nomor 4

Profil Pemecahan Masalah

Profil mengenai kemampuan pemecahan masalah ini diperoleh dari analisis Rasch Model yaitu nilai *logit* abilitas siswa dikurangi nilai rata-rata *logit* tingkat kesulitan butir soal (Astuti *et al.*, 2020). Grafik mengenai profil pemecahan siswa dapat dilihat pada Gambar 9.



Gambar 9. Profil Kemampuan *Problem-Solving* Siswa

Dari hasil analisis di atas dapat disimpulkan bahwa kemampuan *problem-solving* siswa kategori sangat tinggi sebesar 26% yaitu 10 dari 38 siswa, kategori tinggi sebesar 29% yaitu 11 dari 38 siswa, kategori rata-rata sebesar 16% yaitu 6 dari 38 siswa, kategori rendah sebesar 13% yaitu 5 dari 38 siswa, kategori sangat rendah sebesar 16% yaitu 6 dari 38 siswa. Artinya secara keseluruhan kemampuan pemecahan masalah siswa lebih banyak dalam kategori tinggi yaitu sebanyak 29% dari keseluruhan siswa.

PENUTUP

Nilai *person reliability* didapatkan sebesar 0,78, kriteria cukup, dan nilai *item reliability* sebesar 0,85, kriteria bagus. Nilai *alpha Cronbach* sebesar 0,82, kriteria bagus sekali. Butir soal yang tidak *fit* atau *outliers* berdasarkan hasil analisis *item fit order* adalah butir soal nomor 13 dan 5. Berdasarkan analisis *item measure* terlihat tingkat kesulitan butir soal tertinggi adalah butir soal nomor 13 dengan nilai *logit* +2,81 *logit*. Sedangkan butir dengan tingkat kesulitan yang rendah adalah butir soal nomor 8 dengan nilai *logit* sebesar -2,33 *logit*. Responden 01L memiliki nilai *logit* tertinggi yaitu sebesar +2,00 *logit*. Berdasarkan hasil analisis Rasch model ditemukan ada dua siswa yang pola responnya tidak ideal yaitu siswa 12P dan 37P, yang terindikasi menebak dan tidak cermat. Butir soal nomor 4 terdeteksi DIF sehingga harus diperbaiki sebelum melakukan uji coba selanjutnya. Hasil penelitian menunjukkan profil kemampuan *problem-solving* siswa dengan kategori sangat tinggi sebesar 26% yaitu 10 dari 38 siswa, kategori tinggi sebesar 29% yaitu 11 dari 38 siswa, kategori rata-rata sebesar 16% yaitu 6 dari 38 siswa, kategori rendah sebesar 13% yaitu 5 dari 38 siswa, kategori sangat rendah sebesar 16% yaitu 6 dari 38 siswa. Instrumen tes ini mampu mengukur kemampuan pemecahan masalah siswa dan dapat membantu pelaksanaan tes hasil belajar.

DAFTAR PUSTAKA

- Afnia, P. N., & Istiyono, E. (2020). The Development of Two-Tier Multiple Choice Instruments to Measure Higher Order Thinking Skills Bloomian, 397, 1038–1045.
- Aprilia, N., Susilaningsih, E., Priatmoko, S., & K. (2020). Desain Instrumen Tes Pemahaman Konsep Berbasis HOT dengan Analisis Model Rasch. *Chemistry in Education*, 9(2), 1–8.
- Astuti, A. T., Supahar, Mundilarto, & Istiyono, E. (2020). Development of assessment instruments to measure problem solving skills in senior high school. *Journal of Physics: Conference Series*, 1440(012063), 1–8.
- Bohori, M., & Winny, L. (2019). Analisis penguasaan konsep siswa menggunakan Rasch Model pada materi usaha dan energi. *Prosiding Seminar Nasional Fisika*, 5, 138–143.
- Boone, W.J., Staver, J.R., & Yale, M.S. (2014). *Rasch Analysis in the Human Sciences*. Dordrecht: Springer.
- Hamdu, G., Fuadi, F. N., Yulianto, A., & Akhirani, Y. S. (2020). Items Quality Analysis Using Rasch Model To Measure Elementary School Students' Critical Thinking Skill On Stem Learning. *JPI (Jurnal Pendidikan Indonesia)*, 9(1), 61–74.
- Hilaliyah, H., Agustin, Y., Setiawati, S., Hapsari, S. N., Rangka, I. B., & Ratodi, M. (2019). Wright-Map to investigate the actual abilities on math test of elementary students. *Journal of Physics: Conference Series*, 1318(012067), 1–7.
- Holster, T. A., & Lake, J. (2016). Guessing and the Rasch Model. *Language Assessment Quarterly*, 13(2), 124–141.
- Lestari, W., Pratama, L. D., & Jailani, J. (2019). Metacognitive Skills in Mathematics Problem Solving. *Jurnal Daya Matematis*, 6(3), 286–295.
- Lian, L.H., & Yew, W. . (2020). Development of an Assessment Literacy Super-item Test for Assessing Preservice Teachers' Assessment Literacy. *International Journal of Innovation, Creativity and Change*. *Www.Ijicc.Net*, 13(7), 870–889.
- Mariam, S., Saleh, M., Warsono, & Mujiyanto, J. (2018). Using the Rasch Model for the Affective Assessment of EFL Learners. *Arab World English Journal*, 9(2), 441–456.
- Martinkova, P., Drabinová, A., Liaw, Y. L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking equity: Why differential item functioning analysis should be a routine part of developing conceptual assessments. *CBE Life Sciences Education*, 16(2), 1–13.
- Misbah, M., Mahtari, S., Wati, M., & Harto, M. (2018). Analysis of Students' Critical Thinking Skills in Dynamic Electrical Material. *Kasuari: Physics Education Journal*, 1(2), 103–110.
- Muller, M. (2020). Item fit statistics for Rasch analysis: can we trust them? *Journal of Statistical Distributions and Applications*, 7(5), 1–12.
- Noben, I., Maulana, R., Deinum, J. F., & Hofman, W. H. A. (2021). Measuring university teachers' teaching quality: a Rasch modelling approach. *Learning Environments Research*, 24, 87–107.
- Nurwahidah, I., Widiyawati, Y., Sari, D. S., Masykuri, M., & Budiyanto, C. W. (2020). Development of Science Test To Measure Hots and Digital Literacy of Junior High School Students on the Topic of City Noise. *Edusains*, 12(2), 203–213.
- Pratama, D., & Husnayaini, I. (2020). Applying Rasch Model To Measure Students' Reading Comprehension. *JISAE: Journal of Indonesian Student Assessment and Evaluation*, 6(2), 203–209.
- Putri, F.S., Istiyono, E., & Nurcahyanto, E. (2016). Pengembangan Instrumen Tes Keterampilan Berpikir Kritis dalam Bentuk Pilihan Ganda Beralasan

- (Politomus) di DIY. *Unnes Physics Education Journal*, 5(2), 76–84.
- Rachman, T., & Napitupulu, D. B. (2017). Rasch Model for Validation a User Acceptance Instrument for Evaluating E-learning System. *CommIT (Communication and Information Technology) Journal*, 11(1), 9–16.
- Ringo, S. S., Samsudin, A., & Ramalis, T. R. (2020). Utilizing Rasch Model to Analyze A Gender Gap in Students' Cognitive Ability on Simple Harmonic Motion. *Journal of Physics: Conference Series*, 1467(012054), 1–9.
- Rusmansyah, & Almubarak. (2020). Students' Cognitive Analysis Using Rasch Modeling As an Assessment for Planning of Strategies in Chemistry Learning. *JKPK (Jurnal Kimia Dan Pendidikan Kimia)*, 5(3), 222–235.
- Sari, A. N., Rosidin, U., & Abdurrahman, A. (2020). Developing an Instrument of Performance Assessment to Measure Problem-Solving Skills of Senior High School Students in Physics Inquiry-Based Learning. *Scientiae Educatia: Jurnal Pendidikan Sains*, 9(1), 1–13.
- Shuib, N. H., Amat, S., Kari, D. N. P. M., Ishak, N. M., & Ahmad, I. (2020). Validity and Reliability of MEQI-C Item Using Rasch Measurement Model. *International Journal of Academic Research in Business and Social Sciences*, 10(4), 468–481.
- Sihombing, R.U., Naga, D.S., & Rahayu, W. (2019). A Rasch Model Measurement Analysis on Indonesian Science Literacy Test: Smart Way to Improve the Learning Assessment. *Indonesian Journal of Educational Review*, 6(2), 42–53.
- Singaravelu, S. (2017). Problem Solving Ability of Higher Secondary Chemistry Students. *IOSR Journal of Research & Method in Education (IOSR-JRME)*, 7(4), 19–22.
- Stehle, S.M., & Peters-Burton, E. . (2019). Developing student 21 st Century skills in selected exemplary inclusive STEM high schools. *International Journal of STEM Education*, 6(39), 1–15.
- Sudihartinih, E., & Prabawanto, S. (2020). Test Instrument Validation in Plane Geometry using Rasch Model. *Mathematics Education Journals*, 4(2), 102–115.
- Sumintono, B., & Widhiarso, W. (2013). *Aplikasi Model Rasch untuk Penelitian Ilmu-Ilmu Sosial*. Cimahi: Trim Komunikata.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Permodelan Rasch Pada Assessment Pendidikan*. Cimahi: Trim Komunikata.
- Suryani, Y. E. (2018). Aplikasi Rasch Model dalam Mengevaluasi Intelligenz Structure Test (IST). *Jurnal Penelitian Psikologi*, 3(1), 73–100.
- Susilaningsih, E., Nuswowati, M., Aprilia, N., & Luthfiyah, A. (2021). Dissemination of test instruments as product of the development research to measure the problem-solving ability of class X students by online in the pandemic period. *Journal of Physics: Conference Series*, 1918(032024), 1–8.
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159–169.
- Treagust, D.F. (1995) Diagnostic assessment of students' science concepts. In S. Glynn and R. Duit (Eds) *Learning science in the schools: Research reforming practice*. New Jersey: Lawrence Erlbaum Associates, 327–346.
- Widiasih, Permanasari, A., Riandi, & Damayanti, T. (2018). The profile of problem-solving ability of students of distance education in science learning. *Journal of Physics: Conference Series*,

1013(012081), 1–6.

Wigati, I., & Aini, K. (2020). Biology Education Students' Analysis of Reasoning on Plant Taxonomy Coursewhite Rasch Model. *Edusains*, 12(2), 145–153.

Winarti, A., & Mubarak, A. (2020). Rasch Modeling: A Multiple Choice Chemistry Test. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 2(1), 1–9.

Yasin, S. N. T. M., Yunus, M. F. M., & Ismail, I. (2018). The use of rasch measurement model for the validity and reliability. *Journal of Counseling and Educational Technology*, 1(2), 22–27.

Yazdani, K., Nedjat, S., Karimlou, M., Zeraati, H., Mohammad, K., & Fotouhi, A. (2015). Developing a shortened quality of life scale from Persian version of the WHOQOL-100 using the rasch analysis. *Iranian Journal of Public Health*, 44(4), 522–534.