

Pengembangan Instrumen Tes Hasil Belajar Matematika Berbasis Pendekatan Kontekstual Siswa Kelas IV Sekolah Dasar Wilayah II Marioriwawo Kabupaten Soppeng

Sutriani¹, Sukmawati², Rukli³

^{1), 2), 3)} Magister Pendidikan Dasar, Program Pascasarjana
Universitas Muhammadiyah Makassar

Abstrak. Salah satu komponen esensial dalam program pembelajaran adalah evaluasi pembelajaran yang menentukan kondisi, dimana suatu tujuan telah dapat dicapai. Penelitian ini bertujuan untuk mendeskripsikan hasil pengembangan instrumen tes hasil belajar matematika berbasis pendekatan kontekstual di sekolah dasar ditinjau dari validitas isi, tingkat kesukaran, daya beda butir, efektivitas pengecoh, validitas butir dan reliabilitas. Penelitian ini merupakan penelitian pengembangan model 4-D yang melibatkan 65 siswa kelas IV sekolah dasar di kabupaten Soppeng sebagai sampel penelitian. Pengumpulan data dilakukan dengan teknik tes dan dianalisis dengan pendekatan deskriptif kuantitatif. Hasil penelitian ini menunjukkan bahwa instrumen yang terdiri dari 25 butir soal ini memenuhi aspek validitas isi dengan koefisien konsistensi internal sebesar 1. Dari 25 butir soal yang dikembangkan sebanyak 2 butir soal sulit, 9 butir soal sedang, dan 14 butir soal mudah. Selain itu, 22 butir soal yang memiliki daya beda yang baik dan 3 butir soal memiliki daya beda cukup. Lebih lanjut, terdapat 22 butir soal yang memiliki pengecoh yang berfungsi dan 3 butir soal lainnya memiliki pengecoh yang kurang berfungsi. Dari aspek validitas butir, terdapat 22 butir soal yang valid secara empirik sementara 3 butir soal lainnya direvisi karena tidak valid. Untuk aspek reliabilitasnya, instrumen tes hasil belajar ini telah reliabel dengan koefisien reliabilitas 0,820. Dengan demikian, instrumen tes hasil belajar matematika layak digunakan untuk mengukur hasil belajar matematika siswa.

Kata kunci: *Hasil Belajar Matematika, Pendekatan Kontekstual*

A. Pendahuluan

Program pembelajaran merupakan program yang dilaksanakan oleh guru di sekolah untuk mengembangkan kompetensi, indikator serta tujuan-tujuan pembelajaran dalam suatu proses pembelajaran di kelas (Marufi, Ilyas, Salwah & Syamsuddin, 2020). Keberhasilan program pembelajaran sangat tergantung pada sejauh mana kompetensi, indikator dan tujuan-tujuan pembelajaran tersebut dapat tercapai dengan baik (Syamsuddin, 2019). Informasi tersebut dapat diperoleh dengan melakukan evaluasi pembelajaran dengan instrumen dan prosedur-prosedur yang tepat. Salah satu komponen esensial dalam program pembelajaran adalah evaluasi

pembelajaran. Evaluasi merupakan bagian dari proses pembelajaran yang menentukan kondisi, dimana suatu tujuan telah dapat dicapai. Menurut Arifin (2017) evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Sementara Wirawan (2016) menyatakan bahwa evaluasi sebagai riset untuk mengumpulkan, menganalisis dan menyajikan informasi yang bermanfaat mengenai objek evaluasi, menilainya dengan membandingkannya dengan indikator evaluasi untuk pengambilan keputusan.

Oleh karena itu, seyogyanya kegiatan evaluasi harus dilakukan secara terencana, terprogram, dan terpercaya. Dalam evaluasi pendidikan, ada tiga komponen yang saling terkait dan merupakan satu kesatuan yang tidak terpisahkan, yaitu penilaian, pengukuran, dan tes. Penilaian sangat berperan dalam penyelenggaraan evaluasi yang bertujuan dalam pengambilan keputusan karena penilaian yang baik akan menentukan pengambilan keputusan yang baik pula. Pada dasarnya penilaian memberikan tekanan pada usaha guru maupun siswa untuk memperoleh informasi tentang hasil pembelajaran yang telah dilakukan sehingga dapat menjawab pertanyaan tentang sebaik apa hasil atau prestasi belajar seorang siswa (Syamsuddin, Jannah, & Kristiawati, 2019).

Penilaian tidak bisa lepas dari pengukuran sehingga hasil pengukuran akan menjadi dasar atau acuan dalam melakukan penilaian. Pengukuran berhubungan dengan proses pencarian atau penentuan nilai secara kuantitatif. Untuk menentukan nilai secara kuantitatif tersebut diperlukan alat untuk mengukur, maka alat ukur itu adalah tes yang bentuknya beranekaragam dan digunakan dalam lembaga pendidikan (Widoyoko, 2012). Tes sebagai alat untuk mengukur kemampuan peserta didik keberadaannya menjadi sangat penting yang menggambarkan keterampilan siswa dalam memilih salah satu atau lebih diantara beberapa kemungkinan jawaban atau dengan jalan menuliskan jawaban pada tiap-tiap butir tes (Susetyo, 2015). Tes merupakan cara penilaian yang dirancang dan dilaksanakan atau diberikan kepada peserta didik pada waktu dan tempat tertentu serta dalam kondisi yang memenuhi syarat-syarat tertentu (Djaali & Muljono, 2008). Tes biasanya diberikan kepada peserta didik setelah mengikuti kegiatan pendidikan selama selang waktu tertentu (Mardapi, 2008; Supardi, 2015; Sudibyo, Jatmiko & Widodo, 2017).

Dalam Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan dijelaskan lebih jauh tentang pelaksanaan evaluasi. Pada pasal 63 ayat (1)

menyebutkan bahwa penilaian pendidikan pada jenjang pendidikan dasar dan menengah terdiri atas: (a) penilaian hasil belajar oleh pendidik, (b) penilaian hasil belajar oleh satuan pendidikan, dan (c) penilaian hasil belajar oleh pemerintah. Penilaian hasil belajar oleh pendidik dan satuan pendidikan merupakan bentuk evaluasi internal (*internal evaluation*). Berkenaan dengan evaluasi yang dilakukan oleh pemerintah, sifatnya sebagai evaluasi eksternal (*external evaluation*) yang sasarannya adalah siswa, termasuk di dalamnya hasil belajar siswa di sekolah.

Uraian tersebut menunjukkan bahwa begitu pentingnya aspek penilaian dalam meningkatkan kualitas pelaksanaan program pembelajaran di sekolah. Penilaian yang baik akan menghasilkan kualitas pembelajaran yang baik pula. Namun fenomena pelaksanaan penilaian di Wilayah II Kecamatan Marioriwawo Kabupaten Soppeng tampaknya belum sesuai dengan harapan. Dimana hasil pengamatan yang dilakukan oleh peneliti diperoleh bahwa penilaian hasil belajar yang dilakukan oleh guru di sekolah tersebut belum memenuhi kriteria terutama dalam pembakuan instrumen, sehingga hasil yang diperoleh belum maksimal. Hal ini terjadi karena guru belum terbiasa mendesain penilaian walaupun ini merupakan tugas dan tanggung jawab seorang guru untuk melaksanakan proses pembelajaran dan melakukan penilaian.

Bahkan sering ditemukan guru dalam membuat tes hasil belajar hanya mengkopi dari buku paket saja tanpa adanya usaha untuk mengembangkan sendiri sesuai dengan prosedur pengembangan tes hasil belajar yang semestinya. Hal ini terlihat pada saat penulis melakukan observasi dan wawancara terbatas dengan guru kelas di SDN 235 Goarie tampak guru memberikan soal siswa dengan langsung mengutip dari buku paket untuk dijadikan sebagai alat tes untuk mengukur pemahaman siswa. Soal yang disajikan ke siswa pun cenderung tidak berkaitan dengan kejadian yang ada di sekitar kehidupan keseharian siswa baik di sekolah maupun di rumah. Atas dasar tersebut maka penggunaan penilaian hasil belajar matematika berbasis pendekatan kontekstual pada siswa merupakan kebutuhan yang mendesak untuk dikembangkan dalam proses pembelajaran yang merupakan penilaian yang memenuhi Standar Nasional Pendidikan.

Hal lain yang ditemukan berdasarkan hasil pengamatan dan data dari beberapa guru di Wilayah II Kecamatan Marioriwawo Kabupaten Soppeng bahwa selama ini hasil belajar siswa kelas IV masih termasuk rendah khususnya pada mata pelajaran matematika. Hal ini ditunjukkan pada skor nilai mata pelajaran matematika semester yang lalu Tahun Pelajaran 2017/2018 di

Wilayah II Kecamatan Marioriwawo, rata-rata skor Ulangan Harian mencapai 6,98, skor Ulangan Semester 6,62 serta skor Ujian Nasional 7,12. Diduga bahwa belum maksimalnya hasil belajar peserta didik pada mata pelajaran matematika selama ini disebabkan karena pelaksanaan penilaian yang belum maksimal. Dalam hal ini, peneliti menyoroti pada aspek instrumen penilaian yang digunakan dalam menjangkau data hasil belajar siswa.

Temuan lainnya yang mendukung dugaan tersebut berdasarkan wawancara terbatas dengan guru kelas diperoleh informasi bahwa selama ini dalam mengembangkan instrumen penilaian hasil belajar matematika, guru tidak pernah melakukan penilaian hasil belajar siswa yang berbasis pendekatan kontekstual dan analisis kualitas instrumen tes hasil belajar matematika yang dikembangkan tidak memenuhi baik dari aspek kualitatifnya maupun pada aspek kuantitatifnya, sehingga tidak ada kejelasan kualitas dari instrumen yang dikembangkan selama ini.

Pembelajaran kontekstual merupakan pembelajaran yang mengkaitkan materi pembelajaran dengan konteks dunia nyata yang dihadapi siswa sehari-hari baik dalam lingkungan keluarga, masyarakat, alam sekitar dan dunia kerja, sehingga siswa mampu membuat hubungan antara pengetahuan yang dimilikinya dengan penerapannya dalam kehidupan sehari-hari (Agustan, 2016). Dengan pembelajaran kontekstual diharapkan terjadi proses belajar, dimana belajar diartikan sebagai perubahan pada individu yang terjadi melalui pengalaman dan bukan karena pertumbuhan atau perkembangan tubuh manusia serta bukan pula karakteristik yang dibawa sejak lahir. Begitu juga dengan belajar matematika yang membutuhkan pengetahuan sebelumnya untuk mengkonstruksi tercapainya pengetahuan baru (Kharie, La Nani dan Tonra, 2021)

Merujuk pada hasil penelitian Sulianto (2008) bahwa pendekatan pembelajaran matematika di sekolah yang diduga akan sejalan dengan harapan dari kurikulum dan dapat meningkatkan berpikir kritis siswa adalah pendekatan kontekstual. Selain itu, agar dapat berpikir kritis maka siswa dihadapkan pada situasi atau masalah matematika yang menuntut siswa menerima berbagai jenis informasi dan menggunakan pengetahuan yang dimiliki serta membuat keputusan dalam situasi-situasi yang kompleks (Syamsuddin, 2020). Dalam pembelajaran yang menggunakan pendekatan kontekstual materi disajikan melalui konteks yang bervariasi dan berhubungan dengan kehidupan siswa baik di rumah, di sekolah maupun di masyarakat secara luas, dan pengetahuan didapat oleh siswa secara konstruktivis. Sejalan dengan hasil penelitian Situmorang (2013) bahwa

penerapan pendekatan kontekstual dalam pembelajaran matematika dapat meningkatkan aktivitas dan hasil belajar siswa sekolah dasar.

Berdasarkan permasalahan yang diuraikan tersebut menunjukkan perlunya pengembangan instrumen yang valid dan reliabel agar pelaksanaan penilaian hasil belajar matematika berbasis pendekatan kontekstual di masa yang akan datang dapat lebih berkualitas dengan dukungan instrumen penilaian hasil belajar yang berkualitas pula. Pengembangan dilakukan untuk membuat atau memperbaiki instrumen yang sudah ada sehingga menjadi produk yang semakin bermanfaat untuk meningkatkan kualitas dan mutu pendidikan dan pembelajaran yang lebih baik (Wiryokusumo, 2011; Sumarno, 2012). Oleh karena itu, dibutuhkan penelitian lebih lanjut untuk mengembangkan dan memperoleh instrumen penilaian hasil belajar matematika berbasis pendekatan kontekstual yang baku.

B. Metode Penelitian

Penelitian ini merupakan penelitian pengembangan dengan model pengembangan 4-D yang terdiri atas 4 tahap utama yaitu: *define* (pendefinisian), *design* (perencanaan), *develop* (pengembangan) dan *dissemination* (penyebaran) (Thiagarajan, 1974). Prosedur Penelitian dan Pengembangan atau *Research and Development* (R&D) ini merupakan proses atau langkah-langkah untuk mengembangkan dan mendesain instrumen Penilaian Hasil Belajar Matematika. Adapun langkah-langkah pengembangan instrument tes hasil belajar matematika berbasis kontekstual yaitu (1) *define* meliputi (a) melakukan analisis kebutuhan terkait instrumen hasil belajar matematika di Wilayah II Kecamatan Marioriwawo; (b) melakukan analisis konsep untuk merumuskan aspek-aspek atau indikator dari penilaian yang akan dilakukan; (c) melakukan observasi awal terkait dengan kesiapan subjek penelitian dan lokasi pengujian produk pengembangan, (2) *design* meliputi (a) penyusunan kisi-kisi instrumen penilaian; (b) pemilihan format instrumen hasil belajar yang ingin dikembangkan (objektif pilihan ganda); (c) rancangan produk awal dengan cara penulisan butir soal hasil belajar matematika berdasarkan kisi-kisi instrumen, (3) *develop* meliputi (a) validasi secara teoretik (validasi isi) berdasarkan kesepahaman dari dua orang pakar sesuai koefisien konsistensi internal Gregory; (b) validasi secara empirik yaitu analisis butir secara kuantitatif (tingkat kesukaran, daya beda dan efektivitas pengecoh), menghitung validitas butir dan menghitung Reliabilitas tes; (c) ujicoba validasi empirik (sampai

hasil jenuh); dan (d) diperoleh instrumen hasil tes belajar matematika final, (4) *disseminate* meliputi (a) sosialisasi instrumen hasil belajar matematika yang telah dikembangkan; (b) penerapan instrumen hasil belajar matematika yang telah baku.

Subjek uji coba penelitian ini adalah seluruh peserta didik kelas IV di Wilayah II Kecamatan Marioriwawo Kabupaten Soppeng yang berjumlah 65 orang siswa yang ada di Wilayah II Kecamatan Marioriwawo Kabupaten Soppeng Provinsi Sulawesi Selatan yang terdiri atas 4 sekolah dasar yang memiliki kesamaan klasifikasi akreditasi sekolah yaitu akreditasi B, yaitu: SDN 165 Asanae, SDN 149 Amessangeng, SDN 223 Lapince, dan SDN 235 Goarie.

Untuk menjaring data dalam penelitian ini digunakan tes melalui uji coba. Adapun tes yang dimaksud dalam penelitian ini adalah instrumen tes hasil belajar matematika yang berbentuk tes objektif pilihan ganda yang dikembangkan dan dilakukan uji coba terhadap 65 siswa kelas IV sekolah dasar. Data yang diperoleh selanjutnya dianalisis secara kuantitatif dengan pendekatan deskriptif kuantitatif dan kualitatif. Adapun penjabaran analisis yang dimaksud adalah sebagai berikut.

a) Analisis secara kualitatif

Analisis secara kualitatif dilakukan untuk menganalisis instrumen tes hasil belajar matematika dengan menilai hasil validitas isi yaitu relevansi butir dengan kompetensi dasar dan indikator melalui bantuan penilaian pakar (*expert judgment*).

b) Analisis secara kuantitatif

Analisis secara kuantitatif meliputi analisis tingkat kesukaran, daya beda butir, keefektifan pengecoh, validitas butir, dan reliabilitas tes.

Analisisnya dapat diuraikan sebagai berikut.

1) Tingkat kesukaran

Berikut ini rumus yang digunakan untuk menghitung tingkat kesukaran butir soal.

$$p = \frac{\sum x}{S_m N}$$

Keterangan

p = proporsi menjawab benar atau tingkat kesukaran

$\sum x$ = banyaknya peserta tes yang menjawab benar

S_m = skor maksimum

N = banyaknya peserta didik yang menjawab butir soal tersebut.

Menurut Mansyur, Rasyid dan Suratno (2009), kriteria yang digunakan untuk menentukan jenis tingkat kesukaran butir soal adalah sebagai berikut.

$P < 0.30$ = butir soal sukar

$0.3 \leq P \leq 0.70$ = butir soal sedang

$P > 0.70$ = butir soal mudah

2) Daya beda

Daya beda butir dihitung dengan menggunakan rumus berikut ini

$$D = \frac{\sum A}{n_A} - \frac{\sum B}{n_B}$$

Keterangan

D = Indeks daya pembeda

$\sum A$ = Jumlah peserta yang menjawab benar pada kelompok atas

$\sum B$ = Jumlah peserta yang menjawab benar pada kelompok bawah

n_A = Jumlah peserta tes kelompok atas

n_B = Jumlah peserta tes kelompok bawah

Untuk menyatakan bahwa besaran daya beda, maka patokan yang dapat digunakan berdasarkan pendapat Naga (2013) sebagai berikut.

Tabel 1. Kriteria daya beda butir

Indeks Daya Beda	Evaluasi Butir
$D > 0.40$	Butir baik
$0.3 \leq D \leq 0.39$	Butir cukup baik/diterima tanpa revisi
$0.2 \leq D \leq 0.29$	Butir kurang baik/perlu direvisi
$D \leq 0.19$	Butir yang buruk/dibuang atau revisi total

3) Efektivitas Distraktor/Pengecoh

Menurut Mardapi (2008), bahwa penyebaran pilihan jawaban dijadikan dasar dalam penelaahan soal. Hal ini dimaksudkan untuk mengetahui berfungsi tidaknya jawaban yang tersedia. Suatu pilihan jawaban atau pengecoh dapat dikatakan berfungsi apabila pengecoh:

- a. Paling tidak dipilih oleh 5 % peserta didik
- b. Lebih banyak dipilih oleh kelompok peserta didik yang belum paham materi.

4) Validitas Butir

Untuk mengestimasi validitas butir tersebut, maka pada penelitian ini menggunakan rumus berikut ini.

$$r_{pbis} = \frac{(\bar{X}_i - \bar{X}_t)}{S_t} \sqrt{\frac{p}{q}}$$

Keterangan:

X_i = Mean total skor peserta yang memiliki jawaban benar.

X_t = Mean skor total

S_t = Standar deviasi skor total

p = Proporsi peserta ujian yang menjawab benar pada butir tes.

q = Proporsi peserta ujian yang menjawab salah atau $(1 - p)$.

Suatu butir dikatakan valid apabila memiliki sumbangan terhadap skor totalnya, jadi koefisien korelasi butirnya besar terhadap skor totalnya.

5) Reliabilitas tes

Perhitungan reliabilitas instrumen menggunakan Kuder Richardson-20 (KR-20) dilakukan dengan menggunakan rumus berikut:

$$KR - 20 = \left(\frac{k}{k - 1} \right) \left(\frac{s_t^2 - \sum pq}{s_t^2} \right)$$

Keterangan:

k = Jumlah butir

s_t^2 = Varians total

p = Proporsi skor yang diperoleh

q = Proporsi skor maksimum dikurangi skor yang diperoleh

Reliabilitas terkait dengan kekonsistenan alat ukur dalam hal ini tes hasil belajar matematika. Pada penelitian ini indeks reliabilitas dihitung dengan menggunakan formula KR-20 dengan kriteria apabila batas minimal suatu tes memiliki keajegan (konsisten dan stabil) adalah minimal 0,70 sedangkan suatu tes dinyatakan tidak reliabel apabila batas suatu tes $< 0,70$ (Mardapi, 2012).

C. Hasil dan Pembahasan

Penyajian data terkait pengembangan instrumen tes hasil belajar matematika berbasis pendekatan kontekstual di sekolah dasar dengan memperhatikan aspek validitas isi, tingkat kesukaran butir soal, daya beda butir, efektivitas distraktor, validitas butir, dan reliabilitas tes dari instrumen. Dengan menggunakan model pengembangan Thiagarajan yang dikenal dengan model 4-D berikut disajikan hasil pengembangan instrumen tes hasil belajar matematika berbasis kontekstual yang didasarkan pada standar kompetensi dan kompetensi dasar kurikulum 13.

1. Tahap Pendefinisian

Tujuan dari tahap ini adalah menetapkan dan mendefinisikan kebutuhan yang dihadapi dalam pembelajaran matematika. Tahap pendefinisian diuraikan secara berikut :

a. Analisis Kebutuhan

Dari analisis yang dilakukan di sekolah dasar wilayah II Kecamatan Marioriwawo Kabupaten Soppeng diperoleh bahwa siswa kesulitan dalam memahami soal atau instrumen tes hasil belajar yang ada karena topik yang diangkat dalam soal tidak kontekstual..

b. Analisis Konsep

Konsep-konsep pada salah satu kompetensi dasar dikaitkan dengan konsep-konsep pada kompetensi lainnya kemudian disusun kedalam sebuah peta konsep. Peta konsep yang telah disusun digunakan sebagai dasar dalam menyusun sebuah indikator soal.

c. Spesifikasi Tujuan

Perumusan penyusunan kisi-kisi disesuaikan standar kompetensi, kompetensi dasar dan indikator yang akan dicapai dengan memperhatikan ranah kognitif (C). Penyusunan kisi-kisi instrumen hasil belajar matematika berbasis kontekstual disajikan pada tabel berikut.

Tabel 2. Spesifikasi tujuan penyusunan kisi-kisi

Standar Kompetensi	Kompetensi Dasar	Indikator Soal	Ranah Kognitif/Nomor Soal		
			C4	C5	C6
8. Memahami sifat bangun ruang sederhana dan hubungan antar bangun datar.	8.1 Menentukan sifat-sifat bangun ruang sederhana	Siswa dapat membedakan bangun ruang berdasarkan sifatnya dengan tepat	1,2, 3	4	
		Siswa dapat menentukan sifat-sifat bangun ruang sederhana dengan tepat.		5	6

Standar Kompetensi	Kompetensi Dasar	Indikator Soal	Ranah Kognitif/Nomor Soal		
			C4	C5	C6
	8.2 Menentukan jaring-jaring balok dan kubus	Disajikan gambar, siswa dapat menentukan jaring-jaring balok dengan tepat.	7, 8		
		Disajikan gambar, siswa dapat menentukan jaring-jaring kubus dengan tepat.	9	10	
		Siswa dapat menentukan jaring-jaring balok sesuai benda yang ada di lingkungan tempat tinggalnya dengan tepat.			11
		Siswa dapat menentukan jaring-jaring kubus sesuai benda yang ada di lingkungan tempat tinggalnya dengan tepat.		12	
	8.3 Mengidentifikasi benda-benda dan bangun datar simetris	Disajikan gambar, siswa dapat menentukan banyaknya sumbu simetri pada salah satu bangun datar dengan tepat.	13	14	
		Disajikan gambar, siswa dapat menentukan kelompok benda-benda yang simetris dengan tepat.	15	16	
		Disajikan gambar, siswa dapat menentukan kelompok benda-benda yang tidak simetris dengan tepat.	17	18	
		Siswa dapat mengidentifikasi ciri bangun datar yang simetris dengan tepat.	19		
		Disajikan gambar siswa dapat menghitung jumlah simetris dengan tepat.	20		21
		Disajikan gambar siswa dapat melengkapi bangun datar yang simetris dengan tepat.		22	
	8.4 Menentukan hasil pencerminan	Siswa dapat menentukan hasil pencerminan bangun datar dengan tepat.	23	24	

Standar Kompetensi	Kompetensi Dasar	Indikator Soal	Ranah Kognitif/Nomor Soal		
			C4	C5	C6
	suatu bangun datar	Disajikan cerita siswa dapat mengidentifikasi hasil pencerminan dengan tepat.		25	

2. Tahap Perancangan

Instrumen tes hasil belajar disusun berdasarkan indikator yang akan dicapai. Sebelum instrumen penilaian hasil belajar disusun terlebih dahulu dibuatkan kisi-kisi soal berbasis kontekstual. Adapun instrumen tes hasil belajar dalam penelitian ini adalah berupa soal pilihan ganda yang terdiri dari 25 nomor.

3. Tahap Pengembangan

Tahap pengembangan bertujuan untuk menghasilkan instrumen yang layak pakai. Sebelum instrumen disebar ke siswa terlebih dahulu instrumen melalui serangkaian uji yaitu validitas empiris dan reliabilitasnya. Adapun langkah-langkah dalam pengembangan sebagai berikut.

a. Validasi

Validasi dilakukan untuk mengetahui kelayakan suatu instrumen penelitian. Adapun instrumen penelitian yang divalidasi adalah instrumen tes hasil belajar matematika berbasis kontekstual dimana instrumen didasarkan pada kesesuaian antara standar kompetensi, kompetensi dasar dan indikator yang akan dicapai. Validasi dilakukan dengan meminta penilaian dari para ahli di bidang pendidikan matematika dan pendidikan dasar melalui pendekatan *expert judgment* dimana instrumen ini divalidasi oleh dua validator ahli yaitu dosen pendidikan matematika dan guru matematika sekolah dasar.

b. Uji Coba

Setelah dilakukan validasi oleh pakar dengan menggunakan analisis kesahihan (validitas) maka dilakukan ujicoba dikelas IV SDN 165 asanae, SDN 149 Amessangeng, SDN 223 Lapince, dan SDN 235 Goarie.

4. Tahap Penyebaran

Tahap penyebaran yang digunakan dalam mempromosikan produk yang dihasilkan adalah Kolokium Online yang diselenggarakan oleh Program Pascasarjana Universitas Muhammadiyah Makassar pada tanggal 17 Juni 2020. Kegiatan ini diikuti oleh dosen dan mahasiswa.

Data yang diperoleh selanjutnya dianalisis dengan menggunakan pendekatan deskriptif kuantitatif sebagai berikut.

1. Hasil Pengembangan Instrumen Berdasarkan Validitas Isi

Validitas isi pada penelitian ini ditentukan melalui pertimbangan para ahli. Pertimbangan para ahli tersebut dilakukan dengan cara, para ahli diminta untuk mencermati semua item dalam tes yang hendak divalidasi. Kemudian diminta untuk mengoreksi interpretasi item-item yang telah dibuat. Para ahli juga diminta untuk memberikan pertimbangan tentang bagaimana interpretasi tes apakah telah menggambarkan cakupan isi yang hendak diukur. Hasil validasi dari pakar di analisis dengan menggunakan analisis kesahihan (validitas). Relevansi kedua pakar secara menyeluruh merupakan koefisien konsistensi internal Gregory. Untuk memutuskan apakah instrumen yang dikembangkan tersebut memiliki koefisien konsistensi internal yang tinggi, maka digunakan model kesepakatan antara dua validator. Berikut disajikan hasil penelitian berdasarkan penilaian pakar terhadap instrumen hasil belajar matematika berbasis pendekatan kontekstual pada siswa kelas IV Wilayah II Marioriwawo Kabupaten Soppeng yang telah dikembangkan pada tabel berikut.

Tabel 3. Data hasil penilaian pakar

Validator II	Validator I	
	Tidak relevan Skor (1-2)	Relevan Skor (3-4)
	Tidak relevan Skor (1-2)	Relevan Skor (3-4)
Tidak relevan Skor (1-2)	0 Butir	0 Butir
Relevan Skor (3-4)	0 Butir	25 Butir

Berdasarkan Tabel 3 di atas, maka dapat dihitung koefisien konsistensi internal. Hasil perhitungan koefisien konsistensi internal diperoleh sebesar 1. Hal ini menunjukkan bahwa instrumen hasil belajar matematika berbasis pendekatan kontekstual konsisten. Dengan demikian maka dapat disimpulkan bahwa seluruh butir (25 butir) telah valid berdasarkan kontennya. Dengan kata lain keseluruhan butir pada instrumen hasil belajar matematika berbasis pendekatan kontekstual yang telah dikembangkan telah valid secara isi. Selain itu instrumen hasil belajar matematika berbasis pendekatan kontekstual yang telah dikembangkan tersebut telah reliabel

berdasarkan konsistensi internalnya. Untuk itu, instrumen hasil belajar matematika berbasis pendekatan kontekstual telah layak untuk diuji cobakan pada siswa.

Jika temuan empirik tersebut dikaitkan dengan tinjauan teoretis sebagaimana yang dikemukakan oleh Guion dalam Surapranata (2004) bahwa validitas isi adalah derajat di mana sebuah tes mengukur cakupan substansi yang ingin diukur. Dengan demikian, dapat disimpulkan bahwa instrumen hasil belajar matematika berbasis kontekstual telah mengukur cakupan substansi yang ingin diukur yakni standar kompetensi, kompetensi dasar dan indikator materi. Dengan kata lain keseluruhan butir pada instrumen hasil belajar matematika berbasis kontekstual yang telah dikembangkan telah valid secara isi.

2. Hasil Ujicoba Instrumen

a. Hasil Ujicoba Berdasarkan Tingkat Kesukaran butir

Tingkat kesukaran butir pada penelitian ini merupakan proporsi peserta tes yang menjawab benar untuk butir tertentu pada instrumen tes hasil belajar matematika berbasis pendekatan kontekstual yang telah dikembangkan. Data hasil analisis tingkat kesukaran pada ujicoba instrumen dijabarkan pada Tabel 4 berikut ini.

Tabel 4. Hasil analisis tingkat kesukaran butir pada ujicoba

Status Butir	Sulit		Sedang										Mudah												
	5	10	4	6	8	9	12	13	14	15	21	1	2	3	7	11	16	17	18	19	20	22	23	24	25
Np	16	19	41	42	43	40	43	42	40	42	45	61	62	53	60	46	55	52	50	57	50	45	48	53	53
p	0,25	0,29	0,63	0,65	0,66	0,62	0,66	0,65	0,62	0,65	0,69	0,94	0,95	0,82	0,92	0,71	0,85	0,80	0,77	0,88	0,88	0,69	0,74	0,82	0,82
q	0,75	0,71	0,37	0,75	0,34	0,38	0,34	0,35	0,38	0,35	0,31	0,06	0,05	0,18	0,08	0,29	0,15	0,20	0,23	0,12	0,23	0,31	0,26	0,18	0,18
p.q	0,19	0,21	0,23	0,19	0,22	0,24	0,22	0,23	0,24	0,23	0,21	0,06	0,04	0,25	0,07	0,21	0,13	0,16	0,18	0,11	0,18	0,21	0,19	0,15	0,15
Jumlah Butir Soal	2		9										14												

Berdasarkan data hasil analisis tingkat kesukaran butir pada ujicoba, dapat dikemukakan bahwa dari 25 butir instrumen tes hasil belajar matematika berbasis pendekatan kontekstual valid secara isi, sebanyak 2 butir soal berada pada kategori sulit, 9 butir soal berada pada kategori sedang, serta 14 butir soal berada pada kategori mudah. Jika dikaitkan antara temuan empirik dengan tinjauan teoretis maka dapat dijelaskan bahwa tingkat kesukaran butir tes diukur melalui besar kecilnya proporsi peserta yang berhasil menjawab butir itu dengan benar (Naga, 2012).

Azwar (2010) menyatakan secara lebih lengkap bahwa indeks kesukaran butir adalah rasio penjawab butir dengan benar dan banyaknya penjawab butir. Proporsi menjawab benar p (*proportion oruect*) adalah indeks kesukaran soal yang paling sederhana dan sering digunakan dalam menentukan besaran indeks. Dengan demikian ke 25 butir soal ini dapat digunakan lebih lanjut baik secara keseluruhan maupun sebagian butir sesuai dengan tingkat kesukaran yang diinginkan oleh evaluator.

b. Hasil Ujicoba Berdasarkan Daya Bada Butir

Daya beda dalam penelitian ini mengukur sejauh mana butir tes membedakan antara siswa yang mengerjakan tes dengan benar secara keseluruhan dan siswa yang tidak menjawab tes dengan benar secara keseluruhan. Data hasil analisis daya beda butir instrumen dapat dikemukakan pada Tabel 5 berikut ini.

Tabel 5. Hasil analisis daya beda butir pada ujicoba

Daya Bada	Baik																									Cukup Baik			Kurang Baik	Buruk
	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	19	20	22	23	24	25	2	18	21	-	-			
Np	61	53	41	16	42	60	43	40	19	46	43	42	40	42	55	52	57	50	45	48	53	53	62	50	45					
p	0,94	0,82	0,63	0,25	0,65	0,92	0,66	0,62	0,29	0,71	0,66	0,65	0,62	0,65	0,85	0,80	0,88	0,88	0,69	0,74	0,82	0,82	0,95	0,77	0,69					
q	0,06	0,18	0,37	0,75	0,75	0,08	0,34	0,38	0,71	0,29	0,34	0,35	0,38	0,35	0,15	0,20	0,12	0,23	0,31	0,26	0,18	0,18	0,05	0,23	0,31					
p.q	0,06	0,25	0,23	0,19	0,19	0,07	0,22	0,24	0,21	0,21	0,22	0,23	0,24	0,23	0,13	0,16	0,11	0,18	0,21	0,19	0,15	0,15	0,04	0,18	0,21					
Jumlah Butir Soal	22																									3			0	0

Berdasarkan data hasil analisis daya beda butir pada ujicoba, maka dapat dikemukakan bahwa dari 25 butir instrumen hasil belajar matematika berbasis pendekatan kontekstual yang telah valid secara isi, sebanyak 22 butir soal berada pada kategori daya beda baik, 3 butir soal berada dalam kategori daya beda cukup baik, tidak ada soal yang berdaya beda kurang baik dan buruk. Dari hasil analisis tersebut menunjukkan bahwa ketiga butir soal yang berkategori daya beda cukup direvisi.

Secara teoritis dapat dijelaskan bahwa menentukan perbedaan kelompok dalam aspek yang diukur sesuai dengan perbedaan yang ada dalam kelompok itu merupakan salah satu tujuan analisis kuantitatif soal. Indeks yang digunakan dalam membedakan antara peserta tes yang berkemampuan rendah adalah indeks pembeda (*item discrimination*). Hal ini sejalan dengan pendapat Naga (2012), bahwa makin tinggi daya beda butir, maka makin besar perbedaan skor

yang dihasilkan oleh kelompok tinggi dan kelompok rendah. Fungsi dari daya pembeda tersebut adalah mendeteksi perbedaan individual yang sekecil-kecilnya di antara para peserta tes.

Jika temuan empirik tersebut dikaitkan dengan tinjauan teoretis, maka dapat disimpulkan bahwa instrumen hasil belajar matematika berbasis kontekstual yang dikembangkan telah mampu membedakan antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah. Dengan demikian instrumen hasil belajar matematika berbasis kontekstual yang dikembangkan telah memenuhi aspek kualitas daya beda butir.

c. Hasil Ujicoba Berdasarkan Efektivitas Pengecoh

Sebagaimana yang diketahui bahwa instrumen tes hasil belajar matematika berbasis kontekstual yang dikembangkan adalah bentuk objektif pilihan ganda. Oleh karena itu, salah satu aspek yang perlu diperhatikan adalah aspek efektivitas distraktor atau pengecoh. Efektivitas pengecoh diperiksa untuk melihat apakah semua distraktor atau semua pilihan jawaban yang bukan kunci telah berfungsi. Distraktor dinyatakan telah dapat menjalankan fungsinya dengan baik apabila distraktor tersebut sekurang-kurangnya sudah dipilih oleh 5% dari seluruh peserta tes. Berikut disajikan data hasil ujicoba berdasarkan aspek efektivitas pengecoh.

Tabel 6. Data hasil ujicoba berdasarkan efektivitas pengecoh

Keefektifan Pengecoh		Butir Yang Memiliki Pengecoh Berfungsi																									Butir Yang Memiliki Pengecoh Tidak Berfungsi		
		1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	19	20	22	23	24	25	2	18	21			
Np		61	53	41	16	42	60	43	40	19	46	43	42	40	42	55	52	57	50	45	48	53	53	62	50	45			
p		0.94	0.82	0.63	0.25	0.65	0.92	0.66	0.62	0.29	0.71	0.66	0.65	0.62	0.65	0.85	0.80	0.88	0.88	0.69	0.74	0.82	0.82	0.95	0.77	0.69			
q		0.06	0.18	0.37	0.75	0.75	0.08	0.34	0.38	0.71	0.29	0.34	0.35	0.38	0.35	0.15	0.20	0.12	0.23	0.31	0.26	0.18	0.18	0.05	0.23	0.31			
p.q		0.06	0.25	0.23	0.19	0.19	0.07	0.22	0.24	0.21	0.21	0.22	0.23	0.24	0.23	0.13	0.16	0.11	0.18	0.21	0.19	0.15	0.15	0.04	0.18	0.21			
Koef	Opsi																												
	A	0.07	0.07	0.14	0.07	0.14	0.07	0.14	0.07	0.14	0.07	0.50	0.43	0.11	0.18	0.29	0.07	0.14	0.68	0.11	0.07	0.43	0.07	0.036	0.286	0.179			
	B	0.07	0.57	0.46	0.61	0.07	0.11	0.29	0.21	0.07	0.11	0.07	0.11	0.18	0.21	0.14	0.57	0.07	0.11	0.07	0.07	0.07	0.18	0.571	0.17	0			
	C	0.71	0.11	0.11	0.14	0.50	0.64	0.07	0.14	0.07	0.54	0.07	0.14	0.11	0.21	0.11	0.11	0.18	0.07	0.11	0.68	0.32	0.61	0.036	0.143	0.143			
Status	D	0.07	0.11	0.07	0.07	0.14	0.07	0.18	0.07	0.07	0.11	0.29	0.18	0.43	0.11	0.36	0.18	0.04	0.07	0.54	0.07	0.07	0.07	0.321	0.037	0.464			
	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	TB	TB	B			
	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	TB			
	C	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	TB	TB	B			
Jumlah Butir Soal	D	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B			
		22 Butir																									3 Butir		

Keterangan

Berfungsi : B

Tidak Berfungsi : TB

Berdasarkan data hasil analisis efektivitas pengecoh dapat dikemukakan bahwa dari 25 butir soal telah valid secara isi, sebanyak 22 butir soal yang memiliki pengecoh berfungsi. Sebanyak 3 butir soal memiliki pengecoh kurang berfungsi. Butir tersebut adalah soal nomor 2 yang memiliki pengecoh yang tidak berfungsi pada opsi A dan C, soal nomor 18 memiliki pengecoh yang tidak berfungsi pada opsi B dan D, serta butir soal nomor 21 memiliki pengecoh yang tidak berfungsi pada opsi B. Secara teoretis pengecoh yang tidak berfungsi seharusnya ditinjau kembali dan diperbaiki sebelum digunakan lebih lanjut. Namun pada penelitian ini ketiga butir tersebut juga memiliki daya beda yang kategori cukup, sehingga disimpulkan bahwa butir tersebut direvisi saja.

Efektivitas pengecoh secara teoretis diperiksa untuk melihat apakah semua distraktor atau semua pilihan jawaban yang bukan kunci telah berfungsi sebagaimana mestinya, yaitu apakah distraktor-distraktor tersebut telah dipilih oleh lebih banyak (atau semua) siswa kelompok rendah sedangkan siswa dari kelompok tinggi hanya sedikit (atau tidak ada) yang memilihnya (Azwar, 2012). Lebih lanjut Kusaeri dan Suprananto (2012) mengemukakan bahwa suatu pilihan jawaban (pengecoh) dapat dikatakan berfungsi apabila: (1) pengecoh paling tidak dipilih oleh 5% peserta tes atau siswa; (2) pengecoh lebih banyak dipilih oleh kelompok siswa yang belum memahami materi yang diujikan. Hal semakna dikemukakan Ali & Khaeruddin (2012) bahwa distraktor sudah berfungsi dengan baik jika sudah dipilih paling sedikit 5% dari peserta tes ($p > 5\%$). Hasil ujicoba berdasarkan aspek efektivitas pengecoh, instrumen hasil belajar dalam penelitian ini telah memenuhi standar pengecoh yang berfungsi karena soal yang diujikan diperoleh nilai distraktor lebih besar dari 5% ($p > 5\%$).

Jika temuan empirik pada penelitian ini dikaitkan dengan aspek teoretis maka dapat dijelaskan bahwa pengecoh dari instrumen hasil belajar matematika berbasis kontekstual yang dikembangkan telah dapat digunakan untuk mengecoh peserta yang tidak menguasai materi. Hasil tersebut menunjukkan bahwa instrumen hasil belajar matematika berbasis kontekstual yang telah direvisi pada ujicoba telah baku berdasarkan aspek kualitas efektivitas pengecoh.

d. Hasil Ujicoba Berdasarkan Validitas Butir

Validitas butir secara empirik yang dimaksud dalam penelitian ini didasarkan pada kriteria internalnya dengan mengkorelasikan skor butir dengan skor totalnya. Suatu butir dikatakan valid apabila memiliki sumbangan yang besar terhadap skor totalnya dengan kriteria valid bila korelasi

$r > 30$ berdasarkan r tabel *product moment*. Berikut data hasil ujicoba instrumen ditinjau dari aspek validitas butir secara empirik.

Tabel 7. Data hasil ujicoba berdasarkan validitas butir

Validitas Butir	Butir Valid																									Butir Direvisi		
	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	19	20	22	23	24	25	2	18	21			
Np	61	53	41	16	42	60	43	40	19	46	43	42	40	42	55	52	57	50	45	48	53	53	62	50	45			
p	0,94	0,82	0,63	0,25	0,65	0,92	0,66	0,62	0,29	0,71	0,66	0,65	0,62	0,65	0,85	0,80	0,88	0,88	0,69	0,74	0,82	0,82	0,95	0,77	0,69			
q	0,06	0,18	0,37	0,75	0,75	0,08	0,34	0,38	0,71	0,29	0,34	0,35	0,38	0,35	0,15	0,20	0,12	0,23	0,31	0,26	0,18	0,18	0,05	0,23	0,31			
p.q	0,06	0,25	0,23	0,19	0,19	0,07	0,22	0,24	0,21	0,21	0,22	0,23	0,24	0,23	0,13	0,16	0,11	0,18	0,21	0,19	0,15	0,15	0,04	0,18	0,21			
Koef	0,71	0,721	0,6	0,5	0,9	0,7	0,5	0,8	0,4	0,9	0,7	0,5	0,6	0,4	0,6	0,7	0,4	0,4	0,4	0,5	0,6	0,5	0,004	0,075	0,1			
Jumlah Butir Soal	22 Butir																									3 Butir		

Berdasarkan data hasil ujicoba, maka diperoleh informasi bahwa sebanyak 22 butir yang telah valid secara empirik. 3 butir soal yang direvisi karena tidak valid. Butir soal yang direvisi tersebut ialah butir 2, butir 18, dan butir 21. Jika ditinjau dari aspek teoretis, Djaali dan Pudji (2008), Indriati (2016), Solichin (2017) dan Syamsurizal (2020) menyatakan bahwa validitas empirik sama dengan validitas kriteria dimana validitas tersebut ditentukan oleh kriteria baik secara internal maupun eksternal. Kriteria internal adalah instrumen tes itu sendiri dan kriteria eksternal adalah hasil dari pengukuran tes lain.

Yang dimaksud validitas konsistensi internal dalam penelitian ini ialah validitas melalui kriteria internal. Mansyur, Harun dan Suratno (2009) bahwa prosedur validitas butir dengan konsistensi internal dapat dilakukan dengan mengkorelasikan skor butir dengan skor totalnya. Berdasarkan tinjauan teoretis tersebut maka ke 22 butir instrumen hasil belajar matematika berbasis kontekstual yang dikembangkan telah memenuhi kriteria internalnya. Dengan demikian instrumen hasil belajar matematika berbasis kontekstual telah baku berdasarkan aspek kualitas validitas butir. Instrumen tersebut dapat mengukur kompetensi dasar yang hendak diukur.

e. Hasil Ujicoba Berdasarkan Reliabilitas tes

Pada penelitian ini, kriteria reliabilitas instrumen yang digunakan untuk menentukan instrumen tes hasil belajar matematika reliabel atau tidak adalah koefisien r minimal 0,7. Data hasil analisis reliabilitas instrumen pada ujicoba dapat dilihat pada tabel berikut.

Tabel 8. Data hasil ujicoba berdasarkan reliabilitas tes

Hasil Uji Coba Instrumen Berdasarkan	Nomor Item																												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	Xt	Xt ²		
Np	61	62	53	41	16	42	60	43	40	19	46	43	42	40	42	55	52	50	57	50	45	45	48	53	53	1163	22291		
p	0,94	0,95	0,82	0,63	0,25	0,65	0,92	0,66	0,62	0,29	0,71	0,66	0,65	0,62	0,65	0,85	0,80	0,77	0,88	0,88	0,69	0,69	0,74	0,82	0,82				
q	0,06	0,05	0,18	0,37	0,75	0,75	0,08	0,34	0,38	0,71	0,29	0,34	0,35	0,38	0,35	0,15	0,20	0,23	0,12	0,23	0,31	0,31	0,26	0,18	0,18				
p.q	0,06	0,04	0,25	0,23	0,19	0,19	0,07	0,22	0,24	0,21	0,21	0,22	0,23	0,24	0,23	0,13	0,16	0,18	0,11	0,18	0,21	0,21	0,19	0,15	0,15	$\Sigma p.q$	4,40		

Berdasarkan hasil penelitian secara empirik diperoleh koefisien reliabilitas sebesar 0,820. Secara teoretis hasil pengukuran dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subjek yang sama diperoleh hasil yang relatif sama, selama aspek yang diukur dalam diri subjek memang belum berubah (Azwar, 2012). Dalam pengertian ini, relatif sama berarti tetap adanya toleransi terhadap perbedaan-perbedaan kecil diantara hasil beberapa kali pengukuran. Bila perbedaan itu sangat besar dari waktu ke waktu maka hasil pengukuran tidak dapat dipercaya dan dikatakan sebagai tidak reliabel. Mansyur, Harun dan Suratno (2009) bahwa sifat reliabel (keterandalan) dari sebuah alat ukur berkenaan dengan kemampuan alat ukur tersebut memberikan hasil yang konsisten dan stabil bila dilakukan pengukuran yang berulang dalam waktu yang berbeda. Untuk itu hasil pengukuran instrumen hasil belajar matematika pada penelitian ini telah konsisten dan dapat dipercaya.

D. Simpulan

Selain itu, instrumen hasil belajar matematika berbasis kontekstual ini telah baku dimana sebanyak 2 butir soal berkategori sulit, 9 butir soal sedang, dan 14 butir soal mudah. Jika ditinjau dari aspek daya butir, terdapat 22 butir yang telah memiliki daya beda butir yang baik dan 3 butir soal memiliki daya beda cukup sehingga soal tersebut harus direvisi kembali. terdapat 22 butir soal yang memiliki pengecoh yang berfungsi dan 3 butir soal lainnya memiliki pengecoh yang kurang berfungsi untuk semua butir soal. Jika ditinjau dari aspek validitas butirnya, terdapat 22 butir soal yang valid secara empirik sementara 3 butir soal lainnya direvisi karena tidak valid. Untuk aspek reliabilitasnya, instrumen tes hasil belajar ini telah reliabel dengan koefisien reliabilitas 0,820. Dengan demikian, instrumen tes hasil belajar matematika layak digunakan.

Daftar Pustaka

- Agustan, S. (2016). Usage of Contextual Approach to Increase Student's Understanding in Learning Mathematics. *IOSR Journal of Mathematics*, 12(6), 118-122.
- Agustan, S., Juniati, D., & Siswono, T. Y. E. (2017). Profile of Male-Field Dependent (FD) Prospective Teacher's Reflective Thinking in Solving Contextual Mathematical Problem. *AIPC*, 1867(1), 020034.
- Ali, S. & Khaeruddin. (2012). *Evaluasi Pembelajaran*. Makassar: Badan Penerbit UNM.
- Arifin, Z. (2017). *Evaluasi Pembelajaran: Prinsip Teknik Prosedur, Cetakan ke-10*. Bandung: PT. Remaja Rosdakarya.

- Azwar, S. (2012). *Reliabilitas dan Validitas edisi 4*. Yogyakarta: Pustaka Pelajar.
- Departemen Pendidikan Nasional. 2005. *Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan*. Jakarta: Balai Pustaka.
- Djaali, P. M., & Muljono, P. (2008). *Pengukuran dalam Bidang Pendidikan*. Jakarta: Grasindo.
- Indriati, N. U. D. (2016). Instrumen Tes Kecerdasan Logika-Matematika Untuk Anak Usia Dini. *Jurnal Pendidikan Usia Dini*, 10(1), 171-184.
- Kharie, E. S., La Nani, K, & Tonra, W. S. (2021). Pengembangan Lembar Kerja Peserta Didik (LKPD) Berbasis Problem Based Instruction pada Materi Persamaan Linear Satu Variabel. *Jurnal Pendidikan Guru Matematika*, 1(1), 13-23
- Kusaeri, K., & Suprananto, S. (2012). *Pengukuran Dan Penilaian Pendidikan*. Jakarta: Graha Ilmu.
- Ma'rufi, Ilyas, M., Salwah, & Syamsuddin, A. (2020). ADD-CoRE Model Development: Mathematics Teachers Mentoring Based on Pedagogical Content Knowledge and Lesson Study. *Universal Journal of Educational Research (UJER)*. 8(10), 4580–4590. <https://doi.org/10.13189/ujer.2020.081026>
- Mansyur, H. R. (2009). *Asesmen Pembelajaran di Sekolah*, Yogyakarta: Multi Pressindo.
- Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta. Mitra Cendikia.
- Mardapi, D. (2012). *Pengukuran penilaian dan Evaluasi Pendidikan*. Yogyakarta: Nuha Medika.
- Naga, D. S. (2012). *Teori Sekor pada Pengukuran Mental*. Jakarta: PT Nagarani Citrayasa.
- Situmorang, A. S. (2013). Implementasi Pendekatan Pembelajaran Kontekstual Dalam Pengajaran Matematika di SD dengan Menggunakan LKS Berbasis Masalah. *Jurnal Penelitian Bidang Pendidikan*, 20(1), 76-82.
- Syamsurizal, S. (2020). Validitas dan Reliabilitas Alat Ukur. <https://doi.org/10.31219/osf.io/v83eh>
- Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes Dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirasat: Jurnal Manajemen dan Pendidikan Islam*, 2(2), 192-213.
- Sudibyo, E., Jatmiko, B., & Widodo, W. (2017). Pengembangan Instrumen Motivasi Belajar Fisika: Angket. *Jurnal Penelitian Pendidikan IPA*, 1(1), 13-21.
- Sulianto, J. (2008). Pendekatan Kontekstual Dalam Pembelajaran Matematika untuk Meningkatkan Berpikir Kritis pada Siswa Sekolah Dasar. *Pythagoras: Jurnal Pendidikan Matematika*, 4(2), 14-25.
- Sumarno, A. (2012). *Penelitian Kausalitas Komparatif*. Surabaya: elearning UNESA.
- Supardi, U. S. (2015). Peran Berpikir Kreatif dalam Proses Pembelajaran Matematika. *Formatif: Jurnal Ilmiah Pendidikan MIPA*, 2(3), 248-262.

- Surapranata, S. (2004). *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes*. Bandung: Remaja Rosdakarya.
- Susetyo, B. (2015). *Prosedur Penyusunan dan Analisis Tes Untuk Penilaian Hasil Belajar Bidang Kognitif*. Bandung: PT. Refika Aditama.
- Syamsuddin, A. (2019). The Impact of Implementing Of INSTAD Model toward Student's Mathematics Learning Outcome for 5th Grade Elementary School Students. In *Journal of Physics: Conference Series* (Vol. 1318, No. 1, p. 012021). IOP Publishing.
- Syamsuddin, A. (2020). Describing Taxonomy of Reflective Thinking For Field Dependent-Prospective Mathematics Teacher in Solving Mathematics Problem. *International Journal of Scientific & Technology Research*, 9(3), 4418-4421.
- Syamsuddin, A., Jannah, M., & Kristiawati, K. (2019). Penerapan Model Explicit Instruction dalam Pembelajaran Matematika Materi Bilangan Romawi pada Siswa Kelas IV SD Inpres Kapasa Makassar. *MaPan: Jurnal Matematika dan Pembelajaran*, 7(1), 136-154.
- Thiagarajan, S., Semmel, D. S & Semmel, M. I. (1974). *Instructional Development for Training Teachers of Expectional Children*. Minneapolis, Minnesota: Leadership Training Institute/Special Education, University of Minnesota
- Widoyoko, E. P. (2009). *Evaluasi Program Pembelajaran (Panduan Praktis Bagi Pendidik dan Calon Pendidik)*. Yogyakarta: pustaka pelajar.
- Wirawan. (2016). *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi, dan Profesi*. Jakarta: Rajagrafindo Persada.
- Wiryokusumo, I., & Mulyadi, U. (1988). *Dasar-Dasar Pengembangan Kurikulum*. Jakarta: Bina Aksara.