

## **Pengembangan Instrumen Tes untuk Mengukur Kemampuan Berpikir komputasional Siswa dalam Pembelajaran Matematika: Sebuah Tinjauan Sistematis**

**Meliana Sukma Desti<sup>1)</sup>\*, Putri Nur Malasari<sup>2)</sup>**

<sup>1,2</sup> Universitas Islam Negeri Sunan Kudus – Jl. Conge, Ngembalrejo, Bae, Kudus, Jawa Tengah, 50111, Indonesia

\*Penulis Korespondensi : email: [melianasukma704@gmail.com](mailto:melianasukma704@gmail.com)

*Diterima: 28 Desember 2025, Direvisi: 9 April 2026, Disetujui: 6 Mei 2026.*

### **Abstract**

*The integration of computational thinking (CT) into mathematics is highly strategic due to the cognitive overlap between the two fields. However, the validity and reliability of its measurement instruments remain major challenges. This study aims to map the development of CT test instruments in mathematics through a Systematic Literature Review (SLR) of 8 empirical articles (2021-2025). The novelty of this research lies in the critical analysis of psychometric quality and the characteristics of domain-specific mathematics instruments. The results indicate a dominance of development at the elementary school level (50%). Critical findings reveal significant variations in reliability, where low coefficients (0.27 and 0.33) indicate the limitations of classical statistics in measuring multidimensional CT. The conclusion emphasizes the urgency of shifting to modern psychometric analysis (Rasch/IRT) and expanding instrument development to higher education levels to ensure the continuity of student cognitive measurement from basic to advanced levels.*

**Keywords:** *Computational Thinking, Test Instrument, Mathematics Learning, Systematic Literature Review, Assessment.*

### **Abstrak**

*Integrasi berpikir komputasional (CT) dalam matematika sangat strategis karena adanya tumpang tindih kognitif antara dua bidang. Namun, validitas dan reliabilitas instrumen pengukuran masih menjadi tantangan utama. Penelitian ini bertujuan memetakan pengembangan instrumen tes CT dalam matematika melalui metode Systematic Literature Review (SLR) terhadap 8 artikel empiris (2021-2025). Kebaruan penelitian ini terletak pada analisis kritis kualitas psikometrik dan karakteristik instrumen spesifik domain matematika. Hasil menunjukkan dominasi pengembangan pada jenjang SD (50%). Temuan kritis mengungkap variasi reliabilitas signifikan, di mana koefisien rendah (0,27 dan 0,33) mengindikasikan keterbatasan statistik klasik dalam mengukur CT yang multidimensi. Simpulan menekankan urgensi pergeseran ke analisis psikometrik modern (Rasch/IRT) serta perluasan pengembangan instrumen pada jenjang pendidikan tinggi guna menjamin kesinambungan pengukuran kognitif siswa dari level dasar hingga tinggi.*

**Kata Kunci:** *Berpikir komputasional, Instrumen Tes, Pembelajaran Matematika, Systematic Literature Review, Asesmen.*

## **1. PENDAHULUAN**

Di era digital dan revolusi industri 4.0 saat ini, dunia kerja menuntut individu untuk memiliki keterampilan yang adaptif dan interdisipliner [1]. Salah satu keterampilan kunci abad ke-21 yang mendapat perhatian luas adalah berpikir komputasional atau pemikiran komputasional. Menurut Alifah dan Widodo, berpikir komputasional tidak sekadar didefinisikan sebagai kemampuan teknis pemrograman, melainkan sebagai proses pemecahan masalah yang

melibatkan perumusan masalah dan penyusunan solusi sedemikian rupa sehingga dapat dieksekusi oleh agen pemroses informasi, baik manusia maupun komputer [2].

Posisinya yang fundamental dalam pendidikan kontemporer ditegaskan oleh Wing yang mendefinisikannya sebagai keterampilan berpikir dasar yang relevan bagi semua orang, bukan hanya bagi ilmuwan komputer [3]. Namun penelitian ini secara khusus menetapkan kerangka kerja Selby & Woollard sebagai definisi operasional utama. Hal ini dikarenakan kerangka tersebut menyediakan struktur yang jelas dalam membedah pilar-pilar berpikir komputasional, yaitu dekomposisi, pengenalan pola, abstraksi, rancangan algoritma, dan evaluasi, yang sangat relevan dengan penyelesaian masalah matematis [4].

Integrasi *berpikir komputasional* dalam pembelajaran matematika menjadi sangat strategis karena adanya irisan kognitif yang kuat antara kedua bidang tersebut. Irisan kognitif tersebut seperti kemampuan abstraksi, inferensi, dan pemodelan, yang ada pada berpikir komputasional dapat meningkatkan pemahaman konsep matematika siswa [5]. Hubungan ini terlihat ketika siswa menggunakan pilar abstraksi untuk memodelkan variabel aljabar, atau menerapkan desain algoritma dalam menyusun langkah-langkah pembuktian geometri yang sistematis [6], [7]. Matematika dan berpikir komputasional juga berbagi elemen fundamental seperti pemecahan masalah, pemikiran logis, abstraksi, dan pengenalan pola. Tren penelitian mengindikasikan bahwa topik integrasi ini telah berkembang pesat, terutama sejak tahun 2013, yang menunjukkan meningkatnya kesadaran akademik akan pentingnya sinergi antara matematika dan berpikir komputasional dalam kurikulum pendidikan [8].

Meskipun integrasi berpikir komputasional dalam kurikulum telah diupayakan melalui berbagai inovasi pedagogis seperti *unplugged coding* yang efektif pada siswa sekolah dasar dengan keterbatasan infrastruktur maupun *game-based learning* yang meningkatkan motivasi siswa vokasi, tantangan signifikan masih ditemukan dalam aspek penilaian [1], [9]. Evaluasi terhadap keterampilan berpikir komputasional merupakan proses yang kompleks karena metode penilaian tradisional sering kali gagal menangkap kedalaman dan keluasan keterampilan tersebut secara utuh [10]. Diperlukan alat analitik dan metode yang tidak terlalu umum untuk menilai kompetensi baru ini secara akurat.

Belum ada standar yang diterima secara luas untuk mengevaluasi berpikir komputasional di tingkat pendidikan dasar dan menengah, serta banyak kerangka kerja evaluasi yang ada masih belum terstandarisasi dengan baik hingga saat ini [11]. Ketiadaan instrumen penilaian yang akurat dan valid menjadi hambatan bagi pendidik untuk mendiagnosis kemampuan siswa serta mengevaluasi efektivitas strategi pembelajaran yang diterapkan [10]. Selain itu, variasi

instrumen pengukuran berpikir komputasional yang ada saat ini seringkali membatasi konsistensi hasil penelitian. Tanpa alat ukur yang tepat, sulit untuk memastikan apakah intervensi pedagogis yang dilakukan benar-benar efektif dalam meningkatkan kemampuan berpikir komputasional siswa dalam konteks matematika.

Meskipun beberapa kajian bibliometrik telah memetakan tren publikasi berpikir komputasional (CT) secara umum [8], namun tinjauan sistematis yang secara spesifik membedah karakteristik teknis, validitas, dan efektivitas instrumen tes dalam konteks matematika masih sangat terbatas. Kebaruan penelitian ini terletak pada fokus analisis yang mendalam terhadap karakteristik teknis instrumen tes, cakupan materi matematika yang diintegrasikan, serta efektivitas instrumen dalam mengukur pilar kognitif CT pada berbagai jenjang pendidikan. Hal ini penting untuk mengisi kesenjangan literatur mengenai parameter instrumen yang ideal untuk mengukur CT dalam domain matematika. Berdasarkan penjelasan yang dipaparkan di atas, penelitian ini bertujuan untuk mengkaji secara komprehensif pengembangan instrumen tes berpikir komputasional dalam pembelajaran matematika, dengan fokus pada karakteristik instrumen, jenjang pendidikan, dan topik matematika yang menjadi konteks pengembangan.

## 2. METODE PENELITIAN

Studi ini didesain dengan menerapkan metode *Systematic Literature Review* (SLR) yang memungkinkan peneliti untuk melakukan identifikasi, tinjauan, evaluasi, serta interpretasi secara menyeluruh terhadap seluruh literatur yang meneliti tentang pengembangan instrumen tes berpikir komputasional dalam pembelajaran matematika. Metode SLR dipilih karena memungkinkan peneliti untuk meninjau literatur yang ada secara objektif, transparan, dan dapat direplikasi, sehingga menghasilkan temuan yang komprehensif dan terpercaya mengenai topik yang diteliti. Untuk memandu proses *review* sistematis ini, dirumuskan pertanyaan penelitian sebagai berikut:

1. Bagaimana karakteristik dan jenis instrumen tes berpikir komputasional yang dikembangkan untuk konteks pembelajaran matematika?
2. Pada level pendidikan mana saja instrumen tes berpikir komputasional dikembangkan?
3. Topik matematika apa saja yang dikembangkan dalam tes berpikir komputasional?

Objektivitas dan relevansi artikel dalam tinjauan sistematis ini dijamin melalui penetapan protokol seleksi yang ketat. Kriteria inklusi dan eksklusi berfungsi sebagai filter utama dalam menyaring hasil pencarian awal hingga diperoleh artikel final. Rincian kriteria tersebut disajikan dalam Tabel 1 berikut.

**Tabel 1.** Kriteria Inklusi dan Eksklusi

<b>Kriteria Inklusi</b>	<b>Kriteria Eksklusi</b>
Fokus konten: Artikel tentang pengembangan instrumen tes berpikir komputasional	Artikel tidak tersedia <i>full text</i> atau hanya berupa abstrak.
Konteks: Dalam pembelajaran matematika	Artikel duplikat
Jenis Penelitian: Empiris dengan pengembangan instrumen	Periode Waktu: di luar tahun publikasi 2021-2025
Validasi: Melaporkan validitas dan reliabilitas instrumen	Ditulis dalam bahasa selain Inggris atau Indonesia

Protokol yang penulis gunakan di tahap seleksi artikel adalah Protokol PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyzes*) dari Liberati dkk dengan terdapat empat tahapan yaitu *identification*, *screening*, kesesuaian, dan final. Empat tahapan tersebut dijelaskan sebagai berikut [12]:

a. Tahap 1 (*Identification*)

Tahap *identification* diawali dengan pencarian literatur secara sistematis di database Google Scholar dan Eric. Pencarian menggunakan kombinasi string pencarian sebagai berikut:

- 1) "Computational thinking" AND "mathematics" AND ("assessment" OR "test") AND "development"
- 2) "Berpikir komputasional" AND "matematika" AND ("instrumen" OR "tes" OR "asesmen") AND "pengembangan"
- 3) "Berpikir komputasional" AND "matematika" AND "instrumen penilaian"

Pencarian yang dilakukan tanpa batasan bahasa ini berhasil mengidentifikasi 890 artikel potensial yang dipublikasikan dalam periode 2021 hingga 2025 menggunakan aplikasi Publish or Perish. Sebanyak 3 artikel dieksklusi karena artikel duplikasi, sisanya sebanyak 887 artikel kemudian dikumpulkan untuk proses seleksi lebih lanjut guna memastikan kelayakan dan relevansinya dengan fokus penelitian tinjauan sistematis ini.

b. Tahap 2 (*Screening*)

Pada tahap *screening*, dilakukan penyaringan terhadap judul dan abstrak 887 artikel. Sebanyak 880 artikel dikeluarkan karena tidak berfokus pada pengembangan instrumen pengukuran CT dalam konteks matematika, melainkan membahas berpikir komputasional dalam konteks robotika atau pemrograman umum, serta artikel yang hanya fokus pada efektivitas model pembelajaran tanpa menyertakan pengembangan instrumen penilaian matematika yang valid. Hanya 10 artikel yang dinilai memenuhi kriteria awal untuk ditelaah secara mendalam pada tahap kesesuaian.

c. Tahap 3 (Kesesuaian)

Tahap penilaian kesesuaian dilakukan melalui evaluasi mendalam terhadap naskah lengkap 10 artikel yang telah lolos *screening*. Penilaian ini menerapkan kriteria inklusi dan eksklusi yang telah ditetapkan sebelumnya, dengan fokus pada aspek empiris penelitian, konteks pembelajaran matematika, kelengkapan laporan validitas dan reliabilitas instrumen, serta ketersediaan *full-text*. Hasil penilaian menunjukkan bahwa 8 artikel memenuhi kriteria inklusi dan 2 dieksklusi karena terdapat 1 artikel tidak tersedia *full-text* dan 1 artikel lainnya hanya berfokus pada berpikir komputasional secara umum tanpa integrasi konten matematika yang spesifik.

Penelitian ini juga menerapkan prosedur penilaian kualitas (*quality assessment*) untuk memastikan kredibilitas artikel yang dianalisis. Penelitian ini menerapkan prosedur penilaian kualitas berdasarkan kriteria yang diadaptasi dari Kitchenham & Charters [13]. Empat pertanyaan kunci berikut ini digunakan sebagai parameter penilaian:

1. (QA1) Apakah artikel mendeskripsikan model pengembangan instrumen?
2. (QA2) Apakah subjek penelitian dicantumkan dengan jelas?
3. (QA3) Apakah instrumen melalui tahap validasi ahli dan uji coba empiris?
4. (QA4) Apakah dalam menganalisis data menggunakan teknik yang kredibel (misal: Alpha Cronbach, Rasch Model)?

Kriteria pemberian skor dilakukan secara objektif dengan ketentuan berikut. Skor 1 diberikan apabila artikel menyajikan informasi secara lengkap, detail, dan didukung oleh data empiris yang jelas terkait indikator yang ditanyakan. Skor 0.5 diberikan apabila artikel membahas indikator tersebut namun tidak disertai penjelasan mendalam atau data yang kurang komprehensif. Skor 0 diberikan apabila artikel sama sekali tidak menyinggung atau memberikan informasi terkait indikator yang dimaksud. Klasifikasi kualitas ditentukan berdasarkan persentase skor total: Tinggi (>75%), Sedang (50%–75%), dan Rendah (<50%). Rincian skor untuk setiap artikel disajikan pada Tabel 2 berikut.

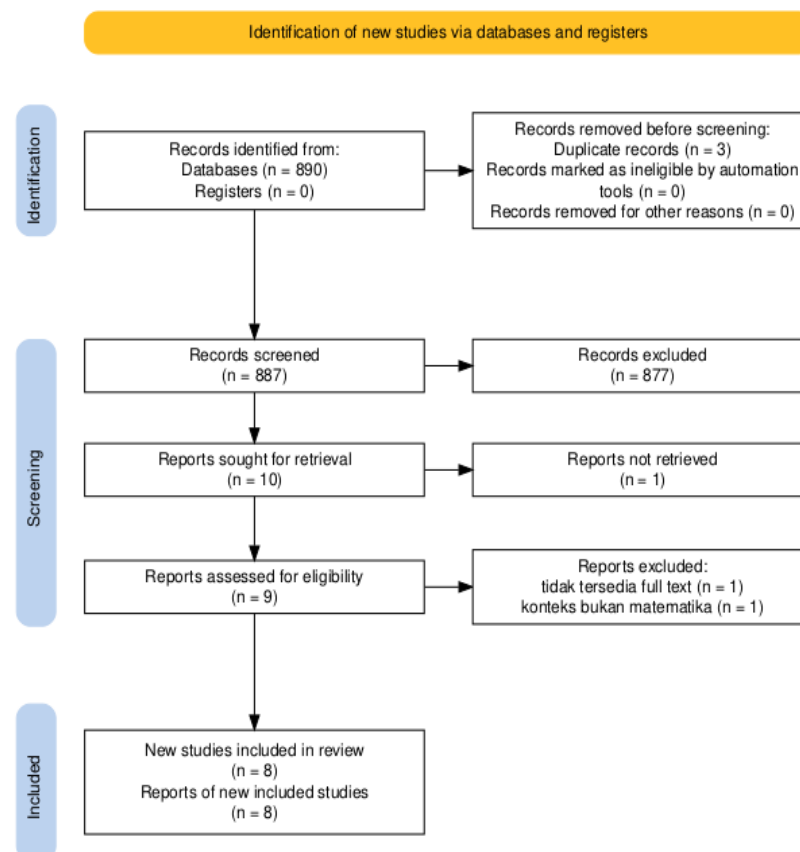
**Tabel 2.** Rincian Skor Penilaian Kualitas Artikel (*Quality Assessment*)

No.	Penulis dan Tahun	QA1	QA2	QA3	QA4	Total skor	Persentase	Kategori
1.	Inasari dkk., 2023 [14]	1	1	1	1	4	100%	Tinggi
2.	Zafrullah dkk., 2024 [15]	1	1	1	1	4	100%	Tinggi
3.	Haniah dan Waluyo, 2024 [16]	1	1	1	1	4	100%	Tinggi
4.	Maksum dkk., 2022 [17]	1	1	1	0.5	3.5	87.5%	Tinggi
5.	Novianto dkk., 2025 [18]	1	1	0.5	1	3.5	87.5%	Tinggi
6.	Munawarah dkk., 2021 [19]	1	1	1	0.5	3.5	87.5%	Tinggi
7.	Dobgenski dkk., 2025 [20]	1	1	1	0.5	3.5	87.5%	Tinggi
8.	Harti dkk., 2025 [21]	0.5	1	0.5	1	3	75%	Sedang

Berdasarkan Tabel 2, diperoleh tujuh artikel berkualitas tinggi dan satu artikel berkualitas sedang. Skor 0,5 pada QA1 merujuk pada penjabaran fase pengembangan yang kurang sistematis. Pada QA3, skor 0,5 diberikan jika data angka validasi ahli (seperti Aiken's V) tidak dipaparkan rinci. Sementara skor 0,5 pada QA4 diberikan karena penggunaan statistik klasik dasar tanpa pelaporan koefisien mendalam. Hal ini membuktikan bahwa *quality assessment* dilakukan secara objektif dan kritis terhadap kedalaman materi artikel.

d. Tahap 4 (Artikel Final)

Pada tahap final, 8 artikel yang telah dinyatakan memenuhi semua kriteria ditetapkan sebagai sampel utama dalam tinjauan sistematis ini. Artikel-artikel tersebut merepresentasikan penelitian-penelitian empiris terkini tentang pengembangan instrumen tes berpikir komputasional dalam konteks matematika yang akan menjadi dasar untuk ekstraksi data, analisis, dan sintesis temuan guna menjawab pertanyaan penelitian yang telah dirumuskan. Proses finalisasi artikel dapat dilihat pada Gambar 1 berikut.



**Gambar 1.** Diagram PRISMA Pengembangan Instrumen Tes Berpikir Komputasional Matematika

Berdasarkan Gambar 1, terlihat bahwa tahapan identifikasi dan seleksi studi untuk pengembangan instrumen tes berpikir komputasional matematika diawali dengan temuan 890

artikel dari berbagai basis data. Setelah menghapus tiga duplikat, sebanyak 887 artikel menjalani penyaringan. Sebanyak 877 artikel kemudian dikeluarkan berdasarkan penilaian judul dan abstrak, menyisakan 10 artikel. Dari penilaian menyeluruh artikel tersebut, 2 artikel lebih lanjut dikecualikan karena 1 artikel tidak tersedia *full-text* dan 1 artikel tidak memuat konteks matematika, sehingga hanya 8 studi yang memenuhi kriteria inklusi dan eksklusi yang akhirnya dimasukkan ke dalam tinjauan sistematis. Jumlah yang terbatas ini memberikan justifikasi kuat bahwa pengembangan instrumen tes CT yang spesifik dan terintegrasi dalam domain matematika masih merupakan bidang penelitian yang sangat baru dan spesifik. Fokus pada 8 artikel ini memungkinkan analisis yang lebih mendalam dan kritis terhadap karakteristik teknis masing-masing instrumen.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Karakteristik dan jenis instrumen tes berpikir komputasional dalam yang dikembangkan dalam pembelajaran matematika

Berdasarkan proses seleksi yang ditunjukkan dalam diagram PRISMA pada Gambar 1, sebanyak 8 studi memenuhi kriteria inklusi. Temuan dari artikel-artikel tersebut disajikan dalam Tabel 3 berikut ini.

**Tabel 3.** Karakteristik dan Jenis Instrumen Tes Berpikir Komputasional

No	Nama penulis dan tahun	Jenis Instrumen	Format Tes	Aspek Berpikir Komputasional yang Diukur	Bukti Validitas	Bukti Reliabilitas
1.	Dobgenski dkk., 2025 [20]	Learning objects (LO) berbasis Scratch	Aktivitas pemrograman Scratch	Abstraction, Decomposition, Algorithmic Design, Pattern Recognition, Evaluation, Generalization (Selby & Woollard, 2013); Konsep CT: variables, loops, conditionals, algorithms (Rachmatullah et al., 2020)	Validasi melalui framework Ye et al. (2023), Silva et al. (2020), dan Rachmatullah et al. (2020). Assessment grid El Mhouthi et al. (2013) menunjukkan kualitas akademik, pedagogik, didaktik, dan teknis yang excellent (skor total 85.9-87.2)	Tidak menggunakan reliabilitas statistik tradisional. Kualitas LO dinilai melalui assessment grid yang menunjukkan konsistensi dan keefektifan sebagai alat pembelajaran
2.	Munawarah dkk., 2021 [19]	Instrumen Tes CT berbasis RME	3 PG + 10 esai	Dekomposisi, Representasi Data, Pengenalan Pola, Algoritma, Generalisasi, Evaluasi	CVR & CVI dinyatakan valid	Cronbach's Alpha: PG = 0.781; Esai = 0.709
3.	Novianto dkk., 2025	Instrumen Penilaian	10 soal pilihan	Dekomposisi, Pengenalan Pola,	Validasi ahli: 86% (sangat layak)	Cronbach's Alpha = 0.85

No	Nama penulis dan tahun	Jenis Instrumen	Format Tes	Aspek Berpikir Komputasional yang Diukur	Bukti Validitas	Bukti Reliabilitas
	[18]	CT	ganda	Abstraksi, Algoritma		
4.	Harti dkk., 2025 [21]	Model Evaluasi Adaptif + Bank Soal Digital	Pre-test adaptif, LK adaptif, portofolio, proyek, ujian	Dekomposisi, Pengenalan Pola, Abstraksi, Algoritma	Validasi pakar (deskriptif)	Tidak disebutkan secara kuantitatif
5.	Zafrullah dkk., 2024 [15]	Instrumen Soal Literasi Numerasi	5 soal uraian	Problem Decomposition, Pattern Recognition, Algorithmic Thinking, Abstraction & Generalization	Aiken's V = 0.81; CFA loading factor > 0.5	CR > 0.6; AVE > 0.5 (kecuali PD = 0.54)
6.	Maksum dkk., 2022 [17]	Instrumen Tes CT pada Matematika	Uraian Terstruktur (dengan panduan jawaban)	Dekomposisi, Pengenalan Pola, Abstraksi & Generalisasi, Perancangan Algoritma	Validitas isi dengan indeks Aiken > 0,8 (tinggi)	Cronbach's Alpha: 0,933 (sangat reliabel)
7.	Haniah dan Waluyo, 2024 [16]	Tes pilihan ganda (20 butir)	Computer Based Test (CBT) via Quizizz	Decomposition, Abstraction, Pattern Recognition, Algorithmic Thinking	Validitas isi (Aiken index $\geq 0.8$ ), validitas butir (Rasch model: outfit ZSTD -2.0 hingga +2.0, MNSQ 0.5-1.5, Pt Measure Corr 0.4-0.85). Semua 20 butir valid.	Reliabilitas tes: 0.62 (cukup), reliabilitas butir: 0.88 (baik)
8.	Inasari dkk., 2023 [14]	Tes soal Bebras (12 butir) terkait matematika	Tidak disebutkan secara eksplisit	Decomposition, Pattern Recognition, Abstraction, Algorithms	Analisis Rasch model (item measure dan fit order). Pada tes ke-1: 59% sukar, 33% mudah, 8% sangat mudah; tes ke-2: 58% sukar, 42% mudah. Beberapa item misfit pada MNSQ tetapi masih layak.	Reliabilitas item tes ke-1: 0.85 (baik), reliabilitas person: 0.27 (lemah). Tes ke-2: reliabilitas item dan person lemah, Cronbach alpha 0.33

Temuan pertama mengungkapkan keragaman yang signifikan dalam hal jenis dan format instrumen tes berpikir komputasional, mulai dari *learning objects* berbasis pemrograman hingga

model evaluasi adaptif digital. Keragaman ini merefleksikan evolusi dalam paradigma asesmen berpikir komputasional, yang bergerak melampaui tes konvensional menuju asesmen kinerja dan berbasis proyek yang lebih autentik. Pergeseran ini selaras dengan argumen Grover dan Pea yang menekankan bahwa karena berpikir komputasional pada dasarnya adalah praktik pemecahan masalah, asesmen yang ideal harus mampu menangkap proses berpikir yang dinamis dan keterampilan konstruksi solusi, yang seringkali tidak terwakili secara memadai oleh format pilihan ganda [22]. Inovasi seperti instrumen berbasis *Scratch* secara langsung mengadopsi semangat ini dengan menilai berpikir komputasional melalui pembuatan artefak digital.

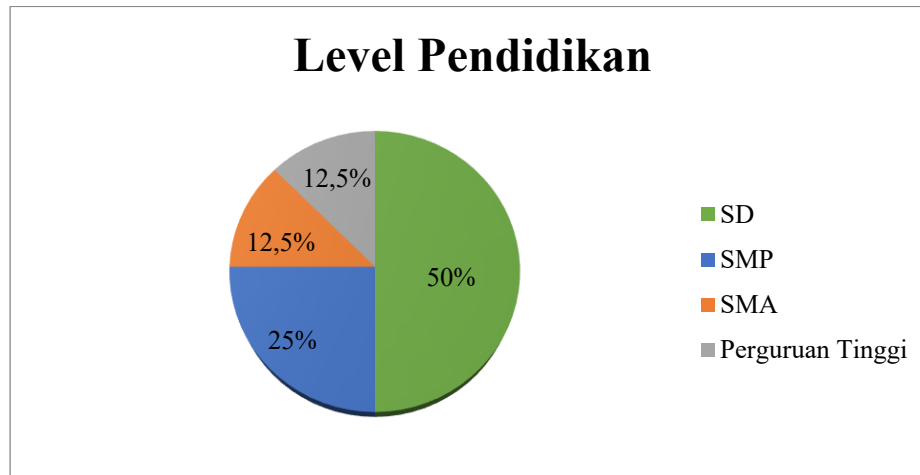
Meskipun formatnya beragam, terdapat konsistensi tinggi dalam aspek berpikir komputasional yang diukur. Kelima pilar utama berpikir komputasional yakni dekomposisi, pengenalan pola, abstraksi, berpikir algoritma, serta generalisasi/evaluasi dominan ada dalam instrumen-instrumen yang dikaji. Konsistensi ini mengindikasikan adanya konsensus di kalangan peneliti mengenai konstruk inti berpikir komputasional, yang sejalan dengan kerangka kerja seminal Selby dan Woollard untuk pendidikan dasar dan menengah [4]. Fokus pada aspek-aspek inti ini memastikan validitas konten dari instrumen yang dikembangkan, meskipun konteks dan format penerapannya berbeda.

Beberapa studi menunjukkan komitmen pada kualitas psikometrik melalui validasi ahli dan analisis statistik. Namun, ditemukan variasi reliabilitas yang kontras yakni dari kategori stabil di atas 0,60 [14], [16] hingga sangat rendah, yakni 0,33 [19], [20]. Rendahnya reliabilitas ini mengkonfirmasi temuan Tang bahwa pengukuran CT jauh lebih kompleks karena sifatnya yang multidimensi [23]. Secara kritis, Li dkk. dan Sung menjelaskan bahwa statistik klasik (Alpha Cronbach) sering kali gagal menangkap konsistensi internal CT karena setiap butir soal menguji pilar kognitif yang berbeda (seperti abstraksi vs algoritma). Meskipun reliabilitas sub-faktor rendah, Sung menegaskan instrumen tersebut tetap valid jika dianalisis melalui pendekatan psikometrik modern seperti Item Response Theory (IRT), yang terbukti lebih presisi dalam validasi kompetensi CT pada anak [24][25].

Dengan demikian, temuan ini memberikan implikasi penting bagi peneliti selanjutnya untuk tidak hanya terpaku pada skor reliabilitas tunggal, melainkan mengadopsi analisis IRT atau Model Rasch guna memisahkan kemampuan prosedural matematika dengan proses berpikir komputasional secara lebih akurat.

3.2. Level pendidikan pengembangan instrumen tes berpikir komputasional dalam pembelajaran matematika

Analisis artikel menunjukkan bahwa pengembangan instrumen tes berpikir komputasional mencakup seluruh spektrum pendidikan formal, dari sekolah dasar hingga pendidikan tinggi (calon guru) seperti digambarkan pada Gambar 2 berikut.



**Gambar 2.** Level Pendidikan Pengembangan Instrumen Tes Berpikir Komputasional dalam Pembelajaran Matematika

Berdasarkan Gambar 2 di atas, hasil pemetaan menunjukkan pola distribusi yang tidak merata, di mana pengembangan instrumen CT didominasi oleh jenjang Sekolah Dasar (50%), jauh melampaui jenjang SMA dan Perguruan Tinggi yang hanya terwakili masing-masing sebesar 12,5%. Dominasi pengembangan instrumen CT di tingkat SD ini dapat dijelaskan oleh beberapa faktor. Pertama, terdapat momentum global untuk mengenalkan konsep pemikiran komputasional dan *coding* sejak dini, sebagai bagian dari literasi dasar abad ke-21. Kajian Bocconi mengenai berpikir komputasional menegaskan pentingnya memulai pengenalan berpikir komputasional di usia dini untuk membentuk *mindset* dan keterampilan dasar [26]. Kedua, kurikulum matematika di SD kaya dengan topik-topik seperti pola bilangan dan aritmatika yang secara alami dapat diintegrasikan dengan aktivitas berpikir komputasional, misalnya melalui pengenalan pola dan penyusunan algoritma sederhana, menciptakan sinergi yang efektif untuk pembelajaran awal. Selain itu pada jenjang SD, CT lebih mudah diukur melalui media visual atau *unplugged* (tanpa komputer). Artikel [20] menunjukkan penggunaan *Scratch* yang bersifat blok visual sangat cocok untuk tahap kognitif anak SD, sehingga peneliti lebih banyak menciptakan instrumen di level ini.

Namun demikian, dominasi pada jenjang SD ini menunjukkan adanya "diskontinuitas kognitif" dalam pengembangan instrumen. Sedikitnya jumlah studi di level SMA dan perguruan tinggi menyisakan celah besar dalam literatur. Dampaknya, kemampuan CT siswa tidak terpantau secara berkelanjutan ketika mereka mulai menghadapi konten matematika yang lebih

abstrak. Padahal CT pada level yang lebih tinggi sangat krusial untuk topik seperti kalkulus atau statistika. Sebagaimana taksonomi Weintrop, praktik tingkat lanjut seperti *modeling* and *simulation* masih belum terjamah oleh instrumen yang ada [6]. Keterbatasan ini mengimplikasikan bahwa instrumen yang ada saat ini belum mampu memotret kematangan berpikir komputasional siswa secara holistik hingga jenjang profesional.

### 3.3 Topik matematika yang dikembangkan dalam tes berpikir komputasional dalam pembelajaran matematika

Pertanyaan penelitian ini bertujuan untuk mengidentifikasi topik-topik matematika yang menjadi konteks dalam pengembangan instrumen tes berpikir komputasional. Hasil analisis menunjukkan bahwa instrumen tes berpikir komputasional selalu diintegrasikan ke dalam materi matematika tertentu. Berikut adalah ringkasan pemetaan topik matematika di setiap artikel yang akan ditunjukkan pada Tabel 4 berikut ini.

**Tabel 4.** Pemetaan Topik Matematika

No.	Penulis (Tahun)	Topik Matematika yang Diintegrasikan
1.	Maksum dkk., 2022 [17]	Pola Bilangan, Generalisasi Pola, Aritmatika
2.	Munawaroh dkk., 2021 [19]	Sistem Persamaan Linear Dua Variabel (SPLDV), Bangun Datar, Lingkaran
3.	Novianto dkk., 2025 [18]	Himpunan Bagian, Perbandingan Berbalik Nilai
4.	Dobgenski dkk., 2025 [20]	Orientasi Geometris, Bidang Kartesius
5.	Zafrullah dkk., 2024 [15]	Bentuk Aljabar, Numerasi
6.	Inasari dkk., 2023 [14]	Konteks Bebras (Numerasi dan Pemecahan Masalah)
7.	Haniah dan Waluyo, 2024 [16]	Aritmatika, Perbandingan
8.	Harti dkk., 2025 [21]	Konteks PISA/AKM (Literasi Numerasi, Pemodelan Matematika)

Integrasi berpikir komputasional (CT) dalam pembelajaran matematika ditemukan tidak terbatas pada satu topik, melainkan tersebar ke dalam tiga domain utama yakni aritmetika, aljabar, dan geometri (Tabel 4).

- a. Aritmetika dan Pola: Topik yang paling dominan adalah yang melibatkan konsep pola dan generalisasi. Tiga studi (menggunakan pola bilangan, pola gambar, dan perbandingan (aritmetika) sebagai inti soal [16], [17], [18]. Topik-topik ini sangat relevan untuk menilai aspek pengenalan pola dan abstraksi dalam berpikir komputasional.
- b. Aljabar: Domain aljabar, seperti bentuk aljabar [15] dan sistem persamaan linear dua variabel (SPLDV) [19], menjadi konteks yang penting. Aljabar dipilih karena secara

inheren membutuhkan kemampuan abstraksi untuk bekerja dengan variabel dan memodelkan masalah, yang merupakan pilar kunci berpikir komputasional.

- c. Geometri dan Koordinat: Topik geometri, termasuk bidang kartesius, orientasi geometris, dan bangun datar [19], [20] digunakan sebagai landasan instrumen. Domain ini sangat efektif untuk mengukur berpikir algoritma, terutama pada instrumen berbasis visual atau pemrograman (coding), di mana siswa harus merancang urutan langkah yang presisi untuk memanipulasi objek.

Selain topik spesifik kurikulum, dua studi menggunakan konteks numerasi umum seperti soal Bebras dan PISA/AKM [14], [21]. Hal ini menunjukkan tren pengembangan instrumen yang mengukur berpikir komputasional dalam pemecahan masalah yang lebih kontekstual dan lintas-topik, daripada terikat pada satu bab materi saja.

Analisis terhadap topik matematika mengungkap bahwa berpikir komputasional terintegrasi penuh dengan konten matematika yang spesifik. Integrasi ini menjadikan asesmen berpikir komputasional lebih kontekstual dan bermakna bagi siswa, sekaligus memperkuat pemahaman matematis.

Pada topik aritmetika dan pola menunjukkan dominasinya yang dikarenakan adanya hubungan simbiosis mutualisme antara berpikir komputasional dan pencarian pola. Aktivitas seperti menggeneralisasi pola bilangan atau gambar secara langsung melatih dan mengukur kemampuan *pattern recognition* dan *abstraction*, yang merupakan jantung dari *computational thinking*. Wing sendiri, dalam artikel seminarnya, mendefinisikan berpikir komputasional sebagai proses pengenalan pola dan abstraksi [3].

Penggunaan topik aljabar, seperti SPLDV dan bentuk aljabar, menegaskan peran berpikir komputasional dalam pemodelan matematika. Nurlaelah menunjukkan bahwa pendekatan berpikir komputasional dapat meningkatkan kemampuan pembuktian matematika, termasuk dalam aljabar abstrak, dengan menekankan aspek logika dan konstruksi solusi yang sistematis [7]. Hal ini memperkuat gagasan bahwa aljabar lebih dari sekadar manipulasi simbol, melainkan proses berpikir komputasional untuk memodelkan dan menyusun solusi matematis yang abstrak dan kompleks.

Topik Geometri dan Koordinat, terutama dalam instrumen berbasis pemrograman memberikan konteks yang ideal untuk mengukur *algorithmic thinking*. Hal ini sejalan dengan penelitian Hidayati yang menunjukkan bahwa pembelajaran geometri berbasis GeoGebra memperkuat kemampuan berpikir komputasional siswa, terutama dalam aspek visualisasi, pengenalan pola, dan algoritmik yang sangat penting dalam pemecahan masalah geometri

kompleks [27]. Penelitian ini juga menemukan bahwa gaya kognitif siswa memengaruhi penguasaan berpikir komputasional, dimana siswa dengan gaya kognitif *Field Independent* (FI) lebih unggul dalam indikator berpikir komputasional seperti abstraksi dan generalisasi dibandingkan siswa *Field Dependent* (FD). Pendekatan ini membuka ruang bagi metode pembelajaran interaktif yang memadukan teknologi dan *scaffolding* untuk memfasilitasi pengembangan berpikir komputasional secara efektif di materi geometri, termasuk topik transformasi dan koordinat.

Selain topik kurikuler yang spesifik, muncul tren penggunaan konteks numerasi umum seperti soal Bebras dan PISA [14], [21]. Tren ini mengindikasikan perluasan fokus dari penguasaan konten menuju penerapan berpikir komputasional dalam pemecahan masalah yang bersifat lintas-disiplin dan kontekstual. Studi sistematis oleh Triantafyllou menegaskan bahwa *Bebras Challenge*, yang menggunakan tantangan yang menarik dan gamifikasi, efektif dalam meningkatkan kemampuan berpikir komputasional siswa, khususnya dalam konteks matematika dan pemecahan masalah abad ke-21 [21]. Penggunaan Bebras ini sejalan dengan tujuan pendidikan modern yang menekankan kemampuan berpikir fleksibel dan aplikatif.

#### 4. KESIMPULAN

Penelitian ini menyimpulkan bahwa pengembangan instrumen berpikir komputasional (CT) dalam matematika saat ini masih menunjukkan dominasi yang sangat kuat pada jenjang sekolah dasar dengan angka mencapai 50%. Dominasi ini mengakibatkan ketergantungan pada topik matematika yang bersifat konkret seperti aritmetika, sehingga efektivitas instrumen pada materi yang lebih abstrak pada jenjang SMA dan perguruan tinggi masih sangat terbatas. Temuan kritis dalam studi ini menyoroti bahwa penggunaan statistik klasik belum menghasilkan reliabilitas yang stabil pada instrumen CT yang bersifat multidimensi, sebagaimana ditunjukkan oleh koefisien rendah pada beberapa artikel yang ditinjau.

Instrumen CT dalam matematika yang ideal harus mengintegrasikan pilar kognitif CT secara langsung ke dalam konten matematika yang kompleks serta mengadopsi analisis psikometrik modern seperti Model Rasch atau IRT guna menjamin ketajaman pengukuran. Secara praktis, guru dan peneliti direkomendasikan untuk beralih dari tes konvensional menuju asesmen berbasis masalah yang mampu memotret alur pikir algoritma siswa secara lebih presisi. Mengingat keterbatasan jumlah literatur yang dianalisis, penelitian masa depan perlu memperluas jangkauan basis data guna memperkuat generalisasi temuan dan menutup celah pengukuran pada jenjang pendidikan yang lebih tinggi.

## 5. DAFTAR PUSTAKA

- [1] E. Raziana, H. Wibawanto, U. I. Indonesia, U. S. Maret, and C. Author, “Pedagogical Innovation of Game-Based Learning for 21st Century Skills in Vocational Education : A Systematic Literature Review,” *ijolii Indones. J. Learn. Instr. Innov.*, vol. 3, no. 1, pp. 19–30, 2025, doi: <https://journal.uns.ac.id/ijolii> Pedagogical.
- [2] A. H. Alifah and S. Widodo, “Membuka Kemampuan *Computational thinking* Sebagai 21 Century Skills Disiplin STEM,” *Fakt. J. Ilm. Kependidikan*, vol. 12, no. 1, pp. 100–108, 2024, doi: <https://doi.org/10.30998/fjik.v11i1.21962>.
- [3] J. M. Wing, “*Computational thinking*,” *Commun. ACM*, vol. 49, no. 3, pp. 33–35, 2006.
- [4] C. C. Selby and J. Woollard, “*Computational thinking* : The Developing Definition,” 2013..
- [5] D. Kaya, A. Ö. Yaşar, İ. Çetin, and T. Kutluca, “The Relationship Between the 21st-Century Skills and *Computational thinking* Skills of Prospective Mathematics and Science Teachers,” *J. Pedagog. Res.*, vol. 9, no. 1, pp. 73–95, 2025, doi: <https://doi.org/10.33902/JPR.202531498>.
- [6] D. Weintrop *et al.*, “Defining *Computational thinking* for Mathematics and Science Classrooms,” *J. Sci. Educ. Technol.*, vol. 25, no. 1, pp. 127–147, 2016, doi: 10.1007/s10956-015-9581-5.
- [7] E. Nurlaelah, A. Pebrianti, M. Taqiyuddin, J. A. Dahlan, and D. Usdiyana, “Improving Mathematical Proof Based on *Computational thinking* Components for Prospective Teachers in Abstract Algebra Courses,” *Infin. J. Math. Educ.*, vol. 14, no. 1, pp. 85–108, 2025, doi: <https://doi.org/10.22460/infinity.v14i1.p85-108>.
- [8] E. Irawan, R. Rosjanuardi, and S. Prabawanto, “Research Trends of *Computational Thinking* in Mathematics Learning : A Bibliometric Analysis from 2009 to 2023,” *EURASIA J. Math. Sci. Technol. Educ.*, vol. 20, no. 3, 2024, doi: <https://doi.org/10.29333/ejmste/14343>
- [9] M. P. Listiowati, T. Astuti, and B. Subali, “Effectiveness of Coding in Enhancing *Computational Thinking* among Elementary School Students : A Literature Review 2020 – 2025 Coding teaching has been introduced since the elementary school ( SD ) level , as conveyed by the Vice President of RI , Gibran,” *Edunesia J. Ilm. Pendidik.*, vol. 6, no. 3, pp. 1766–1784, 2025, doi: <https://doi.org/10.51276/edu.v6i3.1299>.
- [10] J. Bilbao *et al.*, “Analytical Methods and Tools for Evaluating the Development of *Computational Thinking* Abilities,” *Int. J. Educ. Inf. Technol.*, vol. 19, 2025, doi:

- 10.46300/9109.2025.19.6.
- [11] Ö. Korkmaz and X. Bai, “Adapting *Computational Thinking* Scale ( CTS ) for Chinese High School Students and Their Thinking Scale Skills Level,” *Particip. Educ. Res.*, vol. 6, no. 1, pp. 10–26, 2019, doi: <http://dx.doi.org/10.17275/per.19.2.6.1>.
- [12] A. Liberati *et al.*, “The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration,” *PLoS Med*, vol. 6, no. 7, 2009, doi: 10.1371/journal.pmed.1000100.
- [13] B. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering,” 2007.
- [14] L. Inasari, D. Lidinillah, and A. Prehanto, “Pengembangan Instrumen Tes *Computational Thinking* Siswa Sekolah Dasar melalui Analisis RASCH Model,” *COLLASE (Creative Learn. Students Elem. Educ.*, vol. 6, no. 1, pp. 102–110, 2023, doi: <https://doi.org/10.22460/collase.v1i1.16188>.
- [15] Z. Zafrullah, S. Hamdi, A. Wahyuni, R. Safitri, R. N. Gunawan, and Y. Istiawanto, “Development of Numerical Literacy Question Instrument based on *Computational Thinking* for Mathematics Learning,” *AL-ISHLAH J. Pendidik.*, vol. 16, no. 4, pp. 4489–4502, 2024, doi: 10.35445/alishlah.v16i4.5985..
- [16] R. N. Haniah and M. Waluyo, “Development of Computational Mathematic Thinking Test Instruments Based On Computerized Based Test,” *Mathline J. Mat. dan Pendidik. Mat.*, vol. 9, no. 3, pp. 675–690, 2024, doi: 10.31943/mathline.v9i3.648.
- [17] K. Maksum, N. Afifah, M. Ardiyaningrum, and S. Sukati, “Pengembangan Instrumen Tes Keterampilan Berpikir Komputasi pada Pelajaran Matematika Sekolah Dasar (SD)/Madrasah Ibtida’iyah (MI),” *Model. J. Progr. Stud. PGMI*, vol. 09, no. 01, pp. 39–53, 2022, [Online]. Available: <http://jurnal.stitnualhikmah.ac.id/index.php/modeling/article/view/1038%0Ahttps://www.jurnal.stitnualhikmah.ac.id/index.php/modeling/article/download/1038/708>.
- [18] A. Novianto, I. L. Il Maknun, N. Y. N. Aliyah, and A. N. Khalifah, “Pengembangan Instrumen Penilaian *Computational Thinking* pada Pembelajaran Matematika SD,” *Kalam Cendekia J. Ilm. Kependidikan*, vol. 13, no. 1, 2025, doi: 10.20961/jkc.v13i1.93072.
- [19] Munawarah, S. Z. Thalhah, A. D. Angriani, F. Nur, and A. Kusumayanti, “Development of Instrument Test *Computational Thinking* Skills IJHS/JHS Based RME Approach,” *Math. Teaching-Research J.*, vol. 13, no. 4, pp. 202–220, 2021.
- [20] J. Dobgenski, M. E. B. Brito Prado, and A. da F. Garcia Silva, “Development and

- Validation of Learning Objects Aimed at Mathematical Instructions Based On *Computational Thinking*,” *Int. Electron. J. Math. Educ.*, vol. 20, no. 2, p. em0812, 2025, doi: 10.29333/iejme/15816.
- [21] Y. P. Harti, A. Agustin, L. Sari, and Mukarom, “Model Evaluasi Adaptif Berbasis *Computational Thinking* untuk Meningkatkan Kompetensi Pedagogik Calon Guru di Era Digital (Studi Kasus Pada Mahasiswa PPG),” *Econ. Educ. J.*, vol. 7, no. 2, pp. 426–442, 2025, doi: 10.33503/education.v7i2.1847.
- [22] Shuchi Grover and Roy Pea, “*Berpikir komputasional* in K–12: A Review of the State of the Field,” *Educ. Res.*, vol. 42, no. 1, pp. 38–43, Jan. 2013, doi: 10.3102/0013189X12463051.
- [23] X. Tang, Y. Yin, Q. Lin, R. Hadad, and X. Zhai, “Computers & Education Assessing *Computational Thinking* : A systematic review of empirical studies,” *Comput. Educ.*, vol. 148, no. December 2019, p. 103798, 2020, doi: 10.1016/j.compedu.2019.103798.
- [24] Y. Li, S. Xu, and J. Liu, “Development and Validation of *Computational Thinking* Assessment of Chinese Elementary School Students,” *Pacific Rim Psychol.*, vol. 15, no. 22, 2021, doi: 10.1177/18344909211010240..
- [25] J. Sung, “Assessing young Korean children ’ s *Computational Thinking* : A validation study of two measurements,” *Educ. Inf. Technol.*, vol. 27, pp. 12969–12997, 2022, doi: 10.1007/s10639-022-11137-x.
- [26] S. Bocconi, A. Chiocciariello, G. Dettori, A. Ferrari, K. Engelhardt, and Y. Punie, *Developing Computational Thinking in Compulsory Education*. Luxembourg: Publications Office of the European Union, 2016. doi: 10.2791/792158.
- [27] Y. M. Hidayati, B. A. Sukowati, and W. Hastuti, “Kemampuan Berpikir Komputasi pada Pembelajaran Geometri Berbasis Geogebra Ditinjau dari Gaya Kognitif,” *Histogram J. Pendidik. Mat.*, vol. 9, no. 1, pp. 74–89, 2025, doi: <http://journal.stkip-andi-matappa.ac.id/index.php/histogram/index>.