

PELATIHAN GURU MENGEMBANGKAN SOAL LITERASI MATEMATIKA BERBASIS BUDAYA PAPUA BERBANTUAN ARTIFICIAL INTELLIGENCE

Rian Efendi^{1*}, Elieser Kulimbang², Firman Rizki Alfianto³

^{1,2,3}Pendidikan Matematika, FKIP, Universitas Cenderawasih, Jayapura, Indonesia

*E-mail: rianefendi07@gmail.com

ABSTRAK

Rendahnya capaian literasi matematika siswa Indonesia, skor 366 pada PISA 2022 terhadap rerata OECD 472, menempatkan guru sebagai simpul pembenahan, terutama di wilayah yang kekayaan budayanya belum digarap menjadi konteks belajar. Pengabdian ini bertujuan menguatkan pemahaman guru mengenai literasi matematika sekaligus melatih keterampilan menyusun soal literasi bermuatan budaya Papua dengan bantuan platform Asisten Guru berbasis AI. Mitranya sepuluh guru SMP Negeri 10 Kota Jayapura, Distrik Muara Tami, kawasan yang berbatasan dengan Papua Nugini. Kegiatan berbentuk pelatihan partisipatif empat fase: persiapan, pelaksanaan inti, evaluasi ketercapaian, dan pemantauan. Pelaksanaan inti menempuh empat tahap, yaitu pemantapan konsep, penggalian konteks budaya, praktik perintah terstruktur enam unsur dengan verifikasi wajib, serta telaah silang berpasangan dan revisi kolaboratif. Data tes awal, tes akhir, rubrik telaah butir, dan lembar observasi diolah secara deskriptif dengan gain ternormalisasi. Seluruh indikator tercapai dan tiga di antaranya terlampaui. Rerata skor bergerak dari 41,00 ke 93,50 dengan gain 0,89 berkategori tinggi, seluruh peserta meningkat, dan simpangan baku menyusut 80,7 persen sehingga pemahaman lebih merata. Terhimpun 100 butir berkonteks Papua; 64 butir lolos telaah putaran pertama dan seluruhnya layak sesudah satu putaran revisi. Seluruh kebutuhan perbaikan terpusat pada komponen konteks: mesin menghasilkan konteks yang tepat secara kewilayahan namun belum akrab bagi siswa, sehingga titik kritis pekerjaan guru bergeser dari memproduksi menjadi mengurasi. Tindak lanjut yang disepakati mencakup pembentukan kelompok pengembang soal dengan jadwal telaah berkala, uji coba butir kepada siswa untuk memperoleh daya beda dan tingkat kesukaran, serta pendampingan daring dan kunjungan berkala tim pengabdian.

Kata Kunci: Asesmen Berbasis Budaya; Etnomatematika Papua; Kecerdasan Buatan; Literasi Matematika; Pelatihan Guru.

ABSTRACT

Indonesian students' low mathematical literacy, 366 in PISA 2022 against an OECD average of 472, places teachers at the centre of reform, especially where local cultural wealth remains untapped as a learning context. This community service programme aimed to strengthen teachers' understanding of mathematical literacy and train them to

construct Papuan-culture literacy items using Asisten Guru, an artificial intelligence (AI) platform. The partners were ten teachers at SMP Negeri 10 Kota Jayapura in Muara Tami, a district bordering Papua New Guinea. The participatory training comprised four phases: preparation, implementation, evaluation, and monitoring. Core implementation ran in four stages: consolidating concepts, mapping local cultural contexts, practising six-element structured prompting with mandatory verification, and paired peer review with collaborative revision. Pre-test and post-test scores, an item review rubric, and observation sheets were analysed descriptively using normalised gain. All indicators were met and three were exceeded. The mean score rose from 41.00 to 93.50 (normalised gain 0.89, high category), every participant improved, and the standard deviation narrowed by 80.7 per cent, indicating more even understanding. Participants compiled 100 Papuan-context items; 64 passed the first review round and all were deemed usable after one revision cycle. Every revision concerned the context component: the AI produced contexts regionally accurate yet unfamiliar to students, shifting the crux of teachers' work from producing to curating. The agreed follow-up comprises establishing an item development team with scheduled reviews, trialling items with students to obtain discrimination and difficulty indices, and online and on-site mentoring.

Keywords: Artificial Intelligence; Culture-Based Assessment; Mathematical Literacy; Papuan Ethnomathematics; Teacher Training.

Article History:	
Diterima	: 25-07-2026
Disetujui	: 15-08-2026
Diterbitkan Online	: 15-09-2026

PENDAHULUAN

1. Analisis Situasi

Literasi matematika menuntut seseorang membaca situasi nyata, memodelkannya secara matematis, mengerjakan model tersebut, lalu menimbang kewajaran hasilnya terhadap keadaan semula (OECD, 2023a). Kerangka (OECD, (2023a) menguraikan kemampuan tersebut ke dalam tiga proses yang berurutan, yakni merumuskan situasi secara matematis, menerapkan konsep dan prosedur, serta menafsirkan hasilnya kembali ke konteks semula. Ketiganya membedakan literasi matematika dari kecakapan berhitung biasa. Siswa dapat lancar mengoperasikan bilangan namun tersendat begitu persoalan menuntut perumusan dan penafsiran. Kemampuan ini berbeda dari kecakapan berhitung biasa. Siswa dapat lancar mengoperasikan bilangan namun bisa saja tersendat begitu persoalan menuntut pemodelan. Masalah yang dialami adalah capaian matematika siswa di Indonesia pada *Programme for International Student Assessment* (PISA) tahun 2022 sebesar 366, jauh di bawah rerata negara anggota *Organisation for Economic Co-operation and Development* (OECD), yaitu sebesar 472 (OECD, 2023b, 2023c). Selama butir yang beredar di ruang kelas dikuasai soal prosedural, ruang siswa melatih penalaran kontekstual tetap sempit dan sulit berkembang (Basriannor et al., 2023; Putra et al., 2024). Secara teoretis, kegiatan ini berpijak pada tiga bangunan gagasan yang saling menopang. Pertama, kerangka literasi matematika di atas menempatkan tuntutan proses berpikir, bukan panjang narasi maupun kehadiran latar cerita, sebagai penentu mutu sebuah butir soal. Kedua, etnomatematika menempatkan praktik budaya sebagai sumber gagasan matematis yang sah. (D'ambrosio, (1985) memandangnya sebagai cara berpikir matematis yang tumbuh di dalam

suatu kelompok budaya, sedangkan (Bishop, (1988) menegaskan bahwa aktivitas matematis melekat pada praktik dan nilai masyarakat penuturnya. Ketiga, pemanfaatan kecerdasan buatan dalam pembelajaran menuntut kecakapan guru pada dimensi etika, pedagogi, dan aplikasi (Miao & Cukurova, 2024) yang dalam kerangka *Technological Pedagogical Content Knowledge* (TPACK) berarti irisan utuh antara pengetahuan teknologis, pedagogis, dan konten (Mishra & Koehler, 2006). Ketiganya bertemu pada satu tema besar, dimana mutu asesmen kelas ditentukan oleh kemampuan guru menautkan konten matematis, proses berpikir, dan konteks yang benar-benar dijalani siswa. Pertautan terakhir inilah yang menjadi fokus kegiatan, sebab justru pada titik itu bantuan teknologi paling mudah keliru dan paling sulit menggantikan pengetahuan guru setempat.

SMP Negeri 10 Kota Jayapura berdiri di Distrik Muara Tami, kawasan yang berbatasan langsung dengan Papua Nugini. Observasi dan wawancara awal memunculkan dua keadaan. *Pertama*, literasi matematika dimaknai sebatas kecakapan menyelesaikan soal, sehingga butir yang disusun berhenti pada perhitungan langsung tanpa stimulus yang menuntut penalaran. *Kedua*, guru belum terbiasa merumuskan perintah yang tajam bagi *artificial intelligence* (AI), memeriksa ketepatan keluarannya, maupun menimbang kesesuaiannya dengan capaian pembelajaran. Keadaan kedua tersebut menyangkut integrasi teknologi, pedagogi, dan konten yang dalam kerangka *Technological Pedagogical Content Knowledge* (TPACK) dipandang sebagai satu kesatuan pengetahuan profesional (Mishra & Koehler, 2006).

Keadaan tersebut berlangsung justru ketika lingkungan sekitar sekolah menyimpan bahan konteks berlimpah, misalnya pola geometris anyaman noken, transaksi di pasar rakyat, hingga irama melaut nelayan.. Konteks semacam itu menopang numerasi siswa (Kartika et al., 2024) dan sah diperlakukan sebagai sumber belajar matematika (Dhiki & Bantas, 2021). Pemanfaatannya tetap menyimpan jebakan. D'ambrosio (1985) menempatkan etnomatematika sebagai penelusuran cara berpikir matematis yang tumbuh dalam suatu kelompok budaya, sedangkan Bishop (1988) menegaskan aktivitas matematis melekat pada praktik serta nilai masyarakatnya. Bila unsur budaya sekadar ditempelkan sebagai latar cerita, yang lahir adalah kontekstualisasi permukaan yang menambah beban membaca tanpa mempertajam penalaran (Rozi & Budiarto, 2022; Wardani et al., 2026).

Sebaliknya, *artificial intelligence* generatif mempercepat pencarian gagasan konteks, penulisan stimulus, penyusunan pertanyaan berjenjang, sampai perancangan rubrik. Pemanfaatannya disyaratkan berpusat pada manusia dan bertanggung jawab serta menuntut kecakapan guru pada sisi etika, pedagogi, dan aplikasi (Miao & Holmes, 2023; Miao & Cukurova, 2024). Di Indonesia teknologi ini terbukti memangkas waktu penyusunan asesmen, namun sebagian besar guru masih berada pada tahap awal pemanfaatan (Muhajir et al., 2026). Platform Asisten Guru keluaran Pusat Riset Pendidikan, Badan Riset dan Inovasi Nasional dipilih karena dirancang untuk konteks asesmen Indonesia.

2. Permasalahan Mitra dan Solusi yang Ditawarkan

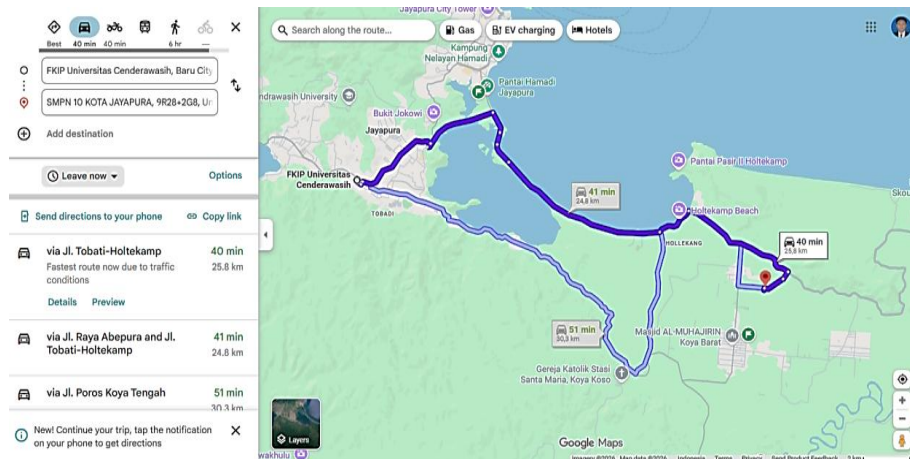
Berdasarkan analisis situasi, dirumuskan dua permasalahan prioritas mitra. Permasalahan pertama yaitu: *Pertama*, pemaknaan literasi matematika yang menyempit. Literasi dimaknai sebatas kecakapan menyelesaikan soal, sehingga butir yang disusun berhenti pada perhitungan langsung tanpa stimulus yang menuntut penalaran. *Kedua*, Pemanfaatan kecerdasan buatan yang belum terarah. Guru belum terbiasa merumuskan perintah yang tajam, memeriksa ketepatan keluaran, maupun menimbang kesesuaiannya dengan capaian pembelajaran. Berdasarkan dua permasalahan tersebut, maka

ditawarkan solusi yang dapat menyelesaikan kedua masalah tersebut. Dua solusi yang ditawarkan, yaitu: *Pertama*, pemaknaan literasi yang sempit dijawab dengan pemantapan konsep disertai latihan mengklasifikasi butir menurut proses merumuskan, menerapkan, dan menafsirkan. *Kedua*, pemanfaatan *artificial intelligence* yang belum optimal dijawab dengan pelatihan perintah terstruktur enam unsur berikut kewajiban memverifikasi setiap keluaran melalui telaah silang antarsejawat. Keduanya disatukan dalam satu alur kerja dan diikat pada konteks sekitar sekolah, berpijak pada temuan bahwa pelatihan guru yang melekat berciri fokus pada konten, pembelajaran aktif, dan kolaborasi antarsejawat (Desimone, 2009). Kegiatan bertujuan menguatkan pemahaman konsep, melatih penyusunan stimulus berbasis budaya Papua, membekali pemanfaatan platform AI secara kritis, dan menghasilkan bank soal awal. Hal ini bermanfaat bagi guru sebagai kerangka kerja asesmen, bagi sekolah sebagai perangkat yang dapat dikembangkan mandiri, dan bagi siswa sebagai butir yang dekat dengan kesehariannya.

METODE PELAKSANAAN KEGIATAN

1. Lokasi, Waktu, dan Peserta Kegiatan

Kegiatan PKM ini berlangsung di SMP Negeri 10 Kota Jayapura, Distrik Muara Tami, pada 18–21 Mei 2026. Sasaran ditetapkan purposif sebanyak sepuluh guru yang terlibat langsung dalam perencanaan pembelajaran dan penyusunan asesmen: 2 guru Matematika dan 8 guru mata pelajaran lain, dengan masa kerja 1 sampai 6 tahun (rerata 3,9 tahun; median 4,0; simpangan baku 2,23). Tim pelaksana beranggotakan dua dosen Pendidikan Matematika dan satu mahasiswa yang menangani pendampingan teknis sekaligus dokumentasi. Kegiatan memperoleh izin Kepala Sekolah dan seluruh peserta menyatakan kesediaan sesudah memperoleh penjelasan mengenai tujuan serta rencana publikasinya. Identitas peserta disamarkan dengan kode G01 hingga G10.



Gambar 1. Peta Jarak Lokasi Kampus Tim PkM menuju Lokasi PkM.

Gambar 1 menunjukkan peta jarak kampus tim PkM (FKIP Universitas Cenderawasih) menuju lokasi kegiatan PkM di SMP Negeri 10 Kota Jayapura. Jarak lokasi adalah sekitar 24,8 Km dengan waktu tempuh kurang lebih 41 menit.

2. Instrumen Kegiatan

Untuk mengukur ketercapaian program ini, ada tiga instrumen yang digunakan, yaitu: *Pertama*, **tes pemahaman konsep literasi matematika** berentang skor 0–100 yang diberikan sebelum dan sesudah kegiatan dengan naskah sama; tes mengukur penguasaan kerangka konseptual dan kecakapan memilah butir menurut ketiga proses literasi. *Kedua*, **rubrik telaah butir**

berupa daftar periksa empat kriteria, kebenaran isi menurut bidang studi penyusun serta keterpenuhan komponen konten, proses, dan konteks, yang dinilai dikotomis, dengan butir dinyatakan layak apabila keempatnya terpenuhi. Karena delapan dari sepuluh penyusun bukan guru Matematika, telaah kebenaran isi oleh sejawat diperiksa ulang dua dosen Pendidikan Matematika. *Ketiga, lembar observasi partisipasi* berikut catatan lapangan. Data kuantitatif diolah secara deskriptif memakai rerata, simpangan baku, persentase, dan indeks gain ternormalisasi menurut rumus Hake (1998), yaitu selisih skor akhir dengan skor awal dibagi selisih skor maksimum ideal dengan skor awal, dihitung per peserta lalu dirata-ratakan, dengan kategori rendah ($g < 0,30$), sedang ($0,30 \leq g < 0,70$), dan tinggi ($g \geq 0,70$). Koefisien korelasi Pearson dilaporkan sebagai pelengkap deskriptif tanpa pengujian inferensial mengingat $n = 10$, sedangkan data kualitatif direduksi, disajikan, lalu disimpulkan. Empat indikator keberhasilan ditetapkan: sekurang-kurangnya 75% peserta memperoleh skor akhir lebih tinggi; tersusun sekurang-kurangnya 10 butir dengan sekurang-kurangnya 80% memenuhi kriteria telaah; sekurang-kurangnya 80% peserta menyusun perintah enam unsur lengkap; serta sekurang-kurangnya 80% butir dinyatakan layak berikut kunci jawabannya.

3. Tahapan Kegiatan

Kegiatan dilaksanakan melalui empat fase, yakni persiapan, pelaksanaan inti, evaluasi ketercapaian, serta tindak lanjut dan pemantauan. Pendekatan yang dipakai bersifat partisipatif: peserta berkedudukan sebagai mitra kerja sejak pengenalan kebutuhan sampai penelaahan produk, bukan sebagai penerima materi. Rincian setiap fase disajikan pada Tabel 1.

Tabel 1. Fase, Kegiatan, Luaran, dan Penanggung jawab Pelaksanaan.

Fase	Waktu	Kegiatan pokok	Luaran fase	Pelaksana
I. Persiapan	Sebelum 18 Mei 2026	Observasi dan wawancara awal; perizinan kepala sekolah; penetapan sasaran secara purposif; penyusunan tes, rubrik telaah, dan lembar observasi; penyiapan lembar langkah perintah enam unsur.	Instrumen siap pakai; kesepakatan pelaksanaan	Tim pengabdi
II. Pelaksanaan inti	18–21 Mei 2026	Tahap 1 pemantapan konsep; Tahap 2 penggalan konteks budaya setempat; Tahap 3 praktik perintah enam unsur dengan verifikasi wajib; Tahap 4 telaah silang dan revisi kolaboratif. Tes akhir; rekapitulasi rubrik telaah butir; rekapitulasi lembar observasi partisipasi; pengukuran empat indikator keberhasilan.	100 butir soal; kemampuan menyusun perintah	Tim dan peserta
III. Evaluasi ketercapaian	Akhir 21 Mei 2026	Serah terima empat luaran; pembentukan kelompok pengembang soal; pendampingan daring dan kunjungan berkala; agenda uji coba butir kepada siswa.	Rekapitulasi capaian indikator	Tim pengabdi
IV. Tindak lanjut dan pemantauan	Sesudah kegiatan		Kelompok pengembang soal; agenda pendampingan	Tim dan sekolah

Kegiatan dirancang dengan pendekatan pelatihan partisipatif, yakni peserta tidak diposisikan sebagai penerima materi melainkan sebagai mitra kerja yang terlibat sejak identifikasi kebutuhan hingga penelaahan produk. Pelaksanaan berlangsung dalam empat tahap yang saling berurutan.

Tahap pertama, penguatan konsep. Peserta memperoleh pemaparan mengenai definisi literasi matematika, perbedaan karakteristik soal rutin dan soal literasi, serta pemetaan butir terhadap proses *formulate*, *employ*, dan *interpret*. Sesi ini diakhiri dengan latihan mengklasifikasi contoh soal menggunakan rubrik sederhana (Efriani & Wachyuni, 2024). Pada sesi ini kerangka proses PISA ditautkan dengan level kognitif AKM Numerasi *knowing*,

applying, dan *reasoning*, agar peserta dapat menghubungkan kerangka internasional dengan perangkat asesmen yang berlaku di sekolahnya.

Tahap kedua, penggalian konteks budaya lokal. Peserta memetakan aktivitas budaya dan keseharian masyarakat di sekitar sekolah yang memuat unsur matematis, kemudian menyusunnya menjadi stimulus soal. Pemetaan diarahkan agar keterkaitan antara praktik budaya dan konsep matematika bersifat struktural, bukan sekadar penamaan objek.

Tahap ketiga, praktik pemanfaatan AI. Peserta berlatih menyusun perintah terstruktur pada aplikasi Asisten Guru yang memuat enam unsur, yaitu konteks, kompetensi yang diukur, level kognitif, bentuk soal, kunci jawaban, dan rubrik penskoran (kerangka lengkap disajikan pada Lampiran 2). Setiap keluaran aplikasi wajib diperiksa ulang oleh peserta sebelum digunakan, dengan penekanan bahwa aplikasi berperan sebagai alat bantu dan bukan pengganti pertimbangan pedagogis guru.

Tahap keempat, telaah dan revisi kolaboratif. Butir soal yang dihasilkan ditelaah secara silang berpasangan antarpeserta menggunakan rubrik empat kriteria, diverifikasi oleh tim pengabdian, lalu direvisi dan dihimpun menjadi bank soal awal sekolah.

HASIL KEGIATAN DAN PEMBAHASAN

1. Pemantapan Pemahaman Konsep Literasi Matematika

Tes awal menghasilkan rerata 41,00 dengan skor terendah 20, tertinggi 70, dan simpangan baku 17,45. Sebaran selebar itu tidak seiring dengan masa kerja: koefisien korelasi keduanya hanya 0,09, dan tiga dari lima peserta bermasa kerja terpanjang justru berada di bawah rerata kelompok. Kerangka literasi matematika karenanya tidak terbentuk sendirinya melalui akumulasi jam mengajar, sehingga penetapan sasaran pelatihan berdasar senioritas kehilangan pijakan. Kekeliruan paling sering muncul pada soal bercerita panjang yang sejatinya hanya menuntut satu prosedur, namun dianggap soal literasi semata-mata karena bernarasi. Hal ini sejalan dengan penegasan OECD (2023a) bahwa penentu sebuah butir mengukur literasi adalah tuntutan proses berpikirnya, bukan kehadiran konteksnya.



Gambar 2. Kiri: Sesi Pemantapan Konsep Literasi Matematika Bersama Pemateri dan Guru SMP Negeri 10 Kota Jayapura; **Kanan:** Praktik Penyusunan Butir Berbantuan Platform Asisten Guru.

Gambar 2 menunjukkan bahwa sesudah kegiatan berakhir, rerata skor naik menjadi 93,50 dengan simpangan baku 3,37, selisih 52,50 poin dengan rerata pre tes sebelum kegiatan, dan indeks gain ternormalisasi 0,89 berkategori tinggi (Tabel 1). Tiga hal layak digarisbawahi. Seluruh peserta, 10 dari 10 atau 100%, memperoleh skor akhir lebih tinggi sehingga indikator pertama terlampaui 25 poin persentase. Tidak seorang pun jatuh di bawah kategori tinggi karena indeks individual terentang 0,75 sampai 1,00, pertanda keberhasilan merata dan bukan ditopang segelintir peserta. Simpangan baku pun menyusut 80,7%, pertanda kesenjangan pemahaman awal dipersempit secara nyata.

Tabel 2. Rekapitulasi Skor Tes Awal, Tes Akhir, dan Gain Ternormalisasi Peserta ($n = 10$).

Data peserta	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	Rerata	SB
Masa kerja (th)	2	6	2	6	1	2	2	6	6	6	3,9	2,23
Tes awal	60	70	30	25	30	50	30	35	60	20	41,00	17,45
Tes akhir	100	95	95	90	90	95	95	95	90	90	93,50	3,37
Gain ternormalisasi	1,00	0,83	0,93	0,87	0,86	0,90	0,93	0,92	0,75	0,88	0,89	0,07

Keterangan: SB = simpangan baku; seluruh gain berkategori tinggi (Sumber: Data primer, 2026).

Tabel 2 menunjukkan rekapitulasi skor tes awal, tes akhir, dan gain ternormalisasi peserta dengan $n = 10$). Tabel ini menunjukkan bahwa kenaikan itu tidak semata bersifat angka: pada latihan penutup tahap pertama peserta tidak lagi menakar butir dari panjang narasinya, melainkan dari tuntutan proses berpikirnya. Besarannya tetap perlu dibaca proporsional, sebab skor awal yang rendah mencerminkan belum dikenalnya kerangka literasi sebagai alat kerja, bukan lemahnya kompetensi matematis, dan materi tes berhimpitan dengan materi pelatihan. Korelasi masa kerja dengan gain bernilai $-0,52$; dengan $n = 10$ angka ini merupakan pola yang layak diuji pada skala lebih luas, bukan bukti hubungan.

2. Penggalan Konteks Budaya dan Produksi Butir Berbantuan *Artificial Intelligence*

Pemetaan konteks sebagian besar bertumpu pada kegiatan ekonomi keseharian, terutama jual beli pinang. Diskusi memuncak ketika tim bertanya tentang unsur matematis apa yang benar-benar hidup di dalam praktik itu, bukan yang sekadar ditempelkan padanya. Rancangan awal beberapa kelompok masih menjadikan objek budaya sebagai latar cerita. Misalnya penjumlahan harga yang kebetulan menyebut noken sebagai barang dagangan, lalu diarahkan pada hubungan struktural seperti pola pengulangan simpul anyaman noken yang dapat dimodelkan sebagai barisan (Wurdani & Budiarto, 2021; Sitepu & Purba, 2026).

Perbandingan di dalam sesi memperlihatkan jurang mutu antara perintah umum dan perintah terstruktur. Permintaan semacam “Buatkan soal tentang noken” menghasilkan butir dangkal yang nyaris seragam antarpeserta, sedangkan perintah enam unsur menghasilkan draf yang jauh lebih dapat dikerjakan dengan unsur budaya yang terikat struktural pada konsep yang diukur. Pada penghujung sesi seluruh peserta (100%) mampu menyusun perintah berunsur lengkap sehingga indikator ketiga terlampaui 20 poin persentase. Setiap peserta merampungkan sepuluh butir dalam satu sesi dan seorang di antaranya menghasilkan sekitar lima puluh butir, sehingga terkumpul 100 butir atau sepuluh kali lipat ambang indikator kedua. Latar belakang delapan peserta yang bukan guru Matematika justru melahirkan cakupan konteks lebih luas. Hal ini menegaskan bahwa literasi matematika bersifat lintas mata pelajaran (OECD, 2023a).

3. Telaah dan Revisi: dari Konteks Kewilayahan ke Konteks Proksimal

Telaah berlangsung silang berpasangan: setiap peserta menelaah sepuluh butir milik rekannya memakai rubrik, lalu hasilnya diverifikasi tim, termasuk pemeriksaan kebenaran matematis oleh dua dosen Pendidikan Matematika. Pada putaran pertama, 64 dari 100 butir memenuhi keempat kriteria sekaligus. Dari 36 butir sisanya, tidak satu pun bermasalah pada kebenaran isi, komponen konten, maupun komponen proses; seluruh kebutuhan perbaikan terpusat pada komponen konteks dan berasal dari delapan peserta. Sesudah satu putaran revisi kolaboratif, seluruh 100 butir dinyatakan layak digunakan (Tabel 3).

Tabel 3. Kebutuhan Perbaikan Butir Menurut Kriteria Telaah ($n = 100$ butir).

Kriteria telaah	Perlu perbaikan	Persen	Catatan
Kebenaran isi bidang studi	0	0%	Diverifikasi ulang tim pengabdi
Komponen konten	0	0%	Tanpa perbaikan
Komponen proses	0	0%	Terpenuhi, bukan kedalamannya
Komponen konteks	36	36%	Dari 8 peserta; benar kewilayahan, belum proksimal
Lolos telaah putaran I	64	64%	Tanpa perbaikan
Layak sesudah satu revisi	100	100%	Keempat kriteria terpenuhi

Sumber: Data primer kegiatan pengabdian, 2026.

Sifat perbaikannya tidak menyentuh kebenaran matematis maupun keautentikan budaya, melainkan kedekatan konteks dengan pengalaman siswa. Contoh yang berulang adalah stimulus penjual pinang yang semula ditempatkan di Pasar Youtefa, lalu dipindahkan peserta ke depan SMP Negeri 10 Kota Jayapura tanpa mengubah tuntutan prosesnya. Menjual pinang memang praktik yang hidup di Papua dan Pasar Youtefa memang lokasi nyata; persoalannya, pasar itu berada di pusat kota sementara siswa bermukim di Muara Tami yang berjarak cukup jauh.

Pola ini menyingkap keterbatasan mesin yang lebih halus daripada kekeliruan faktual. Aplikasi cenderung memilih penanda budaya yang paling menonjol dalam sumber informasi umum tentang suatu daerah, sehingga konteksnya benar secara kewilayahan namun belum tentu dekat secara keseharian. Yang dibutuhkan soal literasi justru kontekstualisasi proksimal, yakni yang berangkat dari lingkungan yang benar-benar dijalani siswa, dan selisih keduanya tak dapat dikenali tanpa pengetahuan orang yang tinggal serta mengajar di sana. Pembedaan ini merupakan sumbangan konseptual utama kegiatan dan berlaku melampaui kasus Papua. Tim juga tidak menemukan keluaran yang menampilkan budaya Papua secara stereotipis, sehingga kelemahannya bersifat kewilayahan, bukan representasional. Temuan ini menguatkan pesan bahwa *artificial intelligence* berkedudukan sebagai penunjang pertimbangan profesional guru, bukan penggantinya Miao & Cukurova (2024), sekaligus argumen Bishop (1988) bahwa dimensi budaya melekat pada praktik yang dijalani masyarakat.

Dari sini lahir pelajaran utama kegiatan. Penghalang terbesar guru menyediakan soal literasi selama ini adalah waktu penyusunan. Ketika mesin menurunkan ongkos produksi secara drastis, penghalang itu lenyap, tetapi persoalannya tidak selesai. Titik kritis bergeser dari kemampuan memproduksi butir menjadi kemampuan menyeleksi, memverifikasi, dan mengurasi. Guru yang sanggup menghasilkan lima puluh butir belum tentu sanggup menilai mana yang benar-benar mengukur literasi. Pelatihan pemanfaatan *artificial intelligence* karenanya perlu memberi porsi menelaah keluaran setidaknya setara dengan porsi menyusun perintah: menyusun perintah berada pada irisan pengetahuan teknologis, sedangkan mengurasi menuntut irisan penuh teknologi, pedagogi, dan konten (Mishra & Koehler, 2006).

4. Kendala Pelaksanaan, Partisipasi, dan Strategi Tutor Sebaya

Kendala paling nyata lahir dari berimpitnya jadwal pelatihan dengan program Makan Bergizi Gratis (MBG) di sekolah. Sejumlah peserta harus meninggalkan ruangan untuk mengatur pendistribusian makanan; empat peserta terlibat cukup intensif dan seorang lagi membantu sebentar, sehingga separuh peserta terdampak. Dampaknya paling terasa pada sesi praktik penyusunan perintah yang berurutan.

Tim tidak menghentikan kegiatan, melainkan menerapkan tutor sebaya: peserta yang lebih dahulu menguasai langkah penyusunan perintah dipasangkan dengan rekan yang tertinggal, berbekal lembar langkah satu halaman agar penjelasan tetap seragam. Pemilihan tutor bersandar pada hasil kerja sesi sebelumnya, bukan senioritas, melainkan keputusan yang dibenarkan secara empiris oleh nihilnya korelasi masa kerja dengan pemahaman awal. Hasilnya, seluruh peserta terlibat penuh pada keempat tahap dan menuntaskan setiap tugas keluarannya, tidak seorang pun jatuh di bawah kategori gain tinggi, dan separuh peserta mengambil peran sebagai tutor. Hal ini sejalan dengan tinjauan Topping (2005) bahwa manfaat pembelajaran antarsejawat dirasakan pula oleh pihak yang membantu, dengan temuan bahwa kolaborasi antarsejawat merupakan ciri pelatihan yang paling melekat (Desimone, 2009), serta dengan gagasan bahwa pengulangan mikro membantu menyusun kembali urutan langkah yang terputus (Sweller et al., 2019). Pelatihan di dalam jam sekolah karenanya sebaiknya dirancang modular dan menyiapkan skema tutor sebaya sejak awal, bukan sebagai tindakan darurat.

5. Ketercapaian Indikator, Luaran, dan Dampak Kegiatan

Seluruh indikator keberhasilan tercapai dan tiga di antaranya terlampaui (Tabel 4). Satu catatan diperlukan agar pembacaannya tepat: ambang 80% pada kriteria telaah baru terpenuhi sesudah satu putaran revisi kolaboratif, yang memang merupakan tahap keempat rancangan dan bukan tindakan di luar prosedur. Kedua angka dilaporkan terbuka karena justru 64% paling bermakna secara diagnostik: ia menakar nilai tambah tahap telaah silang. Capaian 100% pun dibaca sebagai keberhasilan proses telaah-revisi, bukan bukti mutu butir yang berdiri sendiri.

Tabel 4. Ketercapaian Indikator Keberhasilan Kegiatan.

Indikator	Ambang	Capaian	Status
Peserta dengan skor tes akhir lebih tinggi	$\geq 75\%$	100% (10/10)	Terlampaui
Jumlah butir soal tersusun	≥ 10 butir	100 butir	Terlampaui
Butir memenuhi kriteria telaah	$\geq 80\%$	64% (putaran I); 100% (sesudah revisi)	Tercapai
Peserta menyusun perintah enam unsur	$\geq 80\%$	100% (10/10)	Terlampaui
Butir dinyatakan layak sesudah telaah	$\geq 80\%$	100% (100/100)	Terlampaui

Empat luaran diserahkan kepada sekolah: bank soal awal berisi 100 butir lolos telaah lengkap dengan kunci jawaban dan rubrik penskoran, panduan penyusunan perintah enam unsur beserta lembar langkah ringkas, rubrik telaah butir empat kriteria, serta materi pelatihan digital; di luar itu tercatat satu luaran kelembagaan berupa kesepakatan membentuk kelompok kecil pengembang soal. Dampaknya terbaca pada empat tataran. Pada tataran individu, dampak terkuat bukan kenaikan skor melainkan bergesernya kriteria peserta menilai butir, dari menakar panjang narasi menjadi menakar tuntutan proses berpikir, pergeseran yang bersifat generatif karena berlaku pada butir apa pun. Pada tataran praktik kelas, 100 butir layak pakai berakar konteks Muara Tami menghapus hambatan klasik “tidak sempat menyusun soal literasi”. Pada tataran kelembagaan, kelompok pengembang soal berikut perangkat pendukungnya menjadi artefak agar praktik bertahan tanpa kehadiran tim (Desimone, 2009). Pada tataran tata kelola teknologi, peserta keluar dengan norma kerja eksplisit bahwa tidak ada keluaran yang dipakai tanpa diperiksa.

6. Keberlanjutan dan Keterbatasan

Tim membuka pendampingan lanjutan secara daring maupun kunjungan berkala dengan agenda uji coba butir pada siswa serta pengukuran retensi pemahaman. Empat keterbatasan perlu dinyatakan terbuka. Kegiatan tidak memakai kelompok pembanding dan melibatkan peserta terbatas ($n = 10$) sementara instrumen tes berhimpitan dengan materi pelatihan, sehingga besaran kenaikan tidak dapat digeneralisasi. Rubrik telaah bersifat dikotomis sehingga mendeteksi keterpenuhan komponen dan bukan kedalamannya, terlebih karena seluruh butir berbentuk pilihan ganda. Kelayakan butir ditetapkan berdasarkan pertimbangan ahli tanpa uji coba empiris. Tingkat kesepakatan antarpemelaah juga tidak dihitung secara formal. Keempatnya tidak membatalkan capaian, tetapi menetapkan batas yang tepat bagi klaimnya.

KESIMPULAN DAN SARAN

Kegiatan berhasil menguatkan pemahaman guru SMP Negeri 10 Kota Jayapura mengenai konsep literasi matematika sekaligus keterampilan mengembangkan soal literasi berbasis budaya Papua berbantuan platform Asisten Guru. Seluruh indikator tercapai dan tiga di antaranya terlampaui: rerata skor naik dari 41,00 menjadi 93,50 dengan gain ternormalisasi 0,89 berkategori tinggi, seluruh peserta meningkat, dan simpangan baku menyusut 80,7%. Kegiatan menghasilkan 100 butir soal berbasis konteks budaya Papua, 64 di antaranya lolos telaah putaran pertama dan seluruhnya layak sesudah satu putaran revisi. Kegiatan ini juga melengkapi sekolah dengan bank soal awal, panduan penyusunan perintah, dan rubrik telaah. Dua pelajaran bersifat metodologis: kemudahan produksi menggeser titik kritis pekerjaan guru dari memproduksi menjadi mengurasi, dan seluruh kebutuhan perbaikan terpusat pada kedekatan konteks dengan keseharian siswa sehingga keabsahan kontekstual tidak dapat dialihkan kepada teknologi.

Sekolah mitra disarankan segera menjalankan kelompok pengembang soal dengan jadwal telaah berkala dan mengujicobakan butir kepada siswa untuk memperoleh data daya beda serta tingkat kesukaran. Bagi penyelenggara kegiatan sejenis, pelatihan di dalam jam sekolah sebaiknya dirancang modular dengan skema tutor sebaya sejak awal; porsi waktu menelaah keluaran mesin setidaknya setara dengan porsi menyusun perintah; dan bentuk soal uraian perlu dilibatkan dengan rubrik berskala bertingkat. Bagi peneliti lanjutan, pengukuran retensi serta pelibatan kelompok pembanding akan memperkuat dasar bukti efektivitas model pelatihan ini.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada Fakultas Keguruan dan Ilmu Pendidikan (FKIP) Universitas Cenderawasih atas dukungan pendanaan melalui dana Penerimaan Negara Bukan Pajak tahun anggaran 2026, serta kepada Kepala Sekolah dan guru SMP Negeri 10 Kota Jayapura atas kesediaan menjadi mitra.

DAFTAR PUSTAKA

- Basriannor, A., Zulkarnain, I., & Hidayanto, T. (2023). Pengembangan Soal Kontekstual Berbasis Kearifan Lokal pada Pembelajaran Matematika SMA/MA. *JURMADIKA*, 3(3), 23–32. <https://doi.org/10.20527/jurmadika.v3i3.1886>
- Bishop, A. J. (1988). Mathematical Enculturation: A Cultural Perspective on Mathematics Education. In *Mathematical Enculturation* (6th ed.). Kluwer Academic Publishers. <https://doi.org/10.1007/978-94-009-2657-8>

- D'ambrosio, U. (1985). Ethnomathematics and its Place in the History and Pedagogy of Mathematics. *For the Learning of Mathematics*, 5(1), 44–48. <https://www.jstor.org/stable/40247876>
- Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>
- Dhiki, Y. Y., & Bantas, M. G. D. (2021). Eksplorasi Etnomatematika sebagai Sumber Belajar Matematika di Kabupaten Ende. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 10(4), 2698–2709. <https://doi.org/10.24127/ajpm.v10i4.4254>
- Efriani, A., & Wachyuni, D. R. (2024). Pengembangan Soal Matematika Model PISA Konteks Ramadhan dengan Konten Uncertainty and Data. *Jurnal Pendidikan Matematika*, 15(1), 60–69. <https://doi.org/https://doi.org/10.36709/jpm.v15i1.80>
- Hake, R. R. (1998). Interactive-Engagement Versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Kartika, K., Ummah BK, Muh. K., & Rudini, M. (2024). Model Discovery Learning dengan Menggunakan Media Konkret dalam Meningkatkan Numerasi Siswa pada Pembelajaran Matematika di Sekolah Dasar. *Jurnal Basicedu*, 8(4), 2908–2923. <https://doi.org/10.31004/basicedu.v8i4.8364>
- Miao, F., & Cukurova, M. (2024). AI Competency Framework for Teachers. In *AI competency framework for teachers*. UNESCO. <https://doi.org/10.54675/ZJTE2084>
- Miao, F., & Holmes, W. (2023). Guidance for Generative AI in Education and Research. In *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record: The Voice of Scholarship in Education*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Muhajir, M., Rismawati, R., Marlis, M. I., Faisal, S. F. R., Vanezal, V., Raldiastari, S., & Hakil, M. (2026). Pemanfaatan Kecerdasan Buatan (AI) dalam Asesmen Pembelajaran Mendalam di SDIT Darul Fikri Makassar. *JURNAL PETISI (Pendidikan Teknologi Informasi)*, 7(2), 302–310. <https://doi.org/10.36232/jurnalpetisi.v7i2.4417>
- OECD. (2023a). *PISA 2022 Assessment and Analytical Framework*. OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>
- OECD. (2023b). *PISA 2022 Results: Factsheets – Indonesia*. https://www.oecd.org/en/publications/pisa-results-2022-volume-iii-factsheets_041a90f1-en/indonesia_a7090b49-en.html
- OECD. (2023c). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Putra, K. D. P., Wibawa, K. A., & Noviantari, P. S. (2024). Kemampuan Literasi Matematis Siswa Dalam Menyelesaikan Soal Pisa Konten Change And Relationship. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, 4(1), TIDAK ADA NOMOR HALAMAN. <https://doi.org/10.51574/kognitif.v4i1.1211>
- Rozi, M. F., & Budiarto, M. T. (2022). Literasi Matematis Berbasis Budaya Jombangan dalam Perspektif Etnomatematika. *MATHEdunesa*, 11(1), 58–69. <https://doi.org/10.26740/mathedunesa.v11n1.p58-69>

- Sitepu, R. A. S., & Purba, C. (2026). Dari Anyaman Noken Ke Pola Bilangan: Pendekatan Etnomatematika Papua Untuk Meningkatkan Kemampuan Literasi Numerasi. *Science and Education Journal (SICEDU)*, 5(2), 1382–1390. <https://doi.org/10.31004/sicedu.v5i2.363>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, 31(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Topping, K. J. (2005). Trends in Peer Learning. *Educational Psychology*, 25(6), 631–645. <https://doi.org/10.1080/01443410500345172>
- Wardani, T. K., Rahmawati, D., & Haryadi, H. (2026). Analisis Kontekstualisasi Soal Cerita Matematika Kelas V Berbasis Literasi Numerasi Siswa Sekolah Dasar. *Walada: Journal of Primary Education*, 5(2), 1602–1611. <https://doi.org/10.61798/wjpe.v5i2.712>
- Wurdani, W. P. A., & Budiarto, M. T. (2021). Etnomatematika Usaha Kerajinan Anyaman Rotan Masyarakat Gresik dalam Perspektif Literasi Matematis. *Jurnal Pendidikan Matematika*, 12(1), 94. <https://doi.org/10.36709/jpm.v12i1.15255>