

PENGEMBANGAN INSTRUMEN *HIGHER ORDER THINKING SKILL (HOTS)* MATEMATIKA

N.K.L.A. Rahayu¹, N.K. Widiartini², D.G.H. Divayana³

¹²³Program Studi Penelitian dan Evaluasi Pendidikan
Universitas Pendidikan Ganesha
Singaraja, Indonesia

e-mail: lindaastiti4@gmail.com¹, ketut.widiartini@undiksha.ac.id²,
hendra.divayana@undiksha.ac.id³

Abstrak

Penelitian ini berfokus dan bertujuan guna untuk menghasilkan sebuah produk yakni kisi-kisi instrumen dan sebuah instrumen HOTS matematika pada kelas IV SD. Penelitian ini merupakan penelitian pengembangan yang menggunakan model formatif evaluasi. Penelitian ini dilakukan di SD Gugus V Sukawati yang menggunakan 30 responden sebagai small group dan 60 responden pada field test. kisi-kisi instrumen berbasis HOTS Matematika kelas IV SD memuat satu Kompetensi Dasar (KD) dan 4 pengembangan indikator yang dengan 1 indikator termasuk C4/K2, 2 indikator termasuk C5/K3 serta 1 indikator termasuk C6/K4. Kemudian untuk Instrumen yang dikembangkan pada penelitian ini adalah instrumen dengan jenis soal uraian yang dalam menentukan kalitas instrumen tersebut harus melalui beberapa analisis. Berdasarkan hasil uji pakar diperoleh nilai CVR = 1 maka dari itu instrumen ini memiliki validitas isi yang tinggi sementara itu pada uji one to one diperoleh nilai kepraktisan instrumen yakni 91% yang berkategori sangat praktis. Sementara itu kualitas tes instrumen matematika berbasis HOTS jika dilihat dari hasil uji validitas dan reliabilitas menunjukkan bahwa kesepuluh instrumen tersebut berkategori valid serta memiliki tingkat reliabilitas yang tinggi yakni 0,79. Sementara itu kualitas tes jika dilihat dari pengujian uji daya beda menghasilkan 7 tes berkategori baik dan 3 tes berkategori cukup. Dan kualitas tes melalui uji tingkat kesukaran mendapatkan hasil bahwa 3 soal berkategori mudah, 6 soal berkategori sedang dan 1 soal berkategori sulit serta pada analisis partial credit model diperoleh output yakni tahap 1 memiliki tingkat kesulitan tahap sebesar - 0,47 (tinggi) yang artinya individu dengan kemampuan dibawah 0,47 memiliki probabilitas untuk mendapatkan skor 0 sebaliknya jika diatas 0,47 memiliki probabilitas mendapat skor 1, kemudian pada tahap 2 memiliki tingkat kesulitan tahap sebesar 0,08 (tinggi) yang artinya untuk individu yang memiliki kemampuan dibawah 0,08 memiliki probabilitas untuk mendapatkan skor 1 sebaliknya jika diatas 0,08 maka memiliki probabilitas mendapatkan skor 2. Sementara itu untuk tahap 3 memiliki tingkat kesulitan tahap sebesar 0,39 (sedang) yang artinya individu dengan kemampuan dibawah 0,39 maka memiliki probabilitas untuk mendapatkan skor 2 tetapi jika diatas 0,39 maka memiliki probabilitas mendapatkan nilai maksimal yakni 3

Kata kunci: *Higher Order Thinking Skill*; Penelitian Pengembangan

Abstract

This research focuses and aims to produce a product, namely an instrument grid and a mathematical HOTS instrument in the fourth grade of elementary school. This research is a development research that uses a formative evaluation model. This research was conducted at SD Cluster V Sukawati using 30 respondents as a small group and 60 respondents in the field test. The HOTS Maths Grade IV Elementary School-based instrument grid contains one Basic Competence (KD) and 4 development indicators with 1 indicator including C4/K2, 2 indicators including C5/K3 and 1 indicator including C6/K4. Then for the instrument developed in this study, it is an instrument with a description of the type of question which in determining the quality of the instrument must go through several analyzes. Based on the results of the expert test, the value of CVR = 1, therefore this instrument has high content validity, meanwhile in the one to one test, the practicality value of the instrument is 91% which is categorized as very practical. Meanwhile, the quality of the HOTS-based mathematical instrument test when viewed from the results The validity and reliability test shows that the ten instruments are categorized as valid and have a high level of

reliability, namely 0.79. Meanwhile, the quality of the test when viewed from the test of differentiating power resulted in 7 tests in the good category and 3 tests in the sufficient category. And the quality of the test through the level of difficulty test got the results that 3 questions were categorized as easy, 6 questions were categorized as medium and 1 question was categorized as difficult and in the partial credit model analysis the output was obtained, namely stage 1 had a stage difficulty level of - 0.47 (high) which means individual with abilities below 0.47 have a probability of getting a score of 0 otherwise if above 0.47 has a probability of getting a score of 1, then at stage 2 has a stage difficulty level of 0.08 (high) which means for individuals who have abilities below 0.08 has a probability of getting a score of 1 otherwise if it is above 0.08 then it has a probability of getting a score of 2. Meanwhile for stage 3 it has a stage difficulty level of 0.39 (medium) which means that individuals with abilities below 0.39 have a probability of getting a score of 2 but if it is above 0.39 then it has a probability of getting a maximum value of 3.

Keywords: Higher Order Thinking Skill; Development Research

PENDAHULUAN

Pendidikan di jenjang sekolah dasar yakni pendidikan paling penting dan berguna pada perkembangan diri seorang peserta didik. Hal ini disebabkan karena SD merupakan sumber pendidikan dasar untuk seorang anak dalam rangka mendapatkan ilmu pengetahuan, setelah anak dididik oleh kedua orang tua dan setelah meninggalkan taman kanak-kanak. Pada tahap pendidikan formal, SD sebagai pondasi awal (dasar) seorang siswa akan melanjutkan pendidikannya. Hal ini sesuai dengan Undang-Undang RI Nomor 20 Tahun 2003 pasal 17 ayat 1 bahwa "pendidikan dasar merupakan pendidikan awal yang melandasi jenjang pendidikan menengah". Pendidikan di Indonesia berupaya untuk menciptakan bangsa yang cakap, beriman Bertaqwa kepada Tuhan serta Memiliki Pengetahuan yang baik dan wawasan kebangsaan (Sujana, 2019). Oleh karena itu diperlukan suatu penilaian atau asesmen berskala nasional dan lokal sebagai alat ukur ketercapaian tujuan pendidikan tersebut (Purwati, et al., 2021). Berkaitan dengan pentingnya pelaksanaan evaluasi ataupun asesmen dalam pendidikan, saat ini Menteri Pendidikan di Indonesia telah mencetuskan ide dalam pendidikan nasional di Indonesia, yaitu dengan meniadakan ujian nasional (UN) dan digantikan dengan pelaksanaan asesmen nasional berbasis komputer (ANBK) (Muflikhah et al., 2021). ANBK (Asesmen Nasional Berbasis Komputer)

menjadi salah satu diantara 4 kebijakan program merdeka belajar.

Dalam rangka pengukuran kompetensi berpikir serta bernalar pesera didik ketika membaca teks (literasi) maupun dalam menghadapi persoalan atau masalah yang memerlukan pengetahuan matematika (numerasi). ANBK pada jenjang sekolah dasar diikuti oleh siswa yang berada pada kelas V. Dalam pelaksanaan ANBK ada 3 jenis materi yang disajikan yakni literasi, numerasi serta survei karakter. ANBK merupakan inovasi baru pemerintah indonesia untuk meningkatkan kemampuan literasi dan numerasi pelajar indonesia serta bertujuan agar pendidikan di Indonesia menjadi semakin baik lagi serta bisa berkompetisi secara internasional (Muflikhah et al., 2021).

Namun dalam pelaksanaan ANBK yang telah berlangsung, ada beberapa permasalahan yang terjadi di lingkungan sekolah. Berdasarkan hasil observasi yang telah dilakukan sebelumnya diperoleh informasi bahwa beberapa sekolah di SD N Gugus V Sukawati mengalami beberapa kendala terkait dengan pelaksanaan ANBK. Kendala tersebut diantaranya adalah masih adanya beberapa peserta didik yang harus dibimbing dalam pelaksanaan ANBK terutama pada materi numerasi (matematika). Hal ini menandakan bahwa pemahaman beberapa siswa terkait soal-soal numerasi masih awam. Hasil wawancara dengan beberapa guru juga menyatakan bahwa para guru

belum sepenuhnya memberikan soal-soal latihan numerasi (matematika) dengan jenis soal *HOTS*.

Sejalan dengan permasalahan yang ditemukan peneliti saat observasi, ada beberapa hasil penelitian juga menemukan permasalahan terkait pelaksanaan ANBK seperti pada temuan penelitian dari (Purwati, et al., 2021) bahwa terjadi salah pemahaman bahwasannya ANBK dianggap hanya tanggung jawab guru wali kelas 5 saja padahal ANBK merupakan tanggung jawab semua guru khususnya guru-guru yang menjadi wali kelas 1,2,3,4 karna pada jenjang kelas inilah harusnya anak-anak terus dilatih dengan soal-soal yang terkait ANBK baik literasi maupun numerasi. Disamping itu ada juga beberapa temuan lain dari penelitian Purwati, et al., (2021) serta penelitian Novita, (2021) yakni sama-sama menemukan masalah terkait masih rendahnya pemahaman calon gurumaupun guru-guru terkait pelaksanaan ANBK. Hingga ada salah satu penelitian yang mendapatkan hasil atau temuan yakni harus ada peningkatan di bidang kompetensi seorang pendidik serta tenaga kependidikan dalam penyusunan butir soal-soal yang berorientasi dan mengacu pada literasi dan numerasi di setiap mata pelajaran (Ahmad, 2022).

Mengacu pada materi numerasi (matematika) ANBK yang soal-soalnya berkiblat pada soal-soal *Programme for Internaional Student Assesment (PISA)* dengan *Trend in International Matematicsand Science Study (TIMSS)* (Anas et al., 2021). Oleh karena itu pemberian soal-soal latihan atau ulangan harian di sekolah harusnya di susun atau dimodifikasi menjadi soal-soal yang membutuhkan keterampilan bernalar serta berpikir tingkat tinggi.

Namun beberapa hasil penelitian menemukan bahwasannya beberapa guru belum mampu dalam rangka mengembangkan sebuah instrumen tes berupa soal dalam rangka mengukur *HOTS* serta belum adanya panduan terkait dengan penyusunan soal untuk menyusun *HOTS* (Suhady et al., 2020)

hasil kajian selanjutnya oleh (Retnawati et al., 2018) menunjukkan bahwa pengetahuan beberapa guru terkait dengan *HOTS*, kemampuan dalam rangka meningkatkan *HOTS* peserta didik, menyelesaikan masalah berorientasi *HOTS* dan kegiatan dalam rangka mengukur kemampuan *HOTS* peserta didik masih cukup rendah. Temuan yang hampir mirip juga diperoleh (Driana, 2019) yakni pendidik di sekolah dasar yang turut serta dalam berpartisipasi di penelitiannya masih belum memiliki beberapa pemahaman yang cukup komprehensif terkait instrumen *HOTS*. Padahal *HOTS* ini sangat penting dilatih pada peserta didik karena melalui soal-soal *HOTS* peserta didik akan terbiasa dalam rangka mengasah kemampuan penalaran serta berpikir tingkat tinggi.

Berpikir tingkat tinggi yakni salah satu kemampuan dalam aspek kognitif yang saat ini merupakan poin penting di kurikulum 2013 (Masitoh & Aedi, 2020). Hal ini hampir sama dengan beberapa hal yang disampaikan oleh menteri pendidikan yakni satu diantara banyak kemampuan, kemampuan berpikir tingkat tinggi merupakan salah satunya yang penting (Kurniasi & Arsisari, 2020). Berpikir tingkat tinggi ini juga menjadi satu dari banyak tujuan dari pembelajaran di matematika berdasarkan kurikulum 2013 pada poin pertama yakni meningkatkan kemampuan intelektual, khususnya kemampuan tingkat tinggi siswa. Begitu pentingnya poses berpikir tingkat tinggi maka ada salah satu temuan yang memperoleh hasil bahwasannya terdapat korelasi yang positif antara keterampilan berpikir tingkat tinggi peserta didik dengan prestasi belajar peserta didik (Tanujaya et al., 2017). Menurut (Kurniasi & Arsisari, 2020) kemampuan serta ketrampilan terkait berpikir tingkat tinggi peserta didik bisa mengalami suatu perkembangan dengan dilakukan suatu pembiasaan dan latihan. Latihan atau pembiasaan ini bisa dilakukan dengan pemberian evaluasi kepada siswa secara berkala dengan jenis soal *HOTS* karena *HOTS* yakni sebuah instrumen

pengukuran yang dipergunakan dalam rangka mengukur kemampuan berpikir tingkat tinggi yakni kemampuan yang pada dasarnya sekedar mengingat serta mengungkapkan kembali ataupun merujuk tanpa dilakukannya analisis (Lestari, 2019). Hal ini sejalan dengan pendapat (Safi'i & Amar, 2019) *HOTS* yakni tujuan dalam pendidikan yang merupakan salah satu poin teratas yang bisa meningkatkan prestasi belajar peserta didik (Pratama & Retnawati, 2018). Pemilihan serta penggunaan beberapa model dalam pembelajaran yang tepat serta butir soal latihan *HOTS* bisa menunjang kemampuan *HOTS* peserta didik (Arifin & Retnawati, 2017). Mengacu pada hal itu maka sudah seharusnya seorang pendidik dirasa perlu untuk mempersiapkan instrumen yang sesuai supaya dapat menyiapkan peserta didik untuk bisa berpikir tingkat tinggi berupa butir soal-soal yang digunakan dalam rangka mengukur *HOTS* peserta didik.

Ada beberapa hasil temuan yang membahas terkait soal-soal *HOTS* dalam dunia pendidikan salah satunya penelitian dari (Risalah & Sandie, 2021) yang menghasilkan suatu instrumen *HOTS* matematika dengan materi program linier di kelas 11 jenjang SMA serta penelitian dari (Fauziah, 2020) yang juga menghasilkan instrumen *HOTS* matematika pada sub bahasan yaitu relasi dan fungsi di kelas VIII jejang SMP. Kedua penelitian tersebut sama-sama mengembangkan instrumen *HOTS* tetapi hanya pada jenjang SMA dan SMP. Berdasarkan beberapa uraian terkait masalah yang ditemukan peneliti saat observasi serta beberapa kajian penelitian yang relevan, maka dikembangkan sebuah instrumen tes matematika berbasis serta berorientasi *HOTS* yang bisa dipergunakan sebagai acuan oleh pendidik dalam rangka menyusun instrumen tes yang berorientasi *HOTS* matematika di jenjang SD. Maka karena itu, penulis melakukan sebuah penelitian pengembangan dengan judul yakni "Pengembangan Instrumen Tes *Higher Order Thinking*

Skill (HOTS) Matematika SD Kelas IV di Gugus V Sukawati"

METODE

Penelitian ini termasuk penelitian pengembangan. Samsu, (2017) menyatakan bahwa penelitian pengembangan adalah suatu pendekatan pada suatu penelitian yang dipergunakan dalam rangka pengembangan lebih lanjut terkait dengan hasil penelitian ataupun produk penelitian. Biasanya produk penelitian dari hasil penelitian mempunyai beberapa kekurangan yang harus terus menerus dikembangkan supaya lebih cepat memiliki nilai guna. Maka dari itulah penelitian pengembangan yakni penelitian yang panjang. Sejalan dengan hal itu (Ilyas & Pd, 2015) juga menganggap bahwa penelitian pengembangan yakni sebuah proses maupun sebuah langkah dalam rangka pengembangan sebuah produk baru ataupun penyempurnaan dari produk yang sebelumnya sudah ada. Hardware seperti (modul, buku, alat antu pembelajaran) maupun software seperti program-program pembelajaran di seolah baik laboratorium, perpustakaan maupun di kelas.

Berdasarkan beberapa pendapat para ahli tersebut maka dapat dirangkum bahwa penelitian pengembangan adalah sebuah penelitian yang dilakukan berdasarkan prosedur-prosedur tertentu yang nantinya akan mendapatkan atau menghasilkan sebuah produk maupun pengembangan atau modifikasi produk yang sudah ada yang dapat berupa hardware ataupun software. Model pengembangan yang digunakan dalam mengembangkan produk ini adalah model pengembangan tipe *formative research* yang alurnya terdiri dari tahap *preliminary* dan tahap *formatif evaluation*

HASIL DAN PEMBAHASAN

Kisi-kisi instrumen tes *HOTS* pada penelitian ini dibuat berdasarkan Kompetensi Inti (KI) dan Kompetensi dasar (KD) pada kurikulum 2013. Penyusunan kisi-kisi pada penelitian ini diawali dengan menganalisis kompetensi

Dasar (KD) matematika kelas IV semester 2. Kompetensi Dasar (KD) yang digunakan adalah Kompetensi Dasar yang memenuhi tingkat kognitif C4 (Menganalisis), C5 (mengevaluasi) serta C6 (Mencipta).

KD yang memenuhi syarat tersebut kemudian dikembangkan menjadi beberapa indikator yang akan nantinya dijadikan sebuah instrumen. Berdasarkan hasil analisis Kompetensi Dasar (KD), ada 1 kompetensi dasar dari 6 Kompetensi dasar yang memenuhi kriteria soal HOTS yakni “Menganalisis segi banyak beraturan dan segi banyak tidak beraturan”. Kompetensi Dasar ini kemudian dikembangkan menjadi 4 indikator yang mengacu pada tingkat kognitif C4, C5 serta C6. Selanjutnya disusunlah kisi-kisi tersebut yang memuat beberapa komponen yakni sebagai berikut Indikator Pencapaian Kompetensi (IPK), Dimensi Proses Berpikir, Dimensi Pengetahuan, Cakupan Materi serta nomor butir soal.

Kisi-kisi yang telah disusun kemudian diserahkan kepada para judges untuk diuji pakar. Berdasarkan analisis para judges kisi-kisi yang telah disusun sudah memenuhi dan sesuai kriteria karena para judges tidak memberikan masukan ataupun komentar terkait kisi-kisi yang telah disusun. Nilai r_{hitung} yang didapatkan lalu dibandingkan dengan nilai r_{tabel} . R_{hitung} terbilang valid jika $r_{hitung} > r_{tabel}$. Perhitungan yang dilakukan berdasarkan jumlah responden dengan taraf sigmifikansi 5% dengan dk = 60 maka diperoleh nilai r_{tabel} adalah 0,25 berdasarkan perhitungan tersebut dari 10 butir soal yang dujikan kesepuluh butir soal tersebut terbilang valid. Dipeoleh pula bahwa nilai r_{11} yang didapatkan adalah 0,79 dan jika dibandingkan dalam acuan patokan yang digunakan maka nilai r_{11} tersebut tergolong dalam kategori derajat reliabilitas yang tinggi. Sementara itu pada uji daya beda dapat disimpulkan 7 soal yang berkategori baik dan 3 soal yang berkategori cukup dan pada uji tingkat kesukaran dapat disimpulkan ada 1 soal dengan kriteria sukar, ada 6 soal dengan kriteria sedang serta ada 3 soal

dengan kriteria mudah. Diperkuat lagi dengan hasil uji pada rasch model

SUMMARY OF 60 MEASURED PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	17.3	10.0	.26	.41	.99	.0	1.01	.1
S.D.	6.0	.0	.94	.11	.44	1.0	.48	.9
MAX.	28.0	10.0	2.46	.99	2.47	2.9	2.89	3.1
MIN.	1.0	10.0	-3.01	.34	.38	-2.1	.38	-1.8

REAL RMSE	.45	TRUE SD	.82	SEPARATION	1.84	PERSON RELIABILITY	.77
MODEL RMSE	.42	TRUE SD	.84	SEPARATION	1.99	PERSON RELIABILITY	.80
S.E. OF PERSON MEAN = .12							

PERSON RAW SCORE-TO-MEASURE CORRELATION = .98
 CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .79

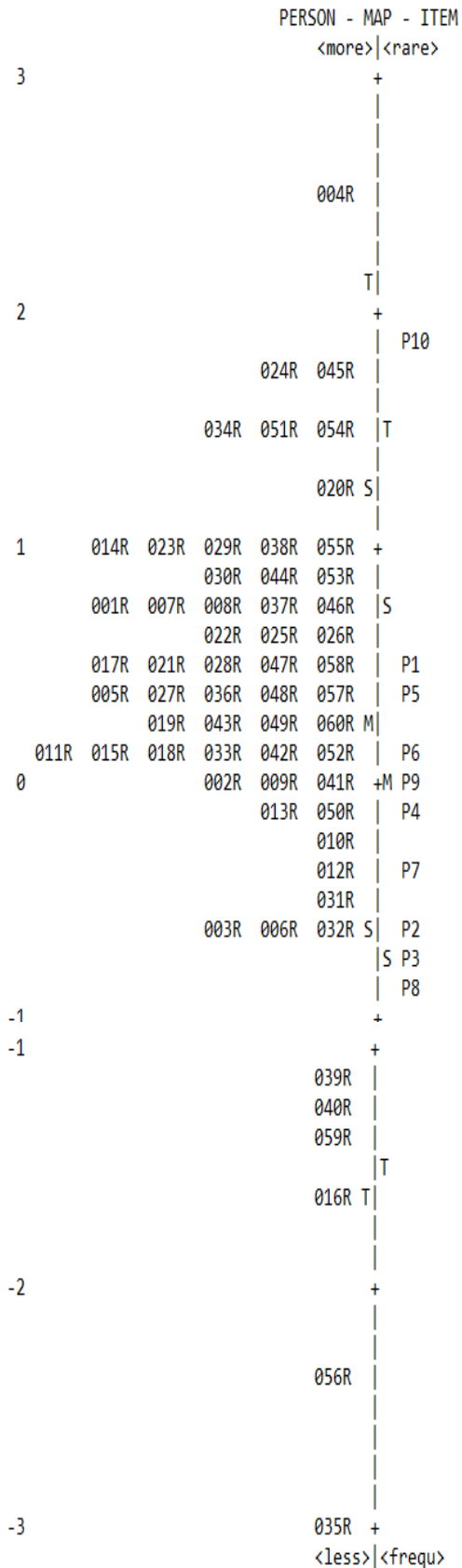
SUMMARY OF 10 MEASURED ITEM

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	103.8	60.0	.00	.16	1.03	.0	1.01	.0
S.D.	31.4	.0	.76	.02	.26	1.4	.28	1.3
MAX.	138.0	60.0	1.85	.20	1.64	2.6	1.54	1.9
MIN.	28.0	60.0	-.84	.14	.71	-2.1	.66	-1.9

REAL RMSE	.17	TRUE SD	.74	SEPARATION	4.35	ITEM RELIABILITY	.95
MODEL RMSE	.16	TRUE SD	.74	SEPARATION	4.59	ITEM RELIABILITY	.95
S.E. OF ITEM MEAN = .25							

Gambar 1. Rangkuman Output

Gambar diatas merupakan hasil output berupa rangkuman dari hasil analisis yang telah dilakukan. Berdasarkan output rangkuman tentang item tes dan juga tentang subjek tersebut didapat bahwa tes mempunyai *separation* 4,35 dan reliabilitas tes 0,95. Hal ini artinya tes telah berfungsi cukup baik dikarenakan mempunyai tingkat kesulitan yang beragam. Sementara itu subjek hanya mempunyai *separation* 1,84 dan reliabilitas subjek 0,77. Berdasarkan hasil tersebut artinya subjek kurang variatif karena hanya memiliki range ability yang sempit



Gambar 2. Item Map

Berdasarkan peta persebaran item dan subjek diatas, pada bagian kiri yakni persebaran kemampuan subjek, sementara itu pada bagian kanan yakni persebaran item. Berdasarkan persebaran pada peta tersebut diketahui bahwa item yang paling sulit adalah item nomer 10 (P10) yang berada di posisi paling atas. Secara teoritis berdasarkan peta persebaran diatas soal P10 hanya subjek no 004R yang memiliki probabilitas menjawab benar karena posisinya ada diatas dari soal (P10). Sementara itu soal yang paling mudah adalah soal (P8) karena sebagian besar subjek memiliki probabilitas menjawab dengan benar kecuali bagi subjek no 039R,040R,059R,016R,056R dan 035R karena posisinya masih berada di bawah dari soal (P8)

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	OBSERVED SCORE	OBSVD COUNT	SAMPLE %	INFINIT AVRG	OUTFIT EXPECT	ANDRICH MISQ	CATEGORY THRESHOLD	MEASURE
0	0	129	22	-1.09	-1.11	1.10 1.17	NONE	(-1.89) 0
1	1	118	20	-.10	-.09	.94 .84	-.47	-.50 1
2	2	139	23	.51	.55	1.09 1.20	.08	.51 2
3	3	214	36	1.10	1.09	.93 .92	.39	(1.86) 3

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	50% CUM. PROBABILITY	COHERENCE M->C C->M	ESTIM RMSR	DISCR
0	NONE		(-1.89) -INF	-1.21	79% 46%	1.0680	0
1	-.47	.13	-.50 -1.21	.01	-.87 34%	39% .7885	.94 1
2	.08	.11	.51 .01	1.21	.03 29%	55% .6031	.96 2
3	.39	.10	(1.86) 1.21	+INF	.85 76%	43% .7968	1.07 3

!->C = Does Measure involve Category?

Gambar 3. Rating Scale

Rating scale pada data politomi ini bisa digolongkan berfungsi dengan baik karena berdasarkan data observed count setiap category label terdapat beberapa subjek yang mendapatkan nilai itu (0,1,2,3) selain itu rating scale juga dikatakan baik jika dilihat dari category measure yakni terjadinya peningkatan yang bertahap dari -1,89 menjadi -0,50 lalu 0,51 hingga 1,86

Berdasarkan dari nilai Andrich Threshold diperoleh hasil sebagai hasil tersebut didapat data sebagai berikut tahap 1 memiliki tingkat kesulitan tahap sebesar - 0,47 (tinggi), tahap 2 memiliki tingkat kesulitan tahap sebesar 0,08 (tinggi), tahap 3 memiliki tingkat kesulitan tahap sebesar 0,39 (sedang)

Tingkat kesulitan tahap dapat dianggap sebagai tingkat kesulitan tahap yang berkaitan dengan transisi atau perpindahan dari satu kategori ke kategori berikutnya. Jadi, menemukan bahwa individu yang memiliki kemampuan di bawah δ_1 (delta 1) atau dibawah - 0,47 memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor 0, dan sebaliknya apabila individu memiliki kemampuan di diatas δ_1 (delta 1) atau diatas - 0,47 memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor 1. Sementara itu individu yang memiliki kemampuan di bawah δ_2 (delta 2) atau dibawah 0,08 memiliki probabilitas atau peluang yang untuk memperoleh skor 1 sebaliknya apabila individu memiliki kemampuan di diatas δ_2 (delta 2) atau diatas 0,08 memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor 2. Kemudian individu yang memiliki kemampuan di bawah δ_3 (delta 3) atau dibawah 0,39 memiliki probabilitas atau peluang untuk memperoleh skor 2 dan sebaliknya apabila individu memiliki kemampuan di diatas δ_3 (delta 3) atau diatas 0,39 memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor maksimal yakni 3.

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	PTMEA CORR.	ITEM
10	0	0	42	70	-.01	.14	1.0	-.44	P10
	1	1	9	15	.56	.10	.5	.14	
	2	2	8	13	1.17	.28	.9	.38	
	3	3	1	2	1.47		.7	.17	
1	0	0	19	32	-.44	.27	1.4	-.51	P1
	1	1	10	17	.51	.22	1.6	.12	
	2	2	20	33	.49*	.09	.7	.18	
	3	3	11	18	.81	.24	1.3	.28	
5	0	0	15	25	-.58	.27	.8	-.51	P5
	1	1	15	25	-.06	.16	.6	-.20	
	2	2	19	32	.69	.12	.6	.32	
	3	3	11	18	1.08	.19	.7	.42	
6	0	0	12	20	-1.01	.28	.6	-.68	P6
	1	1	16	27	.24	.09	.6	-.01	
	2	2	14	23	.38	.15	.8	.07	
	3	3	18	30	1.03	.14	.7	.54	
9	0	0	4	7	-.14	.18	1.1	-.11	P9
	1	1	24	40	-.12	.21	1.5	-.33	
	2	2	16	27	.35	.19	1.9	.06	
	3	3	16	27	.83	.21	1.2	.37	
4	0	0	10	17	-.72	.36	1.0	-.47	P4
	1	1	12	20	-.33	.22	.6	-.31	
	2	2	14	23	.39	.15	.7	.08	
	3	3	24	40	.88	.12	.8	.54	
7	0	0	6	10	-1.52	.44	.5	-.63	P7
	1	1	16	27	.02	.14	.9	-.15	
	2	2	8	13	.44	.22	1.0	.08	
	3	3	30	50	.69	.13	1.1	.46	
2	0	0	7	12	-1.19	.47	1.1	-.56	P2
	1	1	4	7	-.60	.29	.3	-.24	
	2	2	19	32	.34	.16	1.1	.06	
	3	3	30	50	.66	.12	.9	.42	
3	0	0	7	12	-1.10	.48	1.4	-.53	P3
	1	1	4	7	-.86	.37	.3	-.32	
	2	2	16	27	.59	.21	2.5	.21	
	3	3	33	55	.52*	.09	1.0	.31	
8	0	0	7	12	-.61	.49	2.3	-.34	P8
	1	1	8	13	-.50	.39	1.1	-.32	
	2	2	5	8	.22	.28	.7	-.01	
	3	3	40	67	.57	.11	1.1	.46	

* Average ability does not correspond with category score

Gambar 4. Item Measure

Berdasarkan tabel item measure ada beberapa hasil yang bisa diinterpretasikan yakni hasil dari rata-rata ability peserta didik dengan contoh penjabaran pada butir soal nomor 10 yakni peserta didik yang memiliki rata-rata ability -0,01 memiliki probabilitas mendapatkan skor 0, kemudian peserta didik yang memiliki rata-rata ability 0,56

memiliki probabilitas mendapatkan skor 1, lalu peserta didik dengan rata-rata ability 1,17 memiliki probabilitas mendapatkan skor 2 sementara itu peserta didik yang memiliki rata-rata ability 1,47 memiliki probabilitas mendapatkan skor 3 dan begitu seterusnya untuk butir soal yang lain tentunya disesuaikan dengan hasil pada output

PENUTUP

Berdasarkan pada hasil observasi yang telah dilakukan peneliti di SD N Gugus V Sukawati ada beberapa permasalahan yakni yang ditemukan yakni Soal-soal yang diberikan guru sebagai ulangan harian atau latihan masih bersifat hapalan rumus-rumus belum mengacu pada jenis soal HOTS. Para siswa tidak dibiasakan untuk menyelesaikan jenis-jenis soal yang membutuhkan analisis tingkat tinggi sehingga pada saat pelaksanaan ANBK ada beberapa siswa yang memerlukan bimbingan guru. Hal inilah yang mendasari penelitian pengembangan instrumen HOTS ini dilakukan. Penelitian ini adalah penelitian pengembangan atau research and Development. Penelitian pengembangan ini bertujuan dalam rangka menghasilkan sebuah produk yakni instrumen tes matematika kelas IV SD yang berbasis HOTS. Oleh karena penelitian pengembangan ini menggunakan model Formatif Evaluation maka ada beberapa tahapan yang dilakukan pada penelitian ini sehingga bisa menghasilkan instrumen tes matematika berbasis HOTS yakni tahap preliminari dan tahap formatif evaluation. Setelah dilakukan penelitian maka diperoleh hasil atau produk berupa kisi-kisi dan instrumen tes.

Berdasarkan hasil penelitian yang diperoleh ada beberapa hal yang bisa disimpulkan yakni kisi-kisi instrumen berbasis HOTS Matematika kelas IV SD memuat satu Kompetensi Dasar (KD) dan 4 pengembangan indikator yang dengan 1 indikator termasuk C4/K2, 2 indikator termasuk C5/K3 serta 1 indikator termasuk C6/K4.

Sementara itu kualitas tes instrumen matematika berbasis HOTS jika dilihat dari hasil uji validitas dan reliabilitas menunjukkan bahwa kesepuluh instrumen tersebut berkategori valid serta memiliki tingkat reliabilitas yang tinggi yakni 1,1. Sementara itu kualitas tes jika dilihat dari pengujian uji daya beda menghasilkan 7 tes berkategori baik dan 3 tes berkategori cukup. Dan kualitas tes melalui uji tingkat kesukaran mendapatkan hasil bahwa 3 soal berkategori mudah, 6 soal berkategori sedang dan 1 soal berkategori sulit serta pada analisis partial credit model diperoleh output yakni tahap 1 memiliki tingkat kesulitan tahap sebesar $-0,47$ (tinggi), tahap 2 memiliki tingkat kesulitan tahap sebesar $0,08$ (tinggi), tahap 3 memiliki tingkat kesulitan tahap sebesar $0,39$ (sedang). Berdasarkan grafik PCM individu yang memiliki kemampuan dibawah $\delta 1$ (delta 1) atau dibawah $-0,47$ memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor 0, dan sebaliknya apabila individu memiliki kemampuan di diatas $\delta 1$ (delta 1) atau diatas $-0,47$ memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor 1. Sementara itu individu yang memiliki kemampuan di bawah $\delta 2$ (delta 2) atau dibawah $0,08$ memiliki probabilitas atau peluang yang untuk memperoleh skor 1 sebaliknya apabila individu memiliki kemampuan di diatas $\delta 2$ (delta 2) atau diatas $0,08$ memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor 2. Kemudian individu yang memiliki kemampuan di bawah $\delta 3$ (delta 3) atau dibawah $0,39$ memiliki probabilitas atau peluang untuk memperoleh skor 2 dan sebaliknya apabila individu memiliki kemampuan di diatas $\delta 3$ (delta 3) atau diatas $0,39$ memiliki probabilitas atau peluang yang tinggi untuk memperoleh skor maksimal yakni 3.

Berdasarkan beberapa analisis yang dilakukan maka dapat disimpulkan diperolehnya kisi-kisi instrumen tes HOTS matematika kelas IV serta instrumen HOTS matematika kelas IV yang memiliki kualitas.

DAFTAR RUJUKAN

- Ahmad, A. (2022). Pemantauan Pelaksanaan Asesmen Nasional Berbasis Komputer (ANBK) Sekolah Dasar Binaan Kecamatan Kopang Kabupaten Lombok Tengah Tahun 2021. *Jurnal Paedagogy*, 9(1), 34-44. <https://doi.org/10.33394/jp.v9i1.4619>
- Anas, M., Muchson, M., & Forijati, R. (2021). Pengembangan kemampuan guru ekonomi di Kediri melalui kegiatan pelatihan asesmen kompetensi minimum (AKM). In *Rengganis Jurnal Pengabdian Masyarakat*, 1(1), 48-57. <https://mathjournal.unram.ac.id/index.php/Rengganis/index>
- Anselmus, Risalah, D., & Sandie. (2021). Pengembangan Instrumen Tes HOTS dalam Mengukur Kemampuan Berpikir Analitis Siswa pada Materi Program Linear di Kelas XI SMA Sungai Kehidupan. In *Journal for Research in Mathematics Learning*, 4(4), 371-378. <https://ejournal.uin-suska.ac.id/index.php/juring/article/view/14295/7086>
- Arifin, Z., & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X. *PYTHAGORAS: Jurnal Pendidikan Matematika*, 12(1), 98-108. <https://doi.org/10.21831/pg.v12i1.14058>
- Driana, E. & P. U. M. H. J. (2019). *Teachers' Understanding And Practices in Assessing Higher Order Thinking Skills At Primary Schools*. In *Acitya: Journal of Teaching & Education* 1(2), 110-118. <http://journals.umkt.ac.id/index.php/acitya>
- Fauziah. (2020). *Pengembangan Instrumen Penilaian Berbasis Hots Pada Pokok Bahasan Relasi Dan Fungsi*. Universitas Muhammadiyah Surakarta
- Hermawati, H., Jumroh, J., & Sari, E. F. P. (2021). Analisis Kemampuan Pemecahan Masalah Matematis pada Materi Kubus dan Balok di SMP. *Mosharafa: Jurnal Pendidikan Matematika*, 10(1), 141-152. <https://doi.org/10.31980/mosharafa.v10i1.874>
- Ilyas, M., & Pd, M. (2015). *Metodologi Penelitian Pendidikan Matematika*. Pustaka Ramadhan.
- Kurniasi, E. R., & Arsisari, A. (2020). Pengembangan Instrumen Pengukur Higher Order Thinking Skills (Hots) Matematika Pada Siswa Sekolah Menengah Pertama. *Aksioma: Jurnal Program Studi Pendidikan Matematika*, 9(4), 1213-1222. <https://doi.org/10.24127/ajpm.v9i4.3162>
- Lestari. (2019). Pengembangan Instrumen Asesmen Higher Order Thinking SKILL (HOTS) pada Materi Himpunan Kelas VII SMP. *Jurnal Kajian Pendidikan Matematika*, 4(2), 111-120. <http://journal.lppmunindra.ac.id/index.php/jkpm/>
- Masitoh, L. F., & Aedi, W. G. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (Hots) Matematika Di SMP Kelas VII. *Jurnal Cendekia: Jurnal Pendidikan Indonesia*. 4(2), 886-897. <https://j-cup.org/index.php/cendekia/article/view/328/196>
- Muflikhah, I. K., Rahmawati, A. D., & Wahyuningsih, S. (2021). Prosiding SEMAI Seminar Nasional PGMI 2021 Analisis Karakteristik Siswa MI/SD dalam Menghadapi Asesmen Kompetensi Minimum (AKM). *SEMAI Seminar Nasional PGMI 2021*, 1(1), 302-321. <http://proceeding.iainpekalongan.a>

- c.id/index.php/semair-302-
- Purwati, P. D., Faiz, A., Widiyatmoko, A., & Maryatul, S. (2021). Asesmen Kompetensi Minimum (AKM) kelas jenjang sekolah dasar sarana pemacu peningkatan literasi peserta didik. *Jurnal Kajian Pendidikan Umum*, 19(4), 13-24. <https://ejournal.upi.edu/index.php/SosioReligi/article/view/39347/16421>
- Purwati, P.D., Widiyatmoko, A., & Kiptiyah, S.M. (2021). Pembekalan Guru SD Gugus Sindoro Blora Melalui Workshop Asesmen Nasional Menghadapi AKM Nasional. *Journal of Community Empowerment*, 1(1), 32-40. <http://journal.unnes.ac.id/sju/index.php/jce>
- Pratama, G. S., & Retnawati, H. (2018). Urgency of Higher Order Thinking Skills (HOTS) Content Analysis in Mathematics Textbook. *Journal of Physics: Conference Series*, 1097(1), 1-8. <https://doi.org/10.1088/1742-6596/1097/1/012147>
- Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, 76(2), 215–230. <https://doi.org/10.33225/pec/18.76.215>
- Safi'i, I., & Amar, F. (2019). Pelatihan Penyusunan Instrumen Evaluasi Berstandar Hots Bagi Guru-Guru SD di Wilayah Banyudono. *Abdimas Dewantara*, 2(2), 149–157. <https://doi.org/10.30738/ad.v2i2.4176>
- Samsu. (2017). *Metode Penelitian*. Pusaka.
- Setiawati, W., Asmira, O., Ariyana, Y., Reisky, B., & Pudjiastuti, A. (2018). *Buku Penilaian Berorientasi Higher Order Thinking Skills*. Direktorat Jenderal Guru dan Tenaga Kependidikan.
- Suhady, W., Roza, Y., & Maimunah, M. (2020). Pengembangan Soal untuk Mengukur Higher Order Thinking Skill (HOTS) Siswa. *Jurnal Gantang*, 5(2), 143–150. <https://doi.org/10.31629/jg.v5i2.2518>
- Sujana, I. W. C. (2019). Fungsi Dan Tujuan Pendidikan Indonesia. *Adi Widya: Jurnal Pendidikan Dasar*, 4(1), 29-39. <https://doi.org/10.25078/aw.v4i1.927>
- Tanujaya, B., Mumu, J., & Margono, G. (2017). The Relationship between Higher Order Thinking Skills and Academic Performance of Student in Mathematics Instruction. *International Education Studies*, 10(11), 78-85. <https://doi.org/10.5539/ies.v10n11p78>