

Feedback Pedagógico e Adaptativo em Jogos Sérios: Uma Revisão Sistemática com foco em Pensamento Computacional e Equidade de Gênero

Pedagogical and Adaptive Feedback in Serious Games: A Systematic Review on Computational Thinking and Gender Equity.

Ivan Ferreira Martins, Sergio Teixeira de Carvalho, Luciana de Oliveira Berretta

¹ Instituto de Informática (INF) – Universidade Federal de Goiás (UFG)
Alameda Palmeiras, Quadra D, Câmpus Samambaia
CEP 74690-900 – Goiânia – GO – Brasil

ivan.martins@discente.ufg.br

{sergiocarvalho, luciana.berretta}@ufg.br

Abstract. Introduction: Over the past decades, serious games have evolved from peripheral applications into relevant tools in educational, corporate, and training contexts. **Objective:** This article aims to investigate the integration of pedagogical feedback and adaptive feedback in serious games, focusing on the development of computational thinking and gender equity. **Methodology:** To achieve this objective, a Systematic Literature Review was conducted following the PRISMA 2020 guidelines. **Results:** The results indicate an increasing adoption of artificial intelligence techniques, such as stealth assessment, conversational agents, telemetry analysis, and frustration detection, to provide real-time personalized feedback. However, the review also revealed limited explicit pedagogical grounding, heterogeneity in the metrics used to assess computational thinking, and a low incorporation of variables related to gender equity.

Keywords Serious games, Adaptive feedback, Pedagogical feedback, Computational thinking; Gender equity.

Resumo. Introdução: Nas últimas décadas, os jogos sérios passaram de aplicações periféricas a instrumentos relevantes em contextos educacionais, corporativos e formativos. **Objetivo:** Assim, este artigo tem como objetivo investigar a integração entre feedback pedagógico e feedback adaptativo em jogos sérios, com foco no desenvolvimento do pensamento computacional e na equidade de gênero. **Metodologia:** Para atingir esse objetivo, foi realizada uma Revisão Sistemática da Literatura, utilizando as diretrizes PRISMA 2020. **Resultados:** Os resultados indicam o crescimento do uso de técnicas de inteligência artificial, como stealth assessment, agentes conversacionais, análise de telemetria e detecção de frustração, para a personalização do feedback em tempo real. Entretanto, observou-se fundamentação pedagógica explícita limitada, heterogeneidade nas métricas de avaliação do pensamento computacional e baixa incorporação de variáveis relacionadas à equidade de gênero.

Palavras-Chave Jogos sérios, Feedback adaptativo, Feedback pedagógico, Pensamento computacional, Equidade de gênero.

1. Introdução

Nas últimas décadas, os jogos sérios passaram de aplicações periféricas a instrumentos relevantes em contextos educacionais, corporativos e formativos. Seu crescimento está associado à capacidade de combinar desafio, interatividade, narrativa e experimentação em ambientes que favorecem a motivação e a persistência [Mao et al. 2024]. Em áreas como computação, onde estudantes frequentemente enfrentam barreiras cognitivas e afetivas relacionadas à abstração, ao erro recorrente e à ansiedade de desempenho, tais ambientes têm sido considerados especialmente promissores [Akkaya et al. 2022].

No ensino de programação e de pensamento computacional, jogos sérios podem contribuir para o desenvolvimento da capacidade de decompor problemas, do raciocínio algorítmico, do reconhecimento de padrões e da depuração de soluções. Entretanto, a simples inserção de elementos lúdicos não garante uma aprendizagem efetiva. A literatura tem demonstrado que o impacto educacional desses ambientes depende menos da estética do jogo e mais da qualidade do design instrucional que sustenta a experiência [Varajão et al. 2026, Calvo-Morata et al. 2026].

Nesse contexto, o *feedback* emerge como componente central. Em perspectiva educacional, *feedback* consiste em informações fornecidas ao aprendiz acerca de sua situação atual, dos objetivos esperados e dos caminhos possíveis para avançar [Hattie e Timperley 2007]. Em ambientes digitais, entretanto, o *feedback* assume também uma dimensão computacional: algoritmos podem interpretar comportamentos, inferir dificuldades e adaptar intervenções em tempo real [Min et al. 2020].

Embora o avanço técnico seja evidente, muitos estudos permanecem centrados na eficiência algorítmica, negligenciando fundamentos pedagógicos explícitos. Adicionalmente, as dimensões de equidade de gênero ainda ocupam um espaço marginal, mesmo em domínios historicamente marcados por desigualdades na participação, como a computação. Assim, torna-se pertinente compreender como a literatura recente tem integrado essas dimensões.

Neste sentido, este trabalho tem por objetivo documentar uma pesquisa bibliográfica do tipo Revisão Sistemática da Literatura, que visa explorar, identificar e analisar como *feedback* pedagógico e *feedback* adaptativo vêm sendo integrados em jogos sérios, e em que medida tais estudos contemplam pensamento computacional e equidade de gênero. Assim, torna-se pertinente compreender como a literatura recente tem integrado essas dimensões.

Este artigo está organizado em quatro seções, além desta introdução. A Seção 2 apresenta a metodologia da revisão sistemática, incluindo protocolo de busca, critérios de seleção e processo PRISMA. A Seção 3 descreve os resultados, incluindo a caracterização dos estudos, os tipos de *feedback*, as mecânicas de jogo e as métricas de avaliação. A Seção 4 discute criticamente os achados à luz do pensamento computacional, da fundamentação pedagógica e da equidade de gênero. Por fim, a Seção 5 apresenta as conclusões, limitações e direções para trabalhos futuros.

2. Metodologia da Pesquisa

Esta revisão sistemática da literatura seguiu as diretrizes PRISMA 2020 [Page et al. 2021] e o protocolo de Kitchenham e Charters [Kitchenham e Charters 2007] para revisões

sistemáticas em engenharia de software, adaptados ao contexto de tecnologias educacionais. O escopo da revisão foi definido utilizando o framework PICOC (População, Intervenção, Comparação, Resultados, Contexto), conforme apresentado na Tabela 1

Tabela 1. Framework PICOC da Revisão Sistemática

Dimensão	Descrição
População	Estudantes de ensino fundamental, médio e superior; educadores; designers de jogos educacionais.
Intervenção	Jogos sérios que integram feedback pedagógico e/ou adaptativo para o desenvolvimento do pensamento computacional.
Comparação	Abordagens tradicionais de ensino de PC; jogos sem feedback adaptativo; jogos sem consideração explícita de equidade de gênero.
Resultados	Aprendizagem de PC (decomposição, abstração, algoritmos, reconhecimento de padrões); engajamento; motivação; desempenho desagregado por gênero; percepções de autoeficácia; redução das disparidades de gênero.
Contexto	Ambientes formais e informais de aprendizagem; contextos escolares, extraescolares e online; estudos publicados entre 2020 e 2026.

Nota: O framework PICOC foi adaptado para focar na interseção entre feedback adaptativo e equidade de gênero.

2.1. Questões de Pesquisa Secundárias

Seis questões de pesquisa (QP) orientaram a coleta e análise de dados: **QP1** - Como as mecânicas de *feedback* de jogos são integradas a algoritmos de IA para criar *feedback* adaptativo? - **QP2** - Quais modelos teóricos de *feedback* pedagógico fundamentam os sistemas de *feedback* em jogos sérios? - **QP3** - Como o *feedback* formativo é articulado com elementos de design de jogo (imersão, desafio, regras)? - **QP4** - De que maneira o design do *feedback* mediado por IA aborda variáveis de equidade de gênero no ensino de tecnologia? - **QP5** - Quais métricas são utilizadas para avaliar o impacto do *feedback* adaptativo no desenvolvimento do pensamento computacional? - **QP6** - Quais princípios de design instrucional orientam a integração do *feedback* adaptativo em jogos sérios?

2.2. Estratégia de Busca

A busca foi realizada em cinco bases de dados eletrônicas (IEEE Xplore, ACM Digital Library, Scopus, Web of Science e ERIC) em janeiro de 2026. Os artigos foram acessados por meio do Portal de Periódicos CAPES, garantindo acesso ao texto completo das publicações restritas. A frase de busca foi construída combinando elementos do PICOC:

```
("students" OR "K-12" OR "higher education" OR "vocational education" OR "female students" OR "underrepresented groups" or "women in STEM" ) AND ("serious games" OR "educational games" OR "game-based learning" OR "game mechanics" OR "digital learning games") AND ("feedback" OR "formative feedback" OR "adaptive feedback" OR "personalized feedback" OR "feed up" OR "feed forward" OR "pedagogical feedback" OR "automated feedback" OR "adaptive systems" OR "AI-mediated feedback") AND ("computational thinking" OR "programming" OR "computer science education" OR "computer science education" OR "algorithmic thinking" OR "computational literacy" OR "programming skills")
```

Filtros aplicados: artigos publicados entre 2020 e 2026, em inglês ou em português, revisados por pares.

2.3. Critérios de Inclusão e Exclusão

Os critérios de inclusão (CI) e de exclusão (CE) foram definidos para garantir a relevância e a qualidade, conforme apresentado na Tabela 2.

Tabela 2. Critérios de Inclusão (CI) e Exclusão (CE)

Critérios de Inclusão (CI)	Critérios de Exclusão (CE)
CI1: Realizar comparação com outras formas de <i>feedback</i> (apenas estudos comparativos).	CE1: Falhas metodológicas graves, sem dados amostrais ou de duplicata.
CI2: Envolver estudantes em contextos educacionais com dados demográficos mínimos.	CE2: Participantes não humanos ou populações clínicas especializadas.
CI3: Serem estudos empíricos primários revisados por pares.	CE3: Publicados antes de 2020, em idiomas não cobertos ou sem texto completo disponível.
CI4: Apresentar <i>feedback</i> adaptativo com fundamentação pedagógica.	CE4: Apenas métricas técnicas ou sem metodologia de avaliação.
CI5: Abordar jogos sérios/educacionais em contextos STEM e de computação.	CE5: <i>Feedback</i> estático ou não adaptativo.
CI6: Ter pelo menos um desfecho educacional com metodologia descrita.	CE6: Gamificação pura, jogos comerciais não educacionais ou sem <i>feedback</i> .
CI7: Publicados em português, espanhol ou inglês.	CE7: Não atendem aos critérios de inclusão.
CI8: Publicados entre 2020 e 2026.	CE8: Não empíricos, revisões, protocolos, resumos, teses.

2.4. Seleção dos Estudos e Fluxo PRISMA

A seleção dos estudos foi realizada em quatro etapas: identificação, triagem, elegibilidade e inclusão. As etapas iniciais de identificação e triagem foram conduzidas com os seguintes resultados: a busca nas bases de dados resultou em 152 registros; após a remoção de duplicatas (n=27), 125 registros foram triados com base no título e no resumo sendo excluídos 89 registros por não atenderem aos critérios CI1–CI6.. Após essa etapa, a avaliação final de elegibilidade com 36 artigos, a definição dos estudos incluídos 13 estudos e a extração dos dados foram conduzidas seguindo os critérios de inclusão e exclusão previamente estabelecidos no protocolo da revisão. Os resultados de cada etapa estão detalhados na Figura 1:

Dos estudos incluídos foram extraídas as seguintes informações: referência completa; contexto educacional e público-alvo; delineamento metodológico; fundamentos teóricos; tipo e mecanismo de *feedback* implementado; mecânicas de jogo; métricas de avaliação; relação com pensamento computacional; relação com equidade de gênero; limitações; contribuições e lacunas.

A revisão pode ter sido limitada pelo escopo linguístico (português, espanhol e inglês) e pelo recorte temporal (2020–2026). O possível viés de publicação foi mitigado pela inclusão de cinco bases de dados diversificadas. Entretanto, a decisão final de elegibilidade e a extração dos dados foram realizadas por um único pesquisador, o que pode introduzir viés de interpretação, embora tenham sido seguidos critérios previamente definidos no protocolo da revisão.

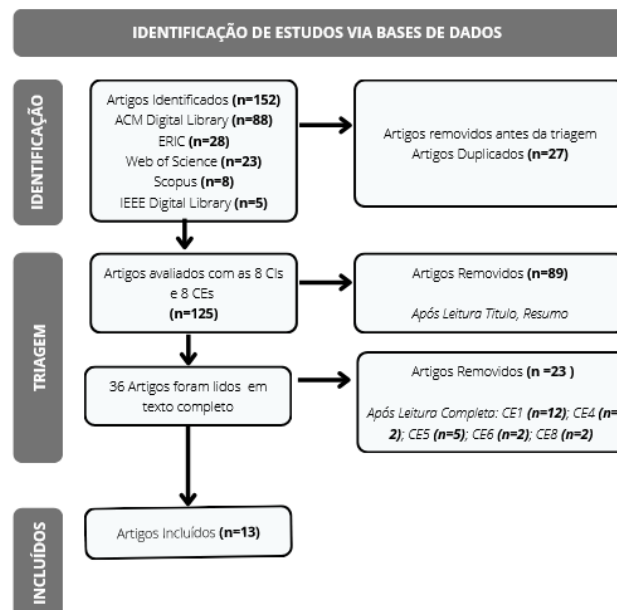


Figura 1. Fluxograma PRISMA 2020 detalhando o processo de seleção

3. Resultados

Os estudos analisados revelam um campo em expansão, porém ainda heterogêneo em termos de objetivos, métodos e profundidade teórica. Esta seção apresenta os achados organizados em subseções temáticas. Os estudos incluídos na revisão foram organizados na Tabela 3, com uma caracterização completa, incluindo: código do estudo, criado para maior fluidez no texto; autor; ano; contexto; público; método; e foco do *feedback*.

O corpus apresenta predominância de investigações experimentais e quase experimentais (8 estudos), frequentemente realizadas em contextos de programação introdutória, ensino médio ou graduação em computação [ES-1], [ES-2], [ES-4], [ES-5], [ES-11], [ES-12], [ES-13]. Também foram identificados estudos orientados por *Design Science Research* (DSR) [ES-8] e *Design-Based Research* (DBR) [ES-7], sobretudo quando o foco estava na construção de artefatos educacionais.

O tamanho das amostras apresentou ampla variação entre os estudos, abrangendo de 24 participantes [ES-10] a 632 estudantes [ES-3], com mediana de 150 participantes. Predominaram estudos com amostras entre aproximadamente 70 e 220 participantes, embora também tenham sido identificadas investigações de grande escala, com 400 [ES-12], [ES-13] e 632 participantes [ES-3]. Os estudos foram conduzidos predominantemente em ambientes controlados, como laboratórios e salas de aula. Quanto à duração das intervenções, observaram-se tanto estudos de sessão única [ES-2], [ES-11], [ES-12] quanto intervenções desenvolvidas ao longo de múltiplas semanas [ES-1], [ES-4], [ES-5], [ES-7], [ES-8], [ES-9], [ES-10], [ES-13].

3.1. Fundamentos Teóricos Mobilizados

Conforme apresentado na Tabela 4, que sintetiza os fundamentos teóricos identificados nos estudos, observa-se que uma parte significativa dos estudos mobilizou referenciais voltados à mensuração ou à motivação, como *Evidence-Centered Design* [ES-2], *Flow*

Tabela 3. Caracterização dos Estudos do Corpus

Nº	Autor	Contexto	Público	Método	Foco do <i>Feedback</i>
[ES-1]	[Mao et al. 2024].	Programação	Ensino Superior	Experimento unifatorial (4 condições)	<i>Feedback</i> estático vs. dinâmico, simples vs. detalhado
[ES-2]	[Min et al. 2020]	Ciência da Computação.	Ensino Fundamental	Comparação de modelos (LSTMs, FFNNs, CRFs)	<i>Stealth assessment</i> com redes neurais
[ES-3]	[Lu et al. 2025]	Hidrologia	Ensino Fundamental	Análise de logs com Redes Bayesianas	Predição de ganhos de aprendizagem via telemetria
[ES-4]	[Chen et al. 2025]	Pensamento Computacional	Ensino Médio	Experimento randomizado (5 condições)	Recompensas e competição
[ES-5]	[Xu et al. 2024]	TI	6ª série	Análise de sequência atrasada	Chatbot de IA integrado ao jogo
[ES-6]	[Maqsood e Chiasson 2021]	Cibersegurança	K-12	Design Participativo	<i>Feedback</i> via consequências simuladas
[ES-7]	[Parthasarathy e Joshi 2026]	Acessibilidade Web	Graduação	<i>Design-Based Research</i> (DBR)	<i>Feedback</i> formativo + <i>analytics</i>
[ES-8]	[Varajão et al. 2026]	Métodos Ágeis	Profissionais	<i>Design Science Research</i> (DSR)	<i>Feedback</i> via minijogos e questionários
[ES-9]	[Liverman et al. 2025]	Socioemocionais	K-12	<i>Design crossover</i> fatorial	<i>Feedback</i> emocional imerso na narrativa
[ES-10]	[Gakingo e Kabia 2025]	Engenharia	Ensino Superior	Aprendizagem baseada em projetos	<i>Feedback</i> implícito via resolução de tarefas
[ES-11]	[Lee e Chiou 2020]	Programação Inicial	Adultos	Experimento (3 condições de dicas)	Dicas estáticas, carrossel de imagens, vídeos
[ES-12]	[Lee 2020a]	Algoritmos	Novatos online	Experimento com controle	Deteccção de frustração + <i>feedback</i> adaptativo
[ES-13]	[Lee 2020b]	Programação	Novatos online	Experimento com controle	Geração procedural de níveis adaptativos

Theory [ES-7] e *Cognitive Load Theory* [ES-1]. Entretanto, modelos clássicos de *feedback* formativo, como o de [Hattie e Timperley 2007], apareceram de forma menos explícita. Em diversos casos, o *feedback* foi tratado como uma funcionalidade de interface ou um mecanismo de suporte, e não como uma estratégia pedagógica estruturada.

3.2. *Feedback* Adaptativo e Tecnologias Empregadas

Em relação ao *feedback* adaptativo, observou-se forte presença de abordagens baseadas em dados comportamentais, conforme apresentado na Tabela 5. Sistemas de *stealth assessment* utilizaram logs de interação, sequências de ações e indicadores temporais para inferir competências sem interromper o fluxo do jogo [ES-3], [ES-4]. Em alguns casos, redes neurais profundas [ES-3] e redes bayesianas [ES-4] foram empregadas para prever o desempenho ou estimar a progressão da aprendizagem em tempo real.

Tabela 4. Síntese dos Fundamentos e Métricas dos Estudos

Nº	Fundamentos Teóricos	Métricas de Avaliação
[ES-1]	FIT, ITFM, Teoria Carga Cognitiva	Desemp. acadêmico, engajamento, seq. atraso.
[ES-2]	ECD, <i>Stealth Assessment</i>	Acurácia modelos, inferência competência.
[ES-3]	<i>Stealth Assessment</i> , NGSS, ATMSG	Ganhos aprendizagem, logs telemetria.
[ES-4]	Recompensa Intrínsecas/Extrínsecas, DGBL	PC (Bebras), tensão/ansiedade, imersão.
[ES-5]	DGBL, HOTS (incluindo PC)	Desempenho acadêmico, teste PC, uso chatbot.
[ES-6]	<i>Procedural Rhetoric</i> , Princípios Gee, Aprendizagem Ativa	Qualitativa (design participativo, feedback).
[ES-7]	<i>Sociocultural Learning Theory</i> , <i>Flow Theory</i>	Pré/pós-testes, <i>gameplay</i> , Grupos Focais.
[ES-8]	PBL, Collab. Learning, Microlearning	Desempenho minijogos, questionários.
[ES-9]	SEL, Trauma, Inteligencia Emocional, CRAFT	SEL, qualitativa (foco e entrevistas).
[ES-10]	<i>Game-based Learning</i> , Bloom	Sucesso formulação/resolução de tarefas.
[ES-11]	Aprendizagem Ativa, Visualização, <i>Idea Garden</i>	Persistência, acesso ao sistema de dicas.
[ES-12]	Deteção/Intervenção de Frustração	Persistência, marcadores comportamentais.
[ES-13]	CAT, Geração Procedural, <i>Scaffolding</i>	Persistência, engajamento, níveis concluídos.

Outra tendência relevante foi a incorporação de agentes conversacionais. *Chatbots* integrados ao ambiente lúdico foram utilizados para responder a dúvidas, orientar tarefas e oferecer *feedback* imediato [ES-5]. Em contextos de ensino de tecnologia, tais agentes demonstraram efeitos positivos no desempenho acadêmico, na motivação e em indicadores de pensamento computacional.

Três estudos exploraram a detecção de frustração como gatilho para *feedback* adaptativo [ES-6], [ES-12], [ES-13]. Marcadores comportamentais (tempo ocioso, execuções de código com erros repetidos, variações de desempenho) acionavam dicas contextuais ou a geração procedural de novos níveis de prática.

3.3. Mecânicas de Jogo e *Feedback*

No que se refere às mecânicas de jogo, o *feedback* apareceu associado principalmente a: Dicas progressivas: Sistemas de suporte contextual que oferecem dicas estáticas, carrossel de imagens ou vídeos animados [ES-11]; Sistemas de recompensa: *Power-ups*, customização de personagens, pontos [ES-4]; Progressão por níveis: Desbloqueio de conteúdo, *lock-and-key* [ES-8]; Competição por ranking: *Leaderboards* [ES-4]; Narrativa contextual: *Feedback* imerso na narrativa do jogo [ES-6], [ES-9]; geração procedural de desafios: criação automática de níveis adaptativos [ES-13].

Estudos focados em pensamento computacional sugeriram que recompensas bem calibradas podem reduzir a tensão e favorecer a persistência, enquanto a competição excessiva tende a deslocar o foco do raciocínio para o ganho imediato [ES-4].

3.4. Métricas de Avaliação

As métricas utilizadas concentraram-se em quatro grandes grupos, conforme já demonstrado na Tabela 4. O desempenho acadêmico é avaliado por meio de pré e pós-testes e de avaliações de conhecimento [ES-1], [ES-3], [ES-5], [ES-7]. O engajamento baseado no tempo de uso, nos níveis concluídos e na permanência [ES-1], [ES-11], [ES-12], [ES-13]. Já o pensamento computacional utilizou instrumentos específicos, como

Tabela 5. Tipos de *feedback* Detalhamento Técnico e Pedagógico dos Estudos Seleccionados

Nº	Tipo de <i>Feedback</i>	Detalhes Técnicos	Detalhes Pedagógicos
[ES-1]	Estático vs. Dinâmico; Simples vs. Detalhado	Análise de sequência de atraso.	<i>Feedback</i> detalhado para tarefas fáceis pode reduzir o engajamento; explicações passo a passo para tarefas complexas.
[ES-2]	<i>Stealth Assessment</i>	LSTMs e FFNNs analisam ações de baixo nível no código em tempo real.	Avaliação oculta sem interrupção do fluxo do jogo.
[ES-3]	Predição de Ganhos	Redes Bayesianas Estáticas cruzam logs comportamentais (ferramentas, tempo).	Predição em tempo real dos ganhos de aprendizagem.
[ES-4]	Recompensas e Competição	Experimento com <i>power-ups</i> vs. customização; <i>leaderboards</i> .	Recompensas intrínsecas reduzem tensão/ansiedade; <i>leaderboards</i> prejudicam estratégia.
[ES-5]	Chatbot de IA	Agente conversacional (LLM) integrado ao jogo.	Alunos de alto desempenho usam seletivamente; alunos de baixo desempenho entram em diálogos livres.
[ES-6]	Consequências Simuladas	Regras transmitem um argumento instrucional.	<i>Feedback</i> como consequência da escolha; <i>microlearning</i> e projetos.
[ES-7]	Formativo + <i>Analytics</i>	Mecânicas de <i>Platformer</i> e progressão restrita (<i>lock-and-key</i>).	Aprendizagem ativa, social e situada por meio de <i>role-play</i> .
[ES-8]	Via Mini-Jogos	Desempenho em mini-jogos e questionários.	<i>Feedback</i> inferido do desempenho.
[ES-9]	Emocional Narrativo	Regulação da ansiedade imersa na narrativa <i>sci-fi</i> .	<i>Feedback</i> sobre regulação emocional mantém a imersão.
[ES-10]	Implícito	Sucesso ou falha na formulação e na resolução.	<i>Feedback</i> implícito no resultado da tarefa.
[ES-11]	Dicas Multimodais	Estáticas, carrossel de imagens, vídeos animados.	Vídeos/animações melhoram a persistência em níveis difíceis.
[ES-12]	Deteção de Frustração	Marcadores de granulação grossa (tempo ocioso, execuções consecutivas).	<i>Feedback</i> contextual (<i>Idea Garden approach</i>).
[ES-13]	Níveis Adaptativos	Geração de Conteúdo Procedural (Síntese de Programa).	Prática adaptativa extra mediante deteção de frustração.

a *Bebras Challenge* e tarefas algorítmicas [ES-4], [ES-5]. A telemetria comportamental foi medida com o uso de dicas, tempo entre ações, padrões de navegação [ES-2], [ES-3], [ES-5], [ES-12]. Essa diversidade confirma a riqueza metodológica, mas também dificulta comparações diretas entre os estudos.

3.5. Pensamento Computacional

Seis estudos abordaram explicitamente o pensamento computacional [ES-1], [ES-2], [ES-4], [ES-5], [ES-10], [ES-11], [ES-12], [ES-13]. Destes, três utilizaram instrumentos validados para avaliar pensamento computacional [ES-4], [ES-5], enquanto os demais inferiram o desenvolvimento de pensamento computacional com base no desempenho em tarefas de programação ou de resolução de problemas algorítmicos.

Os estudos sugeriram que *feedback* adaptativo pode apoiar o desenvolvimento de pensamento computacional quando: identifica dificuldades específicas, como erros de sintaxe, lógica ou estratégia [ES-2], [ES-12]; oferece dicas progressivas que promovem autorregulação [ES-11], [ES-12]; reproduz ansiedade e tensão, favorecendo persistência [ES-4]; e ainda fornece prática adicional adaptativa quando a frustração é detectada [ES-13].

Entretanto, a maioria dos estudos não articulou explicitamente de que forma o *feedback* adaptativo se alinha às dimensões específicas do pensamento computacional (decomposição, abstração, reconhecimento de padrões, raciocínio algorítmico, depuração).

3.6. Equidade de Gênero

A equidade de gênero foi a dimensão menos explorada do corpus. Apenas dois estudos abordaram explicitamente questões de gênero. O estudo de [Min et al. 2020], [ES-2]. No texto, é apresentado o ambiente ENGAGE, que implementou o formato *driver/navigator* e o trabalho em pares no PC, para fechar a lacuna de confiança (*confidence gap*) de meninas em Ciência da Computação.

O trabalho [Maqsood e Chiasson 2021], [ES-6], possui um design que intencionalmente submeteu personagens masculinos às mesmas vulnerabilidades e assédios online tradicionalmente associados às meninas, o que foi fundamental para o engajamento dos meninos na empatia e das meninas na superação de estigmas.

Nos demais estudos, o gênero não foi considerado como variável de design ou de avaliação. Quando presente, o tema apareceu indiretamente em escolhas narrativas inclusivas [ES-6], estratégias colaborativas [ES-2] ou preocupação com a redução da ansiedade [ES-4]. Em termos gerais, a literatura ainda não trata o gênero como variável central de design ou de avaliação.

4. Discussão

Os resultados desta revisão sistemática revelam uma tensão central entre a sofisticação computacional dos sistemas de *feedback* adaptativo e sua fundamentação pedagógica explícita. Embora 23,08% dos estudos empreguem técnicas avançadas de IA/*Machine Learning*, como redes neurais, redes bayesianas, *chatbots* com LLMs, apenas 30,77% citam explicitamente teorias clássicas de *feedback* formativo. A maioria dos estudos trata *feedback* como funcionalidade de interface ou mecânica de jogo, sem articulação clara com modelos pedagógicos estabelecidos, como o modelo de [Hattie e Timperley 2007], que distingue *feedback* sobre a tarefa (*feed back*), sobre o processo (*feed forward*) e sobre a autorregulação (*feed up*).

Em seu trabalho [Mao et al. 2024] demonstra que a eficácia do *feedback* depende do nível da tarefa: *feedback* detalhado para tarefas fáceis pode diminuir o engajamento e criar dependência da máquina, enquanto *feedback* simples para tarefas complexas pode ser insuficiente. [Chen et al. 2025] mostrou que *leaderboards* e competição podem fazer com que o *feedback* perca o efeito educacional, pois os alunos abandonam o pensamento estratégico para focar em ganhar pontos rapidamente. Esses achados reforçam a importância de fundamentar o design de *feedback* adaptativo em princípios pedagógicos sólidos, e não apenas na sofisticação técnica.

No que se refere ao pensamento computacional, 38,46% dos estudos abordam explicitamente essa habilidade, mas as métricas são fragmentadas e não padronizadas, o que dificulta as comparações entre os estudos. [Chen et al. 2025] utilizaram uma adaptação do *Bebras Challenge*, [Xu et al. 2024] empregaram um teste não especificado, e [Min et al. 2020] inferiram competência via *stealth assessment*. A ausência de métricas

consensuais limita a capacidade de avaliar o impacto real do *feedback* adaptativo no desenvolvimento do pensamento computacional.

A lacuna mais crítica identificada nesta revisão é a ausência de atenção à equidade de gênero. Apenas 7,69% dos estudos abordam o gênero diretamente como variável de design [Maqsood e Chiasson 2021], e 84,62% não mencionam o gênero em suas análises. Considerando que a computação é uma área historicamente masculinizada, com sub-representação feminina persistente, a ausência de gênero como variável de design é uma limitação estrutural do campo. [Maqsood e Chiasson 2021] demonstram que o design anti-estereótipo pode ser eficaz para engajar meninos na empatia e meninas na superação de estigmas, sugerindo que intervenções intencionais são necessárias.

Outras limitações identificadas incluem: (a) amostras pequenas (mediana de 105 participantes), (b) contextos únicos (maioria dos estudos em um único país ou instituição), (c) ausência de estudos longitudinais e (d) heterogeneidade metodológica que dificulta comparações diretas.

5. Conclusões

Os resultados desta Revisão Sistemática indicam que a integração entre *feedback* pedagógico e adaptativo em jogos sérios é uma direção promissora para a Educação em Computação, especialmente no ensino de programação e pensamento computacional. Evidências apontam que recursos baseados em inteligência artificial, como *stealth assessment*, agentes conversacionais, análise de telemetria e detecção de frustração — favorecem engajamento, persistência, personalização e desempenho acadêmico. Assim, quando associados a mecanismos de *feedback* bem estruturados, jogos sérios podem ir além da motivação, atuando como ambientes capazes de acompanhar trajetórias de aprendizagem e oferecer suporte oportuno.

Observou-se, contudo, que o avanço técnico nem sempre é acompanhado de fundamentação pedagógica equivalente. Em uma parte significativa dos estudos, o *feedback* é tratado como uma resposta automática ao comportamento do jogador, sem explicitação de objetivos formativos nem ancoragem em modelos clássicos de *feedback* educacional. Além disso, embora o pensamento computacional seja recorrente, há inconsistências conceituais e metodológicas em sua avaliação, o que dificulta comparações e sínteses mais robustas. Esses resultados indicam que adaptação inteligente, por si só, não garante qualidade educacional, exigindo atenção ao como, quando e para que o *feedback* é oferecido.

A lacuna mais evidente refere-se à equidade de gênero, uma dimensão ainda marginal na literatura analisada. Poucos estudos incorporam o gênero como variável de design, de personalização ou de avaliação, apesar das desigualdades historicamente presentes nas áreas tecnológicas. Assim, recomenda-se que pesquisas futuras avancem para modelos interdisciplinares que integrem inteligência artificial, teorias de *feedback* formativo e princípios de design inclusivo, além de adotarem estudos longitudinais, amostras mais diversas e instrumentos validados.

Referências

Akkaya, R., Özhan, U., e Altun, A. (2022). Experiential serious-game design for development of knowledge of object-oriented programming and computational

- thinking skills. *Computer Science Education*, 32(2):197–225.
- Calvo-Morata, A., Alonso-Fernández, C., Freire, M., Martínez-Ortiz, I., e Fernández-Manjón, B. (2026). Metodología para el diseño de juegos serios: Articoding, un juego para desarrollar el pensamiento computacional a través de la programación. *Revista de Educación a Distancia (RED)*, 26(82).
- Chen, C.-H., Shih, C.-C., e Law, V. (2025). Impact of rewards in a digital game on computational thinking: A structural equation modeling approach. *Education and Information Technologies*, 30:1–23.
- Gakingo, G. K. e Kabia, M. (2025). Games in engineering: Spreadsheet-based sudoku for teaching optimization. *Computer Applications in Engineering Education*, 33(1):e22767.
- Hattie, J. e Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1):81–112.
- Kitchenham, B. e Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Joint Report EBSE 2007-001, Keele University and Durham University, Keele, UK.
- Lee, M. J. (2020a). Auto-generated game levels increase novice programmers' engagement. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education (SIGCSE 2020)*, pages 1–7. ACM.
- Lee, M. J. (2020b). (re)engaging novice online learners with adaptive feedback in an introductory programming game. *ACM Transactions on Computing Education*, 20(4). Article 29.
- Lee, M. J. e Chiou, G. (2020). Animated hints help novices complete more levels in a game that teaches introductory programming. *Journal of Educational Computing Research*, 58(8):1523–1548.
- Liverman, R., Demmans Epp, C., Munoz, K., Patel, R., e Antle, A. N. (2025). Game-based social-emotional learning for middle school students: A mixed-methods study of brain agents. *International Journal of Child-Computer Interaction*, 43:100651.
- Lu, W., Boulrier, K., Mott, B., Rowe, J., Azevedo, R., e Lester, J. (2025). From in-game behaviors to learning gains: Predicting middle school students' science learning with static bayesian networks. *Proceedings of the ACM on Human-Computer Interaction (PACM HCI)*, 9(CHI PLAY).
- Mao, P., Cai, Z., He, J., e Chen, X. (2024). The effects of dynamic and static feedback under tasks with different difficulty levels in digital game-based learning. *The Internet and Higher Education*, 60:100929.
- Maqsood, S. e Chiasson, S. (2021). Cybersecurity, privacy, and digital literacy game: A design participatory study with children and teachers. In *Proceedings of the 20th Annual International Conference on Interaction Design and Children (IDC '21)*, pages 275–287. ACM.
- Min, W., Frankosky, M. H., Mott, B. W., Rowe, J. P., Smith, A., Wiebe, E., Boyer, K. E., e Lester, J. C. (2020). Deepstealth: Game-based learning stealth assessment with deep neural networks. *IEEE Transactions on Learning Technologies*, 13(2):312–325.

- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., e Moher, D. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372:n71.
- Parthasarathy, S. e Joshi, A. (2026). Accessquest: A game-based approach to teach web accessibility. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024)*, pages 1–7. ACM.
- Varajão, J., Magalhães, L., Silva, H., e Pinto, F. (2026). Game-based learning for information systems agile development: Design and evaluation of a course. *Information Systems Education Journal*, 24(1):4–18.
- Xu, Y., Huang, Y.-M., Shadiev, R., e Huang, Y.-T. (2024). The impact of a digital game-based ai chatbot on college students' academic performance and higher-order thinking skills. *Applied Sciences*, 14(21):11135.