

PENGEMBANGAN INSTRUMEN KEMAMPUAN BERPIKIR REFLEKTIF MATEMATIS SISWA MADRASAH ALIYAH

Anggi Anggraini¹, Muntazhimah^{2*}

^{1,2*} Universitas Muhammadiyah Prof. Dr. Hamka

*Corresponding author.

E-mail: anggianggraini192@gmail.com¹⁾
muntazhimah@uhamka.ac.id^{2*)}

Received 14 September 2021; Received in revised form 23 November 2021; Accepted 27 December 2021

Abstrak

Tujuan penelitian ini adalah untuk mendapatkan sebuah instrumen kemampuan berpikir reflektif matematis siswa Madrasah Aliyah yang valid dan reliabel. Penelitian ini merupakan sebuah penelitian pengembangan (*Research & Development*) dengan model pengembangan 4D Thiagarajan dengan tahapan *define, design, development, dan dissemination*. Data yang diperoleh pada penelitian ini merupakan data kualitatif berupa komentar atau saran perbaikan dari *expert*, dan data kuantitatif berupa skor yang diperoleh siswa pada saat uji coba instrumen tes ini. Uji coba penelitiannya dilakukan kepada 60 siswa kelas 10 MAN 12 Jakarta untuk mengetahui validitas, reliabilitas dan tingkat kesukaran instrumen. Analisis hasil uji coba dilakukan dengan pendekatan model Rasch dibantu oleh *software Winsteps*. Penelitian ini menghasilkan 5 butir soal uraian valid, reliabel, dan tingkat kesukaran yang baik. Uji validitas menggunakan tabel *item dimensionality* dan tabel *item fit order* dimana *Undimensionality* berdasarkan nilai *raw variance by measure* dan kesesuaian butir soal sesuai salah satu kriteria, yakni $0,50 < MNSQ < 1,5$; $-2,0 < ZSTD < +2,0$ dan $0,4 < Pt Mean Car < 0,85$. Uji reliabilitasnya berdasarkan tabel *summary statistic* pada nilai *person dan item reliability* sebesar 0,73 dan 0,76 yang berkriteria cukup dan nilai *alpha Cronbach* sebesar 0,92 yang berkriteria bagus sekali. Lalu tingkat kesukaran instrumen terbagi tiga tingkatan, yaitu sulit, sedang dan mudah dengan menggunakan tabel *item measure*. Berdasarkan hal tersebut, maka instrumen tes ini layak digunakan untuk mengukur kemampuan berpikir reflektif matematis siswa Madrasah Aliyah.

Kata kunci: Instrumen; kemampuan berpikir reflektif; model Rasch.

Abstract

The purpose of this study was to obtain a valid and reliable instrument of mathematical reflective thinking ability of Madrasah Aliyah students. This research is a development research (*Research & Development*) with a 4D Thiagarajan development model with the stages of *define, design, development, and dissemination*. The data obtained in this study are qualitative data in the form of comments or suggestions for improvement from experts, and quantitative data in the form of scores obtained by students when testing this test instrument. The research trial was conducted on 60 10th grade students of MAN 12 Jakarta to determine the validity, reliability and level of difficulty of the instrument. The analysis of the test results was carried out using the Rasch model approach assisted by the *Winsteps* software. This study resulted in 5 items of valid, reliable, and good level of difficulty. The validity test uses the *item dimensionality* table and the *item fit order* table where *Undimensionality* is based on the *raw variance by measure* value and the suitability of the items according to one of the criteria, namely $0.50 < MNSQ < 1.5$; $-2.0 < ZSTD < +2.0$ and $0.4 < Pt Mean Car < 0.85$. The reliability test is based on the statistical summary table on the person and item reliability values of 0.73 and 0.76 which have sufficient criteria and Cronbach's alpha value of 0.92 which has excellent criteria. Then the instrument difficulty level is divided into three levels, namely difficult, medium and easy using the *item measure* table. Based on this, this test instrument is suitable to be used to measure the mathematical reflective thinking ability of Madrasah Aliyah students.

Keywords: Instrument; Rasch model; mathematical reflective thinking skill.



This is an open access article under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

PENDAHULUAN

Dunia pendidikan tidak dapat dipisahkan dari kegiatan evaluasi. Evaluasi adalah salah satu aspek dasar pada sistem pendidikan yang musti dilaksanakan dengan terencana dan sistematis sebagai alat untuk mengukur ketercapaian target dalam proses pembelajaran (L, 2019). Evaluasi dilakukan sebagai pengendalian mutu pendidikan secara nasional sebagai bentuk akuntabilitas penyelenggaraan pendidikan kepada pihak-pihak yang berkepentingan seperti yang dijelaskan pada Bab XVI tentang evaluasi, akreditasi dan sertifikasi pasal 57 ayat 1 dalam UU No. 20 Tahun 2003 (Kemendikbud, 2016). Pada proses evaluasi digunakan suatu alat ukur yang salah satunya dinamakan instrumen tes. Tes adalah cara (yang dapat dipergunakan) atau prosedur yang (yang perlu ditempuh) dalam rangka pengukuran dan penilaian di bidang pendidikan (Magdalena, Syariah, Mahromiyati, & Nurkamillah, 2021)

Salah satu evaluasi pembelajaran yang perlu dilakukan yaitu, evaluasi pada pembelajaran matematika. Matematika merupakan mata pelajaran yang terdapat di setiap silabus, baik dari tingkat dasar hingga tingkat perguruan tinggi (Mutmainna, Mania, & Sriyanti, 2018). Matematika merupakan hal yang wajib dipelajari, karena dengan mempelajarinya siswa dapat meningkatkan kemampuan berpikir dan pemecahan masalah (Thahir, Komarudin, Hasanah, & Rahmahwaty, 2019). Salah satu kemampuan berpikir adalah kemampuan berpikir reflektif matematis. Memiliki kemampuan berpikir reflektif matematis akan membuat siswa lebih giat dan hati-hati dalam mempertimbangkan sesuatu yang diperolehnya (Muntazhimah, 2019a). Dengan berpikir reflektif siswa dapat

mengetahui apa yang diperlukannya dalam proses belajar (Pamungkas, Mentari, & Nindiasari, 2018). Individu akan berpikir terarah dalam menganalisis, mengevaluasi, memotivasi, mendapatkan makna yang mendalam dan menggunakan strategi yang tepat ketika mereka memiliki kemampuan berpikir reflektif (Gurol, 2011). Sehingga dapat disimpulkan bahwa kemampuan berpikir reflektif matematis sangatlah penting bagi siswa. Oleh sebab itu diperlukan pula instrumen yang berkualitas untuk mengukur kemampuan berpikir reflektif matematis siswa.

Penelitian terdahulu terkait pengembangan instrumen kemampuan berpikir reflektif matematis siswa masih belum banyak ditemukan. Beberapa penelitian antara lain : penelitian pengembangan bahan ajar untuk mengembangkan kemampuan berpikir reflektif matematis siswa SMA (Hendriana, Putra, & Hidayat, 2019; Nindiasari, 2011), Pengembangan model pembelajaran penemuan terbimbing dan menguji efektivitasnya pada keterampilan berpikir reflektif matematis (Andreas, Noer, & Asmiati, 2019), serta penelitian Muntazhimah, (2019) mengenai pengembangan instrumen tes berpikir reflektif matematis untuk siswa kelas 8 SMP, dan pengembangan instrumen untuk meningkatkan kemampuan berpikir reflektif matematis melalui pendekatan scientific disertai strategi what if not oleh (Putra, 2016).

Dari penelitian terdahulu diatas, referensi terkait pengembangan instrumen tes berpikir reflektif matematis siswa madrasah aliyah masih minim. Di sisi lain, penelitian pengembangan instrumen tes berpikir reflektif matematis yang sudah terlaksana masih memakai teori tes

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

klasik dalam analisisnya. Padahal ada pendekatan lain yang memiliki kelebihan dibandingkan dengan tes teori klasik dalam menganalisis instrumen yakni model Rasch. Model Rasch dianggap mampu melakukan prediksi terhadap data yang hilang (missing data), yang didasarkan pada pola respon yang sistematis (An & Yu, 2021; Nur, Nurani, Suryana, & Ahmad, 2020) dan pemodelan Rasch mampu menghasilkan nilai pengukuran standar error untuk instrumen yang digunakan yang dapat meningkatkan ketepatan perhitungan (Sumintono & Widhiarso, 2015).

Oleh sebab itu, dilakukan penelitian pengembangan ini dengan tujuan agar menghasilkan instrumen tes kemampuan berpikir reflektif matematis siswa madrasah aliyah yang bermutu yang analisis terkait validitas dan reabilitasnya menggunakan pendekatan model Rasch. Instrumen tes berpikir reflektif siswa madrasah aliyah diperoleh dari penelitian ini diharapkan bisa digunakan untuk mengukur kemampuan berpikir reflektif siswa sekolah aliyah serta penelitian selanjutnya terkait pembelajaran yang bisa memfasilitasi peningkatan kemampuan berpikir reflektif matematis siswa madrasah aliyah.

METODE PENELITIAN

Penelitian pengembangan ini menggunakan model pengembangan 4D (*Four D*) dari Thiagarajan (1974) dengan tahapan pendefinisian (*define*), perancangan (*design*), pengembangan (*development*), dan terakhir penyebaran (*disseminate*).

Pada tahap pendefinisian (*define*) yaitu menentukan kebutuhan dalam proses pembelajaran yang ditetapkan dan didefinisikan, serta mengumpulkan informasi yang berkaitan dengan produk yang dikembangkan. Pada tahap

perancangan (*design*) yaitu untuk merancang bentuk dasar dari instrumen tes kemampuan berpikir reflektif matematis siswa madrasah aliyah. Dimulai dari merancang kisi-kisi, membuat instrumen awal, membuat kunci jawaban, dan membuat pedoman penskoran.

Tahap *development* merupakan tahap lanjutan untuk menyempurnakan konsep soal sebelumnya menjadi versi akhir soal. Langkah-langkah yang dilakukan pada tahap ini yaitu, penilaian parah ahli, uji coba produk dan analisis butir soal. Pengujian instrumen oleh para ahli dilakukan untuk mengetahui kelayakan dari instrumen tes yang dikembangkan. Instrumen ini di tinjau dari tiga aspek yaitu konten, konstruk dan bahasa. Untuk mengetahui dan menghitung penilaian validasi dari para ahli dapat dilihat pada tabel 1, penilaian menggunakan skala Likert dengan 4 kategori (sangat tidak sesuai, kurang sesuai, sesuai, dan sangat sesuai). Mengenai rumus yang digunakan dalam perhitungan kuantitas yaitu (Sari, Noer, & Asmiati, 2020):

$$P = \frac{\sum x_i}{\sum x} \times 100\% \quad (1)$$

Keterangan:

P : persentase penilaian

$\sum x_i$: total nilai yang di peroleh

$\sum x$: total nilai seluruhnya

Tabel 1 Kriteria persentase nilai ahli

Persentase (%)	Kriteria
76-100	Valid
56-75	Cukup Valid
40-55	Kurang Valid
0-39	Tidak Valid

(Sari et al., 2020)

Instrumen layak digunakan dengan minimal persentase penilaian 56%. Setelah instrumen dinyatakan valid

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

oleh ahli, instrumen dilakukan uji coba produk ke siswa. Uji coba skala kecil kepada 30 siswa untuk mengetahui keterbacaan soal dan uji coba skala besar 60 siswa untuk mengetahui validitas dan reliabilitas instrumen serta tingkat kesukarannya dengan melakukan analisis menggunakan *Rasch model* bantuan *software Winsteps*.

Tahap akhir yang dilakukan adalah penyebaran (*disseminate*). Tentunya tahapan ini dilakukan setelah dilakukan revisi atas masukan para ahli, siswa serta didasarkan juga oleh hasil analisis uji coba. Tahap ini bertujuan untuk menyebarkan instrumen yang telah berkualitas. Penyebaran dilakukan secara terbatas kepada guru bidang studi matematika kelas 10 MAN 12 Jakarta.

HASIL DAN PEMBAHASAN

Hasil Penelitian akan dibahas sesuai dengan tahapan pengembangan instrumen yang dilakukan pada penelitian ini yang terdiri dari empat tahapan yaitu 1) *Define* (Pendefinisian), 2) *Design* (Perancangan), 3) *Development* (Pengembangan), 4) *Disseminate* (Penyebaran).

Tahap *Define* (Pendefinisian)

Pada analisis awal dilakukan dengan studi pendahuluan dengan wawancara pada guru bidang studi matematika kelas 10 di MAN 12 Jakarta. Berdasarkan hasil wawancara dengan guru matematika kelas 10 bahwa kurikulum yang digunakan merupakan kurikulum 2013 revisi dengan metode pembelajaran yang digunakan adalah metode ceramah. Adapun buku pegangan yang digunakan oleh guru dalam proses pembelajaran yaitu buku matematika wajib kelas 10 SMA/MA kurikulum 2013 revisi penerbit erlangga dan buku PKS matematika wajib kelas 10 SMA/MA

kurikulum 2013 revisi sebagai salah satu sumber untuk memberikan materi dan latihan soal. Selain itu, diperoleh pula informasi bahwa kemampuan matematika yang dimiliki oleh siswa kelas 10 berbeda-beda. Baik pada siswa kelas 10 MIPA, 10 IPS maupun 10 Agama. Dimana kemampuan siswa pada setiap kelasnya terdapat kemampuan tinggi, sedang dan rendah. Namun untuk kemampuan berpikir reflektif matematis belum pernah diukur sebelumnya karena kemampuan matematis ini belum secara serius terfasilitasi melalui pembelajaran matematika, sehingga instrumen tes berpikir reflektif matematis siswa madrasah aliyah dirasa penting untuk dikembangkan dalam sebuah penelitian.

Selanjutnya dilakukan analisis konsep dengan cara membaca berbagai macam referensi mengenai kemampuan berpikir reflektif matematis melalui beberapa artikel, jurnal dan membaca buku matematika kelas 10 SMA/MA/SMK/MAK kurikulum 2013 edisi revisi 2017 mengenai materi fungsi. Berdasarkan hal tersebut dibuatlah instrumen kemampuan berpikir reflektif matematis dengan materi fungsi pada bagian fungsi komposisi dan fungsi invers.

Pada analisis tugas dilakukan dengan mengidentifikasi Kompetensi Inti (KI), Kompetensi Dasar (KD) dan kegiatan pembelajaran mengenai materi fungsi pada bagian fungsi komposisi dan fungsi invers (KD 3.6 dan KD 3.9). Berdasarkan Kompetensi Inti (KI) dan Kompetensi Dasar (KD) dibuatlah instrumen untuk mengukur kemampuan berpikir reflektif matematis siswa dengan menggunakan materi fungsi komposisi dan fungsi invers.

Pada analisis tujuan pembelajaran yaitu untuk memodifikasi dari analisis konsep dan analisis tugas menjadi

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

tujuan pembelajaran. Adapun tujuan pembelajaran materi fungsi pada fungsi komposisi dan fungsi invers yaitu yang didapat yaitu menunjukkan rasa ingin tahu dalam memahami dan menyelesaikan masalah fungsi, menentukan hasil operasi aritmatika pada fungsi komposisi dan invers, menentukan invers suatu fungsi, dan menyelesaikan masalah kontekstual yang berkaitan dengan materi fungsi. Selanjutnya definisi kemampuan berpikir reflektif matematis yang dipakai pada penelitian ini adalah kemampuan dimana siswa secara giat, sungguh-sungguh dan dengan penuh kehati-hatian untuk meninjau sesuatu dengan pengetahuan yang telah diperolehnya terhadap persoalan yang diberikan. Dengan indikator menurut yaitu 1. *Reacting* (berpikir reflektif untuk aksi), 2. *Comparing* (berpikir reflektif untuk evaluasi), 3. *Contemplating* (berpikir reflektif untuk inquiri kritis).

Tahap Design (Perancangan)

Tujuan pembelajaran telah terdefiniskan dengan baik, begitu pula definisi kemampuan berpikir reflektif matematis beserta indikatornya, juga telah terdefiniskan dengan jelas. Tahap selanjutnya adalah perancangan. Tahapan ini dilakukan dengan membuat butir-butir soal tes kemampuan berpikir reflektif matematis yang diturunkan dari indikator yang telah ditetapkan. Butir soal yang dibuat berkisar pada materi fungsi komposisi dan fungsi invers. Perancangan butir-butir soal ini disusun dalam suatu kisi-kisi soal dan tersusunlah lima butir soal yang sesuai dengan kisi-kisi yang ada.

Selanjutnya dirancang pula alternatif penyelesaian soal dan pedoman penskoran. Sejalan dengan pedoman penskoran di susun pula

lembar penilaian yang digunakan untuk meminta penilaian dari para ahli berdasarkan aspek konten, konstruk dan bahasa untuk mengetahui kualitas instrumen yang telah dirancang.

Tahap Development (Pengembangan)

Pada tahapan ini terdiri dari penilaian para pakar atau ahli, revisi berdasarkan masukan para ahli, uji coba produk dan melakukan analisis butir soal. Penilaian para ahli dilakukan oleh 2 dosen pendidikan matematika dan 1 guru bidang studi matematika. Setelah direvisi berdasarkan saran dari para ahli, selanjutnya dilakukan rekapitulasi dari penilaian para ahli yang hasilnya dapat dilihat pada Tabel 2.

Tabel 2 Rekapitulasi penilaian para ahli

Para Ahli	Persentase Penilaian	Kriteria
Ahli 1	96,43%	Valid
Ahli 2	67,86%	Cukup Valid
Ahli 3	92,86%	Valid

Berdasarkan tabel 2 bahwa instrumen tes kemampuan berpikir reflektif matematis layak diuji cobakan. Selanjutnya dilakukan uji coba produk secara dua tahap, tahap pertama dengan melakukan uji coba skala kecil dilakukan dengan melibatkan 30 siswa untuk mengetahui keterbacaan soal. Siswa mengerjakan soal dengan waktu yang telah ditentukan yaitu 90 menit dan siswa diminta untuk mengerjakan dengan fokus, teliti dan sungguh-sungguh.

Setelah uji coba skala kecil dilakukan, maka secara bahasa butir soal yang disusun ternyata tidak terdapat masalah yang berarti dan dapat dilanjutkan dengan melakukan uji coba skala besar. Uji coba skala besar dilakukan kepada 60 siswa kelas 10. Pada uji skala besar dilakukan uji

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

validitas, reliabilitas, dan tingkat kesukaran pada soal. Analisis butir soal pada tahapan ini dilakukan dengan *Rasch model* bantuan *software winsteps*.

Validitas merupakan suatu bentuk yang menunjukkan tingkat kevalidan atau kesahihan suatu instrumen. Suatu instrumen dapat dikatakan valid jika mampu mengukur apa yang ingin diukur (Sundayana, 2018). Validitas instrumen tes ini dianalisis dengan pendekatan model Rasch. Output dari *software winstep* yang digunakan adalah

tabel *item dimensionality* dan tabel *item fit order*.

Output tabel *dimensionality* digunakan untuk mengetahui apakah instrumen yang digunakan dapat mengukur apa yang seharusnya diukur yang dilihat berdasarkan nilai pada *kolom raw variance explained by measures* dengan kriteria 20% cukup, >40% bagus dan >60% istimewa (Sumintono & Widhiarso, 2014). Hasil *output* tersebut dapat dilihat pada Gambar 1.

TABLE 23.0 Reflektif Matematis ZOU688WS.TXT Sep 8 19:16 2021
INPUT: 60 Person 5 Item REPORTED: 60 Person 5 Item 10 CATS WINSTEPS 3.73

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		-- Empirical --	Modeled
Total raw variance in observations	=	10.3 100.0%	100.0%
Raw variance explained by measures	=	5.3 51.5%	52.4%
Raw variance explained by persons	=	1.9 18.5%	18.8%
Raw Variance explained by items	=	3.4 33.0%	33.6%
Raw unexplained variance (total)	=	5.0 48.5%	47.6%
Unexplned variance in 1st contrast	=	1.8 17.6%	36.3%
Unexplned variance in 2nd contrast	=	1.6 15.4%	31.7%
Unexplned variance in 3rd contrast	=	1.0 10.0%	20.7%
Unexplned variance in 4th contrast	=	.6 5.5%	11.3%
Unexplned variance in 5th contrast	=	.0 .0%	.1%

Gambar 1 Output Tabel Item Dimensionality

Berdasarkan gambar 1 *output dimensionality*, diperoleh bahwa nilai *kolom observed raw variance explained by measures* yaitu 51,5%. Berdasarkan kriteria nilai *raw variance explained by measures* maka instrumen tes ini berkriteria bagus.

Setelah tabel *dimensionality*, dilanjutkan dengan tabel *item fit order*. Tabel *item fit order* digunakan untuk mengetahui kesesuaian butir soal, dimana dapat diketahui apakah soal

sudah sesuai. Kesesuaian butir soal dapat dilihat dengan menggunakan nilai *Outfit Mean Square (MNSQ)*, *Outfit Z-Standard (ZSTD)*, dan *Point Measure Correlation (Pt Mean Car)*, dimana jika memenuhi salah satu dari kriteria dapat dikatakan valid (Sumintono & Widhiarso, 2015)

Output tabel *item fit order* dapat dilihat pada gambar 2.

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

TABLE 10.1 Reflektif Matematis ZOU688WS.TXT Sep 8 19:16 2021
INPUT: 60 Person 5 Item REPORTED: 60 Person 5 Item 10 CATS WINSTEPS 3.73

Person: REAL SEP.: 1.66 REL.: .73 ... Item: REAL SEP.: 1.76 REL.: .76

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	EXACT MATCH OBS	EXACT MATCH EXP	Item	
3	439	58	-.33	.10	1.59	2.8	1.54	2.6	A .72	69	19.0	24.9	E3
4	381	57	.16	.10	1.17	.9	1.18	.9	B .72	70	33.3	31.2	E4
1	406	59	.04	.09	.99	.0	1.03	.2	C .69	69	28.8	28.5	E1
2	427	60	-.14	.10	.80	-1.1	.78	-1.2	b .75	73	23.3	27.2	E2
5	370	57	.26	.10	.61	-2.2	.56	-2.4	a .66	71	42.1	32.3	E5
MEAN	404.6	58.2	.00	.10	1.03	.1	1.02	.0			29.3	28.8	
S.D.	26.2	1.2	.21	.00	.34	1.7	.34	1.7			8.0	2.7	

Gambar 2 Output Item Fit Order

Berdasarkan rekapitulasi hasil pada tabel *item fit order* pada gambar 2 bahwa semua butir soal memenuhi salah satu dari 3 kriteria *item fit order* berikut $0,50 < MNSQ < 1,5$; $-2,0 < ZSTD < +2,0$ dan $0,4 < Pt Mean Car < 0,85$ (Sumintono & Widhiarso, 2014). Sehingga semua butir soal dapat digunakan. Berdasarkan hasil dari output tabel *item dimensionality* dan tabel *item fit order* diperoleh hasil bahwa instrumen tes ini valid.

Langkah selanjutnya adalah menguji reliabilitas instrumen. Uji reliabilitas dilakukan guna mendapatkan hasil instrumen yang sama (konsisten, ajeg). Reliabilitas instrumen dianalisis dengan melihat nilai *item reliability* dan *person reliability* pada *table summary statistics*. Sedangkan *alpha Cronbach* pada *table summary statistics* untuk melihat reliabilitas interaksi *item* dan *person* (Sumintono & Widhiarso, 2015). Selanjutnya, *output tabel item summary* dapat dilihat pada Gambar 3 dan 4.

TABLE 3.1 Reflektif Matematis ZOU688WS.TXT Sep 8 19:16 2021
INPUT: 60 Person 5 Item REPORTED: 60 Person 5 Item 10 CATS WINSTEPS 3.73

SUMMARY OF 60 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	33.7	4.9	.57	.34	.99	.0	1.01	.0
S.D.	8.4	.6	.74	.04	.67	1.1	.73	1.1
MAX.	45.0	5.0	1.78	.50	3.09	2.4	3.67	2.5
MIN.	5.0	2.0	-1.85	.25	.07	-2.2	.07	-2.2
REAL RMSE	.38	TRUE SD	.63	SEPARATION	1.66	Person RELIABILITY	.73	
MODEL RMSE	.34	TRUE SD	.66	SEPARATION	1.92	Person RELIABILITY	.79	
S.E. OF Person MEAN	= .10							

Gambar 3. Output tabel item summary statistics (bagian 1)

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

TABLE 3.1 Reflektif Matematis ZOU688WS.TXT Sep 8 19:16 2021
 INPUT: 60 Person 5 Item REPORTED: 60 Person 5 Item 10 CATS WINSTEPS 3.73

SUMMARY OF 60 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	33.7	4.9	.57	.34	.99	.0	1.01	.0
S.D.	8.4	.6	.74	.04	.67	1.1	.73	1.1
MAX.	45.0	5.0	1.78	.50	3.09	2.4	3.67	2.5
MIN.	5.0	2.0	-1.85	.25	.07	-2.2	.07	-2.2
REAL RMSE	.38	TRUE SD	.63	SEPARATION	1.66	Person RELIABILITY	.73	
MODEL RMSE	.34	TRUE SD	.66	SEPARATION	1.92	Person RELIABILITY	.79	
S.E. OF Person MEAN = .10								

Gambar 4 Output tabel item summary statistics (bagian 2)

Berdasarkan Gambar 3, dapat dilihat kekonsistenan dari instrumen (*item*) dan responden (*person*) sekaligus interaksi antara instrumen (*item*) dan responden (*person*). Kriteria *item reliability & person reliability* dapat dilihat pada tabel 3 dan kriteria *alpha Cronbach* dapat dilihat pada tabel 4 (Sumintono & Widhiarso, 2015).

Tabel 2 *Item reliability & person reliability*

Item Reliability & Person Reliability	Kriteria
< 0,67	Lemah
< 0,67 – 0,80	Cukup
0,81 – 0,90	Bagus
0,91 – 0,94	Bagus Sekali
> 0,94	Istimewa

Tabel 3 Kriteria *Alpha Cronbach*

Alpha Cronbach	Kriteria
< 0,5	Buruk
0,5 – 0,6	Jelek
0,6 – 0,7	Cukup
0,7 – 0,8	Bagus
> 0,8	Bagus Sekali

(Sumintono & Widhiarso, 2015)

Rekapitulasi dari tabel *summary statistics* memperlihatkan bahwa *item reliability* dan *person reliability* berkriteria cukup dan hubungan antar keduanya berkriteria bagus sekali.

Analisis dilanjutkan dengan mengkaji tingkat kesukaran butir soal. Tingkat kesukaran adalah keadaan suatu butir soal dipandang sulit, sedang atau mudah dalam mengerjakannya (Sundayana, 2018). Tingkat kesukaran pada output *software Winsteps* ditunjukkan pada tabel *item measure*. Pada tabel *item measure*, *item* logit pada batas bawah merupakan kelompok butir soal mudah. Sehingga kelompok butir soal sulit terdapat pada batas atas dan kelompok butir sedang terdapat diantara batas atas dan batas bawah. Tingkat kesukaran juga dapat dilihat pada nilai *measure* dalam tabel *item measure* semakin tinggi nilainya maka semakin sulit pula soalnya (Sumintono & Widhiarso, 2015). Output nilai *measure* dapat dilihat pada Gambar 4.

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

TABLE 13.1 Reflektif Matematis ZOU688WS.TXT Sep 8 19:16 2021
INPUT: 60 Person 5 Item REPORTED: 60 Person 5 Item 10 CATS WINSTEPS 3.73
Person: REAL SEP.: 1.66 REL.: .73 ... Item: REAL SEP.: 1.76 REL.: .76

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	Item
5	370	57	.26	.10	.61	-2.2	.56	-2.4	.66	.71	42.1	32.3	E5
4	381	57	.16	.10	1.17	.9	1.18	.9	.72	.70	33.3	31.2	E4
1	406	59	.04	.09	.99	.0	1.03	.2	.69	.69	28.8	28.5	E1
2	427	60	-.14	.10	.80	-1.1	.78	-1.2	.75	.73	23.3	27.2	E2
3	439	58	-.33	.10	1.59	2.8	1.54	2.6	.72	.69	19.0	24.9	E3
MEAN	404.6	58.2	.00	.10	1.03	.1	1.02	.0			29.3	28.8	
S.D.	26.2	1.2	.21	.00	.34	1.7	.34	1.7			8.0	2.7	

Gambar 3 output tabel item measure

Berdasarkan Gambar 4, tingkat kesukaran terbagi menjadi 3 yaitu sulit, sedang dan mudah. Soal sulit pada *item* 5 dan *item* 4, soal sedang pada *item* 1 dan soal mudah pada *item* 2 dan *item* 3. Pada tahapan ini sudah dapat diperoleh hasil kajian bahwa instrumen tes yang dirancang telah memenuhi uji validitas dan reliabilitas. Instrumen dinyatakan reliabel dan valid serta mengandung butir soal dengan tingkat kesukaran mudah, sedang dan sulit sehingga layak untuk disebarkan pada tahapan *disseminate*.

Tahap *Disseminate* (Penyebaran)

Tahap terakhir ini bertujuan untuk menyebarluaskan instrumen yang berkualitas. Sehingga pada tahapan ini dilakukan dengan menyebarluaskan instrumen yang dikembangkan dan sudah berkualitas. Penyebaran dilakukan dengan memberikan instrumen tes yang telah dikembangkan kepada guru bidang studi matematika kelas 10 MAN 12 Jakarta. Sehingga guru dapat menggunakan instrumen untuk mengetahui kemampuan berpikir reflektif matematis siswa pada materi fungsi.

Berdasarkan hasil dari tahap *define, design, development*, penilaian

para ahli dan uji coba lapangan bahwa instrumen kemampuan berpikir reflektif matematis siswa madrasah aliyah yang dikembangkan ini sudah valid dan reliabel serta memiliki tingkat kesukaran yang baik sehingga layak untuk disebarkan pada tahapan akhir pengembangan instrumen.

Validitas uji coba instrumen kemampuan berpikir reflektif menghasilkan 5 soal yang valid. Reliabilitas instrumen ini berkriteria bagus sekali. Uji coba tingkat kesukaran menunjukkan 2 soal berkriteria sulit, 1 soal berkriteria sedang dan 2 soal berkriteria mudah. Berdasarkan tingkat kesukaran bahwa instrumen yang dibuat berkriteria baik karena dapat membagi tiga macam kriteria yaitu sulit, sedang dan mudah.

Kelebihan penelitian pengembangan ini yaitu keseluruhan langkah pengembangan yang direncanakan telah dilaksanakan, tes yang diperoleh telah melewati proses penilaian dari para ahli dan diujicobakan kepada siswa pada skala kecil berjumlah 30 siswa dan skala besar berjumlah 60 siswa. Kelebihan lainnya yaitu instrumen ini dianalisis dengan menggunakan pendekatan model Rasch hingga diperoleh tes

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

kemampuan berpikir reflektif matematis siswa madrasah aliyah yang valid, reliabel serta memiliki tingkat kesukaran yang beragam. Disisi lain, Kekurangannya yaitu instrumen yang dihasilkan dibatasi oleh materi fungsi komposisi dan fungsi invers.

Penelitian ini memiliki kesamaan dengan penelitian yang dilakukan antara lain oleh Muntazhimah, (2019) mengenai pengembangan instrumen tes berpikir reflektif matematis untuk siswa kelas 8 SMP, dan pengembangan instrumen untuk meningkatkan kemampuan berpikir reflektif matematis melalui pendekatan scientific disertai strategi what if not oleh (Putra, 2016). Juga penelitian (Nindiasari, 2011) tentang pengembangan bahan ajar dan Instrumen untuk meningkatkan berpikir reflektif matematis berbasis pendekatan metakognitif pada siswa sekolah menengah Atas (SMA), (Hendriana et al., 2019) yang juga mengembangkan instrumen kemampuan berpikir reflektif siswa SMA. Beberapa penelitian yang sudah disebutkan, masih sangat sulit mencari referensi penelitian terdahulu tentang pengembangan tes berpikir reflektif matematis siswa madrasah aliyah dan keseluruhan penelitian pengembangan instrumen tes yang sudah dilaksanakan masih memakai teori tes klasik. Oleh sebab itu, penelitian ini bertujuan untuk mengembangkan instrumen tes kemampuan berpikir reflektif siswa madrasah aliyah yang reliabel dan divalidasi menggunakan model Rasch.

Implikasi hasil penelitian ini secara teoritis dapat menambah teori ataupun literatur terkait bagaimana mengembangkan instrumen tes yang dapat mengukur kemampuan berpikir reflektif siswa madrasah aliyah. Secara praktis, tes yang dihasilkan dari penelitian ini dapat digunakan untuk

mengukur kemampuan berpikir reflektif siswa calon guru matematika. Dalam pembelajaran matematika, tidak hanya kemampuan berpikir reflektif saja yang perlu diukur. Ada banyak kemampuan berpikir matematis dan perlu pula dilakukan kajian mengenai kemampuan berpikir matematis lainnya. Oleh karena itu, pengembangan instrumen tes berpikir matematis lainnya selanjutnya dapat dilakukan.

KESIMPULAN DAN SARAN

Berdasarkan hasil analisis pada penelitian pengembangan instrumen kemampuan berpikir reflektif matematis siswa madrasah aliyah, dapat disimpulkan bahwa instrumen yang dikembangkan dengan model 4D Thiagarajan (1974) layak digunakan karena telah melewati serangkaian tahapan pengembangan instrumen dan sudah memenuhi unsur validitas dari para ahli, hasil validitas dan reliabilitas dari uji coba produk ke siswa yang dianalisis dengan pendekatan model Rasch.

Berdasarkan hasil penelitian dan pengembangan instrumen kemampuan berpikir reflektif dengan menggunakan *rasch* model, instrumen dapat dilakukan uji coba dengan skala yang lebih besar lagi dan pengembangan instrumen kemampuan berpikir reflektif dapat dikembangkan lebih lanjut dengan menggunakan materi yang berbeda.

DAFTAR PUSTAKA

- 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).
- Andrean, N. J., Noer, S. H., & Asmiati. (2019). Pengembangan Pembelajaran Penemuan

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

- Terbimbing Untuk Meningkatkan Kemampuan Berpikir Reflektif dan Kemandirian Belajar Siswa. *Aksioma: Jurnal Program Studi Pendidikan Matematika*, 8(2), 270–278.
- 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>
- Kemendikbud. (2016). Standar Kompetensi Lulusan Pendidikan Dasar dan Menengah. *Lampiran Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 20 Tahun 2016*.
- L, I. (2019). Evaluasi Dalam Proses Pembelajaran. *Manajemen Pendidikan Isam*, 9(2), 920–935.
- 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.
- Muntazhimah. (2019a). 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).
- 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>
- 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>
- 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).

DOI: <https://doi.org/10.24127/ajpm.v10i4.4223>

- 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.
- 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).