



UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM
INSTITUTO DE COMPUTAÇÃO - ICOMP
PROGRAMA PÓS-GRADUAÇÃO EM INFORMÁTICA - PPGI

Elicitação de métricas para avaliação de aspectos de aprendizagem em jogos educacionais

Douglas Silva de Melo

Manaus - AM

Março de 2026

Douglas Silva de Melo

Elicitação de métricas para avaliação de aspectos de aprendizagem em jogos educacionais

Dissertação de mestrado submetida à avaliação, como requisito, para a obtenção do título de Mestre em Informática no Programa de Pós-Graduação em Informática, Instituto de Computação da Universidade Federal do Amazonas.

Orientador

David Braga Fernandes de Oliveira, Dr.

Coorientadoras

Marcela Sávia Picanço Pessoa Bastos, Dra.

Fernanda Gabriela de Souza Pires, Dra.

Universidade Federal do Amazonas - UFAM

Instituto de Computação - IComp

Manaus - AM

Março de 2026

Ficha Catalográfica

Elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

-
- M528e Melo, Douglas Silva de
 Elicitação de métricas para avaliação de aspectos de
 aprendizagem em jogos educacionais / Douglas Silva de Melo. -
 2026.
 154 f. : il., color. ; 31 cm.
- Orientador(a): David Braga Fernandes de Oliveira.
 Coorientador(a): Marcela Sávia Picanço Pessoa Bastos.
 Coorientador(a): Fernanda Gabriela de Souza.
 Dissertação (mestrado) - Universidade Federal do Amazonas,
 Programa de Pós-Graduação em Informática, Manaus, 2026.
1. avaliação de aprendizagem. 2. game learning analytics. 3.
 dados de estudantes. 4. learning analytics. 5. métricas. I. Oliveira,
 David Braga Fernandes de. II. Bastos, Marcela Sávia Picanço
 Pessoa. III. Souza, Fernanda Gabriela de. IV. Universidade Federal
 do Amazonas. Programa de Pós-Graduação em Informática. V.
 Título
-



Ministério da Educação
Universidade Federal do Amazonas
Coordenação do Programa de Pós-Graduação em Informática

FOLHA DE APROVAÇÃO

"ELICITAÇÃO DE MÉTRICAS PARA AVALIAÇÃO DE ASPECTOS DE APRENDIZAGEM EM JOGOS EDUCACIONAIS"

DOUGLAS SILVA DE MELO

Dissertação de Mestrado defendida e aprovada pela banca examinadora constituída pelos Professores:

Prof. Dr. David Braga Fernandes de Oliveira - **Presidente**

Prof. Dr. Marcelo da Silva Hounsell - **Membro Externo**

Prof. Dr. Luiz Antonio Lima Rodrigues - **Membro Externo**

Manaus, 26 de Março de 2026.



Documento assinado eletronicamente por **Marcelo da Silva Hounsell, Usuário Externo**, em 09/04/2026, às 09:58, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Luiz Antonio Lima Rodrigues, Usuário Externo**, em 10/04/2026, às 06:30, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **David Braga Fernandes de Oliveira, Professor do Magistério Superior**, em 10/04/2026, às 08:11, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site

[https://sei.ufam.edu.br/sei/controlador_externo.php?](https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0)

[acao=documento_conferir&id_orgao_acesso_externo=0](https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3105786** e o código CRC **98B936D3**.

Avenida General Rodrigo Octávio, 6200 - Bairro Coroados I Campus Universitário Senador Arthur Virgílio Filho, Setor Norte - Telefone: (92) 3305-1181 / Ramal 1193
CEP 69080-900, Manaus/AM, coordenadorppgi@icomp.ufam.edu.br

Referência: Processo nº 23105.008538/2026-37

SEI nº 3105786

Dedico este trabalho a todos aqueles que acreditam na educação e na aprendizagem como caminhos para transformar realidades.

AGRADECIMENTOS

Primeiramente, agradeço a Deus por ter me sustentado e guiado durante toda essa caminhada.

Agradeço ao meu orientador, professor David Fernandes, por todo o apoio durante este período e por ter contribuído para que essa jornada fosse um tempo de aprendizado e crescimento, trazendo leveza ao processo.

Agradeço também às minhas coorientadoras, professora Marcela Pessoa, que sempre me incentivou a ser alguém melhor e foi uma referência acadêmica para mim, especialmente na forma de dialogar e resolver situações. Agradeço, de forma muito especial, à professora Fernanda, que esteve presente em muitos momentos, contribuindo não apenas com a pesquisa, mas também respondendo mensagens até de madrugada, seja sobre questões acadêmicas, projetos ou até assuntos pessoais. Mais do que coorientadoras, foram amigas e colegas de profissão que caminham comigo desde a graduação.

Gostaria de agradecer de forma muito especial à minha esposa Larissa Pessoa, que esteve ao meu lado em todos os momentos, me lembrando sempre que sou capaz e nunca duvidando do meu potencial, me apoiando nos momentos difíceis da vida e da vida acadêmica, sendo um verdadeiro refúgio de paz e força.

Aos meus pais, Domingos Rodrigues e Iramaia Dias, minha profunda gratidão por sempre me incentivarem e por acreditarem na educação como uma verdadeira fonte transformadora.

Aos meus amigos que fizeram parte dessa caminhada e sempre me incentivaram, em especial Rafaela Melo, João Bernardo, Carlos Duarte, Fabrizio Honda, Jenniffer Macena, Luiz Fábio e tantos outros que estiveram comigo ao longo da vida e da vida

acadêmica. Obrigado por cada apoio e cada incentivo.

Agradeço também ao laboratório ThinkTEd e a todos os seus membros, que continuaram firmes mesmo nos momentos em que eu já não pude me dedicar ao grupo como gostaria.

Agradeço a todos os professores que fizeram parte da minha formação. Aos membros da banca, Marcelo Hounsell e Luiz Rodrigues, que me desafiaram a pensar além do óbvio e contribuíram para o aprimoramento desta pesquisa.

"Seja você mesmo, mas não seja sempre o mesmo."

Elicitação de métricas para avaliação de aspectos de aprendizagem em jogos educacionais

Autor: Douglas Silva de Melo

Orientador: David Braga Fernandes de Oliveira, Dr.

Coorientadora: Marcela Sávia Picanço Pessoa Bastos, Dra.

Fernanda Gabriela de Souza Pires, Dra.

Resumo

Jogos educacionais são utilizados para apoiar processos de ensino e aprendizagem por meio de experiências interativas, mas ainda é desafiador avaliar, de forma sistemática, o que os estudantes mobilizam durante o *gameplay*. Grande parte dos estudos da literatura realiza avaliações predominantemente subjetivas, baseadas em questionários, e frequentemente desconsidera os dados de interação registrados durante o uso do jogo. Esta dissertação propõe uma abordagem para identificar indícios de aprendizagem a partir de *logs* de interação de estudantes em jogos educacionais, visando verificar se o jogo atinge os objetivos educacionais para os quais foi desenvolvido.

Para isso, foram definidas nove métricas relacionadas a aspectos cognitivos, motores, de engajamento e ao conteúdo curricular, envolvendo habilidades como resolução de problemas, adaptação de estratégias, controle motor e participação nas atividades propostas. As métricas foram concebidas a partir (i) da seleção de aspectos de aprendizagem relevantes ao contexto de jogos educacionais e (ii) da operacionalização desses aspectos em indicadores mensuráveis com base em eventos do *gameplay*, o que demandou adaptações na estrutura de dados e na instrumentação de jogos via *GLBoard*. A pesquisa foi

conduzida por meio da metodologia *Design-Based Research* (DBR), combinando análises qualitativas e quantitativas exploratórias em contexto educacional real.

Dois experimentos foram realizados: o primeiro com o jogo "Robô Euroi", envolvendo cinco estudantes de graduação, e o segundo com o jogo "My Name", com quinze estudantes do curso de Licenciatura em Computação. As métricas foram aplicadas aos dados de interação conforme a compatibilidade com o *design* de cada jogo, e os resultados indicam coerência entre os valores mensurados e os registros de interação, fornecendo indícios relevantes para a avaliação da aprendizagem. Além disso, a comparação entre métricas e autoavaliações revelou discrepâncias entre percepções subjetivas e indicadores baseados em dados. Por fim, a avaliação por professores, analistas de dados e agentes inteligentes com enfoque psicométrico indicou aceitação conceitual geral das métricas e apontou limitações e ajustes necessários.

Como contribuição, a dissertação sistematiza um conjunto de métricas e uma estrutura de instrumentação baseada em *logs*, ampliando a base de indícios disponíveis para a análise da aprendizagem em jogos educacionais e apoiando processos de avaliação baseados em dados.

Palavras-chave: avaliação de aprendizagem, *game learning analytics* e dados de estudantes.

Elicitação de métricas para avaliação de aspectos de aprendizagem em jogos educacionais

Autor: Douglas Silva de Melo

Orientador: David Braga Fernandes de Oliveira, Dr.

Coorientadora: Marcela Sávia Picanço Pessoa Bastos, Dra.

Fernanda Gabriela de Souza Pires, Dra.

Abstract

Educational games are used to support teaching and learning processes through interactive experiences; however, it remains challenging to systematically assess what students mobilize during *gameplay*. Much of the literature conducts predominantly subjective evaluations based on questionnaires and often disregards interaction data recorded during gameplay. This dissertation proposes an approach to identify indicators of learning from students' interaction *logs* in educational games, aiming to verify whether the game achieves the educational objectives for which it was designed.

To this end, nine metrics were defined, related to cognitive, motor, engagement, and curricular content aspects, involving skills such as problem solving, strategy adaptation, motor control, and participation in proposed activities. The metrics were conceived based on (i) the selection of learning aspects relevant to the context of educational games and (ii) the operationalization of these aspects into measurable indicators derived from *gameplay* events, which required adaptations to the data structure and the instrumentation of games through *GLBoard*. The research was conducted using the *Design-Based Research* (DBR) methodology, combining exploratory qualitative and quantitative analyses in a real educational context.

Two experiments were conducted: the first with the game "Robô Euroi", involving five undergraduate students, and the second with the game "My Name", involving fifteen students from a Computer Science Teacher Education program. The metrics were applied to interaction data according to compatibility with the *design* of each game, and the results indicate coherence between measured values and recorded interaction patterns, providing relevant indicators for learning assessment. In addition, the comparison between metrics and self-assessments revealed discrepancies between subjective perceptions and data-based indicators. Finally, the evaluation conducted by teachers, data analysts, and intelligent agents with a psychometric perspective indicated overall conceptual acceptance of the metrics and identified limitations and necessary adjustments.

As a contribution, this dissertation systematizes a set of metrics and a *log*-based instrumentation framework, expanding the available evidence base for learning analysis in educational games and supporting data-driven assessment processes.

Keywords: learning assessment, *game learning analytics*, and student data.

LISTA DE ILUSTRAÇÕES

Figura 1 – Diagrama da abordagem metodológica DBR utilizada na pesquisa. . .	28
Figura 2 – Arquitetura de um sistema de GLA, baseada em Freire et al. 2016. . .	42
Figura 3 – Visão geral de um sistema <i>Game Learning Analytics</i> (FREIRE et al., 2016). .	43
Figura 4 – Estrutura do sistema GLBoard.	55
Figura 5 – Estrutura de dados do sistema GLBoard.	57
Figura 6 – Exemplo de integração do jogo com o Pacote.	61
Figura 7 – Tela de cadastro no sistema <i>web</i>	63
Figura 8 – Tela de listagem de jogos associados no sistema <i>web</i>	64
Figura 9 – Tela de listagem de jogadores.	64
Figura 10 – <i>Dashboard</i> de análises gerais de um jogo.	65
Figura 11 – Processo de desenvolvimento, aplicação e ajustes das métricas. . . .	68
Figura 12 – Métricas propostas, aspectos de aprendizagem e dados utilizados. .	79
Figura 13 – Estrutura de dados remodelada do sistema GLBoard.	98
Figura 14 – Fase II do jogo Robô Euroi.	111
Figura 15 – Gráficos por métricas em cada fase do jogo "Robô Euroi".	115
Figura 16 – Fase I do jogo <i>MyName</i>	121
Figura 17 – Gráficos por métricas em cada fase do jogo. <i>MyName</i>	124
Figura 18 – Gráficos de desempenho do Estudante 1 no jogo <i>MyName</i>	128
Figura 19 – Gráficos de desempenho do Estudante 9 no jogo <i>MyName</i>	128
Figura 20 – Fluxo de geral do desenvolvedor no sistema GLBoard.	163

LISTA DE TABELAS

Tabela 1 – Alguns benefícios da utilização de jogos educacionais para a aprendizagem.	34
Tabela 2 – Descrição dos campos da classe <i>Player Data</i>	58
Tabela 3 – Descrição dos campos da classe <i>Game Data</i>	59
Tabela 4 – Descrição dos campos da classe <i>Phase Data</i>	59
Tabela 5 – Descrição dos campos da classe <i>Session Data</i>	60
Tabela 6 – Tabela de dados comuns a jogos educacionais.	78
Tabela 7 – Perguntas do questionário para autoavaliação dos estudantes sobre os aspectos de aprendizagem.	99
Tabela 8 – Perguntas relacionadas as métricas (Analistas de dados e Psicométricos).	103
Tabela 9 – Perguntas relacionadas às métricas (Professores).	103
Tabela 10 – Descrição das telas do jogo Robô Euroi.	111
Tabela 11 – Valores únicos e média das respostas dos estudantes em cada aspecto de aprendizagem no jogo "Robô Euroi".	117
Tabela 12 – Descrição das telas do jogo <i>MyName</i>	120
Tabela 13 – Valores únicos e média das respostas dos estudantes em cada aspecto de aprendizagem no jogo <i>MyName</i>	126
Tabela 14 – Resultados do coeficiente Aiken's V para as respostas do questionário aplicado aos professores.	129
Tabela 15 – Resultados do coeficiente Aiken's V para as respostas do questionário aplicado aos analistas de dados.	134
Tabela 16 – Resultados das respostas do questionário aplicado aos psicométricos.	138

LISTA DE ABREVIATURAS E SIGLAS

API *Application Programming Interface*

DBR *Design-Based Research*

EDM *Educational Data Mining*

FPS *First-Person Shooter*

GA *Game Analytics*

GLA *Game Learning Analytics*

JAC *Coeficiente de Similaridade de Jaccard*

JSON *JavaScript Object Notation*

LA *Learning Analytics*

LPA *Latent Profile Analysis*

MB *é uma unidade de medida de informação que equivale a 10^6 bytes*

ML *Machine Learning*

MOODLE *Modular Object-Oriented Dynamic Learning Environment*

PeTRA *Performance Tracing Report Assistant*

STEM *Science Technology Engineering and Mathematics*

UEA *Universidade do Estado do Amazonas*

UFAM *Universidade Federal do Amazonas*

UUID *Universally Unique Identifier*

VR *Virtual Reality*

SUMÁRIO

1	INTRODUÇÃO	22
1.1	Objetivos	26
1.1.1	Objetivo Geral	27
1.1.2	Objetivos Específicos:	27
1.2	Metodologia	27
1.3	Organização do Trabalho	30
2	REFERENCIAL TEÓRICO	32
2.1	Jogos Educacionais	32
2.1.1	Aspectos de Aprendizagem	34
2.1.2	Avaliação de Aprendizagem	36
2.2	<i>Game Analytics</i>	38
2.3	<i>Learning Analytics</i>	39
2.4	<i>Game Learning Analytics</i>	40
2.5	Trabalhos relacionados	44
2.5.1	Avaliação da aprendizagem por instrumentos externos ao jogo	45
2.5.2	Abordagens baseadas em dados de interação para avaliação da aprendizagem	47
3	GLBOARD: SISTEMA DE CAPTURA E ANÁLISE DE DADOS EM JOGOS EDUCACIONAIS	53
3.1	<i>GLBoard</i> - Arquitetura de captura de dados	57
3.1.1	<i>Player Data</i>	58
3.1.2	<i>Game Data</i>	58
3.1.3	<i>Phase Data</i>	59

3.1.4	<i>Session Data</i>	59
3.1.5	<i>GLBoard - Pacote Unity</i>	60
3.1.6	<i>GLBoard - API</i>	62
3.1.7	<i>GLBoard - Web</i>	63
4	PROJETO E IMPLEMENTAÇÃO DA PROPOSTA	67
4.1	Seleção de aspectos de aprendizagem	68
4.2	Seleção de dados disponíveis em jogos educacionais	69
4.3	Desenvolvimento das métricas e integração aos jogos	69
4.4	Teste e aplicação das métricas (Etapa III)	70
4.5	Refinamento e validação das métricas (Etapa IV)	71
5	PROPOSTA	72
5.1	Aspectos de aprendizagem selecionados	72
5.1.1	Engajamento	73
5.1.2	Desenvolvimento cognitivo	74
5.1.3	Desenvolvimento motor	75
5.1.4	Unidade curricular	76
5.2	Seleção de dados comuns a jogos educacionais	77
5.3	Métricas propostas	78
5.3.1	Métricas para mensurar desenvolvimento cognitivo	80
5.3.1.1	Assertividade	80
5.3.1.2	Flexibilidade cognitiva	84
5.3.1.3	Habilidade lógica	87
5.3.2	Métricas para mensurar engajamento	89
5.3.2.1	Exploração	89
5.3.2.2	Tempo de <i>Gameplay</i>	92
5.3.2.3	Tentativas até o sucesso	93
5.3.3	Métricas para mensurar desenvolvimento motor	94
5.3.3.1	Habilidade mecânica	95
5.3.4	Métricas para mensurar unidade curricular	96

5.3.4.1	Pró-eficiência curricular	96
5.4	Adaptação da arquitetura de dados do GLBoard	97
5.5	Questionário de autoavaliação dos estudantes	98
5.6	Processo de validação das métricas	100
5.6.1	Procedimento de validação das respostas dos avaliadores	102
5.6.2	Procedimento de validação das respostas dos avaliadores	105
6	RESULTADOS E DISCUSSÕES	108
6.1	Experimento I	109
6.1.1	Robô Euroi	109
6.1.2	Mapeamento de dados para integração das métricas	110
6.1.3	Análise de desempenho dos jogadores nas métricas	114
6.1.4	Análise comparativa entre o desempenho nas métricas e o ques- tionário autoavaliativo	117
6.2	Experimento II	118
6.2.1	<i>MyName</i>	118
6.2.2	Mapeamento de dados para integração das métricas	121
6.2.3	Análise de desempenho dos jogadores nas métricas	124
6.2.4	Análise comparativa entre o desempenho nas métricas e o ques- tionário autoavaliativo	126
6.2.5	Validação com especialistas	128
6.2.6	Grupo professores	129
6.2.7	Grupo analistas os dados	133
6.2.8	Grupo psicométricos - Agentes Inteligentes	138
6.3	Discussão	141
6.3.1	Como avaliar aspectos de aprendizagem a partir da interação de estudantes com jogos educacionais? (QP1)	142
6.3.1.1	Validação das métricas de aprendizagem com professores, analistas e agen- tes inteligentes	143

6.3.2	Como a autoavaliação de desempenho dos estudantes se com- para aos dados de interação coletados em jogos educacionais? (QP2)	144
7	CONSIDERAÇÕES FINAIS	147
7.1	Limitações	149
7.2	Contribuições e Trabalhos Futuros	150
	Referências	152
APÊNDICE A	APÊNDICES	162
ANEXO A	ANEXOS	163

1

INTRODUÇÃO

Jogos digitais têm ganhado destaque não apenas como forma de entretenimento, mas também como ferramentas de apoio ao ensino. Sua popularização, impulsionada pelo amplo acesso a dispositivos como *smartphones*, computadores e *videogames*, tem ampliado o interesse por suas aplicações em contextos educacionais ([ENGELSTÄTTER; WARD, 2022](#)). O alto nível de engajamento proporcionado por esses ambientes interativos desperta o interesse de educadores e pesquisadores, uma vez que manter o estudante motivado é um dos principais desafios da prática pedagógica ([LONGHI et al., 2021](#)).

Nesse cenário, os jogos educacionais vêm sendo explorados como estratégias capazes de integrar motivação e aprendizagem. A combinação de mecânicas de jogo com conteúdos curriculares têm potencial para desenvolver competências cognitivas, sociais e emocionais, além de promover o aprendizado de diferentes disciplinas ([PLASS; MAYER; HOMER, 2020](#)). Jogos bem planejados podem se adaptar ao ritmo do estudante, oferecer *feedback* imediato, estimular a resolução de problemas e favorecer a aprendizagem significativa por meio de experiências interativas e práticas ([ASMAWATI, 2023](#)).

Apesar desse potencial, a adoção de jogos educacionais como ferramenta pedagógica enfrenta limitações importantes, especialmente no que diz respeito à avaliação da aprendizagem. Grande parte dos jogos disponíveis carece de mecanismos capazes de fornecer evidências confiáveis sobre o progresso dos estudantes ([ARCAGÖK, 2021](#)). Além disso, os métodos de avaliação tradicionalmente utilizados, como questionários

e testes aplicados antes e depois do jogo, tendem a capturar apenas parcialmente os efeitos da experiência de aprendizagem, pois dependem da percepção dos participantes e não consideram, de forma direta, as interações ocorridas durante a experiência, o que limita a compreensão dos processos cognitivos envolvidos (TAHIR; WANG, 2017; PETRI; WANGENHEIM; BORGATTO, 2019).

Nesse contexto, torna-se relevante explorar formas alternativas de avaliação baseadas nos dados gerados durante a interação dos estudantes com o jogo. A área de pesquisa conhecida como *Game Learning Analytics* (GLA) propõe o uso de técnicas analíticas e computacionais para coletar, processar e interpretar esses dados, com o objetivo de extrair indícios sobre o comportamento e a aprendizagem dos jogadores (FREIRE et al., 2016; HOOSHYAR; YOUSEFI; LIM, 2019). Essa abordagem permite não apenas acompanhar o desempenho em tempo real, mas também identificar estratégias de resolução adotadas, dificuldades enfrentadas e padrões de aprendizagem emergentes.

Contudo, a maior parte das soluções encontradas na literatura é construída de forma *ad hoc*, dificultando a generalização e a reutilização em diferentes contextos (ALONSO-FERNANDEZ et al., 2017). A ausência de modelos padronizados para a coleta e análise de dados ainda representa um obstáculo para a consolidação dos jogos como ferramentas de avaliação em larga escala (LI; CHEN; DENG, 2024b). Assim, propõe-se, nesta pesquisa, o desenvolvimento de métricas baseadas em dados de interação, aplicáveis a jogos educacionais que registrem os eventos necessários para seu cálculo, visando apoiar a avaliação de aspectos da aprendizagem de forma sistemática e complementar aos instrumentos tradicionais, ampliando as evidências consideradas no processo avaliativo.

A integração de jogos digitais às práticas de avaliação escolar ainda representa um desafio para os educadores. Um dos principais entraves é a baixa transparência das interações realizadas pelos estudantes durante o jogo, o que dificulta o acompanhamento do desempenho e a avaliação precisa dos conhecimentos adquiridos (GRIER; LINDT; MILLER, 2021). Avaliações detalhadas dessas interações são essenciais não apenas para monitorar o progresso dos estudantes nas habilidades desenvolvidas, mas

também para gerar *insights* sobre seus comportamentos e preferências, possibilitando ajustes no *game design* que favoreçam os objetivos educacionais (FREIRE et al., 2016).

Essa limitação está frequentemente associada à predominância de métodos tradicionais de avaliação, baseados em heurísticas e questionários aplicados antes e após a experiência de jogo (TAHIR; WANG, 2017; WALSH; BOKHOVE, 2021). Tais instrumentos costumam focar na percepção dos estudantes acerca de aspectos como usabilidade e *feedback*, como é o caso do questionário Meega+ (PETRI; WANGENHEIM; BORGATTO, 2019), priorizando a avaliação da experiência com o jogo e não da aprendizagem em si. Além disso, essas abordagens apresentam limitações importantes: por serem subjetivas e não estarem diretamente vinculadas às ações realizadas durante o jogo, deixam de captar dados valiosos sobre o processo de aprendizagem. Muitos questionários, inclusive, concentram-se em aspectos de entretenimento, o que restringe a compreensão do impacto educacional real promovido pelos jogos (WALLNER; KRIGLSTEIN, 2013).

Dado esse cenário, têm-se ampliado as investigações que buscam incorporar dados das interações dos estudantes com jogos educacionais como base para a avaliação da aprendizagem. Uma alternativa discutida na literatura envolve o uso dos dados gerados durante a utilização do jogo para mapear a trajetória do jogador e compreender com maior precisão suas ações e decisões ao longo da experiência (HOOSHYAR; YOUSEFI; LIM, 2019). Embora ferramentas como *Google Analytics*, *Firebase Analytics* e *Unity Analytics* sejam amplamente utilizadas no mercado de jogos comerciais, oferecendo suporte à coleta e análise de dados, seus focos estão em métricas como engajamento e retenção. Em jogos educacionais, por outro lado, os dados devem refletir os objetivos pedagógicos, considerando a integração entre as mecânicas do jogo e os conteúdos propostos (GUI et al., 2023).

Nesse contexto, destaca-se a área de pesquisa *Game Learning Analytics* (GLA), proposta por Freire et al. (2016), que investiga como a análise de dados pode contribuir para o processo de aprendizagem e para o aprimoramento do *game design*. Essa abordagem apoia-se em áreas como *Educational Data Mining* (EDM) (ROMERO; VENTURA, 2010), *Learning Analytics* (LA) (SIEMENS, 2013), *Machine Learning* e *Big Data Analytics*. Os dados são organizados em *dashboards* que auxiliam educadores e pesquisadores na

interpretação dos padrões de comportamento registrados. A partir dessas análises, é possível observar como elementos do jogo, como níveis de dificuldade, usabilidade ou escolhas estratégicas influenciam os resultados de aprendizagem (HAUGE et al., 2014). Trabalhos como os de Melo et al. 2020, Lima et al. 2017 e Calvo et al. (2025) ilustram essa abordagem ao analisarem o impacto do *game design* com base em dados quantitativos extraídos das interações dos estudantes.

Apesar do avanço da GLA como campo de investigação, ainda observam-se lacunas no que diz respeito à sua aplicação em avaliações somativas. Em muitos casos, as análises realizadas concentram-se em aspectos descritivos ou exploratórios, sem estabelecer modelos avaliativos voltados diretamente à aprendizagem. Embora existam contribuições, como a de Souza et al. (2021), que utilizam dados de interação para realizar uma avaliação somativa, essas soluções costumam ser desenvolvidas especificamente para o jogo analisado, sem generalização para outros contextos.

A preferência por soluções *ad hoc* dificulta a padronização dos processos de coleta e análise de dados em jogos educacionais (ALONSO-FERNANDEZ et al., 2017). A diversidade de estruturas e elementos presentes nesses jogos torna complexa a definição de um modelo unificado, o que resulta no desenvolvimento de abordagens específicas para cada contexto. No entanto, essa fragmentação limita a aplicação de melhorias consistentes e a comparação entre diferentes jogos. A adoção de práticas padronizadas pode beneficiar a comunicação entre pesquisadores e desenvolvedores, além de reduzir custos por meio da reutilização de metodologias e ferramentas. Tal sistematização contribuiria para o aprimoramento do *game design* e dos processos de avaliação associados ao uso dos jogos.

A escassez de métodos baseados em dados, tanto específicos (*ad hoc*) quanto genéricos, voltados à avaliação da aprendizagem em jogos educacionais, representa um entrave para sua validação como Objetos de Aprendizagem em larga escala (LI; CHEN; DENG, 2024b). Nesse cenário, propõe-se o desenvolvimento de métodos baseados na análise das interações dos estudantes com os jogos, concebidos para aplicação em diferentes contextos educacionais compatíveis. A análise dessas interações permite: i) acompanhar o progresso dos estudantes ao longo da experiência; ii) identificar e men-

surar aspectos da aprendizagem desenvolvidos, subsidiando avaliações por parte dos professores; iii) detectar dificuldades específicas, apontando habilidades que requerem reforço; e iv) realizar ajustes no jogo e no processo pedagógico com base nos dados observados. Além disso, esses dados possibilitam a comparação entre a percepção dos estudantes sobre seu desempenho e os registros coletados, revelando possíveis discrepâncias e fornecendo subsídios para intervenções mais alinhadas às necessidades identificadas.

Esses métodos baseiam-se na exploração de dados de interação, com o objetivo de extrair padrões e relações que permitem identificar indícios de aprendizagem. A partir dessa exploração, são construídas métricas.

Nesta pesquisa, as métricas propostas têm como finalidade fornecer indícios sobre o desempenho e o progresso dos estudantes em determinados aspectos da aprendizagem, com base nas interações registradas durante os jogos. As métricas foram projetadas para serem aplicáveis a diferentes jogos, respeitando suas estruturas e objetivos. Além disso, uma avaliação de suas propriedades conceituais foi conduzida com especialistas, a fim de verificar a clareza e a pertinência dos aspectos considerados. Adicionalmente, o estudo busca compreender a relação entre os resultados obtidos por meio das métricas e as autoavaliações realizadas pelos estudantes, evidenciando possíveis divergências. As seguintes questões de pesquisa orientaram este trabalho:

QP1: Como avaliar aspectos de aprendizagem a partir da interação de estudantes com jogos educacionais?

QP2: Como a autoavaliação de desempenho dos estudantes se compara ao uso de dados de interação coletados em jogos educacionais?

1.1 Objetivos

Esta seção apresenta o objetivo geral e os objetivos específicos que orientaram esta dissertação, a fim de responder as questões de pesquisa.

1.1.1 Objetivo Geral

Propõe-se a criação de métricas baseadas nas interações dos estudantes com jogos educacionais, de modo que os dados dessas interações permitam gerar indícios para a avaliação do desempenho dos jogadores em aspectos específicos da aprendizagem, aplicáveis a diferentes contextos de jogos educacionais.

1.1.2 Objetivos Específicos:

Para alcançar o objetivo geral proposto, serão considerados os seguintes objetivos específicos:

- Propor métricas para avaliar aspectos do desempenho dos estudantes relacionados à aprendizagem em jogos educacionais;
- Reestruturar uma ferramenta para capturar dados das interações dos estudantes com jogos educacionais;
- Conduzir um estudo comparativo entre a autoavaliação dos estudantes e os dados coletados de suas interações nos jogos educacionais;
- Validar as métricas propostas por meio da análise de especialistas nas áreas de educação, análise de dados e psicometria, utilizando métodos tradicionais e agentes inteligentes.

1.2 Metodologia

Para orientar esta pesquisa e alcançar os objetivos propostos, adotou-se o método *Design-Based Research* (DBR), uma abordagem metodológica flexível que integra *design* e métodos científicos para desenvolver e testar soluções educacionais eficazes em contextos reais (EASTERDAY; LEWIS; GERBER, 2014). Suas principais características incluem a combinação de métodos qualitativos e quantitativos, além da adaptabilidade a pesquisas envolvendo artefatos tecnológicos, considerando aspectos práticos e teóricos

(WANG; HANNAFIN, 2005). A seguir, são detalhadas as fases da DBR e a forma como foram aplicadas neste estudo.

A Figura 1 apresenta as etapas da abordagem DBR conforme aplicadas neste estudo. Na *Fase 1*, Definição do Problema e Análise de Contexto, foram definidos os objetivos, o problema e as questões de pesquisa descritos neste capítulo de introdução. Nessa fase, também foi realizada a catalogação dos principais aspectos de aprendizagem promovidos por jogos educacionais, que está sendo apresentada na Seção 5.1, bem como dos dados comumente gerados pelas interações nesses jogos, elencados na Seção 5.2.

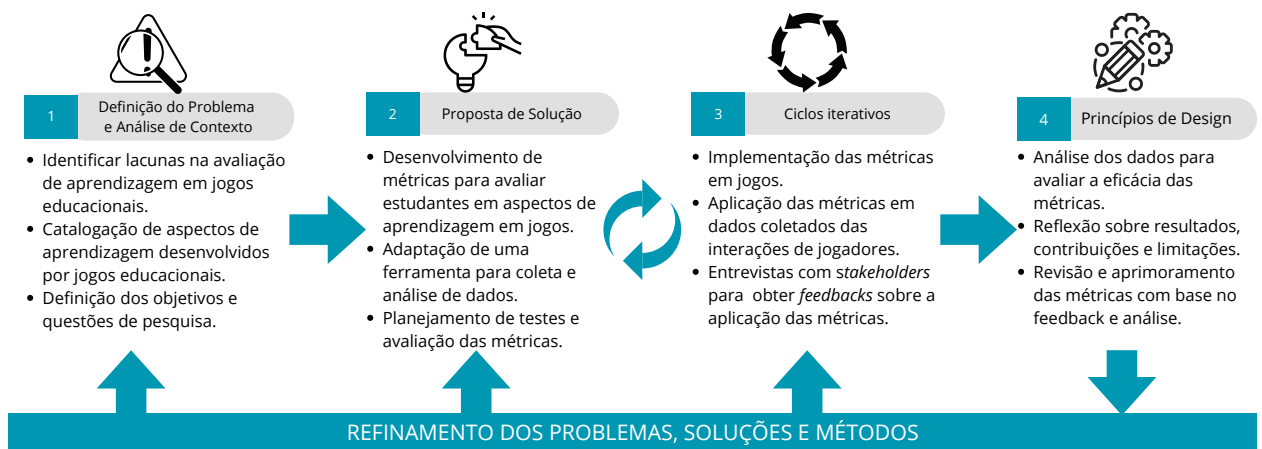


Figura 1 – Diagrama da abordagem metodológica DBR utilizada na pesquisa.

Na *Fase 2*, Proposta da Solução, foi desenvolvido o artefato central deste mes-trado. Considerando que em pesquisa anterior foi desenvolvido o *GLBoard* (SILVA et al., 2022), descrito no Capítulo 3, sistema cujo objetivo é padronizar a captura, o armazenamento e a análise de dados provenientes da interação de estudantes com jogos educacionais, observa-se que sua concepção inicial esteve voltada principalmente à extração de métricas relacionadas à avaliação do *game design* e do comportamento geral dos jogadores. Como consequência dessa orientação original, os dados e análises disponibilizados pelo sistema fornecem informações limitadas sobre aspectos específicos da aprendizagem, não contemplando, de forma sistemática, a avaliação de indícios de aprendizagem dos estudantes.

Diante dessa limitação, o objetivo desta fase foi propor um método baseado em métricas capazes de extrair indícios de aprendizagem a partir dos dados de interação, bem como adequar o *GLBoard* para suportar a captura, o processamento e a visuali-

zação desses dados. Para isso, a plataforma foi adotada como base tecnológica, sendo estendida e remodelada de forma a viabilizar a aplicação do método proposto.

Essa fase compreendeu as seguintes etapas: (i) a seleção dos aspectos de aprendizagem-alvo (Seção 5.1); (ii) a seleção e análise dos dados provenientes de jogos educacionais necessários para sua observação (Seção 5.2); (iii) o desenvolvimento das métricas a partir do cruzamento entre aspectos de aprendizagem e dados de *gameplay* (Seção 5.3), incluindo sua definição conceitual e operacionalização, bem como a análise do estado atual do *GLBoard* e a proposição de ajustes com foco na remodelagem da arquitetura de dados para viabilizar a captura, o processamento e a visualização dos dados requeridos (Seção 5.4); e (iv) o teste e a aplicação das métricas, contemplando a elaboração do questionário de autoavaliação dos estudantes (Seção 5.5) e o processo de validação, incluindo os instrumentos de coleta e os critérios de aplicação e análise (Seção 5.6).

Na *Fase 3*, a solução é aplicada em ciclos iterativos de implementação, teste, coleta e análise de dados, a partir das interações dos estudantes com jogos educacionais. Os resultados dessa fase são apresentados e discutidos no Capítulo 6, onde são analisados os dados empíricos obtidos com os jogos, a comparação com a autoavaliação dos estudantes e a validação das métricas com especialistas. Os resultados obtidos são analisados em conjunto com *stakeholders*, permitindo ajustes sucessivos nas métricas e na solução proposta. Essa fase inclui, ainda, a comparação entre os indicadores gerados pelas métricas e a autoavaliação dos estudantes, bem como a validação dessas métricas por especialistas das áreas de educação, análise de dados e psicometria.

Por fim, a *Fase 4*, descrita a partir da Seção 6.3, corresponde à análise crítica dos resultados acumulados, envolvendo a reflexão sobre as contribuições, limitações e implicações teóricas e práticas da pesquisa. Nessa etapa, são consolidados princípios de *design* e propostas de refinamento das métricas com base nos resultados obtidos e nos *feedbacks* coletados.

1.3 Organização do Trabalho

Esta dissertação de mestrado, além deste capítulo de introdução, está organizada da seguinte forma:

- **Capítulo 2 - Fundamentos Teóricos e Trabalhos Relacionados:** apresenta os conceitos teóricos e os estudos da literatura relacionados ao tema desta pesquisa, abrangendo jogos educacionais, aspectos da aprendizagem, avaliação da aprendizagem, *Learning Analytics*, *Game Analytics* e *Game Learning Analytics*. O capítulo também contempla uma revisão de trabalhos relacionados à avaliação da aprendizagem em jogos educacionais, apresentada na Seção 2.5, que aborda tanto o uso de instrumentos baseados em questionários quanto abordagens fundamentadas em dados de interação.
- **Capítulo 3 - GLBoard:** Este capítulo apresenta o sistema *GLBoard*, desenvolvido pela equipe desta pesquisa para padronizar os processos de captura, análise e visualização de dados provenientes da interação de estudantes com jogos educacionais. O sistema foi utilizado como plataforma para a implementação das métricas propostas.
- **Capítulo 4 - Desenho e Implementação da Proposta:** apresenta o desenho e a implementação da proposta desenvolvida na Fase 2 da abordagem *Design-Based Research*, detalhando a seleção dos aspectos de aprendizagem, o desenvolvimento das métricas, a remodelagem do *GLBoard* e o planejamento dos testes.
- **Capítulo 5 - Proposta:** Este capítulo detalha o método desenvolvido e as modificações requeridas no *GLBoard* para a implementação das métricas elaboradas.
- **Capítulo 6 - Resultados e Discussões:** apresenta e discute os resultados obtidos a partir da aplicação das métricas nos jogos educacionais, contemplando a análise dos dados coletados, a comparação com as autoavaliações dos estudantes e os principais achados do processo de validação.

- **Capítulo 7 - Considerações finais:** discute os principais achados da pesquisa, suas implicações teóricas e práticas, as limitações do estudo e as contribuições da proposta, bem como apresenta as considerações finais da dissertação.

2

REFERENCIAL TEÓRICO

Neste capítulo são apresentados os conceitos e fundamentos necessários para compreender o restante deste trabalho. Inicialmente, são discutidos os jogos educacionais, os aspectos de aprendizagem descritos na literatura como passíveis de desenvolvimento por meio desses jogos e as formas de avaliação da aprendizagem. Em seguida, abordam-se os conceitos de *Game Analytics*, *Learning Analytics* e *Game Learning Analytics*.

2.1 Jogos Educacionais

Jogos educacionais são uma subcategoria de jogos sérios, desenvolvidos com finalidades que vão além do entretenimento, abrangendo áreas como treinamento militar, simulações, saúde e contextos educacionais formais e informais ([BECKER, 2021](#)). O objetivo primário dos jogos sérios é promover aprendizagem estruturada, além de contribuir para a sensibilização acerca de questões sociais ou para a indução de mudanças comportamentais ([CARRION et al., 2019](#)). Embora concebidos para propósitos educativos ou sociais, a distinção entre jogos sérios e jogos cujo objetivo primário é o entretenimento nem sempre é clara, pois estes também podem promover formas indiretas de aprendizagem, ainda que tal aprendizagem não constitua sua finalidade principal.

Nesse contexto, os jogos educacionais, por sua vez, são projetados especificamente para atender a objetivos pedagógicos no contexto curricular, com foco no apoio

à aprendizagem de disciplinas como matemática, português, inglês, linguagens de programação, entre outras (CARRION et al., 2019). Todos os jogos educacionais são considerados jogos sérios, mas nem todos os jogos sérios são educacionais. Há, por exemplo, jogos sérios voltados ao treinamento técnico ou profissional em áreas específicas, como simulações cirúrgicas ou de emergência, que enfatizam o desenvolvimento de habilidades práticas (CARRION et al., 2019). Já os jogos educacionais se destacam por incorporar elementos dos jogos, como desafios, recompensas e narrativas, com o intuito de favorecer o engajamento dos estudantes e ampliar a assimilação dos conteúdos.

O diferencial dos jogos educacionais em relação a outras estratégias pedagógicas interativas reside na utilização dos elementos de jogo para criar experiências de aprendizagem imersivas. Em geral, esses jogos apresentam enredos, mecânicas e narrativas alinhadas ao conteúdo curricular, favorecendo um processo de aprendizagem mediado pelo envolvimento lúdico do jogador.

O interesse crescente pelo uso de jogos educacionais como ferramentas digitais de aprendizagem é amplamente documentado na literatura (GUEDES, 2021). Tal interesse é motivado pela capacidade desses jogos de proporcionar ambientes nos quais os estudantes aprendem por meio de experimentação, tentativa e erro, com desafios graduais e *feedback* contínuo. Além dos conteúdos curriculares, esse tipo de jogo pode contribuir para o desenvolvimento de habilidades cognitivas, motoras e sociais (REVELLE, 2013). Para isso, o *game design* deve articular adequadamente os elementos lúdicos com os objetivos pedagógicos (BLUMBERG; BLUMBERG, 2014).

Este trabalho concentra-se em jogos educacionais, uma vez que as métricas propostas foram concebidas para análise a partir dos dados de interação gerados especificamente nesse tipo de ambiente.

A Tabela 1 resume alguns dos benefícios pedagógicos dos jogos educacionais, incluindo motivação, facilitação do aprendizado, desenvolvimento cognitivo, aprendizagem por descoberta, experimentação de novas identidades e socialização (SAVI; ULBRICHT, 2008; PIRES et al., 2019; PIRES et al., 2020).

Diante das contribuições apontadas, surge a necessidade de compreender como os benefícios associados aos jogos educacionais se manifestam em evidências concretas

Tabela 1 – Alguns benefícios da utilização de jogos educacionais para a aprendizagem.

Benefícios	Descrição
Efeito motivador	Provoca interesse no jogador, que se sente desafiado e curioso (HSIAO, 2007; PLASS; MAYER; HOMER, 2020).
Facilitador da aprendizagem	Coloca o estudante no papel de tomador de decisões e permite a ele realizar quantas tentativas quiser para alcançar um objetivo (FABRICATORE, 2000; MITCHELL; SAVILL-SMITH, 2004).
Desenvolvimento de habilidades cognitivas	Promove a criação de estratégias para vencer o jogo, estimulando a resolução de problemas e o pensamento crítico (PLASS; MAYER; HOMER, 2020; BALASUBRAMANIAN; WILSON, 2006).
Aprendizagem por descoberta	Possibilita a exploração de diversos cenários, dando ao jogador a possibilidade de experimentar coisas novas sem que seja julgado caso erre (MITCHELL; SAVILL-SMITH, 2004; PLASS; MAYER; HOMER, 2020).
Experiência de novas identidades	Oferece ao estudante a possibilidade de se colocar no papel de uma nova identidade, como por exemplo um astronauta, enfrentando os desafios e dilemas da vida desses profissionais (HSIAO, 2007).
Socialização	Permite ao jogador interagir com outros, compartilhando informações do jogo, experiências que vivenciaram e dicas que podem ajudar novos jogadores (HSIAO, 2007).

de aprendizagem. Para isso, é importante examinar os chamados aspectos de aprendizagem, que funcionam como indicadores dos ganhos educacionais adquiridos pelos estudantes durante a interação com o jogo, como descrito na próxima seção.

2.1.1 Aspectos de Aprendizagem

A aprendizagem envolve um conjunto de competências e habilidades desenvolvidas ao longo do processo educacional, que contribuindo para o entendimento de como os indivíduos assimilam, processam e aplicam conhecimentos em diferentes contextos (GEE, 2003). Esses processos podem ser organizados em categorias amplas, referidas nesta pesquisa como aspectos de aprendizagem. Esses aspectos incluem:

- **Cognitiva:** refere-se ao desenvolvimento de capacidades intelectuais, incluindo pensamento crítico, resolução de problemas, memória e atenção (PHELPS, 2006). Jogos que apresentam enigmas e desafios podem favorecer o raciocínio lógico e analítico (ASMAWATI, 2023; LAW, 2016).
- **Socioemocional:** abrange habilidades como autoconsciência, controle emocional, empatia, resiliência, motivação e engajamento (PHELPS, 2006). Ambientes colaborativos e tarefas desafiadoras podem contribuir para o exercício dessas

habilidades, incentivando a persistência diante de dificuldades e a interação com outros participantes.

- **Desenvolvimento Motor:** envolve a coordenação, precisão e controle dos movimentos, desde habilidades motoras grossas (como correr ou pular) até finas (como manipular objetos pequenos) (SUTAPA et al., 2021). Jogos que exigem respostas rápidas ou navegação com coordenação visual e motora estão associados a esse aspecto (VUKIĆEVIĆ et al., 2019).
- **Metacognição:** diz respeito à capacidade de monitorar, avaliar e ajustar estratégias cognitivas (FLAVELL, 1979). Situações em que os jogadores precisam reconsiderar ações após tentativas mal-sucedidas podem estimular esse processo, contribuindo para uma aprendizagem mais autônoma.

Embora tais aspectos estejam presentes nas experiências de aprendizagem mediadas por jogos educacionais, nem sempre são considerados nos processos de avaliação educacional. Isso se deve, em parte, à dificuldade de mensurar tais habilidades com precisão (MISLEVY et al., 2016; WALSH; BOKHOVE, 2021). Além disso, ainda há resistência em reconhecer os jogos como instrumentos válidos de avaliação (ANDRADE; NEGRÃO; VILAÇA, 2021; VIDENOVİK et al., 2023).

Essa resistência pode estar relacionada a preocupações sobre a precisão, confiabilidade e objetividade dos jogos como ferramentas avaliativas. Em contextos educacionais tradicionais, a avaliação ocorre geralmente por meio de provas, testes, trabalhos escritos e observações. Nos jogos educacionais, por outro lado, o desempenho dos estudantes pode ser analisado a partir de como enfrentam desafios relacionados ao conteúdo, quais estratégias utilizam e como constroem novos conhecimentos ao longo da interação.

Os dados de interação registrados durante o jogo oferecem a possibilidade de observar esse progresso e interpretar o desempenho dos estudantes. No entanto, o rastreamento e a análise desses dados apresentam desafios significativos, que serão discutidos na próxima subseção, dedicada aos métodos existentes para avaliação da aprendizagem.

2.1.2 Avaliação de Aprendizagem

O desenvolvimento de jogos educacionais envolve diversos desafios, entre os quais se destaca a necessidade de alinhar o design do jogo aos objetivos de aprendizagem. Isso exige práticas centradas no conteúdo pedagógico, equilíbrio entre engajamento e aprendizagem e estratégias de avaliação contínua (TROYER, 2017).

A avaliação cumpre um papel central na verificação do desempenho e na personalização da experiência de aprendizagem. Ele permite monitorar o progresso dos estudantes, ajustar a dificuldade das tarefas e contribuir para a melhoria da usabilidade dos jogos (CHIOTAKI; POULOPOULOS; KARPOUZIS, 2023). Por exemplo, em um jogo de matemática, a avaliação contínua pode identificar dificuldades específicas, como a resolução de equações de segundo grau, e adaptar os desafios conforme o desempenho do estudante, oferecendo conteúdos adicionais ou ajustando sua complexidade. Além das adaptações, os dados coletados também podem auxiliar o trabalho docente. Ao acompanhar os registros de desempenho, o professor pode identificar padrões de dificuldade e ajustar suas estratégias pedagógicas em sala de aula, promovendo intervenções mais direcionadas.

A avaliação contínua também gera *feedback* sobre o uso dos jogos como ferramentas educacionais, contribuindo para uma melhor compreensão do processo de aprendizagem e do desenvolvimento de habilidades (WISNIEWSKI; ZIERER; HATTIE, 2020; SMIDERLE et al., 2020). Na literatura, os métodos mais utilizados para avaliação em jogos educacionais são apresentados a seguir, listados em ordem de recorrência (TAHIR; WANG, 2017; PETRI; WANGENHEIM, 2017; WALSH; BOKHOVE, 2021):

- i) **Pós-questionários:** aplicados após a experiência de jogo, captam percepções sobre engajamento, usabilidade e satisfação. Instrumentos como o Meega+ (PETRI; WANGENHEIM; BORGATTO, 2019) são amplamente utilizados, com foco na experiência do jogo, e não necessariamente no desempenho individual (WALSH; BOKHOVE, 2021).
- ii) **Pré e pós-testes:** verificam a aquisição de conhecimento ao comparar resultados antes e depois da interação com o jogo. Apesar de úteis na avaliação de avanços cognitivos, não consideram os dados gerados durante a interação, limitando a

compreensão do processo de aprendizagem (WILSON et al., 2016). Também não abrangem aspectos como habilidades socioemocionais ou metacognitivas.

- iii) **Coleta e análise de dados de interação:** utiliza registros gerados durante o jogo para identificar estratégias, decisões e padrões de comportamento, permitindo inferências sobre a aprendizagem.

A coleta contínua de dados permite avaliações formativas e somativas, conforme Souza et al. (2021), ampliando a compreensão sobre a trajetória de aprendizagem. Essa abordagem possibilita observar o domínio de conteúdos e o desenvolvimento de habilidades e estratégias ao longo da interação.

Apesar das potencialidades, esse tipo de avaliação ainda é menos adotado em comparação com métodos tradicionais. Entre os fatores limitantes, destacam-se a complexidade técnica para registrar e analisar dados em tempo real (FREIRE et al., 2016), os desafios na interpretação das interações devido à sua variedade e volume (GLORIA; BELLOTTI; BERTA, 2014), e a falta de padronização nos processos de coleta e análise, o que dificulta comparações entre estudos e a reutilização de abordagens propostas na literatura (ALONSO-FERNANDEZ et al., 2017).

Cada método de avaliação possui objetivos específicos, apresentando diferentes vantagens e limitações. A combinação de diferentes abordagens pode gerar análises mais abrangentes, permitindo ajustes pedagógicos e o aprimoramento dos jogos educacionais (BELLOTTI et al., 2013).

Ainda assim, a avaliação baseada em dados de interação enfrenta barreiras que limitam sua adoção em larga escala. A superação desses obstáculos requer o fortalecimento da área por meio de estudos que proponham novas soluções e ampliem a compreensão sobre o uso dos dados de interação (BUCKLEY et al., 2021; WALSH; BOKHOVE, 2021). Esse campo de investigação insere-se no escopo do *Analytics*, voltado à coleta, processamento e análise de grandes volumes de dados com o objetivo de extrair *insights* relevantes. Destacam-se, nesse contexto, as áreas: *Learning Analytics*, voltada à análise de dados em ambientes educacionais, e *Game Learning Analytics*, focada na análise das interações geradas em jogos educacionais (FREIRE et al., 2016).

As próximas seções aprofundam os conceitos e técnicas de *analytics* aplicadas a jogos educacionais, com ênfase no uso dessas abordagens para enfrentar os desafios da avaliação da aprendizagem a partir das interações dos estudantes com os jogos.

2.2 *Game Analytics*

Game Analytics (GA) refere-se à análise de dados gerados por jogadores com o intuito de aprimorar o *game design*, identificando falhas e oportunidades de melhoria na experiência e na retenção de usuários (MEDLER, 2009).

Ao contrário do *Learning Analytics*, que tem foco na aprendizagem em ambientes educacionais, o GA está voltado principalmente para interesses comerciais, como o aumento de receita e a fidelização de jogadores (LOH; SHENG; IFENTHALER, 2015). Em um mercado competitivo, compreender os padrões de comportamento dos usuários é importante para ajustes no design do jogo, de modo a maximizar o engajamento e as vendas.

As análises são aplicadas em diferentes etapas do desenvolvimento do jogo (ALONSO-FERNÁNDEZ et al., 2022; FREIRE et al., 2016). Durante a produção, o foco recai sobre aspectos técnicos do funcionamento do jogo, como detecção de *bugs*, taxa de quadros e uso de memória. Na fase de testes, passam a ser analisados elementos relacionados ao *game design* e após o lançamento, os dados gerados pela interação contínua com o jogo permitem identificar padrões de uso, ajustar a dificuldade e planejar conteúdos com base no comportamento do jogador (ECK, 2006).

A abordagem é amplamente empregada em jogos comerciais, onde a análise orienta decisões de mercado. No entanto, nos jogos educacionais, o propósito é distinto, estando voltado ao acompanhamento e à promoção da aprendizagem dos estudantes. Embora compartilhem técnicas semelhantes, as finalidades e critérios de avaliação entre GA e jogos educacionais divergem.

2.3 *Learning Analytics*

Learning Analytics (LA) é um campo emergente voltado à medição, coleta, tratamento e análise de dados educacionais com o objetivo de compreender e melhorar os processos de aprendizagem. LA busca identificar tendências nos dados gerados em ambientes virtuais de aprendizagem (SIEMENS; LONG, 2011). Nesse sentido, trata-se de uma abordagem de exploração de dados voltada à educação, com ênfase na compreensão humana (BAKER; INVENTADO, 2014), em contraste com a Mineração de Dados Educacionais, que se concentra na extração automatizada de padrões usando técnicas estatísticas e de aprendizagem de máquina (KLÖSGEN; ZYTKOW, 2002).

Siemens e Long (2011) descrevem LA como o processo de análise de dados obtidos em plataformas educacionais, como sistemas de gestão da aprendizagem (LMS) e sistemas tutoriais inteligentes. As análises visam responder a questões centrais do processo educacional, como: os cursos estão atendendo aos objetivos de aprendizagem? Os estudantes estão aprendendo? Como o sistema pode apoiar os estudantes em suas dificuldades? (GRELLER; DRACHSLER, 2012).

O crescimento de ambientes virtuais de aprendizagem, como o *MOODLE*, tem contribuído para a popularização da LA, pois essas plataformas registram uma quantidade expressiva de interações (*logs*) (CAPRARA; CAPRARA, 2022; ÇEKEN; TAŞKIN, 2022; PORTO; DIAS; BATTESTIN, 2023). Esses dados permitem, por exemplo, identificar padrões de comportamento associados a dificuldades acadêmicas, como a ausência prolongada ou o não acesso aos materiais. A partir dessas análises, os sistemas podem emitir alertas ou sugerir materiais de apoio conforme o desempenho e o engajamento do estudante.

Duas abordagens principais são utilizadas na análise dos dados: a visualização de informações e o uso de técnicas de aprendizagem de máquina. A primeira estrutura os dados coletados em gráficos e indicadores para facilitar sua interpretação, ainda que dependa fortemente da leitura subjetiva do observador (CHAUDY; CONNOLLY; HAINEY, 2014; FARIAS, 2019). Já a segunda aplica modelos preditivos, capazes de identificar tendências e oferecer diagnósticos com base nos padrões encontrados nos dados.

Apesar dos benefícios, há certa resistência por parte de professores em adotar avaliações baseadas em algoritmos. Essas análises, muitas vezes automatizadas e implementadas por sistemas com funcionamento do tipo “caixa-preta”, podem gerar desconfiança quanto à transparência dos critérios utilizados (ZHANG et al., 2023). Como os professores são os principais responsáveis por justificar os resultados das avaliações, torna-se importante que os sistemas de LA apresentem seus critérios de forma clara, a fim fortalecer a confiança no sistema e apoiar a tomada e a tomada de decisões pedagógicas.

2.4 *Game Learning Analytics*

A fusão entre *Learning Analytics* (LA) e *Game Analytics* (GA) originou a área de *Game Learning Analytics* (GLA), que combina a análise de dados educacionais com a análise de interações nos jogos digitais (FREIRE et al., 2016). Enquanto LA foca na aprendizagem e GA na experiência do jogador e no desempenho comercial dos jogos, GLA busca compreender como os elementos do jogo influenciam o processo de aprendizagem.

Essa abordagem possibilita avaliar se a mecânica do jogo está alinhada se as mecânicas do jogo estão alinhadas aos objetivos educacionais, utilizando evidências em vez de depender exclusivamente de pressupostos teóricos (ALONSO-FERNÁNDEZ et al., 2022; ALONSO-FERNÁNDEZ et al., 2023). Assim, a análise não se restringe à aplicação de questionários, mas considera dados de interação gerados durante o uso do jogo, para entender como a aprendizagem ocorre dentro do jogo (FREIRE et al., 2016; WANG et al., 2022).

Com os dados obtidos por GLA, é possível monitorar o progresso do estudante e o desempenho do próprio jogo. Caso os objetivos definidos, como o domínio de conteúdo ou o engajamento contínuo dos estudantes, não sejam atingidos, as evidências coletadas podem orientar ajustes no design ou no conteúdo do jogo (CHEUNG; NG, 2021).

A análise em GLA considera os aspectos de jogabilidade e os de aprendizagem, dado que os jogos educacionais geralmente integram elementos lúdicos e pedagógicos

(MELO et al., 2020; TLILI; CHANG, 2019). A ênfase de cada um desses elementos varia conforme o propósito do jogo: alguns priorizam a diversão, mantendo a aprendizagem em segundo plano, enquanto outros têm um foco explícito em objetivos educacionais.

Na arquitetura de GLA, as visualizações são projetadas para atender a diferentes perfis de usuários. Professores podem acessar *dashboards* com indicadores de progresso, tarefas concluídas e áreas de maior dificuldade, o que auxilia no planejamento pedagógico. Desenvolvedores visualizam dados técnicos, como tempos de resposta, erros e interações com os elementos do jogo, úteis para otimizações. Já os jogadores recebem estatísticas sobre seu próprio desempenho e conquistas, favorecendo uma experiência mais personalizada e motivadora (ALONSO-FERNANDEZ et al., 2017).

Dessa forma, GLA organiza os dados de modo que diferentes *stakeholders* recebam informações relevantes conforme suas necessidades e objetivos. Na Figura 10, apresenta-se um fluxo básico de funcionamento de um sistema de GLA, com base na arquitetura proposta por Freire et al. (2016), no qual o processamento e a circulação dos dados ocorrem da seguinte forma:

- *Design*: etapa responsável pelo planejamento do jogo, com a definição dos dados que serão extraídos para acompanhar o processo de aprendizagem dos jogadores.
- *Jogo*: fonte primária dos dados, gerados a partir das interações dos jogadores. Para sua coleta, integra-se ao jogo uma ferramenta ou *plugin* dedicado à captura dos eventos relevantes.
- *Controlador*: responsável por receber os dados, organizá-los e armazená-los em um banco de dados. Também pode realizar análises automáticas, cujos resultados são disponibilizados no *dashboard* ou enviados de volta ao jogo, oferecendo, por exemplo, *feedback* ao estudante sobre sua trajetória de aprendizagem.
- *Dashboard*: interface visual que apresenta os dados analisados por meio de gráficos e tabelas. Serve de apoio a professores e desenvolvedores e, opcionalmente, pode ser acessado pelos próprios jogadores.

De acordo com Freire et al. 2016, um sistema de GLA é composto por um conjunto de artefatos que seguem padrões comuns e geralmente devem ser implementados:

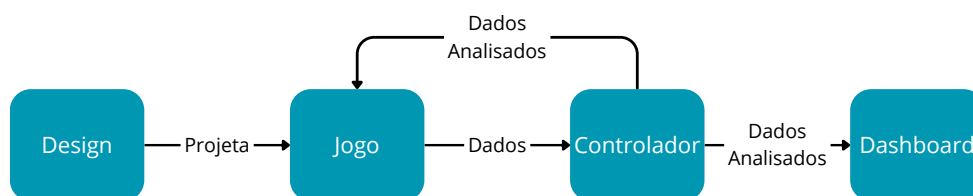


Figura 2 – Arquitetura de um sistema de GLA, baseada em Freire et al. 2016.

- **Instrumentação:** responsável pela captura e envio dos dados gerados durante a execução do jogo para um servidor remoto. Devido ao alto volume de eventos, recomenda-se o envio em lotes, reduzindo o número de requisições. Os dados devem ser armazenados localmente quando não houver conexão com a internet.
- **Coleta e persistência dos dados:** componente do lado do servidor encarregado de organizar e armazenar os dados recebidos, seguindo uma estrutura padronizada que facilite análises posteriores.
- **Análise em tempo real:** permite o monitoramento das interações dos estudantes durante a execução do jogo, especialmente em contextos síncronos com as aulas. Ao oferecer respostas imediatas, viabiliza a identificação de dificuldades e a realização de intervenções pedagógicas pontuais. As análises são simplificadas para evitar sobrecarga computacional.
- **Análise agregada em lotes:** executa análises mais complexas com base em grandes volumes de dados, relacionando múltiplas variáveis associadas aos jogadores. Por não ocorrer em tempo real, exige maior capacidade de processamento.
- **Indicadores de desempenho:** fornecem dados quantitativos relevantes para educadores, como percentuais de conclusão de fases, número de tentativas por tarefa e desempenho por etapa. Esses indicadores devem resultar de cruzamentos significativos entre os dados disponíveis, apresentando informações claras e úteis.
- **Analytics Dashboard:** interface desenvolvida, em geral, ao final do processo. Permite a visualização dos indicadores e análises geradas, com configurações adaptáveis às necessidades de diferentes perfis de usuário, como professores, desenvolvedores e jogadores.

A Figura 3 apresenta uma visão geral de um sistema de *Game Learning Analytics* (GLA) (FREIRE et al., 2016). O jogo envia os dados por meio de um componente de rastreamento para um sistema coletor, responsável por armazená-los. Em seguida, os dados são organizados e analisados por um agregador, que alimenta o relator. Esse componente gera relatórios que podem ser utilizados em tempo real ou armazenados para consultas posteriores. O avaliador interpreta os dados produzidos para verificar se os objetivos educacionais estão sendo atingidos. Por fim, o adaptador envia instruções de retorno ao jogo, possibilitando sua adaptação ao comportamento do jogador.

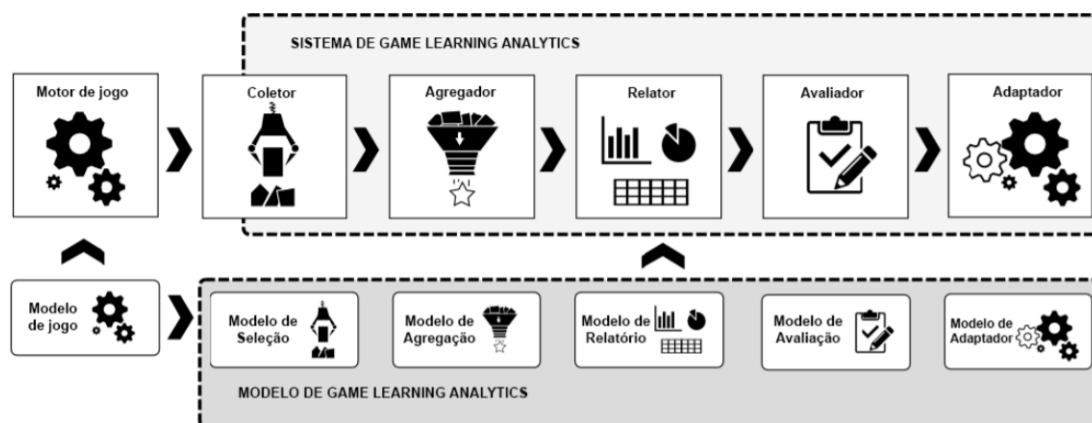


Figura 3 – Visão geral de um sistema *Game Learning Analytics* (FREIRE et al., 2016).

Essa arquitetura favorece o uso contínuo dos dados para ajustes no projeto do jogo, permitindo a validação dos objetivos educacionais, e adaptações durante a sua utilização, como o ajuste automático da dificuldade das fases.

Entretanto, Alonso et al. (2017) destacam um desafio recorrente na área de GLA: a ausência de padrões amplamente aceitos para a estruturação de sistemas. Por se tratarem, em geral, de soluções *ad hoc*, cada jogo define seu próprio conjunto de dados e métodos de análise, dificultando a criação de modelos reutilizáveis. Além disso, a documentação frequentemente limitada desses sistemas compromete a sua replicação e implementação. Como consequência, novos jogos exigem adaptações específicas, incluindo ajustes nos componentes de rastreamento, nas análises e nas visualizações, o que eleva os custos e o tempo de desenvolvimento.

As principais contribuições da sistematização dos processos de GLA incluem:

- Reutilização do componente de captura de dados em diferentes jogos, eliminando a necessidade de reimplementação por parte dos desenvolvedores;
- Estabelecimento de um padrão de comunicação entre o jogo e a plataforma analítica, permitindo que dados sejam interpretados por diferentes ferramentas compatíveis. Isso viabiliza a substituição da plataforma sem alterações no jogo;
- Utilização de ativos de análise e visualização padronizados e modulares, como *dashboards*, adaptáveis aos dados específicos de cada jogo, mas capazes de gerar informações relevantes sem configurações adicionais, desde que o padrão de dados seja seguido.

Considerando a importância da avaliação, tanto do desempenho dos estudantes quanto da eficácia dos jogos educacionais, a seção seguinte apresenta trabalhos relacionados ao tema. São abordadas pesquisas que empregam métodos tradicionais de avaliação, bem como estudos que utilizam GLA a partir de dados de interação dos jogadores.

2.5 Trabalhos relacionados

A revisão está organizada em duas abordagens principais: estudos que avaliam a aprendizagem por meio de instrumentos externos ao jogo, como questionários e testes de desempenho, e pesquisas que utilizam dados de interação para a análise do desempenho dos estudantes. Embora fatores como engajamento, motivação e diversão influenciem a experiência do usuário (PETRI; WANGENHEIM; BORGATTO, 2019; PETRI; WANGENHEIM, 2017), a aprendizagem é compreendida como a aquisição de conhecimentos ou habilidades específicas (NOEMÍ; MÁXIMO, 2014), não sendo necessariamente garantida pelo entretenimento proporcionado pelo jogo.

2.5.1 Avaliação da aprendizagem por instrumentos externos ao jogo

Esta seção inicial concentra-se em estudos que avaliam a aprendizagem com base em instrumentos aplicados fora do ambiente do jogo. A seguir, são apresentadas algumas dessas abordagens.

Fu et al. (2009) propuseram o *EGameFlow*, um questionário baseado em escala *Likert* para avaliar a satisfação de estudantes com jogos de *e-learning*. O instrumento considera dimensões como concentração, objetivos claros, *feedback*, desafio, autonomia, imersão, interação social e aprimoramento do conhecimento. Com base em 166 respostas de estudantes de um curso on-line em Taiwan, o instrumento foi validado por meio de testes de validade e análise fatorial, indicando ser útil para mensuração do engajamento.

O modelo *MEEGA+* (PETRI; WANGENHEIM; BORGATTO, 2019) é uma evolução do *MEEGA* e foi projetado para avaliar a qualidade e a eficácia de jogos educacionais no ensino de computação. O modelo organiza os objetivos educacionais em medidas específicas, avaliadas por questionário que contempla dimensões como atenção focada, diversão, desafio, interação social, confiança, relevância, satisfação, usabilidade, aprendibilidade, estética e aprendizagem percebida. A aprendizagem é analisada tanto em termos técnicos quanto no desenvolvimento de habilidades, como resolução de problemas, comunicação e liderança. O instrumento foi validado com 843 estudantes em dezoito jogos, apresentando alta confiabilidade (alfa de Cronbach $\alpha = 0,928$). Entre as limitações, destacam-se a amostragem não probabilística e a restrição ao contexto da computação, o que pode comprometer a generalização dos resultados.

Oliveira et al. (2019) propuseram uma metodologia baseada na pesquisa de Mourão (2012), voltada à avaliação de jogos educacionais por meio de questionários de perfil, reações ao jogo, e pré e pós-testes. A metodologia foi aplicada no jogo “Expedição Antártica”, voltado ao ensino interdisciplinar de ciências, matemática e educação científica entre licenciandos. O estudo envolveu quatorze alunos da disciplina “Introdução à Programação de Jogos”, divididos em dois grupos, cada um com uma versão distinta do jogo. A avaliação incluiu: i) questionário de perfil; ii) pré-teste; iii) interação com o jogo; iv) questionário de reação; v) pós-teste. Os resultados indicaram melhor desempenho

da segunda versão, com maiores índices de satisfação e aprendizagem, atribuídos à maior clareza das instruções e à estrutura mais eficiente das fases.

De forma complementar, De Oliveira et al. (2020) avaliaram o jogo educacional “Livro do Conhecimento” utilizando múltiplos critérios: usabilidade, experiência do usuário, requisitos pedagógicos e aceitação como ferramenta de apoio ao ensino de Língua Portuguesa. A análise foi fundamentada na Teoria Cognitiva da Aprendizagem Multimídia e na Teoria da Carga Cognitiva (MAYER, 2009), com testes realizados por desenvolvedores. A avaliação incluiu o modelo *E-Game Flow*, que obteve bons resultados em concentração, clareza de objetivos e *feedback*, embora tenha apontado fragilidades em desafio e imersão. Também foi utilizado o modelo MEEGA+, que apresentou boa aceitação nos quesitos aprendizagem percebida, diversão, confiança e usabilidade. Os resultados levaram ao aprimoramento do jogo, com sugestões para novas fases e versão para dispositivos móveis.

Embora as pesquisas baseadas em questionários ou metodologias específicas ofereçam contribuições relevantes para a avaliação de jogos educacionais, esses métodos geralmente capturam apenas momentos isolados após a interação do jogador com o jogo. Consequentemente, a análise acaba restrita à percepção subjetiva do estudante, sem contemplar a jornada do jogador. Essa limitação é ressaltada por Tahir et al. (2017), que demonstram que, mesmo com a diversidade de elementos avaliados, como usabilidade, aprendizagem, experiência, imersão e mecânica de jogo, a maioria das avaliações é desenvolvida de forma *ad hoc* e não se conecta diretamente ao progresso da aprendizagem.

O estudo de Tahir et al. (2017) analisou 58 artigos relacionados à avaliação em jogos educacionais e identificou dezenove abordagens distintas, sendo a maioria baseada em questionários. De forma complementar, Petri et al. (2017) revisaram 117 estudos e constataram que a aprendizagem foi o aspecto mais frequentemente avaliado, com perguntas como: “Como você avalia a aprendizagem no jogo?” e “O conteúdo curricular está claro?”. Embora úteis, essas questões tendem a gerar respostas generalistas, sem refletir o desenvolvimento de habilidades ao longo das fases do jogo. Além disso, 101 dos estudos analisados por Petri et al. utilizaram pontuações de pós-testes ou autoa-

valiações, reforçando a tendência de avaliações centradas no jogo como um todo, em detrimento de análises baseadas em interações durante a experiência. Para superar essas limitações, seria necessário adotar instrumentos avaliativos mais granulares, aplicados em diferentes momentos do jogo. No entanto, aplicar questionários durante a partida pode comprometer a continuidade da experiência, afetando negativamente o engajamento e a imersão dos jogadores.

2.5.2 Abordagens baseadas em dados de interação para avaliação da aprendizagem

Diante desse cenário, é necessário buscar alternativas que atendam às demandas avaliativas de jogos educacionais, superando as limitações dos métodos tradicionais de abordagem caixa-preta, que pouco revelam sobre os detalhes das interações dos estudantes com o jogo. Assim, esta seção foca em trabalhos que propõem abordagens quantitativas baseadas na coleta de dados de interação, como feito pelo presente estudo.

Walsh et al. (2021) investigam a eficácia dos jogos educacionais como ferramentas de avaliação, analisando o progresso dos estudantes, suas respostas e a forma como resolvem os desafios ao longo do jogo. Um dos principais desafios identificados é a dificuldade de alinhar os dados gerados pelos jogos com os modelos tradicionais de avaliação, dado o caráter variável dos jogos e o contexto específico de cada interação. Os dados coletados incluem tempo de resposta, escolhas feitas pelos jogadores e progresso em tarefas repetidas. Os autores destacam que as primeiras tentativas tendem a ser menos representativas das habilidades dos estudantes, sugerindo seu descarte para análises mais consistentes. Além disso, recomendam que o desempenho seja avaliado com base no progresso, repetição e enfrentamento de novos desafios, e não apenas em respostas corretas. O estudo também alerta que medir habilidade exclusivamente pelo tempo de resposta pode ser impreciso, pois nem sempre reflete compreensão conceitual. A principal contribuição está na proposta de novos paradigmas avaliativos, mais alinhados aos dados dos jogos educacionais. Os autores reforçam o potencial dos jogos como ferramentas inovadoras de avaliação, ressaltando a importância de

estratégias analíticas sensíveis ao contexto dos jogos e à natureza das interações dos jogadores.

Souza et al. (2021) desenvolveram o jogo educacional *Castaways*, voltado à aprendizagem de Linguagens Regulares e Autômatos Finitos Determinísticos. O jogo integra avaliações formativas e somativas por meio da coleta automatizada de dados, permitindo *feedback* em tempo real e ajustes pedagógicos ao longo do processo de aprendizagem. As métricas analisadas incluíram o tempo para resolução de desafios, número de reinícios e eficiência das soluções adotadas. Essas informações possibilitaram a identificação das principais dificuldades dos estudantes e a avaliação da efetividade pedagógica de cada parte do jogo.

Os resultados obtidos contribuíram para ajustes no conteúdo e melhorias no engajamento. De Souza et al ilustram como a coleta e análise de dados de interação podem ampliar a compreensão sobre o processo de aprendizagem em jogos, permitindo avaliações mais precisas e contextualizadas. Por exemplo, altos índices de reinícios em determinados níveis podem indicar instruções pouco claras ou conteúdo excessivamente complexo, sugerindo a necessidade de ajustes no *game design* para facilitar a compreensão dos objetivos educacionais. Por outro lado, a resolução consistente de desafios com soluções eficientes aponta que os conceitos estão sendo assimilados adequadamente pelos estudantes.

Os resultados reforçam o potencial do *Castaways* como ferramenta de engajamento, e como recurso de avaliação e adaptação pedagógica. A análise contínua dos dados permite aos educadores realizar intervenções pontuais, ajustando o conteúdo às necessidades observadas, promovendo melhorias tangíveis na experiência e nos resultados de aprendizagem.

Alonso-Fernández et al. (2023) utilizaram a biblioteca *xAPI-SG* para capturar e analisar dados de interação no jogo *Conectado*, voltado à conscientização sobre o *bullying*. A *xAPI* registra ações do jogador em formato padronizado (*JSON*), como: ator (estudante), verbo (escolheu), objeto (questão), resultado (sucesso ou falha) e resposta ('1'). Essas ações são armazenadas em servidor e utilizadas para análises de engajamento (tempo logado, cliques), desempenho (acertos, tempo de resposta) e

progressão (tentativas por fase).

A partir dos dados, variáveis analíticas foram extraídas para avaliar mudanças na percepção dos estudantes. A regressão Bayesiana apresentou alta acurácia preditiva, demonstrando o potencial da abordagem para mensurar aprendizagem em jogos sérios. Diferentemente da proposta da *xAPI-SG*, que adota uma estrutura baseada em ações genéricas para análises de aprendizagem e *game design*, a pesquisa apresentada nesta dissertação propõe métricas voltadas a elementos particulares dos jogos, com foco na avaliação direta de aspectos específicos da aprendizagem. Essa abordagem visa compreender a evolução dos estudantes e a influência do jogo nesse processo.

O estudo de Mangaroska et al. (2024) investigou o impacto de elementos de design no engajamento em jogos educacionais, utilizando o jogo *Letter Link* com 39 participantes. Foram aplicadas técnicas de *Game Learning Analytics* para coletar dados de interação dos jogadores, como pontos de experiência (XP), conquistas e número de partidas jogadas, complementadas por questionários que avaliaram a percepção dos participantes. A análise indicou que elementos como XP e conquistas influenciam positivamente o engajamento cognitivo, afetivo e o estado de fluxo. O estudo destaca a relevância do uso de dados analíticos no aprimoramento do *game design*, reforçando a importância da integração de elementos que promovam engajamento contínuo. Os resultados evidenciam o potencial das ferramentas de *GLA* para gerar *insights* que orientem o desenvolvimento de experiências de aprendizagem mais atrativas e efetivas.

Uma pesquisa relevante para a área de *GLA*, fortemente relacionada a este estudo, é a de Garmen et al. (2022), que investigou o uso de jogos sérios para avaliar habilidades cognitivas com base na Teoria das Múltiplas Inteligências de Gardner. O estudo envolveu 209 adultos, entre 19 e 59 anos, utilizando dois instrumentos: o jogo digital *Cutie Cuis*, parte do software *TOI*, e um questionário de autoavaliação baseado em instrumentos consolidados na literatura. O jogo incluiu dez testes lúdicos voltados à avaliação de seis tipos de inteligências (visual-espacial, lógico-matemática, musical-rítmica, linguístico-verbal, naturalista e cinestésica corporal), a partir de métricas como tempo de resposta, número de erros e acertos, e velocidade de reação. O questionário, com 42 itens organizados em dois blocos (“o que você gosta de fazer” e “em que você

é bom”), utilizou uma escala *Likert* de 4 pontos. Os dados foram analisados por meio da técnica de *Latent Profile Analysis* (LPA), que identificou quatro perfis cognitivos com base nos dados do jogo e seis a partir do questionário. Embora ambos apresentem boa acurácia classificatória, os perfis diferem entre si, indicando que jogo e questionário capturam dimensões distintas do perfil cognitivo. Os resultados evidenciam o potencial dos jogos sérios como ferramenta complementar de avaliação, ao gerar dados baseados na interação e no desempenho real dos participantes, ampliando as possibilidades de diagnóstico educacional.

O estudo de Sanchez et al. (2024) investigou o uso de jogos sérios para o desenvolvimento de competências relacionadas à linguagem, atenção e memória em estudantes em situação de vulnerabilidade socioeducacional. A pesquisa foi realizada com 75 estudantes entre 9 e 12 anos, distribuídos em quatro centros de educação primária na Espanha. Durante a intervenção, os participantes interagiram diariamente por 15 minutos com os jogos educacionais *Leobien* e *Walinwa*, que continham tarefas adaptativas voltadas ao estímulo de atenção e memória, ajustadas dinamicamente de acordo com o desempenho individual dos estudantes. Foram coletados dados como precisão das respostas, erros, tempo de resposta e velocidade de reação. A análise revelou avanços significativos nas competências avaliadas, além de indícios de que o desempenho nos jogos poderia prever o desempenho acadêmico. Como limitação, os autores destacam o uso exclusivo de dados quantitativos, sugerindo a inclusão de abordagens qualitativas para melhor compreensão de fatores contextuais.

Rowe et al. (2015) analisaram o uso de jogos sérios para avaliar a aprendizagem implícita de conceitos científicos, com foco na compreensão das Leis de Newton. Utilizando o jogo *Impulse*, projetado para promover a internalização de conceitos sem instrução direta, os autores propuseram uma abordagem avaliativa baseada na análise de padrões comportamentais, por meio de técnicas de mineração de dados educacionais. Foram registrados eventos como início e fim de níveis, cliques, colisões e estados do jogo, além de vídeos de tela, expressões faciais e diálogos dos participantes. Detectores automatizados foram desenvolvidos para identificar estratégias compatíveis com os princípios físicos abordados. O experimento, com 69 estudantes do ensino médio, de-

monstrou que aqueles que adotaram estratégias coerentes com os conceitos avaliados apresentaram melhor desempenho. As limitações incluem a necessidade de validação em amostras maiores e o uso de métodos qualitativos complementares. O estudo reforça o potencial dos dados de interação como recurso avaliativo alternativo, especialmente na identificação de aprendizagens emergentes durante a *gameplay*.

O estudo de Loh et al. (2013) propôs o uso de métricas de similaridade de *string* para identificar especialistas em jogos sérios, por meio da comparação entre o desempenho dos jogadores e trajetórias ideais definidas por especialistas. A abordagem envolveu a criação de um Índice de Similaridade de Especialistas, baseado na comparação entre sequências de ações dos jogadores e padrões referenciais dos especialistas. Foram capturados dados como coordenadas de movimento, carimbos de tempo, número de aldeões resgatados e eventos específicos do jogo. As ações foram monitoradas com a ferramenta *Performance Tracing Report Assistant* (PeTRA), totalizando 104.233 registros. A análise utilizou o Coeficiente de Similaridade de Jaccard (JAC) para mensurar o alinhamento entre o comportamento dos participantes e o modelo especialista. Os resultados indicaram que a métrica foi eficaz para distinguir jogadores mais experientes daqueles com desempenho inferior. Entre os 31 participantes universitários, apenas um apresentou JAC = 1, sendo classificado como especialista; os demais demonstraram valores inferiores, sugerindo menor aderência ao padrão de referência.

Embora as pesquisas mencionadas proponham métodos baseados em dados para a avaliação de desempenho dos estudantes, com foco em avaliações formativas e somativas e apresentando estratégias relevantes, tais métodos são geralmente desenvolvidos de forma *ad hoc*, voltados para jogos específicos ou não diretamente relacionados à avaliação da aprendizagem dos jogadores. Como resultado, apresentam baixa generalização e pouca reutilização em outros contextos. Em contraste, esta pesquisa propõe o cruzamento de diferentes tipos de dados para a formulação de métricas mais versáteis, com potencial de aplicação em múltiplos jogos educacionais.

Silva (2022) propõe a padronização do processo de captura, armazenamento e análise de dados por meio da ferramenta *GLBoard*. A solução tem como foco principal o apoio à avaliação do *game design*, oferecendo subsídios a desenvolvedores, professores e

jogadores por meio da visualização de dados de interação e indicadores de uso do jogo.

Entretanto, por ter sido concebido prioritariamente sob a perspectiva de *Game Analytics*, o *GLBoard* não explicita como os dados coletados podem ser convertidos em métricas e interpretados como indícios de aprendizagem, o que limita sua utilização como ferramenta de apoio à avaliação educacional em jogos.

Diante dessa lacuna, o objetivo desta dissertação é propor e integrar ao *GLBoard* métricas de aprendizagem capazes de relacionar dados de interação a aspectos e indícios de aprendizagem em jogos educacionais, viabilizando sua análise e interpretação para fins avaliativos. Para melhor compreensão da estrutura da ferramenta e do ponto de inserção da proposta desta pesquisa, o Capítulo 3 apresenta a arquitetura do *GLBoard*.

O próximo capítulo aborda os módulos do *GLBoard*, desenvolvidos pelos membros desta pesquisa, destacando sua importância para o estudo.

3

GLBoard: SISTEMA DE CAPTURA E ANÁLISE DE DADOS EM JOGOS EDUCACIONAIS

Este capítulo apresenta o *GLBoard* ([SILVA et al., 2022](#)), um sistema desenvolvido para padronizar a captura, o armazenamento e a análise de dados em jogos educacionais.

A concepção inicial de sua arquitetura foi idealizada por Melo et al. (??), no contexto de uma estratégia de *Game Learning Analytics* voltada à avaliação de *level design*. Essa primeira proposta estabeleceu as bases para a estruturação da captura e organização de dados provenientes da interação de estudantes com jogos educacionais.

Posteriormente, após ajustes arquiteturais e com o objetivo de generalizar o processo de captura e modelagem dos dados, bem como desenvolver uma biblioteca destinada a auxiliar desenvolvedores na implementação da coleta de dados, ampliando sua aplicabilidade a diferentes jogos e contextos educacionais, o sistema foi aprimorado e formalizado na versão apresentada por Silva et al. ([SILVA et al., 2022](#)). Nessa etapa, consolidou-se uma arquitetura genérica de dados, orientada à reutilização, à interoperabilidade e à sistematização das análises.

O objetivo principal do *GLBoard* é sistematizar a gestão desses dados, tornando o sistema reutilizável em diferentes contextos educacionais e jogos. Neste trabalho, o termo “genérico” refere-se principalmente a essa arquitetura e ao modelo de dados do

GLBoard; as métricas propostas, por sua vez, são aplicáveis a jogos que registrem os eventos necessários ao seu cálculo.

Um dos pilares dessa sistematização é a definição de um modelo genérico de dados capturáveis, capaz de representar interações comuns entre estudantes e jogos educacionais. Esses dados viabilizam análises detalhadas sobre o comportamento dos jogadores e os processos de aprendizagem envolvidos. A principal característica que confere ao GLBoard um caráter padronizado é o uso desses dados genéricos, entre os quais destacam-se:

- **Timestamps:** registram o tempo total de jogo, períodos de inatividade e o tempo necessário para concluir fases;
- **Mudanças de nível:** permitem monitorar o progresso e verificar o alcance dos objetivos;
- **Tentativas:** indicam persistência, dificuldade e evolução do jogador frente aos desafios;
- **Interações do usuário:** incluem cliques, transições de tela e frequência de uso de botões, fornecendo informações sobre o comportamento e o percurso do jogador.

Para que esses dados possam ser interpretados por sistemas externos, é essencial que estejam organizados conforme um modelo de dados estruturado. Mesmo sem conhecer os valores específicos gerados por cada jogo, o sistema analítico compreende os tipos e a estrutura dos dados, viabilizando análises relevantes. Essa padronização também facilita a replicação de estudos por outros pesquisadores, mesmo em contextos distintos ou com jogos diferentes.

Além da arquitetura genérica, o sistema inclui componentes voltados ao gerenciamento e à análise dos dados coletados. Com base na arquitetura proposta por Freire et al. (2016), o GLBoard é composto por quatro módulos principais, apresentados na Figura 4 e descritos a seguir:

1. **GLBoard Pacote:** biblioteca desenvolvida para a *game engine Unity*, responsável pela padronização e envio dos dados conforme a arquitetura definida. Essa bibli-

oteca facilita a integração do sistema aos jogos educacionais. O GLBoard Pacote está melhor detalhado na Subseção 3.1.5.

2. **GLBoard API**: responsável por gerenciar o recebimento, o processamento e a análise dos dados, além de armazená-los em um banco de dados estruturado, como descrito na Subseção 3.1.6.
3. **Banco de Dados**: armazena as interações registradas, de acordo com o modelo de dados estabelecido pelo sistema.
4. **Dashboard**: interface web responsável por apresentar as análises geradas pela API. A plataforma permite a criação de contas por desenvolvedores, educadores e estudantes. Desenvolvedores podem registrar jogos para habilitar o envio de dados; educadores monitoram turmas; estudantes acompanham seu desempenho individual. A Subseção 3.1.7 apresenta com mais detalhes essa interface.

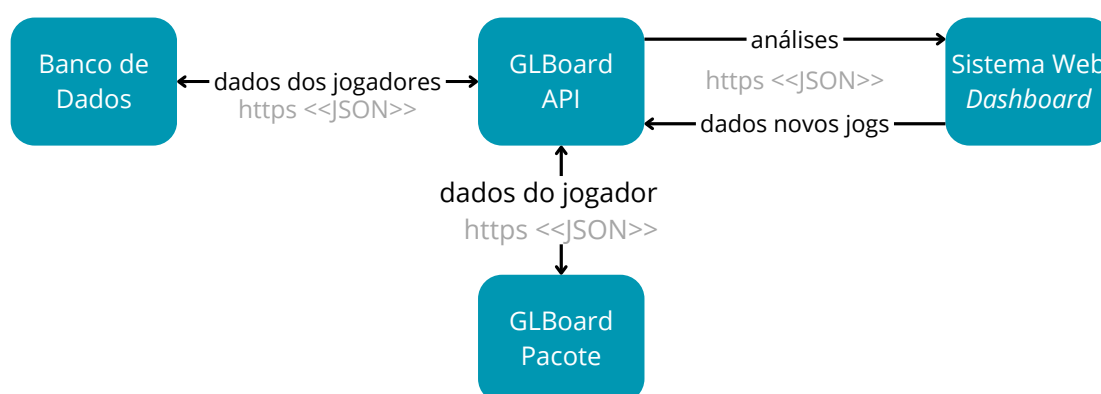


Figura 4 – Estrutura do sistema GLBoard.

Uma diferença relevante em relação à arquitetura proposta por Freire et al. (2016) é a ausência do módulo *Adaptador*, responsável por ajustar dinamicamente o jogo com base em avaliações pedagógicas. No *GLBoard*, optou-se por não incorporar esse componente à arquitetura central, a fim de preservar a generalidade do sistema e evitar o acoplamento direto entre os mecanismos de análise e a lógica interna dos jogos. A inclusão de um módulo adaptativo implicaria em uma maior complexidade de integração para desenvolvedores, uma vez que exigiria modificações estruturais no código do jogo para permitir intervenções dinâmicas.

Ao concentrar-se na padronização da captura, organização e análise dos dados, o *GLBoard* mantém uma arquitetura modular e reutilizável, permitindo que eventuais mecanismos adaptativos sejam implementados externamente, conforme as necessidades específicas de cada jogo. Dessa forma, a estrutura permanece robusta e flexível, possibilitando a aplicação da lógica de análise tanto para o refinamento do *game design* quanto para o apoio a decisões pedagógicas, sem impor complexidade adicional ao processo de desenvolvimento dos jogos.

Professores também se beneficiam do sistema ao terem acesso a dados granulares sobre o desempenho dos estudantes, facilitando a identificação de lacunas de aprendizagem, o planejamento de intervenções pedagógicas e o fornecimento de *feedback* personalizado. Além disso, esses dados podem fomentar discussões em sala de aula, integrando o conteúdo do jogo ao currículo formal. Jogadores, por sua vez, também têm acesso aos próprios dados de interação, podendo acompanhar seu progresso. Esse recurso estimula a autorreflexão e contribui para o desenvolvimento da autonomia no processo de aprendizagem.

Cada um desses *stakeholders* desempenha um papel específico e ativo no sistema. O fluxo do desenvolvedor, por envolver mais etapas, é representado em um diagrama de atividades disponível na Figura 20, apresentado no Anexo A. De forma geral:

- **Desenvolvedor:** cria uma conta no sistema web e cadastra um jogo, informando o nome e a descrição. Após o cadastro, recebe uma chave de *API* exclusiva, utilizada para integrar o jogo ao Pacote Unity. Essa integração exige a importação do pacote, a configuração das funcionalidades de captura de dados e a inserção da chave de *API*. A partir disso, as interações dos jogadores possam ser registradas, estruturadas segundo o modelo padronizado e enviadas à *API*, que armazena e processa as informações. O desenvolvedor pode acompanhar as análises gerais e individuais pelo sistema web.
- **Professor:** pode criar uma conta na plataforma para acessar as análises de dados. Para isso, solicita ao desenvolvedor a chave de *API* do jogo e a vincula à sua conta, obtendo acesso às análises disponibilizadas na plataforma.

- **Jogador:** interage com o jogo e visualiza suas análises no sistema web. Para isso, cria uma conta e solicita ao desenvolvedor uma chave de jogador. Cada jogador possui um ID único vinculado ao jogo, o que permite o acesso aos seus dados e ao acompanhamento de seu desempenho.

As seções a seguir descrevem os componentes do *GLBoard* e suas respectivas funções no sistema.

3.1 GLBoard - Arquitetura de captura de dados

Para garantir que o *GLBoard* atendesse às necessidades de um sistema genérico, desenvolveu-se um modelo de dados baseado em pesquisas existentes na literatura (ALENCAR; MELO; PESSOA, 2020; ALONSO-FERNANDEZ et al., 2019; MELO et al., 2020). Esse levantamento permitiu identificar os tipos de dados mais frequentemente coletados em jogos educacionais, bem como suas implicações para a aprendizagem dos jogadores (OWEN et al., 2014; ALONSO-FERNANDEZ et al., 2019; ALENCAR; MELO; PESSOA, 2020; HOOSHYAR; YOUSEFI; LIM, 2019). O modelo resultante apresenta quatro classes principais (*PlayerData*, *GameData*, *Phase*, *Session*), organizando dados do perfil do jogador e seu progresso no jogo, como ilustrado na Figura 5 e descrito a seguir.

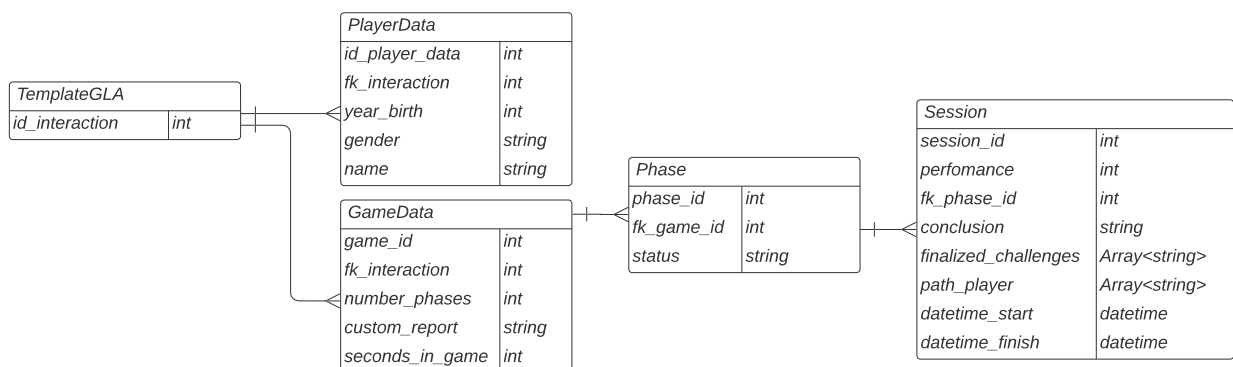


Figura 5 – Estrutura de dados do sistema GLBoard.

3.1.1 *Player Data*

Em estudos sobre jogos educacionais, é comum a captura de um conjunto de dados relacionados ao perfil do jogador. Esse conjunto de dados, no GLBoard, é armazenado em uma estrutura denominada *Player Data*. Os dados presentes nessa estrutura, descritos na Tabela 2, incluem informações como data de nascimento e gênero. Eles são importantes para analisar como diferentes grupos de jogadores interagem com o jogo, identificando padrões de comportamento, preferências e necessidades variadas. Análises demográficas podem revelar diferenças no engajamento, fornecendo subsídios para ajustes no *game design* e na entrega de conteúdo para atender melhor a esses diferentes perfis de jogadores.

Tabela 2 – Descrição dos campos da classe *Player Data*.

Campo	Tipo	Descrição
<i>day_birthday</i>	<i>STRING(FORMAT DATE-TIME)</i>	Data de nascimento do jogador.
<i>gender</i>	<i>STRING(FEMININO, MASCULINO, OUTROS)</i>	Gênero do jogador, ao invés de “OUTROS” pode especificar o gênero.
<i>id</i>	<i>STRING</i>	ID único do jogador. No jogo não deve haver outro jogador com o mesmo ID.
<i>name</i>	<i>STRING</i>	Nome do jogador, pode existir outro jogador com o mesmo nome.

3.1.2 *Game Data*

A classe *Game Data* organiza dados gerados durante a *gameplay* dos jogadores em fases, conforme detalhado na Tabela 3. Essa organização por fases é comum em jogos e permite uma correspondência direta entre cada fase do objeto *Game Data* e cada fase do jogo, onde são registradas as interações do jogador. Importante destacar que a definição de “fase” adotada no modelo é flexível, isto é, ela não se restringe apenas a jogos com fases explícitas. Em jogos contínuos, por exemplo, segmentos entre *checkpoints* ou tentativas de superar desafios podem ser tratados como fases. Assim, o sistema se adapta bem tanto a jogos com divisões claras de fases quanto àqueles mais fluídos, garantindo a aplicabilidade do modelo em diversos contextos de jogos.

Tabela 3 – Descrição dos campos da classe *Game Data*.

Campo	Tipo	Descrição
<i>number_phases</i>	INT	Total de fases no jogo.
<i>phases_unlocked</i>	INT	Quantidade de fases concluídas pelo jogador.
<i>player_minutes_in_game</i>	INT	Tempo total, em minutos, do jogador no jogo.
<i>phases</i>	ARRAY(PHASES)	Array contendo dados de cada fase, incluindo informações específicas.
<i>date_last_login</i>	STRING(DATE TIME)	Data e hora do último login do jogador.

3.1.3 Phase Data

A classe *Phase* desempenha papel estruturante no modelo de dados, ao compor o *array* de *Phases* presente no campo *phases* (Subseção 3.1.2). Sua estrutura é detalhada na Tabela 4, que apresenta a descrição de cada campo pertencente à classe *Phase Data*.

Tabela 4 – Descrição dos campos da classe *Phase Data*.

Campo	Tipo	Descrição
<i>phase_id</i>	STRING	Identificador único de uma fase. Deve ser exclusivo e não necessariamente sequencial (por exemplo, “Fase dos caminhos perdidos”, “Fase da viagem”).
<i>status</i>	STRING(NAO_FINALIZADA, FINALIZADA)	Indica se uma fase foi concluída (“FINALIZADA”) ou não (“NAO_FINALIZADA”) pelo jogador.
<i>Sessions</i>	ARRAY(SESSION)	Contém um <i>array</i> de <i>sessions</i> , cada uma com informações específicas relacionadas à fase.

Essa estrutura permite o detalhamento do progresso do jogador dentro de cada fase do jogo, abrangendo desde a identificação única da fase, o *status* de conclusão, até as sessões/tentativas específicas que ocorrem na fase (Subseção 3.1.4).

3.1.4 Session Data

A classe *Session Data* é utilizada para armazenar cada sessão ou tentativa que um jogador realiza dentro de uma fase. Essas sessões são armazenadas como elementos em um *array* chamado *session* dentro de uma *Phase*. Cada *Session* contém informações detalhadas sobre a tentativa, conforme descrito na Tabela 5. Por exemplo, em um jogo, uma fase pode envolver várias tentativas de resolver um quebra-cabeça; cada tentativa, bem sucedida ou não, é registrada como uma sessão independente na fase.

Essa estrutura permite uma análise detalhada do desempenho do jogador em cada sessão, abrangendo desde o resultado final da tentativa (vitória, derrota, desis-

Tabela 5 – Descrição dos campos da classe *Session Data*.

Campo	Tipo	Descrição
<i>conclusion</i>	<i>STRING(DERROTA, VITORIA, DESISTENCIA)</i>	Resultado da sessão, podendo ser vitória, derrota ou desistência.
<i>dateTimeStart</i>	<i>DATETIME</i>	Início da sessão.
<i>dateTimeFinish</i>	<i>DATETIME</i>	Término da sessão.
<i>finalized_challenges</i>	<i>ARRAY(STRING)</i>	Desafios concluídos pelo jogador durante a sessão, detalhando vitórias ou <i>checkpoints</i> alcançados.
<i>path_player</i>	<i>ARRAY(STRING)</i>	Percurso realizado pelo jogador, indicando interações e trajetória detalhada, por exemplo, a sequência de plataformas percorridas.
<i>performance</i>	<i>INT</i>	Campo que quantifica a performance do jogador na sessão, em uma escala de 0-100. A definição do valor fica a critério do desenvolvedor, considerando diversos fatores de influência.

tência), o tempo gasto, os desafios enfrentados e superados, até o caminho percorrido e a performance geral. As informações coletadas aqui são importantes para entender o comportamento do jogador, avaliar a dificuldade das sessões e subsidiar ajustes no *game design*, para melhorar a experiência educativa.

3.1.5 GLBoard - Pacote Unity

GLBoard - Pacote Unity é uma biblioteca em *C-Sharp*, denominada “Pacote”, compatível com a engine *Unity*. A escolha pela *Unity* deve-se à sua ampla adoção entre desenvolvedores *indie* e à possibilidade de criação gratuita de jogos educacionais sem fins lucrativos (SILVA et al., 2022).

O módulo é composto por classes e funções que organizam os dados de interação dos jogadores e os enviam à *API*. A Figura 6 apresenta um exemplo de integração utilizando o Pacote.

Na linha 2, a classe `GLBoard` é instanciada, passando a representar o ponto de integração do jogo com a arquitetura de dados do sistema. Na linha 5, é inicializada com dois parâmetros: a *Chave API* e o ID do usuário. A *Chave API* é um *UUID* gerado automaticamente após o cadastro do jogo no sistema web, responsável por vincular o jogo ao *GLBoard* e assegurando a correta associação das interações no banco de dados. O ID do usuário, obtido na tela de login (por exemplo, e-mail ou *nickname*), identifica cada jogador de forma única.

```
1 public class Initialize: MonoBehaviour {
2     public GLBoard api;
3
4     public async void Start() {
5         api = new GLBoard(game_id: "API_KEY", userId: user.email);
6         await api.LOAD_USER_DATA();
7         api.SetPlayerData(
8             name: "Player 1",
9             day_birthday: "21/04/1999",
10            gender: GENDER.MASCULINO
11        );
12        api.SetQuantPhaseGame(13);
13        api.AddPhaseGame("Fase 1 - Dragões");
14        api.AddSectionInPhase(
15            phase_id: "Fase 1 - Dragões",
16            conclusion: STATUS_SECTION.DERROTA,
17            performance: 10,
18            dateTimeStartSection: initSection,
19            dateTimeFinishSection: DateTime.Now,
20            finalized_challenges: new List < string > () {
21                "Derrotou inimigo 1",
22                "Derrotado pelo inimigo 2"
23            },
24            path_player: new List < string > () {
25                "Plataforma 1",
26                "Plataforma 3"
27            },
28        );
29        StartCoroutine(api.SEND_USER_DATA());
30    }
31 }
```

Figura 6 – Exemplo de integração do jogo com o Pacote.

Na linha 6, os dados do jogador são carregados. O pacote consulta a *API* para verificar a existência de registros prévios do jogador no jogo em questão. Se existirem, esses dados são recuperados e incorporados à instância da classe *GLBoard*, evitando sobrescritas. Assim, novas tentativas são somadas ao histórico já existente do jogador.

Nas linhas 7 a 9, são inseridos os dados de *Player Data* (Seção 3.1.1). Na linha 12, define-se o número total de fases do jogo. Em seguida, a linha 13 registra uma fase específica por meio da função `api.AddPhaseGame`, que deve ser chamada no início de cada fase.

Na linha 14, uma nova sessão (ou tentativa) é registrada. A função recebe o nome da fase e os dados da tentativa (Seção 3.1.4). Essa função deve ser chamada sempre que o jogador concluir uma tentativa, independentemente do resultado (vitória, derrota ou desistência).

Na linha 29 a função `api.SEND_USER_DATA` envia os dados consolidados para a *API*. Em implementações reais, as funções devem ser chamadas em momentos distintos do ciclo de execução do jogo, e não de forma concentrada, como no exemplo.

3.1.6 GLBoard - API

A *GLBoard-API* atua como o controlador central do sistema. Desenvolvida em *Python*, utilizando o *framework Flask-API*, sua principal função é receber os dados de interação dos jogadores, enviados pelo Pacote Unity (Seção 3.1.5), e armazená-los de forma estruturada em um banco de dados baseado no *Firebase* (Figura 4). Cada jogo possui uma lista de jogadores vinculados, cujas interações são organizadas e registradas conforme o modelo de dados.

A *GLBoard-API* disponibiliza duas categorias principais de análise, acessíveis por meio de *endpoints*:

1. **Análises Gerais do Jogo:** agregam os dados de todos os jogadores de um determinado jogo, fornecendo informações como: (i) Número total de jogadores; (ii) Distribuição por gênero; (iii) Percentual e número de fases completadas; (iv) Tempo total e médio de jogo; (v) Número médio de tentativas por fase; (vi) Tempo médio por fase; (vii) Taxas média e geral de vitória/derrota por fase; e (viii) Desempenho médio por fase.
2. **Análises de Grupo:** utilizam técnicas de clusterização para identificar padrões de comportamento entre os jogadores, considerando métricas como: (i) Relação entre tempo de jogo e número de tentativas; e (ii) Relação entre número de derrotas e tentativas.

Diferentemente das abordagens tradicionais de *Game Analytics*, essas análises não se limitam apenas ao desempenho do jogador no jogo. Ao integrar a avaliação de estratégias de aprendizagem e padrões de comportamento, aproximam-se dos objetivos do *Learning Analytics* (LA), oferecendo uma visão mais abrangente do processo de aprendizagem dos jogadores.

Os resultados podem ser acessados em formato *JSON* por meio das rotas da API, e também visualizados em gráficos e tabelas na interface de *dashboard* do sistema web.

3.1.7 GLBoard - Web

O sistema *Web* tem como objetivo apresentar os resultados das análises realizadas na API (Seção 3.1.6) através de um *dashboard*. Além disso, permite que desenvolvedores cadastrem jogos e que, junto aos professores, visualizem as análises das interações dos jogadores, enquanto os estudantes podem visualizar seu próprio desempenho. O processo de login e cadastro (Figura 7) é comum a todos os usuários, sendo necessário apenas especificar o tipo de usuário durante o cadastro.

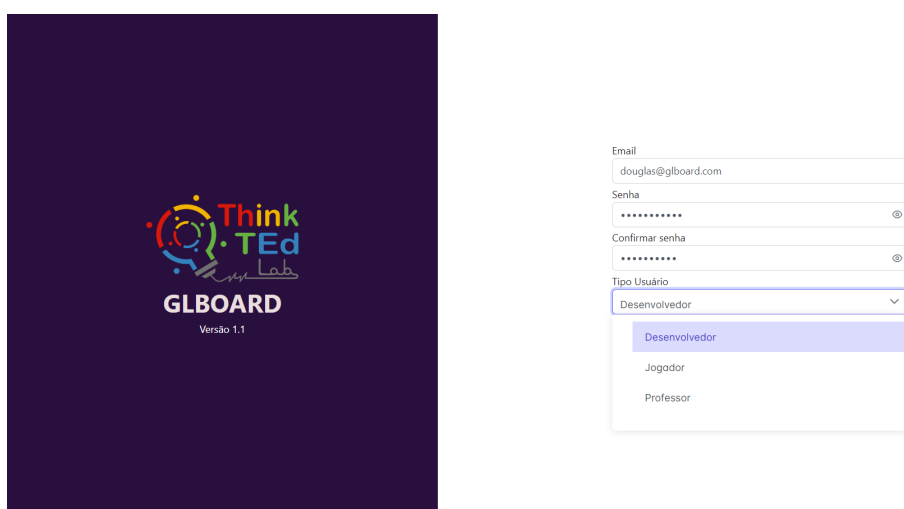


Figura 7 – Tela de cadastro no sistema *web*.

Após o registro e *login*, o desenvolvedor acessa a lista de jogos cadastrados (Figura 8). Essa lista é visualizada através de uma tabela que inclui o nome do jogo, a quantidade de jogadores e a *Chave API* do jogo. A *Chave API* é necessária para a transmissão de dados da *game engine Unity* para a API via Pacote.

O botão “Registrar novo” permite cadastrar novos jogos, adicionando-os à lista do desenvolvedor após informar o nome e a descrição, gerando uma nova chave única para o jogo.

Os professores podem visualizar as análises das interações dos jogadores importando a *Chave API* de um jogo cadastrado pelo desenvolvedor. Para isso, é necessário solicitar essa chave ao desenvolvedor. Na interface, os professores têm acesso a uma lista similar à dos desenvolvedores. Clicando em “Associar Jogo” (no mesmo local onde os desenvolvedores veem “Registrar Novo”), o professor insere a *Chave API* e o jogo é inserido na lista.

Nome	Chave	Número de Jogadores	Ação
MyName	_25QFu6_u34tDLpx2Sv-JA	4	[Add] [Edit] [Delete]
Robô Euroi	38-g3GfgIwUPqMzh3PQLbA	3	[Add] [Edit] [Delete]
Game Teste	45X1SnmfsZX-DM0Ezj-XtQ	13	[Add] [Edit] [Delete]

Figura 8 – Tela de listagem de jogos associados no sistema *web*.

Os estudantes podem visualizar seu próprio desempenho por meio da interface de listagem de jogos, na qual também está disponível o botão “Associar Jogo”. Para realizar essa associação, é necessário informar a *Chave Jogador*, fornecida pelo desenvolvedor ou pelo professor. Essa chave pode ser obtida na tela de listagem de usuários do jogo (Figura 9), acessível ao desenvolvedor ou professor, que apresenta informações como os jogadores cadastrados, a última sessão registrada, o nível de conclusão do jogo (quantidade de fases finalizadas) e a respectiva *Chave Jogador*.

Nome do Jogador	Última Sessão	% Conclusão	Chave Jogador
	31/07/2023 21:17:32	7	NDVYTMVNubWZzeigtREDwRXpqlVh0USQkLSQkQWlyd89uX99eaXZlaXJh
	30/07/2023 15:55:58	10	NDVYTMVNubWZzeigtREDwRXpqlVh0USQkLSQkRmYaaXB0X9pFUIVTX0IB0QVETw==
	31/07/2023 19:36:48	9	NDVYTMVNubWZzeigtREDwRXpqlVh0USQkLSQkRmYaaXB0X0pIc3Vz
	27/07/2023 22:11:29	3	NDVYTMVNubWZzeigtREDwRXpqlVh0USQkLSQkR2FcmllBf90b3N0YQ==
	02/08/2023 19:37:30	4	NDVYTMVNubWZzeigtREDwRXpqlVh0USQkLSQkR2FcmllBf9MaWiH
	01/08/2023 21:15:37	11	NDVYTMVNubWZzeigtREDwRXpqlVh0USQkLSQkUGVkcma8FTG9wZXM=
	31/07/2023 23:39:21	6	NDVYTMVNubWZzeigtREDwRXpqlVh0USQkLSQkVmfz2ChzZWxfUzIdmE=

Figura 9 – Tela de listagem de jogadores.

Na tela de listagem de jogos, ao selecionar um jogo, desenvolvedores e professores são direcionados ao *dashboard* de análises gerais do jogo, que inclui:

- **Distribuição de Gênero dos Jogadores:** um gráfico de pizza que mostra a proporção de jogadores masculinos e femininos;
- **Quantidade Total de Jogadores:** exibe o número total de jogadores no jogo;
- **Total de Fases:** mostra a quantidade de fases disponíveis no jogo;
- **Horas Jogadas:** apresenta a soma das horas jogadas por todos os jogadores;

- **Porcentagem de Conclusão de Fases:** um gráfico de barras que indica a porcentagem de jogadores que completaram cada fase;
- **Média e tentativas por Fase:** gráfico de barras mostrando o número de tentativas feitas pelos jogadores em cada fase;
- **Média de desempenho em Fase:** gráfico de barras que exibe a média de desempenho dos jogadores em cada fase;
- **Tempo Médio e geral em Cada Fase:** gráfico de barras com o tempo total gasto pelos jogadores em cada fase;
- **Sessões Realizadas Por Dia:** gráfico de linha que mostra o número de sessões jogadas por dia da semana;
- **Derrota/Vitória por Fase:** gráfico de barras empilhadas que compara o número de sessões que resultaram em vitória *versus* derrota em cada fase.

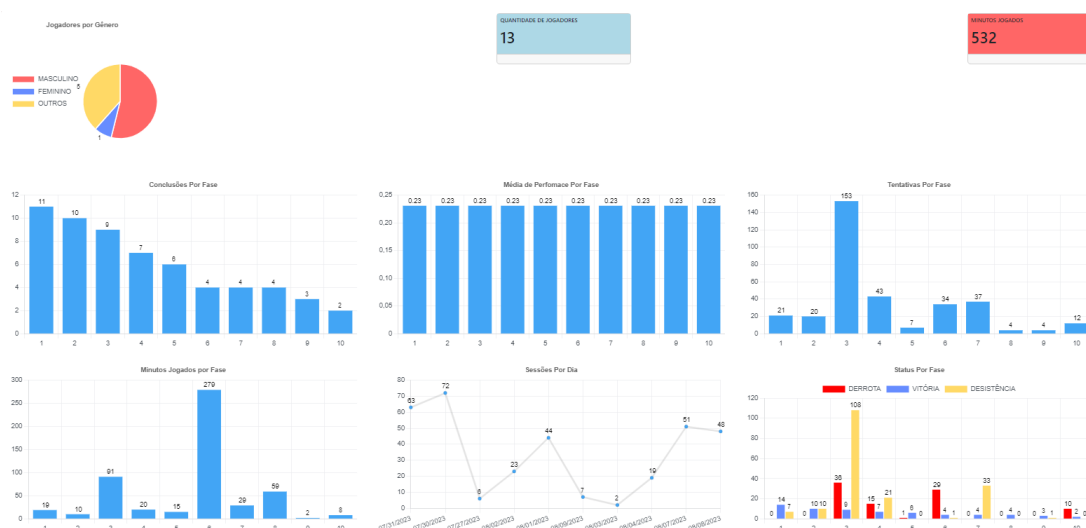


Figura 10 – *Dashboard* de análises gerais de um jogo.

Os estudantes conseguem visualizar praticamente as mesmas informações mostradas nos gráficos, porém restritas apenas aos dados de suas interações.

Dado o contexto e as funcionalidades que o *GLBoard* proporciona, esse sistema foi utilizado nesta pesquisa como a ferramenta para capturar os dados de interação dos

jogadores com os jogos educacionais, sendo esses dados necessários para a aplicação do método proposto.

Apesar do *GLBoard* proporcionar análises importantes para a avaliação de *game design* em jogos educacionais, ele se limita a esse tipo de análise e não se aprofunda nas análises de aprendizagem. Um dos objetivos desta pesquisa é remodelar a arquitetura de dados do *GLBoard* para integrar o método proposto, permitindo a análise e avaliação do desempenho dos estudantes em diferentes aspectos de aprendizagem.

Os ajustes propostos para o *GLBoard* nesta dissertação consistem: (i) na integração de novos atributos à arquitetura de dados, conforme descrito na Seção 5.4; e (ii) na proposição de novos gráficos no *dashboard*, destinados à visualização das análises derivadas das métricas desenvolvidas (Figuras 15 e 17). Esses ajustes são detalhados no Capítulo 5.

Esses ajustes visam transformar o *GLBoard* em uma ferramenta capaz de servir como fonte de dados, aplicar o método desenvolvido nos jogos e fornecer um *dashboard* para a visualização dos resultados, proporcionando uma análise mais abrangente e detalhada do processo de aprendizagem dos estudantes.

Dessa forma, este capítulo apresenta a estrutura, os módulos e as funcionalidades do *GLBoard*, evidenciando seu papel como infraestrutura de captura e análise de dados na pesquisa, bem como as limitações que motivam sua remodelagem. No próximo capítulo, é detalhada a Fase 2 da abordagem *Design-Based Research*, conforme apresentada na Seção 1.2. São descritos os procedimentos adotados para o desenvolvimento das métricas, a remodelagem do *GLBoard* e o planejamento da aplicação e validação da solução proposta.

4

PROJETO E IMPLEMENTAÇÃO DA PROPOSTA

Este capítulo detalha o desenho e a implementação da proposta desenvolvida na Fase 2 da abordagem *Design-Based Research*, apresentada na Seção 1.2. São descritos os procedimentos adotados para a seleção dos aspectos de aprendizagem e dados comuns a jogos educacionais, o desenvolvimento das métricas, a remodelagem do *GLBoard*, bem como o planejamento dos testes. Também são apresentados os instrumentos de coleta, bem como os critérios de aplicação e validação da solução proposta, os quais são detalhados no Capítulo 5.

Conforme discutido na Seção 1.2, abordagens centradas exclusivamente no *Game Design* ou em questionários tradicionais mostram-se insuficientes para atender aos objetivos dos jogos educacionais no apoio ao processo de aprendizagem. Torna-se, portanto, pertinente estruturar métodos fundamentados nos dados gerados pelas interações dos jogadores, viabilizando a construção de métricas capazes de representar, de forma mais precisa, indícios do processo de aprendizagem.

Métricas são medidas quantitativas utilizadas para avaliar, comparar e rastrear o desempenho ou progresso em relação a determinados objetivos. No contexto de jogos educacionais, podem ser derivadas de dados isolados ou compostos, fornecendo informações sobre o comportamento e a aprendizagem dos estudantes. Uma métrica pode consistir em um dado único, como o tempo gasto em uma tarefa, ou em uma combinação de elementos, como número de tentativas, erros e tempo de resposta. A

partir desses dados, podem ser aplicadas fórmulas ou algoritmos que geram métricas mais complexas. Por exemplo, a relação entre tempo de resposta e precisão pode indicar o desempenho cognitivo, oferecendo indícios sobre o processo de aprendizagem.

Nesta pesquisa, as métricas foram desenvolvidas com o objetivo de relacionar dados das interações aos aspectos de aprendizagem, permitindo interpretações sobre o progresso dos estudantes. A Figura 11 apresenta as etapas envolvidas no desenvolvimento, aplicação e refinamento das métricas, correspondentes à Etapa 2 da metodologia. As subseções a seguir detalham cada uma dessas etapas.



Figura 11 – Processo de desenvolvimento, aplicação e ajustes das métricas.

4.1 Seleção de aspectos de aprendizagem

Conforme descrito na Seção 2.1.1, aspectos de aprendizagem referem-se a habilidades desenvolvidas ao longo do processo educacional, como engajamento, motivação e compreensão conceitual. Jogos educacionais podem contribuir para o desenvolvimento de diferentes habilidades, como pensamento crítico, funções executivas e engajamento. A

mensuração desses aspectos oferece indícios relevantes sobre o processo de aprendizagem.

Nesta etapa, foi realizada uma investigação exploratória para identificar e catalogar aspectos da aprendizagem, com base na análise de estudos e jogos existentes, como apresentado na Seção 5.1. Esse levantamento permitiu mapear os aspectos de aprendizagem mais recorrentes em jogos educacionais e serviu de referência para a definição das métricas propostas. Também foram identificados os dados comumente disponíveis nesses jogos (Seção 5.2), os quais fundamentaram a formulação das métricas propostas nesta pesquisa.

4.2 Seleção de dados disponíveis em jogos educacionais

Após a identificação dos aspectos de aprendizagem (Seção 5.1), foi realizada uma análise com o objetivo de identificar dados de interação passíveis de coleta em diferentes jogos educacionais, capazes de subsidiar a construção das métricas.

Essa análise envolveu uma revisão da literatura sobre dados utilizados na avaliação de aprendizagem em jogos, bem como a análise da arquitetura de dados do *GLBoard*, buscando mapear informações recorrentes, como dados de sessão, perfil de estudantes e registros de interação (*gameplay*). A partir desse procedimento, foi definido um conjunto de dados comuns que fundamentou o desenvolvimento das métricas apresentadas na Seção 5.3. Os resultados detalhados desse mapeamento são apresentados na Seção 5.2.

4.3 Desenvolvimento das métricas e integração aos jogos

Para cada aspecto identificado, os dados coletados nos jogos foram analisados em busca de padrões e correlações que permitissem sua mensuração no contexto educacional. O processo resultou na formulação de métricas voltadas à avaliação de subaspectos dos aspectos de aprendizagem. Conforme a natureza das interações registradas, diferentes métricas foram definidas para representar um mesmo aspecto.

Durante o desenvolvimento, estabeleceu-se a distinção entre dois tipos principais

de métricas:

- **Métricas Globais:** voltadas à avaliação do desempenho geral do estudante, como o engajamento ao longo do jogo.
- **Métricas Específicas:** focadas em sessões, fases ou *checkpoints*, que variam em termos de complexidade, mecânicas e dinâmicas. Permitem analisar como os estudantes enfrentam desafios pontuais, identificando padrões de desempenho e dificuldades específicas. Essas informações subsidiam estratégias pedagógicas direcionadas à melhoria da aprendizagem.

O próximo capítulo apresenta as métricas desenvolvidas, os respectivos aspectos de aprendizagem representados e os dados utilizados em sua formulação.

Para a coleta dos dados, foi utilizado o sistema *GLBoard* (Capítulo 3), responsável por capturar interações dos jogadores por meio do *Pacote Unity*, aplicar as métricas via *GLBoard API* e apresentar os resultados por meio do módulo de *dashboard*.

Nesta fase inicial, apenas a captura de dados foi realizada pelo *GLBoard*, uma vez que a integração das métricas à API e ao *dashboard* demandaria desenvolvimento adicional. Assim, a aplicação das métricas foi conduzida por meio de um *Colab Notebook* (Python), que processou os dados no formato *JSON* para gerar as análises de desempenho.

Para viabilizar essa coleta, foi necessário remodelar a arquitetura de dados original do *GLBoard*, anteriormente voltada à análise centrada no *game design*, a fim de atender aos requisitos das métricas propostas. A nova arquitetura será detalhada no capítulo seguinte (Capítulo 5).

4.4 Teste e aplicação das métricas (Etapa III)

Após o desenvolvimento das métricas, iniciou-se sua aplicação a dados reais provenientes das interações de estudantes com jogos educacionais. Para isso, foi selecionado um conjunto de jogos nos quais as métricas foram aplicadas a partir dos dados coletados, permitindo a avaliação do desempenho dos participantes.

Paralelamente, para responder à QP2, foi elaborado um questionário de autoavaliação com perguntas em escala Likert (Tabela 7), voltadas à percepção dos estudantes sobre seu desempenho em aspectos específicos de aprendizagem, considerando a experiência com o jogo. Os aspectos avaliados no questionário correspondem àqueles representados pelas métricas. Por exemplo, para uma métrica associada ao *Raciocínio Lógico*, a pergunta correspondente poderia ser: “Quando eu estava resolvendo os desafios lógicos, sempre busquei as melhores soluções”.

Esse cruzamento entre os dados das métricas e as autoavaliações permite comparar a percepção dos estudantes com os resultados extraídos das interações registradas.

4.5 Refinamento e validação das métricas (Etapa IV)

Esta etapa ocorreu em paralelo ao desenvolvimento da solução, envolvendo o refinamento das métricas a partir de *feedbacks* de especialistas das áreas de educação, análise de dados e psicometria. Foram realizadas reuniões e entrevistas com os especialistas para apresentar as métricas e discutir sua clareza, coerência e relação com aspectos de aprendizagem em jogos educacionais. No caso da psicometria, foram utilizados agentes inteligentes como apoio no processo de análise. A coleta de contribuições incluiu sugestões de ajustes nos nomes, descrições e formas de aplicação das métricas, bem como comentários sobre possíveis limitações e melhorias. Esse processo buscou assegurar que as métricas apresentem consistência, compreensibilidade e relevância para diferentes contextos de uso, permitindo que possam ser utilizadas como apoio na análise da aprendizagem de estudantes em ambientes mediados por jogos educacionais.

Em síntese, este capítulo descreveu o encadeamento de procedimentos que compõem a Fase 2 da abordagem *DBR* (Seção 1.2). Com base nessa estrutura, o próximo capítulo (Capítulo 5) apresenta a proposta, detalhando as métricas formuladas para responder à QP1, as adaptações na arquitetura de dados do *GLBoard*, o questionário de autoavaliação associado à QP2 e o protocolo adotado para validação e refinamento das métricas.

5

PROPOSTA

O presente capítulo descreve as métricas formuladas para avaliar estudantes em diferentes aspectos de aprendizagem, com o objetivo de responder à QP1: “Como avaliar aspectos de aprendizagem a partir da interação de estudantes com jogos educacionais?”. Também apresenta as adaptações realizadas na estrutura de dados do sistema *GLBoard*, utilizadas para capturar as informações necessárias à aplicação dessas métricas durante a interação dos estudantes com jogos educacionais. Além disso, o capítulo detalha o questionário de autoavaliação aplicado aos estudantes, voltado aos aspectos de aprendizagem relacionados às métricas propostas, com o intuito de responder à QP2: “Como a autoavaliação de desempenho dos estudantes se compara ao uso de dados de interação coletados em jogos educacionais?”. Por fim, descreve o processo de validação das métricas realizado com especialistas das áreas de educação, análise de dados e psicometria, visando assegurar a clareza, coerência e aplicabilidade das métricas em diferentes contextos de uso.

5.1 Aspectos de aprendizagem selecionados

As métricas desenvolvidas visam avaliar o desempenho dos jogadores em aspectos específicos da aprendizagem, alinhados às habilidades potencialmente estimuladas por jogos educacionais. O mapeamento e a mensuração desses aspectos permitem gerar *insights* sobre o processo de aprendizagem, possibilitando compreender tanto o progresso dos estudantes, quanto seu desempenho em habilidades gerais e no conteúdo

abordado no jogo.

Para esta etapa, foram selecionados quatro aspectos de aprendizagem: engajamento (Subseção 5.1.1), desenvolvimento cognitivo (Subseção 5.1.2), desenvolvimento motor (Subseção 5.1.3) e unidade curricular (Subseção 5.1.4). Esses aspectos foram escolhidos com base em sua relevância para o contexto dos jogos educacionais e na viabilidade de sua mensuração por meio de dados de interação. Aspectos com menor aderência pedagógica ou que demandam instrumentos avaliativos adicionais, mais complexos, foram reservados para fases futuras da pesquisa.

5.1.1 Engajamento

Engajamento é um componente essencial da aprendizagem, refletindo o envolvimento emocional, cognitivo e comportamental do estudante com a atividade educacional (FREDRICKS; BLUMENFELD; PARIS, 2004). Em jogos educacionais, esse aspecto é particularmente relevante, pois influencia diretamente a motivação, a persistência diante de desafios e a qualidade da interação com os conteúdos do jogo (NADEEM; OROSZLANYOVA; FARAG, 2023; LI; CHEN; DENG, 2024b).

Nesta pesquisa, o engajamento é considerado um aspecto de aprendizagem que pode ser mensurado por meio de dados de interação. Ao invés de focar apenas em percepções subjetivas, busca-se analisar padrões objetivos de comportamento, como tempo contínuo de jogo, frequência de acessos e superação de fases, com o intuito de identificar níveis de envolvimento ao longo da experiência.

Mensurar o engajamento permite compreender em quais momentos os estudantes demonstram maior dedicação e interesse, bem como detectar pontos de evasão ou de queda de motivação ao longo da interação com o jogo. Esses dados contribuem para o ajuste do *game design* e o aprimoramento da experiência pedagógica, promovendo uma aprendizagem mais significativa.

A escolha do engajamento como métrica nesta pesquisa se justifica por seu potencial impacto no sucesso da aprendizagem e por sua viabilidade de mensuração a partir dos dados capturados no jogo, possibilitando análises contínuas e automatizadas

do comportamento dos estudantes durante a atividade.

5.1.2 Desenvolvimento cognitivo

O desenvolvimento cognitivo envolve habilidades mentais, como pensamento lógico, memória de trabalho, resolução de problemas e flexibilidade cognitiva (FLAVELL, 1977). Em jogos educacionais, essas habilidades são estimuladas por desafios que exigem raciocínio estratégico, adaptação e tomada de decisão, enquanto a natureza lúdica do jogo favorece o engajamento e a assimilação dos conteúdos.

Jogos educacionais têm mostrado potencial para estimular aspectos de aprendizagem, como pensamento computacional e funções executivas, ao demandarem do estudante raciocínio, tomada de decisão e adaptação de estratégias durante o *gameplay*. O jogo “Lightbot”, por exemplo, introduz conceitos de programação como *loops* e procedimentos, exigindo raciocínio sequencial e reconhecimento de padrões (SOUZA et al., 2018). Já “Cadê Minha Pizza?” explora algoritmos de caminho mínimo em mapas representados por grafos, promovendo o exercício do pensamento computacional (HONDA et al., 2022). Krause, Hounsell e Gasparini (2020) propõem um modelo que relaciona diferentes mecânicas de jogo a funções executivas como controle inibitório e flexibilidade cognitiva, habilidades fundamentais para o raciocínio adaptativo.

Pesquisas empíricas têm reforçado a relação entre a prática de jogos e o desempenho em funções executivas e outros indicadores de aprendizagem, especialmente em experiências que exigem resolução de problemas e tomadas de decisão. Chaarani et al. (2022), por exemplo, observaram associações com ganhos em controle inibitório e memória de trabalho em crianças que jogavam com maior frequência.

Apesar dessas evidências, ainda há desafios na avaliação sistemática desses aspectos durante a interação com o jogo. Em muitos contextos, a mensuração do desenvolvimento cognitivo depende de instrumentos externos, como testes padronizados ou questionários, que nem sempre capturam os processos que ocorrem ao longo do *gameplay*. Torna-se, portanto, relevante investigar formas de avaliar indícios de aprendizagem diretamente a partir dos dados gerados pelas interações dos estudantes com o

jogo.

A base conceitual dessa abordagem fundamenta-se na premissa de que processos cognitivos podem ser inferidos a partir de comportamentos observáveis durante a resolução de tarefas. Em ambientes digitais interativos, como jogos educacionais, as ações realizadas pelos jogadores — como tentativas, erros, revisões de estratégia e tempo de resposta — constituem registros objetivos do processo de resolução de problemas, permitindo a identificação de padrões associados a habilidades cognitivas específicas.

Nesta pesquisa, a avaliação ocorre a partir da forma como os jogadores enfrentam desafios durante o *gameplay*, permitindo identificar estratégias de resolução de problemas, adaptação a dificuldades e resposta a erros, considerados indicadores de flexibilidade cognitiva e persistência. Estudantes que ajustam suas abordagens após falhas demonstram indícios de pensamento adaptativo, enquanto padrões repetitivos de erro podem indicar necessidade de suporte adicional.

A análise contínua das interações também pode orientar ajustes pedagógicos e refinamentos no *game design*, promovendo intervenções mais alinhadas às necessidades dos estudantes. Esses aspectos justificam a escolha do desenvolvimento cognitivo como um dos focos desta pesquisa.

5.1.3 Desenvolvimento motor

O desenvolvimento motor envolve o aprimoramento de habilidades físicas relacionadas à coordenação, controle e precisão de movimentos (AKBARI et al., 2009). Jogos educacionais podem favorecer esse desenvolvimento ao incorporar mecânicas que exigem respostas motoras, como movimentar os dispositivos ou interagir com controladores sensíveis ao movimento, comuns em jogos móveis e de Realidade Virtual (VR) (VUKIĆEVIĆ et al., 2019). Esses desafios promovem o controle motor fino e grosso, integrando o aprendizado com a ação física.

Na literatura, destacam-se jogos como o *Parkrun*, que utiliza o acelerômetro de dispositivos móveis para controlar personagens em movimento, exigindo precisão manual e promovendo coordenação olho-mão (PINHEIRO et al., 2022). Outro exemplo é

o *GameLab*, voltado ao desenvolvimento da motricidade fina em pessoas com deficiência motora, no qual os jogadores movimentam uma bola por um labirinto através de inclinações do aparelho Pinheiro, Segundo e Sousa (2020).

A análise do desempenho motor durante o jogo permite detectar padrões de dificuldade e identificar estudantes que possam precisar de suporte adicional. Além disso, pode revelar falhas de *game design* — por exemplo, quando a exigência motora se torna excessiva e compromete a aprendizagem. Esses dados não apenas contribuem para ajustar o nível de dificuldade e acessibilidade do jogo, mas, quando combinados com outros aspectos avaliativos, fortalecem a análise sobre o processo de aprendizagem. Por isso, o desenvolvimento motor é incluído como um dos aspectos-chave a serem mensurados nesta pesquisa.

5.1.4 Unidade curricular

O aspecto denominado *Unidade Curricular* refere-se ao domínio dos conteúdos abordados no jogo educacional, alinhados aos objetivos pedagógicos da disciplina.

Estudos têm descrito o potencial dos jogos educacionais para promover a aprendizagem de conteúdos específicos (BARZILAI; BLAU, 2014). Barzilai et al. (2014, 2018, 2020), por exemplo, observaram melhorias em habilidades matemáticas com jogos educativos. De forma semelhante, Byun e Joung (2018), em uma revisão de 17 estudos com estudantes da educação básica K-12, confirmaram impactos positivos de jogos de matemática no desempenho acadêmico. Hao-Jan e Hsiao-Ling (2020) destacaram o jogo *Slave Trade* como eficaz no ensino de vocabulário e história, ao proporcionar um contexto repetitivo e imersivo para estudantes de inglês como segunda língua.

Esses achados reforçam que jogos educacionais podem complementar avaliações formativas e somativas, oferecendo dados sobre o progresso dos estudantes (OLIVEIRA et al., 2022). Avaliar o conteúdo por meio da *gameplay* permite compreender não só os resultados alcançados pelos estudantes, mas também o processo de aprendizagem, revelando dificuldades e avanços conceituais.

O cruzamento dos dados curriculares com aspectos como engajamento, cognição

e desempenho motor amplia os *insights* sobre o processo de aprendizagem, favorecendo o aprimoramento do *game design* e das práticas pedagógicas. Além disso, a comparação desses dados com avaliações tradicionais também ajuda a validar os jogos como instrumentos eficazes de ensino.

Esse aspecto é central nesta pesquisa por seu alinhamento com o objetivo educacional dos jogos: avaliar o quanto os estudantes aprendem sobre o conteúdo. A análise do desempenho nos desafios curriculares indica o nível de compreensão e possíveis intervenções pedagógicas.

5.2 Seleção de dados comuns a jogos educacionais

Após a definição dos aspectos a serem mensurados, foi necessário criar um corpo de conhecimento sobre dados comuns aos jogos educacionais. Esses dados serviram como base para desenvolver métricas capazes de avaliar os aspectos selecionados. Durante essa definição, realizou-se também uma análise exploratória empírica, observando relações entre os dados disponíveis para gerar *insights* sobre os aspectos de aprendizagem.

A definição de dados comuns foi importante para criar métricas potencialmente transferíveis entre jogos educacionais com instrumentação e características compatíveis. Portanto, tais dados deveriam permitir análises relacionadas a indícios de aprendizagem. A Tabela 6 apresenta esses dados, suas descrições, possíveis *insights* e respectivas fontes.

Em alguns casos, a fonte corresponde ao próprio *GLBoard*, que já realiza a captura desses dados, ou a estudos que destacam direta ou indiretamente sua relevância. Os dados foram estruturados segundo o modelo do *GLBoard*, incluindo dados demográficos, dados de *gameplay*, fases e sessões, uma organização comum para representar a interação dos estudantes em diferentes jogos.

Nem todos os dados mapeados foram utilizados nas métricas atuais, porém podem ser explorados na construção de métricas futuras.

Após a etapa de mapeamento, alguns dos dados mapeados que foram utilizados para o desenvolvimento das métricas foram integrados ao *GLBoard*. Essa integração está sendo apresentada na Seção 5.4. A integração foi feita para que os dados pudessem ser

Tabela 6 – Tabela de dados comuns a jogos educacionais.

Categoria	Dado	Descrição	Insights	Fonte
Demográficos	Data de Nascimento	Data de nascimento do jogador.	Permite realizar análises de faixas etárias, identificando a distribuição de idade entre os jogadores.	GLBoard
	Gênero	Gênero do jogador.	Ajuda a realizar análises de gênero, permitindo a identificação de possíveis disparidades na participação e desempenho.	GLBoard
	Localização	Localização geográfica do jogador.	Fornece insights sobre o alcance geográfico do jogo, permitindo análises de desempenho regional.	(WU; LIU; SHUKLA, 2020) (GUARDIA; ARRIGO; GIUSEPPE, 2012)
	Nível de Escolaridade	Nível de escolaridade atual do jogador.	Correlaciona o desempenho no jogo com o nível educacional, identificando padrões de aprendizado relacionados à escolaridade.	(FUSTER-GUILLÓ et al., 2019) (MARIN-VEGA et al., 2020)
Fase	Fases Desbloqueadas	Quantidade de fases desbloqueadas pelo jogador.	Visualiza o estado da jornada do jogador, ajudando a entender seu progresso geral no jogo.	GLBoard
	Tempo em Jogo	Tempo que o jogador fica logado no jogo, interagindo.	Analisa o engajamento e performance do jogador, identificando o tempo de interação com o jogo.	GLBoard
	Último Login	Data do último login do jogador.	Útil para análises de engajamento, ajudando a identificar a frequência de uso do jogo.	GLBoard
	Pontuação	Pontuação total obtida pelo jogador, como "Estrelas" obtidas em uma fase ou desafio, dependendo do contexto do jogo.	Avalia o desempenho geral do jogador, permitindo comparações com outros jogadores.	GLBoard
	Caminho Ideal	Melhor sequência de ações ou decisões para completar uma fase ou desafio.	Identifica a eficiência do jogador, comparando suas ações com a estratégia mais eficiente ou ideal.	(MELO; PIRES; FREITAS, 2019) (GOUWS; BRADSHAW; WENTWORTH, 2013)
Sessão de Fase	Conclusão	Conclusão da tentativa: (Derrota, Vitória, Abandono). Considera-se desistência o ato de o estudante fechar o jogo ou sair da fase após iniciar uma tentativa em uma fase.	Identifica em quais fases ou desafios os estudantes têm bons resultados e onde estão desinteressados ou frustrados.	GLBoard
	Caminho	Caminhos ou ações que o jogador executou dentro de uma fase para atingir seu objetivo.	Visualiza as decisões do jogador, identificando estratégias mais comuns e áreas de falha ou sucesso. Podendo cruzar com o "Caminho Ideal" na fase para identificar o quão assertivo o jogador foi.	GLBoard
	Data e Hora de Início	Tempo que o estudante iniciou a sessão de uma fase.	Permite avaliar a agilidade e eficiência do estudante ao completar a fase, cruzando com dados de caminhos e estratégias.	GLBoard
	Data e Hora de Finalização	Tempo que o estudante finalizou a sessão de uma fase.		GLBoard
	Pedidos de Ajuda	Ajudas ou dicas solicitadas pelo jogador: alguns jogos possuem um sistema de dicas que auxilia os estudantes a concluir determinados desafios.	Identifica pontos de dificuldade, além de permitir visualizar se a ajuda disponibilizada realmente auxiliou o jogador a progredir.	(INVENTADO et al., 2018) (CONATI; JAKES; MUIR, 2013)
	Tempo em Desafios	Tempo gasto em cada tarefa ou desafio específico dentro do jogo.	Avalia a complexidade e dificuldade das tarefas, indicando áreas que podem precisar de ajustes.	GLBoard
	Feedback Recebido	Tipo e quantidade de feedbacks fornecidos ao jogador.	Avalia como o feedback afeta o aprendizado e o progresso do jogador.	(BANG; LI; FLYNN, 2023) (DELACRUZ, 2012)
	Nível de Dificuldade	Nível de dificuldade escolhido pelo jogador (fácil, médio, difícil).	Permite analisar como diferentes níveis de dificuldade dos jogos impactam o aprendizado, observando, por exemplo, se o modo difícil do jogo resultou em melhores desempenhos ao longo da gameplay.	(KIM; RUIPÉREZ-VALIENTE, 2020) (GRAPPIOLO et al., 2011)
	Ações Colaborativas	Interações do jogador com outros jogadores, como trabalho em equipe.	Avalia habilidades sociais e de colaboração, importantes em muitos jogos educacionais.	(MARKEY; LEEDER, 2011) (LI; CHEN; DENG, 2024a)
	Desafios	Informações sobre desafios lógicos e mecânicos tentados e finalizados na sessão. Exemplo: derrotar um inimigo (mecânico), resolver um puzzle (lógico), responder uma pergunta (conteúdo).	Entende quais tipos de desafios os jogadores enfrentam, os resultados obtidos e como lidam com diferentes obstáculos.	(CARDOSO et al., 2019)
	Itens Coletados	Todos os itens coletados durante a sessão.	Avalia o nível de exploração e interação do jogador com o ambiente do jogo durante uma sessão específica.	(HENDRIX et al., 2018) (LI; CHEN; DENG, 2024a)
	Nível do Jogador	Nível atual do jogador durante a sessão.	Identifica o quanto o jogador precisou progredir para passar de uma fase e se outros jogadores conseguiram chegar à mesma fase com um nível menor ou maior.	(ZAGAL; ALTIZER, 2014) (CARSTENS DOTIR; KLEINMAN; EL-NASR, 2019)
	Conquistas Coletadas	Conquistas ou troféus obtidos durante a sessão.	Indica níveis de dedicação e engajamento do jogador.	(MARKEY; LEEDER, 2011) (LI; CHEN; DENG, 2024a)

capturados nos jogos educacionais, permitindo a aplicação das métricas desenvolvidas.

5.3 Métricas propostas

Finalizada a etapa de seleção de aspectos e dados comuns a jogos educacionais, foram propostas as métricas com base no que foi levantado. Nas seções a seguir, serão descritas as métricas desenvolvidas, ao todo oito, abordando o aspecto de aprendizagem alvo, sua relevância, os dados usados na formulação, a fórmula e um exemplo de aplicação da métrica, quando necessário. As métricas são agrupadas em quatro dimensões de aprendizagem, como dito na Seção 5.1: Engajamento, Cognitiva, Motora e Unidade Curricular. Para cada uma dessas dimensões, foi definido um conjunto específico de

métricas.

A Figura 12 sintetiza a relação entre (i) as dimensões/aspectos de aprendizagem, (ii) as métricas definidas em cada dimensão e (iii) as categorias de dados necessárias para o seu cálculo, funcionando como um mapa de dependências. A interpretação pode ser feita da esquerda para a direita: a partir de uma dimensão (aspecto de aprendizagem), identifica-se o conjunto de métricas associado (Engajamento, Desenvolvimento Cognitivo, Desenvolvimento Motor e Unidade Curricular) e, em seguida, observam-se, na base da figura, as variáveis de registros do jogo que sustentam o cálculo de cada métrica. A estrutura apresentada na figura 12 é descrita a seguir.



Figura 12 – Métricas propostas, aspectos de aprendizagem e dados utilizados.

A escolha dessas métricas foi baseada em suas relações diretas com jogos educa-

cionais. Como mostrado na Tabela 6, já existem investigações abordando esses aspectos de aprendizagem em jogos educacionais e oferecendo exemplos de como podem ser avaliados (HOOKHAM; NESBITT, 2019; LI et al., 2022; MARTINEZ; MENÉNDEZ-MENÉNDEZ; BUSTILLO, 2022). Entretanto, essas investigações não visam a mensuração desses aspectos de maneira abrangente, mas sim de forma *ad hoc* para cada jogo educacional. No contexto do aspecto “Desenvolvimento Cognitivo”, foram elaboradas três métricas, as quais serão demonstradas na seção subsequente.

Para facilitar a compreensão e a reprodutibilidade do processo de construção das métricas, cada métrica proposta é apresentada seguindo uma estrutura padronizada em três partes: **(a) Mecanismo de ação**, que descreve o princípio de funcionamento da métrica e o que ela busca capturar no *gameplay*; **(b) Formulação**, que explicita as variáveis consideradas, a equação de cálculo e a interpretação do resultado; e **(c) Exemplo de aplicação**, que ilustra, com dados fictícios ou reais do jogo, como a métrica é calculada em uma fase e/ou por tentativa/sessão.

5.3.1 Métricas para mensurar desenvolvimento cognitivo

Esta subseção apresenta cada uma das métricas que compõem o desenvolvimento cognitivo, que são: “Assertividade”, “Habilidade Lógica” e “Flexibilidade Cognitiva”.

5.3.1.1 Assertividade

Uma das habilidades relacionadas ao *Desenvolvimento Cognitivo* é a capacidade de buscar soluções eficientes, considerando não apenas a quantidade de ações ou o custo envolvido na resolução de um problema, mas também o processo de exploração e tomada de decisão ao longo da atividade (SHI; QU, 2022). Para avaliar essa habilidade, foi criada a métrica *Assertividade*, que mede a eficácia com que os estudantes escolhem caminhos e solucionam desafios em jogos educacionais. Diferentemente da simples contagem de acertos e erros, essa métrica considera o processo adotado pelo estudante durante a resolução dos desafios, incluindo suas tentativas, escolhas e estratégias. O objetivo é

fornecer uma análise do progresso cognitivo, destacando aspectos como planejamento, tomada de decisão e uso de estratégias ao longo da resolução das atividades.

Para os educadores, o cruzamento da métrica *Assertividade* com outras, como o tempo de resolução, permite a obtenção de *insights* sobre as estratégias de decisão adotadas pelos estudantes (LIANG, 2024). Por exemplo, ao analisar o tempo necessário para chegar a uma solução eficaz, é possível identificar padrões de comportamento durante a resolução dos desafios. Um estudante metódico pode levar mais tempo, mas ser mais assertivo, enquanto outro, que adota estratégias de tentativa e erro, pode chegar mais rapidamente à solução, porém com menor grau de assertividade, por depender de múltiplas tentativas. Essas informações podem auxiliar os educadores na compreensão das estratégias utilizadas e na proposição de atividades que reforcem essa habilidade.

Para os desenvolvedores de jogos educacionais, a métrica *Assertividade* oferece subsídios para ajustes no *game design* e na complexidade dos desafios. Caso os dados indiquem escolhas frequentes de soluções ineficazes ou longos tempos para resolução, pode ser necessária a revisão do nível de dificuldade do jogo. Ajustes adequados permitem equilibrar o desafio, estimulando o pensamento crítico sem tornar a experiência frustrante ou trivial.

a) Mecanismo de ação

O mecanismo de ação da métrica “Assertividade” centra-se em avaliar as escolhas e trajetórias tomadas pelos estudantes em jogos educacionais, comparando-as com uma ou mais soluções ideais, aqui denominados de “Caminhos Perfeitos”. Esses caminhos representam diferentes sequências de ações que conduzem à forma mais eficaz de alcançar os objetivos do jogo, fundamentando-se nas mecânicas e nas soluções mais eficientes possíveis. A métrica considera o “Caminho Perfeito” que estiver mais próximo do “Caminho do Jogador”, que reflete as decisões reais tomadas pelo estudante durante a jogatina, permitindo uma análise objetiva da eficiência das soluções dos jogadores frente aos desafios propostos.

b) Formulação

A fórmula a seguir representa a métrica “Assertividade”, que pode ser mensurada por fase e por tentativas/sessões dentro das fases do jogo. A métrica permite múltiplos caminhos ideais, ajustando-se ao caminho perfeito mais próximo do realizado pelo estudante:

A Fórmula 5.1 ilustra como a métrica “Assertividade” é calculada, considerando:

- C_P : caminho perfeito;
- C_{Ps} : conjunto de caminhos perfeitos;
- C_E : caminho do estudante;
- DistLev: distância de Levenshtein;

A fórmula para a métrica "Assertividade" é dada por:

$$A = 1 - \frac{\min_{C_P \in C_{Ps}} \text{DistLev}(C_E, C_P)}{\max(|C_P|, |C_E|)} \quad (5.1)$$

- A: representa a assertividade em um intervalo de 0 a 1, sendo 1 um indicativo de alto valor da métrica.
- *Caminho do Estudante*: ações ou caminhos tomados pelo estudante durante a *game-play* em uma determinada fase.
- *Caminhos Perfeitos*: conjunto de sequências ideais que o estudante pode seguir para alcançar o objetivo de forma otimizada, considerando múltiplas alternativas corretas.
- **Distância Levenshtein (LEVENSHTEIN et al., 1966)**: uma métrica que calcula o número mínimo de operações de inserção, deleção ou substituição necessárias para transformar uma sequência em outra, considerando não apenas o tamanho das sequências, mas também a ordem e o conteúdo específico dos elementos dentro delas. Recebe como entrada duas sequências (*arrays*) e retorna um valor inteiro que representa a quantidade mínima de alterações necessárias para que as duas sequências se tornem exatamente iguais.

- **min:** operador que seleciona o menor valor da *Distância Levenshtein* entre o *Caminho do Estudante* e todos os *Caminhos Perfeitos*, garantindo que o caminho ideal mais próximo seja considerado.
- $\max(|C_P|, |C_E|)$: determina o maior valor entre o comprimento do caminho ideal e o do caminho do estudante, servindo como denominador para normalizar a distância de *Levenshtein* em relação ao comprimento do percurso.

Essa métrica apresenta uma limitação em jogos com geração procedural, isto é, jogos cujos elementos são gerados aleatoriamente por algoritmos, em vez de serem definidos de forma fixa pelos designers. A impossibilidade de definir previamente um caminho perfeito, devido a variabilidade dos elementos a cada tentativa, compromete o cálculo da *assertividade*. Como mostra a fórmula, a assertividade compara o caminho realizado pelo jogador (C_E) com um ou mais caminhos perfeitos (C_P) usando a distância de *Levenshtein*. Sem caminhos ideais definidos, não há base confiável para essa comparação. Além disso, como a *Flexibilidade Cognitiva*, métrica apresentada a seguir, é calculada a partir da assertividade, sua aplicação também se torna inviável nessas condições. Ambas as métricas, portanto, exigem cenários com trajetórias conhecidas e comparáveis.

c) Exemplo de aplicação

- **Caminho Perfeito:** o desenvolvedor para a Fase 1 informou que existe um caminho ideal que seria ["P1", "P4", "P5", "P6"], sendo P interpretado como "Plataforma".
- **Caminhos do Estudante:** sequências de ações realizadas pelo estudante em suas tentativas de completar a fase. Os caminhos dos estudantes nas tentativas foram:
 - Tentativa 1: ["P1", "P3", "P2", "P4", "P5", "P6"];
 - Tentativa 2: ["P1", "P4", "P2", "P5", "P6"];
 - Tentativa 3: ["P1", "P2", "P4", "P5", "P6"].

Para calcular a assertividade de um estudante em uma fase específica, concatena-se, inicialmente, todas as tentativas realizadas pelo estudante em um único *array*, sendo

este o “Caminho do Estudante”, formando um conjunto que pode ser exemplificado por: [“P1”, “P3”, “P2”, “P4”, “P5”, “P6”, “P1”, “P4”, “P2”, “P5”, “P6”, “P1”, “P2”, “P4”, “P5”, “P6”].

Considerando o “Caminho Perfeito” e o “Caminho do Estudante”, o valor da assertividade para a fase seria:

$$A = 1 - \frac{12}{\max(4, 16)} = 1 - \frac{12}{16} = 0.25$$

A assertividade pode também ser calculada individualmente para cada tentativa. Nesse cenário, o “Caminho do Estudante” utilizado no cálculo da assertividade refere-se somente ao trajeto realizado durante uma única tentativa.

5.3.1.2 Flexibilidade cognitiva

Outra habilidade associada ao *Desenvolvimento Cognitivo* é a capacidade de alternar abordagens e aplicar conhecimentos em novos contextos, refletindo adaptação e dinamismo na resolução de problemas (CARVALHO, 2011). Para mensurar essa habilidade, foi desenvolvida a métrica *Flexibilidade Cognitiva*, que avalia o grau de adaptação dos estudantes ao enfrentarem desafios variados em jogos educacionais. Essa métrica permite identificar se os alunos estão explorando diferentes estratégias diante de problemas complexos, evidenciando sua capacidade de ajustar abordagens diante de grandes desafios.

Para os educadores, a análise dessa métrica possibilita identificar padrões de comportamento, como a repetição de abordagens ineficazes, que podem indicar resistência à mudança ou limitação no repertório de estratégias utilizadas pelos estudantes. Esses *insights* viabilizam intervenções personalizadas que incentivem a experimentação e a adaptação.

Para os desenvolvedores, a métrica oferece subsídios para aprimorar o *game design*, promovendo a criação de cenários que exijam múltiplas abordagens. Situações em que soluções anteriores deixam de ser eficazes podem estimular o pensamento crítico e a flexibilidade dos jogadores.

a) Mecanismo de ação

Essa métrica funciona através da verificação de diferentes ações realizadas pelo jogador em cada tentativa dentro de uma determinada fase, analisando como essas ações variam para adaptar-se a novos desafios e situações. De forma geral, a métrica é calculada buscando compreender se houve alterações significativas entre a tentativa de solução anterior e a atual, indicando possíveis mudanças de estratégia durante a resolução do desafio.

b) Formulação

Para avaliar a Flexibilidade Cognitiva, utiliza-se a fórmula 5.2, que incorpora a distância de *Levenshtein* para medir as variações entre tentativas distintas de um jogador dentro de uma fase:

Sejam i e j duas tentativas diferentes do jogador.

$$\sum d_{ij} = \text{Levenshtein}(\text{tentativa}_i, \text{tentativa}_j)$$

A distância de *Levenshtein* entre todas as tentativas de uma fase é calculada representando a variação estratégica entre elas. Em seguida, definimos:

$$D_{\text{média}} = \frac{\sum d_{ij}}{N}, \text{ onde } N \text{ é o número total de comparações únicas.}$$

Para um total de T tentativas, o número N de comparações únicas pode ser calculado como:

$$N = \frac{T \times (T - 1)}{2}$$

Assim, $D_{\text{média}}$ representa a média das distâncias de *Levenshtein* calculadas entre todas as combinações únicas de pares de tentativas. A média do comprimento das tentativas, $L_{\text{média}}$, é calculada como a soma dos comprimentos individuais de cada tentativa dividida pelo número total de tentativas realizadas pelo jogador na fase.

$$L_{\text{média}} = \frac{\sum_{i=1}^T |tentativa_i|}{T}$$

Finalmente, a Flexibilidade Cognitiva, M_{exp} , é determinada por:

$$M_{\text{exp}} = \min \left(\frac{D_{\text{média}}}{L_{\text{média}}}, 1 \right) \quad (5.2)$$

- M_{exp} : medida de Flexibilidade Cognitiva, indicando a capacidade do jogador de variar suas estratégias e adaptar-se com base em experiências anteriores. A função $\min$ é utilizada para assegurar que a métrica não exceda 1, normalizando a pontuação dentro de um intervalo entre 0 e 1.
- $\text{Levenshtein}(tentativa_i, tentativa_j)$: calcula a distância de Levenshtein entre as tentativas i e j , quantificando o grau de variação entre elas.
- $D_{\text{média}}$: média das distâncias de Levenshtein, oferecendo uma visão geral da diversidade nas estratégias adotadas pelo jogador ao longo das tentativas.
- N : total de pares únicos de tentativas comparadas, fornecendo a base para o cálculo da média.
- $L_{\text{média}}$: média dos comprimentos das tentativas, utilizada para normalizar a métrica de Flexibilidade Cognitiva.

c) Exemplo de aplicação

Em um jogo educacional, as ações de um estudante em diferentes tentativas são registradas para calcular sua Flexibilidade Cognitiva:

- **Tentativas:**
 - Tentativa 1: ["Ação1", "Ação3", "Ação2"];
 - Tentativa 2: ["Ação1", "Ação2", "Ação4"];
 - Tentativa 3: ["Ação2", "Ação3", "Ação5"].
- **Distâncias de Levenshtein:**

- Distância(1,2) = 2:
 - * “Ação3” → “Ação2”
 - * “Ação2” → “Ação4”
- Distância(1,3) = 2:
 - * “Ação1” → “Ação2”
 - * “Ação2” → “Ação5”
- Distância(2,3) = 3:
 - * “Ação1” → “Ação2”
 - * “Ação2” → “Ação3”
 - * “Ação4” → “Ação5”

- **Cálculos:**

- $D_{\text{média}} = \frac{2+2+3}{3} = 2,333$
- $L_{\text{média}} = \frac{3+3+3}{3} = 3$
- $M_{\text{exp}} = \min\left(\frac{2,333}{3}, 1\right) = 0,777$

- **Interpretação:** O valor 0,77 indica uma capacidade moderadamente alta de adaptação, com variação significativa nas ações entre as tentativas, refletindo bom nível de Flexibilidade Cognitiva.

5.3.1.3 Habilidade lógica

Diferente das métricas de “Assertividade” e “Flexibilidade Cognitiva”, que consideram o conjunto de ações realizadas pelos jogadores, a métrica “Habilidade Lógica” está relacionada aos desafios lógicos resolvidos, ou não, dentro do jogo. Exemplos desses desafios incluem resolver um quebra-cabeça, responder a uma pergunta lógica ou tomar uma decisão lógica relevante. Essa métrica usa esses desafios para avaliar de forma clara o desempenho dos estudantes na lógica, levando em conta quantos desafios lógicos foram acertados ou não, desconsiderando outras ações.

Para os educadores, essa métrica fornece *insights* diretos sobre o desempenho dos estudantes em tarefas que envolvem raciocínio lógico. O valor resultante da métrica pode ser utilizado diretamente para avaliar os alunos nessa habilidade, além de permitir identificar pontos onde os estudantes precisam de ajuda, promovendo intervenções para melhorar as habilidades.

Para os desenvolvedores de jogos, ela orienta o ajuste de dificuldade, visando um equilíbrio que desafie adequadamente os jogadores sem desencorajá-los.

a) Mecanismo de ação

A métrica é aplicada tanto por sessão quanto por fase do jogo. Nessa avaliação, analisa-se o desempenho dos estudantes em desafios lógicos, observando-se a proporção entre os problemas resolvidos e o total de tentativas realizadas na sessão e na fase.

b) Formulação

A habilidade lógica é quantificada para sessões individuais e para a fase completa, conforme a Fórmula 5.3:

$$H_{\text{lógica, sessão/fase}} = \frac{N_{\text{Resolvidos}}^{\text{lógica}}}{N_{\text{TotalTentadas}}^{\text{lógica}}} \quad (5.3)$$

onde:

- $N_{\text{Resolvidos}}^{\text{lógica}}$ representa o número de desafios lógicos resolvidos na sessão/fase.
- $N_{\text{TotalTentadas}}^{\text{lógica}}$ é o número total de tentativas para resolver desafios lógicos na sessão/fase.

c) Exemplo de aplicação

Como exemplo, em um jogo educacional voltado à resolução de quebra-cabeças, considere um estudante que realizou três tentativas (ou sessões) para concluir uma determinada fase:

- **Tentativa 1:** resolveu 1 desafio lógico e tentou resolver 4 (incluindo o que resolveu);
- **Tentativa 2:** resolveu 2 desafios e tentou 4;
- **Tentativa 3:** resolveu 3 desafios e tentou 3.

Nesse exemplo para calcular a métrica de habilidade lógica para cada sessão e para a fase completa, utiliza-se a seguinte fórmula:

- **Para cada seção (tentativa):**

$$H_{\text{lógica, seção}} = \frac{\text{Número de Desafios Resolvidos}}{\text{Número Total de Tentativas na Seção}}$$

$$\text{– Tentativa 1: } H_{\text{lógica, seção 1}} = \frac{1}{4} = 0.25$$

$$\text{– Tentativa 2: } H_{\text{lógica, seção 2}} = \frac{2}{4} = 0.5$$

$$\text{– Tentativa 3: } H_{\text{lógica, seção 3}} = \frac{3}{3} = 1.0$$

- **Para a fase completa:**

$$H_{\text{lógica, fase}} = \frac{\text{Número de Desafios Resolvidos}}{\text{Número Total de Tentativas na Fase}}$$

$$H_{\text{lógica, fase}} = \frac{1 + 2 + 3}{11} = \frac{6}{11} \approx 0.545$$

5.3.2 Métricas para mensurar engajamento

As métricas usadas para avaliar o engajamento incluem “Exploração”, “Tempo de *Gameplay*” e “Tentativas até o sucesso” que serão descritas a seguir:

5.3.2.1 Exploração

A exploração é um elemento comum em jogos, frequentemente expressa por salas secretas, itens e locais opcionais, diálogos extras e outros elementos acessíveis durante

a navegação. Em jogos educacionais, esse processo pode contribuir tanto para a aquisição de novos conhecimentos quanto para o aumento do engajamento no processo de aprendizagem.

A métrica *Exploração* foi criada para mensurar a variedade e a extensão dos itens e locais acessados pelos jogadores, refletindo seu envolvimento em descobrir e aprender de forma ativa.

Para os educadores, essa métrica permite identificar áreas de maior interesse ou de baixa interação, orientando intervenções pedagógicas mais precisas. Por exemplo, áreas frequentemente ignoradas podem indicar conteúdos pouco atrativos ou de difícil acesso, sugerindo a necessidade de revisão ou suporte adicional aos estudantes.

Para os desenvolvedores, a métrica oferece subsídios para ajustar o *game design*, aprimorando a apresentação e integração de elementos opcionais. Isso estimula uma exploração mais ativa e amplia a motivação dos jogadores em interagir com todo o conteúdo disponível.

a) Mecanismo de ação

A métrica envolve o cálculo da proporção entre o número total de itens opcionais únicos explorados pelo jogador e o total de itens/locais únicos possíveis na fase do jogo. Quanto maior for essa proporção, maior será a pontuação de exploração do jogador, indicando um maior interesse e disposição para explorar novos desafios e conteúdos.

b) Formulação

A métrica de Exploração é definida por:

$$E = \frac{|S|}{\text{Número Total de Itens/Locais/Conquistas Únicas Possíveis na Fase}} \quad (5.4)$$

Seja $S_1, S_2, \dots, S_n$ os conjuntos de itens/locais explorados em cada tentativa. A união desses conjuntos nos dá o total de itens únicos explorados:

$$S = S_1 \cup S_2 \cup \dots \cup S_n.$$

Na Equação 5.4:

- $S_1, S_2, \dots, S_n$ representam os conjuntos de itens/locais explorados em cada tentativa.
- $S = S_1 \cup S_2 \cup \dots \cup S_n$ representa o conjunto total de itens únicos explorados, obtido pela união dos conjuntos de todas as tentativas (a união de conjuntos elimina duplicatas).
- $|S|$ representa a cardinalidade do conjunto S , ou seja, o número de elementos no conjunto S .
- E representa a métrica de Exploração.

c) Exemplo de aplicação

Como exemplo, considere um jogo no qual existem doze tipos diferentes de conquistas, itens, caminhos e outros elementos que podem ser explorados. Em cada tentativa, o jogador pode coletar e/ou conquistar esses itens. Considere o seguinte exemplo com três tentativas:

- **Tentativa 1:** o jogador coleta os itens A, B, C ;
- **Tentativa 2:** o jogador conquista as insígnias D, E, F ;
- **Tentativa 3:** o jogador coleta os itens C, G, H e conquista as insígnias I, J .

Nesse exemplo, para calcular a métrica de Exploração (E), são realizados os passos abaixo:

1. Primeiramente, definem-se os conjuntos para cada tentativa:

$$S_1 = \{A, B, C\} \quad (\text{Itens coletados na tentativa 1});$$

$$S_2 = \{D, E, F\} \quad (\text{Itens coletados na tentativa 2});$$

$$S_3 = \{C, G, H, I, J\} \quad (\text{Itens coletados na tentativa 3}).$$

2. Em seguida, calcula-se o conjunto total de itens únicos explorados (união dos conjuntos):

$$\begin{aligned} S &= S_1 \cup S_2 \cup S_3 \\ &= \{A, B, C, D, E, F, G, H, I, J\} \quad (\text{Cardinalidade: } 10) \end{aligned}$$

Aplicando a fórmula:

$$E = \frac{|S|}{\text{Número Total de Itens/Locais/Conquistas Únicas Possíveis na Fase}} = \frac{10}{12} \approx 0.83$$

Portanto, a métrica de “Exploração” para essas tentativas é aproximadamente 0.83, indicando que o jogador explorou cerca de 83% dos tipos únicos de conquistas, itens ou caminhos disponíveis no jogo.

5.3.2.2 Tempo de *Gameplay*

Dados de *timestamps* em jogos digitais permitem identificar o tempo dedicado pelo jogador ao jogo, à conclusão de tarefas ou à tomada de decisões. Patriarca et al. (2019) utilizaram essa métrica em um jogo de simulação para treinamento em segurança industrial, mensurando o tempo de reação a emergências. Respostas lentas indicavam a necessidade de ajustes no jogo ou nas atividades formativas.

Dutra Júnior et al. (2021) analisaram *logs* temporais em um jogo sobre plantas industriais, correlacionando o tempo de configuração de componentes ao desempenho nas fases.

Neste trabalho, a métrica *Tempo de Gameplay* mensura tanto o engajamento quanto o tempo para concluir jogos, fases, tentativas ou desafios. Embora amplamente utilizada na análise de jogos, aqui busca-se extrair *insights* adicionais, indo além da simples avaliação do envolvimento.

Para educadores, a métrica indica o grau de participação dos estudantes e possíveis necessidades de mediação. Quando associada ao desempenho, permite identificar padrões comportamentais. Para desenvolvedores, revela a duração média das sessões e os pontos de abandono, orientando ajustes no *design* e no equilíbrio da dificuldade.

a) Mecanismo de ação

Diferentemente de métricas normalizadas (como de 0 a 1), o “Tempo de *Gameplay*” é variável e depende do contexto do jogo e do perfil do jogador (WALSH; BOKHOVE, 2021). Sua aplicação permite o cruzamento com outras métricas, gerando análises mais abrangentes. Por exemplo, associada à “Assertividade”, pode indicar perfis estratégicos mais cautelosos, nos quais há maior ponderação antes das ações.

Essa métrica é aplicada em quatro níveis:

- **Jogo completo:** mede o tempo total de interação com o jogo, do início ao encerramento da aplicação.
- **Fase:** contabiliza a duração de cada fase, possibilitando observar a adaptação do jogador aos desafios propostos.
- **Tentativas:** registra o tempo gasto em cada tentativa dentro de uma fase, revelando estratégias e persistência.
- **Desafios:** mede o tempo necessário para superar desafios pontuais, com base nos registros de início e fim de cada interação.

A aplicação da métrica baseia-se na simples marcação dos tempos de início e término (*datetime*), não exigindo fórmulas complexas. Esses registros permitem calcular com precisão a duração das interações e apoiar análises sobre esforço, estratégias adotadas e possíveis dificuldades enfrentadas pelos jogadores.

5.3.2.3 Tentativas até o sucesso

Com propósito semelhante à métrica “Tempo de *Gameplay*”, amplamente discutida na literatura, esta métrica busca avaliar o engajamento e a persistência dos estudantes diante de desafios propostos pelo jogo. A repetição de tentativas pode indicar esforço contínuo e disposição para superar dificuldades, fornecendo *insights* sobre estratégias adotadas ao longo da interação com o jogo.

Do ponto de vista educacional, a métrica pode auxiliar na identificação de estudantes mais engajados ou que apresentam dificuldades, permitindo intervenções

direcionadas. Já para desenvolvedores, fornece indícios sobre possíveis desajustes no equilíbrio do jogo, como níveis com dificuldade desproporcional ou ausência de orientações suficientes, caso haja elevado número de tentativas sem sucesso.

a) Mecanismo de ação

Sem um intervalo fixo, a métrica “Tentativas até o sucesso” depende do contexto de jogo e do comportamento individual dos jogadores. Pode ser combinada com outras métricas, como “Flexibilidade Cognitiva”, permitindo identificar padrões. Por exemplo, múltiplas tentativas acompanhadas de baixa flexibilidade podem indicar dificuldade na adaptação de estratégias.

A métrica se aplica a dois níveis:

- **Fase:** contabiliza o número de tentativas necessárias para concluir uma fase, permitindo avaliar o grau de persistência do jogador.
- **Desafios:** registra a quantidade de tentativas para superar desafios, revelando pontos críticos no progresso da fase.

Sua implementação é direta, baseada na contagem sequencial das tentativas (primeira, segunda, etc.) até o sucesso. Cada tentativa é registrada como um evento único. Essa simplicidade operacional não compromete sua capacidade de revelar padrões de comportamento relevantes ao longo da jogabilidade.

5.3.3 Métricas para mensurar desenvolvimento motor

Foi elaborada uma única métrica para mensurar o desenvolvimento motor, denominada “Habilidade Mecânica”. Essa delimitação decorre da complexidade inerente à avaliação desse aspecto. No contexto de jogos educacionais, a habilidade motora é frequentemente mensurada por meio de sensores de dispositivos móveis, bem como por dados de interação, como cliques, pressão exercida em botões e outros registros quantitativos (LANDER et al., 2022).

Em vez de utilizar sensores, optou-se pelo desenvolvimento de uma métrica fundamentada em elementos intrínsecos à dinâmica do jogo. Tal decisão encontra respaldo na literatura, uma vez que a pesquisa de Moron et al. (2022) evidencia a existência de diferentes elementos em jogos educacionais associados ao desenvolvimento motor, indicando a viabilidade de abordagens baseadas em mecânicas de jogo.

5.3.3.1 Habilidade mecânica

A “Habilidade Mecânica” é uma métrica usada para avaliar quão bem os jogadores superam desafios mecânicos em jogos educacionais. Esses desafios incluem atividades que demandam coordenação motora e precisão, como reações rápidas, saltos entre plataformas, lançamentos de projéteis e confrontos com chefes, todos ligados ao desenvolvimento motor. Avaliar essa habilidade ajuda os educadores a entender como os alunos lidam com esses elementos mecânicos e também permite aos desenvolvedores de jogos ajustar a dificuldade dessas atividades. A métrica não só mede o progresso motor dos estudantes, mas também fornece dados valiosos para os desenvolvedores ajustarem a dificuldade dos desafios, garantindo que os jogos sejam acessíveis e ao mesmo tempo estimulantes.

a) Mecanismo de ação

Essa métrica utiliza uma abordagem semelhante à “Habilidade Lógica”, mas adapta-se para avaliar competências mecânicas, ao invés de lógicas. Enquanto a “Habilidade Lógica” foca em *puzzles* e decisões estratégicas, a “Habilidade Mecânica” concentra-se em superar desafios físicos e de coordenação, como: derrotar inimigos; saltar sobre plataformas em movimento; utilizar habilidades especiais de movimentação, predominantemente motoras, envolvendo tempo de reação, precisão nos controles e coordenação motora fina; e desviar de projéteis.

Os desafios mecânicos variam conforme o tipo de jogo: jogos de plataforma, por exemplo, podem exigir saltos precisos e esquivar de inimigos, enquanto jogos de nave

espacial podem demandar ações como desviar de projéteis e destruir inimigos.

A formulação dessa métrica é similar à da “Habilidade Lógica”, aplicando-se de maneira correspondente às necessidades específicas dos desafios mecânicos presentes no jogo.

5.3.4 Métricas para mensurar unidade curricular

Assim como no aspecto “Desenvolvimento Motor”, “unidade curricular” também teve apenas uma métrica relacionada. Ela foi desenvolvida de forma geral com o propósito de permitir a compreensão das tarefas relacionadas ao conteúdo do jogo que foram resolvidas pelos jogadores, e a partir disso, avaliar o desempenho nesse aspecto.

5.3.4.1 Pró-eficiência curricular

A “Métrica de Pró-Eficiência Curricular” é projetada para mensurar a habilidade dos estudantes em solucionar desafios diretamente vinculados ao conteúdo curricular. Em um jogo educacional voltado ao ensino de português, por exemplo, esses desafios podem incluir *quizes* ou *puzzles*, que tenham como *background* o conteúdo curricular. Semelhante à métrica de “Habilidade Lógica” e “Habilidade Mecânica”, essa abordagem avalia a capacidade dos alunos de aplicar conhecimentos e habilidades curriculares em contextos práticos fornecidos pelos jogos educacionais. Tal métrica é relevante tanto para educadores, quanto para desenvolvedores de jogos pois fornece dados sobre o quanto bem os conteúdos curriculares estão sendo integrados e assimilados através dos jogos, orientando ajustes para otimizar o aprendizado e torná-lo mais significativo e efetivo.

a) Mecanismo de ação

Essa métrica adota uma abordagem semelhante às de “Habilidade Lógica” e “Habilidade Mecânica”, realizando duas análises complementares: por sessões e por fases. A

distinção da métrica reside na definição de desafios curriculares, ou seja, tarefas que exigem respostas diretamente relacionadas a conteúdos ou a resolução de problemas que incorporam conhecimentos curriculares.

5.4 Adaptação da arquitetura de dados do *GLBoard*

Todas as métricas desenvolvidas dependem de dados de entrada, que são provenientes da interação dos estudantes com os jogos. No entanto, foi observado que alguns desses dados de interação não estavam sendo capturados pelo *GLBoard*, o que poderia comprometer a aplicação das métricas. Diante disso, foi necessário ajustar a arquitetura de dados do *GLBoard*. A versão original, mencionada na Seção 3.1, estava limitada às informações voltadas para análise de *game design*. Assim, foram feitas modificações na arquitetura para integrar os dados exigidos por cada métrica. A seguir estão descritos os dados necessários para cada métrica:

- **Exploração:** foi criada uma estrutura para capturar itens e colecionáveis opcionais coletados, bem como locais opcionais acessados dentro do jogo.
- **Tentativas até o sucesso:** a arquitetura existente já permitia identificar o número de tentativas necessárias para concluir uma fase ou desafio.
- **Tempo de *gameplay*:** inseriram-se dados temporais em diversas estruturas para registrar o tempo total de jogo, o tempo para concluir uma fase e um desafio.
- **Assertividade:** a trajetória do jogador já era capturada, mas adicionou-se um atributo referente ao melhor caminho ou conjunto de ações possíveis em uma fase, preenchido pelo desenvolvedor.
- **Flexibilidade Cognitiva:** não houve necessidade de adicionar novos atributos, apenas o caminho do jogador foi utilizado.
- **Habilidade lógica, mecânica e proficiência curricular:** requereu-se uma estrutura que permitisse identificar os desafios lógicos, mecânicos ou curriculares enfrentados na sessão, indicando se foram tentados e se houve sucesso.

A Figura 13 ilustra a nova estrutura de dados elaborada e integrada ao *GLBoard* para capturar as informações necessárias para a aplicação das métricas. As entidades e atributos em vermelho representam os novos atributos inseridos.

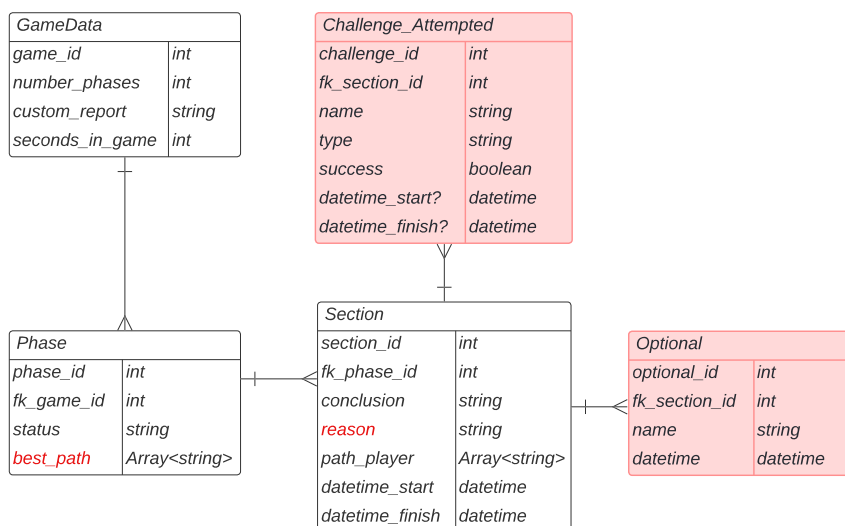


Figura 13 – Estrutura de dados remodelada do sistema *GLBoard*.

Após a implementação dos ajustes na arquitetura de dados do *GLBoard*, foi necessário que os desenvolvedores dos jogos atualizassem o pacote mencionado na Seção 3.1.5. Esse pacote, responsável por coletar e enviar os dados ao banco de dados, agora pode capturar os novos dados adicionados à arquitetura. Após essa atualização, o pacote foi integrado aos jogos.

É importante destacar que o pacote, por si só, não coleta dados automaticamente; ele apenas simplifica e organiza o envio dos dados para o *GLBoard*. Portanto, é necessário um esforço para adaptar os dados à estrutura.

O processo de integração do pacote com o jogo envolve mapear os dados do jogo e como eles podem ser encaixados na arquitetura de dados. Esse processo inclui discussões sobre limitações das métricas e desafios de adequação.

5.5 Questionário de autoavaliação dos estudantes

Além da elaboração das métricas, foi criado um questionário para compreender como os estudantes se autoavaliam em relação ao seu desempenho nos aspectos de aprendizagem, os mesmos que foram mensurados pelas métricas. Esse questionário foi aplicado

ao final de cada interação dos estudantes com um conjunto de jogos previamente selecionados. O objetivo foi verificar se os valores mostrados pelas métricas estão de acordo com a autoavaliação dos estudantes, identificando similaridades ou discrepâncias.

Essa estratégia de comparação entre autoavaliação e métricas computacionais foi adotada com base em abordagens utilizadas em estudos de *learning analytics*, como apresentado por Barba, Oliveira e English (2025), no qual a autoavaliação dos estudantes é utilizada como um meio complementar para validar indicadores comportamentais extraídos automaticamente dos sistemas. Nesse tipo de abordagem, a validação é realizada por meio da comparação direta entre a autoavaliação dos estudantes e os valores gerados pelo sistema, permitindo analisar a convergência ou divergência entre a percepção dos estudantes e os comportamentos observados. Dessa forma, essa estratégia possibilita verificar se as métricas propostas representam adequadamente os aspectos de aprendizagem analisados.

Embora o estudo de Barba, Oliveira e English (2025) tenha sido desenvolvido em ambientes de sistemas de gestão de aprendizagem, a estratégia metodológica de comparação entre autoavaliação e métricas foi adaptada neste estudo para o contexto de jogos digitais educacionais.

As perguntas do questionário incluíram questões demográficas e relacionadas ao desempenho dos estudantes em determinados aspectos, utilizando uma escala *Likert*. Na Tabela 7, é possível visualizar as perguntas e os respectivos aspectos a que se referem.

Pergunta	Aspecto de Aprendizagem
Quando enfrentei dificuldades, busquei soluções alternativas	Flexibilidade Cognitiva
Eu tomei decisões que efetivamente resolveram os desafios apresentados no jogo	Assertividade
Eu precisei me esforçar consistentemente para progredir nas fases do jogo	Tentativas Por Fase
Não tive dificuldade para controlar o personagem e interagir com os desafios mecânicos do jogo, como por exemplo: desviar dos inimigos	Habilidade Mecânica
Sempre busquei tomar as melhores decisões lógicas	Habilidade Lógica

Tabela 7 – Perguntas do questionário para autoavaliação dos estudantes sobre os aspectos de aprendizagem.

No próximo capítulo, serão discutidas as dificuldades de adequação de dois

jogos à arquitetura de dados necessária para a aplicação das métricas. Em seguida, serão analisadas e discutidas a interpretação do desempenho dos estudantes segundo as métricas e a autoavaliação do desempenho dos estudantes.

5.6 Processo de validação das métricas

A validação das métricas propostas para avaliar aspectos de aprendizagem em jogos educacionais foi realizada por meio de um questionário aplicado a três grupos de especialistas: professores, analistas de dados e psicometristas. Esses grupos foram escolhidos para contemplar diferentes perspectivas sobre a clareza, coerência e aplicabilidade das métricas, reunindo a experiência prática de uso de jogos educacionais em sala de aula, a análise técnica dos dados de interação dos estudantes e a mensuração educacional, esta última validada com o apoio de inteligência artificial.

Embora tenha-se buscado a participação de psicometristas como avaliadores humanos na etapa de validação, não foi possível compor um painel com esses profissionais dentro do período de execução do estudo, principalmente devido à disponibilidade limitada de especialistas com experiência em mensuração educacional aplicada ao contexto de jogos educacionais e instrumentos baseados em métricas. Como estratégia de mitigação, manteve-se a validação com professores e analistas de dados e incluiu-se uma etapa complementar de revisão por modelos de inteligência artificial, com enfoque psicométrico.

Os questionários foram compostos por perguntas gerais e por um conjunto de questões específicas de “Avaliação da Métrica”, repetidas para cada métrica proposta, permitindo a análise individualizada de cada uma delas.

A utilização de questionários estruturados aplicados a especialistas foi adotada com o objetivo de obter evidências de validade de conteúdo das métricas propostas. Esse tipo de abordagem é amplamente utilizada na literatura para verificar se os elementos de um instrumento representam adequadamente os construtos que se pretende avaliar, especialmente quando se tratam de definições conceituais e operacionais ainda em processo de refinamento ([HAYNES; RICHARD; KUBANY, 1995](#); [ZAMANZADEH et](#)

al., 2015).

Inicialmente, os questionários passaram por uma pré-validação interna realizada pela equipe de pesquisa, com o objetivo de identificar inconsistências gerais, avaliar a clareza dos enunciados e garantir a coerência entre as perguntas e os objetivos do estudo.

Posteriormente, foi realizada uma pré-validação externa com quatro profissionais: três professores (um doutor e dois doutorandos) e um aluno de pós-graduação, utilizando o Índice de Validade de Conteúdo (IVC) como referência para análise. O IVC é uma metodologia que auxilia na avaliação das perguntas com base nos critérios de clareza e coerência, a partir de notas atribuídas em uma escala Likert de 1 a 4 para cada critério. O cálculo do IVC é feito somando-se as pontuações recebidas e dividindo pelo valor máximo possível, sendo considerado válido quando apresenta valor igual ou superior a 0,78.

Na pré-validação, o questionário para professores, inicialmente com 24 perguntas, apresentou 16 questões abaixo do valor de referência, enquanto o questionário para analistas de dados, também com 24 perguntas, apresentou 18 questões abaixo do valor, apesar de concentrarem notas médias próximas de 0,75. A partir das observações e comentários recebidos, foram realizados ajustes nos enunciados, exclusão de perguntas, subdivisão em itens mais específicos e inclusão de novos itens. Após as revisões, o questionário final ficou com 29 perguntas para professores e 23 para analistas de dados, todas com valores acima do estabelecido para validação.

Também foi utilizado um modelo de inteligência artificial (ChatGPT 4.5) para apoiar esta etapa de validação das perguntas dos questionários de ambos os grupos. As perguntas foram submetidas ao modelo, que atribuiu notas a cada questão nos critérios de clareza e coerência e calculou os respectivos valores de IVC para cada critério, auxiliando na identificação de incoerências.

A etapa final de validação sobre as métricas foi realizada com professores e analistas de dados, com uma avaliação complementar por agentes inteligentes com enfoque psicométrico. Essa etapa teve como objetivo confirmar a compreensão e a relevância das métricas antes de sua aplicação prática; os resultados dessa validação

são apresentados na Seção 6.2.5.

5.6.1 Procedimento de validação das respostas dos avaliadores

A validação das respostas dos especialistas foi realizada com base no coeficiente *V* de Aiken (AIKEN, 1985), utilizado para medir o grau de concordância entre avaliadores quanto à relevância e clareza de cada aspecto avaliado. O *V* de Aiken foi escolhido por considerar toda a escala ordinal de respostas (1 a 5), evitando a perda de informação associada à dicotomização empregada em outros métodos de validação de conteúdo, como o Índice de Validade de Conteúdo (IVC). Além disso, o método apresenta maior sensibilidade na análise de julgamentos em escalas ordinais e é particularmente adequado para estudos com número reduzido de avaliadores, sendo amplamente recomendado em pesquisas que envolvem validação de instrumentos educacionais e psicológicos (PENFIELD; GIACOBBI, 2004; ZAMANZADEH et al., 2015).

Para cada métrica, foram selecionadas as perguntas do grupo “Avaliação da Métrica” que utilizavam escala Likert (Tabela 8 e Tabela 9). Cada pergunta foi transformada em um critério analítico, como por exemplo: “As definições são claras?” (Clareza), “Ajudaria a identificar dificuldades de aprendizagem nos estudantes?” (Identificação de dificuldades) e “O nome representa claramente o que se pretende avaliar?” (Representatividade). Esses critérios foram avaliados numericamente com base no *V* de Aiken, permitindo identificar o grau de consenso entre os especialistas.

As questões binárias (Sim/Não), como “Você compreendeu a lógica da métrica?” (Compreensão), seguiram um critério de classificação baseado na proporção de respostas positivas. Quando a maioria dos avaliadores respondeu “Sim”, o aspecto foi classificado como **Compreensão Plena**; quando houve concordância parcial, como **Compreensão Parcial**; e, na ausência de consenso positivo, como **Não Compreendido**. Esse procedimento é compatível com abordagens descritivas aplicadas à análise de dados categóricos em estudos de validação de conteúdo, permitindo identificar o nível geral de entendimento dos avaliadores em relação aos elementos analisados (ZAMANZADEH et al., 2015).

Pergunta	Aspecto	Tipo de Pergunta
As definições são claras?	Clareza	Likert
Você compreendeu a lógica da métrica?	Compreensão (*)	Binária
Se teve dificuldade, qual foi essa dificuldade?	Vazio	Textual
Ajudaria a identificar dificuldades de aprendizagem nos estudantes?	Utilidade	Likert
O nome representa claramente o que se pretende avaliar?	Representatividade	Likert
Se não, qual nome acharia mais adequado?	Vazio	Vazio
A abordagem metodológica é coerente com os princípios de mensuração educacional?	Coerência	Likert
A fórmula/lógica proposta é adequada para medir o desempenho nos aspectos de aprendizagem?	Adequação	Likert
Há alguma fragilidade na formulação para medir o desempenho dos alunos?	Fragilidade	Binária
Se sim, quais problemas você observa?	Vazio	Textual
Pode ser adaptada para medir aspectos de aprendizagem em outros jogos educacionais?	Generalização	Likert
Se não, por quais motivos?	Vazio	Textual
Os resultados demonstrados refletem o contexto apresentado?	Contextualização	Likert
Os resultados são influenciados por fatores externos (por exemplo, perfil dos jogadores, ambiente de jogo)?	Sensibilidade	Likert
Oferece dados válidos e confiáveis para avaliar os indícios de aprendizagem dos alunos?	Validade	Likert
Há alguma característica ausente que melhoraria a avaliação do aprendizado?	Vazio	Textual

Tabela 8 – Perguntas relacionadas as métricas (Analistas de dados e Psicométricos).

Pergunta	Aspecto	Tipo de Pergunta
O nome da métrica representa claramente o que se pretende avaliar?	Representatividade	Likert
Se não, como você ajustaria o nome da métrica?	Vazio	Vazio
Você compreendeu a relação entre a métrica e o aspecto de aprendizagem que ela pretende mensurar?	Compreensão	Likert
Quanto você considera que compreendeu como a métrica é calculada?	Clareza do cálculo	Likert
Se teve dificuldades para compreender, descreva qual foi a dificuldade encontrada	Vazio	Vazio
Quanto você considera que compreendeu como a métrica é aplicada aos dados de interação do jogador com o jogo educacional?	Aplicabilidade	Likert
Você considera que as medidas e os valores gerados pela métrica refletem de forma confiável o progresso dos estudantes em relação à aprendizagem?	Validade	Likert
Pensando nas informações geradas pela métrica, o que seria necessário para que você tivesse mais confiança em usá-las como indicativo do progresso dos estudantes?	Vazio	Vazio
Você considera que as informações geradas pela métrica podem ajudar a identificar dificuldades de aprendizagem nos estudantes?	Utilidade	Likert
Você consideraria utilizar as informações geradas pela métrica como um dos critérios para compor a nota dos estudantes?	Aplicação pedagógica	Likert
Caso você não considere, por quais motivos?	Vazio	Vazio
Se você tivesse acesso aos dados gerados pela métrica, como pretendia utilizá-los para acompanhar o desenvolvimento ou as dificuldades dos seus alunos?	Vazio	Vazio

Tabela 9 – Perguntas relacionadas às métricas (Professores).

Cada métrica foi avaliada a partir de múltiplos aspectos, como clareza, coerência, aplicabilidade e compreensão, caracterizando uma análise multidimensional das definições propostas. A adoção de diferentes critérios para avaliação de um mesmo elemento é recomendada em processos de desenvolvimento e validação de instrumentos, pois contribui para identificar limitações específicas e aprimorar a qualidade metodológica dos itens avaliados (DEVELLIS, 2016). Essa estratégia também favorece comparações sistemáticas entre métricas e a análise independente das diferentes características associadas à sua utilização em contextos educacionais.

Já as perguntas abertas (dissertativas) foram analisadas qualitativamente, com o objetivo de compreender os motivos das respostas quantitativas, identificar possíveis ambiguidades nas métricas e reunir sugestões de aprimoramento. A inclusão de questões abertas em processos de validação é recomendada na literatura por permitir complementar os resultados quantitativos com evidências interpretativas, contribuindo para o refinamento conceitual dos instrumentos avaliados (HAYNES; RICHARD; KUBANY, 1995).

A interpretação dos resultados seguiu os seguintes limiares de validação: valores de V de Aiken iguais ou superiores a 0,80 foram considerados **aceitos sem restrição**; entre 0,75 e 0,79, **aceitos com restrição**; e inferiores a 0,75, classificados como **revisar**. Esses limiares foram definidos com base em recomendações presentes em estudos recentes que empregaram o coeficiente V de Aiken para validação de instrumentos educacionais e escalas do tipo Likert, indicando valores iguais ou superiores a 0,80 como evidência adequada de validade de conteúdo (PENFIELD; GIACOBBI, 2004; AHMAD et al., 2024; LOUREIRO et al., 2024).

Esse processo foi aplicado de forma sistemática a todas as métricas propostas, considerando as respostas dos três grupos de avaliadores (professores, analistas de dados e psicometristas representados por agentes inteligentes), com o objetivo de assegurar rigor metodológico e consistência na validação dos resultados.

No caso dos agentes inteligentes, o procedimento apresentou diferenças, pois foram utilizados apenas dois modelos, ChatGPT 4.5 e Gemini 2.5 Pro. Nessa etapa, não foi aplicado o índice de Aiken, uma vez que a análise concentrou-se em uma

avaliação qualitativa e comparativa das métricas, com ênfase na clareza conceitual, coerência teórica e propriedades psicométricas das definições e fórmulas apresentadas. Essa abordagem complementar foi adotada como estratégia de apoio interpretativo, permitindo ampliar a análise conceitual das métricas por meio de revisão sistemática baseada em conhecimento especializado disponível em modelos de linguagem.

5.6.2 Procedimento de validação das respostas dos avaliadores

Para ambos os modelos, foi utilizado um *prompt-base* de configuração do papel do agente, definindo-o como um especialista em Psicometria e orientando a análise segundo evidências de validade (conteúdo, construto e critério, incluindo evidências baseadas na estrutura interna) e fidedignidade/precisão da medida (por exemplo, consistência interna e estabilidade temporal), além da consideração de abordagens como Teoria Clássica dos Testes (TCT), Teoria de Resposta ao Item (TRI) e modelos Rasch.

Também se estabeleceu uma postura crítica, objetiva e imparcial, alinhada a boas práticas internacionais (por exemplo, os *Standards for Educational and Psychological Testing*), com explicitação de limitações, riscos metodológicos e oportunidades de aprimoramento.

Ressalta-se que foram necessários ajustes nos prompts, uma vez que, inicialmente, os modelos apresentavam respostas superficiais ou com baixa profundidade crítica. Para superar essa limitação, foram introduzidas instruções adicionais explícitas, orientando os modelos a adotarem maior rigor analítico e fundamentação científica. Entre essas instruções destacam-se comandos como: “lembre, eu espero que você seja uma pessoa rígida, que não esteja apenas querendo me agradar”, “responda com bases científicas e lógicas”, “não tente justificar o que eu fiz, apenas validar o que foi feito”, “imagine que você é um profissional disso, sua missão é essa” e “pense antes de responder e leve em conta múltiplos fatores associados”. Essas orientações tiveram como finalidade induzir respostas mais críticas, analíticas e metodologicamente fundamentadas.

As métricas foram apresentadas aos modelos de forma individual e sequencial, seguindo um procedimento padronizado. Para cada métrica, foi fornecida inicialmente

uma descrição conceitual contendo sua definição teórica, lógica operacional e exemplo de aplicação em contexto de jogo educacional. Em seguida, foi aplicado um conjunto estruturado de questões avaliativas que abordavam aspectos como clareza conceitual, coerência metodológica, adequação da formulação matemática, capacidade de identificar dificuldades de aprendizagem, possíveis fragilidades na mensuração, influência de fatores externos, validade e confiabilidade dos dados gerados, bem como sugestões de aprimoramento metodológico.

Esse procedimento foi repetido sistematicamente para cada uma das métricas analisadas, incluindo Assertividade, Flexibilidade Cognitiva, Habilidade Lógica, Exploração, Habilidade Mecânica e Pró-Eficiência Curricular. A apresentação individualizada das métricas permitiu análises independentes de cada construto, favorecendo maior precisão na identificação de limitações conceituais e operacionais específicas.

Após a introdução dos comandos adicionais de rigor crítico, observou-se uma mudança consistente no padrão das respostas geradas, caracterizada por maior identificação de fragilidades metodológicas, distinção mais clara entre indicadores comportamentais e evidências efetivas de aprendizagem, além de maior explicitação de limitações relacionadas à validade de construto, validade de conteúdo e generalização dos resultados. Esse processo configurou um refinamento iterativo dos prompts, funcionando como um mecanismo de calibração do comportamento analítico dos modelos, permitindo maior alinhamento com princípios psicométricos e boas práticas de avaliação educacional.

Síntese do capítulo

Este capítulo apresentou a proposta de métricas para avaliar indícios de aspectos de aprendizagem a partir de dados de interação em jogos educacionais. Foram definidos os aspectos-alvo (engajamento, desenvolvimento cognitivo, desenvolvimento motor e unidade curricular) e mapeadas categorias de dados comuns aos jogos que sustentam a mensuração desses aspectos. Com base nesse mapeamento, foram descritas oito métricas, incluindo seus mecanismos de ação, formulações e exemplos de aplicação,

além das adaptações realizadas na arquitetura de dados do *GLBoard* para viabilizar a coleta dos registros necessários. Complementarmente, foi delineado o questionário de autoavaliação aplicado aos estudantes e o processo de validação das métricas com especialistas e agentes inteligentes, estabelecendo os procedimentos que fundamentam a análise dos resultados no capítulo seguinte.

6

RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados obtidos na pesquisa. Após a formulação das métricas e a reestruturação da arquitetura de dados do *GLBoard*, iniciou-se a fase de testes das métricas em dois jogos educacionais distintos: “Robô Euroi” (MELO et al., 2018) e “MyName” (NASCIMENTO et al., 2023). A escolha desses jogos considerou suas diferentes mecânicas, abordadas nas seções seguintes, com o intuito de verificar a adequação e os limites das métricas propostas em contextos variados de interação.

As métricas foram aplicadas aos dados gerados durante a interação dos estudantes com os dois jogos citados. Complementarmente, foi realizado um estudo de autoavaliação, descrito no Capítulo 5, por meio de um questionário aplicado ao término das sessões de jogo. O objetivo foi verificar a convergência entre os resultados das métricas e a percepção dos próprios estudantes sobre seu desempenho. O questionário incluiu perguntas de caráter demográfico e itens relacionados aos quatro aspectos de aprendizagem.

Além disso, este capítulo também apresenta e discute os resultados da avaliação realizada por três grupos de avaliadores, sendo eles professores, analistas de dados e psicometristas (estes últimos representados por agentes inteligentes), em relação à clareza, coerência, utilidade e validade das métricas propostas. Essa análise complementa os resultados empíricos obtidos, oferecendo uma validação teórica e técnica sobre a aplicabilidade e consistência das métricas propostas.

A Tabela 7 apresenta a correspondência entre os itens do questionário e os respectivos aspectos avaliados.

Esta pesquisa integra o projeto “Um *framework* para o desenvolvimento de jogos educacionais com base nas teorias de aprendizagem e aspectos cognitivos dos estudantes,” aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal do Amazonas (CEP-UFAM), sob parecer emitido em 26/03/2025. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido (TCLE), autorizando o uso dos dados para fins acadêmicos, com garantia de anonimato e confidencialidade das informações coletadas.

A captura dos dados exigiu a adaptação dos jogos ao pacote do *GLBoard*, processo que envolveu desafios, discutidos ao longo das próximas seções.

Para cada jogo analisado, será apresentada uma breve descrição de suas mecânicas, seguida da análise da integração com o *GLBoard* e da aplicação das métricas. Os resultados serão discutidos com base nas interações dos estudantes, nas respostas dos especialistas e na comparação com os dados provenientes da autoavaliação.

6.1 Experimento I

Esse experimento, que utilizou o jogo “Robô Euroi”, envolveu cinco estudantes do primeiro período de cursos de computação da Universidade do Estado do Amazonas (UEA), sendo três de “Licenciatura em Computação”, um de “Engenharia Naval” e um de “Engenharia Civil”. Entre os participantes, dois eram do gênero masculino e duas do gênero feminino.

Os jogadores foram colocados em uma sala onde puderam jogar livremente pelo tempo que desejassem, sem a obrigação de finalizar o jogo. Apenas um dos participantes não conseguiu concluir o jogo. Na próxima seção, serão explicados o enredo e a mecânica do jogo.

6.1.1 Robô Euroi

“Robô Euroi” é um jogo educacional 2D com mecânica de plataforma, desenvolvido com o propósito de fomentar o pensamento computacional e habilidades matemáticas

em crianças em fase de alfabetização. Esse jogo de estratégia matemática combina a aquisição de conhecimento de maneira implícita com desafios lógicos (MELO et al., 2018).

Na dinâmica do jogo, os jogadores controlam Euroi, um robô encarregado de recuperar peças de uma nave espacial que se acidentou enquanto explorava diversos planetas. A mecânica do jogo é centrada na gestão de energia de Euroi, que é afetada positiva ou negativamente pela interação com diferentes elementos nos cenários. Plataformas contendo blocos vermelhos simbolizam energias negativas que reduzem a energia de Euroi, introduzindo um elemento de risco e estratégia à navegação por plataformas. Em contraste, blocos azuis representam energias positivas que recarregam a energia do robô. Além desses elementos, o personagem também deve enfrentar inimigos que, ao entrar em contato, removem toda a sua vida. Para vencer esses inimigos, o personagem pode disparar projéteis, mas cada disparo remove um ponto de vida. O objetivo é planejar cuidadosamente seus movimentos e fazer escolhas estratégicas para avançar no jogo, sem esgotar seus recursos vitais. Algumas fases do jogo introduzem novos elementos, além de desafios mecânicos e lógicos que devem ser realizados com sucesso. Para concluir a missão (fase) com sucesso, o jogador deve guiar o personagem até a plataforma que contém um “Bloco de Engrenagem”; ao coletar esse elemento, a fase é finalizada com sucesso. Ao todo, o jogo possui 13 fases, descritas na Tabela 10.

6.1.2 Mapeamento de dados para integração das métricas

Como pode ser observado na Tabela 10, de modo geral, as fases apresentam os mesmos elementos base (plataformas), variando apenas a quantidade de plataformas e de decisões a serem tomadas. Em algumas fases, novos elementos são adicionados, como inimigos e plataformas que se movem ou caem. A análise de cada fase serve como base para explicar o processo de “Mapeamento de Dados para Integração de Métricas”, no qual é considerado como os elementos do jogo se alinham à estrutura exigida pelo *GLBoard* para a coleta dos dados necessários à aplicação das métricas.

Dentre as métricas propostas neste trabalho, algumas não puderam ser aplicadas

Tabela 10 – Descrição das telas do jogo Robô Euroi.

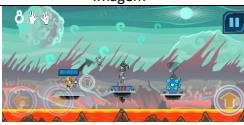
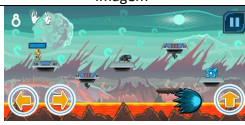
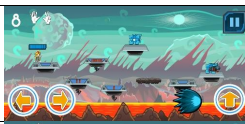
Fase	Imagem	Descrição	Fase	Imagem	Descrição
1		Esta é uma fase tutorial cujo principal objetivo é familiarizar o jogador com a mecânica do jogo. Nela, o jogador deve saltar entre plataformas até chegar ao bloco de engrenagem azul. Durante o percurso, são introduzidos os controles do personagem e as plataformas que removem energia.	8		Nesta fase, dois novos elementos são apresentados ao jogador: um inimigo e a opção de realizar um disparo, que consome um ponto de vida do personagem. O objetivo é acertar e derrotar o inimigo.
2		Nesta fase, o jogador é posto diante de um desafio lógico, no qual deve escolher entre duas plataformas. Além disso, um novo elemento é introduzido, o qual concede energia ao personagem.	9		Nesta fase, o jogador enfrenta o desafio de eliminar um inimigo e fazer uma escolha lógica. Além disso, é introduzido um novo elemento: o jogador deve capturar duas caixas de energia para vencer a fase.
3		Esta fase apresenta dois desafios lógicos que o jogador deve resolver, escolhendo entre seguir pelo caminho superior ou inferior. O caminho superior é o que leva à vitória.	10		Nesta fase, o principal objetivo é resolver o problema da melhor forma possível, preservando o máximo de energia. Além disso, a fase exige habilidade motora e estratégia, pois o jogador deve calcular corretamente os disparos de projéteis e os saltos para concluir a fase com sucesso.
4		Nesta fase, o jogador enfrenta três desafios lógicos, que envolvem fazer escolhas corretas para alcançar o melhor resultado. O objetivo é conservar a maior quantidade de energia do personagem.	11		Nesta fase, o principal objetivo é compreender o funcionamento do novo elemento introduzido: um escudo que permite ao personagem pular em uma plataforma sem perder vida. No entanto, após o uso, o escudo é perdido.
5		Nesta fase, o jogador enfrenta um desafio lógico e é introduzido a um novo elemento: uma plataforma móvel.	12		Nesta fase, o personagem enfrenta dois desafios lógicos, além de desafios mecânicos que incluem saltos e derrotar inimigos. A dificuldade da fase está na escolha do caminho a ser seguido.
6		Nesta fase o jogador é posto a resolver dois desafios lógicos. O foco aqui é fazer com que o estudante possa pensar em escolhas futuras, pois a primeira escolha que ele faz de plataforma implica nas escolhas futuras.	13		Esta é a última fase do jogo, na qual o personagem deve derrotar um chefe que possui uma barra de vida e um ataque de laser capaz de remover 3 pontos de vida do personagem. Nesta fase, a habilidade mecânica do personagem é primordialmente testada, exigindo saltos precisos para desviar dos ataques do chefe, além de acertar os disparos.
7		Nesta fase, um novo elemento é introduzido no jogo: plataformas que começam a cair quando pisadas. Não há desafios lógicos nesta fase; o foco está apenas em demonstrar esse novo elemento.			



Figura 14 – Fase II do jogo Robô Euroi.

ao jogo, dado que ele não possui elementos compatíveis com os aspectos ou métricas específicas. A seguir será descrito como cada métrica se comporta no jogo e como o mapeamento foi realizado:

- **Exploração:** essa métrica não se aplica ao jogo, pois não há elementos de exploração disponíveis, como itens opcionais a serem coletados, insígnias ou locais a descobrir.
- **Assertividade:** aplica-se ao jogo. Para que a métrica funcione, é necessário registrar o caminho que o jogador percorreu e definir um caminho ótimo para cada fase. O caminho do jogador foi mapeado para plataformas que foram acessadas, e o caminho perfeito foi mapeado como a sequência de plataformas que leva ao melhor resultado possível em cada fase, onde o personagem finaliza com a maior quantidade.
- **Flexibilidade Cognitiva:** aplica-se, pois é possível mapear o caminho do jogador, que inclui os saltos entre plataformas, e capturar esses dados em cada tentativa dentro da fase.
- **Habilidade Lógica:** Para aplicá-la, foi necessário definir o que caracteriza um desafio lógico e como identificá-lo no jogo. Neste contexto, considera-se desafio lógico toda decisão crítica que afeta diretamente o desempenho do jogador. A Figura 14 ilustra esse conceito: cada círculo verde representa um desafio lógico. Apenas transições entre plataformas que envolvam escolhas com consequências significativas para o progresso são consideradas.

Nesta fase do jogo (Figura ??), foram identificados dois desafios lógicos:

1. A partir da plataforma “Fixa A”, o jogador deve escolher entre “Plat-4 B” e “Plat-3 C”. A escolha correta é “Plat-3 C”, pois leva ao melhor desfecho. Qualquer outra opção é considerada uma falha.
2. Ao alcançar a plataforma “Móvel”, o jogador deve seguir para “Plat-4 E”. Escolher qualquer outra alternativa é interpretado como erro na resolução do desafio.

A implementação dessa métrica exigiu modificações no código do jogo para permitir a detecção e registro dos desafios lógicos.

- **Habilidade Mecânica:** essa métrica também é aplicável ao jogo e, assim como a Habilidade Lógica, baseia-se no mapeamento de plataformas. Na Figura 14, os círculos laranjas representam os desafios mecânicos. No contexto analisado, um desafio mecânico corresponde à execução bem-sucedida de um salto entre plataformas. Saltos concluídos com êxito em novas plataformas são registrados como acertos, enquanto falhas são caracterizadas pela queda em elementos hostis, como a “Lava”, resultando na eliminação do personagem. Essas falhas estão associadas à imprecisão na coordenação entre o comando de salto e a direção do movimento.

Além disso, o jogo inclui desafios mecânicos relacionados ao disparo de projéteis contra inimigos. Consideram-se erros os disparos que não atingem o alvo ou situações em que o jogador é atingido por projéteis inimigos; por outro lado, desvios bem-sucedidos são contabilizados como acertos. Devido à ênfase do jogo em elementos de ação, observa-se uma maior frequência de desafios mecânicos em relação aos desafios lógicos.

- **Pro- Eficiência Curricular:** não se aplica nesse caso, pois o conteúdo curricular do jogo está voltado ao desenvolvimento do pensamento computacional e de habilidades matemáticas. Como as decisões lógicas avaliadas pela métrica de Habilidade Lógica já refletem esse conteúdo, a aplicação simultânea da métrica de Proficiência Curricular resultaria em sobreposição avaliativa.
- **Tempo de *Gameplay*:** é aplicável, visto que o sistema registra com precisão os instantes de início e término de cada fase e tentativa, além da conclusão de cada desafio lógico ou mecânico. Esses dados permitem o cálculo do tempo total de interação do jogador com o jogo.
- **Tentativas até o Sucesso:** também aplicável, uma vez que o jogo armazena todas as tentativas realizadas em cada fase e desafio, juntamente com os respectivos resultados (sucesso ou falha). A métrica permite identificar a quantidade de tentativas necessárias para alcançar o êxito, servindo como indicativo de progresso e assimilação de estratégias por parte do jogador.

As métricas mais complexas de aplicar foram *Habilidade Lógica*, devido à necessidade de identificar as decisões corretas do jogador, e *Assertividade*, que exigiu a definição e o registro do caminho ideal para cada fase na estrutura de dados.

Algumas métricas não foram aplicadas em fases específicas do jogo, em função das características e objetivos de cada fase. Por exemplo, a métrica de *Habilidade Lógica* não foi utilizada na Fase 7, cujo objetivo era apenas introduzir um novo elemento, sem a inclusão de desafios lógicos. O mesmo ocorreu em fases como a 8 e a 10, nas quais determinadas métricas foram desconsideradas conforme os objetivos definidos para cada fase. Esses aspectos serão detalhados na próxima seção.

6.1.3 Análise de desempenho dos jogadores nas métricas

Após o mapeamento dos dados necessários para capturar as interações no jogo, iniciou-se um experimento com o objetivo de aplicar as métricas definidas, analisar seu comportamento e discutir a aprendizagem dos estudantes com base nos resultados obtidos por essas métricas.

Cada jogador teve suas fases e tentativas avaliadas conforme discutido na Subseção 6.1.2. Quatro gráficos foram elaborados para ilustrar o comportamento das métricas, divididas em grupos devido à alta oscilação. O primeiro gráfico apresenta “Assertividade” e “Flexibilidade Cognitiva”; o segundo, “Habilidade Lógica” e “Habilidade Mecânica”; o terceiro, “Tentativas por Fase”; e o quarto, “Tempo por Fase (em segundos)”. O eixo X representa as fases do jogo, e o Y, os valores das métricas (0 a 1, quando aplicável). Cada linha representa a média dos valores obtidos por fase. Métricas não aplicáveis foram excluídas.

A métrica de Habilidade Mecânica apresentou valores elevados, sugerindo que os jogadores não enfrentaram grandes dificuldades com saltos ou inimigos. Como os participantes não pertencem ao público-alvo (crianças), esse desempenho era esperado. Destacam-se picos nas fases 1 (introdutória) e 7 (novo elemento), ambas com desafios caracterizados como simples. A partir da fase 8, observa-se uma queda nos valores, possivelmente pela introdução de inimigos que exigem precisão. Em seguida, verifica-

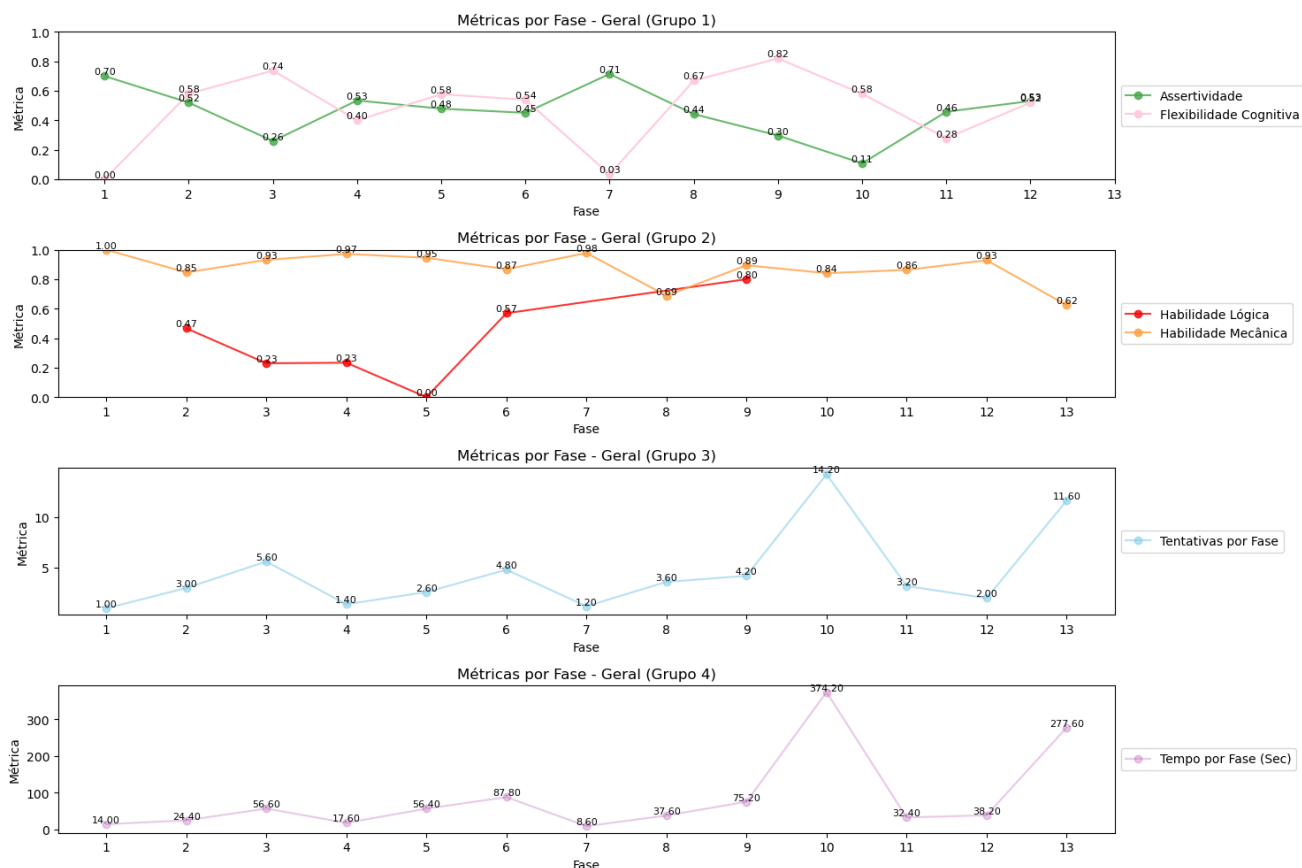


Figura 15 – Gráficos por métricas em cada fase do jogo "Robô Euroi".

se uma recuperação no desempenho, indicando um processo de adaptação às novas exigências do jogo. A fase 13 volta a apresentar desempenho inferior, compatível com sua maior complexidade.

A Habilidade Lógica não é aplicada uniformemente ao longo das fases, revelando lacunas no *design*. Isso decorre da introdução de novos elementos sem desafios lógicos estruturados. As fases iniciais têm desempenho inferior até a fase 5, sugerindo curva de adaptação. Depois disso, há maior domínio das regras lógicas, inclusive em fases mais exigentes.

A Assertividade inicia alta na fase 1 e decresce até a fase 3, acompanhando o aumento da complexidade. A Flexibilidade Cognitiva cresce até a fase 4, indicando adaptação a novas situações, e depois se estabiliza. Na fase 7, a Assertividade cresce e a Flexibilidade cai, em função da linearidade da fase, isto é, uma fase com trajetória mais definida e menor quantidade de alternativas para alcançar o objetivo. Nas fases 8 e 9, ocorre o oposto, com a introdução de desafios não convencionais, nos quais o estudante precisa explorar diferentes possibilidades de solução.

A fase 10 apresenta queda acentuada na Assertividade e elevação da Flexibilidade Cognitiva, sendo uma das mais exigentes. A dificuldade está associada à necessidade de realizar um salto longo e preciso, indicando possível inadequação no *design*, já que ocorre antes do final. Apesar de a fase 13 ter sido projetada como a mais difícil, os maiores sinais de frustração aparecem na fase 10.

Nas fases 11 e 12, observa-se redução gradual da Flexibilidade e leve recuperação da Assertividade, indicando domínio das novas mecânicas. Na fase final, a Assertividade não é mensurada, pois o objetivo é derrotar o chefe, tornando irrelevantes as escolhas de trajeto.

As métricas de Tentativas e Tempo variam ao longo das fases, refletindo o grau de dificuldade. Em fases mais complexas, observa-se um aumento nesses valores; em fases mais simples ou já familiarizadas pelos jogadores, tendem a cair. Após o pico da fase 10, ambas diminuem nas fases 11 e 12, sugerindo desafio menor e maior preparo dos jogadores. Isso é reforçado pelo aumento da Assertividade e queda da Flexibilidade, sinalizando adaptação às mecânicas do jogo.

A análise das fases iniciais mostra alta Assertividade e baixa Flexibilidade Cognitiva, indicando confiança em ambientes previsíveis. Com o aumento da exigência, a Flexibilidade cresce, evidenciando adaptação estratégica.

As variações de Habilidade Mecânica e Lógica apontam momentos específicos de exigência. Fases com baixa Habilidade Mecânica e altos valores de Tentativas e Tempo, como a 10, indicam dificuldades de coordenação motora, como saltos e esquivas.

O número de Tentativas e o Tempo dispendido também indicam persistência e envolvimento. Fases com alta recorrência de tentativas e tempo prolongado, mesmo com baixo desempenho (como na fase 10), sugerem disposição para enfrentar desafios. Esse comportamento reflete níveis significativos de engajamento e contribuindo para uma análise mais ampla da aprendizagem durante a *gameplay*.

6.1.4 Análise comparativa entre o desempenho nas métricas e o questionário autoavaliativo

Os resultados da autoavaliação dos jogadores em relação aos aspectos de aprendizagem estão apresentados na Tabela 11, com os valores normalizados entre 0 e 1, utilizando o método Min-Max, a partir da escala Likert de 1 a 5.

Estudante	Gênero	Flexibilidade Cognitiva	Assertividade	Mecânica	Lógica	Tentativas Por Fase (Esforço)
1	Masculino	1,00	1,00	0,40	1,00	0,60
2	Masculino	0,80	1,00	0,60	1,00	0,80
3	Masculino	0,80	0,80	0,40	0,80	0,60
4	Feminino	1,00	0,60	0,80	0,40	1,00
5	Feminino	0,80	0,80	0,80	0,60	0,60
Média		0,88	0,84	0,60	0,76	0,72

Tabela 11 – Valores únicos e média das respostas dos estudantes em cada aspecto de aprendizagem no jogo "Robô Euroi".

A Flexibilidade Cognitiva foi avaliada como alta (0,88), condizente com os dados do gráfico 15, que mostram tentativas diante de desafios. Já a Assertividade apresentou discrepância: média de 0,84 na autoavaliação, mas valores inferiores a 0,5 nas métricas em grande parte das fases. Isso sugere que os jogadores acreditaram estar fazendo boas escolhas, embora os resultados não tenham refletido isso. A jogadora 4 foi uma exceção, com avaliação mais alinhada às métricas.

Em Tentativas por Fase, a média autorreferida (0,72) foi compatível com os dados empíricos (4,49 tentativas por fase), com aumentos nas fases 10 e 13. Apesar de o questionário não captar dificuldades por fase, observações durante o experimento confirmam os picos de esforço nessas etapas.

A Habilidade Lógica teve média de 0,76 na autoavaliação e 0,45 nas métricas, refletindo baixo desempenho inicial com posterior melhora. A Habilidade Mecânica também apresentou diferença: média autorreferida de 0,87 contra 0,6 nas métricas, possivelmente impactada pelas fases finais. As jogadoras 4 e 5 tiveram percepções mais próximas dos dados empíricos.

6.2 Experimento II

O segundo experimento foi realizado com o jogo *MyName*, contando com a participação de 15 estudantes a partir do terceiro período do curso de Licenciatura em Computação, sendo oito do gênero masculino e sete do feminino. A escolha por estudantes desse período se justifica pela oferta da disciplina “Aprendizagem em Informática – Abordagem Pedagógica”, voltada ao desenvolvimento de objetos educacionais digitais. Nesse contexto, a aplicação e apresentação do jogo atuaram como referência prática, servindo de inspiração para propostas e metodologias que os estudantes poderiam adotar em seus próprios projetos.

A atividade foi conduzida em sala de aula, durante o horário regular da disciplina, com duração de 45 minutos. A participação foi facultativa, e oito estudantes não concluíram o jogo. A próxima seção descreve a mecânica do jogo utilizado no experimento.

6.2.1 *MyName*

O jogo *MyName* é um jogo educacional centrado no Problema da Mochila, incorporando elementos de jogos de aventura com perspectiva *top-down* (NASCIMENTO et al., 2023). Seu objetivo é auxiliar estudantes, especialmente de cursos de computação, a compreender e aplicar esse problema clássico de otimização, que envolve a seleção de itens com pesos e valores distintos para maximizar o valor total sem exceder a capacidade da mochila.

A narrativa do jogo se integra à sua mecânica: o jogador assume o papel da Andarilha, uma personagem responsável por restaurar um universo devastado por monstros que devoraram palavras e seus significados. Cada monstro carrega uma letra e um valor de peso, estabelecendo uma analogia direta com os itens do Problema da Mochila, exigindo do jogador o gerenciamento da capacidade de carga da mochila da personagem.

Em cada fase (com exceção da última), o jogador escolhe uma entre três palavras disponíveis. A palavra selecionada deve ser formada coletando letras na ordem correta.

Por exemplo, ao escolher “CASA”, o jogador deve capturar as letras “C”, “A”, “S” e “A” na sequência indicada. Ao interagir com um monstro portador da letra correta na ordem correta, a letra é capturada e o monstro é derrotado. Caso contrário, o monstro é teleportado e o jogador perde uma vida.

Além disso, podem existir múltiplos monstros com a mesma letra, porém com pesos diferentes, exigindo atenção à soma dos pesos coletados. O valor total deve se manter dentro de um intervalo: deve atingir um valor mínimo e não ultrapassar a capacidade da mochila. Caso a palavra seja formada corretamente, mas o peso não esteja no intervalo exigido, a fase é reiniciada.

O jogo também introduz elementos adicionais, como projéteis lançados pelos inimigos. O jogador deve desviar desses ataques, evitando a perda de vidas. Cada fase permite até três vidas; ao perder todas, a fase é reiniciada, demandando uma revisão da estratégia de coleta e gestão de peso. Cada tentativa é distinta. Ao iniciar ou reiniciar uma fase, os inimigos são gerados de forma pseudoaleatória, variando letras, pesos e quantidades. A palavra-alvo permanece constante, mas o conjunto de inimigos disponíveis altera-se a cada tentativa, o que impõe ao jogador o ajuste contínuo de suas estratégias. Na Tabela 12 é possível visualizar a descrição de cada fase do jogo:

Tabela 12 – Descrição das telas do jogo *MyName*.

Fase	Imagem	Descrição
1		Nesta fase, o jogador tem a opção de formar as palavras MAGA, CAVALEIRA ou GODZILLA. Há dois tipos de inimigos: os azuis, que se lançam contra o personagem causando a perda de um ponto de vida se atingirem o jogador, e os CocoMan, que fogem do personagem.
2		As palavras disponíveis para escolha são DOBRO, XILOFONE e CLAVICÓRDIO. Os inimigos incluem os já conhecidos CocoMan e os novos inimigos Raivin, que também se lançam no personagem. Se o personagem passar perto de um Raivin, ele começará a persegui-lo temporariamente.
3		As opções de palavras são CEDRO, PINHEIRO e CARVALHO. Nesta fase, além dos Raivin, surgem inimigos similares a ogros, que lançam machados bumerangues. Esses machados complicam a mobilidade do jogador, que precisa constantemente desviar dos ataques que vão e voltam.
4		As palavras que podem ser formadas são TUNDRA, ESTEPES e PRADARIA. Esta fase introduz os inimigos Vazim, magos que lançam bolas de fogo. Vários Vazim aparecem simultaneamente, lançando bolas de fogo em conjunto, o que aumenta significativamente a dificuldade em manter a mobilidade do jogador.
5		Nesta última fase, o jogador deve formar a palavra CORAGEM. O chefe desta fase lança inimigos em direção ao personagem. Esses inimigos são os CocoMan, que contêm as letras e os pesos que precisam ser coletados para formar a palavra.

6.2.2 Mapeamento de dados para integração das métricas

Na Figura 16, é possível visualizar uma fase do jogo. De forma geral, todas as fases possuem os mesmos elementos, mudando somente a palavra-alvo e os tipos de inimigos que o jogador deve enfrentar. A análise desta figura serve de base para explicar o processo de Mapeamento de Dados para Integração de Métricas, considerando como os elementos do jogo se alinham à estrutura exigida pelo *GLBoard* para a coleta dos dados necessários à aplicação das métricas.

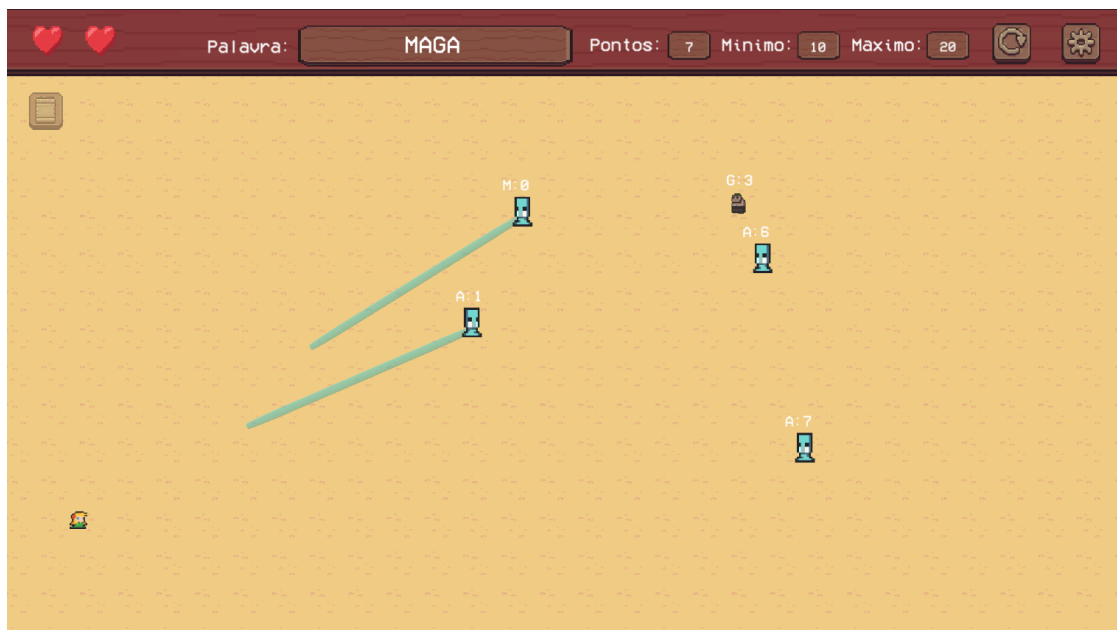


Figura 16 – Fase I do jogo *MyName*.

Dentre as métricas estabelecidas, algumas não puderam ser aplicadas ao jogo, pois não há elementos compatíveis com os respectivos aspectos. A seguir, descreve-se como cada métrica se comporta em *MyName* e como foi realizado o mapeamento:

Exploração: não se aplica, uma vez que o jogo não possui elementos opcionais, como itens a descobrir, insígnias ou rotas alternativas.

Assertividade: não se aplica ao modelo atual do jogo. Para sua mensuração, seria necessário definir o caminho percorrido em cada tentativa e compará-lo a um caminho considerado ótimo. Esse caminho pode ser representado por uma sequência de interações com inimigos, considerando letras e pesos coletados. No entanto, a estrutura atual do *GLBoard* exige que o caminho ótimo (*bestPath*) seja definido previamente por fase. Como os inimigos são gerados aleatoriamente a cada tentativa, essa definição só seria

possível após o início da fase, tornando a aplicação da métrica inviável neste contexto. Além disso, não seria possível comparar duas soluções dos estudantes para a mesma fase, visto que cada solução seria baseada em um caminho perfeito diferente.

Flexibilidade Cognitiva: também não se aplica, pois mede a variação entre os caminhos adotados em tentativas diferentes. Como cada tentativa possui elementos distintos (letras, pesos e quantidades), os caminhos sempre serão diferentes, o que levará a um valor elevado da métrica, independentemente do comportamento do jogador. Isso compromete a sua relevância no contexto do jogo.

Habilidade Lógica: métrica aplicável, diretamente associada às decisões de coleta de letras e ao gerenciamento da capacidade da mochila. Cada decisão representa um desafio lógico, considerado bem-sucedido quando o jogador coleta a letra correta, na ordem correta e com peso que contribua para a otimização do espaço disponível.

Por exemplo, na Fase I (Figura 16), a palavra-alvo é “MAGA” e os inimigos disponíveis são [A1, A7, M0, G3, A6]. O jogador deve priorizar letras que maximizem a ocupação da mochila, como M0, A7, G3 e A6, somando os 16 pontos exigidos, sem ultrapassar o limite de 20. Coletar A1, embora correto quanto à letra, resulta em uso ineficiente do espaço.

Para aplicação da métrica, foi desenvolvido um algoritmo que determina a melhor ordem de coleta por tentativa. Uma estrutura adicional verifica se a letra coletada está na ordem correta e pertence ao conjunto ótimo da fase. Quando ambas as condições são atendidas, a tentativa é registrada como sucesso no *GLBoard*; caso contrário, como falha.

A viabilidade da métrica decorre do fato de que, apesar da aleatoriedade, é possível calcular a solução ótima após a geração dos elementos da fase. Isso a diferencia de métricas como *assertividade* e *flexibilidade cognitiva*, que exigem caminhos ótimos previamente definidos, inviáveis em jogos com múltiplos estados possíveis por tentativa.

Habilidade Mecânica: aplica-se à capacidade do jogador de esquivar-se de ataques. Os desafios mecânicos ocorrem quando projéteis ou inimigos realizam investidas contra o jogador. Considera-se sucesso quando o dano é evitado; considera-se falha quando o

jogador é atingido.

Para viabilizar a mensuração, foram implementadas modificações nos *scripts* dos inimigos com comportamento ofensivo. *Triggers* adicionais identificam ataques ativos e verificam se atingem ou não o jogador. Quando um inimigo acerta o jogador durante um ataque, a falha é mecânica; quando o contato ocorre sem ataque ativo, caracteriza uma falha lógica, indicando escolha inadequada.

Proficiência Curricular: não se aplica, pois o conteúdo curricular está integralmente incorporado nas decisões lógicas do jogo. Assim como no *Robô Euroi*, essa métrica seria redundante, uma vez que a Habilidade Lógica já contempla os aspectos educacionais propostos.

Tempo de Gameplay e Tentativas até o Sucesso: ambas são aplicáveis. Assim como no *Robô Euroi*, o jogo registra o tempo total de interação por fase e o número de tentativas até a conclusão com sucesso, permitindo acompanhar a progressão e o aprendizado do jogador.

No jogo, não foi possível aplicar métricas que dependem da comparação entre o caminho do jogador e um ou mais caminhos ideais previamente definidos, como é o caso da *assertividade* e da *flexibilidade cognitiva*. Isso se deve à aleatoriedade do sistema, que gera diferentes configurações de desafios a cada tentativa, impossibilitando a definição de caminhos perfeitos antecipadamente. Aplicar essas métricas nessas condições poderia resultar em valores distorcidos, como assertividade artificialmente baixa e flexibilidade cognitiva inflada, sem refletir com precisão o desempenho real do jogador.

As métricas mais complexas de mapear foram a “Habilidade Lógica”, que exigiu a verificação das decisões corretas do jogador, e a “Habilidade Mecânica”, que necessitava de ajustes nos *scripts* de cada inimigo para determinar se atingiam ou não o personagem.

Na última fase do jogo, as métricas de Habilidade Mecânica e Lógica não foram aplicadas devido a um *bug* encontrado na etapa. Além disso, a natureza da fase impede a aplicação da habilidade lógica. Isso ocorre porque não é possível determinar se a escolha do jogador é a melhor, dado que os inimigos, juntamente com as letras e os

pesos, são gerados aleatoriamente a cada disparo feito pelo chefe da fase.

6.2.3 Análise de desempenho dos jogadores nas métricas

Após a etapa de mapeamento dos dados necessários para capturar as interações no jogo, iniciou-se um experimento realizado igual ao ocorrido com o jogo “Robô Euroi”. Três gráficos (Figura 17) gerais foram elaborados para ilustrar o comportamento das métricas em relação aos jogadores.

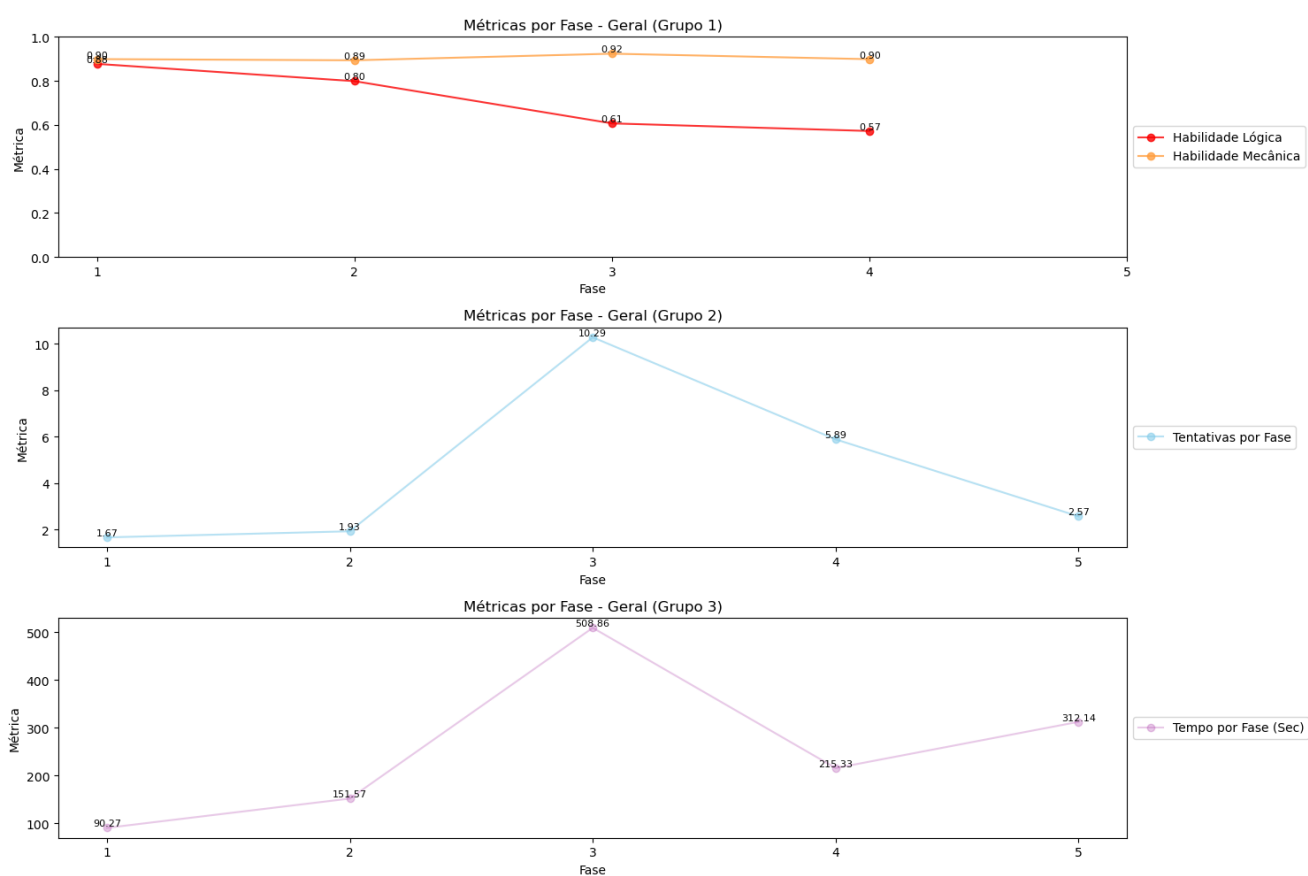


Figura 17 – Gráficos por métricas em cada fase do jogo. *MyName*

Nos gráficos, observa-se um declínio contínuo na “Habilidade Lógica” até a fase 4, sugerindo dificuldades crescentes à medida que os jogadores avançam. Esse comportamento pode estar relacionado ao aumento gradual no número de inimigos, com maior variedade de letras e pesos, ampliando as possibilidades de escolhas disponíveis para o jogador. Sob pressão, os jogadores tendem a optar por soluções mais acessíveis, em detrimento de estratégias otimizadas, o que compromete o desempenho lógico.

Paralelamente, nota-se um aumento progressivo na “Habilidade Mecânica”, indicando que os jogadores passam a lidar com mais facilidade com desafios motores mais exigentes. O crescimento reflete um processo de adaptação, evidenciado pela melhora na navegação e na execução de ações mecânicas, como esquivas e saltos.

A relação entre as duas métricas evidencia uma dinâmica entre exigências cognitivas e motoras. Enquanto a pressão por respostas rápidas pode comprometer a eficiência lógica, os desafios mecânicos ganham protagonismo. Essa interação reforça a importância de um *game design* equilibrado, no qual o aumento das demandas motoras não prejudique o engajamento com os aspectos lógicos do jogo.

As métricas de “Tentativas por Fase” e “Tempo por Fase” corroboram com a análise ao apontar os momentos de maior dificuldade. A fase 3 destaca-se por picos em ambas as métricas, indicando maior número de tentativas e tempo para superação. Essa dificuldade decorre da introdução simultânea de diversos inimigos, o que torna mais complexas as tarefas de esquiva e coleta. Embora não tenha sido originalmente projetada como crítica, a fase impõe desafios lógicos e mecânicos combinados.

Por outro lado, a última fase, concebida como a mais difícil, apresenta número reduzido de tentativas, mas um tempo médio de conclusão mais elevado. Isso ocorre porque, embora os inimigos sejam lentos, a fase exige a formação de uma palavra extensa, cujas letras são disponibilizadas apenas após os disparos do chefe, prolongando a jogabilidade.

De modo geral, os jogadores mantêm desempenho elevado ao longo do jogo, com quedas pontuais nas fases mais complexas. Tais variações são esperadas, especialmente considerando a estrutura do jogo, que apresenta um número reduzido de fases. Essa limitação pode gerar uma progressão de dificuldade abrupta, dificultando a adaptação gradual dos jogadores e impactando seu desempenho nas fases finais.

Estudante	Gênero	Mecânica	Lógica	Tentativas Por Fase (Esforço)
1	Feminino	1,00	1,00	0,60
2	Feminino	0,40	0,80	0,80
3	Masculino	0,40	0,80	0,60
4	Feminino	0,40	0,80	0,80
5	Masculino	0,80	1,00	0,60
6	Masculino	0,40	1,00	0,80
7	Masculino	1,00	1,00	0,80
8	Masculino	0,80	1,00	1,00
9	Masculino	0,20	0,60	0,80
10	Feminino	0,40	0,40	0,80
11	Masculino	0,20	0,60	1,00
12	Feminino	0,40	1,00	1,00
13	Feminino	0,40	0,80	0,80
14	Masculino	1,00	0,60	0,40
15	Feminino	0,40	0,40	0,80
Média		0,55	0,79	0,77

Tabela 13 – Valores únicos e média das respostas dos estudantes em cada aspecto de aprendizagem no jogo *MyName*.

6.2.4 Análise comparativa entre o desempenho nas métricas e o questionário autoavaliativo

Os resultados da autoavaliação dos jogadores em relação aos aspectos de aprendizagem estão apresentados na Tabela 13. O tratamento dos valores na tabela foi realizado conforme mostrado na Subseção 6.1.4, levando em conta somente as autoavaliações relacionadas aos aspectos aplicados pelas métricas.

No jogo *MyName*, três aspectos de aprendizagem foram avaliados: “Habilidade Mecânica”, “Habilidade Lógica” e “Tentativas por Fase”. Assim como no jogo *Robô Euroi*, os estudantes se autoavaliaram com uma “Habilidade Mecânica” baixa. Acredita-se que essa percepção esteja associada principalmente aos resultados obtidos nas últimas fases, que apresentavam maior nível de complexidade. Nessas fases, os jogadores precisaram realizar muitas tentativas e foram derrotados várias vezes, mas cometeram poucos erros mecânicos. Por exemplo, desviaram de muitos inimigos, mas foram atingidos apenas uma vez, o que resultou na perda do jogo. Consequentemente, tiveram mais acertos que erros, mas, devido às derrotas, a percepção geral foi de um desempenho ruim.

Em relação à “Habilidade Lógica”, a autoavaliação dos estudantes indicou um

valor médio de 0,79, sugerindo que eles acreditavam ter feito boas escolhas. No contexto do jogo, isso significa otimizar a mochila. A média das métricas por fase foi de 0,72, resultando em uma diferença de 0,07 pontos. Essa proximidade sugere uma concordância razoável entre a percepção dos jogadores e os dados registrados, especialmente nas fases finais, que tendem a influenciar mais fortemente a autoavaliação por serem mais recentes na memória dos participantes.

Quanto às “Tentativas por Fase”, os estudantes indicaram um valor médio alto de 0,77, mostrando que encontraram dificuldades, mas se esforçaram para resolver os problemas. Se fosse possível analisar a “Flexibilidade Cognitiva” nesse contexto, entenderíamos melhor as dificuldades enfrentadas, avaliando se os estudantes estavam presos a uma única solução ou buscaram alternativas. O esforço pode ser visto negativamente se os estudantes persistirem em uma estratégia ineficaz.

Dois casos foram analisados em mais detalhes, o Estudante 1, ilustrado na Figura 18 e o Estudante 9, ilustrado na Figura 19, e descritos a seguir:

1. **Estudante 1:** avaliou-se como tendo poucas dificuldades nos desafios lógicos e mecânicos, sempre escolhendo as melhores soluções e não enfrentando problemas com os inimigos. As métricas confirmam sua percepção na “Habilidade Mecânica”, com apenas uma tentativa por fase. No entanto, há uma discrepância na “Habilidade Lógica”, onde os dados mostram que suas escolhas nem sempre foram as melhores, evidenciando a subjetividade da avaliação por questionário.
2. **Estudante 9:** sugeriu forte subestimação em sua autoavaliação de habilidade mecânica, atribuindo-se nota 0,2, enquanto as métricas mostraram desempenho consistentemente elevado ao longo das fases. Além disso, indicou um alto esforço (0,8), embora o número de tentativas tenha sido baixo na maior parte do jogo, com exceção da fase 4. A avaliação da habilidade lógica (0,6) esteve próxima dos dados coletados, o que sugere que sua percepção foi mais precisa nesse aspecto. Esse caso reforça como a autoimagem do jogador pode ser influenciada por percepções subjetivas que não correspondem ao desempenho real.

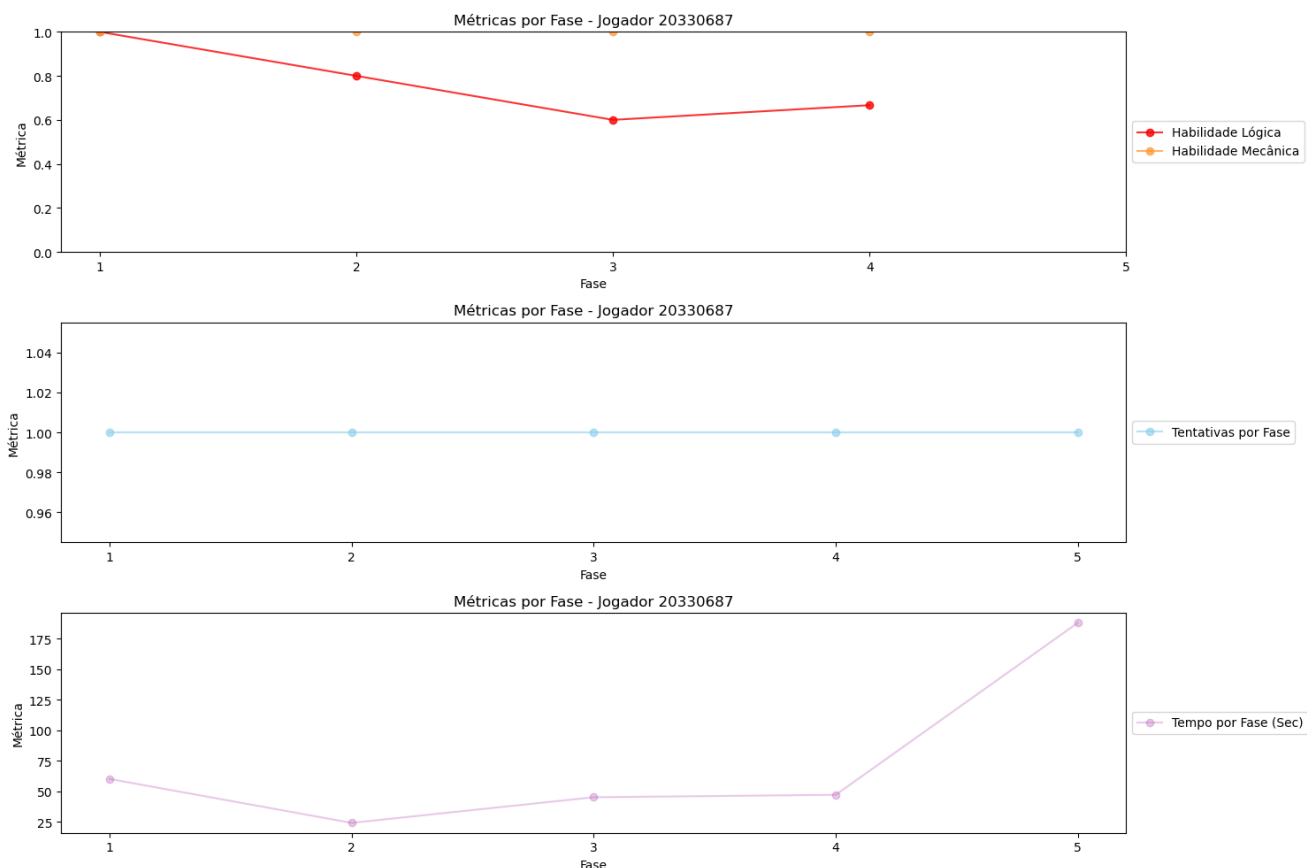


Figura 18 – Gráficos de desempenho do Estudante 1 no jogo *MyName*.

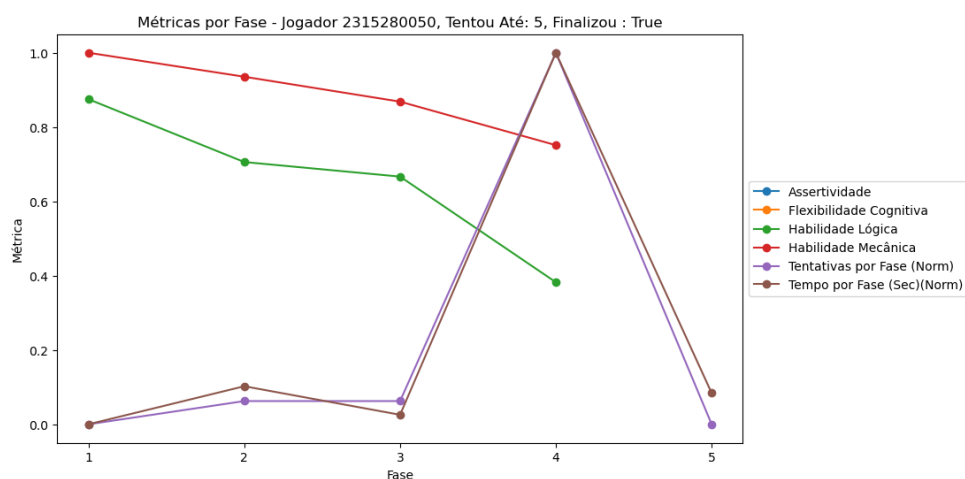


Figura 19 – Gráficos de desempenho do Estudante 9 no jogo *MyName*.

6.2.5 Validação com especialistas

A validação com especialistas seguiu o processo metodológico descrito na Seção 5.6 (Capítulo 4.1). Participaram três especialistas da área de educação, três da área de análise de dados e, de forma complementar, dois agentes inteligentes com enfoque psicométrico (ChatGPT 4.5 e Gemini 2.5 Pro). Os resultados a seguir apresentam, para especialistas

humanos, os valores do coeficiente V de Aiken e, para os agentes inteligentes, uma síntese qualitativa dos comentários e justificativas.

6.2.6 Grupo professores

A Tabela 14 exibe os dados brutos coletados a partir das respostas dos professores, juntamente com o cálculo do coeficiente de Aiken's V .

Ao contrário dos analistas de dados, os professores receberam uma introdução às métricas focada em sua lógica conceitual e em exemplos práticos, em vez de um detalhamento técnico do cálculo. Essa abordagem buscou avaliar a clareza e a relevância pedagógica das métricas do ponto de vista docente, sem exigir um conhecimento técnico profundo dos processos de mensuração. Como resultado, foi percebida uma tendência a valores menores em áreas relacionadas à compreensão do cálculo e à aplicabilidade prática, o que reflete um foco em aspectos pedagógicos em vez de uma limitação técnica.

		ASSERTIVIDADE				
Aspecto	N1	N2	N3	Aiken's V	V-FINAL	
Representatividade	5	5	5	1,00	ACEITO SEM RESTR.	
Compreensão	5	5	5	1,00	ACEITO SEM RESTR.	
Clareza do cálculo	5	5	1	0,67	REVISAR	
Aplicabilidade	5	5	1	0,67	REVISAR	
Validade	5	4	2	0,67	REVISAR	
Utilidade	3	3	5	0,67	REVISAR	
Aplicação pedagógica	5	4	2	0,67	REVISAR	

		FLEXIBILIDADE COGNITIVA				
Aspecto	N1	N2	N3	I-CVI	Aiken's V	V-FINAL
Representatividade	5	5	5	1,00	1,00	ACEITO SEM RESTR.
Compreensão	5	5	4	1,00	0,92	ACEITO SEM RESTR.
Clareza do cálculo	5	5	5	1,00	1,00	ACEITO SEM RESTR.
Aplicabilidade	5	5	4	1,00	0,92	ACEITO SEM RESTR.
Validade	5	4	3	0,67	0,75	ACEITO COM RESTR.
Utilidade	5	3	2	0,33	0,58	REVISAR
Aplicação pedagógica	5	2	2	0,33	0,50	REVISAR

		HABILIDADE LÓGICA				
Aspecto	N1	N2	N3	Aiken's V	V-FINAL	
Representatividade	5	5	5	1,00	ACEITO SEM RESTR.	
Compreensão	5	5	4	0,92	ACEITO SEM RESTR.	
Clareza do cálculo	5	5	1	0,67	REVISAR	
Aplicabilidade	5	5	1	0,67	REVISAR	
Validade	5	5	3	0,83	ACEITO SEM RESTR.	
Utilidade	5	5	3	0,83	ACEITO SEM RESTR.	
Aplicação pedagógica	5	5	3	0,83	ACEITO SEM RESTR.	

		EXPLORAÇÃO				
Aspecto	N1	N2	N3	I-CVI	Aiken's V	V-FINAL
Representatividade	5	5	5	1,00	1,00	ACEITO SEM RESTR.
Compreensão	5	5	5	1,00	1,00	ACEITO SEM RESTR.
Clareza do cálculo	5	5	1	0,67	0,67	REVISAR
Aplicabilidade	5	5	1	0,67	0,67	REVISAR
Validade	5	3	3	0,33	0,67	REVISAR
Utilidade	5	2	2	0,33	0,50	REVISAR
Aplicação pedagógica	5	3	2	0,33	0,58	REVISAR

		HABILADDE MECÂNICA				
Aspecto	N1	N2	N3	Aiken's V	V-FINAL	
Representatividade	5	5	5	1,00	ACEITO SEM RESTR.	
Compreensão	5	5	5	1,00	ACEITO SEM RESTR.	
Clareza do cálculo	5	5	1	0,67	REVISAR	
Aplicabilidade	5	5	1	0,67	REVISAR	
Validade	5	5	3	0,83	ACEITO SEM RESTR.	
Utilidade	5	5	2	0,75	ACEITO COM RESTR.	
Aplicação pedagógica	5	3	2	0,58	REVISAR	

		PRÓ-EFICIENCIA CURRICULAR				
Aspecto	N1	N2	N3	I-CVI	Aiken's V	V-FINAL
Representatividade	5	5	5	1,00	1,00	ACEITO SEM RESTR.
Compreensão	5	5	5	1,00	1,00	ACEITO SEM RESTR.
Clareza do cálculo	5	5	1	0,67	0,67	REVISAR
Aplicabilidade	5	5	1	0,67	0,67	REVISAR
Validade	5	5	2	0,67	0,75	ACEITO COM RESTR.
Utilidade	5	5	3	0,67	0,83	ACEITO SEM RESTR.
Aplicação pedagógica	5	3	2	0,33	0,58	REVISAR

Tabela 14 – Resultados do coeficiente Aiken's V para as respostas do questionário aplicado aos professores.

Assertividade. A métrica apresentou excelente aceitação nos critérios de representatividade ($V=1,00$) e compreensão ($V=1,00$), sendo aceita sem restrições. Os avaliadores reconheceram que o nome da métrica expressa com clareza a precisão das decisões do jogador e sua relação com o desempenho nas atividades.

Aspectos como clareza do cálculo, aplicabilidade, validade, utilidade e aplicação pedagógica ($V \approx 0,67$) foram classificados como necessitando de revisão. As respostas indicaram dúvidas sobre a penalização de erros e o tratamento de escolhas parcialmente corretas, sugerindo ajustes para considerar o progresso do jogador ao longo das tentativas.

Apesar disso, a métrica foi considerada capaz de gerar indícios para a identificação de padrões de erro e para o apoio a intervenções pedagógicas. Observou-se concordância entre os avaliadores de que, isoladamente, não deve compor a nota do aluno, uma vez que o desempenho no jogo depende de múltiplos fatores, além do domínio curricular. De modo geral, a métrica de Assertividade foi considerada clara e pertinente, adequada a jogos lineares, mas requer revisão para refletir melhor o processo de aprendizagem e reduzir vieses associados ao estilo exploratório do jogador.

Flexibilidade Cognitiva. A métrica apresentou alta aceitação geral, com elevados valores nos critérios de representatividade ($V=1,00$), clareza do cálculo ($V=1,00$), compreensão ($V=0,92$) e aplicabilidade ($V=0,92$), todos aceitos sem restrições. Os avaliadores compreenderam a proposta e reconheceram sua pertinência para analisar a capacidade dos estudantes de alterar estratégias e adaptar-se a novas situações.

A validade ($V=0,75$) foi aceita com restrição, indicando dúvidas sobre a fronteira entre flexibilidade e tentativa e erro. A utilidade ($V=0,58$) e a aplicação pedagógica ($V=0,50$) foram revistas, refletindo incerteza quanto ao uso prático em sala de aula. As respostas qualitativas reforçaram a necessidade de distinguir a exploração aleatória da adaptação intencional e de esclarecer o cálculo da métrica.

De modo geral, a Flexibilidade Cognitiva foi considerada clara e relevante, mas requer aprimoramentos conceituais para evidenciar a diferença entre mudança estratégica e simples experimentação.

Habilidade Lógica. A métrica apresentou excelente aceitação, com altos valores

nos critérios de representatividade ($V=1,00$), compreensão ($V=0,92$), validade ($V=0,83$), utilidade ($V=0,83$) e aplicação pedagógica ($V=0,83$), todos aceitos sem restrições. Os avaliadores compreenderam o propósito da métrica e reconheceram sua pertinência para avaliar o raciocínio lógico dos estudantes em situações de resolução de problemas.

A clareza do cálculo e a aplicabilidade ($V=0,67$) foram classificadas como "Revisar", indicando dúvidas sobre a operacionalização e a interpretação prática dos resultados. As observações qualitativas sugerem a necessidade de detalhar de que forma o cálculo reflete a complexidade das ações dos jogadores.

De modo geral, a "Habilidade Lógica" foi considerada clara, útil e bem estruturada, adequada para representar o raciocínio analítico dos estudantes e apoiar a identificação de dificuldades cognitivas, desde que acompanhada de orientações mais precisas para a aplicação, especialmente quanto à clareza do cálculo e à aplicabilidade.

Exploração. A métrica apresentou alta aceitação nos critérios de representatividade ($V=1,00$) e compreensão ($V=1,00$), sendo aceita sem restrições. Os avaliadores reconheceram que ela reflete o comportamento exploratório dos estudantes durante o jogo.

Os demais aspectos apresentaram valores entre $V=0,50$ e $V=0,67$, classificados como "revisar". As respostas qualitativas mostraram incertezas sobre a relação entre exploração e aprendizagem, apontando confusão entre comportamento exploratório e desempenho. Um avaliador questionou se maior engajamento implicaria menor assertividade, destacando que a exploração pode refletir perfil e não necessariamente progresso cognitivo.

Apesar das dúvidas, os professores reconheceram o potencial pedagógico na métrica, especialmente para identificar perfis de jogadores e ajustar estratégias de ensino, embora tenham concordado que ela não deve compor as notas. Em síntese, a Exploração foi considerada relevante e clara, mas requer revisão conceitual para explicitar sua relação com o aprendizado e diferenciar o engajamento da experimentação.

Habilidade Mecânica. A métrica apresentou altos níveis de aceitação nos critérios de representatividade ($V=1,00$), compreensão ($V=1,00$) e validade ($V=0,83$), sendo aceita sem restrições. O aspecto de utilidade ($V=0,75$) foi aceito com restrição, enquanto

a clareza do cálculo, aplicabilidade e aplicação pedagógica ($V \in [0,58; 0,67]$) foram classificadas como "Revisar".

Os avaliadores reconheceram o potencial da métrica para mensurar a precisão e o desempenho técnico, mas observaram que sua relevância pedagógica é limitada por não se vincular diretamente à aprendizagem de conteúdos curriculares. Um avaliador destacou que “habilidade mecânica não é algo que costumo medir em minhas disciplinas”.

Ainda assim, foi apontada como útil para identificar dificuldades operacionais e ajustar estratégias de ensino em atividades que envolvem coordenação, tempo de resposta e domínio de interface. De forma geral, a Habilidade Mecânica foi considerada clara e confiável, mas requer ajustes para ampliar sua pertinência pedagógica e integração com competências cognitivas e curriculares.

Pró-Eficiência Curricular. A métrica obteve altos índices de representatividade ($V=1,00$) e compreensão ($V=1,00$), sendo aceita sem restrição. A validade ($V=0,75$) foi aceita com restrição, enquanto a clareza do cálculo, a aplicabilidade e a aplicação pedagógica ($V \in [0,58; 0,67]$) foram classificadas como a. A utilidade apresentou bom desempenho ($V=0,83$), sendo aceita sem restrição.

Os avaliadores reconheceram seu potencial para indicar o domínio de conteúdos curriculares, mas solicitaram maior clareza nos procedimentos de cálculo e uso. As respostas qualitativas destacaram sua utilidade para identificar lacunas de aprendizagem e apoiar o planejamento pedagógico, como o reforço de conteúdos e a reorganização de atividades.

De modo geral, a “Pró-Eficiência Curricular foi considerada clara, relevante e alinhada ao acompanhamento do desempenho escolar, exigindo apenas ajustes conceituais para ampliar sua aplicabilidade e evitar sobreposição com métricas de raciocínio lógico.

Síntese geral. As análises indicam que as métricas apresentam boa clareza conceitual e relevância pedagógica, com destaque para os critérios de representatividade e compreensão ($V \approx 1,00$). Dimensões como aplicabilidade, clareza de cálculo e uso pedagógico mostraram menor consistência ($V \in [0,58; 0,67]$), sugerindo ajustes para

torná-las mais intuitivas e adequadas ao contexto escolar.

Os professores reconheceram o potencial das métricas para identificar dificuldades de aprendizagem e orientar intervenções, embora tenham apontado dúvidas quanto à interpretação dos resultados e ao seu vínculo com os conteúdos curriculares. Também ressaltaram a importância de explicitar o tipo de aprendizagem avaliado, seja cognitiva, procedimental ou curricular, para alinhar melhor os resultados aos objetivos educacionais.

Houve consenso de que as métricas são complementares e devem ser aplicadas de forma integrada, oferecendo uma visão mais ampla do desempenho dos estudantes. As combinações mais valorizadas foram “Pró-Eficiência Curricular” e “Habilidade Lógica”, seguidas por “Flexibilidade Cognitiva” e “Assertividade”, que relacionam estratégia cognitiva e domínio conceitual. A “Exploração” foi associada principalmente ao engajamento e não diretamente à aprendizagem.

Em síntese, o conjunto foi considerado promissor e pedagogicamente relevante, com boa base conceitual e alto potencial de aplicação. Recomenda-se aprimorar a clareza dos cálculos e incorporar dimensões temporais, como tempo de resolução e progressão, para ampliar seu valor diagnóstico em jogos educacionais.

6.2.7 Grupo analistas os dados

A Tabela 15 apresenta os dados brutos coletados das respostas dos analistas de dados, acompanhados do cálculo do coeficiente de V de Aiken.

Assertividade. A métrica apresentou excelente aceitação entre os avaliadores, com altos valores de clareza ($V=0,92$), representatividade ($V=1,00$), coerência ($V=0,83$) e adequação ($V=0,92$), todos classificados como aceitos sem restrição. As respostas qualitativas confirmaram a compreensão da métrica e sua coerência metodológica, especialmente em jogos com trajetórias lineares, como *Robô Euroi* e *MyName*.

Os aspectos de “utilidade” ($V=0,67$) e “sensibilidade” ($V=0,75$) foram aceitos com restrição, indicando que a métrica é eficaz em contextos diretos, mas menos sensível em jogos com múltiplos caminhos ou decisões estratégicas. Um dos avaliadores destacou

ASSERTIVIDADE						FLEXIBILIDADE COGNITIVA				
Aspecto	N1	N2	N3	Aiken's V	V-FINAL	N1	N2	N3	Aiken's V	V-FINAL
Clareza	4	5	5	0,92	ACEITO SEM RESTR.	4	5	5	0,92	ACEITO SEM RESTR.
Compreensão (*)	1	1	1	1	Compreensão Plena	1	1	1	1	Compreensão Plena
Utilidade	4	4	3	0,67	REVISAR	5	5	3	0,83	ACEITO SEM RESTR.
Representatividade	5	5	5	1,00	ACEITO SEM RESTR.	5	5	5	1,00	ACEITO SEM RESTR.
Coerência	5	4	4	0,83	ACEITO SEM RESTR.	5	5	4	0,92	ACEITO SEM RESTR.
Adequação	5	4	5	0,92	ACEITO SEM RESTR.	5	5	4	0,92	ACEITO SEM RESTR.
Fragilidade (*)	0	1	0	-0,17	Fragilidade Parcial	0	1	0	-0,17	Fragilidade Parcial
Generalização	3	4	5	0,75	ACEITO COM RESTR.	3	5	5	0,83	ACEITO SEM RESTR.
Contextualização	4	5	5	0,92	ACEITO SEM RESTR.	4	5	5	0,92	ACEITO SEM RESTR.
Sensibilidade	5	3	4	0,75	ACEITO COM RESTR.	5	3	4	0,75	ACEITO COM RESTR.
Validade	4	4	4	0,75	ACEITO COM RESTR.	5	4	4	0,83	ACEITO SEM RESTR.
EXPLORAÇÃO						HABILIDADE MECÂNICA				
Aspecto	N1	N2	N3	Aiken's V	V-FINAL	N1	N2	N3	Aiken's V	V-FINAL
Clareza	5	5	5	1,00	ACEITO SEM RESTR.	5	5	5	1,00	ACEITO SEM RESTR.
Compreensão (*)	1	1	1	1	Compreensão Plena	1	1	1	1	Compreensão Plena
Utilidade	3	4	5	0,75	ACEITO COM RESTR.	4	4	5	0,83	ACEITO SEM RESTR.
Representatividade	5	5	5	1,00	ACEITO SEM RESTR.	5	5	5	1,00	ACEITO SEM RESTR.
Coerência	4	4	4	0,75	ACEITO COM RESTR.	4	4	5	0,83	ACEITO SEM RESTR.
Adequação	2	4	3	0,50	REVISAR	4	4	5	0,83	ACEITO SEM RESTR.
Fragilidade (*)	1	0	1	-0,08	Fragilidade Parcial	0	1	0	-0,17	Fragilidade Parcial
Generalização	5	4	5	0,92	ACEITO SEM RESTR.	5	4	5	0,92	ACEITO SEM RESTR.
Contextualização	5	4	5	0,92	ACEITO SEM RESTR.	4	4	4	0,75	ACEITO COM RESTR.
Sensibilidade	5	4	4	0,83	ACEITO SEM RESTR.	5	4	4	0,83	ACEITO SEM RESTR.
Validade	2	4	4	0,58	REVISAR	4	4	4	0,75	ACEITO COM RESTR.
HABILIDADE LÓGICA						Pró-Eficiência Curricular				
Aspecto	N1	N2	N3	Aiken's V	V-FINAL	N1	N2	N3	Aiken's V	V-FINAL
Clareza	5	5	5	1	ACEITO SEM RESTR.	5	5	5	1	ACEITO SEM RESTR.
Compreensão (*)	1	1	1	1	Compreensão Plena	1	1	1	1	Compreensão Plena
Utilidade	5	4	4	0,83	ACEITO SEM RESTR.	4	5	5	0,92	ACEITO SEM RESTR.
Representatividade	5	5	5	1,00	ACEITO SEM RESTR.	4	5	5	0,92	ACEITO SEM RESTR.
Coerência	3	4	5	0,75	ACEITO COM RESTR.	3	5	4	0,75	ACEITO COM RESTR.
Adequação	2	4	4	0,58	REVISAR	3	5	4	0,75	ACEITO COM RESTR.
Fragilidade (*)	1	0	1	-0,08	Fragilidade Parcial	1	0	1	-0,08	Fragilidade Parcial
Generalização	3	4	5	0,75	ACEITO COM RESTR.	2	5	4	0,67	REVISAR
Contextualização	4	4	5	0,83	ACEITO SEM RESTR.	3	5	4	0,75	ACEITO COM RESTR.
Sensibilidade	5	4	4	0,83	ACEITO SEM RESTR.	5	3	4	0,75	ACEITO COM RESTR.
Validade	3	4	4	0,67	REVISAR	2	5	4	0,67	REVISAR

Tabela 15 – Resultados do coeficiente Aiken's V para as respostas do questionário aplicado aos analistas de dados.

que escolhas menos assertivas podem gerar vantagens futuras, sugerindo que a métrica pode ser complementada pela de “Exploração”.

De modo geral, a “Assertividade” foi considerada clara, objetiva e confiável, adequada para jogos lineares, mas requer cautela em aplicações com estruturas abertas.

Flexibilidade Cognitiva. A métrica apresentou alta aceitação geral, com índices elevados nos critérios de clareza ($V=0,92$), representatividade ($V=1,00$), coerência ($V=0,92$) e adequação ($V=0,92$), sendo aceita sem restrições nesses aspectos. Os avaliadores compreenderam a lógica da métrica e reconheceram seu potencial para medir a capacidade de adaptação e a mudança de estratégia dos jogadores.

Os comentários qualitativos reforçaram sua relevância conceitual, mas apontaram limitações práticas: a variação de ações nem sempre reflete a aprendizagem, podendo indicar apenas tentativas e erros aleatórios. Essa preocupação foi refletida nos valores de sensibilidade e validade ($V=0,75$), classificados como aceitos com restrição.

De modo geral, a Flexibilidade Cognitiva foi considerada uma métrica clara, pertinente e metodologicamente sólida, adequada para jogos lineares ou com caminhos definidos. Recomenda-se, contudo, aprimorar a distinção entre exploração significativa e experimentação aleatória, além de considerar o uso combinado com métricas como “Assertividade” para análises mais robustas.

Habilidade Lógica. A métrica foi bem compreendida e apresentou avaliação positiva nos critérios de clareza, representatividade e adequação ($V \geq 0,83$), sendo aceita sem restrições nesses aspectos. Os avaliadores destacaram que ela mede de forma direta o raciocínio lógico do jogador e é facilmente aplicável.

Entretanto, foram identificados pontos de melhoria em coerência ($V=0,75$) e validade ($V=0,67$), considerados aceitos com restrição. As respostas qualitativas indicam que a métrica, ao se basear apenas em acertos, pode simplificar excessivamente o raciocínio lógico, desconsiderando fatores como tempo, dificuldade e estratégia.

Foi sugerido incluir um fator de ponderação de desempenho, atribuindo pesos a desafios mais complexos ou a acertos parciais, e combinar essa métrica com outras, como “Flexibilidade Cognitiva” ou “Assertividade”, para capturar melhor o processo de aprendizagem.

De forma geral, a “Habilidade Lógica” foi considerada clara, útil e aplicável, mas necessita de ajustes para refletir com maior precisão a complexidade do raciocínio e

evitar aproximações excessivas a avaliações tradicionais.

Exploração. A métrica foi bem avaliada em clareza, representatividade e potencial de generalização ($V \geq 0,90$), sendo aceita sem restrições nesses aspectos. Os avaliadores compreenderam a lógica da métrica e destacaram sua utilidade para observar o comportamento exploratório e o engajamento dos jogadores, sobretudo em jogos que ofereçam possibilidades de exploração registradas nos dados de interação.

Entretanto, adequação ($V=0,50$) e validade ($V=0,58$) receberam avaliações mais baixas, indicando a necessidade de revisão. As análises qualitativas mostraram que, embora a métrica capture bem o interesse e a curiosidade, ela nem sempre reflete a aprendizagem em si, especialmente em jogos lineares ou quando o design limita as opções de exploração.

Os especialistas sugeriram introduzir pesos para ações de maior complexidade e diferenciar a exploração reflexiva da simples navegação superficial. De modo geral, a métrica foi considerada clara e relevante para medir o engajamento, mas requer ajustes conceituais e analíticos para representar com maior precisão a aprendizagem significativa.

Habilidade Mecânica. A métrica obteve alta consistência, com índices elevados de clareza, representatividade, coerência e adequação ($V \geq 0,83$), sendo aceita sem restrições na maioria dos aspectos. Os avaliadores compreenderam sua lógica e destacaram a utilidade da métrica para mensurar o desempenho técnico dos jogadores em diferentes contextos.

Apesar disso, a fragilidade parcial identificada por um avaliador e os valores limítrofes de validade ($V=0,75$) indicam a necessidade de ajustes. As análises qualitativas apontaram que o modelo atual tende a supervalorizar o número de sucessos, sem considerar a precisão, o tempo ou o esforço, o que pode gerar viés contra jogadores com menor experiência ou limitações motoras.

Recomenda-se a inclusão de pesos diferenciados para ações de maior complexidade e a incorporação de indicadores, como tempo de resposta e evolução entre tentativas. Em síntese, a “Habilidade Mecânica” foi considerada clara, confiável e aplicável, mas requer refinamentos para capturar nuances de desempenho e reduzir vieses

relacionados à experiência prévia com jogos digitais.

Pró-Eficiência Curricular. A métrica foi avaliada positivamente quanto à clareza ($V=1$) e representatividade ($V=0,92$), sendo aceita sem restrições nesses aspectos. Os avaliadores compreenderam seu propósito e reconheceram seu potencial para mensurar o domínio de conteúdos explicitamente vinculados ao currículo.

Entretanto, a coerência, a adequação e a validade, apresentaram valores entre $V=0,67$ e $V=0,75$, classificando-se como aceitas com restrição. As análises qualitativas indicam que a métrica pode se sobrepor à “Habilidade Lógica” em jogos baseados em raciocínio explícito, além de subestimar dimensões não curriculares, como criatividade, colaboração ou pensamento crítico.

Os especialistas sugeriram ajustar o modelo para permitir maior flexibilidade de aplicação, atribuindo pesos às ações conforme o tipo de conteúdo abordado e considerando sua integração com outras métricas. De forma geral, a “Pró-Eficiência Curricular” foi considerada clara e relevante, mas necessita de refinamentos conceituais para evitar redundâncias e ampliar sua abrangência interpretativa.

Síntese geral. As análises indicam que as métricas apresentam clareza conceitual, coerência metodológica e boa consistência entre os avaliadores. O coeficiente de Aiken confirmou consenso nos critérios de clareza, representatividade e coerência ($V \geq 0,80$), enquanto dimensões como utilidade, validade e adequação apresentaram valores próximos ao limiar, sugerindo ajustes conceituais e analíticos pontuais.

As respostas qualitativas reforçaram que as métricas são complementares e devem ser aplicadas de forma integrada, ampliando a compreensão da aprendizagem. Combinações como “Assertividade” + “Flexibilidade Cognitiva” e “Pró-Eficiência Curricular” + “Flexibilidade Cognitiva” foram destacadas por permitirem observar a adaptação e a relação entre desempenho e estratégia.

Os avaliadores sugeriram o desenvolvimento de métricas derivadas, como “Foco” e “Inatividade”, voltadas para a análise de engajamento e atenção. Ressaltaram ainda o potencial diagnóstico das métricas, úteis para identificar evidências de aprendizagem e falhas no design pedagógico. Em síntese, o conjunto foi considerado válido e promissor, compondo uma base consistente para o monitoramento da aprendizagem em jogos

Aspecto	ASSERTIVIDADE		Flexibilidade Cognitiva		Habilidade Lógica		Exploração		Habilidade Mecânica		Pró-Eficiência Curricular	
	GPT	GEMINI	GPT	GEMINI	GPT	GEMINI	GPT	GEMINI	GPT	GEMINI	GPT	GEMINI
Clareza	4	4	4	5	4	5	4	5	4	5	4	5
Compreensão (*)	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Utilidade	4	3	3	4	3	4	2	2	2	4	4	5
Representatividade	3	3	4	5	4	2	5	5	5	4	3	5
Coerência	4	3	4	5	4	3	3	3	3	3	4	4
Adequação	3	3	5	4	4	2	3	3	3	3	4	3
Fragilidade	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Generalização	4	3	4	5	4	5	4	4	4	5	5	5
Contextualização	5	5	5	5	5	5	5	5	5	5	5	3
Sensibilidade	4	5	4	4	4	5	5	5	4	4	4	5
Validade	4	3	4	4	3	2	2	2	3	3	3	3

Tabela 16 – Resultados das respostas do questionário aplicado aos psicométricos.

educacionais, desde que aplicado de forma integrada e continuamente aprimorado.

6.2.8 Grupo psicométricos - Agentes Inteligentes

A Tabela 16 apresenta os dados brutos coletados a partir das respostas dos instrumentos psicométricos. As respostas qualitativas (textuais) não estão incluídas na tabela; contudo, foram consideradas na realização das análises subsequentes.

Assertividade. A métrica foi considerada clara, coerente e aplicável a jogos lineares, demonstrando bom potencial para representar a precisão nas decisões do jogador e sua relação com o desempenho. Também foi reconhecida como útil para identificar padrões de erro e apoiar intervenções pedagógicas quando interpretada no contexto adequado.

Os avaliadores destacaram a ambiguidade conceitual do termo “assertividade”, mais ligado ao comportamento social do que à eficiência estratégica, sugerindo alternativas como “Eficiência Estratégica” ou “Aderência ao Caminho Ideal”. Apontaram ainda que a métrica é sensível a microcorreções e à definição do “caminho perfeito”, o que pode gerar ruído. O Gemini observou que a agregação das tentativas elimina a dimensão temporal, prejudicando a análise da aprendizagem ao longo do tempo.

Em síntese, a métrica mede a eficiência em desafios, mas requer ajustes conceituais e técnicos. Recomenda-se incorporar ponderação temporal, controle de ruído e integração com outras métricas (como tempo e flexibilidade cognitiva) para aumentar a validade interpretativa e reduzir vieses decorrentes do design do jogo.

Flexibilidade Cognitiva. A métrica foi avaliada como clara, coerente e adequada para representar a variação estratégica entre tentativas. Ambos os avaliadores con-

sideraram o nome pertinente, embora o GPT tenha sugerido “Variação Estratégica” ou “Diversidade de Tentativas” como alternativas mais precisas. Reconheceu-se a sua utilidade para observar processos de adaptação e mudança de estratégia durante a aprendizagem.

Os principais ajustes envolvem a ausência de critérios qualitativos que diferenciem variações produtivas de variações aleatórias, além da sensibilidade a ruídos gerados por ações redundantes. O Gemini apontou ainda a instabilidade dos resultados quando o número de tentativas é reduzido. Ambos recomendaram associar os resultados à eficácia das tentativas (sucesso/falha) para evitar interpretações ambíguas.

De modo geral, a métrica representa adequadamente a adaptação cognitiva, mas requer aperfeiçoamentos para fortalecer sua validade. Sugere-se incluir ponderações por tipo de ação, controle de ruído e integração com métricas de sucesso e tempo para aprimorar a interpretação dos resultados.

Habilidade Lógica. A métrica foi considerada clara e compreensível, com coerência metodológica satisfatória. Os avaliadores reconheceram sua aplicabilidade para medir desempenho em desafios lógicos e seu potencial para identificar padrões de acerto em tarefas cognitivas.

Os principais ajustes dizem respeito à inadequação conceitual do nome, pois “Habilidade Lógica” remete a um traço psicológico amplo, enquanto a métrica mensura apenas o desempenho observável. Também foram apontadas limitações metodológicas, como a ausência de controle temporal, a falta de ponderação por dificuldade e o risco de distorções em desafios com diferentes níveis de complexidade.

No geral, a métrica é funcional e útil para descrever o desempenho lógico, mas necessita de aprimoramento psicométrico. A adoção de modelos da Teoria de Resposta ao Item (TRI), a inclusão de componentes temporais e a integração com métricas complementares podem torná-la mais sensível à aprendizagem processual.

Exploração. A métrica foi avaliada como clara e de fácil aplicação, representando bem o comportamento exploratório em jogos com elementos opcionais e registros de navegação/interação adequados. Destacou-se a sua utilidade para descrever o engajamento e a interação dos jogadores com o ambiente de jogo.

Os avaliadores, contudo, ressaltaram que ela mede cobertura e não aprendizagem, apresentando validade de construto limitada. O GPT observou a dependência do design do jogo, já que fases com mais áreas opcionais elevam o escore artificialmente. O Gemini destacou que a métrica não distingue a exploração intencional da navegação aleatória, o que compromete sua interpretação como indicador de aprendizagem.

Em síntese, a métrica descreve o engajamento, mas não os processos de aprendizagem. Recomenda-se normalizar pelo esforço (ex.: itens explorados por minuto), ponderar itens pela relevância pedagógica e integrar os resultados a métricas cognitivas, distinguindo a exploração metódica da exploração aleatória.

Habilidade Mecânica. A métrica foi considerada clara, coerente e útil para mensurar o desempenho motor, com potencial de aplicação em diferentes jogos que registrem eventos de execução e tempo de resposta de forma comparável. Foi reconhecida como adequada para avaliar a precisão, a coordenação e o tempo de resposta dos jogadores.

Os principais ajustes concentram-se na ausência de ponderação por dificuldade, apontada como a limitação psicométrica mais severa. A métrica trata todos os desafios como equivalentes, ignorando diferenças de complexidade e exigência motora. Além disso, pode mascarar o progresso entre tentativas e inflar os resultados com repetições sucessivas do mesmo desafio.

De modo geral, a métrica descreve bem o desempenho motor, mas requer maior sensibilidade para inferir a aprendizagem. Sugere-se aplicar pesos diferenciados por complexidade, incorporar medidas de tempo e consistência e integrar resultados com métricas cognitivas para capturar a relação entre habilidades motoras e desempenho geral.

Pró-Eficiência Curricular. A métrica foi avaliada como clara e metodologicamente coerente, representando, de forma satisfatória, o desempenho em desafios relacionados ao conteúdo curricular. Foi considerada útil para indicar o domínio conceitual e a aplicação prática do conhecimento.

As principais fragilidades estão na ausência de ponderação por dificuldade e na abordagem binária de avaliação (acerto/erro), que reduz a sensibilidade da métrica. O

GPT destacou que perguntas fáceis e difíceis têm o mesmo peso, e o Gemini reforçou a necessidade de calibrar desafios usando parâmetros da Teoria de Resposta ao Item (TRI) para garantir precisão na estimativa de proficiência.

Em síntese, a métrica é adequada para mensurar o desempenho curricular, mas requer ajustes psicométricos. A aplicação da TRI e a inclusão de níveis graduais de desempenho podem aprimorar sua validade e transformar o índice em uma medida mais justa e sensível da aprendizagem dos estudantes.

Síntese geral. As análises dos agentes inteligentes mostraram que as métricas formam um sistema complementar, abrangendo dimensões cognitivas, motoras e curriculares da aprendizagem em jogos educacionais. GPT e Gemini destacaram que o valor analítico do modelo depende da integração entre indicadores de processo e de resultado, reforçando o caráter multidimensional do modelo.

As métricas mais relevantes foram *Pró-Eficiência Curricular*, *Assertividade* e *Flexibilidade Cognitiva*, por refletirem domínio conceitual, eficiência cognitiva e adaptação estratégica. A Exploração e a Habilidade Mecânica foram vistas como complementares, úteis para contextualizar o desempenho, mas com menor validade direta para inferir aprendizagem.

As combinações entre métricas foram apontadas como o principal valor analítico do modelo, permitindo distinguir padrões de aprendizagem e desempenho. No conjunto, as métricas favorecem avaliações mais diagnósticas e interpretativas, apoiando professores e alunos na compreensão do processo de aprendizagem.

6.3 Discussão

Esta seção apresenta a discussão sobre os resultados alcançados com os dois estudos realizados para responder as questões de pesquisa propostas na introdução.

6.3.1 Como avaliar aspectos de aprendizagem a partir da interação de estudantes com jogos educacionais? (QP1)

Nesta pesquisa, foi apresentado um conjunto de métricas destinadas a auxiliar na avaliação do desempenho dos estudantes em aspectos específicos da aprendizagem, considerando suas interações com jogos educacionais. No presente capítulo, foram realizadas análises e discussões sobre os resultados parciais da aplicação dessas métricas em dois jogos: “Robô Euroi” e “MyName”. De forma geral, os estudantes obtiveram bons resultados nas métricas, especialmente nas habilidades mecânicas, nas quais os jogadores são frequentemente desafiados a resolver esse tipo de problema. As análises dos *logs* dos estudantes e os valores apresentados pelas métricas mostraram-se consistentes, evidenciando uma alta relação entre as interações e os valores mensurados.

As métricas permitiram a visualização baseada em dados sobre a evolução dos estudantes em determinados aspectos de aprendizagem, identificando momentos em que esses aspectos foram mais exigidos e onde surgiram dificuldades. Além disso, as métricas ajudaram a identificar pontos de melhoria nos jogos, como na fase 10 do “Robô Euroi1” e na fase 3 do “MyName”, que não foram projetadas para serem complexas. Observou-se também que os *designers* nem sempre consideram as habilidades que serão trabalhadas em cada fase do jogo durante o desenvolvimento. Isso foi evidenciado no caso do “Robô Euroi”. Embora o jogo se proponha a desenvolver habilidades lógicas, a métrica “Habilidade Lógica” não pôde ser aplicada em algumas fases, indicando possíveis falhas no *game design*.

Entretanto, as métricas possuem limitações. Elas não se aplicam a todos os contextos de jogos, o que pode ocorrer devido à ausência do dados relacionados à métrica no jogo, por exemplo, nos jogos relatados nos estudos, que não possuem aspectos de exploração. No caso discutido na Seção 6.2.2, não foi possível integrar as métricas de flexibilidade cognitiva e assertividade devido a restrições na aplicação das métricas. Isso resultou na perda de dados que poderiam fornecer mais *insights* para a validação das métricas. Outro ponto de atenção é a dependência da correta integração das métricas pelo desenvolvedor. Caso a integração não seja realizada corretamente, podem surgir valores incorretos. Isso ocorreu no caso das habilidades mecânica e lógica,

onde uma má definição dos desafios pode resultar em valores inconsistentes com a interação real do jogador com o jogo.

6.3.1.1 Validação das métricas de aprendizagem com professores, analistas e agentes inteligentes

A avaliação das métricas por parte de professores e analistas de dados fortaleceu a viabilidade do método proposto nesta pesquisa para a avaliação de aspectos de aprendizagem a partir da interação de estudantes com jogos educacionais. Em termos gerais, as métricas demonstraram elevados níveis de clareza conceitual, coerência metodológica e representatividade ($V \geq 0,80$), o que indica que os avaliadores compreenderam os construtos envolvidos e reconheceram sua pertinência na mensuração de dimensões cognitivas e curriculares no contexto dos jogos.

As análises qualitativas revelaram que professores e analistas consideram as métricas como ferramentas complementares, capazes de proporcionar uma visão integrada da aprendizagem. As combinações entre “Pró-Eficiência Curricular”, “Habilidade Lógica”, “Flexibilidade Cognitiva” e “Assertividade” foram especialmente valorizadas, pois permitem associar o desempenho do jogador à adaptação estratégica e ao domínio de conteúdo. Destacou-se também a utilidade diagnóstica das métricas, que podem auxiliar os professores na identificação de lacunas de aprendizagem, no ajuste de atividades e na compreensão de diferentes perfis de jogadores. Nesse sentido, as métricas não têm a finalidade de substituir avaliações subjetivas (por exemplo, observação docente e questionários), mas de sistematizar evidências registradas durante o *gameplay* e ampliar a base de informações disponível para interpretação.

A etapa de validação conduzida com os agentes inteligentes (ChatGPT 4.5 e Gemini 2.5 Pro) complementou essas análises, fornecendo uma avaliação psicométrica mais detalhada das propriedades conceituais e técnicas das métricas. Os resultados indicaram clareza nas definições e coerência metodológica, mas também revelaram limitações, tais como a ausência de ponderação por dificuldade, perda de informação temporal e ambiguidades em termos como “assertividade” e “habilidade lógica”. As

análises sugeriram aprimoramentos envolvendo controle de ruído, integração entre métricas e adoção de modelos baseados na Teoria de Resposta ao Item (TRI) para aumentar a validade de construto e a sensibilidade à aprendizagem processual.

De forma integrada, as três instâncias de validação (professores, analistas e agentes inteligentes) confirmam que as métricas desenvolvidas capturam indícios de aprendizagem de modo consistente, ainda que requeiram ajustes conceituais e técnicos para aprimorar sua precisão interpretativa. Tais resultados fortalecem a resposta à QP1, demonstrando que é viável avaliar aspectos de aprendizagem em jogos educacionais por meio de métricas baseadas em dados, desde que estas sejam sustentadas por critérios psicométricos e orientações pedagógicas e interpretadas em conjunto com outros instrumentos avaliativos.

6.3.2 Como a autoavaliação de desempenho dos estudantes se compara aos dados de interação coletados em jogos educacionais? (QP2)

Através dos experimentos realizados, foi possível observar que, em alguns casos, a autoavaliação dos jogadores, obtida por meio de questionários, distoa dos valores gerados pelas métricas. Isso pode ter ocorrido por diferentes fatores. Primeiramente, os estudantes podem considerar apenas uma etapa do jogo e não sua trajetória como um todo, focando geralmente em momentos de muita facilidade ou dificuldade. Isso pode influenciar a percepção do próprio desempenho e gerar respostas que não correspondem integralmente aos registros observáveis do *gameplay*.

Um exemplo é o estudante 9, que utilizou o jogo “MyName”. Embora tenha apresentado desempenho mecânico elevado em todas as fases, com valores próximos ao máximo na métrica de habilidade mecânica, essa habilidade foi autoavaliada como muito baixa (0,2). Essa subestimação contrasta com os dados coletados.

Além disso, foi indicado alto esforço (0,8), apesar de o número de tentativas ter permanecido baixo na maior parte do jogo, com exceção da fase 4. A habilidade lógica foi autoavaliada em 0,6, valor relativamente próximo ao desempenho observado.

Esse caso evidencia que a autoavaliação pode não apenas superestimar o desempenho, mas também refletir autocrítica excessiva, revelando diferentes vieses perceptivos entre os participantes.

É relevante destacar que divergências entre autoavaliação e métricas não implicam necessariamente erros por parte dos participantes. Parte dessas diferenças pode decorrer de limitações nas próprias métricas, seja quanto a sua sensibilidade às diferentes estratégias adotadas pelos jogadores, seja em relação à forma como foram operacionalizadas. Ainda que baseadas em fundamentos teóricos e testadas em jogos com estruturas distintas, reconhece-se que essas métricas podem não capturar aspectos subjetivos da experiência do jogador, como esforço percebido ou decisões intuitivas.

Síntese do capítulo

Este capítulo apresentou os resultados da aplicação de métricas de aprendizagem integradas ao *GLBoard* em dois jogos educacionais com dinâmicas distintas: “Robô Euroi” e *MyName*. Foram descritos o mapeamento de dados e as adaptações necessárias à captura das interações, evidenciando que a aplicabilidade das métricas depende do *game design* e da instrumentação do jogo.

No “Robô Euroi”, seis das oito métricas foram aplicadas; *Exploração* e *Pró-Eficiência Curricular* não foram utilizadas por incompatibilidade e redundância. As maiores complexidades concentraram-se em *Assertividade* (definição do caminho ótimo por fase) e *Habilidade Lógica* (delimitação de decisões críticas). Os resultados por fase acompanharam a progressão de dificuldade e indicaram possíveis desequilíbrios de *design*, com destaque para a fase 10. Também se observou distribuição irregular de desafios lógicos, limitando a aplicação contínua da *Habilidade Lógica*.

Em *MyName*, a aleatoriedade por tentativa inviabilizou *Assertividade* e *Flexibilidade Cognitiva*. Em contrapartida, *Habilidade Lógica* foi viabilizada por um algoritmo que calcula a solução ótima em cada tentativa, e *Habilidade Mecânica* foi mensurada por ajustes em *scripts* e *triggers* para identificar ataques. Observou-se queda do desempenho lógico até a fase 4, melhora gradual do desempenho mecânico e picos de tentativas

e tempo em fases específicas. Na fase final, a análise foi limitada por um *bug* e pela dificuldade de definir a melhor escolha sob aleatoriedade contínua.

A comparação entre métricas e questionário autoavaliativo mostrou convergências parciais e discrepâncias associadas a vieses perceptivos, efeito de recência e avaliação baseada em momentos específicos do jogo, reconhecendo-se também que limitações das métricas podem contribuir para essas diferenças. A validação com professores e analistas (coeficiente *V* de Aiken) indicou alta aceitação conceitual, mas apontou necessidade de refinamento em clareza de cálculo, aplicabilidade e uso pedagógico em parte das métricas. As análises psicométricas de agentes inteligentes reforçaram limitações técnicas (ausência de ponderação por dificuldade, perda de informação temporal e ambiguidades terminológicas) e sugeriram controle de ruído, integração entre métricas e abordagens inspiradas na TRI.

Em síntese, os resultados sustentam a viabilidade de avaliar aspectos de aprendizagem a partir de dados de interação em jogos educacionais, delimitando condições de aplicação e ajustes necessários para ampliar a precisão interpretativa e a utilidade educacional das métricas.

7

CONSIDERAÇÕES FINAIS

A utilização de jogos educacionais tem sido ampliada no contexto escolar, por possibilitar experiências interativas mediadas por tecnologias digitais. Contudo, ainda são limitadas as abordagens sistemáticas voltadas à análise do que os estudantes mobilizam durante o *gameplay*. Em muitos casos, a avaliação concentra-se em percepções gerais ou em resultados finais, com menor atenção aos dados produzidos ao longo da interação. Essa lacuna dificulta a identificação de padrões de desempenho, dificuldades recorrentes e possíveis indícios de aprendizagem ao longo da experiência.

Diante desse contexto, esta dissertação foi orientada pelo seguinte problema de pesquisa: como avaliar, de maneira objetiva e reprodutível, indícios de aprendizagem em jogos educacionais a partir de dados de interação? O objetivo geral consistiu em propor um conjunto de métricas que, a partir dos *logs* gerados pelas interações dos estudantes com jogos educacionais, permitam gerar indícios para a avaliação de desempenho em determinados aspectos de aprendizagem, aplicáveis em diferentes contextos.

Como objetivos específicos, buscou-se: (i) propor métricas para avaliar o desempenho dos estudantes em aspectos de aprendizagem em jogos educacionais, com base em dados registrados nos *logs*; (ii) reestruturar uma ferramenta para capturar e organizar os *logs* das interações dos estudantes com jogos educacionais; (iii) conduzir um estudo comparativo entre a autoavaliação dos estudantes e os dados coletados nos *logs* de suas interações nos jogos educacionais; e (iv) validar as métricas propostas por meio da análise de especialistas das áreas de educação, análise de dados e psicometria, utilizando métodos tradicionais e agentes inteligentes.

Os objetivos foram atendidos em grande medida por meio de fundamentação teórica, definição operacional das métricas e aplicação de uma metodologia *Design-Based Research*, na qual a solução foi concebida, refinada e testada em ciclos. Nesse processo, o *GLBoard* foi remodelado para suportar a instrumentação requerida pelas métricas, com adaptações na arquitetura de dados para incorporar registros que não eram contemplados na versão original (por exemplo, itens/locais opcionais, informações temporais, referência de melhor caminho por fase e identificação estruturada de desafios lógicos, mecânicos e curriculares), além da atualização do pacote de integração aos jogos. Quanto ao objetivo de validação, foi realizada uma validação conceitual com especialistas e análises complementares por agentes inteligentes; contudo, não se pretende afirmar uma validação completa das métricas, a qual demanda estudos adicionais com amostras ampliadas, mais jogos e procedimentos estatísticos específicos.

No experimento com o jogo “Robô Euroi”, a maior parte das métricas pôde ser aplicada, e os valores obtidos mostraram coerência com os registros de interação, evidenciando variações associadas a fases específicas (incluindo fases em que o nível de desafio observado divergiu do esperado pelo *design*). Também foi identificada uma distribuição irregular de desafios compatíveis com alguns aspectos, o que limitou a aplicação contínua de determinadas métricas em parte das fases. No jogo “MyName”, também foram observadas relações consistentes entre indicadores e *logs*; entretanto, características do *design* (como aleatoriedade por tentativa) restringiram a aplicação de métricas que dependem de trajetórias ideais estáveis, e a análise de um trecho final foi parcialmente limitada por problemas de execução do jogo. Adicionalmente, emergiram discrepâncias entre autoavaliação e indicadores baseados em dados, sugerindo influência de vieses perceptivos e de momentos marcantes da experiência.

Do ponto de vista teórico, os resultados indicam que dados de interação podem ser utilizados para identificar indícios de aprendizagem, desde que haja correspondência entre eventos do *gameplay* e os aspectos analisados, que os jogos sejam instrumentados de forma compatível e que as métricas sejam interpretadas de maneira articulada. Ressalta-se, contudo, que as métricas não mensuram aprendizagem de forma direta, mas oferecem evidências que podem contribuir para sua interpretação, especialmente

quando consideradas em conjunto com outros elementos de análise.

No âmbito prático, a remodelação do *GLBoard* e a definição das métricas configuram uma base para análise de jogos com instrumentação adequada e com dados padronizados, mitigando a dependência de soluções *ad hoc*. Em particular, o *GLBoard* passa a apoiar a coleta, o armazenamento e a análise dos dados necessários às métricas, possibilitando sua utilização em jogos distintos mediante mapeamento explícito dos eventos do jogo para a arquitetura proposta.

Professores podem utilizar os indicadores como apoio à identificação de padrões de interação, dificuldades e perfis de resposta; desenvolvedores e *designers* podem localizar pontos do *design* associados a maior concentração de tentativas, tempo ou falhas e ajustar progressão e desafios; e analistas podem trabalhar com dados estruturados para fins de acompanhamento e comparação. Assim, a pesquisa indica a viabilidade de estruturar a análise de indícios de aprendizagem a partir de *logs* com suporte tecnológico e critérios interpretativos definidos.

7.1 Limitações

Entre as limitações metodológicas, destaca-se o uso de questionários de autoavaliação como instrumento complementar, uma vez que tais respostas refletem percepções subjetivas. Além disso, os experimentos foram realizados em contextos específicos e em períodos delimitados, o que restringe inferências sobre processos de aprendizagem em longo prazo. Nesse sentido, os resultados devem ser interpretados como indícios derivados da interação, e não como medidas diretas de aprendizagem.

Há também limitações associadas ao *game design*. A estrutura das fases, a distribuição de desafios e os tipos de interação disponíveis influenciam os eventos registrados e, conseqüentemente, os indicadores derivados. Isso afeta a comparabilidade entre jogos com mecânicas distintas e contribui para que algumas métricas não sejam aplicáveis em determinados jogos ou fases (por exemplo, ausência de elementos de exploração ou de desafios lógicos em fases específicas).

No jogo *MyName*, a aleatoriedade por tentativa também impôs restrições à

aplicação de métricas que dependem de caminhos ideais pré-definidos, o que evidencia que parte do conjunto de métricas está condicionada ao tipo de estrutura do jogo e ao modelo de geração de desafios. Adicionalmente, a indisponibilidade de métricas na fase final, em função de um *bug* e das características da fase, limitou a análise completa do desempenho nesse trecho do jogo.

As restrições amostrais constituem outro limite, pois o número de participantes e de jogos analisados foi reduzido. Além disso, no experimento com “Robô Euroi”, os participantes não pertenciam ao público-alvo original do jogo, o que tende a influenciar a dificuldade percebida e, por consequência, o comportamento dos indicadores. Assim, os resultados devem ser interpretados considerando o contexto em que foram obtidos.

No plano psicométrico, embora a validação com especialistas tenha indicado clareza e representatividade conceitual, o procedimento adotado não permite generalizações amplas. As análises complementares realizadas por agentes inteligentes tiveram caráter descritivo e não substituem estudos com amostras ampliadas e procedimentos estatísticos adicionais. Também se observa que parte das decisões de modelagem (por exemplo, definição do que constitui um desafio lógico ou mecânico, e a delimitação de escolhas críticas) depende de critérios implementados no jogo, o que pode introduzir variação entre implementações.

Por fim, a aplicação das métricas depende da qualidade da instrumentação e da integridade dos *logs*. A ausência de determinados registros, observada em um dos experimentos, limitou a aplicação de parte das métricas, evidenciando que a análise está condicionada à disponibilidade e consistência dos dados e à forma como os eventos são mapeados para a estrutura do *GLBoard*.

7.2 Contribuições e Trabalhos Futuros

Como contribuição científica, a pesquisa sistematiza um conjunto de métricas voltadas à identificação de indícios de aprendizagem a partir de dados de interação, explicitando a relação entre eventos do *gameplay* e aspectos analisados. No plano metodológico, apresenta um processo de definição, validação e aplicação dessas métricas em con-

texto real, bem como a articulação entre indicadores derivados de *logs* e autoavaliação, evidenciando convergências e discrepâncias relevantes para a interpretação.

Como contribuição prática, a remodelação do *GLBoard* oferece uma estrutura de padronização para captura e análise de dados, favorecendo sua utilização em jogos com características distintas, desde que devidamente instrumentados. Os resultados também indicam que a própria análise por métricas pode apoiar a identificação de trechos do jogo com maior concentração de tentativas, tempo ou falhas, contribuindo para ajustes no *design* e para uso formativo em contextos educacionais.

Como possibilidades de continuidade, sugerem-se: (i) ampliar a aplicação das métricas em diferentes jogos, conteúdos e públicos, incluindo o público-alvo original dos jogos avaliados, para observar o comportamento dos indicadores em perfis mais diversos; (ii) investigar formas de considerar níveis de dificuldade na modelagem dos indicadores, inclusive com ponderações por tipo de desafio; (iii) incorporar análises que contemplem a dimensão temporal do *gameplay* (por exemplo, evolução entre tentativas e progressão intra-fase), reduzindo a perda de informação processual; (iv) explorar aproximações com modelos psicométricos, como a Teoria de Resposta ao Item, quando pertinente, para refinar interpretações e calibrações; (v) aperfeiçoar diretrizes de instrumentação no *GLBoard* e criar verificações automáticas de integridade dos *logs*, reduzindo perda de dados e inconsistências; (vi) estender a arquitetura do *GLBoard* para lidar com jogos com aleatoriedade por tentativa, permitindo, quando possível, registrar referências ótimas calculadas dinamicamente; e (vii) combinar os indicadores com outros instrumentos de avaliação, como pré/pós-testes, observações sistemáticas e análise qualitativa do *gameplay*, ampliando a base de evidências para a interpretação dos indícios de aprendizagem.

REFERÊNCIAS

- AHMAD, N. A. et al. Learning management system instrument development based on aiken's v technique. *International Journal of Evaluation and Research in Education*, v. 13, n. 5, p. 3211–3219, 2024. [104](#)
- AIKEN, L. R. Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, v. 45, n. 1, p. 131–142, 1985. [102](#)
- AKBARI, H. et al. The effect of traditional games in fundamental motor skill development in 7-9 year old boys. *Iranian journal of pediatrics*, Tehran University of Medical Sciences Publications, v. 19, n. 2, p. 123–129, 2009. [75](#)
- ALENCAR, L.; MELO, R.; PESSOA, M. Uma proposta de análise de dados exploratoria para um jogo educacional de matemática. 2020. [57](#)
- ALONSO-FERNANDEZ, C. et al. Systematizing game learning analytics for serious games. In: IEEE. *2017 IEEE global engineering education conference (EDUCON)*. [S.l.], 2017. p. 1111–1118. [23](#), [25](#), [37](#), [41](#), [43](#)
- ALONSO-FERNANDEZ, C. et al. Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education*, Elsevier, v. 141, p. 103612, 2019. [57](#)
- ALONSO-FERNÁNDEZ, C. et al. Game learning analytics:: Blending visual and data mining techniques to improve serious games and to better understand player learning. *Journal of Learning Analytics*, v. 9, n. 3, p. 32–49, 2022. [38](#), [40](#)
- ALONSO-FERNÁNDEZ, C. et al. Evidence-based evaluation of a serious game to increase bullying awareness. *Interactive Learning Environments*, Taylor & Francis, v. 31, n. 2, p. 644–654, 2023. [40](#), [48](#)
- ANDRADE, A. N. de; NEGRÃO, F. da C.; VILAÇA, A. L. de A. O ensino remoto emergencial no amazonas nas lentes dos professores: Inclusão ou exclusão? 2021. [35](#)
- ARCAGÖK, S. The impact of game-based teaching practices in different curricula on academic achievement. *International Online Journal of Education and Teaching*, ERIC, v. 8, n. 2, p. 778–796, 2021. [22](#)
- ASMAWATI, L. The development of puzzle games for early childhood based on the banten local culture. *Jurnal Ilmiah Peuradeun*, v. 11, n. 2, p. 531–550, 2023. [22](#), [34](#)

- BAKER, R. S. J. de; INVENTADO, P. S. Chapter x: Educational data mining and learning analytics. *Computer Science*, p. 1–16, 2014. 39
- BALASUBRAMANIAN, N.; WILSON, B. G. Games and simulations. In: CITESEER. *Society for information technology and teacher education international conference*. [S.l.], 2006. v. 1. 34
- BANG, H. J.; LI, L.; FLYNN, K. Efficacy of an adaptive game-based math learning app to support personalized learning and improve early elementary school students' learning. *Early Childhood Education Journal*, Springer, v. 51, n. 4, p. 717–732, 2023. 78
- BARBA, P. G. de; OLIVEIRA, E. A.; ENGLISH, N. Development and validation of a learning analytics rubric for self-regulated learning. *Educational Technology Research and Development*, v. 73, p. 3223–3245, 2025. 99
- BARZILAI, S.; BLAU, I. Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, Elsevier, v. 70, p. 65–79, 2014. 76
- BECKER, K. What's the difference between gamification, serious games, educational games, and game-based learning. *Academia Letters*, v. 209, n. 2, 2021. 32
- BELLOTTI, F. et al. Assessment in and of serious games: An overview. *Advances in Human-Computer Interaction*, Wiley Online Library, v. 2013, n. 1, p. 136864, 2013. 37
- BLUMBERG, F. C.; BLUMBERG, F. *Learning by playing: Video gaming in education*. [S.l.]: OUP US, 2014. 33
- BUCKLEY, J. et al. Game-based assessment for education. *OECD Digital Education Outlook 2021 Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*, OECD Publishing, p. 195, 2021. 37
- BYUN, J.; JOUNG, E. Digital game-based learning for k–12 mathematics education: A meta-analysis. *School Science and Mathematics*, Wiley Online Library, v. 118, n. 3-4, p. 113–126, 2018. 76
- CALVO-MORATA, A. et al. Full software support for game learning analytics. *Learning Analytics Summer Institute Spain*, 2025. 25
- CAPRARA, L.; CAPRARA, C. Effects of virtual learning environments: A scoping review of literature. *Education and Information Technologies*, Springer, v. 27, n. 3, p. 3683–3722, 2022. 39
- CARDOSO, F. et al. Estudo sobre dados relevantes para o acompanhamento de participações em jogos educacionais. In: . [S.l.: s.n.], 2019. p. 922. 78
- CARRION, M. et al. Serious game, gamified applications, educational software: A comparative study. In: IEEE. *2019 International Conference on Information Systems and Software Technologies (ICI2ST)*. [S.l.], 2019. p. 55–62. 32, 33
- CARSTENSDOTTIR, E.; KLEINMAN, E.; EL-NASR, M. S. Player interaction in narrative games: structure and narrative progression mechanics. In: *Proceedings of the 14th international conference on the foundations of digital games*. [S.l.: s.n.], 2019. p. 1–9. 78

- CARVALHO, A. A. A. A teoria da flexibilidade cognitiva e o modelo múltiplas perspectivas. Universidade Federal Rural de Pernambuco (UFRPE), 2011. 84
- CASTRO, S. S.; SEVILLANO, M. P.; CADAVIECO, J. F. Learning analytics in serious games as predictors of linguistic competence in students at risk. *Technology, Knowledge and Learning*, Springer, p. 1–27, 2024. 50
- ÇEKEN, B.; TAŞKIN, N. Multimedia learning principles in different learning environments: A systematic review. *Smart Learning Environments*, Springer, v. 9, n. 1, p. 19, 2022. 39
- CHAARANI, B. et al. Association of video gaming with cognitive performance among children. *JAMA network open*, American Medical Association, v. 5, n. 10, p. e2235721–e2235721, 2022. 74
- CHAUDY, Y.; CONNOLLY, T.; HAINEY, T. Learning analytics in serious games: A review of the literature. In: *European Conference in the Applications of Enabling Technologies 2014*. [S.l.: s.n.], 2014. 39
- CHEN, H.-J. H.; HSU, H.-L. The impact of a serious game on vocabulary and content learning. *Computer Assisted Language Learning*, Routledge, v. 33, n. 7, p. 811–832, 2020. 76
- CHEUNG, S. Y.; NG, K. Y. Application of the educational game to enhance student learning. In: FRONTIERS MEDIA SA. *Frontiers in Education*. [S.l.], 2021. v. 6, p. 623793. 40
- CHIOTAKI, D.; POULOPOULOS, V.; KARPOUZIS, K. Adaptive game-based learning in education: a systematic review. *Frontiers in Computer Science*, Frontiers Media SA, v. 5, p. 1062350, 2023. 36
- CONATI, C.; JAQUES, N.; MUIR, M. Understanding attention to adaptive hints in educational games: an eye-tracking study. *International Journal of Artificial Intelligence in Education*, Springer, v. 23, p. 136–161, 2013. 78
- DELACRUZ, G. C. Impact of incentives on the use of feedback in educational videogames. cressst report 813. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*, ERIC, 2012. 78
- DEVELLIS, R. F. *Scale Development: Theory and Applications*. 4. ed. [S.l.]: SAGE Publications, 2016. 104
- EASTERDAY, M. W.; LEWIS, D. R.; GERBER, E. M. Design-based research process: Problems, phases, and applications. In: . [S.l.]: Boulder, CO: International Society of the Learning Sciences, 2014. 27
- ECK, R. V. Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE review*, Educause, v. 41, n. 2, p. 16, 2006. 38
- ENGELSTÄTTER, B.; WARD, M. R. Video games become more mainstream. *Entertainment Computing*, Elsevier, v. 42, p. 100494, 2022. 22
- FABRICATORE, C. Learning and videogames: An unexploited synergy. 2000. 34

- FARIAS, L. L. d. *Utilização de Game Learning Analytics para verificação do aprendizado em jogo sério voltado ao ensino de zoologia*. Dissertação (B.S. thesis) — Brasil, 2019. 39
- FLAVELL, J. H. *Cognitive development*. [S.l.]: Prentice-Hall, 1977. 74
- FLAVELL, J. H. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American psychologist*, American Psychological Association, v. 34, n. 10, p. 906, 1979. 35
- FREDRICKS, J. A.; BLUMENFELD, P. C.; PARIS, A. H. School engagement: Potential of the concept, state of the evidence. *Review of educational research*, Sage Publications Sage CA: Thousand Oaks, CA, v. 74, n. 1, p. 59–109, 2004. 73
- FREIRE, M. et al. Game learning analytics: learning analytics for serious games. In: *Learning, design, and technology*. [S.l.]: Springer Nature Switzerland AG, 2016. p. 1–29. 13, 23, 24, 37, 38, 40, 41, 42, 43, 54, 55
- FU, F.-L.; SU, R.-C.; YU, S.-C. Egameflow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, Elsevier, v. 52, n. 1, p. 101–112, 2009. 45
- FUSTER-GUILLÓ, A. et al. Evaluating impact on motivation and academic performance of a game-based learning experience using kahoot. *Frontiers in Psychology*, Frontiers, v. 10, p. 498489, 2019. 78
- GARMEN, P. et al. Can serious games measure your cognitive profile in adults? an innovative proposal to evaluate and stimulate cognitive skills. *Universal Access in the Information Society*, Springer, p. 1–12, 2022. 49
- GEE, J. P. What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, ACM New York, NY, USA, v. 1, n. 1, p. 20–20, 2003. 34
- GLORIA, A. D.; BELLOTTI, F.; BERTA, R. Serious games for education and training. *International Journal of Serious Games*, v. 1, n. 1, 2014. 37
- GOUWS, L. A.; BRADSHAW, K.; WENTWORTH, P. Computational thinking in educational activities: an evaluation of the educational game light-bot. In: *Proceedings of the 18th ACM conference on Innovation and technology in computer science education*. [S.l.: s.n.], 2013. p. 10–15. 78
- GRAPPIOLO, C. et al. Towards player adaptivity in a serious game for conflict resolution. In: IEEE. *2011 Third International Conference on Games and Virtual Worlds for Serious Applications*. [S.l.], 2011. p. 192–198. 78
- GRELLER, W.; DRACHSLER, H. Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, JSTOR, v. 15, n. 3, p. 42–57, 2012. 39
- GRIER, D.; LINDT, S. F.; MILLER, S. C. Formative assessment with game-based technology. *International Journal of Technology in Education and Science*, ERIC, v. 5, n. 2, p. 193–202, 2021. 23
- GUARDIA, D. L.; ARRIGO, M.; GIUSEPPE, O. D. A location-based serious game to learn about the culture. *Proc. FOE*, 2012. 78

- GUEDES, S. F. O uso de jogos como objetos digitais de aprendizagem: Um breve review. *INTERNATIONAL JOURNAL EDUCATION AND TEACHING (PDVL) ISSN 2595-2498*, v. 4, n. 2, p. 116–133, 2021. [33](#)
- GUI, Y. et al. Effectiveness of digital educational game and game design in stem learning: a meta-analytic review. *International Journal of STEM Education*, Springer, v. 10, n. 1, p. 36, 2023. [24](#)
- HAUGE, J. B. et al. Implications of learning analytics for serious game design. In: IEEE. *2014 IEEE 14th international conference on advanced learning technologies*. [S.l.], 2014. p. 230–232. [25](#)
- HAYNES, S. N.; RICHARD, D. C. S.; KUBANY, E. S. Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, v. 7, n. 3, p. 238–247, 1995. [100](#), [101](#), [104](#)
- HENDRIX, M. et al. Implementing adaptive game difficulty balancing in serious games. *IEEE Transactions on Games*, IEEE, v. 11, n. 4, p. 320–327, 2018. [78](#)
- HONDA, F. et al. Cadê minha pizza? um jogo para exercitar matemática e pensamento computacional através de grafos. In: SBC. *Anais Estendidos do XXI Simpósio Brasileiro de Jogos e Entretenimento Digital*. [S.l.], 2022. p. 876–885. [74](#)
- HOOKHAM, G.; NESBITT, K. A systematic review of the definition and measurement of engagement in serious games. In: *Proceedings of the australasian computer science week multiconference*. [S.l.: s.n.], 2019. p. 1–10. [80](#)
- HOOSHYAR, D.; YOUSEFI, M.; LIM, H. A systematic review of data-driven approaches in player modeling of educational games. *Artificial Intelligence Review*, Springer, v. 52, n. 3, p. 1997–2017, 2019. [23](#), [24](#), [57](#)
- HSIAO, H.-C. A brief review of digital games and learning. In: IEEE. *2007 First IEEE International Workshop on Digital Game and Intelligent Toy Enhanced Learning (DIGITEL'07)*. [S.l.], 2007. p. 124–129. [34](#)
- INVENTADO, P. S. et al. Contextual factors affecting hint utility. *International journal of STEM education*, Springer, v. 5, p. 1–13, 2018. [78](#)
- JÚNIOR, N. C. d. S. D.; BILLA, C. Z.; ADAMATTI, D. F. Descoberta de conhecimento em um jogo sério para o ensino de plantas industriais: um estudo de caso utilizando árvores de decisão. *Revista Brasileira de Computação Aplicada*, v. 13, n. 1, p. 98–111, 2021. [92](#)
- KIM, Y. J.; RUIPÉREZ-VALIENTE, J. A. Data-driven game design: The case of difficulty in educational games. In: SPRINGER. *Addressing Global Challenges and Quality Education: 15th European Conference on Technology Enhanced Learning, EC-TEL 2020, Heidelberg, Germany, September 14–18, 2020, Proceedings 15*. [S.l.], 2020. p. 449–454. [78](#)
- KLÖSGEN, W.; ZYTKOW, J. M. *Handbook of data mining and knowledge discovery*. [S.l.]: Oxford University Press, Inc., 2002. [39](#)
- KRAUSE, K. K. G.; HOUNSELL, M. da S.; GASPARINI, I. Um modelo para inter-relação entre funções executivas e elementos de jogos digitais. *Revista Brasileira de Informática na Educação*, v. 28, p. 596–625, 2020. [74](#)

- LANDER, N. et al. Accuracy vs. practicality of inertial measurement unit sensors to evaluate motor competence in children. *Frontiers in Sports and Active Living*, Frontiers, v. 4, p. 917340, 2022. [94](#)
- LAW, B. Puzzle games: a metaphor for computational thinking. In: ACADEMIC CONFERENCES INTERNATIONAL LIMITED. *European Conference on Games Based Learning*. [S.l.], 2016. p. 344. [34](#)
- LEVENSHTEIN, V. I. et al. Binary codes capable of correcting deletions, insertions, and reversals. In: SOVIET UNION. *Soviet physics doklady*. [S.l.], 1966. v. 10, n. 8, p. 707–710. [82](#)
- LI, S. et al. Are active video games useful in the development of gross motor skills among non-typically developing children? a meta-analysis. *BMC Sports Science, Medicine and Rehabilitation*, Springer, v. 14, n. 1, p. 140, 2022. [80](#)
- LI, Y.; CHEN, D.; DENG, X. The impact of digital educational games on student's motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. *PloS one*, v. 19, p. e0294350, 01 2024. [78](#)
- LI, Y.; CHEN, D.; DENG, X. The impact of digital educational games on student's motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. *Plos one*, Public Library of Science San Francisco, CA USA, v. 19, n. 1, p. e0294350, 2024. [23](#), [25](#), [73](#)
- LIANG, Z. More haste, less speed? relationship between response time and response accuracy in gamified online quizzes in an undergraduate engineering course. In: FRONTIERS MEDIA SA. *Frontiers in Education*. [S.l.], 2024. v. 9, p. 1412954. [81](#)
- LIMA, I.; BUENO, D.; PERRY, G. T. Avaliação de usabilidade e de experiência de jogo digital educacional: uma experiência com o suscity. *RENOTE*, v. 15, n. 1, 2017. [25](#)
- LOH, C. S.; SHENG, Y. Performance metrics for serious games: Will the (real) expert please step forward? In: IEEE. *Proceedings of CGAMES'2013 USA*. [S.l.], 2013. p. 202–206. [51](#)
- LOH, C. S.; SHENG, Y.; IFENTHALER, D. Serious games analytics: Theoretical framework. In: *Serious games analytics*. [S.l.]: Springer, 2015. p. 3–29. [38](#)
- LONGHI, M. T. et al. Aspectos socioafetivos na educação a distância: um olhar sobre o engajamento e a evasão. *Hard Tech*, 2021. [22](#)
- LOUREIRO, A. C. et al. Validade de conteúdo por avaliação de especialistas para medir as competências digitais de estudantes de mestrado em educação especial. *Texto Livre: Linguagem e Tecnologia*, v. 17, n. 4, p. e52564, 2024. [104](#)
- MANGAROSKA, K. et al. Understanding engagement through game learning analytics and design elements: Insights from a word game case study. In: *Proceedings of the 14th Learning Analytics and Knowledge Conference*. [S.l.: s.n.], 2024. p. 305–315. [49](#)
- MARIN-VEGA, H. et al. A brief review of game engines for educational and serious games development. *Language Learning and Literacy: Breakthroughs in Research and Practice*, IGI Global, p. 447–469, 2020. [78](#)

- MARKEY, K.; LEEDER, C. The effect of scoring and feedback mechanisms in an online educational game. *Proceedings of the American Society for Information Science and Technology*, Wiley Online Library, v. 48, n. 1, p. 1–10, 2011. 78
- MARTINEZ, K.; MENÉNDEZ-MENÉNDEZ, M. I.; BUSTILLO, A. A new measure for serious games evaluation: gaming educational balanced (geb) model. *Applied Sciences*, MDPI, v. 12, n. 22, p. 11757, 2022. 80
- MAYER, R. *Multimedia Learning*. Cambridge University Press, 2009. (Multimedia Learning). ISBN 9780521514125. Disponível em: <<https://books.google.com.br/books?id=5g0AM1CHysgC>>. 46
- MEDLER, B. Generations of game analytics, achievements and high scores. *Eludamos. Journal for Computer Game Culture*, v. 3, n. 2, p. 177–194, 2009. 38
- MELO, D. et al. Uma estratégia de game learning analytics para avaliar level design em um jogo educacional. In: SBC. *Anais do XXXI Simpósio Brasileiro de Informática na Educação*. [S.l.], 2020. p. 622–631. 25, 41, 57
- MELO, D.; PIRES, F.; FREITAS, R. de. As aventuras de biguió: um jogo educacional sobre coberturas e caminhos de rainhas, torres e cavalos. In: *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*. [S.l.: s.n.], 2019. v. 8, n. 1, p. 1310. 78
- MELO, D. et al. Robô euroi: Game de estratégia matemática para exercitar o pensamento computacional. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2018. v. 29, n. 1, p. 685. 108, 110
- MISLEVY, R. J. et al. Psychometrics and game-based assessment. *Technology and testing: Improving educational and psychological measurement*, Routledge New York, NY, p. 23–48, 2016. 35
- MITCHELL, A.; SAVILL-SMITH, C. The use of computer and video games for learning. *A review of the literature*, LSDA, 2004. 34
- MORON, V. B. et al. Executive functions, motor development, and digital games applied to elementary school children: A systematic mapping study. *Education Sciences*, MDPI, v. 12, n. 3, p. 164, 2022. 95
- MOURÃO, L.; MENESES, P. Construção de medidas em td&e. *Medidas de Avaliação em Treinamento, Desenvolvimento e Educação*, Artmed Porto Alegre, p. 50–63, 2012. 45
- NADEEM, M.; OROSZLANYOVA, M.; FARAG, W. Effect of digital game-based learning on student engagement and motivation. *Computers*, MDPI, v. 12, n. 9, p. 177, 2023. 73
- NASCIMENTO, L. T. et al. Myname: um jogo educacional inspirado no problema da mochila. In: SBC. *Anais Estendidos do XII Congresso Brasileiro de Informática na Educação*. [S.l.], 2023. p. 145–148. 108, 118
- NOEMÍ, P.-M.; MÁXIMO, S. H. Educational games for learning. *Universal Journal of Educational Research*, ERIC, v. 2, n. 3, p. 230–238, 2014. 44

- OLIVEIRA, R. et al. Pro-avaliajs: Protocolo para planejamento e execução da avaliação da reação e aprendizagem de jogos sérios. In: *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*. Porto Alegre, RS, Brasil: SBC, 2022. p. 517–527. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/sbie/article/view/22436>>. 76
- OLIVEIRA, R. et al. Avaliações em jogos educacionais: instrumentos de avaliação da reação, aprendizagem e comparação de jogos. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2019. v. 30, n. 1, p. 972. 45
- OWEN, V. E. et al. Capturing learner trajectories in educational games through adage (assessment data aggregator for game environments): a click-stream data framework for assessment of learning in play. In: *American Educational Research Association Annual Meeting*. [S.l.: s.n.], 2014. p. 1–7. 57
- PATRIARCA, R. et al. Serious games for industrial safety: An approach for developing resilience early warning indicators. *Safety science*, Elsevier, v. 118, p. 316–331, 2019. 92
- PENFIELD, R. D.; GIACOBBI, P. R. Applying a score confidence interval to aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, v. 8, n. 4, p. 213–225, 2004. 102, 104
- PETRI, G.; WANGENHEIM, C. G. von. How games for computing education are evaluated? a systematic literature review. *Computers & education*, Elsevier, v. 107, p. 68–90, 2017. 36, 44, 46
- PETRI, G.; WANGENHEIM, C. Gresse von; BORGATTO, A. F. Meega+: Um modelo para a avaliação de jogos educacionais para o ensino de computação. *Revista Brasileira de Informática na Educação*, v. 27, n. 3, 2019. 23, 24, 36, 44, 45
- PHELPS, E. A. Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.*, Annual Reviews, v. 57, p. 27–53, 2006. 34
- PINHEIRO, A. G. P.; SEGUNDO, F. C. G. da S.; SOUSA, R. R. de. Gamelab: Um jogo para auxílio na motricidade fina. *Anais do Encontro de Computação do Oeste Potiguar ECOP/UFERSA (ISSN 2526-7574)*, n. 4, 2020. 76
- PINHEIRO, A. G. P. et al. Parkrun: Um jogo educacional voltado para o estímulo e melhora da capacidade motora e cognitiva de crianças com síndrome de down. *Conjecturas*, v. 22, n. 15, p. 571–585, 2022. 75
- PIRES, F. et al. A game proposal to develop computational thinking and environmental awareness. *Journal on Computational Thinking (JCThink)*, v. 3, n. 1, p. 111, 2019. 33
- PIRES, F. G. de S. et al. O livro do conhecimento: um serious game educacional para aprendizagem de ortografia da língua portuguesa. *Revista Brasileira de Informática na Educação*, v. 28, p. 436–460, 2020. 33, 46
- PLASS, J. L.; MAYER, R. E.; HOMER, B. D. *Handbook of game-based learning*. [S.l.]: Mit Press, 2020. 22, 34
- PORTO, B.; DIAS, D. M.; BATTESTIN, V. Tendências de learning analytics em moodle: uma revisão sistemática. *EaD em Foco*, v. 13, n. 1, p. e2070–e2070, 2023. 39

- REVELLE, G. Applying developmental theory and research to the creation of educational games. *New directions for child and adolescent development*, Wiley Online Library, v. 2013, n. 139, p. 31–40, 2013. [33](#)
- ROMERO, C.; VENTURA, S. Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (applications and reviews)*, Ieee, v. 40, n. 6, p. 601–618, 2010. [24](#)
- ROWE, E.; ASBELL-CLARKE, J.; BAKER, R. S. Serious games analytics to measure implicit science learning. *Serious games analytics: Methodologies for performance measurement, assessment, and improvement*, Springer, p. 343–360, 2015. [50](#)
- SAVI, R.; ULBRICHT, V. R. Jogos digitais educacionais: benefícios e desafios. *Renote*, v. 6, n. 1, 2008. [33](#)
- SHI, Y.; QU, S. The effect of cognitive ability on academic achievement: The mediating role of self-discipline and the moderating role of planning. *Frontiers in psychology*, Frontiers, v. 13, p. 1014655, 2022. [80](#)
- SIEMENS, G. Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, Sage Publications Sage CA: Los Angeles, CA, v. 57, n. 10, p. 1380–1400, 2013. [24](#)
- SIEMENS, G.; LONG, P. Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, ERIC, v. 46, n. 5, p. 30, 2011. [39](#)
- SILVA, D. et al. Glboard: um sistema para auxiliar na captura e análise de dados em jogos educacionais. In: SBC. *Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames)*. [S.l.], 2022. p. 959–968. [28](#), [51](#), [53](#), [60](#)
- SMIDERLE, R. et al. The impact of gamification on students' learning, engagement and behavior based on their personality traits. *Smart Learning Environments*, Springer, v. 7, n. 1, p. 3, 2020. [36](#)
- SOUZA, D. et al. Lightbot logicamente: um game lúdico amparado pelo pensamento computacional e a matemática. In: *Anais do Workshop de Informática na Escola*. [S.l.: s.n.], 2018. v. 24, n. 1, p. 61–69. [74](#)
- SOUZA, H. C. de; JUNIOR, M. M. C.; FELINTO, A. S. Inclusão das avaliações formativa e somativa no processo de desenvolvimento de um jogo educacional: Um estudo de caso no projeto do jogo castaways. In: SOCIEDADE BRASILEIRA DE COMPUTAÇÃO (SBC). *Anais Estendidos do XX Simpósio Brasileiro de Jogos e Entretenimento Digital*. [S.l.], 2021. p. 445–454. [25](#), [37](#), [48](#)
- SUTAPA, P. et al. Improving motor skills in early childhood through goal-oriented play activity. *Children*, MDPI, v. 8, n. 11, p. 994, 2021. [35](#)
- TAHIR, R.; WANG, A. I. State of the art in game based learning: Dimensions for evaluating educational games. In: ACADEMIC CONFERENCES INTERNATIONAL LIMITED. *European Conference on Games Based Learning*. [S.l.], 2017. p. 641–650. [23](#), [24](#), [36](#), [46](#)
- TLILI, A.; CHANG, M. *Data analytics approaches in educational games and gamification systems: Summary, challenges, and future insights*. [S.l.]: Springer, 2019. [41](#)

- TROYER, O. D. Towards effective serious games. In: IEEE. *2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. [S.l.], 2017. p. 284–289. [36](#)
- VIDENOVİK, M. et al. Game-based learning in computer science education: a scoping literature review. *International Journal of STEM Education*, Springer, v. 10, n. 1, p. 54, 2023. [35](#)
- VUKIĆEVIĆ, S. et al. A demonstration project for the utility of kinect-based educational games to benefit motor skills of children with asd. *Perceptual and motor skills*, Sage Publications Sage CA: Los Angeles, CA, v. 126, n. 6, p. 1117–1144, 2019. [35](#), [75](#)
- WALLNER, G.; KRIGLSTEIN, S. Visualization-based analysis of gameplay data—a review of literature. *Entertainment Computing*, Elsevier, v. 4, n. 3, p. 143–155, 2013. [24](#)
- WALSH, C.; BOKHOVE, C. Targeting data collection in games based assessment. *Computers and Education Open*, Elsevier, v. 2, p. 100054, 2021. [24](#), [35](#), [36](#), [37](#), [47](#), [93](#)
- WANG, F.; HANNAFIN, M. J. Design-based research and technology-enhanced learning environments. *Educational technology research and development*, Springer, v. 53, n. 4, p. 5–23, 2005. [28](#)
- WANG, L.-H. et al. Effects of digital game-based stem education on students' learning achievement: a meta-analysis. *International Journal of STEM Education*, Springer, v. 9, n. 1, p. 26, 2022. [40](#)
- WILSON, D. W. et al. Serious games: an evaluation framework and case study. In: IEEE. *2016 49th Hawaii International Conference on System Sciences (HICSS)*. [S.l.], 2016. p. 638–647. [37](#)
- WISNIEWSKI, B.; ZIERER, K.; HATTIE, J. The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in psychology*, Frontiers, v. 10, p. 487662, 2020. [36](#)
- WU, X.; LIU, S.; SHUKLA, A. Serious games as an engaging medium on building energy consumption: A review of trends, categories and approaches. *Sustainability*, MDPI, v. 12, n. 20, p. 8508, 2020. [78](#)
- ZAGAL, J. P.; ALTIZER, R. Examining' rpg elements': Systems of character progression. In: *FDG*. [S.l.: s.n.], 2014. [78](#)
- ZAMANZADEH, V. et al. Design and implementation content validity study. *Nursing and Midwifery Studies*, 2015. [100](#), [101](#), [102](#)
- ZHANG, C. et al. Acceptance of artificial intelligence among pre-service teachers: a multigroup analysis. *International Journal of Educational Technology in Higher Education*, Springer, v. 20, n. 1, p. 49, 2023. [40](#)

A

APÊNDICES

A

ANEXOS

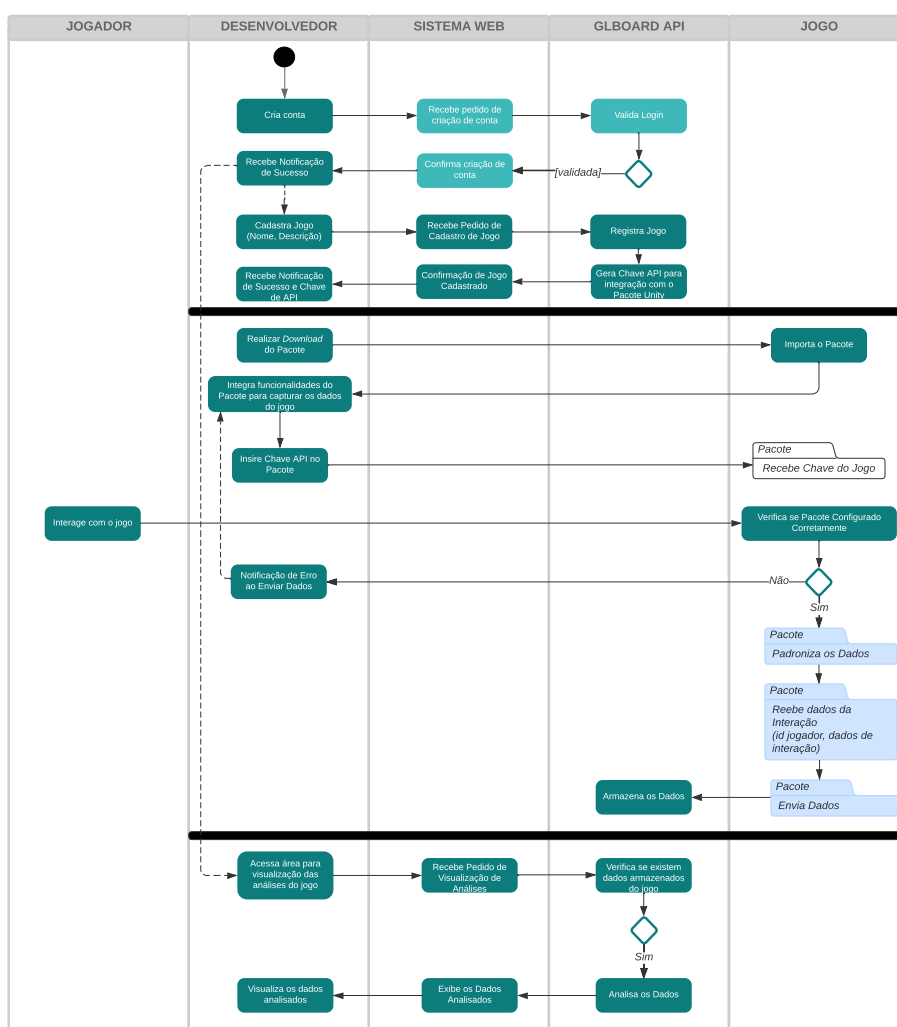


Figura 20 – Fluxo de geral do desenvolvedor no sistema GLBoard.