



Model Rasch: Pengembangan Instrumen Penelitian Pendidikan

Muntazhimah



Model Rasch di dalam sebuah penelitian dapat digunakan sebagai alat untuk menganalisis instrumen. Model ini mampu menganalisis sampai ke tingkat butir soal dan individu. Dibandingkan teori tes klasik, model ini mampu melakukan prediksi terhadap data yang hilang karena didasarkan pada pola respons yang sistematis. Melalui buku ini, penulis berusaha menjabarkan langkah-langkah praktis menggunakan model Rasch untuk menguji sebuah instrumen penelitian. Penulis juga menjabarkan beberapa model Rasch yang berhasil diterapkan melalui laporan-laporan ilmiah dari berbagai artikel.

Model Rasch: Pengembangan Instrumen Penelitian Pendidikan

Penerbit Deepublish (CV BUDI UTAMA)
Jl. Kaliurang Km 9,3 Yogyakarta 55581
Telp/Fax : (0274) 4533427
Anggota IKAPI (078/DIY/2012)

● ce@deepublish.co.id
● Penerbit Deepublish
● [@penerbitbuku_deepublish](https://www.instagram.com/penerbitbuku_deepublish)
● www.penerbitdeepublish.com



Model Rasch: Pengembangan Instrumen Penelitian Pendidikan

UU No 28 tahun 2014 tentang Hak Cipta

Fungsi dan sifat hak cipta Pasal 4

Hak Cipta sebagaimana dimaksud dalam Pasal 3 huruf a merupakan hak eksklusif yang terdiri atas hak moral dan hak ekonomi.

Pembatasan Pelindungan Pasal 26

Ketentuan sebagaimana dimaksud dalam Pasal 23, Pasal 24, dan Pasal 25 tidak berlaku terhadap:

- i. Penggunaan kutipan singkat Ciptaan dan/atau produk Hak Terkait untuk pelaporan peristiwa aktual yang ditujukan hanya untuk keperluan penyediaan informasi aktual;
- ii. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk kepentingan penelitian ilmu pengetahuan;
- iii. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk keperluan pengajaran, kecuali pertunjukan dan Fonogram yang telah dilakukan Pengumuman sebagai bahan ajar, dan
- iv. Penggunaan untuk kepentingan pendidikan dan pengembangan ilmu pengetahuan yang memungkinkan suatu Ciptaan dan/atau produk Hak Terkait dapat digunakan tanpa izin Pelaku Pertunjukan, Produser Fonogram, atau Lembaga Penyiaran.

Sanksi Pelanggaran Pasal 113

1. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000 (seratus juta rupiah).
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

Model Rasch: Pengembangan Instrumen Penelitian Pendidikan

Muntazhimah



**MODEL RASCH:
PENGEMBANGAN INSTRUMEN PENELITIAN PENDIDIKAN**

Muntazhimah

Editor :
Ahmad Khanafi

Desain Cover :
Rulie Gunadi

Sumber :
www.shutterstock.com

Tata Letak :
Werdiantoro

Proofreader :
Mira Muarifah

Ukuran :
viii, 61 hlm, Uk: 15.5x23 cm

ISBN :
978-623-02-5823-7

Cetakan Pertama :
Januari 2023

Hak Cipta 2023, Pada Penulis

Isi diluar tanggung jawab percetakan

Copyright © 2023 by Deepublish Publisher
All Right Reserved

Hak cipta dilindungi undang-undang.
Dilarang keras menerjemahkan, memfotokopi, atau
memperbanyak sebagian atau seluruh isi buku ini
tanpa izin tertulis dari Penerbit.

PENERBIT DEEPUBLISH
(Grup Penerbitan CV BUDI UTAMA)

Anggota IKAPI (076/DIY/2012)
Jl. Rajawali, G. Elang 6, No 3, Drono, Sardonoharjo, Ngaglik, Sleman
Jl. Kaliurang Km.9,3 – Yogyakarta 55581
Telp/Faks: (0274) 4533427
Website: www.deepublish.co.id
www.penerbitdeepublish.com
E-mail: cs@deepublish.co.id

Kata Pengantar Penerbit

Assalamualaikum, w.r. w.b.

Segala puji kami haturkan ke hadirat Allah Swt., Tuhan Yang Maha Esa, atas limpahan rahmat dan karunia-Nya. Tak lupa, lantunan selawat serta salam senantiasa tercurahkan kepada Nabi Besar Muhammad saw.

Dalam rangka mencerdaskan dan memuliakan umat manusia dengan penyediaan serta pemanfaatan ilmu pengetahuan dan teknologi untuk menciptakan industri *processing* berbasis sumber daya alam (SDA) Indonesia, Penerbit Deepublish dengan bangga menerbitkan buku dengan judul ***Model Rasch: Pengembangan Instrumen Penelitian Pendidikan***.

Terima kasih dan penghargaan terbesar kami sampaikan kepada penulis yang telah memberikan kepercayaan, perhatian, dan kontribusi penuh demi kesempurnaan buku ini. Semoga buku ini bermanfaat bagi semua pembaca, mampu berkontribusi dalam mencerdaskan dan memuliakan umat manusia, serta mengoptimalkan pemanfaatan ilmu pengetahuan dan teknologi di tanah air.

Wassalamualaikum, w.r. w.b.

Hormat Kami,
Penerbit Deepublish

Daftar Isi

| | |
|--|------|
| Kata Pengantar Penerbit | v |
| Daftar Isi | vi |
| Daftar Tabel | vii |
| Daftar Gambar | viii |
| Bab 1 Analisis Instrumen berbasis Model Rasch..... | 1 |
| Bab 2 Validasi Instrumen Tes Berpikir Reflektif..... | 17 |
| Bab 3 Validasi Instrumen Angket Resiliensi Matematis | 29 |
| Bab 4 Validasi Instrumen Berpikir Reflektif Matematis Siswa Madrasah Aliyah | 41 |
| Daftar Pustaka..... | 53 |
| Profil Penulis..... | 61 |

Daftar Tabel

| | |
|---|----|
| Tabel 1. Definisi, Indikator dan Contoh Butir Soal | 20 |
| Tabel 2. Pemaknaan Reliabilitas Berdasarkan Nilai <i>Alpha Cronbach</i> | 22 |
| Tabel 3. Ringkasan Uji Reliabilitas..... | 23 |
| Tabel 4. Ringkasan Uji Validitas..... | 24 |
| Tabel 5. Kriteria <i>Item Fit</i> | 25 |
| Tabel 6. Pemaknaan Uji <i>Item Fit Order</i> | 26 |
| Tabel 7. Kisi-Kisi Instrumen Resiliensi Matematis Calon Guru Matematika..... | 32 |
| Tabel 8. Interpretasi Uji Reliabilitas Berdasarkan Nilai <i>Cronbach Alpha</i> | 36 |
| Tabel 9. Hasil Pengolahan Uji Reliabilitas | 37 |
| Tabel 10. Hasil Pengolahan Validitas Instrumen..... | 38 |
| Tabel 11. Nilai Kriteria Kesesuaian Butir Soal | 39 |
| Tabel 12. Hasil Pengolahan <i>Item Fit Order</i> | 40 |
| Tabel 13. Kriteria Persentase Penilaian Ahli | 43 |
| Tabel 14. Rekapitulasi Penilaian Para Ahli | 46 |
| Tabel 15. <i>Item Reliability & Person Reliability</i> | 50 |
| Tabel 16. Kriteria <i>Alpha Cronbach</i> | 50 |

Daftar Gambar

| | | |
|-------------------|---|----|
| Gambar 1. | <i>Output Summary Statistics Winstep 4.4.3</i> | 22 |
| Gambar 2. | <i>Output Item Undimensionality Winstep 4.4.3</i> | 23 |
| Gambar 3. | <i>Output Item Fit Order Winstep 4.4.3</i> | 25 |
| Gambar 4. | <i>Output Summary Statistics pada Winstep</i> | 33 |
| Gambar 5. | <i>Output Item Undimensionality pada Winstep</i> | 34 |
| Gambar 6. | <i>Output Item Fit Order pada Winstep</i> | 35 |
| Gambar 7. | <i>Output Tabel Item Dimensionality</i> | 47 |
| Gambar 8. | <i>Output Item Fit Order</i> | 48 |
| Gambar 9. | <i>Output Tabel Item Summary Statistics</i> | 49 |
| Gambar 10. | <i>Output Tabel Item Measure</i> | 51 |

Bab 2

Validasi Instrumen Tes Berpikir Reflektif

Artikel ini dimuat di dalam Jurnal Elemen Vol 8 No.1 tahun 2022, dengan judul '*The development and validation of mathematical reflective thinking test for prospective mathematics teachers using the Rasch Model*'. Artikel ini merupakan hasil dari eksperimen tentang pengembangan instrumen tes berpikir reflektif matematis bagi mahasiswa calon guru Matematika menggunakan model Rasch masih terbatas (Muntazhimah, 2022). Langkah pengembangan yang dilakukan antara lain:

1. penyusunan *blueprint* (cetak biru);
2. penulisan butir soal;
3. *me-review* soal;
4. melakukan uji coba;
5. menganalisis hasil uji coba; dan
6. melakukan revisi.

Tes yang dikembangkan berupa uraian yang sudah divalidasi oleh lima orang *expert* meliputi dua orang Profesor ahli di bidang evaluasi kemampuan matematis, satu orang dosen ahli dalam bidang kemampuan matematis, serta dua orang dosen ahli dalam materi teori grup. Tes diujicobakan pada 26 mahasiswa dan berhasil melaporkan bahwa instrumen ini memiliki kriteria "sangat reliabel" dan 12 dari 13 *item* butir tes dikatakan "valid". Karena sudah memenuhi semua tahapan pengembangan instrumen dan dinyatakan valid serta reliabel, maka penelitian kemampuan berpikir reflektif matematis mahasiswa calon guru Matematika dapat diukur dengan menggunakan instrumen tes ini.

Riset yang dilaporkan melalui artikel ini dilatarbelakangi oleh

adanya regulasi dari Negara Republik Indonesia melalui Undang-Undang Republik Indonesia Nomor 14, Tahun 2005 tentang Guru dan Dosen mengatur bahwa kompetensi pedagogik, kompetensi kepribadian, kompetensi sosial, dan kompetensi profesional adalah kompetensi yang wajib dimiliki oleh seorang guru maupun dosen. Lebih lanjut bahwa Peraturan Menteri Pendidikan Nasional Nomor 16, Tahun 2007 tentang Standar Kualifikasi Akademik dan Kompetensi Guru menyebutkan bahwa salah satu kompetensi inti dalam kompetensi pedagogik, yakni menyelenggarakan penilaian dan evaluasi proses dan hasil belajar. Salah satu penjabaran implementasi penguasaan kompetensi ini ialah kemampuan untuk mengembangkan instrumen penilaian dan evaluasi proses dan hasil belajar. Salah satu alat untuk mengukur ketercapaian tujuan pembelajaran adalah instrumen dalam bentuk tes (Prabowo & Dahlan, 2020; Suharman, 2018; Wahyudi, 2012). Betapa pentingnya posisi suatu tes di dalam dunia pendidikan seperti diamanatkan dalam undang-undang di atas. Hal inilah yang menjadi dasar bahwa pendidik harus mau dan mampu mengembangkan instrumen tes.

Sayangnya, intensitas guru dalam mengembangkan instrumen tes masih minim (Osnal, *et al.*, 2015), yang mengakibatkan masih banyak instrumen tes belum mengikuti kaidah sebagai tes yang berkualitas (Prabowo *et al.*, 2018; Wardhani & Putra, 2016). Untuk mendapatkan instrumen tes yang baik, perlu dilakukan analisis kualitas instrumen (Arifin, 2016). Analisis instrumen inilah yang nantinya harus melewati kriteria kualitas secara substansi, konstruksi dan bahasa, serta memenuhi unsur validitas dan reliabilitas. Tahapan-tahapan ini juga berlaku pada pengembangan instrumen penelitian dalam bidang pendidikan Matematika.

Kemampuan matematis menjadi isu penting di dalam pendidikan profesionalitas guru sebagai kemampuan berpikir reflektif (Amidu, 2012); (Yuen Lie Lim, 2011). Kemampuan berpikir reflektif matematis adalah salah satu *Higher Order Thinking Skill* (HOTS) atau kemampuan berpikir tingkat tinggi. HOTS merupakan proses berpikir di mana terjadi aktivitas merefleksikan ide, masalah atau informasi yang diterima atau proses membuat pemaknaan dari satu

pengalaman ke depan dengan membuat pemahaman yang lebih dalam dengan menghubungkan pengalaman atau ide-ide yang lain. (Muin *et al.*, 2018; Clarà, 2015). Kemampuan berpikir reflektif matematis menjadi salah satu kemampuan matematis yang harus dimiliki oleh seluruh peserta didik, Melalui kemampuan ini, peserta didik secara aktif, sungguh-sungguh dan penuh kehati-hatian untuk mempertimbangkan sesuatu dengan pengetahuan yang telah diperolehnya terhadap persoalan Matematika yang diberikan.

Berbagai penelitian tentang kemampuan berpikir reflektif matematis telah banyak dilakukan, antara lain tentang kemampuan berpikir reflektif mahasiswa calon guru di Aceh berdasarkan gender dan pengetahuan awalnya (Rahmadhani *et al.*, 2020), adaptasi *Reflective Thinking Questionnaire* (RTQ) ke dalam versi Indonesia dan menginvestigasi kualitasnya pada mahasiswa calon guru Kimia di Tanjungpinang (Sabekti *et al.*, 2020), serta proses berpikir reflektif seorang mahasiswa perempuan dengan gaya kognitif tipe independen dalam memecahkan masalah aljabar (Agustan *et al.*, 2017). Pengembangan riset terkait bahan ajar dan instrumen kemampuan berpikir reflektif matematis sayangnya masih belum banyak ditemukan. Sejauh ini hanya sebatas pada pengembangan bahan ajar untuk mengembangkan kemampuan berpikir reflektif matematis siswa SMA (Hendriana *et al.*, 2019; Nindiasari *et al.*, 2016), serta penelitian dari Muntazhimah, (2019) mengenai penelitiannya melakukan pengembangan instrumen tes berpikir reflektif matematis untuk siswa kelas 8 SMP.

Riset di dalam artikel ini dilakukan karena dilatarbelakangi oleh sulitnya mencari referensi tentang pengembangan tes berpikir reflektif matematis mahasiswa calon guru Matematika, dan keseluruhan penelitian pengembangan instrumen tes yang sudah dilaksanakan masih memakai teori tes klasik. Di sisi lain, Pemodelan Rasch memiliki beberapa kelebihan dibandingkan dengan tes teori klasik, antara lain mampu melakukan prediksi terhadap data yang hilang (*missing data*), yang didasarkan pada pola respons yang sistematis (Nur *et al.*, 2020; An & Yu, 2021). Hal ini jelas menjadikan hasil analisis statistik yang lebih akurat. Selain itu, pemodelan Rasch juga

mampu menghasilkan nilai pengukuran standar error untuk instrumen yang digunakan yang dapat meningkatkan ketepatan perhitungan (Sumintono & Widhiarso, 2015).

Metodologi yang diadopsi di dalam artikel ini mengacu pada Spaan (2006), dan Prabowo & Dahlan (2020) yang meliputi: penyusunan *blueprint* (cetak biru), penyusunan butir soal, *me-review* soal, uji coba, analisis hasil uji coba, dan merevisi sehingga menghasilkan instrumen tes berpikir reflektif matematis mahasiswa calon guru Matematika yang efektif, efisien dan berkualitas. Rowe (2001) menyatakan bahwa *blueprint* bisa dijadikan petunjuk efektif untuk menulis butir soal. *Blueprint* disebut juga sebagai kisi-kisi tes. Pada kisi-kisi disebutkan definisi kemampuan berpikir reflektif matematis, pemilihan materi atau mata kuliah yang relevan, serta penyesuaian sub capaian mata kuliah dengan indikator berpikir reflektif matematis. Selanjutnya adalah menulis butir soal dengan menurunkan indikator kemampuan berpikir reflektif menjadi beberapa butir soal. Tabel 1 menjelaskan tentang definisi dan indikator, serta contoh butir soal yang sudah disusun.

Tabel 1. Definisi, Indikator dan Contoh Butir Soal

| Definisi | Indikator | Contoh Butir Soal Tes |
|--|---|---|
| Berpikir dengan penuh pertimbangan dengan menerapkan pengetahuan matematis dan pengalaman matematis yang sudah diperoleh sehingga bisa menganalisis, mengevaluasi dan mendapatkan makna yang mendalam dalam menyelesaikan permasalahan Matematika. | <ol style="list-style-type: none"> 1. Menganalisis kebenaran pertanyaan/solusi/ analogi atau generalisasi Matematika, 2. Mengidentifikasi konsep atau formula Matematika yang digunakan dalam soal Matematika tidak sederhana 3. Membedakan antara data relevan dan tidak relevan, 4. Mengevaluasi keabsahan argumen berdasarkan konsep atau sifat yang digunakan, 5. Menemukan beragam strategi dalam penyelesaian masalah Matematika | <p>Contoh butir soal ini mewakili indikator kelima. Butir soalnya berbunyi: Misalkan $H = \{ \}$. Dan adalah sebuah grup.</p> <p>Pertanyaan</p> <ol style="list-style-type: none"> a. Temukan ada berapa strategi untuk membuktikan bahwa H merupakan subgrup dari b. Berikan penjelasan konsep, prinsip, sifat-sifat atau aturan yang dipakai oleh strategi-strategi tersebut c. Pilih salah satu strategi dan buktikan bahwa H adalah subgrup dari |

Setelah instrumen tes terdiri dari soal yang mewakili setiap indikator, selanjutnya adalah *me-review* soal dengan memintakan *judgement* dari 5 orang *expert* untuk memvalidasi konten materi. *Expert* terdiri dari dua orang dosen yang mengampu mata kuliah yang sama dan berasal dari perguruan tinggi berbeda, dua orang profesor yang ahli di bidang evaluasi kemampuan matematis, serta satu orang dosen yang ahli dalam bidang kemampuan matematis. Setelah dilakukan revisi, instrumen tes diminta untuk dilakukan *review* juga secara kepada lima orang mahasiswa calon guru Matematika untuk di cek keterbacaannya. Selanjutnya adalah uji coba instrumen kepada 26 orang mahasiswa kemudian hasilnya dianalisis menggunakan model Rasch dengan *software* Winstep Rasch 4.4.3. Hasil pengujian kemudian dijadikan dasar untuk revisi. Butir soal yang sudah direvisi selanjutnya disusun menjadi sebuah instrumen tes berpikir reflektif matematis mahasiswa calon guru Matematika.

Software Winstep Rasch 4.4.3 digunakan untuk melihat reliabilitas dan validitas instrumen tes berpikir reflektif matematis mahasiswa calon guru Matematika. Validitas dan reliabilitas merupakan aspek penting yang harus diperhatikan dalam pengembangan sebuah instrumen baru dalam penelitian (Alavi *et al.*, 2020). Validitas menilai sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Sedangkan reliabilitas menilai sejauh mana hasil suatu pengukuran dapat dipercaya. Reliabilitas dimaksudkan untuk menganalisis apakah instrumen kemampuan berpikir reflektif matematis ini dapat dipakai secara general. Analisis reliabilitas berpatokan pada tabel 3.1 (*Summary statistic*) di program *Winsteps*, seperti pada Gambar 1 berikut ini.

SUMMARY OF 26 MEASURED Person

| | TOTAL SCORE | COUNT | MEASURE | MODEL S.E. | INFIT | | OUTFIT | |
|---------------------------|-------------|---------|---------|------------|-------|--------|-------------|-------|
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD |
| MEAN | 27.2 | 13.0 | -.67 | .30 | 1.02 | -.04 | 1.15 | .21 |
| SEM | 1.5 | .0 | .13 | .00 | .16 | .24 | .21 | .20 |
| P.SD | 7.6 | .2 | .65 | .01 | .78 | 1.21 | 1.04 | 1.02 |
| S.SD | 7.8 | .2 | .66 | .01 | .79 | 1.23 | 1.06 | 1.04 |
| MAX. | 53.0 | 13.0 | 1.40 | .33 | 3.54 | 3.34 | 4.87 | 3.55 |
| MIN. | 16.0 | 12.0 | -1.69 | .29 | .23 | -1.86 | .34 | -1.05 |
| REAL RMSE | .34 | TRUE SD | .55 | SEPARATION | 1.64 | Person | RELIABILITY | .73 |
| MODEL RMSE | .30 | TRUE SD | .57 | SEPARATION | 1.91 | Person | RELIABILITY | .79 |
| S.E. OF Person MEAN = .13 | | | | | | | | |

Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
 CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .71 SEM = 4.11

SUMMARY OF 13 MEASURED Item

| | TOTAL SCORE | COUNT | MEASURE | MODEL S.E. | INFIT | | OUTFIT | |
|-------------------------|-------------|---------|---------|------------|-------|-------|-------------|-------|
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD |
| MEAN | 54.5 | 25.9 | .00 | .25 | 1.01 | .04 | 1.15 | .07 |
| SEM | 9.5 | .1 | .23 | .02 | .09 | .37 | .18 | .37 |
| P.SD | 32.9 | .3 | .80 | .07 | .33 | 1.27 | .64 | 1.28 |
| S.SD | 34.3 | .3 | .83 | .07 | .34 | 1.33 | .66 | 1.33 |
| MAX. | 143.0 | 26.0 | 1.14 | .34 | 1.84 | 2.73 | 3.08 | 2.47 |
| MIN. | 9.0 | 25.0 | -1.52 | .11 | .47 | -2.78 | .47 | -2.74 |
| REAL RMSE | .27 | TRUE SD | .75 | SEPARATION | 2.77 | Item | RELIABILITY | .88 |
| MODEL RMSE | .26 | TRUE SD | .76 | SEPARATION | 2.93 | Item | RELIABILITY | .90 |
| S.E. OF Item MEAN = .23 | | | | | | | | |

Gambar 1. Output Summary Statistics Winstep 4.4.3

Gambar di atas menunjukkan, bahwa nilai *Alpha Cronbach* sebesar 0.71. Dengan merujuk pada Tabel 2 berikut (Sumintono & Widhiarso, 2015) instrumen ini masuk dalam kriteria “Bagus”.

Tabel 2. Pemaknaan Reliabilitas Berdasarkan Nilai *Alpha Cronbach*

| Nilai | Interpretasi |
|-------|--------------|
| | Bagus sekali |
| | Bagus |
| | Cukup |
| | Jelek |
| | Buruk |

Selanjutnya adalah skor *Person Reliability* yang tertera pada output gambar 1 sebesar 0,73. Artinya konsistensi jawaban responden berada dalam kategori “baik”. *Item Reliability* yang menunjukkan skor 0.88 menunjukkan, bahwa mutu instrumen masuk dalam kriteria “Istimewa”. Berdasarkan pemaparan tersebut, instrumen tes

kemampuan ini dapat disimpulkan sebagai instrumen dengan tingkat reliabilitas yang tinggi.

Tabel 3. Ringkasan Uji Reliabilitas

| <i>Cronbach Alpha</i> | Interpretasi | <i>Item Reliability</i> | Interpretasi | <i>Person Reliability</i> | Interpretasi | Kesimpulan |
|-----------------------|--------------|-------------------------|--------------|---------------------------|--------------|------------|
| 0,71 | Bagus | 0,88 | Bagus sekali | 0,73 | bagus | Reliabel |

Validitas instrumen digunakan untuk menilai, apakah instrumen tes kemampuan berpikir reflektif matematis mahasiswa calon guru Matematika ini dapat dipakai sebagai alat mengukur kemampuan berpikir reflektif matematis mahasiswa calon guru Matematika. Output *software Winstepnya* bisa menggunakan tabel 23 (*Item undimensionality*) dan tabel 10 (*Item Fit Order*). Outputnya akan menyajikan *item* mana yang tepat untuk mengukur apa yang seharusnya diukur. Analisis validitas pada pendekatan Rasch dengan *software Winsteps* disebut dengan uji *fit* dan *misfit* (*item* yang valid dan tidak valid). Output yang dimaksud dapat dilihat pada gambar 2 sebagai berikut.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

| | Eigenvalue | Observed | Expected |
|--------------------------------------|------------|----------|----------|
| Total raw variance in observations = | 42.5612 | 100.0% | 100.0% |
| Raw variance explained by measures = | 29.5612 | 69.5% | 70.5% |
| Raw variance explained by persons = | 6.9255 | 16.3% | 16.5% |
| Raw Variance explained by items = | 22.6358 | 53.2% | 54.0% |
| Raw unexplained variance (total) = | 13.0000 | 30.5% | 29.5% |
| Unexplned variance in 1st contrast = | 2.7831 | 6.5% | 21.4% |
| Unexplned variance in 2nd contrast = | 2.3464 | 5.5% | 18.0% |
| Unexplned variance in 3rd contrast = | 2.1617 | 5.1% | 16.6% |
| Unexplned variance in 4th contrast = | 1.2525 | 2.9% | 9.6% |
| Unexplned variance in 5th contrast = | 1.0658 | 2.5% | 8.2% |

Gambar 2. Output *Item Undimensionality Winstep 4.4.3*

Kriteria *item undimensionality* pada model Rasch berdasarkan skor *raw variance explained by measures* menunjukkan nilai sebesar 69.5%. Artinya, jika skor *raw variance explained by measures* > 20% dikatakan terpenuhi, jika skornya > 40% bagus dan jika skor > 60%

dikatakan memiliki kriteria istimewa. Sehingga untuk instrumen ini kriterianya adalah istimewa.

Selanjutnya untuk melihat ada tidaknya butir soal yang masih bermasalah atau tidak sesuai, dapat dilakukan pengecekan kembali skor *eigenvalue* dan *observed* dalam *unexplained variance contrast*. Butir soal dikatakan tidak ada yang bermasalah jika skor *eigenvalue* kurang dari 3 serta untuk butir soal yang sesuai, skor *observed* harus kurang dari 15%.

Tabel 4. Ringkasan Uji Validitas

| <i>Raw variance explained by measures</i> | Interpretasi | <i>unexplained variance contrast</i> ^{1st} | | Interpretasi |
|---|--------------|---|-----------------|---------------------------------------|
| | | <i>Eigenvalue</i> | <i>Observed</i> | |
| 69.5% | istimewa | 2.7831 | 6.5% | Tidak ada <i>item</i> yang bermasalah |

Tabel 4 menunjukkan, bahwa skor *raw variance explained by measures* adalah 69.5 %. Artinya, keseluruhan butir tes berpikir reflektif matematis ber kriteria “Istimewa”. Lebih dalam, skor *observed* pada *unexplained variance contrast* adalah 6.5%. Atau dalam hal ini keseluruhan butir soal telah sesuai dan dapat dipakai, serta skor *eigenvalue* 2.7831 yang bermakna bahwa semua butir soal sudah baik dan tidak ada yang bermasalah. Namun masih dapat dilakukan analisis lanjutan yaitu *item fit order* karena nilai *eigenvalue* yang hampir mendekati angka 3.

Pada tahap analisis *item fit order*, penting untuk melihat skor pada *Outfit Mean Square (MNSQ)*, *Outfit Z-Standard (ZSTD)* serta skor *Point Measure Coorelation (Pt Mean Corr)*. Secara lengkap skor tersebut dapat dilihat pada gambar 3 di bawah ini. Tampak bahwa untuk *entry number* 1 artinya untuk butir nomor 1, skor *outfit MNSQ* adalah 3.08, sedangkan *ZSTD*-nya 1,44 serta *PTMEASURE Corr*-nya -0,50. Pada baris selanjutnya, butir no 13 skornya secara berurutan adalah 1,73 dan 2,47 serta 0,68 dan seterusnya hingga baris ke terakhir.

Item STATISTICS: MISFIT ORDER

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MEASURE | MODEL S. E. | | INFIT | | OUTFIT | | PTMEASUR-AL | | EXACT MATCH | | Item | G |
|--------------|-------------|-------------|---------|-------------|------|-------|------|--------|------|-------------|------|-------------|------|------|---|
| | | | | MNSQ | ZSTD | MNSQ | ZSTD | CORR. | EXP. | OBS% | EXP% | | | | |
| 1 | 9 | 26 | .64 | .29 | .95 | .21 | 3.08 | 1.44 | A | .50 | .61 | 84.6 | 86.5 | E1 | A |
| 13 | 40 | 26 | .76 | .22 | 1.84 | 2.73 | 1.73 | 2.47 | B | .68 | .49 | 26.9 | 41.7 | E13 | E |
| 9 | 43 | 26 | -.27 | .31 | 1.25 | .82 | 1.32 | .94 | C | .30 | .41 | 57.7 | 57.6 | E9 | C |
| 10 | 31 | 26 | .80 | .29 | 1.25 | 1.02 | 1.24 | .98 | D | .56 | .39 | 30.8 | 50.8 | E10 | C |
| 2 | 143 | 26 | -1.06 | .11 | 1.08 | .39 | .90 | -.06 | E | .66 | .68 | 19.2 | 19.7 | E2 | B |
| 4 | 94 | 26 | -1.52 | .18 | 1.08 | .37 | 1.04 | .23 | F | .57 | .50 | 23.1 | 32.8 | E4 | D |
| 7 | 64 | 26 | -.45 | .23 | 1.08 | .34 | 1.07 | .31 | G | .14 | .57 | 50.0 | 47.6 | E7 | E |
| 11 | 42 | 26 | .70 | .34 | .88 | .25 | .94 | .16 | f | .64 | .66 | 65.4 | 72.0 | E11 | F |
| 5 | 76 | 26 | -1.00 | .16 | .91 | -.30 | .89 | -.33 | e | .56 | .56 | 38.5 | 31.1 | E5 | D |
| 3 | 39 | 26 | .10 | .30 | .90 | -.27 | .89 | -.28 | d | .63 | .39 | 73.1 | 55.4 | E3 | C |
| 12 | 52 | 25 | -.20 | .31 | .84 | -.41 | .82 | -.45 | c | .54 | .31 | 68.0 | 63.2 | E12 | G |
| 6 | 48 | 26 | .36 | .22 | .57 | -1.84 | .56 | -1.78 | b | .47 | .51 | 57.7 | 43.0 | E6 | E |
| 8 | 27 | 26 | 1.14 | .29 | .47 | -2.78 | .47 | -2.74 | a | .26 | .39 | 84.6 | 50.4 | E8 | C |
| MEAN | 54.5 | 25.9 | .00 | .25 | 1.01 | .0 | 1.15 | .1 | | | | 52.3 | 50.1 | | |
| P. SD | 32.9 | .3 | .80 | .07 | .33 | 1.3 | .64 | 1.3 | | | | 21.9 | 17.0 | | |

Gambar 3. Output Item Fit Order Winstep 4.4.3

Analisis terakhir yaitu uji kesesuaian untuk menilai kenormalan fungsi butir soal dalam pengukuran atau tidak (*Item fit*). Skor yang akan diperhatikan adalah *outfit means-square*, *Outfit Z-Standard*, dan *Point Measure Correlation* (Boone et al., 2014). Kriteria penilaian *item fit* ditampilkan oleh Tabel 5 (Sumintono, Bambang & Widhiarso, 2015) sebagai berikut.

Tabel 5. Kriteria Item Fit

| Kriteria | Nilai |
|----------------------------------|--------------------------|
| <i>Outfit Mean Square (MNSQ)</i> | 0,5 MNSQ 1,5 |
| <i>Outfit Z-standart (ZSTD)</i> | -2,0 ZSTD +2,0 |
| <i>Point Measure Correlation</i> | 0,4 PT Measure Corr 0,85 |

Apabila ketiga kriteria pada tabel 5 di atas terpenuhi, maka butir soal sudah “sesuai” dan kualitas butir soal tersebut dikatakan bagus serta bisa dipakai. Namun apabila terdapat hanya dua kriteria atau bahkan satu kriteria yang terpenuhi, butir soal masih dapat dipertahankan dan tidak perlu diubah. Apabila tidak memenuhi satu pun kriteria yang telah ditetapkan pada tabel 5 di atas, dipastikan butir soal “tidak sesuai” dan

harus disusun ulang. Pemaknaan hasil uji *Item Fit* tertera pada Tabel 6 sebagai berikut.

Tabel 6. Pemaknaan Uji *Item Fit* Order

| Nomor Soal | Kode Soal | Outfit | | PT Measure Corr. | Status | Interpretasi |
|------------|-----------|--------|-------|------------------|------------|--------------|
| | | MNSQ | ZFTD | | | |
| 1 | P1 | 3,08 | 1,44 | 0,50 | 1 Kriteria | Sesuai |
| 2 | P2 | 0,90 | -0,06 | 0,66 | - | Sesuai |
| 3 | P3 | 0,89 | -0,28 | 0,63 | - | Sesuai |
| 4 | P4 | 1,04 | 0,23 | 0,57 | - | Sesuai |
| 5 | P5 | 0,89 | -0,33 | 0,56 | - | Sesuai |
| 6 | P6 | 0,56 | -1,78 | 0,47 | - | Sesuai |
| 7 | P7 | 1,07 | 0,31 | 0,14 | 1 Kriteria | Sesuai |
| 8 | P8 | 0,47 | -2,74 | 0,26 | 3 Kriteria | Tidak Sesuai |
| 9 | P9 | 1,32 | 0,94 | 0,30 | 1 Kriteria | Sesuai |
| 10 | P10 | 1,24 | 0,98 | 0,56 | - | Sesuai |
| 11 | P11 | 0,98 | 0,16 | 0,64 | - | Sesuai |
| 12 | P12 | 0,82 | -0,45 | 0,54 | - | Sesuai |
| 13 | P13 | 1,73 | 2,47 | 0,68 | 2 Kriteria | Sesuai |

Tabel 6 di atas menunjukkan, hanya ada satu butir soal yang tidak memenuhi satu pun kriteria, yakni butir nomor 8. Butir nomor 8 adalah butir yang tidak valid (*Misfit*) dan harus dikeluarkan dari instrumen tes. Selain itu, butir soal sudah memenuhi minimal satu kriteria sehingga bisa disebutkan bahwa keseluruhan butir lainnya adalah *item fit* atau valid. Secara keseluruhan, instrumen tes berpikir reflektif matematis mahasiswa calon guru Matematika pada penelitian ini dinyatakan reliabel dan valid dengan kriteria “sangat reliabel” dan 12 dari 13 butir tes dikatakan “valid”. Tahapan akhir setelah selesai menganalisis hasil uji coba adalah melakukan revisi agar dihasilkan produk akhir berupa instrumen tes berpikir reflektif matematis yang siap untuk disosialisasikan dan digunakan sebagai instrumen penelitian.

Pada akhirnya kita mengetahui, bahwa tes berpikir reflektif matematis mahasiswa calon guru Matematika telah melalui

keseluruhan langkah pengembangan, serta telah dinyatakan valid dan reliabel menggunakan model Rasch, sehingga diharapkan menghasilkan instrumen tes berpikir reflektif matematis mahasiswa calon guru Matematika yang berkualitas. Penyusunan *blueprint* dilakukan melalui pemilihan materi atau mata kuliah yang relevan dengan kemampuan matematis yang akan diuji. Mata kuliah teori grup dipilih sebagai mata kuliah wajib yang harus diambil oleh seluruh mahasiswa calon guru Matematika. Yenni & Sukmawati (2019) menyatakan bahwa mempelajari materi teori grup atau struktur aljabar dapat memfasilitasi berkembangnya kemampuan berpikir matematis salah satunya adalah kemampuan berpikir reflektif matematis. Penyusunan butir soal dilakukan dengan menetapkan definisi operasional dan indikator kemampuan berpikir reflektif matematis yang dipakai. Indikator ini selanjutnya diturunkan menjadi butir-butir soal sehingga butir soal yang disusun diharapkan bisa digunakan untuk mengukur kemampuan berpikir reflektif matematis mahasiswa calon guru Matematika.

Berdasarkan hasil penilaian *expert*, masih terdapat beberapa soal yang membingungkan, kurang tepat dan butuh perbaikan. Selain itu, beberapa kalimat dalam butir soal masih belum tajam dan membutuhkan perbaikan untuk bisa dijadikan instrumen kemampuan berpikir reflektif matematis. *Review* secara terbatas juga dilakukan kepada lima orang mahasiswa calon guru Matematika untuk tes keterbacaannya. Revisi dilakukan sesuai dengan masukan dan saran dari para *expert* maupun mahasiswa pada uji coba terbatas.

Wibisono (2018) menyatakan, bahwa instrumen yang divalidasi dengan model Rasch lebih memenuhi definisi pengukuran dan akan menghasilkan informasi yang lebih holistik. Hasil analisis data menunjukkan bahwa nilai reliabilitas *Alpha Cronbach* (KR-20) 0.71 dengan reliabilitas *person* 0.73 dan *item* hingga 0.88. Artinya, instrumen memenuhi kriteria reliabel. Akan tetapi, diperoleh pula informasi terkait *item* yang sebaiknya dieliminasi, yaitu *item* P8. *Item* ini dinilai memang kurang bagus karena memiliki kemiripan dengan *item* P9, namun soal P9 lebih mudah dipahami. *Item* ini meminta

penjelasan pemilihan antara data yang relevan dan data yang tidak relevan. Namun pada intinya, instrumen tes berpikir reflektif matematis ini dinyatakan reliabel dan valid dengan kriteria “sangat reliabel” dan 12 dari 13 butir tes dinyatakan “valid”. Oleh karena itu, instrumen tes berpikir reflektif matematis calon guru Matematika ini bisa digunakan dalam penelitian selanjutnya.

Bab 3

Validasi Instrumen Angket Resiliensi Matematis

Artikel ini dimuat di dalam *Jurnal Kajian Pendidikan Matematika* Volume 6 No. 1 tahun 2020 dengan judul “*Using Rasch Model for Validating Mathematics Resilience Instrumen of preservice mathematics teacher*”. Artikel ini ditulis berdasarkan penelitian yang dilatarbelakangi oleh Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional mengamanatkan bahwa Matematika adalah salah satu mata pelajaran wajib dalam pendidikan formal mulai dari sekolah dasar sampai sekolah menengah (Kemdikbud, 2016). Dengan adanya amanat tersebut, setiap siswa akan mengalami belajar Matematika selama bersekolah kurang lebih lima belas tahun (Muntazhimah, 2020). Organisation for Economic Co-operation and Development (OECD) menyebutkan, bahwa kebanyakan populasi dalam suatu negara di dunia merasakan perasaan negatif ketika mempelajari Matematika (Xolocotzin, 2017). Contoh dari perasaan negatif ialah kurang percaya diri ketika dihadapkan dengan pembelajaran Matematika, bahkan merasa frustrasi, tidak memahami, dan merasa bodoh terhadap langkah penyelesaian yang harus dilakukan (Johnston-Wilder *et al.*, 2015a).

Berbagai macam sikap siswa ketika menghadapi tantangan-tantangan dalam pembelajaran Matematika juga berbeda-beda. Ada yang dapat mengatasi dengan baik namun tidak jarang pula memilih untuk menghindari tantangan tersebut (Johnston-Wilder, 2013). Sikap individu yang mampu beradaptasi, kuat dan tangguh menghadapi tantangan, tetap teguh, optimis dan positif dengan kemampuan dirinya untuk dapat mengatasi tantangan tersebut dapat dikatakan

dengan sikap resiliensi (Hendriana *et al.*, 2017); Hutauruk, 2020; Johnston-Wilder *et al.*, 2015b).

Sikap, kepribadian, dan karakter merupakan fenomena kualitatif yang bersifat *latent* (tidak bisa diamati secara langsung), dan melakukan pembobotannya secara kuantitatif dalam ilmu sosial termasuk pendidikan sebagai kesulitan yang mendasar (Cavanagh & Waugh, 2011). Pendekatan yang masih banyak dipakai untuk melakukan penilaian dalam dunia pendidikan adalah teori tes klasik atau CTT (*classical test theory*) (Sumintono, 2016). Asumsi dasar dari CTT yakni, hasil penjumlahan antara skor murni (T) dan eror (E) akan menghasilkan skor tampak (X). Artinya, dalam skor hasil ujian yang diperoleh seorang siswa terdiri dari skor murni dan error. Error merupakan segala kondisi situasional yang bersifat *latent*. CTT hanya menekankan pada skor tampak dari satu ujian atau dalam hal ini sebagai kemampuan (abilitas) seseorang dari ujian yang diikuti. Hal ini kurang relevan karena skor yang dihasilkan bersifat ordinal sehingga tidak dapat diperlakukan sebagaimana bilangan bulat (Wibisono, 2018).

Model Rasch melakukan pengembangan alat ukur pada ilmu sosial sebagai respons atas berbagai kelemahan paradigma CTT (Sumintono & Widhiarso, 2015). Perbedaan mendasar model Rasch apabila dibandingkan CTT terletak pada bagaimana memperlakukan skor mentah dalam proses analisis. Dalam CTT, skor mentah dalam bentuk peringkat (*rating scale*) langsung dianalisis sebagai data yang seolah-olah memiliki karakter bilangan bulat. Sedangkan dalam Model Rasch, data mentah tidak dapat langsung dianalisis, tetapi harus dikonversikan dulu ke dalam bentuk '*odds ratio*' untuk kemudian dilakukan transformasi logaritma menjadi unit logit sebagai manifestasi probabilitas responden dalam merespons suatu *item* (Wibisono, 2018).

Mengacu pada prosedur di atas, maka model Rasch dapat dijadikan sebagai metode mengembalikan data sesuai kondisinya (Sumintono & Widhiarso, 2013). Kondisi alamiah ini mengacu pada karakteristik dasar data kuantitatif, yaitu bersifat kontinum.

Melalui model Rasch, sebuah respons ordinal dapat ditransformasikan ke dalam bentuk rasio dengan tingkat akurasi lebih tinggi dengan mengacu pada prinsip probabilitas. Hal yang membedakan model Rasch dengan CTT adalah, bahwa dalam analisis data dengan model Rasch, data menyesuaikan model, sedangkan dalam CTT, model dipilih berbasis pada data. Artinya, penggunaan model Rasch dalam validasi instrumen akan menghasilkan informasi yang lebih holistik dan memenuhi definisi pengukuran.

Di Indonesia, penelitian resiliensi matematis masih tergolong topik baru, walaupun sudah ada beberapa di sekolah dasar dan menengah maupun perguruan tinggi. Hal ini selaras dengan pengembangan instrumennya. Padahal fungsi instrumen pada kegiatan evaluasi dalam dunia pendidikan, adalah untuk memperoleh hasil yang lebih baik sesuai dengan kenyataan yang dievaluasi (Sappaile, 2007). Oleh karena itu, evaluasi dalam sebuah penelitian bidang pendidikan harus dibuat menjadi alat pengukur yang baik dan berkualitas, sehingga dapat menghasilkan nilai yang objektif dan akurat (Muntazhimah, 2019). Artikel ini berfokus pada mengkaji reliabilitas dan validitas instrumen kemampuan resiliensi matematis mahasiswa calon guru Matematika dengan memakai model Rasch. *Rasch Model* dijalankan melalui program *winsteps*, yakni program khusus untuk *Rasch Model* yang dapat bekerja pada sistem *Microsoft Windows* yang dibuat oleh John Linacre.

Keterlibatan 126 orang mahasiswa program studi pendidikan Matematika dilakukan menggunakan *purposive sampling*. Kemampuan resiliensi matematis mahasiswa calon guru Matematika diwakili oleh empat buah indikator dan setiap indikator akan diturunkan dalam tujuh hingga sembilan pernyataan baik pernyataan positif maupun negatif, sebagai berikut.

Tabel 7. Kisi-Kisi Instrumen Resiliensi Matematis Calon Guru Matematika

| Indikator | No Item | Positif | Negatif | Banyak Butir |
|--|-------------------------------|-------------------------------|-------------------|--------------|
| Memiliki keyakinan bahwa Matematika sebagai sesuatu yang berharga dan layak untuk ditekuni dan dipelajari. | 1 2 3 4 5 6 7 8 9 | 1 3 4 5 6 7 8 9 | 2 | 9 |
| Memiliki kemauan dan kegigihan dalam mempelajari Matematika, walaupun mengalami kesulitan, hambatan dan tantangan. | 10 11 12 13 14 15 16 17 | 11 13 16 17 | 10 12 14 15 | 8 |
| Memiliki keyakinan pada diri sendiri mampu mempelajari dan menguasai Matematika, baik berdasarkan pemahaman atas Matematika, kemampuan menciptakan strategi, bantuan alat dan orang lain, dan juga pengalaman yang dibangun. | 18 19 20 21 22 23 24 | 18 19 20 23 | 21 22 24 | 7 |
| Memiliki sifat bertahan, tidak pantang menyerah serta selalu memberi respons positif | 25 26 27 28 29 30 31 32 | 25 26 27 28 29 30 31 32 | | 8 |
| Total | | 24 | 8 | 32 |

Dalam model Rasch, sebuah nilai tidak dilihat berdasarkan skor mentah, melainkan nilai logit yang mencerminkan probabilitas keterpilihan suatu *item* pada sekelompok responden. Hal ini digunakan sebagai antisipasi skor mentah dari rating Likert yang berbentuk ordinal yang tidak memiliki kesamaan interval antar skornya. Penggunaan model Rasch untuk data politomi dikembangkan oleh Andrich dengan tetap berlandaskan pada dua teorema dasar, yakni tingkat kemampuan/kesetujuan individu dan tingkat kesulitan *item* untuk disetujui (Misbach & Sumintono, 2014). Output untuk analisis data adalah output *summary statistics* (tabel 3.1) guna mendapatkan informasi reliabilitas serta output *item undimensionality* (tabel 23.0) dan *Item Fit Order* (tabel 10) untuk validitas.

Untuk melihat reliabilitas instrumen, digunakanlah tabel 3.1 (*Summary statistic*) di program *Winsteps*. Salah satu informasi yang dapat diinterpretasikan melalui output *Summary Statistics* ini, yaitu nilai untuk pengukuran reliabilitas, baik pengukuran dari sisi

responden (*Person Reliability*) ataupun dari tinjauan *item* pertanyaan (*Item Reliability*) serta interaksi antara responden dengan *item*. Berikut adalah gambar dari Output *Summary Statistics* (3.1) yang menunjukkan informasi reliabilitas instrumen yang digunakan.

| TABLE 3.1 resiliensi matematika | | | | | | | | | |
|---|-------------|---------|---------|------------|--------------------------------|--------------------|--------|-------|--|
| INPUT: 126 Person 32 Item | | | | | ZOU422ws.TXT Jul 28 2020 22:22 | | | | |
| REPORTED: 126 Person | | | | | 32 Item 4 CATS WINSTEPS 4.4.3 | | | | |
| SUMMARY OF 126 MEASURED Person | | | | | | | | | |
| | TOTAL SCORE | COUNT | MEASURE | MODEL S.E. | INFIT | | OUTFIT | | |
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD | |
| MEAN | 90.4 | 32.0 | .76 | .29 | 1.01 | -.44 | 1.01 | -.45 | |
| SEM | .6 | .0 | .05 | .00 | .06 | .24 | .06 | .24 | |
| P. SD | 6.4 | .0 | .54 | .01 | .69 | 2.72 | .68 | 2.69 | |
| S. SD | 6.4 | .0 | .55 | .01 | .69 | 2.73 | .68 | 2.70 | |
| MAX. | 109.0 | 32.0 | 2.46 | .33 | 4.28 | 8.20 | 4.37 | 8.31 | |
| MIN. | 72.0 | 32.0 | -.74 | .28 | .11 | -6.29 | .11 | -6.20 | |
| REAL RMSE | .33 | TRUE SD | .43 | SEPARATION | 1.31 | Person RELIABILITY | | .63 | |
| MODEL RMSE | .29 | TRUE SD | .46 | SEPARATION | 1.58 | Person RELIABILITY | | .71 | |
| S.E. OF Person MEAN = .05 | | | | | | | | | |
| Person RAW SCORE-TO-MEASURE CORRELATION = 1.00 | | | | | | | | | |
| CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .71 SEM = 3.48 | | | | | | | | | |
| SUMMARY OF 32 MEASURED Item | | | | | | | | | |
| | TOTAL SCORE | COUNT | MEASURE | MODEL S.E. | INFIT | | OUTFIT | | |
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD | |
| MEAN | 355.9 | 126.0 | .00 | .15 | 1.00 | -.24 | 1.01 | -.17 | |
| SEM | 10.8 | .0 | .22 | .00 | .06 | .44 | .06 | .45 | |
| P. SD | 60.4 | .0 | 1.20 | .01 | .31 | 2.44 | .32 | 2.52 | |
| S. SD | 61.3 | .0 | 1.22 | .01 | .31 | 2.48 | .32 | 2.56 | |
| MAX. | 440.0 | 126.0 | 2.81 | .16 | 1.64 | 4.27 | 1.70 | 4.61 | |
| MIN. | 212.0 | 126.0 | -1.88 | .14 | .52 | -4.61 | .52 | -4.60 | |
| REAL RMSE | .16 | TRUE SD | 1.19 | SEPARATION | 7.63 | Item RELIABILITY | | .98 | |
| MODEL RMSE | .15 | TRUE SD | 1.19 | SEPARATION | 8.11 | Item RELIABILITY | | .99 | |
| S.E. OF Item MEAN = .22 | | | | | | | | | |

Gambar 4. Output *Summary Statistics* pada *Winstep*

Output dari tabel *Summary Statistics* di atas akan memberikan informasi bagaimana kualitas instrumen (*item*) dan responden (*person*) dalam menjawab, serta interaksi antara *person* dan *item* sekaligus. Yang perlu kita bahas dari gambar 1 ini adalah *Person measure* untuk menunjukkan rata-rata nilai responden dalam instrumen berdasarkan *mean measure* skornya 0,76. Nilai *Alpha Cronbach* sebagai tolak ukur reliabilitas yaitu interaksi antara *person* dan *item* sekaligus diperoleh dari skor "Test" *reliability* yaitu 0,71. Nilai *Person Reliability* untuk melihat konsistensi jawaban dari responden yaitu 0,63, dan *Item Reliability* yang merupakan nilai reliabilitas butir soal untuk mengetahui kualitas *item-item* dalam instrumen tersebut

yaitu 0,98, Serta *INFIT* dan *OUTFIT MNSQ* dan *ZSTD*, dan tidak ketinggalan *separation*-nya.

Selanjutnya adalah validitas instrumen untuk menguji apakah instrumen ini memang bisa digunakan untuk mengukur kemampuan resiliensi matematis mahasiswa calon guru Matematika. Bagian ini merupakan langkah untuk mengetahui kerangka umum dari fungsi Matematika yang secara khusus menjelaskan mengenai interaksi antara subjek (*person*) dengan butir skala/*item* (*test items*). Tabel yang digunakan dalam *software Winstep* adalah Tabel 23 dan Tabel 10 yaitu *Item undimensionality* dan *Item Fit Order*. *Output*-nya akan menampilkan kegunaan untuk menguji *item* instrumen, yaitu *item* instrumen yang tepat untuk mengukur apa yang seharusnya diukur. Analisis validitas instrumen resiliensi matematis dalam program Winsteps disebut dengan uji *fit* dan *misfit* (*item* yang valid dan tidak valid). Kriteria untuk memeriksa *item* tersebut termasuk *fit* atau *misfit* dapat dilakukan dengan menganalisis *output* dari *Item fit order*. Berikut adalah gambar dari *output* kedua tabel yang menunjukkan informasi kriteria validitas instrumen yang digunakan.

TABLE 23.0 Resiliensi Matematis ZOU421WS.TXT Dec 19 2020 21:10
 INPUT: 126 Person 32 Item REPORTED: 126 Person 32 Item 4 CATS WINSTEPS 4.4.3

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

| | Eigenvalue | Observed | Expected |
|--------------------------------------|------------|----------|----------|
| Total raw variance in observations = | 55.1766 | 100.0% | 100.0% |
| Raw variance explained by measures = | 23.1766 | 42.0% | 42.2% |
| Raw variance explained by persons = | 2.7363 | 5.0% | 5.0% |
| Raw variance explained by items = | 20.4403 | 37.0% | 37.2% |
| Raw unexplained variance (total) = | 32.0000 | 58.0% | 57.8% |
| Unexplned variance in 1st contrast = | 4.2286 | 7.7% | 13.2% |
| Unexplned variance in 2nd contrast = | 2.9001 | 5.3% | 9.1% |
| Unexplned variance in 3rd contrast = | 2.5294 | 4.6% | 7.9% |
| Unexplned variance in 4th contrast = | 2.1698 | 3.9% | 6.8% |
| Unexplned variance in 5th contrast = | 1.7981 | 3.3% | 5.6% |

Gambar 5. *Output Item Undimensionality* pada *Winstep*

Gambar 5 di atas menunjukkan, bahwa kolom *raw variance explained by measure* memiliki skor 42% dan kolom *unexplned variance in 1st contrast* memiliki skornya *eigenvalue* 4.2286 serta skor *observed*-nya 7,7%.

Selanjutnya adalah *item fit order*. Pada tahapan ini, yang perlu diperhatikan untuk analisis adalah nilai *Oufit Mean Square (MNSQ)*,

nilai *Outfit Z-Standard (ZSTD)* dan nilai *Point Measure Coorelation (Pt Mean Corr)*. Sebagai contoh pada baris pertama adalah hasil *output* untuk *item* nomor 21 pada instrumen resiliensi matematis, berturut-turut skornya adalah 1,70 dan 4,59 serta -0,10. Baris kedua adalah *item* nomor 9 berturut-turut skornya adalah 1,65 dan 4,61 serta 0,00. Baris ketiga adalah *item* nomor 22 berturut-turut skornya adalah 1,51 dan 3,63 serta 0,13 dan seterusnya hingga baris ke 32. Secara rinci *output item fit order* ada pada gambar 6 sebagai berikut.

TABLE 10.1 resilliansi matematika ZOU422ws.TXT Jul 28 2020 22:22
 INPUT: 126 Person 32 Item REPORTED: 126 Person 32 Item 4 CATS WINSTEPS 4.4.3
 Person: REAL SEP.: 1.31 REL.: .63 ... Item: REAL SEP.: 7.63 REL.: .98

Item STATISTICS: MISFIT ORDER

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | TOTAL MEASURE | MODEL S.E. | INFIIT MNSQ | ZSTD | OUTFIT MNSQ | ZSTD | PTMEAS CORR. | R-AL EXP. | EXACT OBS% | MATCH EXP% | Item |
|--------------|-------------|-------------|---------------|------------|-------------|-------|-------------|-------|--------------|-----------|------------|------------|------|
| 21 | 386 | 126 | -.58 | .15 | 1.64 | 4.27 | 1.70 | 4.59 | A-.10 | .31 | 54.8 | 62.3 | S21 |
| 9 | 406 | 126 | -1.04 | .15 | 1.55 | 3.94 | 1.65 | 4.61 | B.00 | .30 | 62.7 | 60.1 | S9 |
| 22 | 323 | 126 | .72 | .14 | 1.50 | 3.58 | 1.51 | 3.64 | C.13 | .33 | 46.8 | 54.3 | S22 |
| 2 | 371 | 126 | -.25 | .15 | 1.44 | 3.02 | 1.46 | 3.16 | D.28 | .31 | 61.1 | 62.4 | S2 |
| 19 | 377 | 126 | -.38 | .15 | 1.44 | 3.01 | 1.46 | 3.17 | E.17 | .31 | 56.3 | 62.6 | S19 |
| 12 | 212 | 126 | 2.81 | .14 | 1.37 | 3.03 | 1.36 | 2.97 | F-.12 | .32 | 46.0 | 53.3 | S12 |
| 6 | 355 | 126 | .09 | .14 | 1.32 | 2.29 | 1.33 | 2.38 | G.23 | .32 | 53.2 | 60.5 | S6 |
| 13 | 246 | 126 | 2.14 | .14 | 1.26 | 2.16 | 1.25 | 2.16 | H-.06 | .33 | 61.1 | 54.4 | S13 |
| 30 | 391 | 126 | -.69 | .15 | 1.18 | 1.41 | 1.26 | 1.94 | I.12 | .31 | 63.5 | 62.0 | S30 |
| 20 | 383 | 126 | -.51 | .15 | 1.15 | 1.17 | 1.15 | 1.20 | J.40 | .31 | 57.1 | 62.5 | S20 |
| 8 | 369 | 126 | -.21 | .15 | 1.14 | 1.09 | 1.14 | 1.08 | K.50 | .31 | 54.0 | 62.1 | S8 |
| 14 | 268 | 126 | 1.74 | .14 | 1.14 | 1.20 | 1.14 | 1.24 | L.33 | .34 | 61.1 | 54.3 | S14 |
| 23 | 257 | 126 | 1.94 | .14 | 1.06 | .60 | 1.07 | .68 | M.20 | .34 | 60.3 | 54.4 | S23 |
| 4 | 373 | 126 | -.30 | .15 | .96 | -.23 | .98 | -.15 | N.56 | .31 | 57.1 | 62.5 | S4 |
| 10 | 229 | 126 | 2.47 | .14 | .98 | -.12 | .95 | -.39 | O.09 | .33 | 66.7 | 53.7 | S10 |
| 7 | 364 | 126 | -.10 | .15 | .92 | -.61 | .93 | -.54 | P.39 | .32 | 62.7 | 61.6 | S7 |
| 18 | 440 | 126 | -1.88 | .16 | .87 | -1.27 | .87 | -1.27 | Q.40 | .28 | 69.0 | 57.9 | S18 |
| 3 | 388 | 126 | -.62 | .15 | .86 | -1.13 | .86 | -1.16 | O.67 | .31 | 58.7 | 62.2 | S3 |
| 15 | 230 | 126 | 2.45 | .14 | .85 | -1.33 | .86 | -1.25 | N.09 | .33 | 73.0 | 53.8 | S15 |
| 16 | 407 | 126 | -1.06 | .15 | .81 | -1.66 | .82 | -1.58 | M.17 | .30 | 71.4 | 60.0 | S16 |
| 17 | 429 | 126 | -1.60 | .16 | .82 | -1.77 | .81 | -1.85 | L.42 | .29 | 63.5 | 57.1 | S17 |
| 32 | 397 | 126 | -.83 | .15 | .78 | -1.91 | .77 | -2.01 | K.63 | .30 | 63.5 | 61.1 | S32 |
| 27 | 367 | 126 | -.17 | .15 | .77 | -1.93 | .76 | -1.94 | J.38 | .31 | 69.0 | 61.9 | S27 |
| 29 | 378 | 126 | -.40 | .15 | .77 | -1.85 | .77 | -1.86 | I.57 | .31 | 67.5 | 62.6 | S29 |
| 31 | 364 | 126 | -.10 | .15 | .76 | -1.98 | .75 | -2.04 | H.50 | .32 | 73.0 | 61.6 | S31 |
| 5 | 362 | 126 | -.06 | .15 | .75 | -2.11 | .75 | -2.05 | G.51 | .32 | 67.5 | 61.4 | S5 |
| 1 | 390 | 126 | -.67 | .15 | .68 | -2.82 | .69 | -2.75 | F.50 | .31 | 69.0 | 62.0 | S1 |
| 28 | 422 | 126 | -1.42 | .16 | .68 | -3.13 | .69 | -3.09 | E.53 | .29 | 67.5 | 57.6 | S28 |
| 26 | 392 | 126 | -.71 | .15 | .68 | -2.87 | .67 | -2.91 | D.49 | .31 | 69.8 | 61.8 | S26 |
| 25 | 350 | 126 | .19 | .14 | .64 | -3.24 | .64 | -3.19 | C.42 | .32 | 72.2 | 59.5 | S25 |
| 11 | 374 | 126 | -.32 | .15 | .59 | -3.70 | .59 | -3.72 | B.57 | .31 | 73.8 | 62.5 | S11 |
| 24 | 388 | 126 | -.62 | .15 | .52 | -4.61 | .52 | -4.60 | A.46 | .31 | 77.0 | 62.2 | S24 |
| MEAN | 355.9 | 126.0 | .00 | .15 | 1.00 | -.2 | 1.01 | -.2 | | | 63.4 | 59.6 | |
| P.SD | 60.4 | .0 | 1.20 | .01 | .31 | 2.4 | .32 | 2.5 | | | 7.6 | 3.3 | |

Gambar 6. Output Item Fit Order pada Winstep

Seperti yang diketahui, bahwa reliabilitas adalah ketetapan atau keajekan alat dalam menilai apa yang dinilainya. Suatu instrumen dikatakan reliabel apabila hasil dari instrumen akan memberikan hasil yang relatif sama atau bersifat stabil (Sumintono & Widhiarso,

2015). Uji validitas dan reliabilitas merupakan hal yang esensial yang perlu dipenuhi agar instrumen yang digunakan dalam sebuah penelitian dapat memiliki nilai kepercayaan. Melalui gambar 1 yaitu output *Summary Statistics* untuk instrumen tes resiliensi matematis mahasiswa calon guru Matematika yang ada pada pembahasan *Result* di atas, maka akan dicermati beberapa hal sebagai berikut. Pertama adalah *person measure* yang menunjukkan nilai 0.76, karena lebih besar dari logit 0,0. Nilai rata-rata yang lebih besar dari logit menunjukkan kecenderungan abilitas responden lebih besar dari pada tingkat kesulitan soal atau responden cenderung mampu merespons pernyataan-pernyataan yang ada pada instrumen. Kedua adalah nilai *Alpha Cronbach* (mengukur reliabilitas yaitu interaksi antara responden dengan *item*) yang menunjukkan 0.71. Interpretasi skor ini bisa dilihat pada tabel 8 di bawah. Dengan nilai tersebut maka kriterianya termasuk “Bagus”, sehingga instrumen yang digunakan dapat dikatakan reliabel.

Tabel 8. Interpretasi Uji Reliabilitas Berdasarkan Nilai *Cronbach Alpha*

| Nilai | Interpretasi |
|--------------------|--------------|
| $a > 0,8$ | Bagus sekali |
| $0,7 < a \leq 0,8$ | Bagus |
| $0,6 < a \leq 0,7$ | Cukup |
| $0,5 < a \leq 0,6$ | Jelek |
| $a < 0,5$ | Buruk |

(Sumintono & Widhiarso, 2015)

Nilai *Person Reliability* dan *Item Reliability* berturut-turut menunjukkan 0,63 dan 0.98. *Person Reliability* menunjukkan bahwa konsistensi jawaban dari responden dikatakan “Lemah”, namun *Item Reliability* menunjukkan kriteria nilai yang “Istimewa” sehingga dapat dikatakan kualitas *item-item* yang digunakan dalam instrumen sangat reliabel.

Selanjutnya adalah nilai INFIT MNSQ dan OUTFIT MNSQ, serta INFIT ZSTD dan OUTFIT ZSTD. INFIT MNSQ nilai *person*-nya 1.01 dan nilai *item*-nya 1.00, OUTFIT MNSQ nilai *person*-nya 1.01 dan nilai *item*-nya 1.01, terlihat bahwa Nilai yang ditunjukkan pada tabel *person* dan *item* dari INFIT MNSQ dan OUTFIT MNSQ semuanya mendekati kepada nilai 1,00, karena semakin data mendekati nilai 1,00 kualitas dikatakan semakin baik. Kemudian untuk INFIT ZSTD nilai *person*-nya -0.44 dan nilai *item*-nya -0.24, dan INFIT ZSTD nilai *person*-nya -0.45 dan nilai *item*-nya -0.17. Nilai ideal adalah 0,0 dengan artian semakin mendekati nilai ideal maka kualitas semakin baik. Untuk *item*, semua skor cenderung dekat dengan 0,0, namun untuk *person* masih jauh artinya data responden memang ada yang kurang baik dalam menyelesaikan instrumen yang diberikan.

Berikutnya adalah pengelompokan *person* dan *item* dapat diketahui dari nilai *separation*. Semakin besar nilai *separation* maka kualitas instrumen dalam hal keseluruhan responden dan butir semakin baik, karena bisa mengidentifikasi kelompok responden dan kelompok *item*. Untuk *separation person* skornya pada gambar 1 adalah 1,61 dan skor *separation item* adalah 7,63. Berdasarkan tinjauan di atas, maka instrumen kemampuan resiliensi matematis mahasiswa calon guru Matematika dalam hal ini dapat dinilai sebagai instrumen dengan tingkat reliabilitas yang tinggi.

Tabel 9. Hasil Pengolahan Uji Reliabilitas

| <i>Cronbach Alpha</i> | Interpretasi | <i>Item Reliability</i> | Interpretasi | <i>Person Reliability</i> | Interpretasi | Kesimpulan |
|-----------------------|--------------|-------------------------|--------------|---------------------------|--------------|------------|
| 0,71 | Bagus | 0,98 | Istimewa | 0,63 | Lemah | Reliabel |

Berdasarkan Tabel 9, interpretasi *Cronbach Alpha* adalah 'bagus'. Terdapat kesesuaian antara *item* dan *person* (responden). Kemudian konsistensi jawaban dari responden (*Person Reliability*) dapat dikatakan 'lemah' dengan kualitas butir soal instrumen (*Item Reliability*) adalah 'istimewa'. Artinya, instrumen kemampuan resiliensi matematis mahasiswa calon guru yang diteliti ini dapat dikatakan reliabel.

Selanjutnya adalah validitas yang mempermasalahkan seberapa jauh *item-item* tes mampu mengukur apa-apa yang benar-benar hendak diukur sesuai dengan konsep khusus atau definisi konseptual yang telah ditetapkan (Djaali & Muljono, 2008). Dalam analisis Rasch, uji validitas atau *item undimensionality* (unidimensionalitas instrumen) (Sumintono & Widhiarso, 2015) dipilih untuk mengevaluasi apakah instrumen yang dikembangkan mampu mengukur apa yang seharusnya diukur sehingga dapat dikatakan valid. Analisis Rasch menggunakan analisis komponen utama (*principal component analysis*) dari *standardized residual variance (in Eigenvalue units)* (Sumintono & Widhiarso, 2015). Uji validitas berdasarkan *item undimensionality* dapat terlihat dari nilai *raw variance explained by measures*. Interpretasi *item undimensionality* berdasarkan nilai *raw variance explained by measures* ditunjukkan oleh skor > 20% dikatakan terpenuhi, > 40% bagus dan > 60% untuk kriteria istimewa.

Untuk mengetahui ada atau tidaknya butir soal yang bermasalah dan tidak cocok, perlu dilihat dari nilai *eigenvalue* dan *observed* dalam *unexplained variance 1st contrast* dengan nilai *eigenvalue* harus kurang dari 3. Nilai ini menunjukkan, bahwa tidak ada butir soal yang bermasalah dan nilai *observed* harus kurang dari 15% untuk menunjukkan butir soal yang sesuai (*item fit*). Berikut adalah hasil dari pengolahan validitas dengan *software Winstep* versi 4.4.3.

Tabel 10. Hasil Pengolahan Validitas Instrumen

| <i>Raw variance explained by measures</i> | Interpretasi | <i>unexplained variance 1st contrast</i> | | Interpretasi |
|---|--------------|---|-----------------|---------------------------------|
| | | <i>Eigenvalue</i> | <i>observed</i> | |
| 42.0% | Bagus | 4.2286 | 7.7% | Ada <i>item</i> yang bermasalah |

Kita mengetahui bahwa nilai *raw variance explained by measures* di atas *item* tes resiliensi keseluruhan berada pada kategori “bagus”. Kemudian nilai *observed* pada *unexplained variance 1st contrast* menunjukkan tidak ada kecenderungan ketidaksesuaian butir soal sehingga dapat digunakan. Akan tetapi, pada nilai

eigenvalue yang lebih dari 3 mengindikasikan ada butir soal yang bermasalah sehingga dapat dilakukan analisis lebih lanjut dengan analisis *item fit order* untuk menentukan suatu butir soal dapat dipertahankan atau harus diganti.

Item fit digunakan untuk menjelaskan apakah butir soal berfungsi normal untuk melakukan pengukuran atau tidak nilai *outfit means-square*, *Outfit Z-Standard*, dan *Point Measure Correlation*. Ini merupakan kriteria yang digunakan untuk melihat tingkat kesesuaian butir (*item fit*) (Boone *et al.*, 2014; Bond & Fox, 2015). Adapun kriteria yang digunakan untuk memeriksa kesesuaian butir soal dapat dilihat pada Tabel 11 sebagai berikut.

Tabel 11. Nilai Kriteria Kesesuaian Butir Soal

| Kriteria | Nilai |
|----------------------------------|---------------------------------------|
| <i>Outfit Mean Square</i> (MNSQ) | $0,5 < \text{MNSQ} < 1,5$ |
| <i>Outfit Z-standart</i> (ZSTD) | $-2,0 < \text{ZSTD} < +2,0$ |
| <i>Point Measure Correlation</i> | $0,4 < \text{PT Measure Corr} < 0,85$ |

(Sumintono & Widhiarso, 2015)

Berdasarkan Tabel 11, apabila ketiga kriteria terpenuhi pada butir soal maka dapat dikatakan butir soal telah “sesuai”. Dalam hal ini dapat dipastikan, bahwa kualitas butir soal tersebut bagus dan dapat digunakan. Sementara apabila hanya terdapat dua kriteria atau satu kriteria yang terpenuhi, maka butir soal masih dapat dipertahankan dan tidak perlu diubah sehingga dapat dikategorikan “sesuai” dan dapat digunakan. Akan tetapi, apabila ketiga kriteria tidak terpenuhi, maka dapat dikatakan butir soal tidak sesuai dan dapat dipastikan butir soal “tidak sesuai” sehingga perlu diperbaiki ataupun diganti. Adapun contoh hasil dari pengolahan kesesuaian butir soal dapat dilihat pada Tabel 12 yang hanya sampai pada kode pernyataan no 10 pada instrumen. Tabel 12 merupakan hasil yang bisa dilihat dengan berpedoman pada gambar 6.

Tabel 12. Hasil Pengolahan *Item Fit Order*

| Nomor Soal | Kode Soal | Outfit | | PT Measure Corr. | Status tidak memenuhi | Interpretasi |
|------------|-----------|--------|-------|------------------|-----------------------|--------------|
| | | MNSQ | ZFTD | | | |
| 1 | P1 | 0,69 | -2,75 | 0,50 | 1 Kriteria | Sesuai |
| 2 | P2 | 1,46 | 3,16 | 0,28 | 2 Kriteria | Sesuai |
| 3 | P3 | 0,86 | -1,16 | 0,67 | - | Sesuai |
| 4 | P4 | 0,98 | -0,15 | 0,56 | - | Sesuai |
| 5 | P5 | 0,75 | -2,05 | 0,51 | 1 Kriteria | Sesuai |
| 6 | P6 | 1,33 | 2,38 | 0,23 | 2 Kriteria | Sesuai |
| 7 | P7 | 0,93 | -0,54 | 0,39 | 1 Kriteria | Sesuai |
| 8 | P8 | 1,14 | 1,08 | 0,50 | - | Sesuai |
| 9 | P9 | 1,65 | 4,61 | 0,00 | 3 Kriteria | Tidak Sesuai |
| 10 | P10 | 0,95 | -0,39 | 0,09 | 1 Kriteria | Sesuai |

Tabel 6 menunjukkan mekanisme validasi item yang ada pada instrumen. Diperoleh bahwa item nomor 9 tidak sesuai. Dari mekanisme tersebut dan diterapkan ke seluruh item soal pada angket resiliensi matematis mahasiswa calon guru Matematika, diperoleh bahwa item bernomor 9, 21 dan 22 tidak memenuhi tiga syarat kriteria validitas, sehingga dapat dikatakan item tersebut tidak valid (*Misfit*) dan tidak dapat dipertahankan, serta harus dibuang atau diganti. Sedangkan, item lainnya telah memenuhi minimal satu kriteria untuk dikatakan item fit atau valid, dan masih dapat dipertahankan untuk digunakan pada penelitian. Dari 32 item pada angket, diperoleh 29 butir yang dinyatakan valid. Secara keseluruhan diperoleh hasil bahwa instrumen kemampuan resiliensi matematis mahasiswa calon guru Matematika pada penelitian ini dinyatakan reliabel dan valid dengan kriteria “sangat reliabel” dan 29 dari 32 item pernyataan dikatakan “valid”. Artinya, instrumen ini dapat digunakan untuk meneliti kemampuan resiliensi matematis mahasiswa calon guru Matematika.

Daftar Pustaka

- Agustan, S., Juniati, D., & Siswono, T. Y. E. (2017). Reflective thinking in solving an algebra problem: A case study of field independent-prospective teacher. *Journal of Physics: Conference Series*, 893(1), Article 012002. <https://doi.org/10.1088/1742-6596/893/1/012002>
- Alavi, K., Isa, K., & Palpanadan, S. T. (2020). Application of Rasch model on resilience in higher education: An examination of validity and reliability of Malaysian academicians happiness index (MAHI). *International Journal of Higher Education*, 9(4), 261–271. <https://doi.org/10.5430/ijhe.v9n4p261>
- Amidu, AR. (2012). Exploring real estate students' learning approaches, reflective thinking and academic performance. *48th ASC Annual International Conference Proceedings, June 2012*. <http://ascpro0.ascweb.org/archives/cd/2012/paper/CEUE214002012.pdf>
- An, M., & Yu, X. (2021). A Rasch analysis of emerging adults' health motivation questionnaire in higher education context. *PLoS ONE: Rasch Analysis Health Motivation*, 16(3).
- Anggraini, A., & Muntazhimah, M. (2021). Pengembangan Instrumen Kemampuan Berpikir Reflektif Matematis Siswa Madrasah Aliyah. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 10(4), 2465. <https://doi.org/10.24127/ajpm.v10i4.4223>
- Arifin, Z. (2016). *Evaluasi pembelajaran: Prinsip, teknik, dan prosedur [Learning evaluation: Principles, techniques and procedures]*. PT Remaja Rosdakarya.
- Bond, T., & Fox, C. (2015). *Applying the Rasch Model*. In *Applying the Rasch Model*. Routledge. <https://doi.org/10.4324/9781315814698>

- Boone, W. J., Staver, J. R., Yale, M. S., Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Item Measures. *Rasch Analysis in the Human Sciences*, 93–110. https://doi.org/10.1007/978-94-007-6857-4_5
- Cavanagh, R. F., & Waugh, R. F. (2011). Applications of Rasch Measurement in Learning Environments Research. *Applications of Rasch Measurement in Learning Environments Research*. <https://doi.org/10.1007/978-94-6091-493-5>
- Clarà, M. (2015). What is reflection? Looking for clarity in an ambiguous notion. *Journal of Teacher Education*, 66(3), 261–271. <https://doi.org/10.1177/0022487114552028>
- Djaali, & Muljono, P. (2008). *Pengukuran dalam bidang pendidikan*. Jakarta: PT.Grasindo
- Fisher, W.P.Jr. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions*, Vlo. 21 No. 1 dapat diakses dari: <https://www.rasch.org/rmt/rmt211m.htm>
- Gurol, A. (2011). A.Gurol-A8.pdf. *Energy Education Science and Technology Part B: Social and Educational Studies*, 3(3), 387–402.
- Hendriana, H., Putra, H. D., & Hidayat, W. (2019). How to Design Teaching Materials to Improve the Ability of Mathematical Reflective Thinking of Senior High School Students in Indonesia? *Eurasia Journal of Mathematics, Science and Technology Education*, 15(12). <https://doi.org/10.29333/ejmste/112033>
- Hendriana, H., Soemarmo, U., & Rohaeti, E. E. (2017). *Hard Skill dan Soft Skil*. Refika Aditama
- Hutauruk, A. J. B. (2020). Indikator Pembentuk Resiliensi Matematis Mahasiswa Prodi Pendidikan Matematika FKIP. *Sepren*, 1(02), 78–91. <https://doi.org/10.36655/sepren.v1i02.227>
- Idrus, L. (2019). Evaluasi Dalam Proses Pembelajaran. *Manajemen Pendidikan Islam*, 9(2), 920–935.
- Johnston-Wilder, S, Lee, C., Brindley, J., & Garton, E. (2015a). Developing Mathematical Resilience in School-Students

- Who Have Experienced Repeated Failure. *Iceri2015: 8Th International Conference of Education, Research and Innovation*.
- Johnston-Wilder, S, Lee, C., Brindley, J., & Garton, E. (2015b). Developing Peer Coaching for Mathematical Resilience in Post-16 Students Who Are Encountering Mathematics in Other Subjects. *Iceri2015: 8Th International Conference of Education, Research and Innovation*, 6002–6011.
- Johnston-Wilder, Sue. (2013). Measuring mathematical resilience : an application of the construct of resilience to the study of mathematics. *AERA 2013, San Francisco California, University of Warwick*, 1– 15. <http://wrap.warwick.ac.uk/51559>
- Kemendikbud. (2016). Standar Kompetensi Lulusan Pendidikan Dasar dan Menengah. *Lampiran Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 20 Tahun 2016*.
- Magdalena, I., Syariah, E. N., Mahromiyati, M., & Nurkamilah, S. (2021). Analisis Instrumen Tes Sebagai Alat Evaluasi Pada Mata Pelajaran SBdP Siswa Kelas II SDN Duri Kosambi 06 Pagi. *Nusantara: Jurnal Pendidikan Dan Ilmu Sosial*, 3(2), 276–287.
- Misbach, I. H., & Sumintono, B. (2014). Pengembangan dan Validasi Instrumen “Persepsi Siswa Terhadap Karakter Moral Guru” di Indonesia dengan Model Rasch. *PROCEEDING Seminar Nasional Psikometri*, 148–162
- Muin, A., Novianti, L., & Musyrifah, E. (2018). Analysis of mathematical reflective thinking skills based on learning model and mathematical prior knowledge. *3rd International Conferences on Education in Muslim Society (ICEMS 2017)*. <https://doi.org/10.2991/icems-17.2018.5>
- Muntazhimah, M., & Wahyuni, R. (2022). The Development and Validation of Mathematical Reflective Thinking Test Instruments for Prospective Mathematics Teachers Using the Rasch Model. *Jurnal Elemen*, 8(1), 175–186. <https://doi.org/10.29408/jel.v8i1.3981>

- Muntazhimah, M., Putri, S., & Khusna, H. (2020). Rasch Model untuk Memvalidasi Instrumen Resiliensi Matematis Mahasiswa Calon Guru Matematika. *JKPM (Jurnal Kajian Pendidikan Matematika)*, 6(1), 65. <https://doi.org/10.30998/jkpm.v6i1.8144>
- Muntazhimah. (2019). Pengembangan Instrumen Kemampuan Berpikir Reflektif Matematis siswa Kelas 8 SMP. *Imaginer: Jurnal Matematika Dan Pendidikan Matematika*, 1(5), 237–242.
- Mutmainna, D., Mania, S., & Sriyanti, A. (2018). Pengembangan Instrumen Tes Diagnostik Pilihan Ganda Dua Tingkat Untuk Mengidentifikasi Pemahaman Konsep Matematika. *Mapan: Jurnal Matematika Dan Pembelajaran*, 6, 56–69. <https://doi.org/https://doi.org/10.24252/mapan.2018v6n1a6>
- Nindiasari, H. (2011). Pengembangan Bahan Ajar dan Instrumen untuk Meningkatkan Berpikir Reflektif Matematis Berbasis Pendekatan Metakognitif pada Siswa Sekolah Menengah Atas (SMA). *Prosiding Seminar Nasional Matematika Dan Pendidikan Matematika*, (ISBN : 978-979-16353-6-3).
- Nindiasari, H., Novaliyosi, N., & Pamungkas, A. S. (2016). Pengembangan bahan ajar untuk meningkatkan tahapan kemampuan berpikir matematis [Development of teaching materials to improve the stages of mathematical thinking skills]. *JPPM*, 9(1), 109–114
- Nur, L., Nurani, L. A., Suryana, D., & Ahmad, A. (2020). Rasch Model Application on Character Development Instrument for Elementary School Students. *International Journal of Learning, Teaching and Educational Research*, 19(3), 437–459. <https://doi.org/10.26803/ijlter.19.3.24>
- Osnal, Suhartoni, & Wahyudi, I. (2015). Meningkatkan kemampuan guru dalam menyusun tes hasil belajar akhir semester melalui workshop di KKG Gugus 02 Kecamatan Sumbermalang tahun 2014/2015 [Improving teacher ability in preparing final semester learning outcomes tests through workshops at KKG

- Cluster 02 Sumbermalang District 2014/2015]. *Pancaran Pendidikan*, 5(1), 67–82. <https://jurnal.unej.ac.id/index.php/pancaran/article/view/2604>
- Pamungkas, A. S., Mentari, N., & Nindiasari, H. (2018). Analisis Kemampuan Berpikir Reflektif Siswa SMP Berdasarkan Gaya Belajar. *NUMERICAL: Jurnal Matematika Dan Pendidikan Matematika*, 2(1), 69. <https://doi.org/10.25217/numerical.v2i1.209>
- Prabowo, A., & Dahlan, J. A. (2020). Pengembangan tes matematika dengan konteks COVID19 untuk siswa SMP/MTs kelas VIII [Development of mathematics tests with the context of COVID-19 for class VIII SMP/MTs students]. *Jurnal Elemen*, 6(2), 302–317. <https://doi.org/10.29408/jel.v6i2.2115>
- Prabowo, A., Kusdinar, U., & Rahmawati, U. (2018). Pelatihan pengembangan instrumen tes mata pelajaran matematika SMP [Training for development of middle school mathematics subject test instruments]. *International Journal of Community Service Learning*, 2(3), 141–148. <https://doi.org/10.23887/ijcsl.v2i3.14189>
- Putra, H. D. (2016). Pengembangan Instrumen untuk Meningkatkan Kemampuan Berpikir Reflektif Matematis Siswa SMA dengan Pendekatan Scientific Disertai Strategi What If Not. *Prosiding Seminar Nasional Matematika Dan Pendidikan Matematika STKIP Siliwangi*, 4(ISSN 2338-8315).
- Rahmadhani, E., Wahyuni, S., Mandasari, L., & Setiawan, Y. (2020). Reflective thinking ability of teacher candidate students based on ability level and gender. *Companion Proceedings of the SEADRIC 2019 (2020)*, 165–172. <https://doi.org/10.24071/seadr.2019.23>
- Rowe, S. E. (2001). Development of a test blueprint for the National Association of Industrial Technology certification exam [Unpublished Doctoral Dissertation]. Iowa State University. <https://lib.dr.iastate.edu/rtd/668>

- Sabekti, A. W., Khoirunnisa, F., & Liliyasi, L. (2020). Validating the Indonesian version of reflective thinking questionnaire and investigation of the relationship between pre-service teachers' reflective thinking and academic achievement. *Companion Proceedings of the SEADRIC 2019 (2020)*, 138–144. <https://doi.org/10.24071/seadr.2019.19>
- Sappaile, B. I. (2007). Konsep Instrumen Penelitian Pendidikan. *Jurnal Pendidikan Dan Kebudayaan*, 13(66), 379-391. <https://doi.org/10.24832/jpnk.v13i66.356>
- Sari, A. D., Noer, S. H., & Asmiati. (2020). Pengembangan Model Creative Problem Solving (CPS) Untuk Meningkatkan Kemampuan Berpikir Reflektif. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, 04(02), 1115–1128.
- Spaan, M. (2006). Test and item specifications development. *Language Assessment Quarterly*, 3(1), 71– 79. https://doi.org/10.1207/s15434311laq0301_5
- Suharman. (2018). Tes sebagai alat ukur prestasi akademik [Test as a measuring tool for academic achievement]. *Jurnal Ilmiah Pendidikan Agama Islam*, 10(1), 93–115.
- Sumintono, B. (2016). Penilaian Keterampilan Berpikir Tingkat Tinggi : Aplikasi Pemodelan Rasch pada Asesmen Pendidikan. *Seminar Nasional Pendidikan IPA, FKIP Jurusan PMIPA, Universitas Lambung Mangkurat*
- Sumintono, B., & Widhiarso, W. (2013). Model Rasch untuk Penelitian Sosial Kuantitatif. *ITS Surabaya, November 201*, 1–9. <http://eprints.um.edu.my/id/eprint/11414>
- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch Untuk Penelitian Ilmu-ilmu Sosial Edisi Revisi* (Bambang Trim, ed.). Cimahi: Trim Komunikata Publishing House.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Cimahi: Trim Komunikata.
- Sundayana, R. (2018). *Statistika Penelitian Pendidikan*. Bandung: Alfabeta,cv.

- Thahir, A., Komarudin, Hasanah, U. N., & Rahmahwaty. (2019). MURDER Learning Models and Self Efficacy: Impact on Mathematical Reflective Thinking Ability. *Journal for the Education of Gifted Young Scientists*, (December).
- Wahyudi, W. (2012). Assesment pembelajaran berbasis portofolio di sekolah [Portfolio-based learning assessment in schools]. *Jurnal Visi Ilmu Pendidikan*, 2(1), 288–297. <https://doi.org/10.26418/jvip.v2i1.370>
- Wardhani, D., & Putra, A. (2016). Pengembangan instrumen tes standar kognitif pada mata pelajaran IPA kelas 7 SMP di Kabupaten Banjar [Development of cognitive standard test instruments for science subjects for grade 7 junior high schools in Banjar regency]. *Proceeding Biology Education Conference: Biology, Science, Enviromental, and Learning*, 13(1), 75-82. <https://jurnal.uns.ac.id/prosbi/article/view/5658/5026>
- Wibisono, S. (2018). Aplikasi Model Rasch Untuk Validasi Instrumen Pengukuran Fundamentalisme Agama Bagi Responden Muslim. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia (JP3I)*, 5(1). <https://doi.org/10.15408/jp3i.v5i1.9239>
- Wibisono, S. (2018). Aplikasi Model Rasch Untuk Validasi Instrumen Pengukuran Fundamentalisme Agama Bagi Responden Muslim. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia (JP3I)*, 5(1). <https://doi.org/10.15408/jp3i.v5i1.9239>
- Wright, B. D., & Mok, M. M. C. (2004). An Overview of the Family of Rasch Measurement Models. In *Introduction to Rasch Measurement: Theory, Models and Applications* (pp. 1–24). JAM Press.
- Xolocotzin, U. (2017). Understanding Emotions in Mathematical Thinking and Learning. In *Understanding Emotions in Mathematical Thinking and Learning*. <https://doi.org/10.1016/C2014-0-02036-9>
- Yenni, Y., & Sukmawati, R. (2019). Analisis kemampuan berpikir reflektif matematis berdasarkan minat belajar pada mata kuliah struktur aljabar [Analysis of mathematical reflective

thinking ability based on learning interest in algebra structure course]. *Teorema: Teori dan Riset Matematika*, 4(2), 75-82. <https://doi.org/10.25157/teorema.v4i2.2283>

Yuen Lie Lim, LA. A. (2011). A comparison of students' reflective thinking across different years in a problem-based learning environment. *Instructional Science*, 39(2), 171–188. <https://doi.org/10.1007/s11251-009-9123-8>

Profil Penulis

Muntazhimah, M.Pd. lahir di Medan pada 11 Agustus 1988. Penulis menempuh studi S-1 Pendidikan Matematika di Universitas Negeri Medan (2007–2011) dengan skripsi berjudul “Peningkatan Hasil Belajar Matematika dengan Model Pembelajaran Kooperatif Tipe STAD Berbantuan *Chip* Bilangan pada Siswa Kelas VII SMPN 1 Lubuk Pakam”. Jenjang S-2 ditempuhnya di Universitas Pendidikan Indonesia sejak tahun 2013 hingga 2015, dengan tesis berjudul “Pembelajaran Kolaboratif Berbantuan Cabri 3D untuk Meningkatkan Kemampuan Spasial dan Komunikasi Matematis serta Dampaknya terhadap Kecemasan Matematis Siswa SMP” dan S-3 Pendidikan Matematika ditempuhnya di Universitas Pendidikan Indonesia. Penulis dapat dihubungi di alamat *e-mail* muntazhimah@uhamka.ac.id.