

ANALISIS KUALITATIF LITERASI ASESMEN MATEMATIKA GURU SEKOLAH DASAR MELALUI PANDITA

Mira Amelia Amri^{1*}, Yurniwati², Wily Wandari³, Nina Nurhasanah⁴, Apit Dulyapit⁵

Universitas Negeri Jakarta, Jakarta, Indonesia¹²³⁴⁵

E-mail: [^{1*\)}](mailto:miraamelia@unj.ac.id)
[^{2\)}](mailto:yurniwati@unj.ac.id)
[^{3\)}](mailto:wilywandari@unj.ac.id)
[^{4\)}](mailto:nnurhasanah@unj.ac.id)
[^{5\)}](mailto:apitdulyapit@unj.ac.id)

Abstrak

Literasi asesmen matematika diperlukan agar guru mampu menghubungkan bukti pemikiran siswa, interpretasi kesulitan, dan tindak lanjut pembelajaran. Penelitian ini bertujuan mendeskripsikan pola perubahan literasi asesmen matematika guru sekolah dasar sebelum dan setelah penggunaan PANDITA; penelitian tidak dimaksudkan untuk menguji efektivitas atau hubungan kausal PANDITA. Penelitian menggunakan pendekatan kualitatif deskriptif-komparatif dengan sumber data berupa dokumen respons pretest dan posttest 12 guru dari empat sekolah dasar di Kecamatan Sukamakmur. Sepuluh pasangan indikator dipilih berdasarkan kesamaan konstruk dan dikelompokkan ke dalam lima dimensi: diferensiasi adaptif, penggunaan data, tindak lanjut matematika, integrasi TPACK, dan keputusan berbasis bukti. Data dianalisis melalui pemetaan respons, pengodean berbasis kriteria, perbandingan dalam setiap kasus, dan analisis lintas kasus. Kebaruan penelitian terletak pada analisis berpasangan pada guru yang sama untuk membaca konfigurasi hubungan bukti-interpretasi-tindakan lintas lima dimensi, bukan mereduksinya menjadi satu skor total. Respons yang sesuai meningkat pada seluruh dimensi, terutama tindak lanjut matematika; sembilan guru menunjukkan pola penguatan, dua pola stabil dengan konfigurasi berubah, dan satu pola campuran. Temuan menunjukkan bahwa perkembangan literasi asesmen dapat berupa penguatan maupun reorganisasi antardimensi. Implikasinya, pengembangan profesional asesmen matematika perlu menyediakan scaffolding kognitif, ruang refleksi, dan latihan keputusan berbasis bukti agar guru mampu menghubungkan indikator, data, dan tindak lanjut secara konsisten.

Kata Kunci: Literasi Asesmen Matematika, Analisis Kualitatif Komparatif, Guru Sekolah Dasar, PANDITA

Abstract

Mathematics assessment literacy enables teachers to connect evidence of students' thinking, interpretations of learning difficulties, and instructional follow-up. This study aimed to describe patterns of change in elementary school teachers' mathematics assessment literacy before and after using PANDITA; it was not designed to test PANDITA's effectiveness or establish a causal relationship. A descriptive-comparative qualitative approach was employed using paired pretest and posttest response documents from 12 teachers representing four elementary schools in Sukamakmur District. Ten indicator pairs were selected on the basis of construct similarity and organized into five dimensions: adaptive differentiation, data use, mathematics instructional follow-up, TPACK integration, and evidence-based decision making. Data were analyzed through response mapping, criterion-based coding, within-case comparison, and cross-case analysis. The study's novelty lies in its paired within-case analysis of the same teachers, which examines the configuration of evidence, interpretation, and action across five dimensions rather than reducing assessment literacy to a single total score. Aligned responses increased across all dimensions, particularly mathematics instructional follow-up; nine teachers showed strengthening patterns, two stable patterns with reconfigured responses, and one a mixed pattern. These findings indicate that assessment literacy may develop through both strengthening and reorganization across dimensions. The implication is that professional learning in mathematics assessment should provide cognitive scaffolding, structured reflection, and evidence-based decision practice so teachers can connect indicators, data, and instructional follow-up consistently.

Keywords: Mathematics Assessment Literacy, Comparative Qualitative Analysis, Elementary Teachers, PANDITA

Copyright©2026 Mira Amelia Amri, Yurniwati, Wily Wandari, Nina Nurhasanah, Apit Dulyapit

Corresponding Author: Mira Amelia Amri

Email Address: miraamelia@unj.ac.id

Received: 30 Mei 2026, Accepted: 14 Juni 2026, Published: 30 Juni 2026

PENDAHULUAN

Asesmen matematika di sekolah dasar seharusnya membantu guru memahami cara siswa membangun konsep, menggunakan representasi, melakukan prosedur, dan menjelaskan alasan. Sayac dan Veldhuis (2022) menunjukkan bahwa praktik asesmen matematika guru sekolah dasar berkaitan dengan tujuan asesmen, bentuk praktik yang digunakan, dan respons terhadap keragaman siswa. Temuan tersebut menunjukkan bahwa kualitas asesmen tidak hanya ditentukan oleh instrumen, tetapi juga oleh cara guru memaknai informasi yang dihasilkan. Anantharajan (2020) menemukan bahwa perhatian guru terhadap representasi matematika siswa dapat memperluas jumlah dan kedalaman gagasan matematika yang dikenali. Dengan demikian, kemampuan literasi asesmen dalam penelitian ini dianalisis dari sejauh mana guru dapat bergerak dari jawaban benar-salah menuju pembacaan strategi, miskonsepsi, dan kebutuhan bantuan siswa.

Kemampuan menafsirkan bukti tersebut merupakan inti literasi asesmen. Pastore dan Andrade (2019) menjelaskan literasi asesmen sebagai keterpaduan pengetahuan konseptual, keterampilan praktis, dan pertimbangan sosial-emosional. Yan dan Pastore (2022) menambahkan bahwa literasi asesmen formatif menuntut kemampuan menggunakan hasil secara tepat sekaligus mempertimbangkan dampaknya terhadap peserta didik. Kerangka ini berarti bahwa guru yang mengetahui definisi remedial belum tentu memiliki literasi yang utuh apabila ia tidak dapat menjelaskan indikator yang perlu diperbaiki dan bentuk bantuan yang sesuai. Copur-Gencturk dan Rodrigues (2021) juga menunjukkan bahwa kemampuan memperhatikan dan menafsirkan pemikiran matematika merupakan komponen khusus dalam keahlian guru. Oleh karena itu, analisis penelitian ini menempatkan ketepatan respons, konsistensi alasan, dan hubungan antara diagnosis serta tindakan sebagai bukti utama kemampuan literasi asesmen matematika.

Teknologi menyediakan laporan hasil secara cepat, tetapi kecepatan belum menjamin kualitas keputusan. See et al. (2022) menegaskan bahwa teknologi asesmen bermanfaat apabila tugas, umpan balik, dan tindak lanjut dirancang secara pedagogis. Temuan tersebut mengisyaratkan bahwa data digital baru bernilai ketika guru dapat menentukan informasi matematika apa yang perlu dibaca dan bagaimana informasi itu digunakan. Schildkamp et al. (2020) juga menunjukkan bahwa asesmen formatif memerlukan pengetahuan, keterampilan interpretasi, sikap, dan dukungan konteks sekolah. Dalam penelitian ini, kemampuan menggunakan aplikasi karena itu tidak dianalisis sebagai keterampilan teknis tersendiri, tetapi

sebagai bagian dari kemampuan memilih bukti, membaca pola kesalahan, dan menetapkan tindakan pembelajaran yang dapat dipertanggungjawabkan.

PANDITA atau Panduan Digital Asesmen Adaptif Matematika Sekolah Dasar digunakan dalam penelitian ini sebagai kerangka untuk menghubungkan tujuan pembelajaran, penyusunan asesmen, pembacaan hasil, dan keputusan remedial, pengayaan, atau diferensiasi. Keterhubungan tersebut dipakai sebagai lensa analisis karena literasi asesmen matematika tidak muncul dalam satu keputusan yang terpisah, melainkan dalam rangkaian yang dimulai dari tujuan, dilanjutkan dengan pengumpulan bukti dan interpretasi, kemudian berakhir pada tindakan pembelajaran. Kerangka ini juga sejalan dengan TPACK kontekstual yang memandang teknologi, pedagogi, konten, serta kondisi organisasi sebagai satu kesatuan keputusan (Petko et al., 2025). Dengan demikian, respons guru dinilai lebih menunjukkan literasi asesmen ketika pemilihan teknologi disertai pertimbangan tentang konsep matematika yang hendak diungkap, karakteristik siswa, dan tindak lanjut yang mungkin diberikan.

Kajian terdahulu menunjukkan bahwa kemampuan asesmen matematika berkembang melalui pengalaman menafsirkan bukti dan menggunakannya untuk menyesuaikan pembelajaran. Veldhuis dan van den Heuvel-Panhuizen (2020) menemukan bahwa dukungan terhadap penggunaan asesmen kelas yang spesifik pada materi matematika membantu guru memperoleh informasi yang bernilai tentang pemahaman siswa. Lamberg et al. (2020) juga melaporkan pergeseran keputusan guru dari sekadar melihat jawaban benar-salah menuju perhatian terhadap pemikiran siswa dan adaptasi pengajaran. Kedua temuan tersebut memperjelas bahwa indikator literasi yang penting bukan hanya kemampuan memilih jawaban yang benar, tetapi kemampuan menghubungkan bukti matematika dengan perubahan pengajaran. Li et al. (2023) menambahkan bahwa pengembangan literasi asesmen membutuhkan kombinasi teori, praktik, refleksi, dan komunitas belajar. Berdasarkan kajian tersebut, analisis penelitian ini menelusuri pola alasan, ketidakkonsistenan respons, dan hubungan antarkomponen kemampuan guru.

Sebagian penelitian menggambarkan literasi asesmen melalui skor total atau frekuensi praktik, sehingga perubahan hubungan antara bukti, interpretasi, dan tindakan pada guru yang sama dapat tertutupi. Maskos et al. (2025) menekankan bahwa asesmen formatif matematika perlu dijelaskan melalui komponen yang spesifik agar proses yang kompleks tidak disederhanakan menjadi satu skor, sedangkan Copur-Gencturk dan Tolar (2022) menunjukkan bahwa pengetahuan konten, pengetahuan pedagogis, dan kemampuan memperhatikan

pemikiran matematika merupakan komponen keahlian yang berbeda. Kesenjangan penelitian bukan hanya terletak pada terbatasnya pengukuran sebelum-sesudah, tetapi pada kurangnya penjelasan tentang bagaimana komponen literasi tersebut saling terhubung atau justru terfragmentasi pada guru yang sama. Kebaruan penelitian ini terletak pada penggunaan analisis kualitatif-komparatif berpasangan untuk menelusuri konfigurasi lima dimensi literasi asesmen matematika pada tingkat kasus. Pendekatan ini memungkinkan perubahan dibedakan menjadi penguatan, pemantapan, atau reorganisasi antardimensi, sehingga dua guru dengan jumlah respons sesuai yang sama dapat memiliki pola literasi yang berbeda. PANDITA dalam penelitian ini tidak diposisikan sebagai variabel kausal, melainkan sebagai kerangka aktivitas yang menata alur tujuan pembelajaran, bukti asesmen, interpretasi, dan keputusan tindak lanjut. Dengan demikian, kontribusi penelitian diarahkan pada pemahaman baru tentang koherensi literasi asesmen, yaitu konsistensi hubungan antara apa yang diketahui guru dari data, bagaimana data dimaknai secara matematis, dan tindakan apa yang dipilih setelahnya.

Pertanyaan penelitian ini adalah: bagaimana pola perubahan literasi asesmen matematika guru sekolah dasar sebelum dan setelah penggunaan PANDITA ditinjau dari diferensiasi adaptif, penggunaan data, tindak lanjut matematika, integrasi TPACK, dan keputusan berbasis bukti? Penelitian bertujuan mengidentifikasi pola pada setiap dimensi, membandingkan konfigurasi antarpengukuran dan antarpartisipan, serta menjelaskan bagian kemampuan yang menguat, tetap, atau masih terfragmentasi. Tujuan tersebut bersifat deskriptif-interpretatif dan tidak dimaksudkan untuk menyimpulkan efektivitas kausal PANDITA.

METODE PENELITIAN

Penelitian menggunakan pendekatan kualitatif deskriptif-komparatif dengan strategi analisis dokumen berpasangan. Pendekatan ini dipilih karena penelitian tidak hanya ingin mengetahui jumlah respons yang sesuai, tetapi menelusuri perubahan konfigurasi kemampuan pada guru yang sama sebelum dan setelah penggunaan PANDITA. Data numerik berupa frekuensi respons digunakan sebagai penanda kemunculan pola, sedangkan interpretasi utama dibangun dari isi pilihan respons, konsistensi antardimensi, dan pergeseran hubungan antara bukti, interpretasi, serta tindakan. Karena tidak menggunakan kelompok pembanding dan tidak mengendalikan faktor lain yang mungkin memengaruhi respons, desain ini tidak digunakan untuk menyimpulkan efektivitas atau hubungan sebab-akibat PANDITA.

Subjek penelitian terdiri atas 12 guru dari empat sekolah dasar di Kecamatan Sukamakmur. Seluruh guru yang memiliki dokumen pretest dan posttest berpasangan digunakan sebagai kasus analisis, kemudian identitasnya disamarkan menjadi P01 sampai P12. Dokumen pretest merepresentasikan respons awal sebelum penggunaan PANDITA, sedangkan dokumen posttest merepresentasikan respons setelah rangkaian penggunaan PANDITA. Jumlah 12 kasus dipandang memadai untuk tujuan analisis ini karena penelitian tidak diarahkan pada representativitas statistik, melainkan pada pembacaan mendalam atas kasus berpasangan yang memiliki sumber data seragam. Setiap guru menyumbang 20 unit respons terpilih, yaitu 10 respons pada pretest dan 10 respons pada posttest, sehingga tersedia 240 unit respons untuk analisis lintas lima dimensi. Kecukupan kasus ditentukan oleh fokus pertanyaan yang spesifik, kelengkapan seluruh pasangan dokumen, keragaman pola yang telah mencakup penguatan, stabil, dan campuran, serta strategi analisis dalam-kasus dan lintas-kasus. Pertimbangan ini konsisten dengan kajian Hennink dan Kaiser (2022) bahwa kecukupan sampel kualitatif perlu dipertimbangkan bersama tujuan, karakteristik data, dan kedalaman analisis, bukan ditentukan oleh jumlah responden semata.

Paparan terhadap PANDITA berlangsung dalam satu rangkaian pembelajaran profesional pada hari yang sama. Pretest diberikan sebelum PANDITA diperkenalkan. Sebelum posttest, guru mengikuti pembahasan konsep asesmen adaptif, TPACK, dan instructional decision making; menganalisis kasus matematika yang mencakup pecahan serta bangun datar/keliling; mempraktikkan pemetaan indikator dan penggunaan aplikasi interaktif; memperoleh pendampingan saat menelaah pilihan tindak lanjut; kemudian melakukan refleksi sebelum posttest. Dalam rangkaian ini PANDITA digunakan dengan dukungan fasilitator, bukan sebagai penggunaan mandiri jangka panjang. Durasi rinci setiap tahap dan frekuensi penggunaan PANDITA di luar sesi tidak diperlakukan sebagai variabel penelitian, sehingga temuan hanya menggambarkan pola respons yang tampak setelah rangkaian tersebut.

Pretest dan posttest masing-masing terdiri atas 20 butir pilihan ganda yang disusun oleh tim peneliti dalam rangkaian pengembangan PANDITA. Cakupan instrumen meliputi konsep asesmen adaptif, TPACK, penggunaan teknologi asesmen, pengambilan keputusan pembelajaran, asesmen matematika sekolah dasar, fungsi PANDITA, dan implementasi. Pada tahap pengembangan, paket PANDITA beserta instrumen pendukung ditelaah oleh ahli pendidikan matematika sekolah dasar, ahli asesmen, dan praktisi/guru; masukan digunakan untuk merevisi isi, konstruk, keterpakaian, dan kejelasan instrumen pendukung. Uji coba one-

to-one digunakan untuk memeriksa kejelasan, keterpahaman, dan kepraktisan penggunaan panduan. Untuk analisis artikel ini, sepuluh pasangan indikator dipilih sebelum analisis kasus berdasarkan kesamaan fokus konstruk, bukan berdasarkan besar-kecilnya perubahan respons, dan dikelompokkan ke dalam lima dimensi: diferensiasi adaptif, penggunaan data, tindak lanjut pembelajaran matematika, integrasi TPACK, dan keputusan berbasis bukti. Karena konteks dan redaksi butir pretest-posttest tidak seluruhnya identik, penelitian tidak mengklaim kesetaraan psikometrik kedua bentuk tes; perbandingan ditafsirkan pada tingkat konstruk dan pola respons.

Indikator pada setiap dimensi ditetapkan dari definisi operasional yang digunakan dalam kajian teori dan struktur PANDITA. Diferensiasi adaptif menilai hubungan antara variasi capaian dan penyesuaian; penggunaan data menilai fungsi bukti untuk mengenali kebutuhan; tindak lanjut matematika menilai ketepatan remedial/pengayaan terhadap indikator; integrasi TPACK menilai keselarasan teknologi, pedagogi, dan konten; sedangkan keputusan berbasis bukti menilai hubungan data dengan tindakan pembelajaran. Pemetaan indikator disajikan pada Tabel 1.

Tabel 1. Pemetaan Dimensi Dan Pasangan Indikator Pretest-Posttest

Dimensi	Indikator pretest	Indikator posttest
Diferensiasi adaptif (ADP)	Tujuan dan karakteristik penyesuaian asesmen	Tindakan pada variasi penguasaan pecahan dan tindak lanjut diferensial
Penggunaan data (DAT)	Pemanfaatan hasil asesmen untuk keputusan dan perbaikan	Data kebutuhan siswa dan laporan pencapaian indikator
Tindak lanjut matematika (TL)	Respons terhadap kesulitan pecahan dan aspek asesmen matematika	Remedial indikator pecahan serta pengayaan-remedial berdasarkan capaian
Integrasi TPACK (TPK)	Konsep TPACK dan integrasi teknologi dengan strategi-materi	Kasus GeoGebra dan Quizizz dalam pembelajaran matematika
Keputusan berbasis bukti (KBB)	Instructional decision making dan penggunaan teknologi sesuai tujuan	Keputusan berdasarkan bukti dan pemilihan teknologi sesuai materi-kebutuhan

Tabel 2. Aturan Operasional Pengodean Respons

Kode	Kriteria operasional	Contoh interpretasi
SES	Respons menampilkan unsur utama dimensi dan hubungan yang diharapkan antara bukti asesmen, interpretasi, dan tindakan.	Menggunakan capaian indikator untuk menentukan remedial/pengayaan; memilih teknologi dengan mempertimbangkan tujuan, materi, dan kebutuhan siswa.
PAR	Respons mengenali prinsip yang relevan, tetapi satu atau lebih unsur penting belum terhubung atau belum spesifik.	Memilih remedial tanpa merujuk indikator yang belum tercapai; memilih teknologi yang sesuai tetapi tanpa alasan pedagogis-konten yang lengkap.
TST	Respons berlawanan dengan prinsip dimensi, tidak menunjukkan unsur kemampuan yang dicari, atau tidak memberikan jawaban.	Menyeragamkan tindak lanjut tanpa membaca hasil asesmen; memaknai tindakan formatif sebagai penilaian sumatif; tidak menjawab.

Analisis dilakukan dalam enam tahap. Pertama, respons pretest dan posttest disusun dalam matriks berpasangan untuk setiap guru. Kedua, sepuluh pasangan indikator dipetakan ke lima dimensi. Ketiga, setiap respons dikodekan secara deduktif menggunakan aturan pada Tabel 2 sebagai sesuai dengan kriteria indikator (SES), sebagian sesuai (PAR), atau tidak sesuai/tidak teramati (TST). Codebook ditetapkan sebelum perbandingan kasus agar keputusan pengodean tidak dipengaruhi oleh arah perubahan. Keempat, perubahan dalam setiap kasus dibaca dengan membandingkan konfigurasi kode pretest dan posttest, bukan sekadar menjumlahkan jawaban. Kelima, kasus dibandingkan secara lintas partisipan untuk menemukan pola penguatan, pola stabil, dan pola campuran. Keenam, interpretasi diverifikasi dengan kembali pada jawaban asli dan memeriksa kasus yang menyimpang dari kecenderungan umum. Proses reduksi, penyajian matriks, perbandingan kasus, serta verifikasi mengikuti Miles et al. (2014), sedangkan pengembangan dan peninjauan kode mengikuti Saldaña (2021).

Keabsahan interpretasi dijaga melalui audit trail yang memuat keterkaitan butir, indikator, kode, dimensi, dan perubahan setiap peserta; pemeriksaan ulang setiap kode terhadap respons asli; serta analisis kasus negatif untuk menguji apakah interpretasi umum tetap dapat menjelaskan kasus yang menyimpang. Member checking tidak diterapkan karena sumber data berupa respons pilihan terstruktur, bukan narasi pengalaman peserta yang memerlukan konfirmasi makna. Penelitian juga tidak menggunakan dua coder independen, sehingga intercoder agreement tidak dihitung. Sebagai pengganti, kredibilitas pengodean diperkuat melalui codebook yang eksplisit, aturan keputusan yang sama pada pretest dan posttest, serta

penelusuran kembali semua interpretasi ke opsi respons sumber. O'Connor dan Joffe (2020) menegaskan bahwa intercoder reliability merupakan salah satu strategi yang dapat digunakan dalam analisis kualitatif, tetapi penerapannya perlu disesuaikan dengan tujuan analisis dan transparansi kerangka pengodean. Peer debriefing formal tidak menjadi prosedur keabsahan dalam dataset ini; keterbatasan tersebut dipertimbangkan dengan membatasi klaim pada pola yang tampak pada dokumen. Frekuensi respons ditampilkan hanya untuk mendukung pembacaan kecenderungan, bukan untuk pengujian statistik.

HASIL DAN PEMBAHASAN

Gambaran Umum Perubahan Literasi Asesmen Matematika

Analisis dilakukan terhadap 12 guru pada lima dimensi literasi asesmen matematika. Setiap dimensi direpresentasikan oleh dua indikator yang mengukur dua aspek berbeda, tetapi masih berada dalam satu konstruk kemampuan yang sama. Karena setiap guru memberikan satu respons pada masing-masing indikator, satu dimensi menghasilkan 24 unit respons, yaitu 12 guru dikalikan dua indikator. Angka 24 tersebut tidak menunjukkan jumlah peserta, melainkan jumlah keseluruhan respons yang diperiksa untuk menggambarkan ketercapaian satu dimensi. Sebagai contoh, dimensi penggunaan data terdiri atas indikator kemampuan mengidentifikasi kebutuhan belajar siswa dan indikator kemampuan menentukan tindakan berdasarkan hasil asesmen. Dari 12 guru diperoleh masing-masing 12 respons untuk kedua indikator tersebut, sehingga totalnya menjadi 24 unit respons pada dimensi penggunaan data. Setiap unit respons selanjutnya dikategorikan sebagai sesuai dengan kriteria indikator, sebagian sesuai, atau tidak sesuai/tidak teramati berdasarkan ketepatan pemahaman dan tindakan yang ditunjukkan guru.

Berdasarkan unit analisis tersebut, Tabel 3 menyajikan data lengkap respons SES pada lima dimensi. Peningkatan terbesar tampak pada tindak lanjut matematika, dari 15 dari 24 respons (62,5%) pada pretest menjadi 21 dari 24 respons (87,5%) pada posttest, atau bertambah enam respons. Diferensiasi adaptif meningkat dari 17 (70,8%) menjadi 21 respons (87,5%); penggunaan data dari 17 (70,8%) menjadi 20 respons (83,3%); integrasi TPACK dari 17 (70,8%) menjadi 20 respons (83,3%); dan keputusan berbasis bukti dari 16 (66,7%) menjadi 20 respons (83,3%). Frekuensi dan persentase tersebut digunakan sebagai jejak untuk membaca arah perubahan konfigurasi respons, bukan sebagai skor statistik yang berdiri sendiri. Pastore dan Andrade (2019) serta Yan dan Pastore (2022) memandang literasi asesmen sebagai keterpaduan antara pemahaman konseptual dan penggunaan praktis bukti asesmen. Oleh karena

itu, penambahan respons SES ditafsirkan sebagai indikasi semakin kuatnya hubungan antara pemahaman asesmen, interpretasi bukti belajar, dan keputusan pembelajaran matematika.

Tabel 3. Ringkasan Lengkap Respons SES Pada Lima Dimensi

Dimensi	Unit respons per waktu	Pretest SES n (%)	Posttest SES n (%)	Perubahan
Diferensiasi adaptif	24	17 (70,8%)	21 (87,5%)	+4 respons (+16,7 poin persentase)
Penggunaan data	24	17 (70,8%)	20 (83,3%)	+3 respons (+12,5 poin persentase)
Tindak lanjut matematika	24	15 (62,5%)	21 (87,5%)	+6 respons (+25,0 poin persentase)
Integrasi TPACK	24	17 (70,8%)	20 (83,3%)	+3 respons (+12,5 poin persentase)
Keputusan berbasis bukti	24	16 (66,7%)	20 (83,3%)	+4 respons (+16,7 poin persentase)

Pada tingkat kasus, sembilan guru memperlihatkan bertambahnya respons yang sesuai dengan kriteria indikator, dua guru menunjukkan jumlah yang relatif stabil tetapi dengan perubahan pada dimensi tertentu, dan satu guru memperlihatkan pola campuran karena beberapa dimensi menguat sementara dimensi lain melemah. Variasi tersebut menunjukkan bahwa literasi asesmen matematika berkembang secara multidimensi dan tidak selalu bergerak serentak. Seorang guru dapat lebih tepat menggunakan laporan hasil, tetapi belum konsisten menentukan bentuk bantuan matematika; sebaliknya, guru yang sejak awal kuat dalam tindak lanjut dapat masih memisahkan pilihan teknologi dari pertimbangan konten. Meijer et al. (2023) menjelaskan bahwa literasi asesmen terbentuk dalam hubungan antara pengetahuan, konsepsi, dan praktik yang dipengaruhi konteks. Temuan lintas kasus dalam penelitian ini memperjelas konsekuensinya: perubahan perlu dibaca sebagai konfigurasi kemampuan pada setiap guru, bukan diasumsikan sebagai jalur perkembangan yang seragam.

Mekanisme yang mungkin menjelaskan perubahan respons

Pola perubahan dalam data dapat dipahami melalui fungsi PANDITA sebagai cognitive scaffolding atau penyangga kognitif. PANDITA memecah keputusan asesmen yang kompleks menjadi urutan yang lebih eksplisit: menetapkan tujuan dan indikator, mengumpulkan bukti, membaca hasil, menafsirkan kebutuhan, lalu memilih remedial, pengayaan, atau diferensiasi. Struktur seperti ini dapat mengurangi kecenderungan guru langsung melompat dari skor ke tindakan tanpa menjelaskan dasar diagnostiknya. Dalam konteks pendidikan matematika,

Schons et al. (2023) menunjukkan bahwa scaffolding yang terkait dengan konten dan proses diagnosis dapat membantu calon guru menggunakan pengetahuan yang relevan ketika menilai kompetensi serta miskonsepsi matematika. Temuan tersebut tidak membuktikan bahwa scaffolding PANDITA menyebabkan perubahan dalam penelitian ini, tetapi memberi penjelasan teoretis mengapa peningkatan respons terutama muncul pada dimensi yang menuntut penghubungan indikator dengan tindak lanjut.

Mekanisme kedua adalah reflective assessment atau refleksi asesmen. Sebelum posttest, guru tidak hanya menerima paparan konsep, tetapi juga menghadapi kasus matematika, mencoba aplikasi, mendiskusikan pilihan tindak lanjut, dan melakukan refleksi. Rangkaian ini memberi kesempatan untuk membandingkan asumsi awal dengan kriteria keputusan yang lebih eksplisit. Li et al. (2023) menunjukkan bahwa pengembangan literasi asesmen formatif menguat ketika pembelajaran profesional menggabungkan teori, praktik, dan refleksi, sedangkan Vattøy dan Gamlem (2025) menekankan bahwa perkembangan praktik asesmen dalam konteks digital berlangsung melalui proses reflektif dan penyesuaian terhadap pengalaman profesional. Dalam penelitian ini, refleksi tersebut tampak secara tidak langsung pada kasus-kasus yang tidak sekadar berpindah dari salah ke benar, tetapi mengalami reorganisasi antardimensi.

Mekanisme ketiga adalah evidence-based decision making. PANDITA menempatkan data asesmen sebagai penghubung antara diagnosis dan tindakan, sehingga aplikasi digital tidak berhenti pada fungsi koreksi atau pelaporan. Schildkamp et al. (2020) menunjukkan bahwa asesmen formatif membutuhkan kemampuan menginterpretasikan bukti, sedangkan O'Connor dan Park (2023) memperlihatkan bahwa pembahasan data dapat membantu guru menyesuaikan praktik pengajaran. Dalam kerangka ini, laporan indikator, respons siswa, dan informasi capaian berfungsi sebagai dasar argumentasi untuk memilih tindakan. Oleh karena itu, peningkatan pada penggunaan data dan keputusan berbasis bukti dapat ditafsirkan sebagai menguatnya orientasi instruksional terhadap data. Sekali lagi, mekanisme tersebut merupakan interpretasi yang konsisten dengan desain PANDITA dan literatur, bukan pengujian mediasi atau bukti kausal.

Diferensiasi Adaptif Bergerak dari Pengenalan Prinsip Menuju Tindakan yang Lebih Spesifik

Pada pretest, sebagian respons masih memaknai asesmen sebagai pemberian soal yang sama atau belum menghubungkan penyesuaian tingkat kesulitan dengan tindak lanjut. Pada

posttest, lebih banyak guru memilih tindakan yang membedakan tingkat kesulitan, pengayaan, dan remedial sesuai penguasaan siswa. Perubahan ini menunjukkan pergeseran dari pengenalan prinsip adaptif menuju penggunaan hasil asesmen untuk mengatur respons pembelajaran pada situasi operasi pecahan. Namun, P01 memperlihatkan pola yang belum konsisten: ia beralih dari pilihan penyeragaman menuju penyesuaian pada satu indikator, tetapi pada indikator lain masih memilih pengulangan seluruh materi tanpa merujuk pada informasi capaian. Kasus tersebut menunjukkan bahwa mengetahui istilah diferensiasi tidak otomatis menjamin kemampuan menerapkannya secara konsisten pada persoalan matematika yang berbeda.

Sayac dan Veldhuis (2022) menunjukkan bahwa praktik asesmen matematika guru sekolah dasar berkaitan dengan tujuan asesmen dan cara guru merespons keragaman siswa. Maskos et al. (2025) juga menegaskan bahwa asesmen formatif dalam pendidikan matematika memperoleh makna ketika informasi tentang kondisi awal siswa digunakan untuk menyesuaikan pembelajaran dan umpan balik. Kedua kajian tersebut membantu menjelaskan bahwa diferensiasi adaptif tidak berhenti pada pengelompokan siswa atau variasi soal. Dalam data penelitian ini, literasi yang lebih kuat tampak ketika guru menjaga kesinambungan dari identifikasi perbedaan penguasaan menuju tindakan yang berbeda dan tetap berhubungan dengan indikator matematika. Secara mekanistik, alur PANDITA berperan sebagai scaffolding karena guru diarahkan melalui urutan bukti-kebutuhan-tindakan; urutan itu menyediakan titik pemeriksaan yang dapat mencegah penyesuaian dilakukan hanya berdasarkan intuisi. Dengan demikian, ketidakkonsistenan P01 bukan sekadar kesalahan memilih opsi, tetapi menunjukkan bahwa hubungan antara bukti dan tindakan belum sepenuhnya menjadi pola berpikir yang stabil.

Penggunaan Data Semakin diarahkan pada Kebutuhan Belajar, tetapi Diagnosis Belum Sepenuhnya Terlihat

Respons pretest menunjukkan bahwa beberapa guru belum menjawab fungsi hasil asesmen atau hanya memahami data secara umum sebagai bahan perbaikan. Pada posttest, lebih banyak guru menghubungkan data dengan keputusan sesuai kebutuhan siswa dan menggunakan laporan pencapaian indikator untuk merancang pembelajaran berikutnya. P02 dan P11 menjadi contoh penguatan karena keduanya tidak memberikan respons pada salah satu indikator penggunaan data saat pretest, tetapi pada posttest memilih pemanfaatan laporan untuk menetapkan keputusan pembelajaran. Pergeseran tersebut memperlihatkan perubahan dari pandangan administratif data sebagai catatan hasil menuju pandangan instruksional data

sebagai dasar memilih tindakan. Walaupun demikian, pilihan yang tepat pada laporan indikator belum secara otomatis menunjukkan bahwa guru mampu menguraikan miskonsepsi, strategi, atau representasi matematis yang mendasari jawaban siswa.

Schildkamp et al. (2020) mengidentifikasi kemampuan menginterpretasikan bukti sebagai salah satu prasyarat utama asesmen formatif. Dalam pendidikan matematika, interpretasi tersebut bersifat spesifik terhadap konten karena guru perlu memperhatikan strategi dan representasi yang digunakan siswa, bukan hanya hasil akhirnya. Copur-Gencturk dan Rodrigues (2021) menunjukkan bahwa kemampuan memperhatikan aspek penting pemikiran matematika merupakan komponen keahlian yang dapat dibedakan dari pengetahuan umum guru, sedangkan Anantharajan (2020) memperlihatkan bahwa kedalaman perhatian terhadap representasi menentukan banyaknya gagasan matematika yang dapat dikenali. Dikaitkan dengan temuan penelitian ini, penguatan pada dimensi penggunaan data baru mencapai tahap awal diagnosis: guru semakin mampu menentukan untuk apa data digunakan, tetapi dokumen pilihan terstruktur belum menunjukkan seberapa tajam mereka dapat membaca sumber kesalahan matematis. Dari sudut refleksi asesmen, kondisi ini menunjukkan bahwa data baru menjadi bermakna ketika guru berhenti untuk menanyakan indikator apa yang belum tercapai, jenis kesalahan apa yang mungkin terjadi, dan bukti tambahan apa yang dibutuhkan. Oleh karena itu, kemampuan penggunaan data perlu dipahami sebagai jembatan antara informasi hasil dan diagnosis, bukan sebagai pengganti diagnosis itu sendiri.

Tindak Lanjut Matematika Menunjukkan Penguatan Paling Jelas

Dimensi tindak lanjut matematika mengalami peningkatan jumlah respons yang sesuai dengan kriteria indikator paling besar. Pada pretest, sejumlah guru tidak menjawab kasus pecahan atau memilih aspek umum tanpa menghubungkannya dengan tindakan. Pada posttest, sebagian besar guru memilih remedial pada indikator membandingkan pecahan yang belum tercapai dan membedakan pengayaan bagi siswa yang telah tuntas dari remedial bagi siswa yang belum tuntas. P02 dan P11 bergerak dari respons yang belum lengkap menuju keputusan yang lebih spesifik terhadap indikator. Temuan ini menunjukkan perubahan penting dari remedial sebagai pengulangan keseluruhan materi menuju remedial sebagai respons terhadap kesulitan tertentu. Secara pedagogis, perubahan tersebut mempersempit jarak antara informasi asesmen dan keputusan mengajar, karena tindakan tidak lagi diarahkan kepada seluruh kelas tanpa membedakan kebutuhan.

Veldhuis dan van den Heuvel-Panhuizen (2020) menekankan bahwa asesmen kelas matematika harus menghasilkan informasi yang dapat digunakan untuk menentukan langkah pengajaran berikutnya. Lamberg et al. (2020) menemukan bahwa penguatan penggunaan asesmen formatif ditandai oleh pergeseran perhatian guru dari jawaban benar-salah menuju pemikiran siswa dan penyesuaian pengajaran. Telaah Maskos et al. (2025) juga menunjukkan bahwa dampak asesmen formatif bergantung pada kualitas keterkaitan antara bukti, umpan balik, dan respons pembelajaran. Hasil penelitian ini sejalan dengan ketiga kajian tersebut pada tingkat pemilihan keputusan, tetapi belum sepenuhnya pada tingkat desain bantuan. Respons berupa “remedial pada indikator tersebut” telah lebih terarah daripada pengulangan umum, namun belum menjelaskan apakah bantuan akan menggunakan representasi konkret, perbandingan visual, garis bilangan, atau bentuk scaffolding lain yang sesuai dengan kesulitan pecahan. Temuan ini memperlihatkan bahwa evidence-based decision making memiliki setidaknya dua lapis: keputusan harus relevan dengan bukti, kemudian tindakan perlu dijustifikasi secara didaktis berdasarkan karakteristik konsep matematika. PANDITA tampak membantu lapis pertama; kedalaman lapis kedua belum dapat dipastikan dari respons terstruktur.

Integrasi TPACK dan Keputusan Berbasis Bukti Menguat, tetapi Tidak Selalu Stabil pada Setiap Kasus

Pada dimensi TPACK, lebih banyak guru pada posttest mengenali penggunaan GeoGebra dan Quizizz sebagai integrasi teknologi, pedagogi, dan konten. P02 dan P11 yang sebelumnya tidak memberikan respons pada definisi atau penerapan TPACK kemudian memilih integrasi yang sesuai. Pada dimensi keputusan berbasis bukti, mayoritas guru juga beralih menuju keputusan yang didasarkan pada hasil asesmen dan pemilihan teknologi berdasarkan tujuan, karakteristik materi, serta kebutuhan siswa. Meskipun demikian, P10 memilih “TK saja” pada kasus GeoGebra setelah sebelumnya mengenali TPACK sebagai integrasi, sehingga perubahan pada guru ini bersifat campuran. Pola tersebut menunjukkan bahwa pengetahuan deklaratif mengenai TPACK dapat muncul lebih cepat daripada kemampuan menerapkan hubungan teknologi-pedagogi-konten secara konsisten pada kasus matematika tertentu.

Petko et al. (2025) menempatkan TPACK sebagai pengetahuan yang beroperasi dalam konteks keputusan kelas, bukan sekadar penjumlahan tiga jenis pengetahuan. Kajian Li et al. (2025) pada pendidikan matematika sekolah dasar juga menunjukkan bahwa integrasi teknologi perlu tetap terikat pada tujuan matematis, praktik pedagogis, dan kondisi pembelajaran.

Sementara itu, See et al. (2022) mengingatkan bahwa teknologi asesmen tidak otomatis meningkatkan kualitas pembelajaran apabila tugas dan tindak lanjutnya tidak dirancang secara pedagogis. Berdasarkan ketiga rujukan tersebut, penguatan TPACK dalam data ini ditafsirkan bukan dari semakin seringnya aplikasi dipilih, melainkan dari semakin lengkapnya hubungan antara fitur, bukti matematika yang hendak diperoleh, dan tindakan yang mengikuti. PANDITA menyediakan prompt keputusan yang mengarahkan guru untuk memulai dari tujuan dan indikator sebelum memilih aplikasi; prompt tersebut dapat berfungsi sebagai scaffolding yang menahan kecenderungan memilih teknologi karena popularitas atau kemudahan semata. Ketika guru tetap memilih teknologi tanpa mempertimbangkan jenis representasi, tujuan diagnostik, atau kebutuhan siswa, respons tersebut menunjukkan bahwa scaffolding belum terinternalisasi sebagai penalaran TPACK yang stabil.

Perubahan Lintas Kasus Menunjukkan Tiga Pola Perkembangan

Analisis dalam setiap kasus menghasilkan tiga pola. Pertama, pola penguatan ditemukan pada sembilan guru, ditandai oleh bertambahnya keterhubungan antardimensi pada posttest. Kedua, pola stabil ditemukan pada dua guru: jumlah respons yang sesuai relatif tetap, tetapi terjadi perpindahan kekuatan dan kelemahan antarindikator. Ketiga, pola campuran ditemukan pada satu guru yang menunjukkan penguatan penggunaan data, tetapi melemah pada diferensiasi atau integrasi TPACK. Kategori ini bersifat deskriptif dan tidak digunakan untuk memberi peringkat kompetensi. Gotch dan McLean (2019) menunjukkan bahwa hasil pengembangan literasi asesmen dapat berbeda antarpendidik karena pengetahuan dan efikasi diri terhadap tugas asesmen tidak selalu berkembang dengan tingkat yang sama. Temuan penelitian ini memberikan gambaran yang lebih spesifik pada konteks matematika: perbedaan tersebut tampak sebagai jalur perubahan antardimensi yang berlainan, sehingga pola stabil tidak selalu berarti tidak ada perkembangan dan pola penguatan tidak selalu berarti semua komponen telah konsisten.

Kasus P02 memperlihatkan penguatan paling luas, dari empat respons yang sesuai pada pretest menjadi sembilan pada posttest; perubahan terutama tampak pada penggunaan data, tindak lanjut matematika, TPACK, dan keputusan berbasis bukti. P05 telah menunjukkan pola kuat sejak pretest dan menjadi semakin konsisten pada posttest, sehingga perubahan lebih tepat disebut pemantapan daripada lonjakan kemampuan. P01 mempertahankan jumlah respons yang sesuai tetap sama, tetapi konfigurasi kemampuannya bergeser: penggunaan data menguat, sedangkan tindak lanjut matematika menjadi kurang lengkap. P10 menunjukkan pola campuran

karena pemahaman data menguat, tetapi respons diferensiasi dan TPACK tidak selalu konsisten. Meijer et al. (2023) menempatkan konsepsi literasi asesmen sebagai sesuatu yang tertanam dalam praktik dan konteks profesional. Perspektif tersebut membantu menjelaskan mengapa perubahan tidak selalu terlihat sebagai kenaikan total; pada beberapa kasus, perubahan justru muncul sebagai reorganisasi cara guru menghubungkan bukti, tujuan, dan tindakan. Analisis berbasis kasus karena itu memberikan informasi yang akan hilang apabila seluruh respons diringkaskan menjadi satu nilai rata-rata.

Tabel 4. Perbandingan Kasus Representatif pada Pretest dan Posttest

Kode	Perbandingan respons	Makna perubahan
P02	4 respons sesuai pada pretest menjadi 9 pada posttest; menguat pada data, tindak lanjut, TPACK, dan bukti	Pola penguatan luas
P05	9 respons sesuai menjadi 10; konsep yang telah kuat menjadi lebih konsisten pada kasus aplikatif	Pola stabil-kuat dan pementapan
P01	Jumlah tetap 7, tetapi data menguat sementara tindak lanjut matematika menjadi tidak lengkap	Pola stabil dengan konfigurasi berubah
P10	8 respons sesuai menjadi 7; data menguat, tetapi diferensiasi dan TPACK tidak konsisten	Pola campuran/terfragmentasi

Secara keseluruhan, perubahan literasi asesmen matematika tampak pada semakin kuatnya hubungan antara bukti, interpretasi, dan tindakan. Yan dan Pastore (2022) memandang literasi asesmen formatif sebagai konstruk multidimensi, sedangkan Copur-Gencturk dan Tolar (2022) menunjukkan bahwa pengetahuan konten, pengetahuan pedagogis, dan kemampuan memperhatikan pemikiran matematika merupakan komponen keahlian yang berbeda. Data penelitian ini mendukung kedua pandangan tersebut sekaligus memperlihatkan hubungan di antaranya: keputusan remedial menjadi lebih tepat ketika guru menggabungkan pemahaman indikator, pembacaan capaian, dan pilihan tindakan; sebaliknya, keputusan tetap terfragmentasi ketika salah satu komponen tidak terhubung. Kontribusi penelitian ini terletak pada penjelasan bahwa perubahan literasi tidak selalu linier dan tidak cukup direpresentasikan oleh total jawaban. Penguatan yang substantif terlihat ketika respons guru bergerak dari orientasi hasil akhir menuju orientasi pada kebutuhan belajar, tetapi kedalaman literasi baru dapat dinilai sepenuhnya apabila alasan matematis dan bentuk bantuan yang dipilih juga tersedia.

Pemahaman Baru dan Implikasi Penelitian

Pemahaman baru yang dihasilkan penelitian ini adalah bahwa literasi asesmen matematika lebih produktif dibaca sebagai koherensi antarkomponen daripada sebagai jumlah pengetahuan yang dimiliki guru. Koherensi tersebut terlihat ketika bukti belajar mengaktifkan interpretasi matematis yang relevan dan interpretasi itu mengarahkan tindakan yang sesuai. Analisis berpasangan memperlihatkan setidaknya tiga bentuk perkembangan: ekspansi, ketika lebih banyak dimensi menjadi sesuai; pemantapan, ketika konfigurasi yang sejak awal kuat menjadi lebih konsisten; dan reorganisasi, ketika jumlah respons sesuai relatif tetap tetapi lokasi kekuatan dan kelemahan berpindah. Pola P01 dan P10 menunjukkan bahwa skor atau jumlah respons yang sama dapat menyembunyikan perubahan substantif dalam cara guru menghubungkan data, konten, dan keputusan. Dengan demikian, studi ini memperluas pandangan multidimensi tentang literasi asesmen (Yan & Pastore, 2022; Meijer et al., 2023) dengan menekankan koherensi dan konfigurasi lintas dimensi sebagai unit interpretasi.

Implikasi teoretisnya, literasi asesmen matematika tidak cukup didefinisikan sebagai penguasaan konsep, kemampuan membaca data, atau keterampilan teknologi secara terpisah; kualitas literasi terletak pada keterhubungan ketiganya dalam konteks keputusan matematika. Implikasi praktisnya, pengembangan profesional perlu dirancang sebagai siklus yang menggabungkan cognitive scaffolding, reflective assessment, dan evidence-based decision making. Guru perlu diberi kasus yang menuntut pembacaan bukti, prompt untuk menjelaskan alasan, kesempatan membandingkan keputusan, serta umpan balik yang mengembalikan keputusan pada indikator dan karakteristik konsep matematika. PANDITA dapat digunakan sebagai kerangka untuk menyusun siklus tersebut, tetapi aplikasi digital tetap diposisikan sebagai sarana memperoleh dan menampilkan bukti, bukan sebagai indikator literasi itu sendiri.

Penelitian ini memiliki keterbatasan karena data berasal dari respons pilihan terstruktur dan pasangan pretest-posttest tidak menggunakan redaksi yang identik. Oleh sebab itu, perbandingan dilakukan pada tingkat kesamaan konstruk dan tidak dimaksudkan sebagai pembuktian kesetaraan butir maupun efektivitas kausal PANDITA. Pilihan respons dapat menunjukkan arah dan konsistensi kemampuan, tetapi belum sepenuhnya mengungkap alasan matematis guru, cara mereka mendiagnosis pekerjaan siswa, atau bagaimana keputusan diterapkan di kelas. Selain itu, analisis tidak menggunakan dua coder independen dan tidak menerapkan member checking karena karakter data bersifat terstruktur. Temuan perlu dibaca sebagai perubahan pola yang tampak pada dokumen. Penelitian berikutnya dapat menggunakan

respons terbuka, wawancara berbasis kasus, telaah produk asesmen, observasi kelas, refleksi tertulis, dan analisis hasil kerja siswa untuk menguji kedalaman hubungan bukti-interpretasi-tindakan pada situasi autentik.

KESIMPULAN DAN SARAN

Kesimpulan

Analisis kualitatif-komparatif menunjukkan bahwa pola literasi asesmen matematika 12 guru menjadi lebih koheren setelah rangkaian penggunaan PANDITA, terutama pada kemampuan mengubah hasil asesmen menjadi remedial dan pengayaan yang terkait dengan indikator serta kebutuhan siswa. Kontribusi utama penelitian terletak pada temuan bahwa perkembangan literasi tidak selalu berupa kenaikan yang linier: sebagian guru menunjukkan penguatan, sebagian mengalami pemantapan, dan sebagian memperlihatkan reorganisasi atau fragmentasi antardimensi. Hal ini menegaskan pentingnya membaca literasi asesmen melalui hubungan bukti, interpretasi matematis, dan tindakan, bukan melalui skor total saja. Secara praktis, pengembangan profesional perlu menggabungkan scaffolding kognitif, refleksi asesmen, dan latihan keputusan berbasis bukti. Karena desain penelitian bersifat deskriptif-komparatif tanpa kelompok pembandingan, perubahan yang ditemukan tidak dapat ditafsirkan sebagai bukti efektivitas atau pengaruh kausal PANDITA.

Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan, pengembangan PANDITA disarankan untuk memperkuat kemampuan guru dalam menghubungkan bukti asesmen, interpretasi matematis, dan pengambilan keputusan pembelajaran, khususnya dalam merancang remedial dan pengayaan. Pengembangan profesional juga perlu memadukan scaffolding kognitif, refleksi asesmen, dan latihan pengambilan keputusan berbasis bukti. Bagi peneliti selanjutnya, disarankan menggunakan kelompok pembandingan, sampel yang lebih beragam, dan waktu pendampingan yang lebih panjang untuk memperoleh bukti yang lebih kuat mengenai kontribusi PANDITA terhadap literasi asesmen matematika guru.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Negeri Jakarta yang telah memberikan dukungan pendanaan terhadap penelitian ini berdasarkan Kontrak Penelitian Nomor 73/PK/LPPM/III/2026.

DAFTAR PUSTAKA

- Anantharajan, M. (2020). Teacher noticing of mathematical thinking in young children's representations of counting. *Journal for Research in Mathematics Education*, 51(3), 268–300. <https://doi.org/10.5951/jresmetheduc-2019-0068>
- Copur-Gencturk, Y., & Rodrigues, J. (2021). Content-specific noticing: A large-scale survey of mathematics teachers' noticing. *Teaching and Teacher Education*, 101, 103320. <https://doi.org/10.1016/j.tate.2021.103320>
- Copur-Gencturk, Y., & Tolar, T. (2022). Mathematics teaching expertise: A study of the dimensionality of content knowledge, pedagogical content knowledge, and content-specific noticing skills. *Teaching and Teacher Education*, 114, 103696. <https://doi.org/10.1016/j.tate.2022.103696>
- Gotch, C. M., & McLean, C. (2019). Teacher outcomes from a statewide initiative to build assessment literacy. *Studies in Educational Evaluation*, 62, 30–36. <https://doi.org/10.1016/j.stueduc.2019.04.003>
- Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>
- Lamberg, T., Gillette-Koyen, L., & Moss, D. (2020). Supporting teachers to use formative assessment for adaptive decision making. *Mathematics Teacher Educator*, 8(2), 37–58. <https://doi.org/10.5951/mte-2019-0005>
- Li, M., Vale, C., Tan, H., & Blannin, J. (2025). A systematic review of TPACK research in primary mathematics education. *Mathematics Education Research Journal*, 37, 281–311. <https://doi.org/10.1007/s13394-024-00491-3>
- Li, Z., Yan, Z., Chan, K. K. Y., Zhan, Y., & Guo, W. Y. (2023). The role of a professional development program in improving primary teachers' formative assessment literacy. *Teacher Development*, 27(4), 447–467. <https://doi.org/10.1080/13664530.2023.2223595>
- Maskos, K., Schulz, A., Oeksuez, S. S., & Rakoczy, K. (2025). Formative assessment in mathematics education: A systematic review. *ZDM—Mathematics Education*, 57, 679–693. <https://doi.org/10.1007/s11858-025-01696-x>
- Meijer, K., Baartman, L., Vermeulen, M., & de Bruijn, E. (2023). Teachers' conceptions of assessment literacy. *Teachers and Teaching*, 29(7–8), 695–709. <https://doi.org/10.1080/13540602.2023.2190091>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

-
- O'Connor, B., & Park, M. (2023). Exploring the influence of collaborative data-based decision making among teachers in professional learning communities on teaching practice. *Disciplinary and Interdisciplinary Science Education Research*, 5, 17. <https://doi.org/10.1186/s43031-023-00086-1>
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1–13. <https://doi.org/10.1177/1609406919899220>
- Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 128–138. <https://doi.org/10.1016/j.tate.2019.05.003>
- Petko, D., Mishra, P., & Koehler, M. J. (2025). TPACK in context: An updated model. *Computers and Education Open*, 8, 100244. <https://doi.org/10.1016/j.caeo.2025.100244>
- Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE Publications.
- Sayac, N., & Veldhuis, M. (2022). Mathematics assessment practices of primary school teachers in France. *International Journal of Science and Mathematics Education*, 20, 1595–1610. <https://doi.org/10.1007/s10763-021-10229-3>
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103, 101602. <https://doi.org/10.1016/j.ijer.2020.101602>
- Schons, C., Obersteiner, A., Reinhold, F., Fischer, F., & Reiss, K. (2023). Developing a simulation to foster prospective mathematics teachers' diagnostic competencies: The effects of scaffolding. *Journal für Mathematik-Didaktik*, 44, 59–82. <https://doi.org/10.1007/s13138-022-00210-0>
- See, B. H., Gorard, S., Lu, B., Dong, L., & Siddiqui, N. (2022). Is technology always helpful? A critical review of the impact on learning outcomes of education technology in supporting formative assessment in schools. *Research Papers in Education*, 37(6), 1064–1096. <https://doi.org/10.1080/02671522.2021.1907778>
- Vattøy, K.-D., & Gamlem, S. M. (2025). Navigating formative assessment as professional development in digital contexts: Insights from teachers' experiences. *Teacher Development*, 29(2), 343–360. <https://doi.org/10.1080/13664530.2024.2382956>
- Veldhuis, M., & van den Heuvel-Panhuizen, M. (2020). Supporting primary school teachers' classroom assessment in mathematics education: Effects on student achievement. *Mathematics Education Research Journal*, 32, 449–471. <https://doi.org/10.1007/s13394-019-00270-5>
- Yan, Z., & Pastore, S. (2022). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*, 74, 101183. <https://doi.org/10.1016/j.stueduc.2022.101183>