

PENGEMBANGAN DAN VALIDASI INSTRUMEN BERPIKIR KOMPUTASIONAL MATERI SISTEM KOMPUTER

. R. Rahmadhani¹, D. Djuniadi²

¹²Program Studi Penelitian dan Evaluasi Pendidikan
Universitas Negeri Semarang
Semarang, Indonesia

e-mail: rarahmadhani@students.unnes.ac.id¹ djuniadi@mail.unnes.ac.id²

Abstrak

Penelitian ini bertujuan mengembangkan dan memvalidasi instrumen tes berpikir komputasional pada materi Sistem Komputer untuk siswa SMP. Penelitian menggunakan metode *Research and Development* (R&D) dengan model 4D yang meliputi tahap *define*, *design*, dan *develop*. Subjek penelitian terdiri atas tiga validator ahli dan 15 siswa kelas VIII SMP Negeri 9 Langsa pada uji coba terbatas. Data dikumpulkan melalui observasi, wawancara, dokumentasi, lembar validasi ahli, dan tes. Analisis data menggunakan indeks Aiken's V, korelasi *Pearson product-moment*, dan koefisien *Cronbach's alpha*. Hasil penelitian menunjukkan bahwa instrumen memperoleh nilai Aiken's V sebesar 0,967 yang menunjukkan validitas isi yang tinggi. Pada uji coba terbatas, sebanyak 17 dari 20 butir menunjukkan bukti validitas empiris awal, sedangkan tiga butir tidak memenuhi kriteria yang ditetapkan. Koefisien *Cronbach's alpha* sebesar 0,918 diperoleh dari 20 butir pada uji coba awal dan menunjukkan konsistensi internal yang tinggi. Hasil tersebut memberikan bukti awal mengenai validitas isi, validitas empiris butir, dan konsistensi internal instrumen pada uji coba terbatas. Penelitian selanjutnya disarankan melakukan uji coba pada skala yang lebih luas untuk memperoleh bukti empiris yang lebih kuat dan komprehensif.

Kata kunci: Berpikir Komputasional; Instrumen Penilaian; Sistem Komputer; Validitas; Reliabilitas

Abstract

This study aimed to develop and validate a computational thinking test instrument for the Computer Systems topic in junior secondary schools. The study employed a Research and Development (R&D) approach using the 4D model, consisting of the define, design, and develop stages. The participants included three expert validators and 15 eighth-grade students of SMP Negeri 9 Langsa involved in a limited trial. Data were collected through observation, interviews, documentation, expert validation sheets, and tests. Data were analyzed using the Aiken's V index, Pearson product-moment correlation, and Cronbach's alpha coefficient. The results indicated that the instrument achieved an Aiken's V value of 0.967, indicating high content validity. In the limited trial, 17 of the 20 items provided initial evidence of empirical validity, while three items did not meet the predetermined criteria. A Cronbach's alpha coefficient of 0.918 was obtained from the 20 items in the initial trial and indicated high internal consistency. These findings provide initial evidence of the instrument's content validity, empirical item validity, and internal consistency in the limited trial. Future research should conduct larger-scale trials to obtain stronger and more comprehensive empirical evidence.

Keywords: Computational Thinking; Assessment Instrument; Computer Systems; Validity; Reliability

PENDAHULUAN

Pendidikan abad ke-21 menuntut siswa tidak hanya menguasai pengetahuan, tetapi juga memiliki kemampuan berpikir tingkat tinggi, pemecahan masalah, kreativitas, komunikasi, kolaborasi, serta literasi digital yang memadai untuk menghadapi perkembangan teknologi yang semakin pesat (Dilekçi & Karatay, 2023). Binkley (2012)

menjelaskan bahwa pendidikan modern perlu membekali siswa dengan keterampilan yang memungkinkan mereka menggunakan pengetahuan secara kritis dan efektif dalam berbagai situasi kehidupan. Sejalan dengan hal tersebut, Nugroho (2025) menegaskan bahwa keberhasilan siswa pada era digital tidak lagi ditentukan oleh kemampuan mengingat informasi, melainkan oleh kemampuan menganalisis, mengevaluasi, dan memanfaatkan informasi untuk menyelesaikan permasalahan yang kompleks. Oleh karena itu, pembelajaran di sekolah perlu dirancang untuk mengembangkan kemampuan berpikir yang relevan dengan tuntutan abad ke-21 (UNESCO, 2023).

Dalam konteks pendidikan di Indonesia, implementasi Kurikulum Merdeka semakin memperkuat pentingnya pengembangan kompetensi abad ke-21 melalui pembelajaran yang berpusat pada siswa dan berorientasi pada penguatan kompetensi. Salah satu mata pelajaran yang memiliki peran strategis dalam mewujudkan tujuan tersebut adalah Informatika (UNESCO, 2025). Melalui pembelajaran Informatika, siswa tidak hanya belajar menggunakan teknologi, tetapi juga belajar memahami cara kerja sistem, mengolah informasi, serta menyelesaikan masalah secara sistematis. Kemendikbudristek (2021) menegaskan bahwa salah satu elemen penting dalam pembelajaran Informatika adalah kemampuan berpikir komputasional yang menjadi dasar dalam pengembangan kemampuan pemecahan masalah pada era digital.

Berpikir komputasional (*computational thinking*) merupakan proses berpikir untuk menyelesaikan masalah secara logis, sistematis, dan efisien dengan menerapkan prinsip-prinsip ilmu komputer. Wing (2017) menyatakan bahwa berpikir komputasional merupakan keterampilan dasar yang perlu dimiliki setiap individu, tidak hanya bagi mereka yang mempelajari pemrograman. Menurut Shute et al. (2017), berpikir komputasional terdiri atas empat komponen utama, yaitu *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Keempat komponen tersebut memungkinkan siswa mengidentifikasi masalah, mengenali pola yang muncul, menyederhanakan informasi yang kompleks, dan menyusun langkah penyelesaian secara sistematis. Dalam konteks asesmen, keempat komponen tersebut perlu diterjemahkan ke dalam indikator yang dapat diamati melalui respons siswa agar kemampuan berpikir komputasional dapat diukur secara empiris. Dengan demikian, kemampuan berpikir komputasional menjadi salah satu indikator penting yang perlu dikembangkan dan diukur dalam pembelajaran Informatika (Kemendikbudristek, 2022).

Salah satu materi yang memiliki keterkaitan erat dengan kemampuan berpikir komputasional adalah Sistem Komputer. Materi ini membahas hubungan antara perangkat *input*, *process*, *output*, dan *storage* dalam mengolah data menjadi informasi. Menurut Tanenbaum & Austin (2013) sistem komputer bekerja melalui serangkaian proses yang saling berhubungan dan terorganisasi secara logis. Karakteristik tersebut memberikan peluang bagi siswa untuk mengembangkan kemampuan dekomposisi ketika mengidentifikasi komponen sistem, mengenali pola pada alur pemrosesan data, melakukan abstraksi terhadap proses yang kompleks, serta memahami langkah-langkah kerja sistem komputer secara berurutan. Oleh karena itu, materi Sistem Komputer memiliki potensi yang besar untuk digunakan sebagai konteks dalam pengukuran kemampuan berpikir komputasional siswa (Rachmawati et al., 2024). Konteks tersebut juga memungkinkan pengukuran berpikir komputasional melalui aktivitas non-pemrograman, sehingga kemampuan siswa dapat diamati melalui proses menganalisis komponen, hubungan, pola, penyederhanaan informasi, dan urutan kerja sistem komputer.

Namun demikian, hasil observasi awal menunjukkan bahwa pembelajaran Sistem Komputer masih menghadapi berbagai kendala. Proses pembelajaran cenderung berpusat pada guru dan lebih menekankan penguasaan konsep secara teoritis. Siswa umumnya mampu menyebutkan fungsi perangkat komputer, tetapi masih mengalami kesulitan dalam menjelaskan keterkaitan antar komponen sistem komputer maupun menganalisis alur pengolahan data yang terjadi. Kondisi tersebut menunjukkan bahwa kemampuan berpikir komputasional siswa belum berkembang secara optimal. Temuan ini diperkuat melalui diskusi dengan guru-guru Informatika dalam forum MGMP yang mengungkapkan bahwa

instrumen evaluasi yang digunakan selama ini lebih berorientasi pada pengukuran pengetahuan faktual dan belum dirancang untuk mengukur kemampuan berpikir komputasional secara spesifik. Kondisi tersebut menunjukkan adanya kebutuhan terhadap instrumen tes yang dapat memberikan informasi lebih spesifik mengenai kemampuan berpikir komputasional siswa pada materi Sistem Komputer.

Berbagai penelitian telah menunjukkan pentingnya pengembangan instrumen untuk mengukur kemampuan berpikir komputasional. Korkmaz et al. (2017) mengembangkan *Computational Thinking Scales* dan menunjukkan bahwa konstruk berpikir komputasional dapat diukur melalui instrumen yang disusun berdasarkan indikator yang jelas. Shute et al. (2017) juga menekankan pentingnya pengukuran berpikir komputasional berdasarkan komponen dan keterampilan yang dapat diamati. Penelitian El-Hamamsy et al. (2025) selanjutnya menunjukkan bahwa instrumen berpikir komputasional dapat dikembangkan dengan memperhatikan aspek validitas dan reliabilitas sehingga menghasilkan alat ukur yang dapat digunakan untuk memperoleh bukti empiris mengenai kemampuan siswa. Temuan-temuan tersebut menunjukkan bahwa pengembangan instrumen berpikir komputasional perlu memperhatikan kesesuaian antara konstruk, indikator, konteks pembelajaran, dan karakteristik siswa.

Meskipun demikian, masih terdapat ruang untuk mengembangkan instrumen tes berpikir komputasional yang secara khusus menggunakan materi Sistem Komputer sebagai konteks non-pemrograman pada jenjang SMP. Penelitian terdahulu mengenai berpikir komputasional banyak dikembangkan dalam konteks pemrograman, aktivitas pembelajaran umum, atau pengukuran konstruk berpikir komputasional secara umum. Kesenjangan tersebut menunjukkan perlunya instrumen yang secara eksplisit menghubungkan empat komponen berpikir komputasional dengan konteks materi Sistem Komputer, indikator yang dapat diamati, stimulus soal, dan kriteria penskoran. Dalam penelitian ini, *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking* masing-masing dijabarkan ke dalam indikator dan lima butir soal sehingga menghasilkan 20 butir soal uraian.

Berdasarkan kesenjangan tersebut, penelitian ini mengembangkan instrumen tes berpikir komputasional pada materi Sistem Komputer yang disajikan secara digital sebagai sarana pendukung pelaksanaan asesmen. Kebaruan penelitian terletak pada pengoperasian empat komponen berpikir komputasional ke dalam soal uraian berbasis konteks Sistem Komputer sebagai materi non-pemrograman, yang dilengkapi dengan indikator terukur dan rubrik penskoran. Dengan demikian, format digital tidak ditempatkan sebagai konstruk utama penelitian, tetapi sebagai sarana penyajian instrumen agar pelaksanaan asesmen lebih mudah diakses dan dikelola. Penelitian ini bertujuan untuk mengembangkan instrumen tes berpikir komputasional pada materi Sistem Komputer serta memperoleh bukti awal mengenai validitas isi, validitas empiris butir, dan konsistensi internal instrumen melalui uji coba terbatas.

METODE

Penelitian ini merupakan penelitian dan pengembangan (*Research and Development/R&D*) yang menggunakan model 4D yang dikemukakan oleh Thiagarajan et al. (1974). Model pengembangan tersebut terdiri atas tahap *define*, *design*, *develop*, dan *disseminate*. Penelitian ini dibatasi sampai tahap *develop*, karena fokus penelitian adalah menghasilkan instrumen tes berpikir komputasional dengan penyajian digital pada materi Sistem Komputer serta memperoleh bukti awal mengenai validitas isi, validitas empiris butir dan konsistensi internal instrumen.

Penelitian dilaksanakan di SMP Negeri 9 Langsa pada semester genap tahun pelajaran 2025/2026. Subjek penelitian terdiri atas tiga validator ahli, yaitu satu dosen dan dua guru Informatika yang memiliki kompetensi relevan dalam bidang pembelajaran, evaluasi pendidikan, dan media pembelajaran, serta 15 siswa kelas VIII yang terlibat dalam uji coba terbatas. Objek penelitian adalah instrumen tes berpikir komputasional pada materi Sistem Komputer yang disajikan secara digital. Instrumen dirancang untuk mengukur empat

komponen berpikir komputasional, yaitu *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking* (Shute et al., 2017; Wing, 2017).

Tahap *define* dilakukan melalui analisis kebutuhan, analisis kurikulum, analisis karakteristik siswa, dan analisis materi Sistem Komputer. Tahap *design* meliputi penetapan konstruk yang diukur, penyusunan definisi operasional dan indikator, penyusunan kisi-kisi instrumen, penulisan butir soal, penyusunan rubrik penskoran, serta perancangan penyajian instrumen secara digital menggunakan Canva yang diintegrasikan ke dalam platform Liveworksheets. Instrumen dikembangkan dalam bentuk 20 soal uraian, yang terdiri atas lima butir untuk masing-masing komponen berpikir komputasional, yaitu *decomposition* (butir 1–5), *pattern recognition* (butir 6–10), *abstraction* (butir 11–15), dan *algorithmic thinking* (butir 16–20). Setiap butir disusun berdasarkan indikator yang telah ditetapkan dan menggunakan stimulus berupa konteks atau permasalahan yang berkaitan dengan materi Sistem Komputer (Thiagarajan et al., 1974).

Setiap butir soal uraian dinilai menggunakan rubrik penskoran dengan rentang skor 1–4. Penskoran dilakukan berdasarkan tingkat keterpenuhan kriteria respons yang telah ditetapkan untuk setiap butir. Skor 4 diberikan ketika respons siswa memenuhi seluruh kriteria yang dipersyaratkan, sedangkan skor 1 diberikan ketika respons tidak memenuhi atau hanya memenuhi sebagian kecil kriteria. Skor 2 dan 3 menunjukkan tingkat keterpenuhan kriteria di antara kedua kondisi tersebut. Dengan demikian, skor setiap butir merepresentasikan tingkat ketercapaian indikator berpikir komputasional yang diukur.

Tahap *develop* dilakukan melalui validasi ahli, revisi produk berdasarkan masukan validator, dan uji coba terbatas untuk memperoleh bukti awal kualitas instrumen yang dikembangkan (Thiagarajan et al., 1974). Data penelitian dikumpulkan melalui observasi, wawancara, dokumentasi, lembar validasi ahli, dan hasil pengerjaan instrumen oleh siswa. Observasi dan wawancara digunakan untuk mengidentifikasi kebutuhan pengembangan instrumen, sedangkan lembar validasi digunakan untuk memperoleh penilaian ahli terhadap aspek materi, media, dan evaluasi (Korkmaz et al., 2017). Data hasil uji coba terbatas diperoleh dari respons siswa terhadap instrumen yang telah dikembangkan.

Data hasil validasi ahli dianalisis menggunakan indeks Aiken's V untuk memperoleh bukti validitas isi instrumen (Aiken, 1985). Lembar validasi menggunakan skala Likert lima tingkat dengan 20 pernyataan. Tiga validator memberikan penilaian terhadap setiap pernyataan berdasarkan skala 1–5. Nilai Aiken's V dihitung berdasarkan skor yang diberikan validator dengan mempertimbangkan skor terendah, skor tertinggi, dan jumlah validator. Hasil analisis validitas isi kemudian disajikan dalam bentuk nilai Aiken's V sebesar 0,967 berdasarkan penilaian tiga validator terhadap 20 pernyataan.

Validitas empiris butir dianalisis menggunakan korelasi Pearson *product-moment* antara skor setiap butir dengan skor total instrumen. Karena instrumen berbentuk soal uraian, respons siswa terlebih dahulu dinilai menggunakan rubrik penskoran dengan rentang skor 1–4. Analisis dilakukan terhadap 20 butir hasil uji coba pada 15 siswa. Keputusan validitas butir ditentukan dengan membandingkan nilai r hitung dengan r tabel pada taraf signifikansi 5%. Dengan jumlah responden sebanyak 15 siswa, diperoleh r tabel sebesar 0,514. Butir dinyatakan memenuhi kriteria validitas empiris apabila r hitung lebih besar daripada r tabel (Sudijono, 2012).

Selanjutnya, konsistensi internal instrumen dianalisis menggunakan koefisien Cronbach's alpha (Cronbach, 1951). Analisis konsistensi internal dilakukan terhadap 20 butir soal yang diujicobakan pada 15 siswa, sehingga diperoleh informasi mengenai konsistensi internal instrumen pada tahap uji coba awal. Tiga butir yang tidak memenuhi kriteria validitas empiris kemudian tidak dipertahankan dalam instrumen hasil revisi, sedangkan 17 butir yang memenuhi kriteria dipertahankan. Dengan demikian, penggunaan 20 butir pada analisis konsistensi internal merujuk pada instrumen yang diujicobakan sebelum eliminasi butir, sedangkan 17 butir merupakan hasil seleksi berdasarkan analisis validitas empiris.

Pengolahan data dilakukan dengan bantuan Microsoft Excel. Hasil analisis digunakan untuk memberikan bukti awal mengenai kualitas instrumen pada tahap pengembangan.

Validitas isi ditinjau berdasarkan penilaian ahli, validitas empiris ditinjau berdasarkan korelasi skor butir dengan skor total, sedangkan konsistensi internal ditinjau berdasarkan koefisien Cronbach's alpha pada hasil uji coba terbatas. Interpretasi hasil dibatasi pada konteks uji coba terbatas dan tidak dimaksudkan untuk menggeneralisasikan kualitas instrumen kepada populasi yang lebih luas.

HASIL DAN PEMBAHASAN

Pengembangan Instrumen Berpikir Komputasional

Penelitian ini menghasilkan instrumen tes berpikir komputasional pada materi Sistem Komputer untuk siswa kelas VIII SMP. Pengembangan instrumen dilakukan menggunakan model 4D yang meliputi tahap *define*, *design*, *develop*, dan *disseminate* (Thiagarajan et al., 1974). Penelitian ini dilaksanakan sampai tahap *develop* karena berfokus pada pengembangan produk dan memperoleh bukti awal mengenai kualitas instrumen melalui validasi ahli dan uji coba terbatas. Tahap *define* dilakukan untuk mengidentifikasi kebutuhan pengembangan instrumen. Hasil observasi menunjukkan bahwa pembelajaran Informatika pada materi Sistem Komputer masih didominasi oleh metode ceramah dan penugasan yang berorientasi pada penguasaan konsep. Temuan tersebut diperkuat oleh hasil wawancara dengan guru-guru Informatika yang tergabung dalam Musyawarah Guru Mata Pelajaran (MGMP). Guru menyampaikan bahwa belum tersedia instrumen yang secara khusus dirancang untuk mengukur kemampuan berpikir komputasional pada materi Sistem Komputer.

Hasil analisis kebutuhan tersebut menjadi dasar dalam penyusunan rancangan instrumen pada tahap *design*. Pengembangan instrumen diawali dengan analisis capaian pembelajaran dan tujuan pembelajaran pada materi Sistem Komputer untuk memastikan keterkaitan antara kompetensi yang hendak dicapai dengan konstruk yang diukur. Selanjutnya, disusun kisi-kisi instrumen berdasarkan empat komponen berpikir komputasional, yaitu *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking* (Shute et al., 2017). Keempat komponen tersebut menjadi dasar dalam merumuskan indikator dan penyusunan butir instrumen sehingga konstruk berpikir komputasional dapat dioperasionalkan melalui konteks materi Sistem Komputer. Hal ini sejalan dengan pandangan Wing (2017) bahwa berpikir komputasional merupakan keterampilan berpikir yang diterapkan secara luas untuk memecahkan masalah dan tidak terbatas pada aktivitas pemrograman.

Berdasarkan keempat komponen tersebut, disusun *prototype* instrumen yang terdiri atas 20 soal uraian, dengan lima butir untuk masing-masing komponen berpikir komputasional. Pemilihan bentuk soal uraian didasarkan pada pertimbangan bahwa pengukuran kemampuan berpikir komputasional perlu memberikan kesempatan kepada siswa untuk menunjukkan proses berpikir dalam memahami dan menyelesaikan permasalahan. Penggunaan soal uraian juga memungkinkan respons siswa dianalisis berdasarkan indikator yang telah ditetapkan, sehingga jawaban yang diberikan dapat memberikan gambaran mengenai kemampuan siswa dalam menguraikan masalah, mengenali pola, melakukan abstraksi, dan menyusun langkah penyelesaian secara sistematis. Setiap butir instrumen dilengkapi dengan rubrik penskoran yang digunakan untuk memberikan skor berdasarkan kualitas respons siswa terhadap indikator yang diukur. Dengan demikian, instrumen tidak hanya menilai jawaban akhir, tetapi juga memberikan ruang untuk mengidentifikasi proses berpikir yang ditunjukkan melalui jawaban siswa.

Pada tahap *develop*, *prototype* instrumen selanjutnya melalui proses validasi ahli dan revisi berdasarkan masukan validator. Instrumen juga disajikan dalam bentuk digital menggunakan Canva yang diintegrasikan dengan platform Liveworksheets. Penggunaan teknologi digital dalam penelitian ini berfungsi sebagai sarana penyajian dan pelaksanaan tes, sedangkan konstruk utama yang dikembangkan dan diukur tetap kemampuan berpikir komputasional. Pemanfaatan teknologi digital memberikan kemudahan dalam penyajian instrumen dan akses siswa terhadap tes. Dengan demikian, istilah penyajian digital dalam

penelitian ini tidak menunjukkan bahwa *digital learning* merupakan konstruk atau fokus utama penelitian, melainkan menunjukkan bentuk penyajian instrumen tes yang dikembangkan. Hal tersebut sejalan dengan Oktaviani et al. (2023) yang menjelaskan bahwa pemanfaatan teknologi dalam pembelajaran dapat mendukung proses penyajian dan pelaksanaan kegiatan pembelajaran secara lebih fleksibel.

Penggunaan format digital juga memungkinkan instrumen disajikan secara lebih sistematis dan dapat diakses siswa melalui perangkat yang digunakan dalam pelaksanaan uji coba. Namun, aspek digital dalam penelitian ini tidak ditempatkan sebagai tujuan utama pengembangan, melainkan sebagai pendukung pelaksanaan instrumen tes berpikir komputasional. Produk yang dikembangkan tetap berupa instrumen tes yang terdiri atas butir soal, indikator, dan rubrik penskoran, sedangkan Canva dan Liveworksheets digunakan untuk mendukung penyajian dan pelaksanaan instrumen secara digital. Dengan karakteristik tersebut, pengembangan instrumen dalam penelitian ini menekankan keterkaitan antara konstruk berpikir komputasional, indikator, konteks materi Sistem Komputer, bentuk soal uraian, dan kriteria penskoran.

Hasil tahap *define*, *design*, dan *develop* menunjukkan bahwa instrumen dikembangkan berdasarkan kebutuhan yang ditemukan di lapangan, karakteristik materi Sistem Komputer, serta konstruk berpikir komputasional yang terdiri atas *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Penyusunan kisi-kisi, indikator, butir soal, dan rubrik penskoran menjadi bagian penting dalam mengoperasionalkan konstruk tersebut ke dalam instrumen yang dapat diujikan kepada siswa. Selanjutnya, instrumen yang telah dikembangkan melalui validasi ahli dan uji coba terbatas untuk memperoleh bukti awal mengenai validitas isi, validitas empiris butir, dan konsistensi internal instrumen. Hasil validasi dan uji coba tersebut selanjutnya menjadi dasar dalam menentukan kualitas dan penyempurnaan instrumen pada tahap pengembangan (Thiagarajan et al., 1974).

Validasi Ahli dan Revisi Produk

Setelah *prototype* awal instrumen selesai dikembangkan, tahap berikutnya adalah melakukan validasi ahli. Pada penelitian ini, validasi dilakukan oleh tiga validator yang terdiri atas satu dosen dan dua guru Informatika yang memiliki kompetensi relevan dalam bidang pembelajaran, evaluasi pendidikan, dan media pembelajaran. Validasi dilakukan untuk memperoleh penilaian ahli terhadap kualitas instrumen sebelum digunakan dalam uji coba terbatas. Aspek yang dinilai meliputi materi, media, dan evaluasi (Korkmaz et al., 2017).

Proses validasi menggunakan lembar penilaian berbentuk skala Likert lima tingkat yang terdiri atas 20 pernyataan (Likert, 1932). Setiap validator memberikan penilaian dan saran terhadap instrumen yang dikembangkan. Hasil penilaian menunjukkan bahwa validator pertama memperoleh skor 97 dari skor maksimum 100, validator kedua memperoleh skor 98, dan validator ketiga memperoleh skor 97, sehingga skor keseluruhan yang diperoleh adalah 292 dari skor maksimum 300. Selanjutnya, penilaian ahli dianalisis menggunakan indeks Aiken's V untuk memperoleh bukti validitas isi instrumen (Aiken, 1985).

Tabel 1. Rekapitulasi Penilaian Validator

Validator	Skor Diperoleh	Skor Maksimum
Validator 1	97	100
Validator 2	98	100
Validator 3	97	100
Total	292	300

Berdasarkan hasil analisis menggunakan indeks Aiken's V, diperoleh nilai 0,967. Nilai tersebut merupakan nilai agregat/rata-rata indeks Aiken's V berdasarkan penilaian tiga

validator terhadap 20 pernyataan dengan skala penilaian 1-5. Perhitungan tersebut menggunakan tiga validator, 20 pernyataan, dan skala penilaian 1–5.

Tabel 2. Hasil Analisis Validitas Isi Menggunakan Aiken's V

Jumlah Pernyataan	Jumlah Validator	Skala Penilaian	Aiken's V
20	3	1–5	0,967

Nilai Aiken's V sebesar 0,967 menunjukkan tingkat kesesuaian isi yang tinggi berdasarkan penilaian para validator. Aiken (1985) menjelaskan bahwa indeks Aiken's V digunakan untuk mengetahui tingkat kesepakatan para ahli terhadap relevansi atau kesesuaian suatu item berdasarkan kategori penilaian yang diberikan. Nilai indeks yang semakin mendekati 1 menunjukkan semakin tinggi tingkat kesepakatan ahli terhadap item yang dinilai. Dengan demikian, hasil penelitian memberikan bukti awal mengenai validitas isi instrumen berdasarkan penilaian tiga validator.

Hasil penilaian ahli juga menunjukkan bahwa instrumen telah disusun dengan memperhatikan kesesuaian materi Sistem Komputer dan indikator kemampuan berpikir komputasional. Instrumen dirancang untuk mengukur empat komponen berpikir komputasional, yaitu *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Keempat komponen tersebut menjadi dasar dalam penyusunan indikator dan butir soal. Hal ini sejalan dengan Shute et al. (2017) yang menempatkan *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking* sebagai komponen penting dalam pengukuran kemampuan berpikir komputasional. Pandangan tersebut juga sejalan dengan Wing (2017) yang menegaskan bahwa berpikir komputasional merupakan keterampilan berpikir yang dapat diterapkan untuk menyelesaikan permasalahan secara sistematis dan tidak terbatas pada kegiatan pemrograman.

Pada aspek materi, penilaian ahli menunjukkan bahwa isi instrumen telah disesuaikan dengan materi Sistem Komputer dan kompetensi yang menjadi tujuan pembelajaran. Kesesuaian antara materi dengan konstruk yang diukur merupakan bagian penting dalam pengembangan instrumen karena butir soal harus memberikan representasi yang relevan terhadap kemampuan yang hendak diukur. Temuan ini sejalan dengan Rusilowati et al. (2024) yang menekankan pentingnya kesesuaian antara tujuan pengukuran, indikator, dan isi instrumen dalam pengembangan alat ukur pendidikan.

Pada aspek media, penilaian ahli digunakan untuk menilai penyajian instrumen dalam format digital, termasuk keterbacaan, tata letak, dan kemudahan penggunaan. Canva dan Liveworksheets digunakan sebagai sarana untuk menyajikan dan melaksanakan instrumen secara digital. Namun, aspek digital dalam penelitian ini tidak ditempatkan sebagai konstruk yang diukur, melainkan sebagai pendukung penyajian instrumen. Dengan demikian, penilaian media berfungsi untuk memastikan bahwa penyajian instrumen tidak menghambat siswa dalam memahami stimulus dan memberikan respons terhadap soal.

Pada aspek evaluasi, penilaian ahli diarahkan pada kesesuaian butir dengan indikator yang hendak diukur serta kejelasan penyusunan instrumen. Setiap komponen berpikir komputasional direpresentasikan melalui butir soal yang telah dirancang berdasarkan kisi-kisi dan indikator. Penggunaan soal uraian juga memberikan kesempatan kepada siswa untuk menunjukkan respons yang menggambarkan proses berpikirnya. Dengan demikian, hasil validasi ahli pada aspek evaluasi menjadi dasar untuk memastikan bahwa butir yang dikembangkan memiliki keterkaitan dengan tujuan pengukuran.

Meskipun memperoleh nilai Aiken's V yang tinggi, hasil validasi ahli tidak digunakan sebagai satu-satunya dasar untuk menyatakan bahwa seluruh butir instrumen telah memiliki kualitas empiris yang baik. Validitas isi berdasarkan penilaian ahli memberikan informasi mengenai kesesuaian isi instrumen berdasarkan pertimbangan pakar, sedangkan kualitas empiris setiap butir perlu diperiksa melalui respons siswa pada tahap uji coba terbatas. Pemisahan antara bukti validitas isi dan validitas empiris penting dilakukan karena keduanya memberikan informasi yang berbeda mengenai kualitas instrumen.

Selain memberikan penilaian kuantitatif, para validator memberikan beberapa masukan untuk menyempurnakan *prototype* instrumen. Masukan tersebut meliputi perbaikan tata tulis, konsistensi penggunaan huruf, penyempurnaan tata letak, penyesuaian spasi, serta penyederhanaan beberapa kalimat agar lebih mudah dipahami oleh siswa. Masukan tersebut digunakan sebagai dasar dalam melakukan revisi sebelum instrumen digunakan pada tahap uji coba terbatas.

Revisi dilakukan dengan mempertahankan substansi materi dan indikator berpikir komputasional yang telah dirancang, sedangkan perbaikan difokuskan pada kejelasan bahasa, keterbacaan, tata letak, dan penyajian instrumen. Proses revisi tersebut merupakan bagian dari tahap *develop* untuk menyempurnakan instrumen berdasarkan masukan ahli. Setelah proses revisi selesai, diperoleh *prototype* II instrumen tes berpikir komputasional yang siap digunakan pada uji coba terbatas.

Secara keseluruhan, hasil validasi ahli dengan nilai Aiken's V sebesar 0,967 memberikan bukti awal bahwa isi instrumen memiliki tingkat kesesuaian yang tinggi berdasarkan penilaian tiga validator. Nilai tersebut merepresentasikan rekapitulasi penilaian terhadap 20 pernyataan dan tidak digunakan untuk menyimpulkan validitas empiris setiap butir. Bukti mengenai validitas empiris butir dan konsistensi internal instrumen selanjutnya diperoleh melalui uji coba terbatas pada 15 siswa.

Uji Coba Terbatas

Setelah melalui proses validasi dan revisi, instrumen tes berpikir komputasional diujicobakan secara terbatas kepada 15 siswa kelas VIII SMP Negeri 9 Langsa. Uji coba dilakukan untuk memperoleh bukti empiris awal mengenai kualitas instrumen pada tahap uji coba terbatas. Instrumen yang diujicobakan terdiri atas 20 soal uraian yang dikembangkan berdasarkan empat komponen berpikir komputasional, yaitu *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Setiap respons siswa dinilai menggunakan rubrik penskoran yang telah disusun pada tahap pengembangan instrumen.

Pelaksanaan uji coba dilakukan dalam lingkungan pembelajaran yang telah disiapkan oleh guru dan peneliti. Siswa mengakses instrumen melalui perangkat digital dan mengerjakan seluruh soal yang tersedia pada platform Liveworksheets. Sebelum pengerjaan dimulai, siswa diberikan penjelasan mengenai tujuan kegiatan dan tata cara penggunaan instrumen. Penggunaan platform digital dalam tahap ini berfungsi sebagai sarana penyajian dan pelaksanaan tes, sedangkan data yang dianalisis berupa skor respons siswa pada setiap butir instrumen. Dengan demikian, penggunaan platform digital pada uji coba ini tidak dimaksudkan untuk mengukur efektivitas media, tetapi untuk mendukung pelaksanaan instrumen yang dikembangkan.



Gambar 1. Pelaksanaan Uji Coba Terbatas

Selama pelaksanaan uji coba, siswa mengerjakan soal yang menyajikan permasalahan berkaitan dengan materi Sistem Komputer. Bentuk soal uraian memberikan kesempatan kepada siswa untuk menunjukkan proses penyelesaian dan alasan yang mendasari jawaban. Hal tersebut relevan dengan karakteristik berpikir komputasional yang tidak hanya berkaitan dengan hasil akhir, tetapi juga proses perumusan masalah dan penyusunan solusi secara sistematis (Wing, 2017). Shute et al. (2017) juga menjelaskan bahwa *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking* merupakan komponen yang dapat digunakan untuk mengidentifikasi kemampuan berpikir komputasional melalui proses penyelesaian masalah. Oleh karena itu, respons siswa terhadap soal uraian digunakan sebagai dasar untuk memperoleh skor pada masing-masing butir instrumen.

Skor yang diperoleh siswa selanjutnya digunakan untuk menganalisis kualitas empiris instrumen. Analisis dilakukan terhadap 20 butir soal yang diujicobakan untuk memperoleh bukti awal mengenai validitas empiris setiap butir. Analisis dilakukan terhadap setiap butir untuk mengetahui sejauh mana butir mampu menjalankan fungsi pengukurannya sesuai dengan konstruk yang dikembangkan. Hal ini sejalan dengan Azwar (2021) yang menjelaskan bahwa analisis butir diperlukan untuk memperoleh informasi mengenai kualitas setiap butir sebagai bagian dari instrumen pengukuran. Selain validitas empiris butir, skor seluruh siswa juga digunakan untuk menganalisis konsistensi internal instrumen.

Uji coba terbatas dalam penelitian ini tidak dimaksudkan untuk menguji efektivitas penggunaan media digital, peningkatan kemampuan siswa, maupun menggeneralisasikan hasil kepada populasi yang lebih luas. Tahap ini difokuskan pada memperoleh bukti empiris awal mengenai kualitas instrumen tes berpikir komputasional setelah melalui proses validasi ahli dan revisi. Dengan demikian, hasil uji coba selanjutnya digunakan untuk menentukan validitas empiris masing-masing butir dan konsistensi internal instrumen. Hasil analisis tersebut digunakan sebagai dasar untuk menyeleksi butir yang dipertahankan dalam instrumen hasil revisi serta memberikan informasi mengenai konsistensi internal instrumen pada tahap uji coba terbatas.

Validitas Butir Soal

Data hasil uji coba terbatas dianalisis untuk memperoleh bukti empiris awal mengenai kualitas setiap butir instrumen tes berpikir komputasional. Analisis validitas dilakukan menggunakan korelasi Pearson *product-moment* antara skor setiap butir dengan skor total instrumen. Karena berbentuk soal uraian, setiap respons siswa terlebih dahulu dinilai menggunakan rubrik penskoran. Dengan jumlah responden sebanyak 15 siswa, diperoleh r tabel sebesar 0,514 pada taraf signifikansi 5%. Butir dinyatakan valid apabila r hitung lebih besar daripada r tabel (Sudijono, 2012). Azwar (2021) menjelaskan bahwa analisis validitas butir digunakan untuk mengetahui sejauh mana setiap butir mampu mengukur atribut yang menjadi bagian dari konstruk yang diteliti. Dalam penelitian ini, dihasilkan analisis validitas butir ditafsirkan sebagai bukti empiris awal pada konteks uji coba terbatas.

Tabel 3. Hasil Uji Validitas Butir Instrumen Tes Berpikir Komputasional

No. Butir	r hitung	r tabel	Keterangan
1	0,837	0,514	Valid
2	0,708	0,514	Valid
3	0,760	0,514	Valid
4	0,566	0,514	Valid
5	-0,115	0,514	Tidak valid
6	0,685	0,514	Valid
7	0,685	0,514	Valid
8	0,648	0,514	Valid
9	0,612	0,514	Valid
10	0,786	0,514	Valid
11	0,714	0,514	Valid

No. Butir	<i>r</i> hitung	<i>r</i> tabel	Keterangan
12	0,660	0,514	Valid
13	0,099	0,514	Tidak valid
14	0,638	0,514	Valid
15	0,773	0,514	Valid
16	0,595	0,514	Valid
17	0,682	0,514	Valid
18	0,774	0,514	Valid
19	0,117	0,514	Tidak valid
20	0,748	0,514	Valid

Berdasarkan Tabel 3, sebanyak 17 dari 20 butir (85%) memenuhi kriteria validitas empiris awal, sedangkan tiga butir, yaitu nomor 5, 13, dan 19, tidak memenuhi kriteria yang ditetapkan. Butir yang valid memenuhi kriteria memiliki *r* hitung lebih besar dari 0,514. Sebaliknya, butir 5, 13, dan 19 memiliki nilai *r* hitung masing-masing sebesar -0,115, 0,099, dan 0,117 sehingga tidak memenuhi kriteria validitas empiris. Hasil ini memberikan bukti awal bahwa sebagian besar butir memiliki hubungan yang memadai dengan skor total instrumen pada uji coba terbatas.

Secara konseptual, butir dikembangkan berdasarkan *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Shute et al. (2017) menjelaskan bahwa keempat komponen tersebut dapat digunakan untuk merepresentasikan kemampuan berpikir komputasional dalam pemecahan masalah, sedangkan Wing (2017) menekankan bahwa berpikir komputasional berkaitan dengan proses perumusan masalah dan penyusunan solusi secara sistematis. Landasan tersebut menjadi dasar penyusunan butir, sedangkan analisis empiris digunakan untuk memperoleh informasi mengenai kinerja masing-masing butir setelah instrumen diujicobakan kepada siswa.

Tiga butir yang tidak memenuhi kriteria validitas tidak dipertahankan dalam instrumen hasil revisi. Ketidakvalidan butir menjadi dasar untuk meninjau kembali kesesuaian redaksi, stimulus, dan indikator yang diukur (Azwar, 2021). Dengan demikian, dari 20 butir yang diujicobakan, 17 butir dipertahankan untuk analisis selanjutnya, sedangkan tiga butir yang tidak memenuhi kriteria dikeluarkan dari instrumen hasil revisi.

Meskipun demikian, hasil validitas ini perlu ditafsirkan sebagai bukti empiris awal karena uji coba hanya melibatkan 15 siswa. Ukuran sampel yang terbatas menyebabkan hasil analisis belum dapat digunakan untuk menggeneralisasikan kualitas butir pada populasi yang lebih luas. Hasil tersebut juga belum cukup untuk menyatakan validitas konstruk instrumen secara keseluruhan, karena struktur empat komponen berpikir komputasional belum diuji melalui analisis struktur konstruk. Oleh karena itu, 17 butir yang dipertahankan dipandang sebagai hasil seleksi awal berdasarkan kinerja empiris pada uji coba terbatas dan selanjutnya digunakan dalam analisis konsistensi internal instrumen.

Reliabilitas Instrumen

Setelah analisis validitas butir, reliabilitas instrumen dianalisis menggunakan koefisien *Cronbach's alpha* untuk mengetahui konsistensi internal instrumen. Analisis reliabilitas dilakukan terhadap 20 butir soal yang diujicobakan kepada 15 siswa kelas VIII SMP Negeri 9 Langsa, yaitu sebelum tiga butir yang tidak memenuhi kriteria validitas empiris dieliminasi. Cronbach (1951) menjelaskan bahwa *Cronbach's alpha* digunakan untuk menilai konsistensi internal antarbutir dalam suatu instrumen.

Tabel 4. Hasil Uji Reliabilitas Instrumen

Jumlah Butir	<i>Cronbach's alpha</i>	Kategori
20	0,918	Sangat tinggi

Hasil analisis menunjukkan nilai *Cronbach's alpha* sebesar 0,918 dan termasuk kategori sangat tinggi. Nilai tersebut menunjukkan bahwa 20 butir soal yang diujicobakan

memiliki konsistensi internal yang tinggi pada uji coba terbatas. El-Hamamsy et al. (2025) menjelaskan bahwa *Cronbach's alpha* digunakan untuk menilai sejauh mana butir-butir dalam instrumen memiliki keterkaitan dalam mengukur konstruk yang sama.

Namun, nilai *Cronbach's alpha* tidak secara langsung membuktikan validitas konstruk maupun kestabilan hasil pengukuran dari waktu ke waktu. Rizki Hikmawan et al. (2025) menegaskan bahwa *Cronbach's alpha* merupakan ukuran konsistensi internal yang perlu ditafsirkan bersama bukti validitas lainnya. Oleh karena itu, nilai 0,918 dalam penelitian ini diposisikan sebagai bukti awal konsistensi internal instrumen pada tahap uji coba terbatas, bukan sebagai bukti bahwa instrumen telah memiliki reliabilitas yang memadai untuk penggunaan secara luas.

Hasil tersebut juga perlu dipertimbangkan berdasarkan jumlah responden yang hanya 15 siswa. Selain itu, nilai *Cronbach's alpha* yang dilaporkan dalam penelitian ini diperoleh dari 20 butir instrumen sebelum proses eliminasi tiga butir yang tidak memenuhi kriteria validitas empiris. Dengan demikian, nilai 0,918 tidak dimaksudkan untuk merepresentasikan konsistensi internal dari 17 butir yang dipertahankan dalam instrumen hasil revisi. Karena instrumen berbentuk soal uraian, *Cronbach's alpha* juga belum memberikan informasi mengenai konsistensi antarpemilai dalam pemberian skor. Reliabilitas penskoran (*inter-rater reliability*) perlu diuji pada penelitian selanjutnya untuk memperkuat bukti reliabilitas instrumen.

Dengan demikian, nilai *Cronbach's alpha* sebesar 0,918 menunjukkan konsistensi internal yang tinggi pada 20 butir instrumen dalam uji coba terbatas. Hasil ini melengkapi bukti validitas isi dan validitas empiris butir, tetapi perlu ditafsirkan secara terbatas sesuai dengan karakteristik uji coba yang melibatkan 15 siswa. Pengujian dengan sampel yang lebih besar dan analisis reliabilitas penskoran diperlukan untuk memperoleh bukti psikometrik yang lebih kuat.

Pembahasan Kebaruan Temuan dan Implikasi Penelitian

Hasil penelitian menunjukkan bahwa pengembangan instrumen tes berpikir komputasional pada materi Sistem Komputer menghasilkan bukti awal mengenai validitas isi, validitas empiris butir, dan konsistensi internal pada uji coba terbatas. Validitas isi berdasarkan penilaian tiga validator memperoleh nilai Aiken's V sebesar 0,967, sedangkan uji coba terbatas menunjukkan 17 dari 20 butir memenuhi kriteria validitas empiris. Konsistensi internal 20 butir yang diujicobakan sebelum eliminasi berdasarkan *Cronbach's alpha* menunjukkan nilai 0,918. Temuan tersebut memberikan bukti awal yang mendukung kualitas instrumen dalam pengembangan asesmen berpikir komputasional pada pembelajaran Informatika jenjang SMP.

Kebaruan penelitian ini terletak pada penggunaan materi Sistem Komputer sebagai konteks pengukuran berpikir komputasional dalam lingkungan non-pemrograman. Instrumen dikembangkan berdasarkan empat komponen, yaitu *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Shute et al. (2017) menjelaskan bahwa berpikir komputasional dapat direpresentasikan melalui berbagai proses pemecahan masalah, sedangkan Wing (2017) menegaskan bahwa berpikir komputasional merupakan keterampilan yang tidak terbatas pada aktivitas pemrograman. Dengan demikian, penelitian ini memberikan alternatif konteks pengukuran berpikir komputasional melalui materi Sistem Komputer yang merupakan bagian dari pembelajaran Informatika, khususnya melalui soal uraian yang memungkinkan respons siswa diamati berdasarkan indikator yang telah ditetapkan.

Kebaruan lainnya terdapat pada penyajian instrumen dalam format digital menggunakan Canva dan Liveworksheets. Dalam penelitian ini, teknologi digital digunakan sebagai sarana penyajian dan pelaksanaan tes, sedangkan konstruk yang diukur tetap kemampuan berpikir komputasional. Dengan demikian, format digital tidak diposisikan sebagai konstruk utama maupun sebagai variabel yang diuji efektivitasnya, tetapi sebagai sarana pendukung pelaksanaan asesmen. Penggunaan format digital memberikan

kemudahan dalam akses dan administrasi tes, tetapi penelitian ini tidak dirancang untuk menguji efektivitas penggunaan platform digital terhadap hasil belajar siswa.

Implikasi penelitian ini bersifat awal dan terbatas pada penggunaannya sebagai salah satu alternatif asesmen dalam pembelajaran Informatika. Instrumen dapat membantu guru memperoleh informasi awal mengenai respons siswa terhadap permasalahan yang berkaitan dengan *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Informasi tersebut dapat menjadi bahan pertimbangan dalam merancang kegiatan pembelajaran yang memberikan kesempatan kepada siswa untuk menganalisis dan menyelesaikan masalah secara sistematis. Namun, penggunaan instrumen sebagai asesmen perlu mempertimbangkan bahwa bukti empiris yang diperoleh masih berasal dari uji coba terbatas.

Meskipun demikian, hasil penelitian belum dapat menjadi dasar untuk menyatakan bahwa instrumen merupakan alat ukur standar yang dapat digunakan secara luas. Uji coba hanya melibatkan 15 siswa dari satu sekolah, struktur empat komponen belum diuji, reliabilitas penskoran soal uraian belum diketahui, dan belum tersedia norma atau pedoman interpretasi skor. Selain itu, nilai *Cronbach's alpha* sebesar 0,918 diperoleh dari 20 butir sebelum tiga butir yang tidak memenuhi kriteria validitas empiris dieliminasi, sehingga nilai tersebut tidak secara langsung merepresentasikan konsistensi internal 17 butir yang dipertahankan. Oleh karena itu, penelitian selanjutnya perlu melibatkan sampel yang lebih besar dan beragam, menguji struktur konstruk, menganalisis karakteristik butir secara lebih mendalam, serta menguji reliabilitas penskoran (*inter-rater reliability*). Langkah tersebut diperlukan untuk memperoleh bukti psikometrik yang lebih kuat sebelum instrumen digunakan pada konteks yang lebih luas.

PENUTUP

Penelitian ini menghasilkan instrumen tes berpikir komputasional pada materi Sistem Komputer untuk siswa SMP yang dikembangkan melalui tahap *define*, *design*, dan *develop* dalam model 4D. Instrumen disusun berdasarkan empat komponen berpikir komputasional, yaitu *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking*. Hasil validasi oleh tiga ahli memperoleh nilai Aiken's V sebesar 0,967, yang memberikan bukti awal validitas isi yang tinggi berdasarkan penilaian ahli. Uji coba terbatas terhadap 15 siswa menunjukkan bahwa 17 dari 20 butir menunjukkan bukti validitas empiris awal, sedangkan tiga butir tidak memenuhi kriteria dan tidak dipertahankan dalam instrumen hasil revisi. Sementara itu, nilai *Cronbach's alpha* sebesar 0,918 memberikan bukti awal mengenai konsistensi internal instrumen pada uji coba terbatas.

Kebaruhan penelitian ini terletak pada pengembangan instrumen tes berpikir komputasional dengan materi Sistem Komputer sebagai konteks non-pemrograman. Penggunaan konteks tersebut memberikan alternatif dalam pengukuran *decomposition*, *pattern recognition*, *abstraction*, dan *algorithmic thinking* tanpa bergantung pada aktivitas pemrograman. Penyajian instrumen secara digital melalui Canva dan Liveworksheets digunakan sebagai sarana pelaksanaan tes, sedangkan konstruk yang dikembangkan dan diukur tetap kemampuan berpikir komputasional.

Temuan penelitian ini memiliki implikasi awal bagi asesmen Informatika, khususnya sebagai alternatif asesmen formatif dalam konteks pembelajaran Sistem Komputer untuk memperoleh informasi mengenai kemampuan berpikir komputasional siswa. Namun, hasil penelitian ini merupakan bukti awal dari uji coba terbatas dan belum dapat dijadikan dasar untuk menetapkan instrumen sebagai alat ukur standar yang dapat digunakan secara luas. Keterbatasan penelitian meliputi jumlah responden yang terbatas dan berasal dari satu sekolah, belum dilakukannya pengujian struktur konstruk, serta belum tersedianya bukti *inter-rater reliability* untuk penskoran soal uraian.

Penelitian selanjutnya perlu melibatkan sampel yang lebih besar dan beragam, menguji struktur konstruk dan karakteristik butir secara lebih mendalam, serta menganalisis reliabilitas penskoran. Pengujian lanjutan tersebut diperlukan untuk memperkuat bukti

psikometrik instrumen sebelum penggunaannya dipertimbangkan pada konteks pembelajaran yang lebih luas.

DAFTAR RUJUKAN

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Azwar, S. (2021). *Penyusunan skala psikologi* (Edisi 3). Pustaka Pelajar.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 17–66). Springer. https://doi.org/10.1007/978-94-007-2324-5_2
- Buwono, F. A., Djono, Bashir, N. L., & Putri, C. N. (2025). *Computational thinking-based metacognitive learning assessment: Issues and challenges*. *Indonesian Journal of Learning and Instructional Innovation*, 3(2), 102–118. <https://doi.org/10.20961/ijolii.v3i2.2991>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- Dilekçi, A., & Karatay, H. (2023). The effects of the 21st century skills curriculum on the development of students' creative thinking skills. *Thinking Skills and Creativity*, 47, 101229. <https://doi.org/10.1016/j.tsc.2022.101229>
- El-Hamamsy, L., Zapata-Cáceres, M., Martín-Barroso, E., Mondada, F., Zufferey, J. D., Bruno, B., & Román-González, M. (2025). The competent computational thinking test (cCTt): A valid, reliable and gender-fair test for longitudinal CT studies in grades 3–6. *Technology, Knowledge and Learning*, 30(3), 1607–1661. <https://doi.org/10.1007/s10758-024-09777-8>
- Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2021). *Panduan implementasi pembelajaran Kurikulum Merdeka*.
- Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2022). *Mata pelajaran Informatika*.
- Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in Human Behavior*, 72, 558–569. <https://doi.org/10.1016/j.chb.2017.01.005>
- Likert, R. (1932). *A technique for the measurement of attitudes*. *Archives of Psychology*, 22(140), 1–55.
- Nugroho, A. E. J. K., Sugiyanto, S., & Prihadi, S. (2025). The impacts of implementing Merdeka Belajar curriculum on the quality of senior high school education in Indonesia: A scoping review. *Prima Magistra: Jurnal Ilmiah Kependidikan*, 6(4), 525–542. <https://doi.org/10.37478/jpm.v6i4.5652>
- Oktaviani, P. S., Hadiana, O., Heriyana, T., Nurhayati, T., & Cirebon, U. M. (2023). Efektivitas penggunaan media wordwall dan mistar pinter terhadap hasil belajar operasi bilangan bulat pada siswa kelas rendah. *Jurnal Edukasi dan Sains Matematika*, 9(2), 125–136. <https://doi.org/10.25134/jes-mat.v9i2.7814>
- Rachmawati, E., Yulianto, F. A., & Ningsih, L. (2024). Pendampingan implementasi materi berpikir komputasional di mata pelajaran Informatika pada Kurikulum Merdeka tingkat SMP di wilayah Kabupaten Bandung. *I-Com: Indonesian Community Journal*, 4(3), 1723–1735. <https://doi.org/10.33379/icom.v4i3.4899>

- Rizki Hikmawan, Mumu Komaro, Ayi Suherman, & Ayu Permata Sari. (2025). Validity and reliability of computational thinking scales for undergraduate students: Indonesian adaptation. *Didaktika: Jurnal Kependidikan*, 14(3), 5535–5546. <https://doi.org/10.58230/27454312.2910>
- Rusilowati, A., Darsono, T., Sekarningtias, F. O., & Fariyani, Q. (2024). Development of a four-tier e-diagnostic test on the topic of momentum to measure and reduce student misconception. *Jurnal Pendidikan Fisika Indonesia*, 20(2), 173–187. <https://doi.org/10.15294/jpfi.v20i2.1297>
- Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, 22, 142–158. <https://doi.org/10.1016/j.edurev.2017.09.003>
- Sudijono, A. (2012). *Pengantar statistik pendidikan* (Cet 24). Rajagrafindo Persada.
- Tanenbaum, A. S., & Austin, T. (2013). *Structured computer organization* (6th ed.). Pearson.
- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Leadership Training Institute/Special Education, University of Minnesota.
- UNESCO. (2023). *Global education monitoring report 2023: Technology in education: A tool on whose terms?* <https://doi.org/10.54676/UZQV8501>
- UNESCO. (2025). *Digital learning and transformation of education: Preparing learners for a changing world*. <https://www.unesco.org/en/articles/digital-learning-and-transformation-education>
- Wing, J. M. (2017). Computational thinking's influence on research and education for all. *Italian Journal of Educational Technology*, 25(2), 7–14. <https://doi.org/10.17471/2499-4324/922>
- Zainuddin, A., Wasis, Ekohariadi, Anistiyarsari, Y., Rijanto, T., & Habibi. (2026). A systematic review of digital assessment trends in education from 2015-2025. *Multidisciplinary Reviews*, 9(9), 2026433. <https://doi.org/10.31893/multirev.2026433>