Development of A Diagnostic Assessment Instrument For Mathematics Learning For Grade II Elementary School

Sonia Savera Oping[1], Juliana Margareta Sumilat[2], Kartini Ester[3]

[1,2,3]Program Studi Pendidikan Guru Sekolah Dasar, Fakultas Ilmu Pendidikan dan Psikologi, Universitas Negeri Manado, Indonsia
Coresponden author:
Alamat e-mail: soniasaveraoping@gmail.com

Abstract.
This study aims to develop a valid diagnostic assessment instrument for second-grade elementary school mathematics (addition and subtraction of whole numbers). Existing diagnostic instruments mostly target other grade levels or focus on misconceptions, while lower-grade teachers still need a ready-to-use instrument that links students' answers to coding, interpretation, and follow-up actions. This research and development study used a 3D model (Define, Design, Develop) modified from Thiagarajan's 4D model. The product is a 12-item instrument (5 multiple-choice and 7 short-answer items; maximum score 19) with a test blueprint, answer keys, coding guidelines, interpretation, and follow-up plans. Validation by two expert lecturers and one classroom teacher yielded an average validity of 86.67% (very valid). A small-group trial with five students produced a mean score of 83.16; two students were in the "full understanding" category and three in the "partial understanding" category. The least mastered indicators were addition with regrouping and subtraction with borrowing (two of five students each). The instrument is valid and feasible as a teacher's tool for mapping students' initial understanding; its effectiveness and psychometric properties need to be tested with larger samples.
Keywords: *diagnostic assessment instrument, mathematics learning, 3D model, elementary school and grade II.*

I. PENDAHULUAN

Penerapan Kurikulum Merdeka dirancang untuk menciptakan pembelajaran yang lebih kontekstual dan berpusat pada kebutuhan siswa, dengan guru sebagai fasilitator (Mustofa & Mariati, 2022). Namun, implementasinya di lapangan masih menghadapi kendala, terutama pada kesiapan guru merancang instrumen penilaian dan perangkat pembelajaran yang sesuai dengan prinsip pembelajaran berdiferensiasi (Sumarsih dkk., 2022; Wantiana & Mellisa, 2023). Di kelas II, kendala ini terasa penting karena siswa baru membangun fondasi operasi hitung. Penelitian di kelas II menunjukkan bahwa penjumlahan dan pengurangan masih menimbulkan kesulitan bagi siswa, sehingga guru perlu mengenali sumber kesulitan tersebut sejak awal (Panggabean dkk., 2026). Pembelajaran penjumlahan dan pengurangan di kelas rendah juga banyak dikaji melalui penggunaan media dan bahan ajar (Anajjah dkk., 2022; Dethan dkk., 2026), tetapi kajian tersebut berfokus pada perlakuan pembelajaran, bukan pada alat untuk memetakan pemahaman siswa sebelum perlakuan diberikan.

Asesmen diagnostik berfungsi mendeteksi kompetensi awal, kesiapan, dan kesulitan belajar siswa sebelum materi baru diajarkan (Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi [Kemendikbudristek], 2022). Informasi dari asesmen kelas memungkinkan guru menyesuaikan strategi dan memberikan dukungan sesuai kebutuhan siswa (Brookhart, 2023). Kerangka Understanding by Design menekankan keselarasan antara tujuan, aktivitas belajar, dan bukti penilaian (Wiggins & McTighe, 2005); dalam kerangka ini asesmen diagnostik berperan menyediakan bukti kondisi awal siswa sebelum aktivitas belajar dirancang. Penerapan prinsip tersebut pada pembelajaran matematika sekolah dasar juga dilaporkan oleh Mongkau dan Sumilat (2025). Tarigan dkk. (2026) menegaskan bahwa pembelajaran matematika bertujuan mendukung siswa mengembangkan kompetensi berpikir secara sistematis dan logis. Secara perkembangan, siswa kelas II berada pada tahap operasional konkret sehingga penalaran numeriknya bertumpu pada objek dan representasi yang nyata (Fahma & Purwaningrum, 2021), dan perkembangan konsep bilangan mengikuti lintasan belajar yang berurutan (Clements & Sarama, 2020). Kedua pandangan ini mendasari penyusunan butir soal dari konsep nilai tempat, penjumlahan, hingga pengurangan.

Penelitian tentang asesmen diagnostik matematika terus berkembang. Kajian pustaka sistematis menunjukkan bahwa penelitian pengembangan asesmen diagnostik matematika bertambah, tetapi masih memerlukan penguatan pada inovasi, implementasi jangka panjang, dan integrasi teknologi (Andreansyah dkk., 2026); kajian lain menyoroti tantangan penerapan asesmen diagnostik untuk mengidentifikasi kesulitan

belajar (Syah & Febriola, 2026) serta pengalaman guru dalam mengelolanya (Aulia dkk., 2026). Pada tataran instrumen, telah dikembangkan tes diagnostik penjumlahan dan pengurangan dasar untuk siswa sekolah dasar (Homjan dkk., 2022), instrumen diagnostik three-tier multiple choice untuk bilangan cacah di kelas IV (Aprianti & Suminar, 2025), soal strategi hitung mental untuk Fase A (Mulyani dkk., 2026), serta penerapan asesmen diagnostik untuk mengidentifikasi kesulitan operasi hitung (Saleh dkk., 2026).

Berdasarkan penelusuran terhadap penelitian-penelitian tersebut, terdapat tiga celah. Pertama, sejumlah penelitian berfokus pada jenjang, bentuk instrumen, atau tujuan yang berbeda, misalnya kelas IV dan miskonsepsi (Aprianti & Suminar, 2025) atau strategi hitung mental (Mulyani dkk., 2026). Kedua, sebagian kajian berupa tinjauan pustaka atau analisis kesulitan, bukan pengembangan instrumen (Andreansyah dkk., 2026; Syah & Febriola, 2026; Aulia dkk., 2026). Ketiga, dari penelitian yang dirujuk dalam artikel ini, tidak ada yang secara eksplisit memadukan kisi-kisi, kemungkinan jawaban siswa, pengodean, interpretasi, dan tindak lanjut pembelajaran dalam satu paket bagi guru kelas II. Di sekolah tempat penelitian, peneliti mengamati bahwa guru kelas II belum memiliki perangkat asesmen diagnostik matematika yang siap pakai.

Kebaruan penelitian ini terletak pada penyajian instrumen asesmen diagnostik penjumlahan dan pengurangan bilangan cacah kelas II sebagai satu paket yang dilengkapi kisi-kisi, kemungkinan jawaban siswa, pengodean, interpretasi, dan alternatif tindak lanjut pembelajaran.

Berdasarkan uraian tersebut, rumusan masalah penelitian ini adalah: (1) bagaimana proses pengembangan instrumen asesmen diagnostik pembelajaran matematika kelas II sekolah dasar melalui tahap Define, Design, dan Develop; (2) bagaimana kevalidan instrumen berdasarkan penilaian dosen ahli dan guru praktisi; dan (3) bagaimana temuan diagnostik instrumen pada uji coba kelompok kecil. Sejalan dengan itu, penelitian ini bertujuan (1) mendeskripsikan proses pengembangan instrumen tersebut, (2) menganalisis kevalidannya, dan (3) menganalisis temuan diagnostiknya pada uji coba kelompok kecil. Hasilnya diharapkan membantu guru memetakan kemampuan awal dan kesulitan belajar siswa, memberi siswa umpan balik yang jelas, serta menjadi rujukan bagi sekolah dan calon guru dalam menganalisis hasil asesmen untuk merancang pembelajaran yang tepat sasaran.

II. METODE PENELITIAN

Penelitian ini merupakan penelitian dan pengembangan (R&D) yang menghasilkan instrumen asesmen diagnostik pembelajaran matematika kelas II SD. Model yang digunakan adalah modifikasi model 4D Thiagarajan dkk. (1974) menjadi tiga tahap, yaitu Define (pendefinisian), Design (perancangan), dan Develop (pengembangan), seperti pada Gambar 1. Tahap Disseminate tidak dilaksanakan karena keterbatasan waktu dan sumber daya, sehingga penelitian ini berfokus pada pengembangan awal, validasi, revisi produk, dan uji coba kelompok kecil. Penelitian dilaksanakan di SD Katolik Santu Petrus Keroit; validasi dilakukan pada 11 Mei 2026 dan uji coba kelompok kecil pada 18 Mei 2026. Penyusunan produk dan validasi dilakukan bersama dosen Program Studi PGSD Universitas Negeri Manado.



Fig. 1. Alur Penelitian dan Pengembangan Model 3D (dimodifikasi dari Thiagarajan dkk., 1974)

Prosedur penelitian terdiri atas tiga tahap. Pada tahap Define dilakukan analisis awal-akhir, karakteristik peserta didik, konsep, tugas, dan tujuan pembelajaran. Pada tahap Design disusun kisi-kisi berdasarkan capaian dan tujuan pembelajaran, 12 butir soal yang terdiri atas lima pilihan ganda dan tujuh isian singkat, kunci jawaban, kemungkinan jawaban siswa, pedoman pengodean, interpretasi hasil, serta alternatif tindak lanjut.

Pada tahap Develop, produk divalidasi oleh tiga validator, direvisi berdasarkan masukan validator, kemudian diujicobakan dalam kelompok kecil.

Subjek penelitian meliputi tiga validator, yaitu satu dosen ahli materi matematika SD, satu dosen ahli evaluasi/media pembelajaran, dan satu guru kelas II SD Katolik Santu Petrus Keroit. Uji coba kelompok kecil melibatkan lima siswa kelas II di sekolah yang sama dengan kemampuan akademik yang beragam. Validasi dilaksanakan terlebih dahulu, kemudian produk direvisi berdasarkan masukan validator, dan selanjutnya diujicobakan pada kelompok kecil pada 18 Mei 2026. Identitas siswa dirahasiakan dan ditulis sebagai Siswa 1 sampai Siswa 5 (S1–S5).

Data kevalidan dikumpulkan menggunakan angket validasi 14 pernyataan dengan tiga aspek, yaitu kelayakan isi (5 pernyataan), kelayakan penyajian (4 pernyataan), dan komponen asesmen diagnostik (5 pernyataan: kisi-kisi, kemungkinan jawaban, pengodean, interpretasi, dan tindak lanjut), disertai ruang saran perbaikan. Setiap pernyataan diskor 1–5 (5 = Sangat Setuju/Sangat Baik, 4 = Setuju/Baik, 3 = Ragu-ragu/Cukup, 2 = Tidak Setuju/Kurang, 1 = Sangat Tidak Setuju/Sangat Kurang) sehingga skor maksimum ideal per validator adalah 70.

Skor validasi dianalisis dengan rumus berikut, kemudian diinterpretasikan memakai kriteria pada Tabel 1. Persentase yang jatuh pada rentang 81–100% berkategori Sangat Valid dan 61–80% berkategori Valid. Masukan tertulis validator dianalisis secara deskriptif dan menjadi dasar revisi produk.

Persentase Kelayakan (%) = (Skor yang Diperoleh ÷ Skor Maksimum Ideal) × 100

Tabel 1. Kriteria Kelayakan Instrumen (Arikunto & Jabar, 2004)

Persentase Kelayakan	Interpretasi
81–100%	Sangat Layak (Sangat Valid)
61–80%	Layak (Valid)
41–60%	Cukup Layak
21–40%	Tidak Layak
<21%	Sangat Tidak Layak

Skor siswa dihitung dari lima butir pilihan ganda (skor 1 untuk jawaban benar) dan tujuh butir isian singkat (skor 2 untuk jawaban benar, 1 untuk kesalahan kecil, dan 0 untuk jawaban lain), sehingga skor maksimum keseluruhan adalah 5 × 1 + 7 × 2 = 19. Nilai akhir dihitung dengan rumus (skor yang diperoleh ÷ 19) × 100. Kategori pemahaman ditetapkan sebagai berikut: Paham Utuh apabila nilai akhir ≥ 90; Paham Sebagian apabila 60–89; dan Tidak Paham apabila < 60.

Selain nilai akhir, jawaban siswa dikodekan sesuai pedoman pengodean instrumen, yaitu P (paham) untuk jawaban benar, PS (paham sebagian) untuk kesalahan kecil, dan TP (tidak paham) untuk jawaban salah atau jauh dari benar. Pada uji coba ini tidak ada jawaban yang berkode PS, sehingga temuan per indikator hanya memuat kode P dan TP. Untuk indikator yang diukur oleh dua butir, siswa dikodekan P hanya apabila kedua butir berkode P; kesalahan pada satu butir saja sudah dikodekan TP. Jumlah siswa berkode P dan TP pada setiap indikator kemudian dijadikan dasar menentukan indikator yang memerlukan tindak lanjut.

III. Hasil Penelitian dan Pembahasan

C.1 Hasil Penelitian

Tahap Define

Analisis awal-akhir diarahkan untuk mengidentifikasi kesenjangan antara kebutuhan pembelajaran matematika kelas II dan ketersediaan instrumen asesmen diagnostik di sekolah. Peneliti mengamati bahwa guru kelas II belum memiliki perangkat asesmen diagnostik penjumlahan dan pengurangan yang dilengkapi pedoman penafsiran hasil. Analisis karakteristik peserta didik menunjukkan bahwa kemampuan awal siswa kelas II beragam. Analisis konsep menghasilkan urutan materi berupa nilai tempat, penjumlahan tanpa menyimpan, penjumlahan dengan menyimpan, dan pengurangan bilangan cacah. Analisis tugas dan perumusan tujuan pembelajaran menjadi dasar penentuan indikator dan bentuk soal.

Tahap Design

Produk awal berupa instrumen 12 butir (5 pilihan ganda dan 7 isian singkat; skor maksimum 19) beserta kunci jawaban, kemungkinan jawaban siswa, pedoman pengodean, interpretasi, dan rencana tindak lanjut. Sebaran butir terhadap indikator disajikan pada Tabel 2, dan contoh butir serta komponen pendukungnya pada Tabel 3.

Tabel 2. Kisi-kisi Ringkas Instrumen Asesmen Diagnostik

No	Indikator Soal	Bentuk	Nomor Soal	Skor Maks.
1	Menentukan nilai tempat bilangan dua digit	Isian singkat	6	2
2	Menyelesaikan penjumlahan tanpa menyimpan	Isian singkat	7, 8	4
3	Menyelesaikan penjumlahan dengan menyimpan	Isian singkat	9, 10	4
4	Menyelesaikan soal cerita penjumlahan	Pilihan ganda	2, 5	2
5	Menyelesaikan soal cerita pengurangan (dengan meminjam)	Pilihan ganda	1	1
6	Menyelesaikan pengurangan dengan meminjam	Pilihan ganda	3, 4	2
7	Menyelesaikan pengurangan tanpa meminjam	Isian singkat	11, 12	4
	Jumlah		**12 butir**	**19**

Tabel 3. Contoh Butir, Kunci, Kemungkinan Jawaban, Pengodean, dan Tindak Lanjut

Butir	Kunci	Kemungkinan jawaban salah dan artinya	Interpretasi bila benar (P)	Tindak lanjut bila salah (TP)
1. Edo punya 20 balon, 8 terbang. Sisa balon?	C (12)	28: menjumlah, bukan mengurang	Siswa menghubungkan cerita dengan pengurangan	Remedial dengan gambar dan diskusi
2. Ibu membeli 13 telur, lalu 7 lagi. Jumlah?	A (20)	6: mengurang, bukan menambah	Siswa memahami konteks penjumlahan	Remedial dengan benda konkret
3. 15 − 7 = …	C (8)	Jawaban selain 8 (mis. 22 karena menjumlah), menunjukkan kesalahan prosedur meminjam	Siswa paham pengurangan dengan meminjam	Remedial dengan model konkret
6. Nilai tempat angka 4 pada 47	4 puluhan / 40	Menjawab 7 puluhan atau tidak sesuai	Siswa paham nilai tempat	Remedial dengan keping/batang/kartu angka
9. 15 + 5 = …	20	10: mengurang, bukan menjumlah	Siswa paham teknik menyimpan	Remedial dengan media konkret
12. 15 − 5 = …	10	6: kesalahan hitung dasar	Siswa paham pengurangan tanpa meminjam	Pendekatan konkret dan latihan bertahap

Tabel 3 hanya memuat contoh butir; seluruh 12 butir terdapat pada produk instrumen.

Tahap Develop

Validasi dilakukan oleh tiga validator pada tiga aspek (Tabel 4). Rata-rata kevalidan keseluruhan adalah 86,67% (Sangat Valid), dengan rata-rata kelayakan isi 86,67%, penyajian 85,00%, dan komponen asesmen diagnostik 88,00%. Validator I dan III masing-masing memberi 90% (Sangat Valid), sedangkan Validator II memberi 80%, yang menurut Tabel 1 berkategori Valid. Validator I adalah dosen ahli materi, Validator II dosen ahli evaluasi/media, dan Validator III guru kelas II.

Tabel 4. Hasil Validasi Ahli per Aspek

Aspek (skor maks.)	Validator I	Validator II	Validator III (Guru)	Rata-rata (%)	Kategori
Kelayakan Isi (25)	23 (92%)	20 (80%)	22 (88%)	86,67	Sangat Valid
Kelayakan Penyajian (20)	18 (90%)	16 (80%)	17 (85%)	85,00	Sangat Valid
Komponen Asesmen Diagnostik (25)	22 (88%)	20 (80%)	24 (96%)	88,00	Sangat Valid
Total (70)	**63 (90%)**	**56 (80%)**	**63 (90%)**	**86,67**	**Sangat Valid**

Catatan: kategori individual Validator II adalah Valid (80%); kategori pada kolom terakhir merujuk pada rata-rata antarvalidator.

Masukan validator dan revisi yang dilakukan disajikan pada Tabel 5. Validator II memberi penilaian paling rendah di antara ketiga validator (80% pada setiap aspek), dengan masukan yang berkaitan dengan urutan dan ragam soal.

Tabel 5. Masukan Validator dan Revisi Produk

Validator	Masukan	Revisi (sebelum → sesudah)
Validator I (dosen ahli materi)	Judul perangkat diganti dari "Modul" menjadi "Instrumen".	Judul diubah dari Modul Asesmen Diagnostik menjadi Instrumen Asesmen Diagnostik.
Validator II (dosen ahli evaluasi/media)	Soal pilihan ganda disajikan lebih dahulu, diikuti isian singkat, dan ragam soal dibuat lebih bervariasi.	Soal disusun ulang dengan pilihan ganda di bagian awal dan isian singkat di bagian akhir. Soal cerita ditambahkan agar ragam soal lebih bervariasi.
Validator III (guru kelas II)	Bilangan pada soal operasi hitung dibatasi sampai 20 sesuai kemampuan siswa kelas II.	Soal operasi hitung (penjumlahan dan pengurangan, termasuk soal cerita) diganti agar bilangan dan hasil pengerjaannya tidak melebihi 20.

Hasil uji coba kelompok kecil. Setelah direvisi, instrumen diujicobakan pada lima siswa kelas II SD Katolik Santu Petrus Keroit pada 18 Mei 2026. Hasilnya disajikan pada Tabel 6.

Tabel 6. Hasil Uji Coba Kelompok Kecil

Siswa	Skor (maks. 19)	Nilai Akhir	Kategori Pemahaman
S1	19	100	Paham Utuh
S2	13	68,42	Paham Sebagian
S3	19	100	Paham Utuh
S4	15	78,95	Paham Sebagian
S5	13	68,42	Paham Sebagian
Rata-rata	**15,80**	**83,16**	**2 Utuh (40%); 3 Sebagian (60%)**

Berdasarkan data pada Tabel 6, uji coba kelompok kecil pada 18 Mei 2026 di SD Katolik Santu Petrus Keroit menghasilkan rata-rata nilai 83,16. Dua siswa (S1 dan S3; 40%) memperoleh nilai 100 dan berada pada kategori Paham Utuh, sedangkan tiga siswa (60%) berada pada kategori Paham Sebagian, yaitu S4 dengan nilai 78,95 serta S2 dan S5 dengan nilai masing-masing 68,42. Tidak ada siswa pada kategori Tidak Paham.

Dengan lima siswa dan tanpa analisis psikometrik, hasil ini belum cukup untuk menyimpulkan efektivitas atau sensitivitas instrumen. Keterbacaan instrumen oleh siswa juga belum diuji secara terpisah, sehingga variasi nilai belum dapat dipastikan semata-mata mencerminkan perbedaan pemahaman.

Kode jawaban per butir untuk setiap siswa disajikan pada Tabel 7. Butir 1–5 berbentuk pilihan ganda dan butir 6–12 berbentuk isian singkat. Tidak ada jawaban berkode PS pada uji coba ini, sehingga hanya terdapat kode P (paham) dan TP (tidak paham).

Tabel 7. Matriks Kode Jawaban Siswa per Butir

Siswa	1	2	3	4	5	6	7	8	9	10	11	12
S1	P	P	P	P	P	P	P	P	P	P	P	P
S2	P	TP	TP	P	P	TP	P	P	TP	P	P	P

Siswa	1	2	3	4	5	6	7	8	9	10	11	12
S3	P	P	P	P	P	P	P	P	P	P	P	P
S4	P	P	P	P	P	P	TP	TP	P	P	P	P
S5	P	P	TP	TP	P	P	P	P	TP	P	P	TP
Jumlah TP	**0**	**1**	**2**	**1**	**0**	**1**	**1**	**1**	**2**	**0**	**0**	**1**

Catatan: P = paham; TP = tidak paham. Butir 1–5 pilihan ganda (skor 1), butir 6–12 isian singkat (skor 2). Skor S2 = 19 − (2 + 4) = 13; S4 = 19 − 4 = 15; S5 = 19 − (2 + 4) = 13; S1 dan S3 = 19.

Temuan diagnostik per indikator disajikan pada Tabel 8 berdasarkan kode P (paham) dan TP (tidak paham) sesuai aturan pengodean pada bagian metode. Indikator yang diukur oleh dua butir dikodekan TP apabila siswa salah pada salah satu butirnya.

Tabel 8. Temuan Diagnostik per Indikator pada Uji Coba

Indikator	No. Soal	Jumlah siswa (n = 5)	Temuan dan interpretasi
Nilai tempat	6	P: 4; TP: 1	Empat siswa menentukan nilai tempat dengan tepat. Satu siswa (S2) masih memerlukan bimbingan dalam membedakan nilai tempat dan nilai angka.
Penjumlahan tanpa menyimpan	7, 8	P: 4; TP: 1	Empat siswa mampu menyelesaikan penjumlahan dasar tanpa menyimpan dengan benar. Satu siswa (S4) salah pada kedua butir sehingga masih memerlukan penguatan.
Penjumlahan dengan menyimpan	9, 10	P: 3; TP: 2	Dua siswa (S2 dan S5) salah pada butir 9, yang menunjukkan prosedur menyimpan belum lancar atau masih keliru dipahami.
Soal cerita penjumlahan	2, 5	P: 4; TP: 1	Empat siswa memilih operasi penjumlahan yang tepat. Satu siswa (S2) keliru pada butir 2, sehingga kemampuan menerjemahkan soal cerita ke bentuk matematika masih bervariasi.
Soal cerita pengurangan (dengan meminjam)	1	P: 5; TP: 0	Seluruh siswa menjawab butir 1 dengan tepat, sehingga kemampuan menghubungkan cerita dengan operasi pengurangan telah dikuasai kelompok ini.
Pengurangan dengan meminjam (pilihan ganda)	3, 4	P: 3; TP: 2	Dua siswa (S2 dan S5) belum tepat menerapkan prosedur meminjam pada butir hitung langsung (mis. 15 − 7); S5 salah pada kedua butir. Tiga siswa lainnya menjawab dengan tepat.
Pengurangan tanpa meminjam (isian singkat)	11, 12	P: 4; TP: 1	Empat siswa menjawab dengan tepat. Satu siswa (S5) belum tepat pada butir 12 (15 − 5), sehingga ketelitian pengerjaan perlu ditingkatkan.

Berdasarkan Tabel 8, indikator dengan jumlah siswa TP terbanyak adalah penjumlahan dengan menyimpan dan pengurangan dengan meminjam, masing-masing dua siswa. Pola ini dapat menjadi dasar guru memprioritaskan penguatan pada dua indikator tersebut, yang keduanya melibatkan prosedur menyimpan atau meminjam antarnilai tempat.

C.2 Pembahasan

Urutan materi yang dihasilkan pada tahap Define, yaitu dari nilai tempat ke penjumlahan lalu pengurangan, sejalan dengan pandangan bahwa konsep bilangan berkembang melalui lintasan belajar yang berurutan (Clements & Sarama, 2020) dan dengan temuan Mongkau dan Sumilat (2025) tentang perlunya pembelajaran matematika yang dirancang bertahap. Pemilihan bentuk pilihan ganda dan isian singkat dengan bahasa sederhana sesuai dengan karakteristik siswa kelas II yang berada pada tahap operasional konkret (Fahma & Purwaningrum, 2021). Keselarasan antara tujuan pembelajaran, kisi-kisi, butir soal, dan tindak lanjut mencerminkan prinsip keselarasan tujuan dan bukti penilaian dalam Understanding by Design (Wiggins & McTighe, 2005).

Rata-rata validasi sebesar 86,67% berada pada kategori sangat valid. Aspek komponen asesmen diagnostik memperoleh rata-rata tertinggi (88,00%), disusul kelayakan isi (86,67%) dan penyajian (85,00%). Penyajian

memiliki rata-rata terendah, sedangkan Validator II memberi skor terendah di antara validator (80%) pada ketiga aspek; masukannya mengenai urutan dan ragam soal telah diakomodasi pada revisi (Tabel 5).

Nilai siswa bervariasi dari 68,42 hingga 100. Variasi ini menunjukkan bahwa instrumen mampu membedakan tingkat pemahaman siswa pada kelompok kecil ini, tetapi belum dapat dipastikan semata-mata mencerminkan perbedaan pemahaman karena keterbacaan instrumen belum diuji secara terpisah. Secara diagnostik, instrumen memberi gambaran bahwa penjumlahan dengan menyimpan dan pengurangan dengan meminjam merupakan indikator yang paling banyak memerlukan tindak lanjut pada kelompok uji coba ini, sedangkan nilai tempat, penjumlahan tanpa menyimpan, pengurangan tanpa meminjam, dan soal cerita lebih banyak dikuasai. Temuan ini bersifat awal karena hanya berasal dari lima siswa.

Temuan bahwa penjumlahan dengan menyimpan dan pengurangan dengan meminjam menjadi indikator yang paling banyak memerlukan tindak lanjut sejalan dengan arah temuan Panggabean dkk. (2026) pada 28 siswa kelas II di Batangtoru. Pada penelitian tersebut, 46,4% siswa mampu menyelesaikan penjumlahan bersusun dengan menyimpan dan hanya 39,3% yang mampu menyelesaikan pengurangan bersusun dengan teknik meminjam, lebih rendah daripada penjumlahan tanpa menyimpan (71,4%) dan pengurangan sederhana (64,3%). Panggabean dkk. (2026) mengaitkan kesulitan tersebut antara lain dengan rendahnya pemahaman nilai tempat, serta melaporkan bahwa media konkret dan program remedial membantu siswa. Hal ini sejalan dengan tindak lanjut yang dirancang pada instrumen ini, yaitu remedial dengan benda atau model konkret (Tabel 3). Pada kedua penelitian, prosedur menyimpan dan meminjam antarnilai tempat sama-sama lebih sulit daripada penjumlahan dan pengurangan tanpa prosedur tersebut. Adapun Saleh dkk. (2026) menerapkan asesmen diagnostik dengan pendekatan Newman pada operasi hitung campuran di salah satu sekolah dasar (25 siswa, kelas IV pada bagian pembahasan) dan menemukan bahwa kesalahan dominan siswa terjadi pada tahap memahami soal, yang berlanjut ke tahap transformasi dan keterampilan proses. Temuan mereka berfokus pada kesulitan memahami soal dan tidak membahas prosedur menyimpan atau meminjam, sehingga perbandingan dengan penelitian ini terbatas pada kesalahan pada soal cerita. Pada penelitian ini, kesalahan pada soal cerita hanya dialami satu siswa (butir 2), dan kesalahan tersebut berkaitan dengan pemilihan operasi (mengurang, bukan menambah). Perbandingan ini perlu dibaca hati-hati karena subjek, materi, dan jumlah siswa pada ketiga penelitian berbeda, dan penelitian ini belum menggali penyebab kesalahan melalui wawancara.

Tabel 9. Perbandingan dengan Penelitian Sejenis

Penelitian	**Fokus**	**Perbedaan/Posisi Penelitian Ini**
Homjan dkk. (2022)	Tes diagnostik penjumlahan dan pengurangan dasar untuk siswa sekolah dasar	Penelitian ini menambahkan pengodean, interpretasi, dan tindak lanjut per jawaban untuk kelas II.
Aprianti & Suminar (2025)	Instrumen diagnostik *three-tier multiple choice* bilangan cacah kelas IV, fokus miskonsepsi	Penelitian ini menyasar kelas II dengan bentuk pilihan ganda dan isian singkat serta menekankan tindak lanjut bagi guru.
Mulyani dkk. (2026)	Validasi soal strategi hitung mental Fase A	Penelitian ini menilai pemahaman konsep operasi dan nilai tempat serta kaitannya dengan tindak lanjut.
Penelitian ini	**Instrumen 12 butir penjumlahan dan pengurangan kelas II, lengkap dengan pengodean, interpretasi, tindak lanjut**	**Valid (86,67%) dan menghasilkan peta pemahaman pada uji coba kelompok kecil (n = 5).**

Penelitian ini memiliki keterbatasan. Pertama, uji coba hanya melibatkan lima siswa dari satu sekolah sehingga hasilnya tidak dapat digeneralisasikan. Kedua, analisis psikometrik seperti reliabilitas, validitas butir, tingkat kesukaran, dan daya pembeda belum dilakukan, dan temuan diagnostik pada kelima siswa hanya bersifat deskriptif. Ketiga, validasi hanya melibatkan tiga validator dan merupakan validitas isi. Keempat, keterbacaan dan kepraktisan instrumen belum diuji secara terpisah, baik melalui angket siswa maupun angket guru pengguna; penilaian guru praktisi hanya diperoleh sebagai validator. Kelima, beberapa butir dan pernyataan perlu diperbaiki pada pengembangan berikutnya: butir 2 dan 5 bermuatan hampir sama (13 + 7 telur); pernyataan 13 dan 14 pada angket validasi berisi sama; dan butir 11 (10 − 5) secara prosedural

memerlukan penguraian puluhan sehingga kurang sesuai dengan indikator pengurangan tanpa meminjam, sedangkan butir 12 (15 − 5) telah sesuai dengan indikator tersebut. Keenam, efektivitas penggunaan hasil asesmen untuk memperbaiki pembelajaran belum diuji dan tahap Disseminate belum dilaksanakan.

IV. Kesimpulan

Pengembangan melalui tahap Define, Design, dan Develop menghasilkan instrumen asesmen diagnostik matematika kelas II yang terdiri atas 12 butir soal (lima pilihan ganda dan tujuh isian singkat) beserta kisi-kisi, kunci jawaban, pedoman pengodean, interpretasi, dan alternatif tindak lanjut. Instrumen dinilai sangat valid oleh tiga validator dengan rata-rata 86,67%, dengan kelayakan isi 86,67%, penyajian 85,00%, dan komponen asesmen diagnostik 88,00%. Pada uji coba kelompok kecil, instrumen menghasilkan rata-rata nilai 83,16 dengan dua siswa Paham Utuh dan tiga siswa Paham Sebagian. Berdasarkan analisis per indikator, penjumlahan dengan menyimpan dan pengurangan dengan meminjam merupakan indikator yang paling banyak belum dikuasai (masing-masing dua dari lima siswa), sehingga keduanya menjadi prioritas tindak lanjut pembelajaran pada kelompok ini. Hasil ini merupakan bukti awal kevalidan dan kelayakan, bukan bukti efektivitas pembelajaran atau keterbacaan. Penelitian berikutnya perlu melibatkan sampel lebih besar, menguji karakteristik psikometrik dan keterbacaan, memperbaiki butir yang bermuatan serupa, dan mengevaluasi pemanfaatan hasil asesmen dalam perencanaan pembelajaran.

REFERENCE

[1]. Anajjah, S. N., Iriawan, S. B., & Mufliva, R. (2022). Pengembangan bahan ajar berbasis pendekatan concrete-pictorial-abstract (CPA) dalam pembelajaran matematika kelas II sekolah dasar. ***Jurnal Pendidikan Guru Sekolah Dasar***, 7(3), 9–18. https://doi.org/10.17509/jpgsd.v7i3.57139

[2]. Andreansyah, A., Afifah, D. S. N., & Puspasari, R. (2026). Systematic literature review: Tren penelitian pengembangan asesmen diagnostik dalam pembelajaran matematika tahun 2017–2025. Pedagogy: ***Jurnal Pendidikan Matematika***, 11(2), 361–376. https://e-journal.my.id/pedagogy/article/view/8476

[3]. Aprianti, D., & Suminar, T. (2025). Pengembangan instrumen diagnostik bentuk three tier multiple choice pada mata pelajaran matematika tingkat SD. Pendas: ***Jurnal Ilmiah Pendidikan Dasar***, 10(4), 238–250. https://doi.org/10.23969/jp.v10i04.39731

[4]. Arikunto, S., & Jabar, C. S. A. (2004). Evaluasi program pendidikan: Pedoman teoretis praktis bagi mahasiswa dan praktisi pendidikan. Bumi Aksara.

[5]. Aulia, I., Fath, A. M. A., & Hutagalung, E. E. (2026). Eksplorasi pengalaman manajerial guru dalam mengelola asesmen diagnostik pada pembelajaran matematika sekolah dasar: Sebuah tinjauan literatur sistematis. JGK ***(Jurnal Guru Kita),*** 10(3), 329–344. https://doi.org/10.24114/jgk.v10.i3.73935

[6]. Brookhart, S. M. (2023). Classroom assessment essentials. ASCD.

[7]. Clements, D. H., & Sarama, J. (2020). Learning and teaching early math: The learning trajectories approach (3rd ed.). Routledge. https://doi.org/10.4324/9781003083528

[8]. Dethan, E. A., Hendrik, H. M. A., & Bodkova, R. (2026). Pengaruh media konkret terhadap kemampuan numerasi awal siswa sekolah dasar: Studi eksperimen pada pembelajaran penjumlahan dan pengurangan. Pendas: ***Jurnal Ilmiah Pendidikan Dasar***, 11(2), 223–244. https://doi.org/10.23969/jp.v11i02.54191

[9]. Fahma, M. A., & Purwaningrum, J. P. (2021). Teori Piaget dalam pembelajaran matematika. MUST: Journal of Mathematics Education, Science and Technology, 6(1), 31–42. https://doi.org/10.30651/must.v6i1.6966

[10]. Homjan, S., Sri-ngan, K., & Homjan, W. (2022). Construction of diagnostic test in mathematics on addition and subtraction basic for primary students. ***Journal of Education and Learning***, 11(3), 88–94. https://doi.org/10.5539/jel.v11n3p88

[11]. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2022). Panduan pembelajaran dan asesmen. Pusat Kurikulum dan Pembelajaran.

[12]. Mongkau, J. G., & Sumilat, J. M. (2025). Studi kasus penerapan prinsip Understanding by Design pada pembelajaran matematika kelas 5 di SD Negeri 1 Nanasi. Economics and Digital Business Review, 5(2), 520–530.

[13]. Mulyani, E., Putra, Z. H., Alim, J. A., & Afrillia, Y. M. (2026). Validation of mental calculation strategy questions for Phase A elementary school students. Indiktika: ***Jurnal Inovasi Pendidikan Matematika***, 8(2), 380–392. https://doi.org/10.31851/indiktika.v8i2.21085

[14]. Mustofa, M., & Mariati, P. (2022). Implementasi Kurikulum Merdeka Belajar di sekolah dasar: Dari teori ke praktis. Indonesia Berdaya, 4(1), 13–18.

[15]. Panggabean, D. U., Caniago, J. S., Pefriani, & Amir, A. (2026). Analisis kemampuan operasi hitung penjumlahan dan pengurangan pada siswa kelas II SD di Batangtoru: Faktor kesulitan dan upaya guru. Pendas: ***Jurnal Ilmiah Pendidikan Dasar***, 11(2), 226–238. https://doi.org/10.23969/jp.v11i02.47666

[16]. Saleh, S., Hayati, M., & Mappanyompa, M. (2026). Penerapan asesmen diagnostik dalam mengidentifikasi kesulitan belajar operasi hitung matematika siswa sekolah dasar. ***Jurnal Pendidikan Matematika dan Sains***, 14(1), 239–248. https://doi.org/10.21831/jpms.v14i1.94385

[17]. Sumarsih, I., Marliyani, T., Hadiyansah, Y., Hernawan, A. H., & Prihantini. (2022). Analisis implementasi Kurikulum Merdeka Belajar di sekolah penggerak. ***Jurnal Basicedu,*** 6(5), 8248–8258. https://doi.org/10.31004/basicedu.v6i5.3216

[18]. Syah, R. D., & Febriola, E. (2026). Asesmen diagnostik dalam pembelajaran matematika untuk mengidentifikasi kesulitan belajar siswa SD. Pendas: Jurnal Ilmiah Pendidikan Dasar, 11(1), 330–349. https://doi.org/10.23969/jp.v11i1.42519

[19]. Tarigan, A. F., Tarusu, D. T., & Ester, K. (2026). Penerapan model project based learning untuk meningkatkan hasil belajar matematika siswa kelas V SD Katolik I Woloan. ***Indo Green Journal,*** 4(2), 295–298. https://doi.org/10.31004/green.v4i2.11

[20]. Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). Instructional development for training teachers of exceptional children: A sourcebook. Center for Innovation in Teaching the Handicapped, Indiana University.

[21]. Wantiana, I., & Mellisa. (2023). Kendala guru dalam penerapan Kurikulum Merdeka. ***Jurnal Basicedu***, 7(3), 1461–1465. https://doi.org/10.31004/basicedu.v7i3.5149

[22]. Wiggins, G., & McTighe, J. (2005). Understanding by design (Expanded 2nd ed.). ASCD.