

Avaliação do Pensamento Computacional no Ensino Superior em Computação: Uma Revisão Sistemática

Marco Aurélio F. de Souza¹, José Marco M. Nascimento¹, Mikael dos S. Gois¹,
Rone Clay O. Andrade¹, Matheus C. Dos Santos¹, Rafael Jesus¹, Marlos Tacio Silva¹

¹Curso de Ciência da Computação – Instituto Federal de Sergipe (IFS)
Av. Padre Airton Gonçalves de Lima, 1140 – 49500-543 – Itabaiana – SE – Brasil

wise93br@gmail.com,

{jose.nascimento089, mikael.gois077}@academico.ifs.edu.br,

{ronecl1r550, mathxsant17, rafadejesus125}@gmail.com,

marlos.silva@ifs.edu.br

Abstract. Introduction: Computational Thinking (CT) is an essential competency for Computer Science education, although many students enter higher education with deficiencies in this skill. **Objective:** To identify and analyze how CT has been assessed among undergraduate Computer Science students. **Methodology:** A systematic literature review was conducted following the PICO strategy, PRISMA guidelines, and the CASP tool, resulting in the analysis of 22 studies. **Results:** The findings show a predominance of instrument validation studies and a lack of longitudinal research, suggesting that traditional programming-focused curricula may not explicitly foster CT development.

Keywords Computational Thinking, Systematic review, Assessment, Higher Education, Computer Science.

Resumo. Introdução: O Pensamento Computacional (PC) é uma competência essencial para a formação em Computação, embora muitos estudantes ingressem na graduação com deficiências nessa habilidade. **Objetivo:** Identificar e analisar como o PC tem sido avaliado em estudantes de graduação em Computação. **Metodologia:** Foi realizada uma revisão sistemática da literatura seguindo a estratégia PICO, as diretrizes PRISMA e a ferramenta CASP, resultando na análise de 22 estudos. **Resultados:** Os estudos concentram-se na validação de instrumentos de avaliação, com escassez de pesquisas longitudinais, indicando que o ensino tradicional, centrado na sintaxe de programação, pode não promover explicitamente o desenvolvimento do PC. **Palavras-Chave** Pensamento Computacional, Revisão sistemática, Avaliação, Ensino Superior, Ciência da Computação.

1. Introdução

O Pensamento Computacional (PC) é um processo cognitivo fundamental para a formulação e resolução de problemas, permitindo a aplicação de soluções computacionais a desafios complexos do mundo real [Wing 2006]. Essa habilidade transcende a área da computação e constitui um elemento vital para o desenvolvimento de competências em

Ciência, Tecnologia, Engenharia e Matemática (STEM) ao longo da vida acadêmica e profissional. A inserção do PC no currículo desde o ensino básico é defendida na literatura devido à sua relevância na construção do raciocínio lógico [Raabe et al. 2017].

A urgência dessa implementação no cenário educacional brasileiro é evidenciada pelos resultados obtidos no Programa Internacional de Avaliação de Estudantes (PISA). A avaliação de 2022, da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), revelou que o desempenho dos estudantes nacionais situou-se abaixo da média dos países-membros, com pontuações de 379 em matemática, 403 em ciências e 410 em leitura [Organização para a Cooperação e Desenvolvimento Econômico (OCDE) 2023]. Tais indicadores sugerem uma defasagem em habilidades analíticas, de abstração e de resolução de problemas, que são pilares centrais do PC.

As lacunas na formação básica geram consequências diretas no ensino superior. Para a maioria dos estudantes brasileiros, o primeiro contato formal com os fundamentos da programação ocorre na universidade, principalmente em cursos de Ciências Exatas e Tecnológicas. Esse cenário exige a assimilação rápida de conceitos lógicos complexos em um único período letivo, sob risco de retenção escolar. O desafio intensifica-se em cursos de Computação, como Ciência da Computação, Sistemas de Informação e Engenharia de *Software*. Nesses ambientes, demanda-se a transição rápida de uma lógica procedural rudimentar para paradigmas abstratos, como a Programação Orientada a Objetos, o que ocasiona altas taxas de reprovação e evasão.

A carência no estímulo ao desenvolvimento do PC na base, somada às dificuldades na graduação e ao risco de formar profissionais que replicam códigos sem capacidade de abstração, justifica a urgência de investigações sobre a avaliação dessa competência. Torna-se necessário investigar como desenvolver e, principalmente, como avaliar as habilidades de PC em estudantes de cursos superiores.

Este estudo apresenta uma revisão sistemática da literatura para mapear informações estruturadas sobre instrumentos de avaliação do PC na graduação em Computação. A análise investiga as características, as propriedades de validade psicométrica e os contextos de aplicação dessas ferramentas. O intuito é gerar subsídios baseados em evidências científicas para reformular práticas pedagógicas, auxiliando na formação de profissionais preparados para os desafios tecnológicos contemporâneos.

2. Fundamentação Teórica

O conceito de Pensamento Computacional (PC), em sua acepção moderna, foi reintroduzido e popularizado por Wing (2006). Essa abordagem fundamenta-se na resolução de problemas, na compreensão do comportamento humano e no *design* de sistemas apoiados em conceitos basilares da Ciência da Computação. Em sua essência, o PC utiliza a decomposição e a abstração para desvendar desafios em grandes projetos ou tarefas de alta complexidade.

A maturação dessa área de pesquisa impulsionou a evolução do termo de uma definição filosófica para um indicador empírico, utilizado na mensuração de capacidades cognitivas individuais. Essa avaliação visa identificar não apenas a habilidade mecânica de escrita de código-fonte, mas a aptidão latente em STEM.

Os construtos adotados para a mensuração do PC variam conforme o instrumento

e a vertente teórica adotada. A partir do modelo consolidado por Tsai et al. (2021), o PC pode ser operacionalizado e decomposto em cinco dimensões específicas e observáveis:

- **Abstração:** Refere-se à habilidade de processar fluxos de informações, descartando detalhes irrelevantes para isolar o essencial requerido na modelagem da solução.
- **Decomposição:** Consiste na capacidade analítica de fragmentar um problema complexo, monolítico e de difícil resolução imediata, em partes menores, manejáveis e independentes.
- **Pensamento Algorítmico:** Define-se como a habilidade de coordenar a resolução de tarefas em etapas sequenciais lógicas, prevendo fluxos de controle e condições de parada.
- **Avaliação:** Objetiva identificar a intenção crítica do indivíduo ao testar iterativamente e encontrar a solução mais eficiente e otimizada para o problema, considerando restrições de tempo, espaço e contexto situacional.
- **Generalização:** Diz respeito à capacidade de reconhecer padrões subjacentes na resolução de problemas passados e adaptar essas soluções (transferência de aprendizado) para situações novas e correntes.

3. Método

A condução desta revisão sistemática da literatura adotou as diretrizes consolidadas por Kitchenham e Charters (2007). Esse referencial, originário da medicina e adaptado para a Engenharia de Software, visa assegurar rigor científico e transparência ao processo. O fluxo de trabalho englobou o planejamento do protocolo, a execução das buscas, a filtragem dos estudos e a avaliação crítica de qualidade.

A estruturação inicial para a formulação da questão de pesquisa e a extração dos descritores baseou-se na estratégia PICO [Higgins e Thomas 2019]. A modelagem para o escopo educacional em Computação foi delimitada sob as seguintes especificações:

- **(P) População:** Estudantes regularmente matriculados em cursos de graduação na área da Computação (Ciência da Computação, Engenharia de Computação, Sistemas de Informação).
- **(I) Intervenção:** Ferramentas empíricas, testes padronizados, questionários de autopercepção, métodos, rubricas ou quaisquer instrumentos utilizados ativamente para avaliar o PC.
- **(C) Comparação:** Não se aplica de forma restritiva ao escopo desta RSL.
- **(O) Resultado (*Outcome*):** Identificação dos instrumentos adotados, descrição de suas características tipológicas, levantamento de propriedades psicométricas (validade e confiabilidade estatística) e compreensão aprofundada dos contextos sociodemográficos de aplicação.

A Questão de Pesquisa (QP) diretriz foi formulada a partir dessa modelagem:

QP: *Quais são as características, propriedades de validade e contextos de aplicação dos instrumentos utilizados para avaliar o Pensamento Computacional em estudantes do ensino superior na área da computação?*

As buscas primárias foram executadas em cinco bases de dados digitais relevantes para a interseção entre Computação e Educação: ACM Digital Library, IEEE Xplore Digital Library, Periódicos Capes, SBC Open Lib (SOL) e ScienceDirect. A *string* de busca genérica, elaborada e posteriormente adaptada para a sintaxe booleana de cada repositório, consistiu em:

```
("Higher Education" OR "College students") AND  
("Assessment" OR "Evaluation" OR "Instrument" OR  
"Questionnaire" OR "Test") AND  
("Computational Thinking")
```

A janela temporal estabelecida como critério de inclusão abrangeu publicações dos últimos dez anos (janeiro de 2015 a junho de 2025). Os critérios de seleção exigiram artigos empíricos, redigidos em português ou inglês, que descrevessem, criassem ou aplicassem explicitamente uma avaliação de PC em estudantes de graduação em Computação. Foram excluídos estudos de revisão bibliográfica sem contribuição primária, livros, capítulos de livros, pesquisas focadas nos níveis educacionais fundamental ou médio, bem como artigos aplicados a discentes externos ao eixo da Ciência da Computação.

Todo o processo de seleção e filtragem foi documentado visualmente seguindo o fluxograma da declaração PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) [Moher et al. 2009]. A triagem por títulos e resumos, bem como a leitura integral, foram conduzidas por revisão em pares para mitigar vieses de seleção e confirmação.

O diferencial metodológico desta revisão consistiu na submissão de todos os artigos incluídos à avaliação de rigor técnico do *Critical Appraisal Skills Programme* (CASP). Para essa finalidade, utilizou-se o *Checklist* para Ensaaios Clínicos Randomizados [Critical Appraisal Skills Programme 2024]. O instrumento possui 11 questões de caráter metodológico, gerando uma pontuação de 0 a 11. Os itens abrangeram a clareza dos objetivos, a adequação metodológica, a estratégia de recrutamento e a confiabilidade da coleta de dados. Adicionalmente, avaliou-se o rigor na análise de dados, o controle de vieses ou variáveis de confusão, a clareza e precisão dos resultados, a coerência nas conclusões, a discussão de limitações e a aplicabilidade das implicações teóricas. A extração final categorizou 17 variáveis bibliométricas e técnicas para cada estudo selecionado.

Os 22 artigos foram categorizados em quatro eixos temáticos através de análise indutiva realizada independentemente por dois pesquisadores. Cada artigo foi lido integralmente, e sua contribuição principal foi classificada em uma das seguintes categorias: (1) Validação e Adaptação de Instrumentos Psicométricos, referindo-se a estudos que desenvolvem, validam ou adaptam escalas e testes; (2) Avaliação de Intervenções Pedagógicas, compreendendo estudos que aplicam instrumentos para avaliar eficácia de metodologias de ensino; (3) Estudos Comparativos e Demográficos, incluindo pesquisas que comparam PC entre grupos ou exploram variáveis demográficas; e (4) Novos Métodos de Avaliação, abrangendo inovações metodológicas como análise automática de código ou rubricas qualitativas.

4. Resultados

O rastreamento inicial nas bases de dados identificou 651 artigos. A aplicação progressiva dos critérios de inclusão e exclusão reduziu esse montante a 22 estudos elegíveis para a síntese final. O detalhamento do processo de seleção, contemplando as etapas de identificação, triagem e inclusão, está ilustrado no fluxograma da Figura 1.

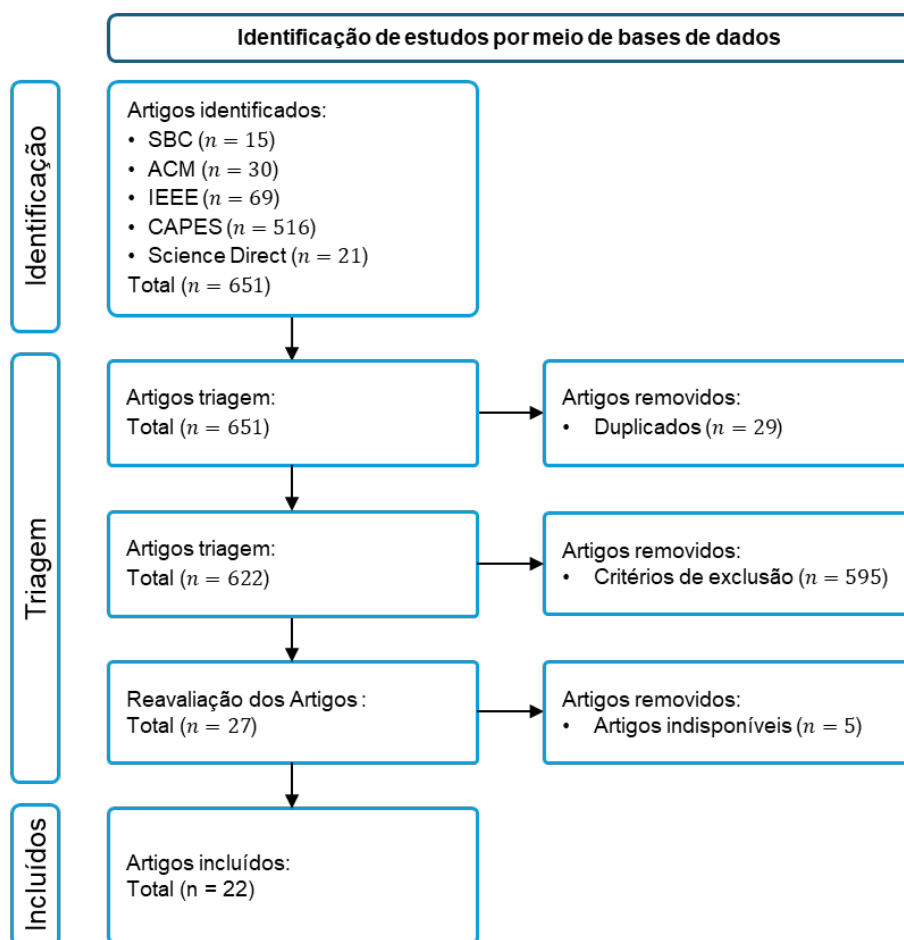


Figura 1. Fluxograma de seleção de artigos.

Fonte: Elaborado pelo autor (2026).

A maior parte das exclusões ocorreu por incompatibilidade com a população-alvo (estudos voltados ao ensino superior em Computação), pela ausência de instrumentos explícitos de avaliação do Pensamento Computacional ou por se tratarem de estudos secundários, como revisões da literatura e estudos de caráter teórico.

A Figura 2 apresenta a síntese dos 22 estudos incluídos nesta revisão, correlacionando os eixos temáticos à sua respectiva qualidade metodológica (escala CASP). O Eixo 1 (Validação e Adaptação) congrega o maior volume de trabalhos ($n = 8$), evidenciando a prioridade da comunidade científica na consolidação e calibração de instrumentos psicométricos robustos. Metodologicamente, a literatura mapeada demonstra elevada maturidade: aproximadamente 77% da amostra ($n = 17$) está alocada no quartil superior da escala de qualidade (CASP entre 9 e 11).

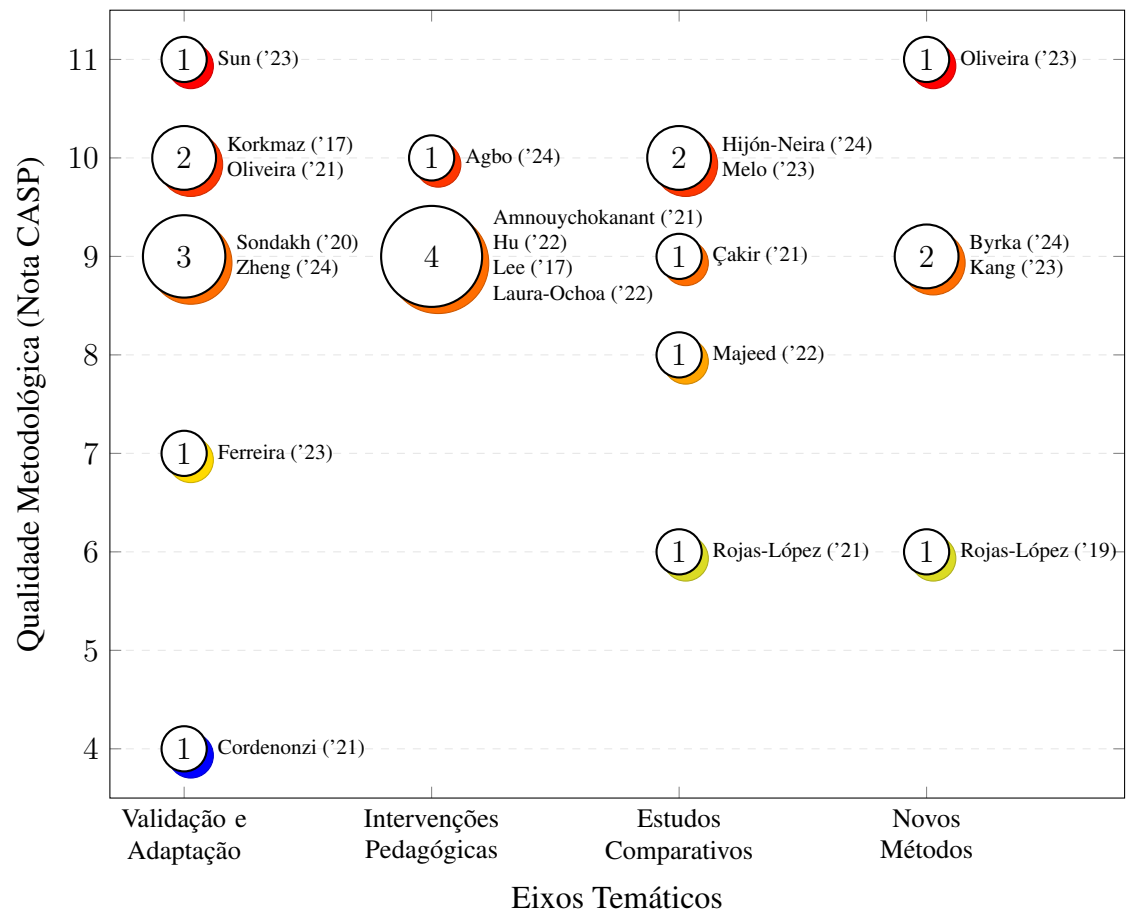


Figura 2. Distribuição dos estudos da revisão sistemática: Eixos Temáticos (X) e Qualidade Metodológica CASP (Y). O tamanho da bolha indica a quantidade de artigos mapeados.

Fonte: Elaborado pelos autores (2026).

A formação de um aglomerado denso e homogêneo de alta confiabilidade destaca-se no Eixo 2 (Intervenções Pedagógicas), no qual a maioria dos ensaios converge para a nota 9. Em contraste, o Eixo 4 (Novos Métodos) desponta como uma fronteira emergente; embora apresente o menor contingente de estudos ($n = 4$) e maior variância em suas notas (CASP 6 a 11), abriga pesquisas de altíssimo rigor investigativo. A Tabela 1 expõe os objetivos, os principais achados e as respectivas pontuações de qualidade (CASP) de todos os 22 estudos para evidenciar a riqueza dos dados coletados.

Tabela 1. Extração consolidada de dados dos 22 estudos incluídos na revisão sistemática, ordenados por qualidade metodológica (Nota CASP).

Estudo e Características
Cordenonzi e Pino (2021) • Resultado: Proposição de um “construto andragógico” para desenvolvimento e avaliação de adultos. O instrumento validado classifica o desempenho em “Alfabetizado em Código (ACod)”, “Pensador Computacional Desplugado (PCD)” ou não classificável. • Nota CASP: 4/11
Rojas-López e García-Penalvo (2019) • Resultado: A taxa de evasão prevista foi menor do que a esperada (aproximadamente 14%). O modelo obteve 92,31% de acerto nas previsões (36 acertos em 39 participantes). • Nota CASP: 6/11
Rojas-López e García-Penalvo (2021) • Resultado: Os resultados refutaram a hipótese de diferença significativa entre os gêneros. • Nota CASP: 6/11
Ferreira et al. (2023) • Resultado: Descrição do processo de adaptação do instrumento. Os 70 participantes indicaram maior uso das habilidades de “Criatividade” e “Cooperatividade” e menor uso de “Resolução de Problemas”. • Nota CASP: 7/11
Majeed et al. (2022) [Majeed et al. 2022] • Resultado: Os estudantes apresentaram nível razoável de PC. Observou-se diferença significativa entre gêneros, com desempenho superior do grupo masculino. • Nota CASP: 8/11
Amnouychokanant et al. (2021) • Resultado: A análise pré e pós-teste demonstrou evolução nas habilidades de PC dos alunos, independentemente de condições prévias. • Nota CASP: 9/11
Byrka et al. (2024) • Resultado: Definição e validação por especialistas de um <i>framework</i> de Pensamento Algorítmico contendo 100 declarações observáveis, distribuídas em: 32 de conhecimentos, 38 de habilidades e 30 de atitudes. • Nota CASP: 9/11
Çakir et al. (2021) • Resultado: Os estudantes turcos obtiveram maior média no construto “Criatividade”, enquanto os indianos superaram nos demais. Ambos os grupos apresentaram melhor <i>performance</i> em “Criatividade” e pior em “Resolução de Problemas”. • Nota CASP: 9/11

Continua na próxima página...

Tabela 1 – Continuação da página anterior

Estudo e Características
Hu et al. (2022) • Resultado: A diferença nas médias do pré-teste (21,80) e pós-teste (22,27) não foi estatisticamente significativa ($p = 0,67$). Houve aumento significativo na “Criatividade” ($p = 0,02$) e declínio na “Cooperação” ($p = 0,01$), este último atribuído aos efeitos da COVID-19. • Nota CASP: 9/11
Kang et al. (2023) • Resultado: Redução de 2 das 14 questões por baixa qualidade. Validação dos 5 construtos multidimensionais via MIRT, com itens apresentando qualidades psicométricas de boas a excelentes e capacidade de diferenciar habilidades entre cursos. • Nota CASP: 9/11
Lee e Cho (2017) • Resultado: Demonstrou-se melhoria estatística significativa ($p = 0,002$) entre as médias do pré-teste (8,26) e pós-teste (8,99). A análise ANCOVA não indicou diferença estatística entre gêneros. • Nota CASP: 9/11
Laura-Ochoa e Bedregal-Alpaca (2022) • Resultado: O método mostrou-se eficaz, com médias do grupo experimental (15,08) superiores às do grupo controle (14,03), apresentando diferença estatisticamente significativa ($p < 0,05$). • Nota CASP: 9/11
Sondakh et al. (2020a) • Resultado: O instrumento foi validado pelos especialistas. O construto “Automação” foi removido, enquanto “Trabalho em Equipe” e “Comunicação” foram destacados como importantes. • Nota CASP: 9/11
Sondakh et al. (2020b) • Resultado: A versão final do instrumento, reduzida de 155 para 114 questões, apresentou confiabilidade e validade, estando apta para avaliação no nível superior. • Nota CASP: 9/11
Zheng et al. (2024) • Resultado: Estabeleceu-se a correlação necessária para o desenvolvimento do instrumento, o qual demonstrou confiabilidade e validade após aplicação empírica. • Nota CASP: 9/11
Agbo et al. (2024) • Resultado: Os resultados indicaram melhor compreensão e confiança dos alunos sobre o conceito de PC após a utilização do método desplugado. • Nota CASP: 10/11

Continua na próxima página...

Tabela 1 – Continuação da página anterior

Estudo e Características
Hijón-Neira et al. (2024) • Resultado: Evidenciou-se correlação entre PC e programação. Não houve diferença significativa entre períodos letivos ou gêneros. Contudo, observou-se diferença significativa por localidade, com melhor desempenho dos alunos espanhóis em comparação aos irlandeses. • Nota CASP: 10/11
Korkmaz et al. (2017) • Resultado: O instrumento CTS foi validado e considerado confiável, constituído por 29 questões distribuídas em 5 construtos. • Nota CASP: 10/11
Melo et al. (2023) • Resultado: Não foi identificada diferença estatística significativa entre novatos e veteranos. • Nota CASP: 10/11
Oliveira et al. (2021) • Resultado: O instrumento demonstrou validade e confiabilidade após a eliminação de 3 questões do construto “Definição” devido à baixa validade. • Nota CASP: 10/11
Oliveira e Pereira (2023) • Resultado: Os 7 docentes participantes consideraram o artefato “Computacionalidade” um meio eficiente e eficaz para análise do PC. • Nota CASP: 11/11
Sun et al. (2023) • Resultado: A versão chinesa da CTS (24 questões) obteve alta validação. As médias variaram de moderadas a altas (15,71 de 25). Fatores pessoais correlacionaram-se com o desempenho, exceto o ano de graduação. • Nota CASP: 11/11

Fonte: Elaborado pelo autor (2026).

5. Discussão

A organização e a interpretação crítica do conhecimento consolidado na Tabela 1 permitiram a identificação de quatro grandes eixos temáticos. Essas categorias estruturam as pesquisas recentes sobre a avaliação do PC no ensino superior.

5.1. Eixo 1: Validação e Adaptação de Instrumentos Psicométricos

Este eixo concentra o maior volume de pesquisas dedicado ao rigor matemático e estatístico das métricas. Como marco fundacional desta categoria, a *Computational Thinking Scale* (CTS) foi desenvolvida por Korkmaz et al. (2017). A validação do instrumento, aplicada a mais de 1.300 estudantes na Turquia, estabeleceu os cinco construtos clássicos no nível de graduação.

A aceitação da CTS gerou ramificações transculturais. Uma versão chinesa da escala foi validada por Sun et al. (2023), agregando camadas teóricas da taxonomia de Marzano com elevada confiabilidade. No contexto brasileiro, o processo de tradução e adaptação para a língua portuguesa foi conduzido por Ferreira et al. (2023), assegurando a validade semântica e operacional do instrumento em graduandos nacionais. Essas iniciativas indicam a superação de barreiras idiomáticas na autoavaliação do PC.

A expansão para novas fronteiras avaliativas também ocorreu. Dimensões de trabalho em equipe, inteligência emocional e comunicação foram incorporadas no instrumento Hi-ACT por Sondakh et al. (2020b, 2020a). Esse instrumento foi reduzido de 155 para 114 questões mediante o método Delphi. Para fundamentar um questionário ancorado em dilemas do cotidiano (*Life Story situations*), a Teoria de Resposta ao Item Multidimensional (TRI-M) foi adotada por Kang et al. (2023). No cenário nacional, escalas focadas simultaneamente em PC e em atitudes comportamentais em relação à computação foram validadas no estudo de Oliveira et al. (2021).

A atenção voltada a públicos específicos motivou o desenvolvimento de abordagens direcionadas ao ensino de adultos. Um construto andragógico para a avaliação do PC nessa população foi proposto por Cordenonzi e Pino (2021). O instrumento validado subsidia a classificação do desempenho dos sujeitos em categorias como “Alfabetizado em Código (ACod)” e “Pensador Computacional Desplugado (PCD)”.

O detalhamento de componentes específicos do PC também foi alvo de investigações instrumentais. Um *framework* voltado ao Pensamento Algorítmico foi definido e validado por especialistas no estudo de Byrka et al. (2024). A ferramenta estrutura-se em 100 declarações observáveis, distribuídas entre dimensões de conhecimentos, habilidades e atitudes.

O panorama deste eixo indica que a engenharia psicométrica em torno do PC apresenta maturidade, disponibilizando instrumentos confiáveis. Entretanto, o foco acentuado na validação de ferramentas frequentemente adia a realização de estudos longitudinais.

5.2. Eixo 2: Avaliação de Intervenções e Estratégias Pedagógicas

A aplicação das escalas para testar metodologias ativas em sala de aula caracteriza as pesquisas do segundo eixo. Uma pluralidade de abordagens instrucionais foi submetida a ensaios quase-experimentais.

O uso da computação desplugada (*unplugged approach*) para alunos recém-ingressantes foi analisado na Nigéria por Agbo et al. (2024). Os resultados indicaram que essa abordagem auxilia na transição de nível educacional, elevando a autoconfiança e a compreensão estrutural dos conceitos. No contexto do ensino remoto emergencial decorrente da pandemia, a programação visual baseada em blocos foi avaliada por Amnouychokanant et al. (2021) na Tailândia. O estudo demonstrou evolução nas habilidades de PC e mitigação de deficits de conhecimento prévio.

A associação da robótica a sistemas de contação de histórias (*storytelling*) foi investigada por Hu et al. (2022). Embora tenha ocorrido expansão nas médias de criatividade, os construtos cooperativos apresentaram declínio. Esse comportamento foi associado aos efeitos do distanciamento social no período de coleta.

O impacto da linguagem Python no ensino de programação foi avaliado nos estudos de Lee e Cho (2017) e Laura-Ochoa e Bedregal-Alpaca (2022). O desenvolvimento de habilidades de PC em estudantes de *liberal arts* submetidos ao aprendizado dessa linguagem foi confirmado por Lee e Cho (2017). Adicionalmente, a associação de Python a plataformas iterativas de exercícios com *feedback* imediato, como o *CodingBat*, foi analisada por Laura-Ochoa e Bedregal-Alpaca (2022). Os dados empíricos indicaram desempenho superior desse método em comparação ao modelo expositivo tradicional.

5.3. Eixo 3: Estudos Comparativos, Transversais e Demográficos

As pesquisas deste eixo são dedicadas a contrapor grupos, nações e trajetórias para mapear as variações da habilidade. Um estudo multicêntrico envolvendo universitários da Espanha e da Irlanda foi conduzido por Hijón-Neira et al. (2024), indicando diferenças no preparo lógico superior com vantagem estatística para os estudantes espanhóis. Em abordagem comparativa desenvolvida por Çakir et al. (2021), observou-se que licenciandos turcos apresentaram maiores médias no construto “Criatividade”, ao passo que os discentes indianos superaram nos demais domínios. Ambos os grupos convergiram no menor desempenho na resolução de problemas.

As investigações sobre disparidades de gênero apresentaram dados divergentes. A hipótese de diferença significativa de habilidades entre gêneros na computação foi refutada no estudo longitudinal de Rojas-López e García-Peñalvo (2021). Por outro lado, um desempenho superior do grupo masculino foi observado por Majeed et al. (2022) no Oriente Médio. Essa divergência sugere o impacto de fatores socioculturais e do acesso prévio a tecnologias sobre os resultados estatísticos.

A análise da maturação temporal intrainstitucional foi realizada por Melo et al. (2023) no sistema universitário brasileiro. O estudo comparou o nível de PC entre alunos ingressantes e veteranos, revelando ausência de diferença estatística significativa entre os grupos. Esse resultado alinha-se aos achados obtidos por Sun et al. (2023) na China, onde também não foram observados avanços associados ao ano de graduação. Essas evidências sugerem que a progressão curricular focada predominantemente na sintaxe de linguagens de programação pode não promover o desenvolvimento pleno das habilidades fundamentais do PC.

5.4. Eixo 4: Novos Métodos Analíticos e Previsibilidade

A alternativa aos instrumentos baseados em formulários de autopercepção (*self-report*) caracteriza as pesquisas do quarto eixo. O modelo P-M-E de avaliação foi proposto por Zheng et al. (2025). O método baseia-se na captura direta e análise estática do código de programação estruturado pelo estudante para mensurar o nível de PC.

A capacidade do PC de atuar como métrica preditiva foi modelada por Rojas-López e García-Peñalvo (2019). O uso de testes diagnósticos nos primeiros dias letivos permitiu o desenvolvimento de um modelo com acurácia de 92,31% para a previsão de retenção e evasão de estudantes. Essa modelagem possibilita a aplicação de avaliações cognitivas no suporte à gestão acadêmica.

Instrumentos de suporte qualitativo e contextual também foram desenvolvidos. No cenário brasileiro, a matriz Computacionalidade foi proposta por Oliveira e Pereira

(2023). O artefato foi projetado para orientar a observação docente sobre indícios práticos de PC manifestados em sala de aula, integrando a avaliação qualitativa ao processo.

De forma geral, os estudos analisados demonstram que a pesquisa sobre avaliação do Pensamento Computacional no ensino superior evoluiu significativamente nos últimos anos, sobretudo no desenvolvimento e validação de instrumentos psicométricos. Entretanto, a predominância de estudos transversais evidencia uma lacuna na literatura quanto ao acompanhamento da evolução dessas habilidades ao longo da formação acadêmica, reforçando a necessidade de pesquisas longitudinais e de abordagens que combinem diferentes estratégias de avaliação.

6. Conclusão

A condução desta revisão sistemática permitiu mapear o panorama investigativo sobre a avaliação do PC no cenário universitário. A fase de validação e estruturação de ferramentas psicométricas apresenta consolidação teórica e eficácia, sustentada por testes globais de ampla amostragem.

As evidências coligidas pelos estudos de aplicação prática indicam que a instrução focada em codificação não equivale, automaticamente, ao desenvolvimento do PC. O ensino baseado em estruturas de linguagens procedurais, funcionais e orientadas a objetos demonstrou-se insuficiente para amadurecer competências de abstração, decomposição e reconhecimento de padrões em estudantes veteranos. A inércia estatística entre ingressantes e alunos de períodos avançados indica que o modelo curricular atual prioriza a escrita de algoritmos em detrimento da formação analítica.

As limitações mapeadas no estado da arte direcionam o desenvolvimento futuro deste campo de pesquisa. Torna-se necessária a transição de produções teóricas focadas em validações transversais para a execução de ensaios longitudinais. Essa mudança requer o monitoramento do mesmo contingente de graduandos do início ao término do curso. Adicionalmente, demanda-se a utilização híbrida de escalas de autopercepção combinadas a testes de competência prática estritamente lógicos. Essas ações oferecem subsídios para reestruturar os currículos e fomentar habilidades analíticas voltadas à era digital.

7. Declaração sobre uso de Inteligência Artificial

Os autores declaram que as ferramentas Claude e Gemini foram utilizadas exclusivamente para fins de revisão gramatical e estrutural do texto. O uso dessas ferramentas não substitui a autoria humana, e os autores assumem total responsabilidade pela integridade, originalidade e exatidão de todo o conteúdo apresentado, garantindo a ausência de plágio.

Referências

- Agbo, F. J. et al. (2024). How can unplugged approach facilitate novice students' understanding of computational thinking? an exploratory study from a nigerian university. *Thinking Skills and Creativity*, 51:101458.
- Amnouychokanant, V. et al. (2021). Online learning using block-based programming to foster computational thinking abilities during the covid-19 pandemic. *International Journal of Emerging Technologies in Learning*, 16(13):227.

- Byrka, M. et al. (2024). Algorithmic thinking in higher education: Determining observable and measurable content. *Information Technologies and Learning Tools*, 104(6):1.
- Çakir, R., Rosaline, S., e Korkmaz, Ö. (2021). Computational thinking skills of turkish and indian teacher candidates: A comparative study. *International Journal of Psychology and Educational Studies*, 8(1):24–37.
- Cordenonzi, W. H. e Pino, J. C. D. (2021). Um constructo andragógico para avaliação do pensamento computacional no adulto. *Revista Electrónica de Investigación en Educación en Ciencias*, 16(1):41–55.
- Critical Appraisal Skills Programme (2024). Checklist para estudos qualitativos. <https://casp-uk.net/>. Acesso em: ago. 2025.
- Ferreira, S. D. N. et al. (2023). Processo de tradução de instrumento diagnóstico de pensamento computacional para graduandos. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação (SBIE)*, pages 58–68, Porto Alegre. Sociedade Brasileira de Computação.
- Higgins, J. P. T. e Thomas, J., editors (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons, Hoboken, 2 edition.
- Hijón-Neira, R. et al. (2024). Computational thinking measurement of cs university students. *Applied Sciences*, 14(12):5261.
- Hu, C.-C. et al. (2022). The effects of constructing robot-based storytelling system on college students' computational thinking skill and technology comprehension. In *Proceedings of the 24th International Conference on Advanced Communication Technology (ICACT)*, pages 496–500. IEEE.
- Kang, C. et al. (2023). Developing college students' computational thinking multidimensional test based on life story situations. *Education and Information Technologies*, 28(3):2661–2679.
- Kitchenham, B. e Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, Keele University and Lincoln University.
- Korkmaz, Ö., Çakir, R., e Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (cts). *Computers in Human Behavior*, 72:558–569.
- Laura-Ochoa, L. e Bedregal-Alpaca, N. (2022). Incorporation of computational thinking practices to enhance learning in a programming course. *International Journal of Advanced Computer Science and Applications*, 13(2).
- Lee, Y. e Cho, J. (2017). The influence of python programming education for raising computational thinking. *International Journal of u-and e-Service, Science and Technology*, 10(8):59–72.
- Majeed, B. H., Jawad, L. F., e Alrikabi, H. T. H. (2022). Computational thinking (ct) among university students. *International Journal of Interactive Mobile Technologies*, 16(10).
- Mélo, Á. M. d. et al. (2023). Avaliação do pensamento computacional em graduandos de cursos de computação: uma disciplina de programação faz diferença? In *Anais do*

- XXXIV Simpósio Brasileiro de Informática na Educação (SBIE), pages 47–57, Porto Alegre. Sociedade Brasileira de Computação.
- Moher, D. et al. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *PLoS Medicine*, 6(7):e1000097.
- Oliveira, C. M. e Pereira, R. (2023). Coleta de evidências do exercício do pensamento computacional no ensino superior em computação: um artefato de apoio. In *Anais do III Simpósio Brasileiro de Educação em Computação (EduComp)*, pages 300–309, Porto Alegre. Sociedade Brasileira de Computação.
- Oliveira, E. C., Correia, R. C. M., e Bittencourt, R. A. (2021). Evaluating a computational thinking and computing attitudes instrument for educational purposes. In *Proceedings of the IEEE Global Engineering Education Conference (EDUCON)*, pages 748–754. IEEE.
- Organização para a Cooperação e Desenvolvimento Econômico (OCDE) (2023). Pisa 2022 results (volume i and ii): Country notes – brazil. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_d6fbcc5-en/brazil_1690648-en.html. Acesso em : 4jun.2025.
- Raabe, A. et al. (2017). Um instrumento para diagnóstico do pensamento computacional. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação (CBIE)*, pages 1172–1183, Porto Alegre. Sociedade Brasileira de Computação.
- Rojas-López, A. e García-Peñalvo, F. J. (2019). Interpretation of computational thinking evaluation results for enrollment prediction. In *Proceedings of the 7th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM)*, pages 13–18. ACM.
- Rojas-López, A. e García-Peñalvo, F. J. (2021). Initial performance analysis in the evaluation of computational thinking from a gender perspective in higher education. In *Proceedings of the 9th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM)*, pages 109–114. ACM.
- Sondakh, D. E., Osman, K., e Zainudin, S. (2020a). A pilot study of an instrument to assess undergraduates’ computational thinking proficiency. *International Journal of Advanced Computer Science and Applications*, 11(11):263–273.
- Sondakh, D. E., Osman, K., e Zainudin, S. (2020b). A proposal for holistic assessment of computational thinking for undergraduate: Content validity. *European Journal of Educational Research*, 9(1):33–50.
- Sun, L. et al. (2023). Evaluation and developmental suggestions on undergraduates’ computational thinking: a theoretical framework guided by marzano’s new taxonomy. *Interactive Learning Environments*, 31(10):6588–6610.
- Tsai, M.-J., Liang, J.-C., e Hsu, C.-Y. (2021). The computational thinking scale for computer literacy education. *Journal of Educational Computing Research*, 59(4):579–602.
- UPMAT Educacional (2024). Desafio bebras brasil. <https://www.bebasbrasil.com.br/>. Acesso em: 26 maio 2025.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3):33–35.

Zhang, W., Zeng, X., e Song, L. (2025). Towards an assessment model of college students' computational thinking with text-based programming. *Education and Information Technologies*, 30(2):1363–1385.