

Universidade Federal do Maranhão
Centro de Ciências Exatas e Tecnológicas
Programa de Pós-Graduação em Engenharia Elétrica

THAMYLA MARIA DE SOUSA LIMA

**ESTUDO DOS MÉTODOS AVALIATIVOS DE JOGOS DIGITAIS
SÉRIOS: ANÁLISE E APLICAÇÃO**

São Luís
2021

THAMYLA MARIA DE SOUSA LIMA

ESTUDO DOS MÉTODOS AVALIATIVOS DE JOGOS DIGITAIS
SÉRIOS: ANÁLISE E APLICAÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Maranhão, **como parte dos requisitos necessários** para obtenção do grau de Mestre em Engenharia Elétrica.

Orientador: Prof^o Dr. Allan Kardec Duailibe Barros Filho

Coorientador: Prof^o Dr. Alex Oliveira Barradas Filho

São Luís

2021

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Lima, Thamyla Maria de Sousa.

Estudo dos métodos avaliativos de jogos digitais sérios: análise e aplicação / Thamyla Maria de Sousa Lima. - 2021.
107 f.

Coorientador(a): Alex Oliveira Barradas Filho.

Orientador(a): Allan Kardec Duailibe Barros Filho.

Dissertação (Mestrado) - Programa de Pós-graduação em Engenharia Elétrica/ccet, Universidade Federal do Maranhão, São Luís, 2021.

1. Avaliação. 2. Jogos sérios. 3. Métodos avaliativos. I. Barradas Filho, Alex Oliveira. II. Barros Filho, Allan Kardec Duailibe. III. Título.

Thamyla Maria de Sousa Lima

Estudo dos métodos avaliativos de jogos digitais sérios: análise e aplicação

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Maranhão, **como parte dos requisitos necessários** para obtenção do grau de Mestre em Engenharia Elétrica.

Trabalho aprovado. São Luís, 18 de Junho de 2021:

**Prof. Dr. Allan Kardec Duailibe
Barros Filho**
Universidade Federal do Maranhão
Orientador

Prof. Dr. Alex Oliveria Barradas Filho
Universidade Federal do Maranhão
Coorientador

**Prof. Dr. Francisco José da Silva e
Silva**
Universidade Federal do Maranhão

Prof. Dr. Rodrigo Pereira dos Santos
Universidade Federal do Estado do Rio de
Janeiro

São Luís
2021

Aos meus pais, Alfredo e Iracema, que sempre me apoiaram.

Agradecimentos

A toda minha família, meus pais e minhas irmãs.

Aos meus orientadores, Allan Kardec e Alex Oliveira, por todas as oportunidades e apoio dado.

A todos os meus professores, sempre dispostos a ensinar e proporcionar uma educação de qualidade.

Aos meus companheiros de laboratório e à todos os meus amigos e colegas da graduação e do mestrado.

A todos que contribuíram direta ou indiretamente na minha formação acadêmica.

*“O maior inimigo do conhecimento não é a ignorância, é a ilusão do conhecimento.” -
Stephen Hawking*

Resumo

Jogos sérios podem ser definidos como aplicações de computador que combinam aspectos sérios (ensino e aprendizado) com aspectos lúdicos. Para avaliar seu poder de impacto sobre os alunos são utilizados os métodos avaliativos. Existe uma variedade de estudos na literatura que abordam os métodos avaliativos de jogos sérios. Estes se diferem na forma de aplicação, de análise dos resultados e tamanho da amostra. Neste trabalho, foi realizado um estudo dos métodos avaliativos, identificando características, tendências, vantagens e desvantagens desses métodos. Para realizar esse estudo, foi conduzido um Mapeamento Sistemático da Literatura. Ainda, a partir de critérios pré-definidos foram selecionados dois métodos de avaliação a serem aplicados em um experimento. Para realização desse experimento, desenvolveu-se dois jogos sérios relacionados a conceitos introdutórios de computação. O acesso a esses jogos se deu através de uma plataforma online que possibilita cadastro de professores e alunos, bem como a criação de salas virtuais. O experimento foi conduzido com 10 voluntários da área de Ciência da Computação e os dados coletados pela plataforma foram analisados e discutidos. Dessa forma, o estudo realizado neste trabalho permitiu identificar os métodos avaliativos utilizados na literatura e suas particularidades. Além disso, a análise de dados realizada permitiu verificar os aspectos mais relevantes a serem avaliados nos jogos digitais sérios. Como trabalhos futuros, cita-se a inserção de técnicas de inteligência artificial na coleta e análise dos dados e a aplicação da plataforma em uma amostra maior de participantes.

Palavras-chaves: jogos sérios, métodos avaliativos, avaliação.

Abstract

Serious games can be defined as computer applications that combine serious aspects (teaching, learning) with playful aspects. To assess their impact on students, evaluation methods are used. There are a variety of articles in the literature that address the methods of evaluating serious games. These differ in the form of application, analysis of results and sample size. In this work, a study of the evaluation methods was carried out, identifying characteristics, trends, advantages and disadvantages of these methods. To carry out this study, a Systematic Literature Mapping was conducted. Also, based on pre-defined criteria, two evaluation methods were selected to be applied in an experiment. To carry out this experiment, two serious games related to basic introductory concepts were developed. Access to these games took place through an online platform that enables the registration of teachers and students, as well as the creation of virtual classrooms. The experiment was conducted with 10 volunteers from the Computer Science area and the data collected by the platform were analyzed and discussed. Thus, the study carried out in this work allowed to identify the evaluation methods used in the literature and their particularities. In addition, the data analysis performed allowed to verify the most relevant aspects to be evaluated in serious digital games. As future works, it can be cited the insertion of artificial intelligence techniques in the collection and analysis of data and the application of the platform in a larger sample of participants.

Keywords: serious games, evaluation methods, evaluation.

Lista de ilustrações

Figura 1 – Fluxograma das etapas do trabalho	22
Figura 2 – <i>Flow chart</i> do mapeamento sistemático realizado	26
Figura 3 – Fluxograma das etapas para escolha dos métodos avaliativos a serem aplicados.	29
Figura 4 – Distribuição dos trabalhos por ano.	30
Figura 5 – Quantidades de usuários utilizadas nos trabalhos.	31
Figura 6 – Pontuação de acordo com critérios de qualidade.	32
Figura 7 – Dimensões/Fatores avaliados pelos métodos e suas frequências.	33
Figura 8 – Quantidade de trabalhos de acordo com o método de avaliação.	35
Figura 9 – Modelos de avaliação baseados em questionários.	36
Figura 10 – Itens extraídos nos trabalhos de logs e suas frequências.	39
Figura 11 – Tipos de análises de dados identificadas nos trabalhos e suas frequências.	41
Figura 12 – Diagrama de classes da plataforma Euklides.	46
Figura 13 – Diagrama de sequência - Login Professor.	47
Figura 14 – Diagrama de sequência - Consulta e cadastro de Salas.	47
Figura 15 – Login/Cadastro.	48
Figura 16 – Página inicial do professor.	48
Figura 17 – Gerenciamento da sala.	49
Figura 18 – Diagrama de sequência - Cadastro de Aluno na plataforma.	49
Figura 19 – Página inicial do aluno.	50
Figura 20 – Quiz Game - Tela de perguntas.	51
Figura 21 – Axio - Jogo iniciado.	52
Figura 22 – Axio - Objetivo do nível.	52
Figura 23 – Equação Teste t para amostras independentes.	53
Figura 24 – Sala criada para o experimento.	54
Figura 25 – Resultados do questionário MEEGA+ no Quiz Game.	55
Figura 26 – Resultados do questionário MEEGA+ no Axio.	57
Figura 27 – Tempo jogado no Quiz Game.	58
Figura 28 – Tempo por partida no Quiz Game.	58
Figura 29 – Taxas de acertos por partida no Quiz Game.	59
Figura 30 – Taxas de erros por partida no Quiz Game.	59
Figura 31 – Taxa de dicas utilizadas por partida no Quiz Game.	60
Figura 32 – Número de tentativas no Quiz Game.	60
Figura 33 – Relação entre número de acertos e o tempo total de jogo.	61
Figura 34 – Relação entre número de acertos e o tempo médio por partida.	62
Figura 35 – Relação entre número de acertos e dicas utilizadas.	62

Figura 36 – Relação entre número de acertos e número de tentativas.	63
Figura 37 – Tempo jogado no Axio.	63
Figura 38 – Tempo médio por partida no Axio.	64
Figura 39 – Número de níveis concluídos no Axio.	64
Figura 40 – Taxa de erros por partida no Axio.	65
Figura 41 – Número de tentativas no Axio.	65
Figura 42 – Relação entre número de níveis alcançados e o tempo total de jogo. . .	66
Figura 43 – Relação entre número de níveis alcançados e o tempo médio por partida.	66
Figura 44 – Relação entre número de níveis alcançados e o número de tentativas. .	67
Figura 45 – Relação entre número de erros e o tempo total.	67
Figura 46 – Relação entre número de erros e o número de tentativas.	68
Figura 47 – Login/Cadastro.	77
Figura 48 – Página inicial do professor.	78
Figura 49 – Criação de uma nova sala.	78
Figura 50 – Gerenciamento da sala - Aba Jogos.	79
Figura 51 – Gerenciamento da sala - Aba Alunos.	79
Figura 52 – Gerenciamento da sala - Aba Alunos - Desempenho detalhado do Aluno.	79
Figura 53 – Gerenciamento da sala - Aba Gerenciar.	80
Figura 54 – Página inicial do aluno.	80
Figura 55 – Questionário para avaliação do jogo.	81
Figura 56 – Quiz Game - Menu principal.	82
Figura 57 – Quiz Game - Tela de perguntas.	82
Figura 58 – Quiz Game - Tela de fim de jogo.	83
Figura 59 – Axio - Menu principal.	83
Figura 60 – Axio - Como Jogar.	84
Figura 61 – Axio - Jogo iniciado.	84
Figura 62 – Axio - Objetivo do nível.	85
Figura 63 – Tempo total x Dimensões MEEGA+	89
Figura 64 – Tempo por partida x Dimensões MEEGA+	90
Figura 65 – Número de acertos x Dimensões MEEGA+	91
Figura 66 – Número de dicas x Dimensões MEEGA+	92
Figura 67 – Tentativas x Dimensões MEEGA+	93
Figura 68 – Tempo por jogada x Dimensões MEEGA+	95
Figura 69 – Tempo total x Dimensões MEEGA+	96
Figura 70 – Níveis alcançados x Dimensões MEEGA+	97
Figura 71 – Número de erros x Dimensões MEEGA+	98
Figura 72 – Número de tentativas x Dimensões MEEGA+	99

Lista de tabelas

Tabela 1 – Questões de pesquisa do mapeamento	23
Tabela 2 – Trabalhos da análise exploratória	24
Tabela 3 – Palavras-chave e sinônimos	24
Tabela 4 – Critérios de inclusão e exclusão	25
Tabela 5 – Critérios de qualidade	27
Tabela 6 – Itens extraídos e respectivas relações com os fatores dos jogos.	40
Tabela 7 – Requisitos funcionais - Plataforma Euklides	45
Tabela 8 – Requisitos não funcionais - Plataforma Euklides	46
Tabela 9 – Análise do tempo total em relação às dimensões do questionário	69
Tabela 10 – Análise do tempo por partida em relação às dimensões do questionário	69
Tabela 11 – Análise de número de acertos em relação às dimensões do questionário	69
Tabela 12 – Análise de número de dicas em relação às dimensões do questionário	70
Tabela 13 – Análise de número de tentativas em relação às dimensões do questionário	70
Tabela 14 – Análise do número de níveis alcançados em relação às dimensões do questionário	71
Tabela 15 – Análise do tempo total em relação às dimensões do questionário	71
Tabela 16 – Análise do tempo por jogada em relação às dimensões do questionário	72
Tabela 17 – Análise da taxa de erros em relação às dimensões do questionário	72
Tabela 18 – Análise do número de tentativas em relação às dimensões do questionário	72
Tabela 19 – Questionário elaborado por Petri, von Wangenheim e Borgatto (2017).	86
Tabela 20 – Continuação - Questionário elaborado por Petri, von Wangenheim e Borgatto (2017).	87

Sumário

1	INTRODUÇÃO	15
1.1	Problemática	16
1.2	Plataforma de jogos sérios	17
1.3	Revisões sistemáticas na área	18
1.4	Organização da dissertação	20
2	OBJETIVOS	21
2.1	Objetivo geral	21
2.2	Objetivos específicos	21
3	METODOLOGIA	22
3.1	Mapeamento Sistemático	22
3.1.1	Questões de pesquisa	23
3.1.2	Estratégia de pesquisa e classificação	23
3.1.3	Avaliação	27
3.1.4	Limitações do Mapeamento Sistemático	27
3.2	Métodos de avaliação aplicados	28
4	ESTUDO DOS MÉTODOS AVALIATIVOS	30
4.1	Resultados do mapeamento sistemático	30
4.2	QP1. Quais os fatores/dimensões que são avaliados nos jogos sérios?	32
4.3	QP2. Quais os métodos utilizados para avaliar jogos digitais sérios?	34
4.3.1	Questionários de experiência do usuário	35
4.3.2	Testes de desempenho	37
4.3.3	Testes A/B	38
4.3.4	Logs do sistema	38
4.3.5	Entrevistas	40
4.3.6	Observações dos aplicadores	40
4.4	QP3. Qual tipo de análise de dados é feita?	40
4.5	Discussão dos resultados	42
5	PLATAFORMA EUKLIDES	45
6	EXPERIMENTOS	53
6.1	Avaliação pelo questionário MEEGA+	55
6.1.1	Quiz Game	55
6.1.2	Axio	56

6.2	Avaliação por logs do sistema	57
6.2.1	Quiz Game	57
6.2.1.1	Relacionamentos entre os dados extraídos	60
6.2.2	Axio	63
6.2.2.1	Relacionamentos entre os dados extraídos	65
6.3	Comparativo entre o questionário e os logs do sistema	68
6.3.1	Quiz Game	68
6.3.2	Axio	70
6.4	Discussão dos resultados	73
7	CONCLUSÃO	74
7.1	Contribuições	75
7.2	Limitações	75
7.3	Trabalhos futuros	76
	APÊNDICE A – MANUAL DO USUÁRIO - PLATAFORMA EU- KLIDES	77
A.1	Quiz Game	81
A.2	Axio	83
	APÊNDICE B – QUESTIONÁRIO DO EXPERIMENTO	86
	APÊNDICE C – TABELAS COMPARATIVAS QUIZ GAME X ME- EGA+	88
	APÊNDICE D – TABELAS COMPARATIVAS AXIO X MEEGA+	94
	REFERÊNCIAS	100

1 Introdução

Nas últimas décadas, as Tecnologias de Informação e Comunicação (TIC) influenciaram o processo de ensino - aprendizagem pela inserção de ferramentas que aprimoram as interações dos alunos e possibilitam ambientes desafiadores com situações específicas (COLLINS; HALVERSON, 2010). Os jogos sérios são um exemplo do emprego de tais ferramentas, atuando no treinamento de habilidades específicas, aprimoramento do aprendizado e resolução de problemas simulados (COLLINS; HALVERSON, 2010; KANG; MIN; QU, 2016; PETRI; VON WANGENHEIM, 2017). No trabalho de (ALVAREZ; DJAOUTI, 2011), jogo sério é definido como uma aplicação de computador que combina aspectos sérios como ensino, aprendizado ou outro tipo de informação com aspectos lúdicos provenientes dos jogos.

O uso dos jogos sérios, como ferramenta auxiliadora no processo de ensino - aprendizagem, pode proporcionar algumas vantagens, como aprendizado com base nos próprios erros em um ambiente controlado, experiências autênticas do mundo real, mas sem riscos à integridade física dos alunos, ambientes competitivos e/ou colaborativos, entre outros (NEWBERY; LEAN; MOIZER, 2016; KANG; MIN; QU, 2016; ERHEL; JAMET, 2013). Ademais, os autores (PETRI; VON WANGENHEIM, 2017) citam também como benefícios do emprego dos jogos sérios, o aumento no interesse, na motivação, bem como na redução no tempo de ensino.

Tais benefícios ocasionaram um rápido aumento no interesse da utilização de jogos sérios para fins educacionais (FREIRE et al., 2016). Pesquisas que aplicam essas ferramentas apontam que os jogos sérios conseguem ao mesmo tempo engajar e melhorar o aprendizado dos alunos e, se mostram como uma ferramenta mais eficiente quando comparadas com outras formas de ensino tradicional (TAHIR; WANG, 2017; RITTERFELD; CODY; VORDERER, 2009).

Contudo, possuir certas características é importante para que o jogo sério obtenha um maior sucesso (MOUAHEB et al., 2012). Citam-se como exemplos: possuir objetivos definidos, ser divertido, transmitir informação, possuir múltiplos objetivos de aprendizagem (ensinar, treinar e educar) e poder ser aplicado em várias áreas e para todas as idades (MOUAHEB et al., 2012). Surge, portanto, a necessidade de avaliar esses fatores (também conhecidos como dimensões). Além disso, segundo (EAGLE, 2009), para identificar a eficácia desses jogos é necessário aplicar uma forma científica de avaliação.

Existem diferentes métodos avaliativos na literatura que permitem observar alguns dos fatores necessários para um jogo sério (LOPES; NUSSBAUM; TSAI, 2015). Os métodos de avaliação estão divididos em quantitativos e qualitativos (LOPES; NUSSBAUM; TSAI,

2015). Nos métodos quantitativos, os dados estão na forma numérica e podem ser usados para gerar estatísticas ao final. Para trabalhos que utilizam essa abordagem é também recomendado que as informações sejam relatadas de tal forma que um leitor possa: replicar o estudo, avaliar o rigor da pesquisa e avaliar a robustez dos resultados (LOPES; NUSSBAUM; TSAI, 2015). Nos métodos qualitativos, o foco está em descobrir como as pessoas se sentem sobre algo e seus dados estão na forma não numérica. Essa abordagem é mais complexa e diversificada, sendo que não se baseia apenas na aplicação de métodos, mas também em pressupostos teóricos e metodológicos dos fenômenos estudados (TWINING et al., 2017).

Vale ressaltar que apesar das metodologias das análises qualitativa e quantitativa serem diferentes, é possível que uma pesquisa adote métodos qualitativos e quantitativos quando estiver lidando com resultados que estiverem tanto na forma numérica como na forma não numérica. A forma como é feita essa avaliação depende da metodologia adotada no trabalho (TWINING et al., 2017).

1.1 Problemática

A partir das características dos métodos avaliativos, nota-se sua importância no processo de aplicação e avaliação dos jogos sérios. Contudo, surge uma questão referente a esse processo avaliativo: o que cada método avaliativo consegue avaliar em um jogo? É também importante identificar como os dados obtidos por meio da aplicação desses métodos devem ser analisados e interpretados. Dessa forma, é possível verificar se esses dados são subjetivos ou se abrem margem para parcialidade.

Assim, com intuito de identificar os métodos avaliativos, a forma como são aplicados e como é feita a análise de dados, foi conduzido um Mapeamento Sistemático da Literatura. Esse Mapeamento teve como foco a identificação dos métodos e modelos avaliativos mais usados atualmente. Além disso, foi possível identificar os métodos avaliativos e suas particularidades, bem como as análises de dados que são realizadas. Assim, para avaliar os resultados encontrados, foi desenvolvida uma plataforma web que permite a aplicação e avaliação de jogos sérios. Essa plataforma, nomeada de Euklides, permite o cadastro de alunos e professores, bem como a criação de salas virtuais, onde é possível adicionar jogos digitais sérios.

Foram desenvolvidos dois jogos sérios para serem inseridos na plataforma. Tendo em vista que o público alvo do experimento seria alunos iniciantes na área de Ciência da Computação, os jogos foram voltados para temas iniciais da área. Para avaliá-los foram introduzidos dois métodos avaliativos na plataforma. O primeiro avalia os jogos por meio de questionários aplicados aos participantes e o segundo pela coleta automática de dados durante o jogo, como tempo gasto, número de acertos, erros e tentativas.

Para aplicar os resultados identificados no Mapeamento Sistemático foi realizado um experimento com 10 voluntários da área de Ciência da Computação que utilizaram a plataforma com os dois jogos disponíveis e responderam o questionário avaliativo disponível. A partir da coleta de dados dos jogos, foi possível realizar a avaliação de ambos os jogos e comparar os resultados dos dois métodos aplicados.

1.2 Plataforma de jogos sérios

Foram identificadas na literatura outras plataformas voltadas para educação e que utilizam jogos sérios. Uma das de maior destaque é a Code.org. Ela tem o objetivo de expandir o acesso à Ciência da Computação, alcançando principalmente mulheres e minorias. É também uma organização sem fins lucrativos apoiada por várias instituições e que disponibiliza recursos para aprender e ensinar Ciência da Computação. Assim, pode ser acessada tanto por alunos como por professores de forma gratuita (CODE.ORG, 2020).

Em (MARTINS; REIS; MARQUES, 2016) foi utilizado um jogo da plataforma Code.org, o Labirinto Clássico, para realizar um experimento com 168 alunos do ensino fundamental. O objetivo do jogo é levar o personagem de um ponto a outro e, para isso, o participante deve resolver 20 desafios utilizando programação. No trabalho de (FIORI; ROCHA; MARQUES, 2019) também foi utilizada a plataforma. Esse trabalho teve como objetivo introduzir Lógica de Programação para 28 alunos de ensino médio e despertar o interesse para o curso de Ciência da Computação. Para tanto, foi utilizado o jogo Aventureiro de Minecraft. Em ambos os trabalhos, para avaliar o impacto dos jogos, foi utilizado o modelo de avaliação MEEGA (SAVI; VON WANGENHEIM; BORGATTO, 2011).

Outra plataforma de aprendizagem online encontrada é a *Kahoot*. Ela permite a criação de jogos baseados em questionários, sendo possível personalizar o formato, o número de questões, bem como adicionar vídeos e imagens (KAHOOT, 2020). Para envolver os alunos, são inseridas características de jogos como pontuação, interação e ranqueamento (MARTINS; GOUVEIA, 2018). No trabalho de (MARTINS; GOUVEIA, 2018), foi utilizada a plataforma para criar uma atividade com 25 perguntas relacionadas ao conteúdo da disciplina. Para avaliar a atividade foi utilizado um questionário com cinco perguntas referentes às impressões gerais do jogo. Em (GUILLÓ et al., 2019), o objetivo foi estudar os benefícios gerados pelo uso de jogos sérios no ensino de Engenharia da Computação, para isso também foi utilizada a plataforma *Kahoot*. Para avaliação do impacto do jogo, os participantes foram separados em grupos e foi utilizado um questionário de sete perguntas referente a experiência do usuário. Outros trabalhos que utilizaram a plataforma *Kahoot* podem ser encontrados em (SANDE; SANDE, 2018; ALVES; ROZA, 2018; SILVA et al., 2018).

No trabalho de (ECONOMOU et al., 2015) foi desenvolvida uma plataforma web para simular um ambiente de ensino. A partir das ferramentas disponíveis, os educadores podem criar simulações baseadas em jogos, observando os alunos em um ambiente controlado e utilizar do sistema de pontuação disponível para atribuir pontos aos alunos quando estes completarem determinados desafios. Já a avaliação da plataforma é feita por meio de entrevistas com alunos e professores. Essa mesma metodologia pode ser encontrada também nos trabalhos de (KLAASSEN et al., 2018; HUSSAAN; SEHABA, 2013). No primeiro, é criada uma plataforma para pacientes com doenças crônicas e avaliação não é feita dentro da plataforma, mas por meio de entrevistas, observações diretas e questionários. No segundo, é desenvolvida a plataforma GOALS que permite que desenvolvedores gerem cenários de aprendizagem e jogos sérios. Para a avaliação são realizados testes antes e depois com os participantes.

Outra plataforma de jogos sérios pode ser encontrada no trabalho de (MARTINS; CARVALHO; SOARES, 2015). Foi desenvolvida uma página web que centraliza o gerenciamento de um ou mais jogos destinados a fisioterapia. Nela, são mostrados os resultados do paciente para cada jogo, número de acessos e inscrições na plataforma, número de jogadores por mês, por gênero, entre outros.

1.3 Revisões sistemáticas na área

No planejamento e no desenvolvimento deste trabalho, foram procuradas revisões e mapeamentos sistemáticos relacionados ao tema. Foram encontrados trabalhos relevantes, disponíveis nas bases, que desenvolveram revisão ou mapeamento na área de avaliação de jogos. Contudo, nenhum destes possuíam os mesmos objetivos deste trabalho. A seguir, os trabalhos encontrados, que conduzem uma revisão ou mapeamento, são caracterizados e diferenciados do mapeamento conduzido neste presente trabalho.

Em (PETRI; VON WANGENHEIM, 2017) foi realizada uma Revisão Sistemática da Literatura em 2015. O objetivo dessa revisão foi identificar o estado da arte relacionado a como os jogos para ensino de computação são avaliados. Foram revisados 3617 trabalhos, dos quais 112 foram classificados como relevantes para o objetivo da pesquisa. Os resultados principais do mapeamento incluíam os itens que são avaliados nos jogos, o número de pesquisas nessa área entre 1999 e 2015, como as avaliações são conduzidas (de forma experimental ou não) e quais os modelos e instrumentos (questionários e testes) utilizados na avaliação. Esse estudo se concentrou nos jogos relacionados ao ensino da computação e, além disso, não teve como foco a forma como os testes são aplicados e como é feita a análise dos dados dos resultados.

Em (FATTA; MAKSOM; ZAKARIA, 2018) foi desenvolvida uma Revisão Sistemática da Literatura de modelos de avaliação da usabilidade de jogos sérios. O objetivo deste

trabalho foi identificar os modelos que avaliam usabilidade, quais os métodos de teste são utilizados para avaliar e como são determinadas as heurísticas para desenvolver um modelo de avaliação de usabilidade. Foram selecionados 213 trabalhos e após sua leitura completa foram selecionados os 30 mais relevantes para a pesquisa. A partir da leitura completa desses trabalhos, alguns achados podem ser destacados: os métodos de avaliação mais utilizados na avaliação da usabilidade, metodologia para desenvolvimento de um modelo de avaliação da usabilidade, entre outros. Uma revisão sistemática semelhante a esta foi a desenvolvida em (CABALLERO-HERNÁNDEZ; PALOMO-DUARTE; DODERO, 2017) que buscou identificar os diferentes métodos de avaliação para avaliar as habilidades adquiridas pelos jogadores durante um jogo sério. Para isso, foram analisados 449 trabalhos. Enquanto na revisão de (FATTA; MAKSOM; ZAKARIA, 2018), o objetivo foi identificar métodos de avaliação da usabilidade, na de (CABALLERO-HERNÁNDEZ; PALOMO-DUARTE; DODERO, 2017), o intuito foi verificar como os trabalhos avaliam o ganho de conhecimento ao utilizar um jogo sério. Ainda nessa vertente, em (HOOKHAM; NESBITT, 2019) foi conduzida uma revisão que buscou compreender como o engajamento tem sido usado, definido e medido no contexto de jogos sérios. Nessa última foram identificados, inicialmente, 1390 trabalhos e destes 107 foram selecionados como aqueles que de fato avaliavam o engajamento.

Outro mapeamento desenvolvido em (SANTOS et al., 2018) focou em jogos para auxiliar no ensino de programação. Os objetivos do trabalho foram identificar os elementos de jogos para ensino de programação e entender como esses elementos são avaliados. Foram identificados 39 estudos primários com 27 elementos distribuídos em 43 jogos sérios. Um dos resultados do trabalho foi o detalhamento de como foi feita a avaliação dos jogos em cada um dos estudos primários selecionados.

Uma Revisão Sistemática da Literatura conduzida em (TAHIR; WANG, 2017) objetivou identificar as dimensões para avaliar jogos educacionais. Foram analisados 58 trabalhos e, a partir destes foram respondidas as perguntas da Revisão Sistemática da Literatura. Como resultado, foram identificadas as abordagens de avaliação de jogos sérios, bem como as dimensões que foram utilizadas nos trabalhos selecionados.

Em (CALDERÓN; RUIZ; O'CONNOR, 2013) foi conduzida uma Revisão Sistemática da Literatura com objetivo de identificar as características de jogos sérios para o ensino de processo de desenvolvimento de software. Foram revisados 190 trabalhos e, desses, seis foram selecionados como estudos primários. Esse trabalho, apesar de não ter como foco a avaliação de jogos, nos seus resultados é analisada a forma como os jogos sérios dos seis trabalhos selecionados são avaliados. Outras revisões ou mapeamentos semelhantes, que fazem uma Revisão Sistemática da Literatura sobre jogos educacionais e tratam da avaliação desses jogos mesmo não tendo essa avaliação como objetivo principal, são os desenvolvidos em (ALFARAH; SCHÜNEMANN; AKL, 2010; AKL et al., 2010;

RODRIGUEZ; TEESSON; NEWTON, 2013; WANG et al., 2016).

A partir da leitura completa desses trabalhos, foi possível perceber que em grande parte não é analisado a fundo como os métodos são aplicados e como é feita a análise dos dados coletados que é um dos objetivos deste trabalho. Ademais, algumas das revisões encontradas focam em apenas uma das dimensões do jogo sério (como usabilidade ou aprendizado). Neste trabalho, pretendeu-se verificar a avaliação dos jogos de uma forma mais geral.

Dessa forma, o mapeamento conduzido nesta pesquisa agrega à literatura no sentido de pontuar não só os métodos utilizados na avaliação de jogos como também quais as características dos jogos que são avaliadas em cada um desses métodos. Assim, a partir dessas características, foi possível identificar qual aspecto do jogo pode ser melhorado de acordo com sua avaliação. Este mapeamento também procurou estudar jogos sérios aplicadas nas mais diversas áreas, permitindo uma visão mais generalista dos modelos de avaliação.

1.4 Organização da dissertação

Esta dissertação está organizada da seguinte forma:

- Capítulo 2 - Objetivos: apresenta o objetivo geral e os objetivos específicos do trabalho;
- Capítulo 3 - Metodologia: apresenta as etapas seguidas para identificar os métodos avaliativos existentes;
- Capítulo 4 - Estudo do métodos avaliativos: apresenta os resultados obtidos com o mapeamento sistemático;
- Capítulo 5 - Plataforma Euklides: apresenta a plataforma Euklides e os jogos desenvolvidos que foram incorporados;
- Capítulo 6 - Experimentos: apresenta o experimento realizado com participantes;
- Capítulo 7 - Conclusão: apresenta as conclusões desta dissertação, incluindo contribuições, limitações e trabalhos futuros.

2 Objetivos

2.1 Objetivo geral

O presente trabalho tem como objetivo geral identificar as características dos métodos de avaliação de jogos sérios existentes.

2.2 Objetivos específicos

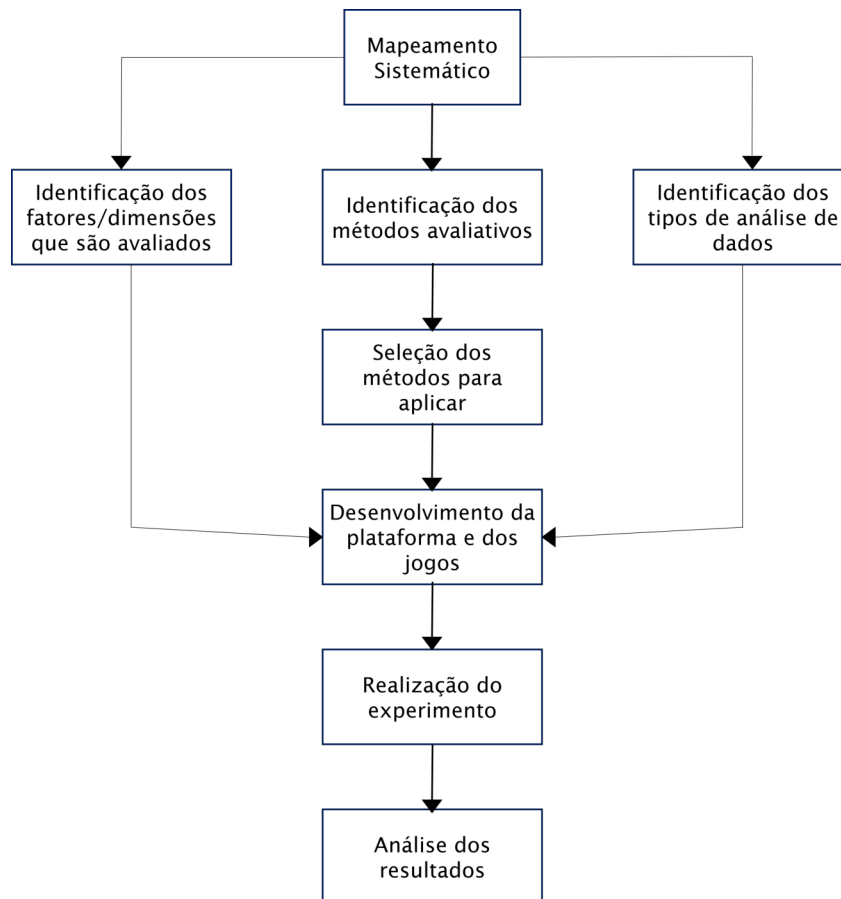
Os objetivos específicos são:

- Identificar o estado da arte em relação aos métodos de avaliação de jogos sérios.
- Aplicar os métodos e as formas de análise de dados para verificar viabilidade desses métodos.
- Discutir as características dos métodos avaliativos, de acordo com os resultados do experimento realizado.

3 Metodologia

Neste capítulo serão descritas as etapas para a realização do trabalho. A partir da condução do mapeamento foram realizadas as etapas de: identificação das dimensões avaliadas em jogos sérios, os métodos avaliativos utilizados e os tipos de análise de dados que são feitas. Em seguida, foram selecionados métodos de avaliação a serem aplicados, de acordo com suas características identificadas. Posteriormente, foram realizadas as etapas de desenvolvimento de jogos e plataforma, realização do experimento e análise dos resultados. A Figura 1 representa um fluxograma com as etapas da metodologia do trabalho.

Figura 1 – Fluxograma das etapas do trabalho



3.1 Mapeamento Sistemático

O propósito do mapeamento sistemático foi caracterizar e determinar o estado da arte em relação aos métodos de avaliação de jogos digitais sérios com intuito de

identificar tendências, possíveis falhas e aspectos que podem ser melhorados nos trabalhos encontrados.

Nesta seção serão descritas as etapas e os processos realizados na condução deste mapeamento.

3.1.1 Questões de pesquisa

As questões de pesquisa foram feitas com objetivo de caracterizar e classificar os métodos de avaliação existentes. Além disso, procurou-se entender o processo de aplicação da avaliação e análise dos dados. Era também um propósito da pesquisa, pontuar quais as dimensões e elementos do jogo que são avaliados.

Foram desenvolvidas três questões de pesquisa, conforme Tabela 1.

Tabela 1 – Questões de pesquisa do mapeamento

	Questões de pesquisa
QP1	Quais os fatores/dimensões que são avaliados nos jogos sérios?
QP2	Quais os métodos utilizados para avaliar jogos digitais sérios?
QP3	Qual tipo de análise de dados é feita?

3.1.2 Estratégia de pesquisa e classificação

A primeira etapa do trabalho consistiu na realização de uma análise exploratória com intuito de obter uma visão geral da área e encontrar as palavras-chave para criar a *string* de busca. Esta análise consistiu na busca informal por trabalhos relacionados ao tema em questão. Estes estudos, mostrados na Tabela 2, realizavam avaliação de um jogo sério ou propunham um modelo de avaliação. Eles foram lidos na íntegra e suas palavras-chave foram extraídas. Após o processo descrito, foram identificados três termos principais: *evaluation*, *serious game* e *education*. A partir destas, foram também identificados seus respectivos sinônimos, conforme Tabela 3.

Na execução das buscas, foram criadas *strings* de busca com as palavras-chave selecionadas. Para os termos principais foi usado o conectivo AND e para os sinônimos, o conectivo OR. A *string* de busca criada pode ser observada abaixo.

((*evaluation OR evaluate OR evaluated OR assess OR assessment OR scale OR validation*) AND ("*serious game*"OR "*educational games*"OR "*e-learning games*") AND (*education OR teaching OR instruction OR learning*))

Tabela 2 – Trabalhos da análise exploratória

Trabalho	Palavras-chave
<i>How games for computing education are evaluated? A systematic literature review</i> (PETRI; VON WANGENHEIM, 2017)	<i>Evaluation; Educational game; Computing education; Systematic literature review; Computer science;</i>
<i>EGameFlow: A scale to measure learners' enjoyment of e-learning games</i> (FU; SU; YU, 2009)	<i>Media in education; Interactive learning environments; Teaching/learning strategy; Application in applications in subject areas; Distance education and tele-learning;</i>
<i>RDU Model dedicated to evaluate needed counsels for Serious Game projects</i> (ALVAREZ et al., 2017)	<i>Counsels; Serious Game; Model; Ecosystem; Typology; Learning;</i>
<i>A serious game for developing computational thinking and learning introductory computer programming</i> (KAZIMOGU et al., 2012)	<i>Serious Game; Game based learning; Computational thinking; Introductory programming; Learning programming with games;</i>
<i>JDoc: A Serious Game for Medical Learning</i> (SLINEY; MURPHY, 2008)	<i>JDoc; Medical Learning; Serious Game; Simulation;</i>
<i>Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project</i> (WRZESIEN; RAYA, 2010)	<i>Interactive learning environments; Virtual reality; Pedagogical issues; Elementary education;</i>

Tabela 3 – Palavras-chave e sinônimos

Termos principais	Sinônimos
<i>evaluation</i>	<i>valuate, evaluated, assess, assessment, scale, validation</i>
<i>serious game</i>	<i>educational games, e-learning games</i>
<i>education</i>	<i>education, teaching, instruction, learning</i>

A estratégia de pesquisa utilizada consistiu na utilização de quatro bases de dados: ACM *Digital Library*, IEEE *Xplore*, *Web of Science* e *Engineering Village*. O idioma escolhido para aceitação dos trabalhos foi inglês por ser a linguagem mais aceita no mundo, tornando o trabalho mais abrangente. Assim, trabalhos que não estavam na língua selecionada eram excluídos, como pode ser observado na Tabela 4. Ademais, o mapeamento buscava encontrar trabalhos completos publicados em jornais, revistas ou conferências, portanto foram excluídos prefácios, teses, dissertações, pôsteres e anais

de evento. Outros critérios de exclusão utilizados foram de trabalhos que não fossem relacionados ao desenvolvimento ou avaliação de jogos sérios ou que não fizessem nenhuma avaliação do jogo desenvolvido ou ainda, cuja avaliação não pudesse ser reproduzida. Esse último aspecto correspondia aos trabalhos que aplicavam uma avaliação mas não disponibilizavam os dados dessa avaliação (por exemplo as questões de um questionário aplicado). Esses critérios foram utilizadas na etapa de leitura de título e resumo de cada estudo.

Tabela 4 – Critérios de inclusão e exclusão

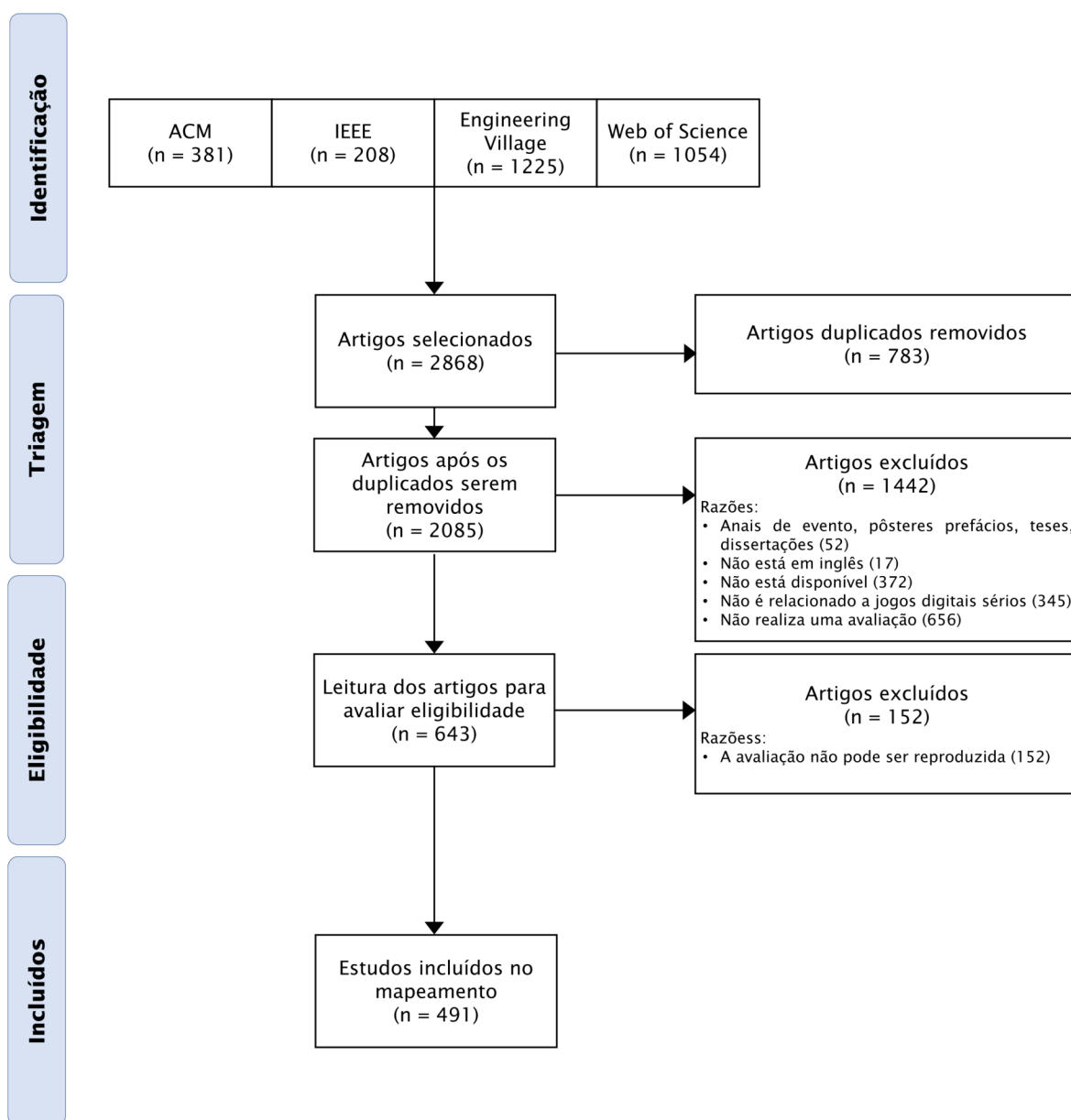
Critérios de inclusão	Critérios de exclusão
<ol style="list-style-type: none"> 1. Estudos com foco na avaliação (métodos e técnicas) dos jogos sérios desenvolvidos. 2. Estudos que apresentarem algum tipo de avaliação de jogos (métodos e técnicas), mesmo que o foco do trabalho seja o jogo criado e não a avaliação. 3. Estudos que apresentarem apenas o modelo para criação de jogos sérios desde que façam algum tipo de avaliação. 	<ol style="list-style-type: none"> 1. Anais de eventos. 2. Prefácios, teses ou dissertações. 3. Trabalhos que não forem em inglês 4. Trabalhos publicados como pôsteres. 5. Trabalhos que não estiverem disponíveis gratuitamente nas bases. 6. Trabalhos que não realizarem nenhuma avaliação do jogo desenvolvido. 7. Trabalhos que não forem relacionados a jogos digitais sérios. 8. Trabalhos cuja avaliação não possa ser reproduzida.

Ao executar a *string* de busca foram encontrados 2868 trabalhos, destes, 381 eram da ACM, 208 da IEEE, 1225 da Engineering Village e 1054 da Web of Science. Após a exclusão dos 783 trabalhos duplicados, resultou-se em um total de 2085 trabalhos. Para a manipulação dos trabalhos foi utilizada a ferramenta online Parsifal (PARSIFAL, 2018) que possibilita o desenvolvimento de uma Revisão Sistemática da Literatura ou Mapeamento Sistemático, organizando os trabalhos e separando-os entre aceitos e rejeitados conforme os critérios de inclusão e exclusão pré-definidos.

A classificação dos trabalhos se deu inicialmente na leitura do título e resumo com objetivo de excluir aqueles não relacionados à presente pesquisa, caso houvesse dúvida na classificação do trabalho, era lida a metodologia do mesmo. Utilizando os critérios de

inclusão e exclusão foram obtidos 1299 trabalhos aceitos e 808 rejeitados. Em seguida, os trabalhos foram lidos com intuito de identificar quais não apresentavam avaliação do jogo desenvolvido ou cuja avaliação não poderia ser reproduzida, ou seja, avaliações que não descreviam os métodos avaliativos utilizados de forma clara. Ao final dessa etapa foram obtidos 491 trabalhos aceitos. A Figura 2 apresenta um *PRISMA flow chart* do mapeamento sistemático realizado.

Figura 2 – *Flow chart* do mapeamento sistemático realizado



3.1.3 Avaliação

Critérios de qualidade são importantes porque existe a necessidade de pontuar a qualidade dos trabalhos aceitos para que as pontuações sirvam como uma base para os resultados finais (OKOLI; SCHABRAM, 2010). Contudo, a avaliação da qualidade é mais essencial em revisões sistemáticas já que permite verificar o rigor e a relevância dos estudos primários (PETERSEN; VAKKALANKA; KUZNIARZ, 2015). Apesar de não ser necessário realizar tal análise em mapeamentos sistemáticos, foram utilizados critérios de qualidade na presente pesquisa com objetivo de ranquear os estudos de acordo com a metodologia de pesquisa adotada. Na avaliação de qualidade, para cada um dos estudos é preenchido um formulário. No caso do mapeamento deste trabalho foram criadas duas perguntas, de acordo com a Tabela 5. Para as duas perguntas propostas, as respostas poderiam assumir três valores: sim (um ponto), parcialmente (0,5 ponto) e não (0).

Tabela 5 – Critérios de qualidade

	Questões de pesquisa
CQ1	Métodos da avaliação e resultados estão explícitos?
CQ2	As dimensões avaliadas são especificadas?

O CQ1 é importante tendo em vista que em uma avaliação cujos métodos e resultados estejam explícitos é possível reproduzi-lá de forma mais completa, bem como utilizar os mesmos métodos avaliativos com intuito de comparar resultados. Já no CQ2, tendo as dimensões especificadas, é possível saber por exemplo, quais perguntas são utilizadas para avaliar determinada dimensão em caso de uma avaliação por questionário.

3.1.4 Limitações do Mapeamento Sistemático

Foram consideradas as limitações de pesquisa abordadas em (AMPATZOGLOU et al., 2019) que coletou dados de mais 100 estudos e identificou as limitações mais comuns em Revisões Sistemáticas da Literatura. Algumas destas são: viés da inclusão e exclusão dos estudos, viés da extração de dados, repetição, construção da *string* de busca, seleção das bibliotecas digitais, viés do pesquisador e viés de publicação.

Limitações relacionadas à inclusão e exclusão dos trabalhos se referem a problemas que podem ter ocorrido na etapa de seleção de trabalhos. Para mitigar tais problemas, o protocolo do mapeamento foi criado de forma metódica e os critérios de inclusão e exclusão foram claramente definidos no protocolo. Buscou-se ainda, não criar critérios de inclusão e exclusão conflituosos e/ou muito genéricos. Vale também destacar que um pesquisador foi responsável pela seleção dos trabalhos. Isso pode ser uma ameaça à validade do mapeamento visto que erros humanos ou opiniões subjetivas são possíveis. Já

problemas na construção da *string* de busca podem provocar um número muito elevado de trabalhos retornados nas buscas, incluindo muitos trabalhos irrelevantes para a pesquisa. Em contrapartida, a construção da *string* de busca deve considerar os sinônimos das palavras-chave com o objetivo de não excluir estudos relevantes (AMPATZOGLOU et al., 2019). Nesse sentido, foram testadas quatro *strings* de buscas diferentes e foi escolhida a que retornava um número viável de trabalhos, mas ao mesmo tempo continha sinônimos das três palavras-chave da pesquisa (*evaluation, serious game education*).

Com relação à extração de dados, esta pode ser causada por perguntas abertas e/ou definidas de forma genérica no protocolo (AMPATZOGLOU et al., 2019). Para esse problema, foram criados formulários de extração que continham perguntas diretas relacionados à população no experimento, a qual método utilizado, quais dimensões avaliadas, entre outras. Além disso, os formulários foram preenchidos e disponibilizados online através da ferramenta utilizada (Parsifal).

Limitações referentes à seleção de bibliotecas digitais podem surgir ao utilizar bases muito específicas, muito amplas ou que não sejam confiáveis (AMPATZOGLOU et al., 2019). Para mitigar esse problema, nesse trabalho foram selecionadas algumas das bases mais utilizadas: ACM, IEEE, Web of Science e Compendex. Contudo, é possível que outras publicações relevantes, mas de repositórios diferentes dos que foram considerados, estejam ausentes nesse mapeamento.

No viés de publicação, a limitação pode ocorrer quando a maioria dos trabalhos são de uma única base, o que pode provocar um direcionamento para o pensamento de uma única comunidade (AMPATZOGLOU et al., 2019). Contudo, observou-se que apesar da base Compendex comportar 202 dos trabalhos aceitos, as demais bases também tiveram números consideráveis de trabalhos aceitos. Ainda, para evitar esse problema, não foi atribuído limite de tempo para as buscas, sendo encontrados trabalhos entre 1980 e 2019.

Por fim, com relação ao problema de replicação do estudo secundário, o protocolo foi desenvolvido de forma metódica para que pudesse ser replicado. Ademais, as etapas de todo o processo foram descritas e a seleção dos estudos de acordo com os critérios de inclusão e exclusão, bem como os formulários do processo de extração estão disponíveis na plataforma Parsifal. Os formulários de extração dos 491 estudos aceitos podem ser acessados em: <<https://drive.google.com/drive/folders/1JqUeWVowdFllIpyi0h3CtUqtpWaR6N8d?usp=sharing>>.

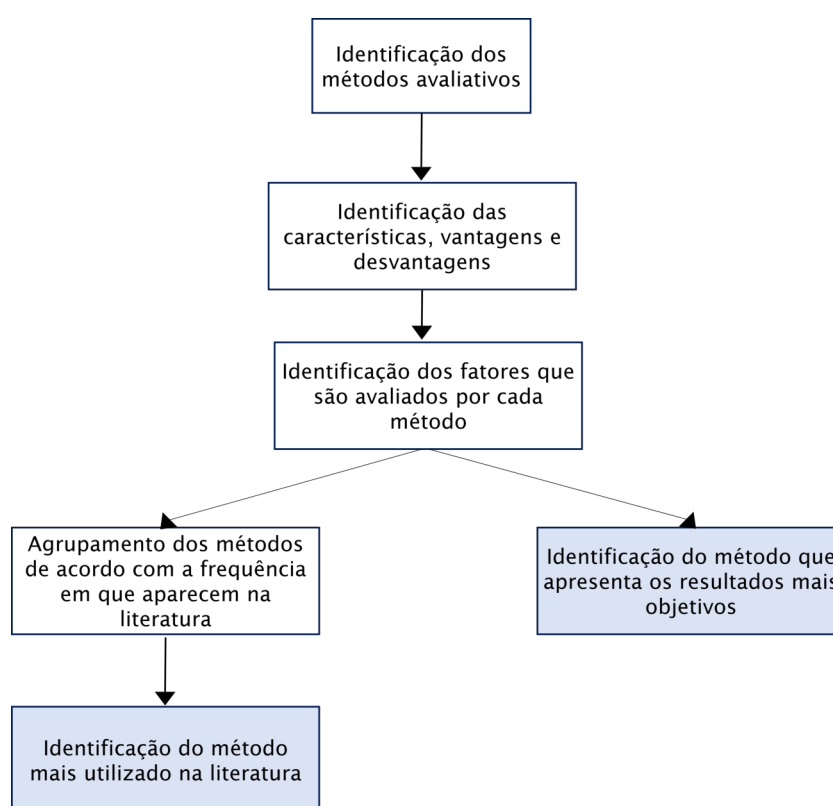
3.2 Métodos de avaliação aplicados

Para selecionar os métodos avaliativos a serem aplicados no trabalho, foram identificados os aspectos principais de cada método encontrado no mapeamento sistemático. Foram verificadas as características principais, as vantagens e desvantagens e os

fatores/dimensões avaliados em cada um. Em seguida, os métodos foram agrupados de acordo com a frequência em que foram aplicados nos trabalhos. O método mais utilizado foi selecionado para ser aplicado. O segundo método selecionado para aplicação foi o que apresentava resultados mais objetivos, no sentido de os dados poderem ser coletados automaticamente e não depender do ponto de vista do usuário. Outro fator relevante para a seleção dos métodos de avaliação foi a viabilidade de sua aplicação diante da situação atual da pandemia do coronavírus.

A Figura 3 mostra um fluxograma do processo descrito.

Figura 3 – Fluxograma das etapas para escolha dos métodos avaliativos a serem aplicados.



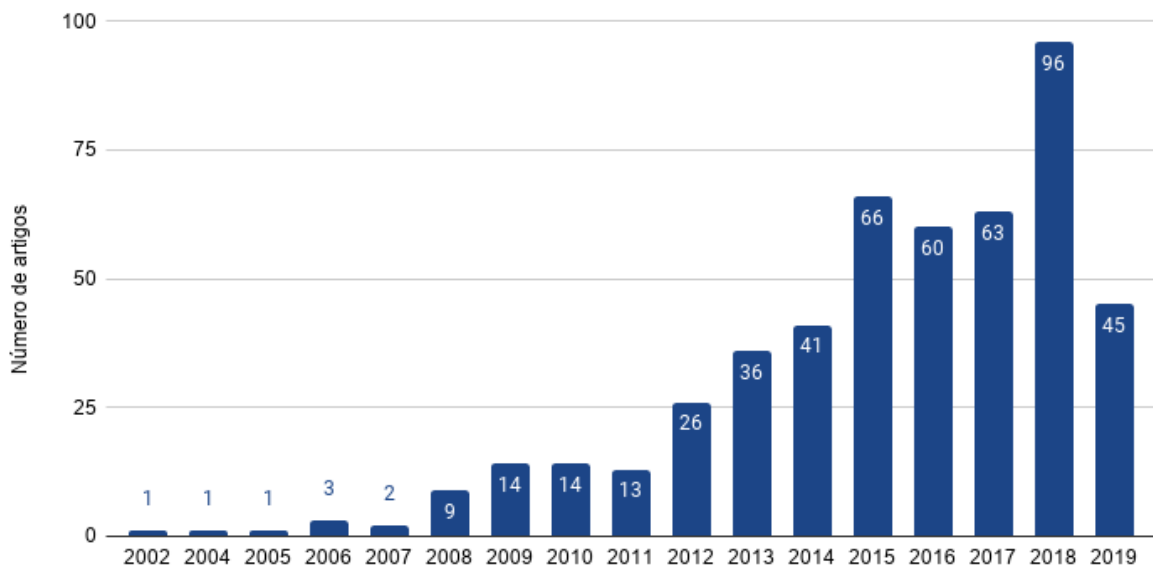
4 Estudo dos métodos avaliativos

Neste capítulo serão apresentados os resultados do mapeamento de acordo com as questões de pesquisa.

4.1 Resultados do mapeamento sistemático

Como pontuado anteriormente, foram aceitos 491 estudos de um total de 2868 a partir dos critérios de inclusão e exclusão. Na Figura 4, os trabalhos aceitos no mapeamento são mostrados de acordo com os anos em que foram publicados. Nota-se que até 2008 o número de trabalhos ficou abaixo da marca de dez. Já em 2012, a quantidade dobrou em relação ao ano anterior e, nos anos subsequentes, o número continuou crescendo atingindo o auge em 2018 com 96 trabalhos. Até agosto de 2019 (mês em que foi realizada a última busca nas bases) o número de trabalhos foi de 45.

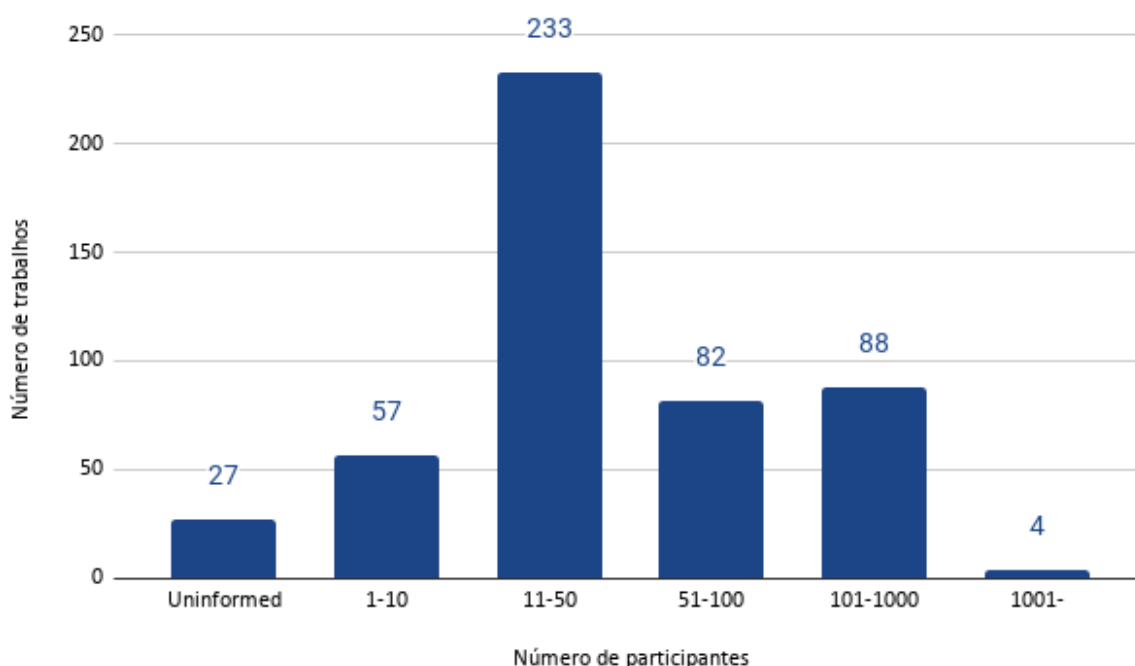
Figura 4 – Distribuição dos trabalhos por ano.



Com relação à quantidade de participantes nos experimentos, 5.5% (27) não informaram a quantidade de usuários utilizados na avaliação. 11.6% (57) dos trabalhos utilizaram amostras de até 10 participantes e 47.5% (233) das avaliações foram executadas com amostras entre 11 e 50 participantes, representando o intervalo mais utilizado. Amostras entre 51 e 100 participantes foram encontrados em 16.7% (82) dos trabalhos e entre 101 e 100 foram 17.9% (88) dos trabalhos. Por fim, amostras com mais de 1000 foram

utilizadas em apenas 4 trabalhos. Ainda, é possível pontuar que o mínimo de participantes utilizados nos experimentos foi o de três e o máximo foi de 99638. Esse último apresenta um número expressivo visto que o jogo foi disponibilizado de forma online, o que possibilitou a adesão de um grande número de usuários (DUNWELL et al., 2014). No gráfico da Figura 5, há a sumarização desses valores.

Figura 5 – Quantidades de usuários utilizadas nos trabalhos.



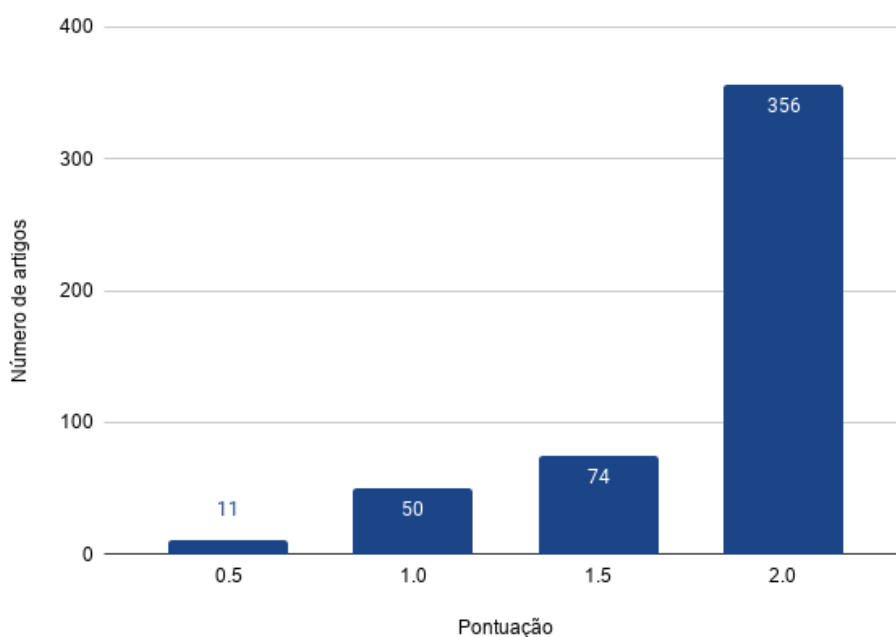
Os trabalhos também foram classificados com pontuação de zero a dois de acordo com os critérios de qualidade descritos na Seção 3.1.3. As pontuações são mostradas na Figura 6. Dos 11 trabalhos que tiveram 0,5 ponto, nove atenderam parcialmente o critério CQ1¹ e não atenderam CQ2². Os outros dois, não atenderam o CQ1 e atenderam parcialmente CQ2. Dos 50 trabalhos que tiveram um ponto, 33 atenderam parcialmente os dois critérios. Os outros 17, atenderam completamente CQ1 e não atenderam CQ2. Dos 74 trabalhos que tiveram pontuação 1,5, 36 atenderam CQ1 completamente e apenas parcialmente CQ2. Os outros 38 atenderam apenas parcialmente CQ1 e totalmente CQ2. Por fim, 356 estudos atenderam ambos os critérios completamente.

Após a classificação, para cada um dos 491 estudos aceitos foram preenchidos formulários de extração com intuito de responder as questões de pesquisa do mapeamento. A partir dos formulários, foi feita a sumarização dos resultados obtidos. Os resultados foram divididos de acordo com a pergunta de pesquisa relacionada.

¹ CQ1 - Métodos da avaliação e resultados estão explícitos?

² CQ2 - As dimensões avaliadas são especificadas?

Figura 6 – Pontuação de acordo com critérios de qualidade.

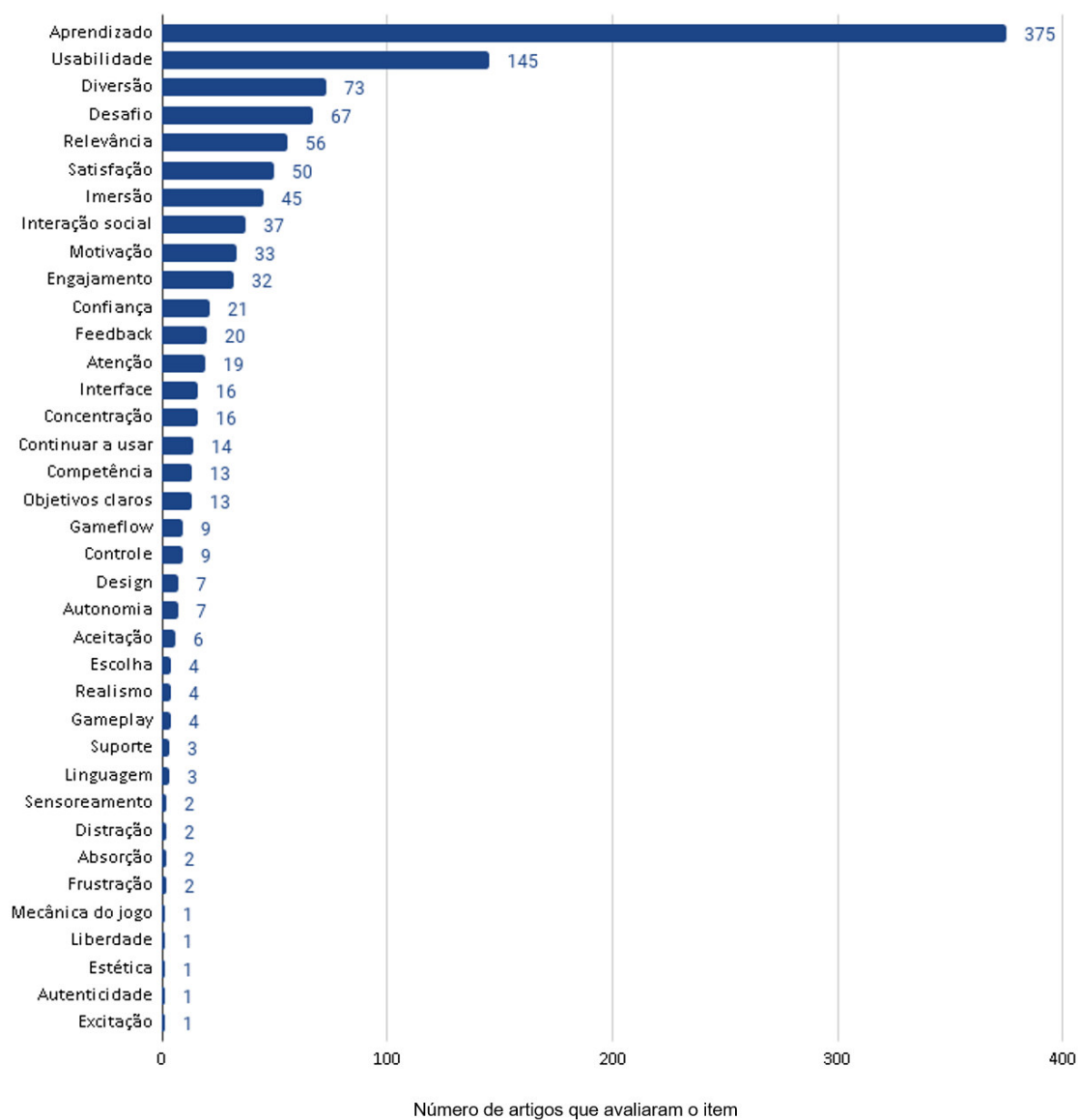


4.2 QP1. Quais os fatores/dimensões que são avaliados nos jogos sérios?

Foram identificados 37 fatores avaliados nos jogos digitais sérios pelos métodos. No gráfico da Figura 7 é possível observar as frequências com que cada um foi analisado nos 491 trabalhos aceitos. A dimensão mais avaliada foi a de Aprendizado com 375 trabalhos. Fatores indicados como Percepção de aprendizado, Conhecimento e Desempenho também foram agrupadas na dimensão Aprendizado.

Outros fatores de destaque são Usabilidade, Diversão, Desafio e Relevância com mais de 50 trabalhos os avaliando. A Usabilidade (avaliada por 145) está relacionada à facilidade de utilizar o jogo. Fatores citados como Facilidade de usar e Navegação foram agrupados na dimensão de usabilidade. Em (ULRICH; HELMS, 2017), os autores afirmam que para criar um jogo que obtenha a fidelidade dos jogadores, os desenvolvedores devem priorizar interfaces fáceis de usar, efeitos de sons imersivos e gráficos e ilustrações detalhistas e bem produzidas. Nesse sentido, o fator Confiança (avaliada por 21) se relaciona com a Usabilidade, pois busca compreender se o usuário sente que consegue utilizar o jogo de forma fácil. Além desse, o fator Interface também está ligado. Já a dimensão Diversão é fundamental tendo em vista que comprovadamente há um relacionamento positivo entre a diversão/motivação e o aumento do aprendizado em um jogo sério (ERHEL; JAMET, 2013). Ligada à Diversão, está a dimensão Gameflow, avaliada por nove trabalhos.

Figura 7 – Dimensões/Fatores avaliados pelos métodos e suas frequências.



Para que o jogo possua desafio, é necessário que haja um equilíbrio entre a dificuldade do jogo e a competência do jogador (PETRI; VON WANGENHEIM, 2017). Assim, para que o usuário não ache o jogo chato e perca o interesse é importante que existam níveis e que a dificuldade do jogo cresça conforme a habilidade do jogador progrida. E a Relevância ou Utilidade assemelha-se à dimensão de Aprendizado, já que verifica se o usuário acredita que o jogo é relevante tanto para os seus interesses como para o de outros.

Satisfação e Motivação, avaliadas por, respectivamente, por 50 e 33 trabalhos, procuram verificar se o usuário encontra-se satisfeito e motivado a continuar jogando. Imersão e Engajamento, avaliadas por, respectivamente, por 45 e 32 trabalhos também

são semelhantes pois verificam se o jogador está em um estado em que sente-se como parte do jogo e temporariamente esquece as tarefas do dia a dia (VALLE et al., 2017). Essa imersão e o envolvimento do jogador são importantes pois estão ligados a continuação do uso do jogo pelo usuário, visto que, por achar a atividade tão agradável, o participante sente a necessidade de se engajar na mesma atividade mais de uma vez (SU; HSAIO, 2015). Assim, é possível perceber que a dimensão Continuar a usar (avaliada por 14) também está diretamente ligada à Imersão/Engajamento. Outras dimensões como Concentração e Atenção também possuem aspectos similares.

Outras dimensões que devem ser pontuadas são *Feedback* e Interação social. A primeira consiste em verificar se é dado *feedback* para o jogador do seu progresso no jogo, como sua pontuação, por exemplo. Em (FU; SU; YU, 2009), os autores afirmam que o *feedback* é importante visto que, permite que o jogador determine o conhecimento necessário que ele precisa adquirir para que possa completar a tarefa. Já a Interação social procura compreender se o o jogo promove um sentimento de ambiente compartilhado, conexão com outros jogadores e competição (PETRI; VON WANGENHEIM, 2017).

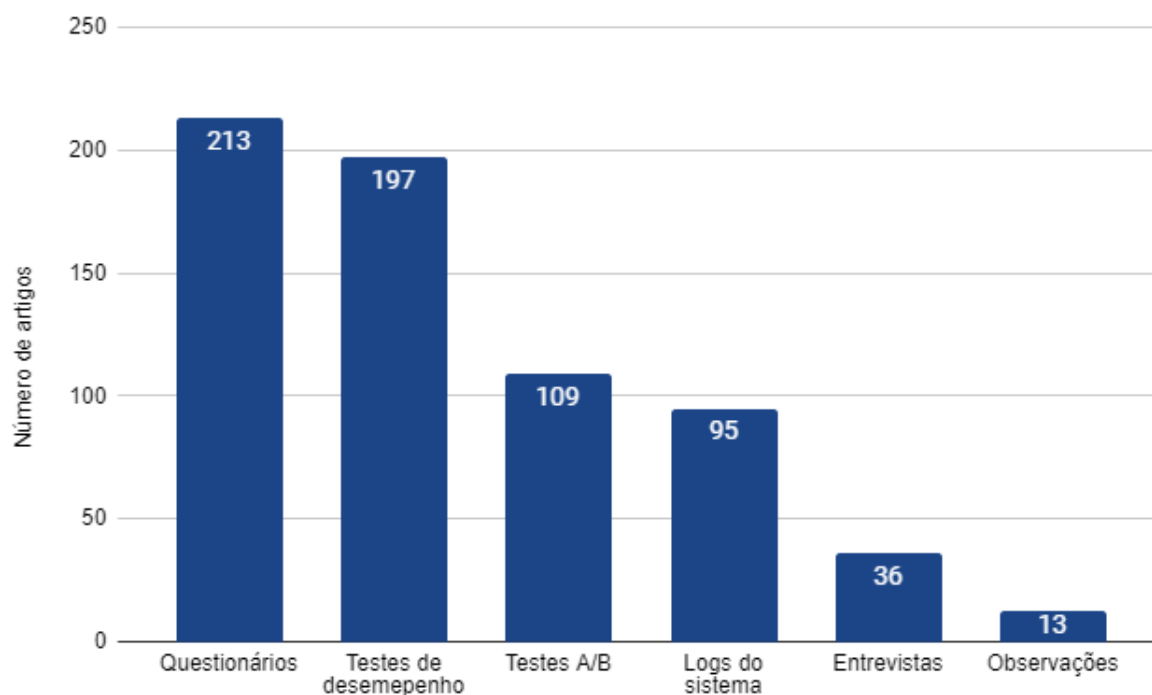
A partir dessa discussão é possível perceber que muitas dessas dimensões podem ser avaliadas em conjunto, como no caso da usabilidade que agruparia os fatores relacionados ao design, confiança, realismo, mecânica do jogo, estética, autenticidade, objetivos claros entre outros.

4.3 QP2. Quais os métodos utilizados para avaliar jogos digitais sérios?

A partir da extração percebeu-se que os estudos se diferenciavam em relação aos métodos aplicados, amostras utilizadas e itens avaliados. Foram identificados seis métodos de avaliação: questionários de experiência do usuário, testes de desempenho, teste A/B, entrevistas, logs do desempenho de cada usuário e observações dos aplicadores. Os trabalhos foram, então, agrupados de acordo com o método utilizado. Contudo, muitos trabalhos aplicaram mais de um método durante sua avaliação.

No gráfico da Figura 8 é mostrado o número de trabalhos para cada método de avaliação. É possível perceber que o método mais utilizado é o de questionários de experiência do usuário com 213 trabalhos, seguido por testes de desempenho (197) e testes A/B (109), sendo que esses dois últimos são geralmente aplicados de forma conjunta. Em seguida, há o método de coleta de logs do sistema com 95 trabalhos, as entrevistas com 36 trabalhos que a utilizam e por fim, as observações que os aplicadores fizeram dos alunos durante o experimento (13).

Figura 8 – Quantidade de trabalhos de acordo com o método de avaliação.

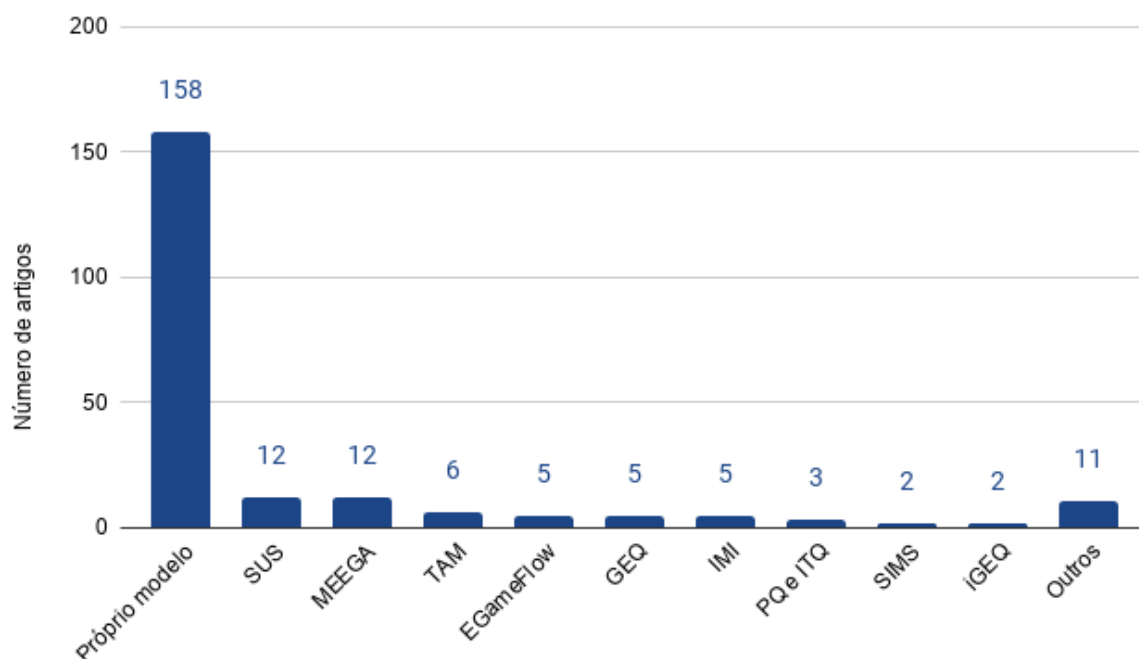


4.3.1 Questionários de experiência do usuário

Esse método consiste na aplicação de questionários ao usuário no final do jogo para verificar aspectos de jogabilidade, percepção de aprendizado, concentração e outros. Nesse tipo de avaliação, os aspectos do jogo são avaliados do ponto de vista do usuário, coletando suas impressões através de perguntas selecionadas. Como ilustrado no gráfico da Figura 9, 63 usam modelos já definidos na literatura como MEEGA (SAVI; VON WANGENHEIM; BORGATTO, 2011), SUS (BROOKE, 1996), TAM (DAVIS, 1989), EGameFlow (FU; SU; YU, 2009), GEQ (BROCKMYER et al., 2009), IMI (RYAN; DECI, 2000), PQ e ITQ (WITMER; SINGER, 1998). Por outro lado, 158 criaram um questionário próprio. É importante pontuar que alguns desses experimentos aplicam mais de um questionário.

MEEGA é um modelo de avaliação que utiliza questionários. Ele permite uma avaliação sistemática da qualidade de um jogo educativo em termos de motivação, experiência do usuário e aprendizado. Também fornece um questionário a ser aplicado aos usuários após a utilização do jogo sério. Nesta primeira versão do modelo, existiam 10 dimensões avaliadas através de 37 questões. As dimensões eram: usabilidade, confiança, desafio, satisfação, interação social, diversão, atenção focada, relevância, aprendizagem de curto prazo e objetivos de aprendizagem. Em (PETRI; VON WANGENHEIM; BORGATTO, 2017), os autores realizaram uma avaliação deste modelo. A análise realizada é baseada na avaliação de 20 jogos diferentes utilizando o modelo MEEGA. Doze dos trabalhos aceitos

Figura 9 – Modelos de avaliação baseados em questionários.



aplicaram o modelo MEEGA para avaliação.

A escala EGameFlow também foi criada para avaliação de jogos. Consiste em 8 dimensões: (1)Concentração: incentivar a concentração enquanto minimiza o estresse referente ao aprendizado. (2)Objetivo claro: atividades devem ser claramente explicadas no início. (3)Feedback: diferença entre o estágio atual e o estágio a ser alcançado. (4)Desafio: fornecer desafio de acordo com as habilidades do jogador. (5)Autonomia: total controle de suas decisões no jogo. (6)Imersão: deve levar o jogador a um estado de imersão. (7) Interação social: jogo deve ser um meio dos jogadores interagirem. (8)Melhoria no aprendizado: o jogo deve aumentar o nível de conhecimento e habilidades do jogador. Ao final, para validar o modelo criado, este foi aplicado em quatro jogos educativos. Para verificar a adequação dos itens, foram utilizados média, desvio padrão e testes de homogeneidade, resultando em um total de 55 questionamentos relevantes para o modelo criado (FU; SU; YU, 2009). Dos trabalhos aceitos no mapeamento, cinco aplicaram o EGameFlow.

Outros modelos de avaliação identificados foram TAM (*Technology Acceptance Model*), SUS (*System Usability Scale*), GEQ (*Game Engagement Questionnaire*) e IMI (*Intrinsic Motivation Inventory*). Esses modelos foram utilizados por 28 dos estudos aceitos no mapeamento. TAM consiste em duas escalas com 14 questões em cada. Avalia a percepção de utilidade e a de facilidade de usar (DAVIS, 1989). SUS é um questionário de 10 itens que avalia usabilidade e pode ser usada em qualquer tipo de sistema (BROOKE, 1996). Já o modelo GEQ contém 19 perguntas e é utilizado para medir o engajamento em

jogos digitais (BROCKMYER et al., 2009). Ademais, o IMI apresenta um questionário com perguntas referentes a diversão, percepção de competência, percepção de escolha e tensão (RYAN; DECI, 2000). O *Presence Questionnaire* (PQ) e o *Immersive Tendencies Questionnaire* (ITQ) são questionários desenvolvidos no trabalho de (WITMER; SINGER, 1998) destinados a avaliação de ambientes virtuais. O primeiro busca medir a presença dos participantes, isto é, a experiência de estar em um ambiente mesmo estando fisicamente situado em outro. Já o segundo propõe medir a imersão dentro desses ambientes.

Outras abordagens utilizaram *User Engagement Scale* (O'BRIEN; TOMS, 2010), *SIMS (Situational Motivation Scale)* (GUAY; VALLERAND; BLANCHARD, 2000), *LES (Learning Effectiveness Survey)* (MOODY; SINDRE, 2003), *QUIS (Questionnaire for User Interface Satisfaction)* (CHIN; DIEHL; NORMAN, 1988), *ARCS* (KELLER, 2009), *SEFQ (Scenario Experience Feedback Questionnaire)* (Weiss et al., 2011), *SFQ (Short Feedback Questionnaire)* (KIZONY; KATZ; WEISS, 2004), *QEF (Quantitative Evaluation Framework)* (ESCUDEIRO; BIDARRA, 2008), *iGEQ (in-Game Experience Questionnaire)* (IJSELSTEIJN et al., 2008), *SDS (Simulation Design Scale)* (JEFFRIES; RIZZOLO, 2006), *SSQ (Simulator Sickness Questionnaire)* (KENNEDY et al., 1993), *UEQ (User Experience Questionnaire)* (HOLZINGER, 2008), *USE (Usefulness, Satisfaction, and Ease of use)* (LUND, 2001) e os instrumentos de avaliação desenvolvidos em (FANG et al., 2010), (MCCOMBS, 2012) e (KIILI, 2006).

Com relação a quantidade de itens (afirmações/perguntas) de cada modelo, esta possuiu um alta variação entre os 213 trabalhos relacionados a questionários. O menor número de questões foi de uma pergunta e o maior número foi de 86 no trabalho de (Tseloudi; Tsiatsos, 2015) que utilizou dois questionários: EGameFlow e USE. A média geral de itens foi de 16,57 por trabalho com desvio padrão relativamente alto de 13,21. Ademais, 58,3% aplicaram menos de 15 questões aos participantes. Outro ponto a ser destacado é que 5,5% não informaram a quantidade de itens aplicados.

4.3.2 Testes de desempenho

Os testes de desempenho correspondem a uma metodologia destinada a avaliar, unicamente, o aprendizado dos alunos com a utilização do jogo. Ele, geralmente, é aplicado juntamente com algum outro método como questionários de experiência do usuário ou testes A/B. Seu processo consiste na aplicação de um teste relacionado aos conhecimentos que o jogo aborda antes da aplicação da atividade. Esse mesmo teste é também aplicado após os participantes utilizarem o jogo. Os resultados dos dois testes são comparados e assim, é possível verificar o nível de aprendizado do jogador.

Em (RANKIN et al., 2008) é desenvolvido um jogo sério para ajudar no aprendizado de uma segunda língua. Para sua avaliação são aplicados testes antes e depois da utilização do jogo. Essa metodologia foi aplicada em 197 dos trabalhos aceitos.

4.3.3 Testes A/B

Os testes A/B também foram utilizados para avaliar o aprendizado dos alunos. Este teste consiste na divisão dos usuários em dois grupos, sendo um grupo de controle e o outro de teste. Para a avaliação de jogos sérios, o grupo de controle corresponde aos usuários que são instruídos pelo método tradicional de ensino (aulas presenciais, por exemplo), já para o grupo de testes, é aplicado o jogo sério. Ao final, são aplicados os mesmos testes de desempenho aos dois grupos e os resultados são comparados.

Em (ASADIPOUR; DEBATTISTA; CHALMERS, 2016) é apresentado um jogo educativo para melhorar a aquisição de habilidade motora. Para avaliar o impacto sobre os participantes, estes foram divididos em dois grupos com o mesmo número de participantes. Questionários de desempenho e de impressões do usuário foram aplicados somente ao final. A mesma abordagem foi utilizada em 109 dos trabalhos aceitos.

Como indicado anteriormente, há ainda trabalhos que utilizam os dois métodos de avaliação de aprendizado. Em (EAGLE; BARNES, 2009), o objetivo principal foi o desenvolvimento de um jogo para auxiliar no ensino de tópicos relacionados a introdução à computação. Este trabalho focou na avaliação do aprendizado dos usuários. Na aplicação do teste A/B, os 55 estudantes que participaram do experimento foram aleatoriamente divididos em dois grupos, sendo que em somente um dos grupos foi aplicado o jogo. Em seguida, foram aplicados testes de desempenho nos dois grupos, antes e depois do experimento. Comparando os resultados de ambos nos testes, foi possível avaliar o desempenho dos usuários. Essa mesma abordagem pode ser encontrada nos trabalhos (CAKIR et al., 2015; KHANE; AHMAD; MALIK, 2017; TRUJILLO et al., 2016).

4.3.4 Logs do sistema

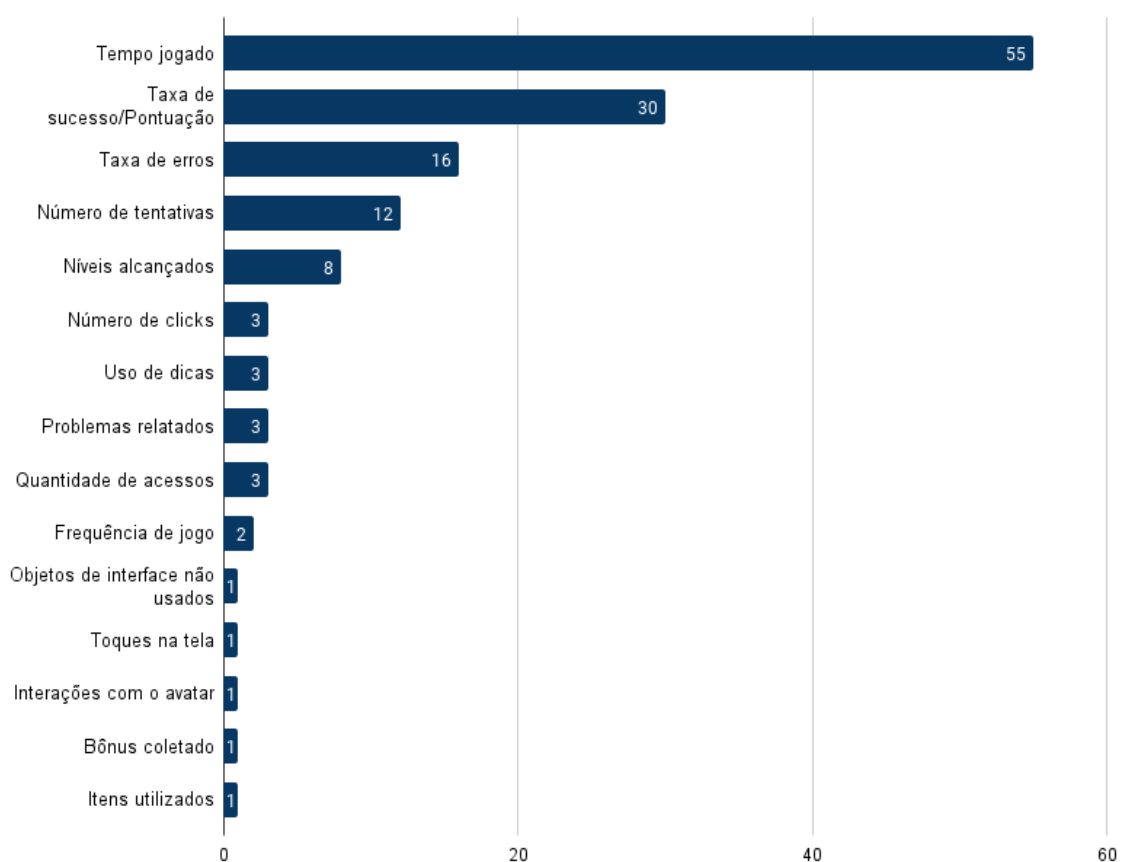
Nesta metodologia são usados logs do sistema, como horas jogadas, acertos e interação entre usuários, para avaliar o jogo. Surge um novo conceito para esse tipo de método derivado de *Learning Analytics*, chamado de *Game Learning Analytics*. *Learning Analytics* é definida em (FREIRE et al., 2016) como exploração de dados de ambientes educacionais para avaliar teorias de aprendizado, *feedback* dos usuários, possíveis problemas nos sistemas e para o desenvolvimento de aplicações de aprendizado. Já o conceito de *Game Learning Analytics* refere-se à medida, coleta, análise e comunicação de dados sobre os participantes com o propósito de entender e otimizar o aprendizado e os ambientes onde ocorre. Em particular, é aplicada a jogos (Calvo-Morata et al., 2018).

Em (GEORGE; CARLOS; DAVID, 2017), é retratado um jogo educativo destinado a auxiliar o ensino de certas áreas da arquitetura. Para avaliação, são usados dados referentes ao tempo gasto no jogo e número de acertos dos participantes. A partir desses dados, o trabalho faz um análise do estado de imersão do participante e o aprendizado adquirido

durante o jogo. Os trabalhos (SCHICKLER et al., 2016; SAFITRI; PRIHATMANTO; RUSMIN, 2015) também utilizam esse mesmo método.

Na extração dos estudos que realizam avaliação por meio de logs do sistema, foram identificados itens que são utilizados para avaliar os jogos sérios. Esses itens são *tempo gasto*, *pontuação*, *taxa de erros*, *número de tentativas*, entre outros. Os itens e suas respectivas frequências podem ser observados no gráfico da Figura 10.

Figura 10 – Itens extraídos nos trabalhos de logs e suas frequências.



Na extração, também foram identificados sete fatores/dimensões avaliados pelos 95 trabalhos que utilizaram logs de sistema. A relação entre cada um desses fatores e os itens extraídos pode ser observada na Tabela 6. Assim, conforme indicado na tabela, o fator aprendizado foi avaliado através dos itens: *número de clicks*, *tempo jogado*, *pontuação* e/ou *níveis alcançados*. Já o fator engajamento foi avaliado por meio dos *níveis alcançados* e *tempo jogado* e assim sucessivamente para os demais fatores.

Tabela 6 – Itens extraídos e respectivas relações com os fatores dos jogos.

Fatores do jogo	Item extraído
Aprendizado	clicks, tempo jogado, pontuação, níveis
Engajamento	níveis, tempo jogado
Usabilidade	tempo jogado, objetos da interface não usados
Dificuldade	tentativas, dicas
Perda de foco	dicas
Motivação	perda de pontos
Controle do jogo	clicks

4.3.5 Entrevistas

As entrevistas fornecem dados qualitativos, isto é, que estão na forma não numérica. Esse método foi utilizado em 36 trabalhos. Ele permite obter uma visão mais específica de cada usuário sobre as impressões gerais do jogo, visto que, as perguntas são abertas permitindo que os usuários respondam livremente. Um exemplo da aplicação desse método ocorre no trabalho de (Ribeiro et al., 2016). Nele, são aplicadas entrevistas individuais com cada usuário após o jogo. São realizadas sete perguntas que buscam compreender a utilidade do jogo, se fornece *feedback*, se é intuitivo e impressões gerais, o que abre espaço para que os usuários possam sugerir melhorias.

4.3.6 Observações dos aplicadores

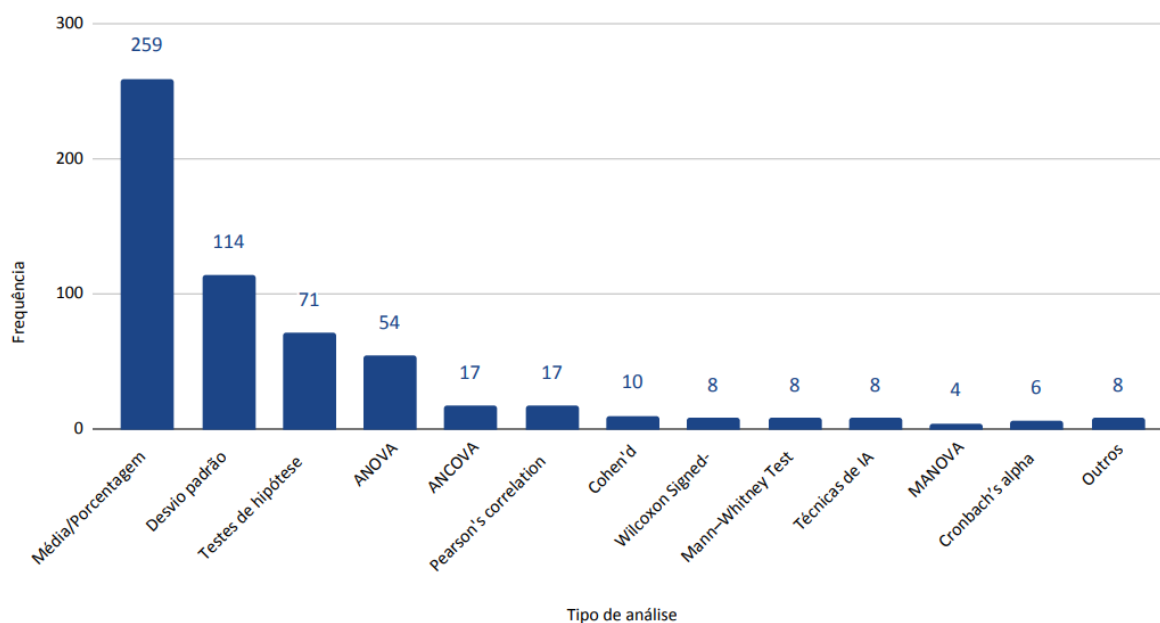
As observações feitas pelos aplicadores também consistem em dados qualitativos. São observadas as interações do jogador tanto com o jogo como com outros jogadores, bem como possíveis distrações. No trabalho de (VIRVOU; KATSIONIS, 2008) foi utilizado esse método. Por meio da observação dos alunos jogando por durante duas horas, os aplicadores conseguiram identificar as oportunidades perdidas durante o jogo (funcionalidades não usadas) e dificuldades gerais. Assim, foi possível que os avaliadores verificassem quanto tempo os usuários desperdiçaram, se desistiram e se houve melhora na sua performance. Esses dados foram então relacionados, por exemplo, quanto mais tempo desperdiçado, menos útil foi o jogo e quanto mais desistências, maior a incidência de problemas de usabilidade.

4.4 QP3. Qual tipo de análise de dados é feita?

A partir da extração foram identificadas diferentes tipos de análises de dados. 74.5% (366) dos trabalhos aplicaram pelo menos um tipo de análise com os dados obtidos.

No gráfico da Figura 11, são mostradas as análises e sua respectivas frequências.

Figura 11 – Tipos de análises de dados identificadas nos trabalhos e suas frequências.



Alguns dos métodos estatísticos identificados foram média, desvio padrão, análise de variância (ANOVA), análise de covariância (ANCOVA), análise multivariada de variância (MANOVA). O cálculo da média foi o mais utilizado, sendo que 70.8% (259) dos 365 trabalhos que fizeram algum tipo de análise a aplicaram. Já o desvio padrão foi utilizado por 31.2% (114). As análises baseadas na variância (ANOVA, ANCOVA e MANOVA), que permite verificar a dispersão estatística dos valores, foram utilizadas em 20.5% (75). Outro tipo de análise que teve destaque foram os testes de hipótese que verificam se os dados amostrais estão de acordo com uma hipótese proposta. Essa análise geralmente foi aplicada nos trabalhos que utilizaram a metodologia de testes de desempenho. No trabalho de (CASANO et al., 2016), por exemplo, os testes de hipótese foram utilizados para comparar se os participantes tiveram um desempenho melhor no teste após o jogo do que no teste feito previamente.

Foram também identificadas técnicas de aprendizado de máquinas e mineração de dados entre os trabalhos. Em (GHALEB et al., 2018) são utilizadas diferentes técnicas de aprendizado de máquinas para identificar o estado afetivo do aluno como tédio, engajamento, frustração em um ambiente de aprendizado. Para isso, é utilizada a Teoria do Flow (um estado mental, onde o participantes se encontra totalmente envolvido em uma atividade). Em (SLIMANI et al., 2018), mineração de dados é definida como um processo de análise de dados através de perspectivas diferentes para que esses dados possam ser sumarizados e transformados em informação útil. Partindo desse pressuposto, os autores aplicam a técnica de *clustering*, que busca descobrir pontos semelhantes proveniente de uma fonte de

dados heterogênea. Outro exemplo da utilização de *clustering* ocorre no trabalho de (KIM et al., 2019). A técnica foi utilizada para identificar grupos de estudantes que compartilham características semelhantes.

4.5 Discussão dos resultados

Os resultados obtidos a partir da extração dos estudos aceitos permitiu que as perguntas de pesquisa fossem respondidas de forma detalhada. Além disso, foi possível coletar dados referentes a distribuição dos trabalhos por ano e as populações utilizadas nos trabalhos. Com relação ao primeiro item, foi possível notar que a avaliação de jogos sérios ganhou interesse nos últimos anos, em especial em 2018, onde foram publicados 96 dos trabalhos aceitos no mapeamento e até agosto de 2019 foram 45 trabalhos. Já em relação ao segundo item, notou-se que o intervalo mais utilizado pelos trabalhos foi o de 11-50 participantes, sendo 47.5% do total de trabalhos que utilizou tal população.

Para a pergunta QP1, foram identificados 37 tipos diferentes de fatores que avaliam jogos sérios. Alguns termos, que possuíam o mesmo conceito no contexto em questão, foram agrupados, estes foram: Aprendizado, Conhecimento e Desempenho; Usabilidade, Facilidade de usar e Navegação; Diversão e Felicidade; Desafio e Dificuldade; e Relevância, Utilidade e Eficácia. Como mencionado anteriormente, a grande parte dos estudos (375) avaliou o aprendizado dos alunos. Esse fator foi avaliado por todos os métodos encontrados, diferentemente das demais dimensões que, em sua maioria, foram avaliadas somente através dos questionários de experiência do usuário.

Com relação a QP2, foi possível especificar cada um dos métodos destacando quais fatores são avaliados, como são aplicados e com que frequência. Os questionários tiveram o maior índice de utilização com 213 estudos, por esse motivo esse método foi selecionado para ser aplicado na plataforma Euklides. Esse tipo de método é fácil de aplicar, pode alcançar um público grande e, como mencionado anteriormente, permite avaliar um grande número de fatores. Os modelos MEEGA e EGameFlow, por exemplo, conseguem avaliar, respectivamente, 10 e 8 dimensões diferentes. Nesse aspecto, os testes de desempenho e os testes A/B podem possuir limitações, visto que em grande parte avaliam somente o aprendizado após determinadas interações ocorrerem. Por esse motivo, tais testes geralmente são aplicados juntamente a questionários de experiência do usuário. Nesse contexto, os testes são utilizados para avaliar o aprendizado e os questionários, que não conseguem coletar de fato o desempenho do jogador e sim, suas impressões, são aplicados para os demais aspectos. Os logs ou relatórios do sistema, apesar de não alcançarem o número de dimensões avaliadas pelos questionários, podem ser usados para inferir desempenho, engajamento, usabilidade, dificuldades na interface, entre outros. Contudo, exigem um análise mais profunda, no sentido de compreender o que os dados

coletados dos usuários podem significar. Como os logs conseguem ser mais objetivos, no sentido de coletar diretamente os dados do usuário, sem necessitar a aplicação de mais um tipo de questionário aos participantes (testes de desempenho e testes A/B necessitariam de tal recurso), o método por avaliação por logs também foi aplicado na plataforma Euklides.

As entrevistas e as observações dos aplicadores são métodos que exigem uma análise de dados mais complexa, visto que seus dados estão na forma qualitativa, exigindo predominantemente uma interpretação dos dados obtidos, ao invés de uma revisão matemática desses resultados. Uma problemática identificada na aplicação de ambos os métodos é a importância de um encontro presencial entre os participantes e os aplicadores do jogo. Diante do contexto mundial da pandemia do coronavírus, o fator citado impediu a utilização desses métodos de avaliação.

Para responder a QP3, buscou-se identificar os tipos de análise utilizados nos trabalhos. Foram encontradas medidas de concentração de dados de uma distribuição, medidas de dispersão e análise de variância e covariância para determinar estatisticamente igualdade entre médias. Outra tendência observada foi a utilização de técnicas de aprendizado de máquinas na análise de dados dos logs de sistema. Tais técnicas foram *clustering* e *naive bayes*. Os dados coletados nos trabalhos de (GHALEB et al., 2018; Yuhana et al., 2017; SLIMANI et al., 2018; KIM et al., 2019) mostram que é uma área que possui potencial para ser utilizada na avaliação de jogos sérios.

É também possível comparar o trabalho realizado com a revisão sistemática de (PETRI; VON WANGENHEIM, 2017), desenvolvida em 2015, que buscou identificar como os jogos sérios voltados para a área da computação são avaliados. Apesar das *strings* de busca terem sido semelhantes, algumas diferenças podem ser elencadas. Tal revisão sumarizou seus resultados através de suas perguntas de pesquisa, constituindo-se de 7 perguntas no total. Foram encontrados 43 fatores de jogos, 4 tipos de estudos (experimental, quasi-experimental, não experimental e *ad-hoc*), os modelos usados na literatura, os métodos de coleta de dados (questionários, testes, entrevistas, observação, desafios/exercícios), populações utilizadas e os tipos de análise de dados realizados (8 tipos). No presente trabalho foram encontrados 37 elementos de jogos avaliados e 12 tipos de análise de dados. Outra importante diferença refere-se aos métodos de coleta de dados encontrados. No mapeamento realizado foi observado que, nos últimos anos, apesar da prevalência de questionários, os trabalhos passaram a utilizar com mais frequência os logs de sistema para avaliar os resultados dos alunos. Isso pode ser justificado pelo fato de os questionários, apesar de serem ferramentas úteis por conseguirem avaliar diversos fatores do jogo, apresentam a perspectiva do usuário com relação ao jogo de uma maneira superficial. Utilizar esse método pode abrir espaços para inconsistências nos resultados e atribuir subjetividade a avaliação de jogos sérios. Os logs do sistema, por sua vez, são os relatórios do próprio jogo podendo agir no sentido de verificar de fato o aprendizado e

o engajamento do aluno. Dessa forma, o método baseado na análise de logs do sistema também foi selecionado para ser aplicado na plataforma Euklides.

5 Plataforma Euklides

Para aplicar os resultados encontrados no mapeamento, optou-se pelo desenvolvimento de uma plataforma web, para disponibilizar os jogos e coletar os dados do desempenho do jogador. Foi necessário utilizar tal abordagem, tendo em vista o cenário atual de pandemia que gerou a necessidade de realizar um experimento remoto, no qual os participantes poderiam acessar os jogos de sua casa. Assim, a plataforma Euklides foi desenvolvida, cumprindo os requisitos das Tabelas 7 e 8.

Tabela 7 – Requisitos funcionais - Plataforma Euklides

	Nome	Descrição
RF1	Cadastro de professores	A plataforma irá permitir o cadastro de professores
RF2	Cadastro de alunos	A plataforma irá permitir o cadastro de alunos somente através de um link de acesso disponibilizado por um professor já cadastrado na plataforma
RF3	Cadastro de salas	Apenas o professor poderá cadastrar novas salas
RF4	Consulta e alteração nos dados das salas	Na tela de gerenciamento da sala, o professor poderá alterar as informações da sua sala e adicionar ou remover jogos
RF5	Permitir o acesso aos jogos sérios disponíveis	Após o cadastro, os alunos poderão acessar os jogos disponíveis para a sua sala
RF6	Monitorar desempenho dos alunos	Os dados dos alunos durante o jogo serão coletados e armazenados em um banco de dados
RF7	Fornecer informações do desempenho dos alunos	Na tela de gerenciamento da sala, será disponibilizada, para o professor, uma visualização geral de todos os alunos e uma visualização detalhada de cada aluno
RF8	Disponibilizar questionário de avaliação do jogo	Os alunos poderão acessar o questionário de avaliação MEEGA+ para avaliar aspectos dos jogos disponíveis

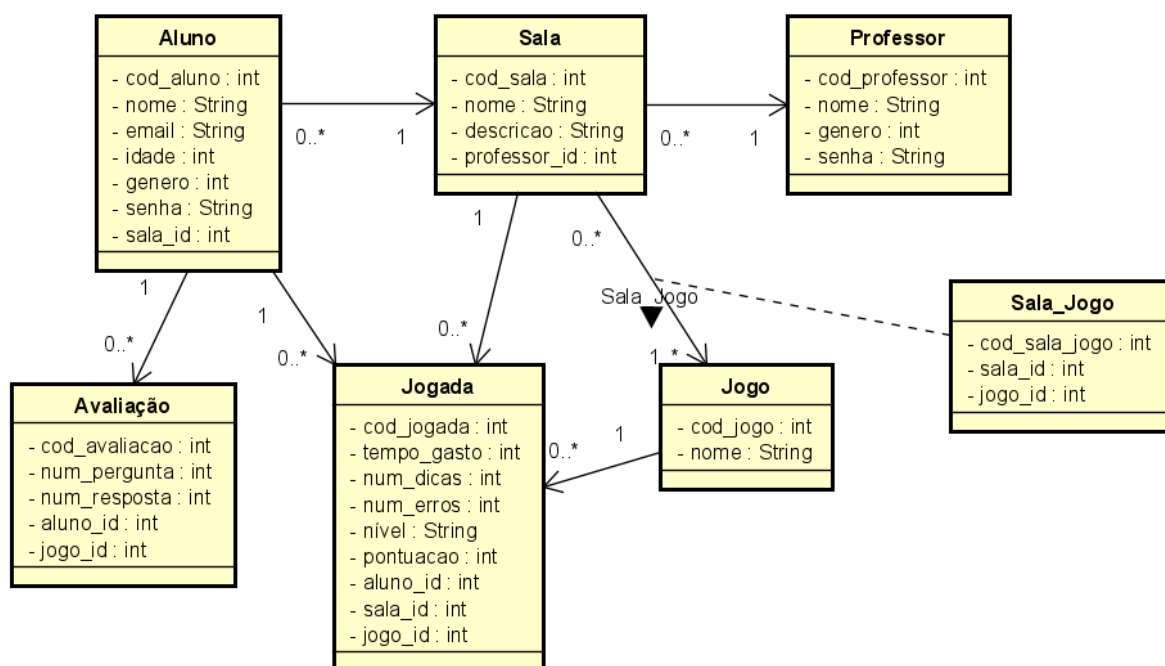
A plataforma foi desenvolvida em PHP, HTML e CSS. O sistema conta com sete classes: Aluno, Professor, Sala, Jogo, Avaliação, Jogada e Sala_Jogo. A Figura 12 ilustra o diagrama de classes. De acordo com o diagrama, tem-se que o Professor pode criar uma ou várias Salas e, nessas Salas, podem estar um ou vários Alunos. Contudo, o Aluno só pode pertencer a uma Sala que, por sua vez, só pode ser criada por um Professor. A partir da relação entre Sala e Jogo, onde várias Salas podem ter vários Jogos, surge a tabela Sala_Jogo. A tabela Jogada está relacionada com três tabelas: Aluno, Sala e Jogo. Nessa tabela, é armazenado um registro quando o participantes finaliza um dos jogos. Já na tabela Avaliação, é armazenado um registro para cada pergunta respondida pelo

Tabela 8 – Requisitos não funcionais - Plataforma Euklides

	Nome	Descrição
RNF1	Armazenamento dos dados	O sistema deverá se comunicar com MySQL Service
RNF2	Escalabilidade	O sistema deverá suportar a inserção de novos jogos digitais sérios sem prejuízos para seu funcionamento
RNF3	Adoção de princípios éticos nas informações dos usuários	As informações dos usuários não serão divulgadas e os dados utilizados na pesquisa não irão conter o nome dos participantes

participante.

Figura 12 – Diagrama de classes da plataforma Euklides.



Após o login, representado pelo diagrama de sequência da Figura 13, o professor acessará sua página, onde irá encontrar as salas que criou. Nessa tela, ele poderá acessar suas salas ou criar novas salas. Essa "sala virtual" permite a adição dos jogos educacionais já cadastrados na plataforma. A plataforma conta, atualmente, com dois jogos. Na Figura 14 é mostrado o diagrama de sequência que representa a consulta e cadastro das salas do professor.

Figura 13 – Diagrama de sequência - Login Professor.

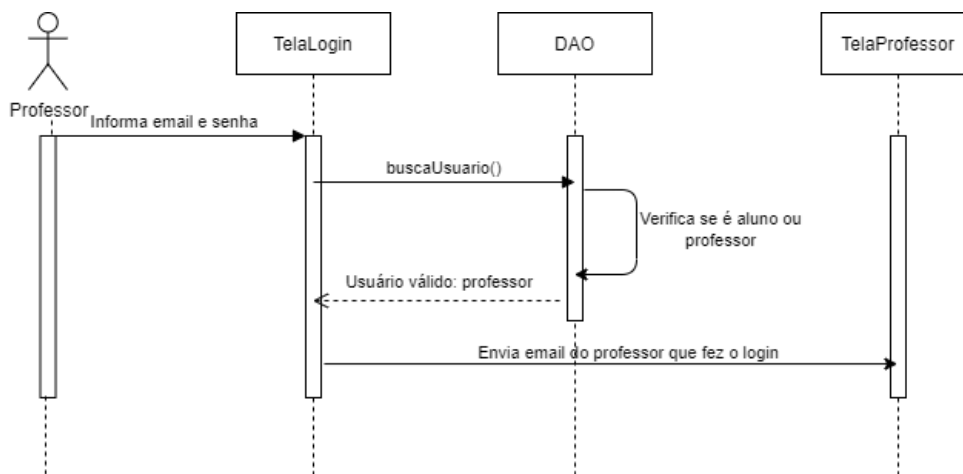
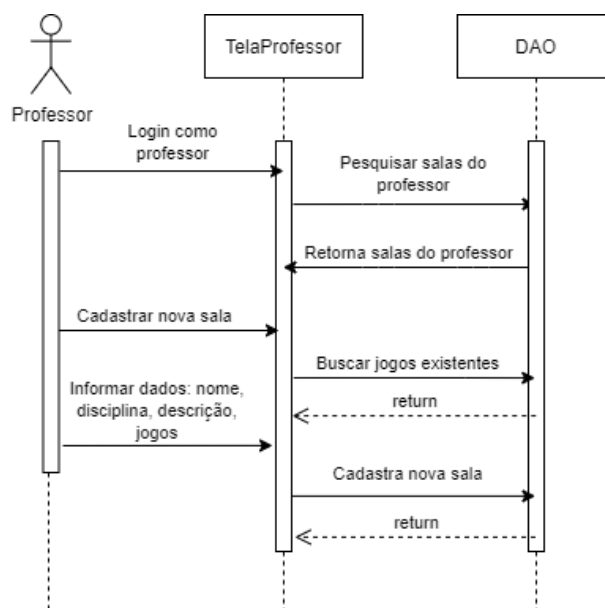


Figura 14 – Diagrama de sequência - Consulta e cadastro de Salas.



A tela inicial da plataforma pode ser observada na Figura 15, na qual é possível realizar login ou cadastro. Já na Figura 16 é mostrada a página do professor, com a visualização de suas salas.

Figura 15 – Login/Cadastro.

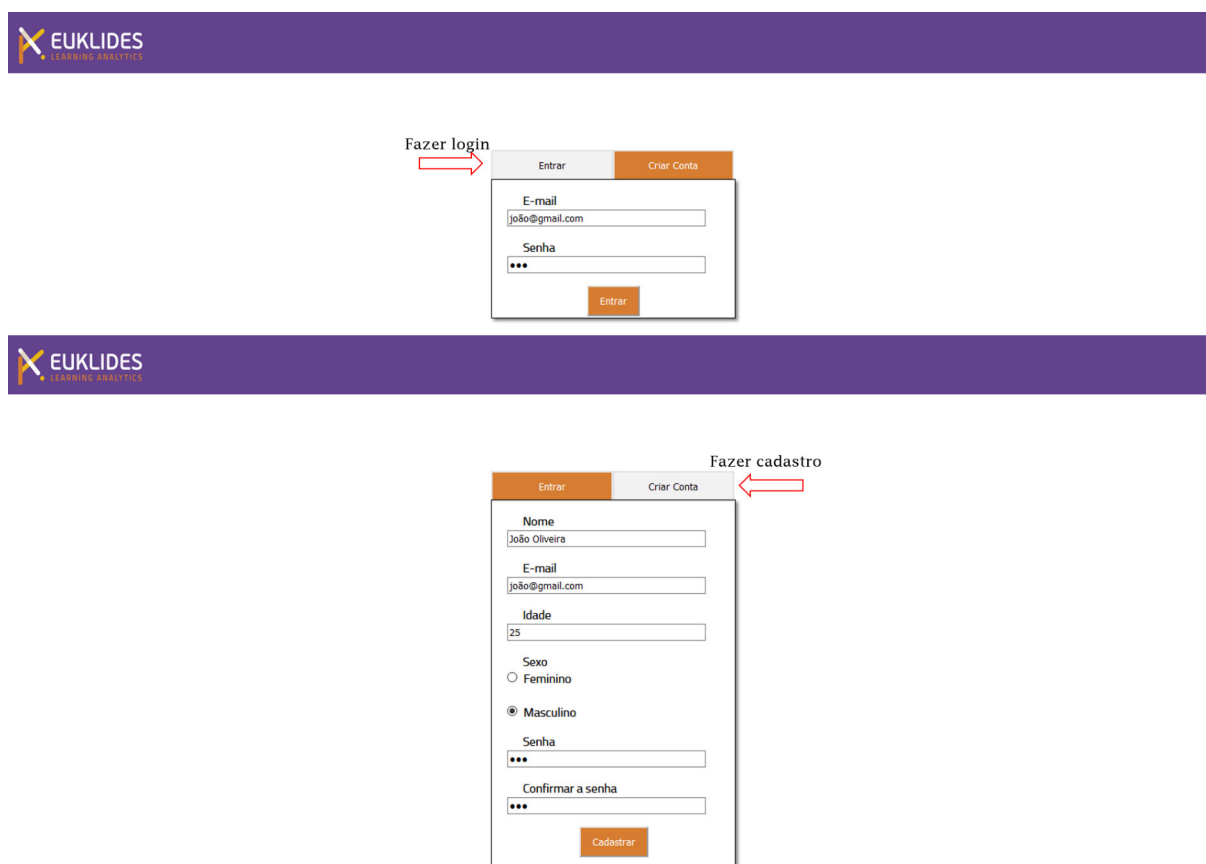
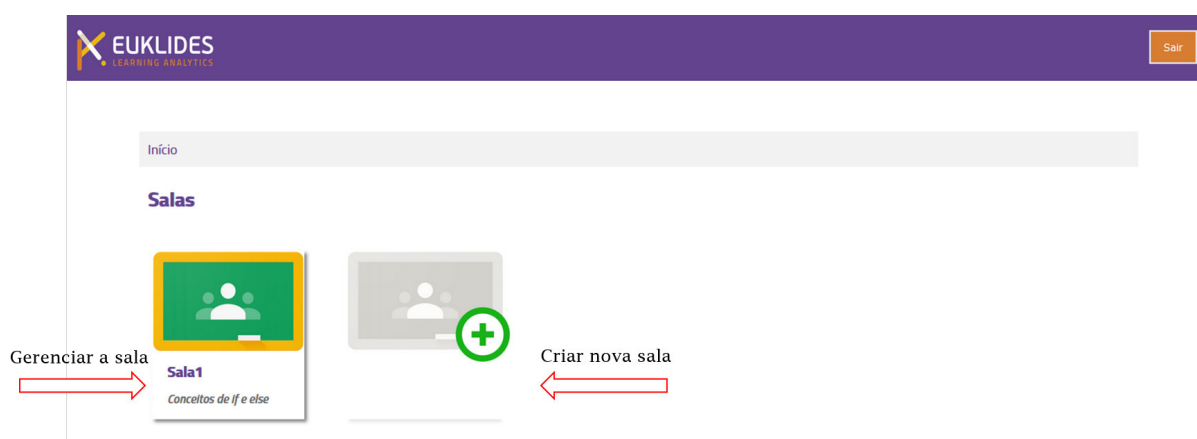


Figura 16 – Página inicial do professor.



Dentro da "sala virtual"(ilustrada na Figura 17), o professor poderá gerenciá-la e disponibilizar o link de cadastro para seus alunos. Os alunos, só poderão fazer seu cadastro através desse link.

Figura 17 – Gerenciamento da sala.



Após o cadastro, os alunos poderão ter acesso aos jogos e ao questionário de avaliação respectivo a cada jogo. A Figura 18, ilustra como é feito o cadastro do aluno na plataforma. A página inicial do aluno (Figura 19) é diferente, visto que o mesmo só pode fazer parte de uma sala, portanto, sua página apresenta os jogos da sala em que está cadastrado. O acesso ao questionário de avaliação é feito através do botão indicado na imagem. O questionário utilizado foi o MEEGA+ (UFSC, 2021), que consiste na versão atualizada do MEEGA (SAVI; VON WANGENHEIM; BORGATTO, 2011).

Figura 18 – Diagrama de sequência - Cadastro de Aluno na plataforma.

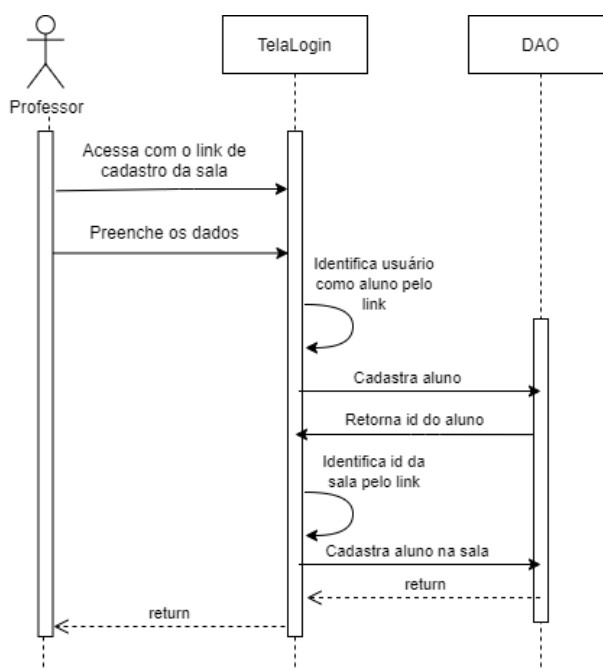


Figura 19 – Página inicial do aluno.

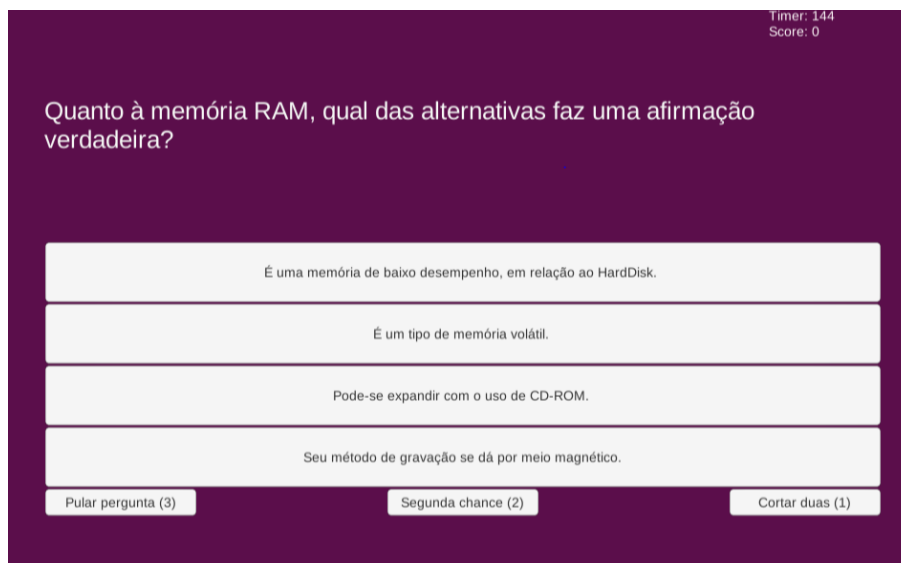


Os dois jogos disponíveis na plataforma são o Quiz Game e o Axio. Eles foram desenvolvidos tendo, como público-alvo, alunos da disciplina de Fundamentos da Computação do curso Bacharelado Interdisciplinar em Ciência e Tecnologia. Tal público é composto, principalmente, por alunos iniciantes na área de Ciência da Computação. Portanto, os jogos foram desenvolvidos com o objetivo de exercitar conceitos iniciais relacionados a área e conhecimentos básicos de lógica matemática.

Quiz é um tipo de jogo no qual os participantes devem responder corretamente as perguntas que são propostas. Permite uma flexibilidade no sentido de poder incluir perguntas de qualquer tema. Como exemplificado no Seção 1.2, já existem plataformas que utilizam questionários como base para a aprendizagem baseada em jogos, como é o caso da *Kahoot*. Nessa plataforma, é possível a criação de testes de múltipla escolha, sendo possível acessar pelo navegador web ou pelo aplicativo Kahoot (KAHOOT, 2020). O quiz criado foi desenvolvido na plataforma Unity (UNITY, 2021).

No Quiz Game, foram atribuídas perguntas de computação básica para revisar conhecimentos relacionados a Fundamentos da Computação. É possível utilizar perguntas de outros temas alterando o arquivo JSON em que elas estão inseridas. Na Figura 20 é mostrada a tela de perguntas do jogo.

Figura 20 – Quiz Game - Tela de perguntas.



O quiz é finalizado quando o jogador responder 10 perguntas ou quando o tempo acabar. O tempo atribuído para cada partida do quiz foi de 150 segundos.

A pontuação do usuário, bem como o tempo gasto para finalizar o quiz e o número de dicas utilizadas pelo jogador, isto é, quantas vezes foram utilizadas as ajudas disponíveis, são armazenados no banco de dados.

Axio foi desenvolvido como um jogo de plataforma em que o personagem deve saltar entre as plataformas, para subir e alcançar seu objetivo. O propósito é exercitar os conhecimentos relacionados a expressões lógicas e despertar o interesse do aluno para o tópico. Axio também foi desenvolvido na plataforma Unity. A tela do jogo pode ser observada na Figura 21.

No Axio, o jogador deve identificar as plataformas com expressões verdadeiras e as com expressões falsas. As com expressões falsas não irão sustentar o personagem e o mesmo irá cair. Se o personagem tocar no chão, perde uma vida. O jogo apresenta quatro níveis que vão aumentando as dificuldades das expressões matemáticas conforme o personagem avança. Quando o personagem alcança a bandeira (Figura 22), o mesmo é automaticamente redirecionado para o próximo nível. Caso o jogador atinja a bandeira e já esteja na última fase (nível 4), o jogo termina. O jogo também termina se o participante perder suas três vidas. O jogador também deve estar atento ao tempo. O tempo atribuído para concluir cada nível foi de 180 segundos, caso esgotasse esse tempo, o participante também perdia uma vida e o cronômetro era reiniciado.

Ademais, os dois jogos foram desenvolvidos, propositalmente, para que apresentassem diferenças entre si. O Quiz é um jogo mais simples, com interface menos robusta, não

Figura 21 – Axio - Jogo iniciado.

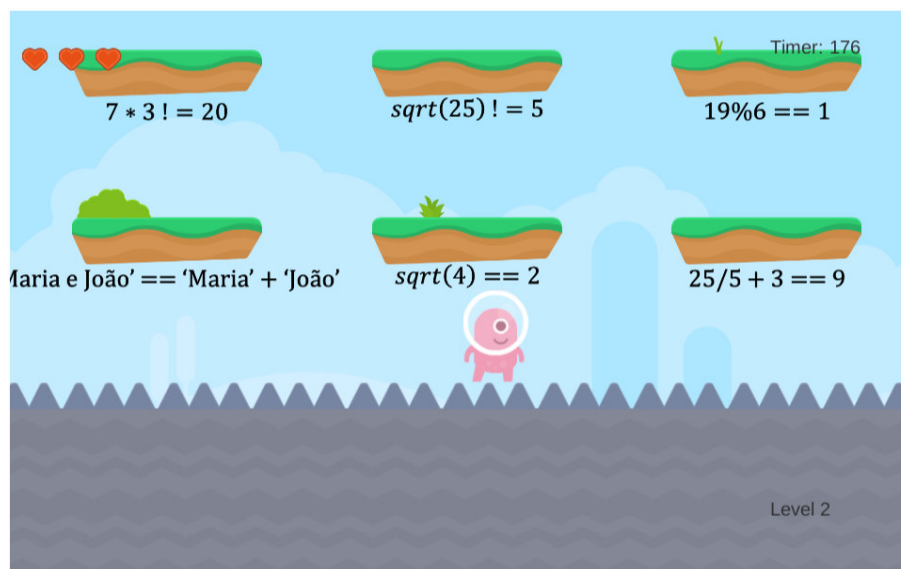
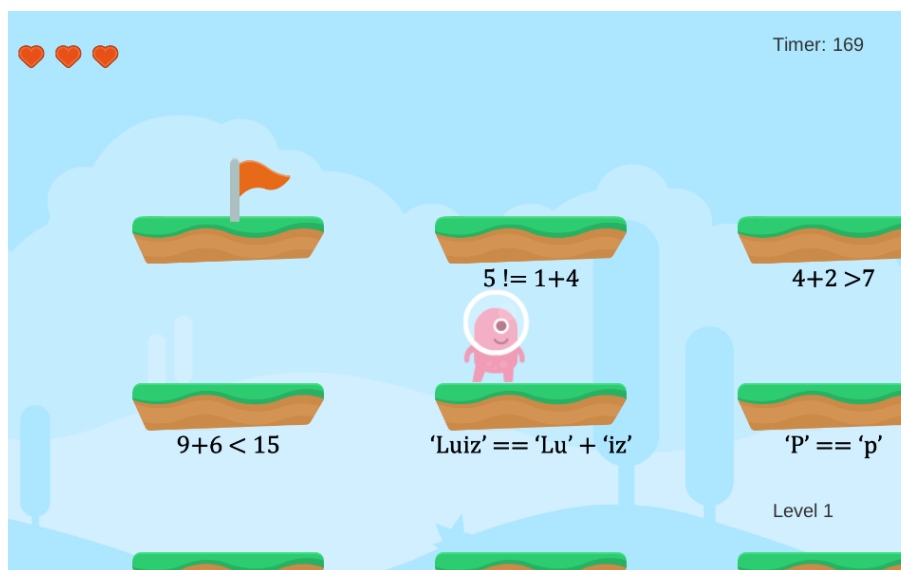


Figura 22 – Axio - Objetivo do nível.



permite interação com personagem e oferece dicas/ajudas. O Axio possui uma interface visualmente mais agradável, interação com personagem e não oferece nenhum tipo de ajuda. Levando em consideração tais diferenças significativas, seria possível verificar o impacto de tais aspectos na avaliação dos jogos.

Um manual do usuário para a plataforma Euklides pode ser encontrado no Apêndice A. Um vídeo do uso da plataforma pode ser visualizado em: <<https://drive.google.com/drive/folders/1JqUeWVowdFllIpyi0h3CtUqtpWaR6N8d?usp=sharing>>.

6 Experimentos

Considerando adversidades que ocorreram no calendário acadêmico e impossibilidade de encontro presencial por conta da pandemia do coronavírus, não foi possível realizar o experimento com o público-alvo original. Assim, o experimento foi realizado com voluntários formados na área de Ciência da Computação. Os participantes foram orientados a jogar os jogos e em seguida, preencher o questionário de avaliação disponível. O MEEGA+ foi utilizado por conta da sua relevância na literatura. O questionário pode ser observado no Apêndice B. MEEGA+ possui 32 afirmações nas quais o participante indica sua concordância em uma escala Linkert de 5 níveis: discordo totalmente; discordo; nem concordo, nem discordo; concordo; concordo totalmente.

Com relação aos logs do sistema, estes foram analisados aplicando testes t. Os testes t são testes de hipótese que podem ser utilizadas para comparar médias, permitindo verificar se a diferença entre elas é significativa. Os testes t permitem calcular o intervalo de confiança da diferença, o que constitui evidência contra a validade da hipótese nula. Assim, a partir do valor-p é possível rejeitar ou não a hipótese nula (BUSSAB; MORETTIN, 2010). No presente trabalho foi adotado um nível de significância $p < 0,05$.

Os três tipos mais comuns de teste t são: para uma amostra, para duas amostras pareadas e para duas amostras independentes. Nesse trabalho foi utilizada a abordagem teste t para amostras independentes com tamanhos e variâncias diferentes (BUSSAB; MORETTIN, 2010). A equação da variável T pode ser observada na Figura 23.

Figura 23 – Equação Teste t para amostras independentes.

$$T = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Com graus de liberdade dados por:

$$gl = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}}$$

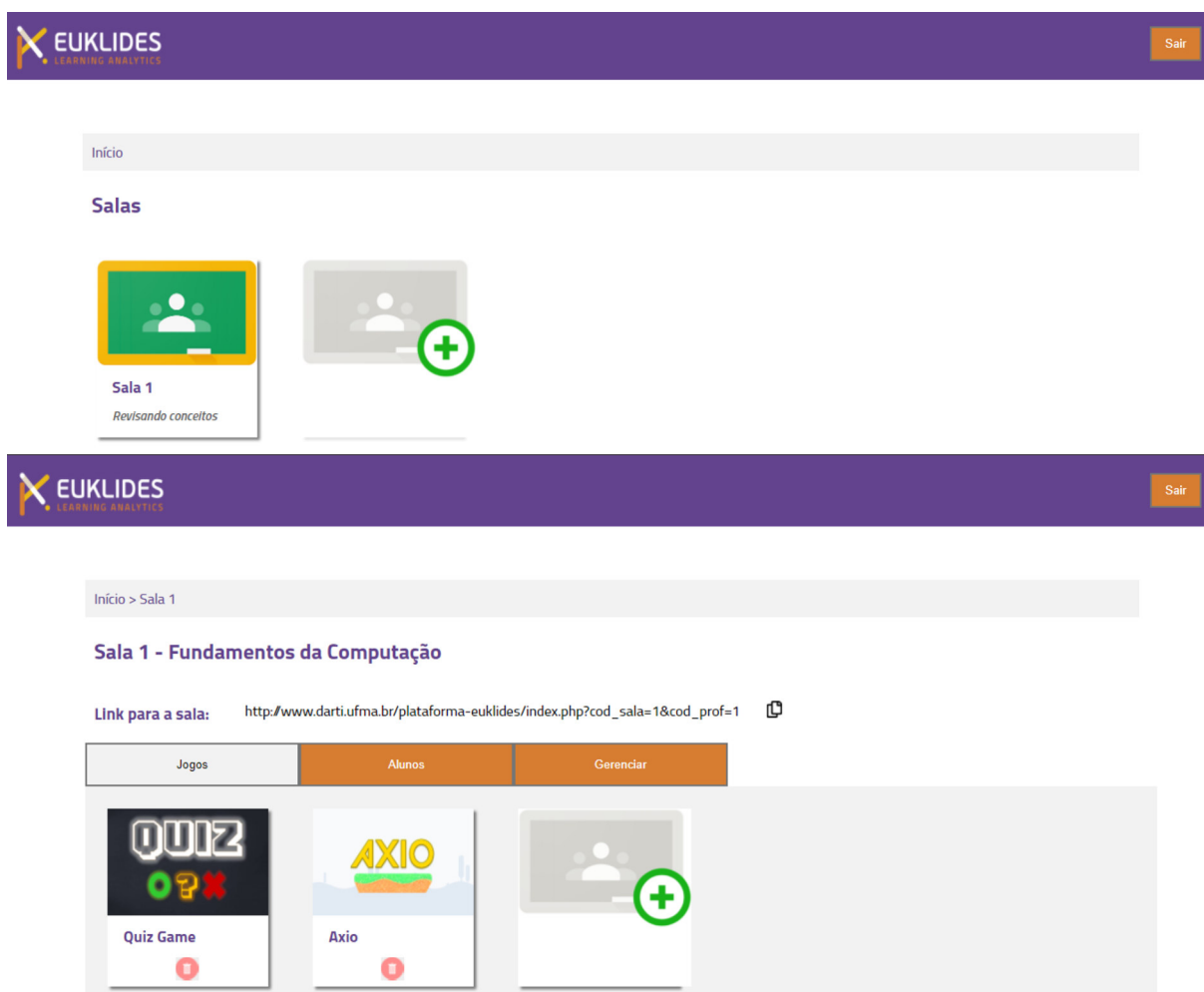
As hipóteses (nula e alternativa) são dadas por:

$$H_0: \mu_I = \mu_{II}$$

$$H_a: \mu_I \neq \mu_{II}$$

Para o experimento, a plataforma foi disponibilizada de forma online através do link: <<http://www.darti.ufma.br/plataforma-euklides/>>. Foi criada uma nova disciplina (Fundamentos da Computação) e nessa disciplina foram adicionados os dois jogos disponíveis na plataforma - Quiz Game e Axio. O link para cadastro de alunos nessa sala (<http://www.darti.ufma.br/plataforma-euklides/index.php?cod_sala=1&cod_prof=1>) foi então enviado para os participantes do experimento. A Figura 24 mostra a sala criada para o experimento.

Figura 24 – Sala criada para o experimento.



Do total de participantes, 8 pertenciam à faixa etária de 18 a 28 anos e 2 à faixa etária de 29 a 39 anos. A aplicação do questionário social também mostrou que 70% dos participantes possuíam o hábito de jogar pelo menos uma vez por semana. Em geral, essas informações indicam a afinidade e o convívio do grupo de alunos com a cultura de jogos, o que lhes permite uma análise crítica mais detalhada do assunto.

Neste capítulo, os resultados referentes aos questionários e aos logs do sistema serão apresentados para cada jogo.

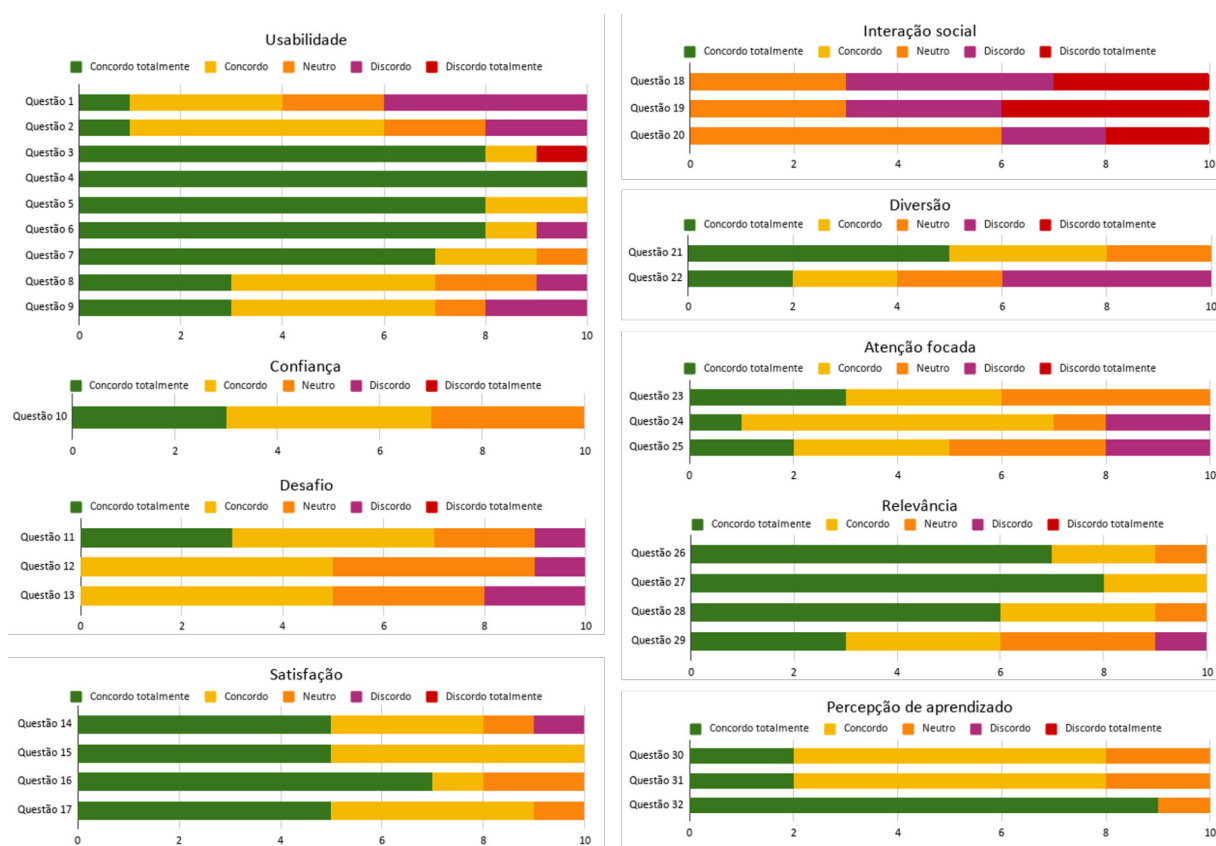
6.1 Avaliação pelo questionário MEEGA+

O questionário MEEGA+ (Apêndice B) avalia as seguintes dimensões: usabilidade, confiança, desafio, satisfação, interação social, diversão, atenção focada, relevância e percepção de aprendizado. Esse método de avaliação foi aplicado nos dois jogos.

6.1.1 Quiz Game

Os 10 participantes jogaram o Quiz Game e o avaliaram através do questionário. Na Figura 25 podem ser observados os níveis de concordância com os itens.

Figura 25 – Resultados do questionário MEEGA+ no Quiz Game.



Com relação à usabilidade as perguntas 1, 2, 8 e 9, referentes ao design do jogo, texto, cores e fontes tiveram os menores índices de concordância na dimensão. As perguntas 3, 4, 5, 6 e 7 relacionadas à facilidade de uso tiveram uma concordância maior, isso pode se dar ao fato de que o quiz é um jogo já bastante conhecido, o que facilita a jogabilidade e a compreensão das regras. Já a dimensão confiança não obteve discordância de nenhum participante, sendo que 70% concordou em partes ou totalmente e 30% optaram pela opção neutra (Nem concordo, nem discordo). Essa confiança se dá principalmente quando o jogo é fácil de usar e o jogador se sente confiante que irá aprender e alcançar os objetivos do jogo.

Na dimensão desafio, apesar de também contar com altos índices de concordância, apenas 30% dos participantes concordaram totalmente com a afirmação 11 (que verifica quão desafiador é o jogo). Com as perguntas 12 e 13 (referentes a novos desafios e monotonia do jogo) nenhum participante concordou totalmente. O quiz, por ser um jogo simples, pode passar a ideia de falta de desafios aos jogadores. Para tentar suprir essa limitação, foram propostos três níveis diferentes aos participantes com graus de dificuldade e pontuação diferentes.

A satisfação, relacionada a realização pessoal do jogador, foi avaliada por quatro perguntas no MEEGA+. Obteve níveis de concordância entre 80 e 100%, sendo que apenas um participante discordou da afirmação de número 14. A interação social, por sua vez, obteve altos índices de discordância. Isso se dá ao fato de que o quiz não conta com recursos de interação social.

A diversão obteve índice de concordância de 80% na afirmação 21 que verificou se os participantes acharam que se divertiram no jogo e índice de concordância de 40% na afirmação 22. Essa última buscou verificar se os participantes sentiram que em algum momento do jogo houve algo que os fizessem sorrir.

A dimensão de atenção focada busca compreender se o participante se engajou no jogo a ponto de perder noção do tempo e do ambiente. As três perguntas utilizadas (23, 24, 25) obtiveram nível de concordância de 50% ou superior. A atenção focada pode ser associada à dimensão de desafio, visto que quanto maiores os desafios, mais o jogador deverá se concentrar para conseguir superá-los.

A relevância do jogo foi avaliada por quatro afirmações e teve concordância superior a 60% em todos os itens. Apenas um participante discordou do item 29 que afirmava que jogar o jogo é melhor que outras formas de ensino. Com relação à percepção de aprendizado, nenhuma das três afirmações obteve discordância por parte dos participantes.

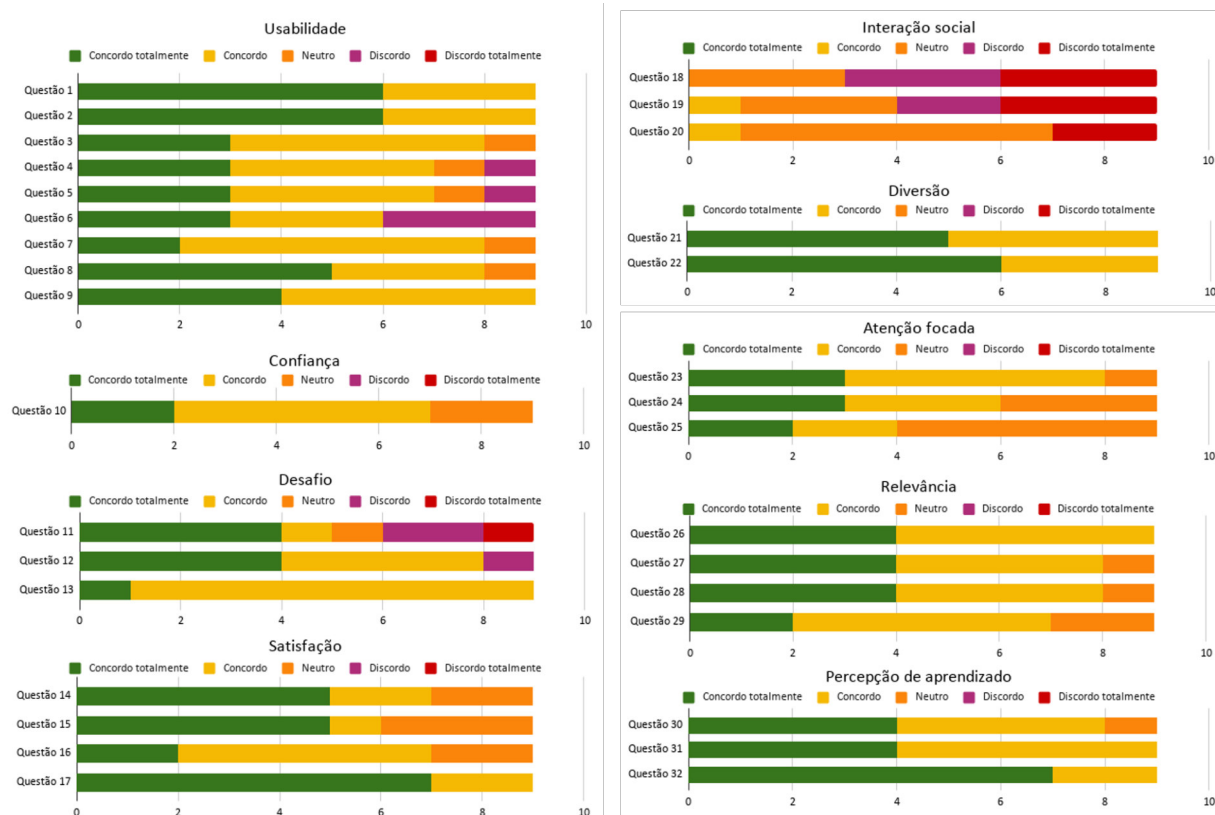
6.1.2 Axio

Dos 10 participantes do experimento, nove jogaram o Axio e o avaliaram através do questionário. Na Figura 26 podem ser observados os níveis de concordância dos itens.

Comparando os resultados do Axio com os do Quiz Game, é possível observar que os itens referentes ao design do jogo (1, 2, 7 e 8) no Axio tiveram concordância de 100% de todos os participantes, enquanto os demais, referentes à facilidade do jogo, tiveram níveis de concordância mais baixos que no Quiz Game. Isso se dá ao fato de que o Axio é um jogo mais complexo e que os participantes não estão tão familiarizados. Tal cenário também impactou nos índices da dimensão de desafio, na qual os participantes avaliaram o jogo como mais desafiador e 100% concordaram que o jogo não se torna monótono.

Nas dimensões confiança, satisfação, interação social e atenção focada, os índices

Figura 26 – Resultados do questionário MEEGA+ no Axio.



foram semelhantes ao Quiz Game. Na satisfação e atenção focada, o Axio se saiu ligeiramente melhor e na interação social, os níveis de discordância foram semelhantes já que o Axio também não oferece ferramentas de socialização. Já na dimensão diversão, 100% dos participantes concordaram parcialmente ou totalmente com as duas afirmações. Relevância e percepção de aprendizado apresentaram níveis de concordância semelhantes ao Quiz Game, mas ligeiramente superiores. Ao comparar os resultados dos dois jogos, nota-se que o Axio teve um melhor desempenho na avaliação por questionários, o que pode estar relacionado ao fato de o mesmo apresentar uma interface mais agradável, desafios mais complexos e permitir uma maior interação com o jogo.

6.2 Avaliação por logs do sistema

Para avaliação por logs foi utilizado o gráfico da Figura 10 do Capítulo 4 como referência.

6.2.1 Quiz Game

Os dados extraídos do Quiz Game foram: tempo jogado, taxa de acertos, taxa de erros, taxa de dicas utilizadas e número de tentativas. O Quiz também permitiu que o

jogador escolhesse entre três temas de perguntas: Software e Hardware (Tema 1); Histórico e Arquitetura dos Computadores (Tema 2); Programação em Python (Tema 3). Todos os temas possuíam 10 perguntas distintas.

A média de tempo jogado pelos participantes foi 380,2 segundos, cerca de 6 minutos e 20 segundos. O desvio padrão foi de 163,73 segundos. O gráfico do tempo jogado pode ser observado na Figura 27. Foi também calculado o tempo médio por partida que foi de 117,45 segundos, com desvio padrão de 19,48 segundos. Os tempos médios por partida de cada participante podem ser observados no gráfico da Figura 28. O tempo máximo atribuído para cada partida foi de 150 segundos, caso o tempo esgotasse, o jogo terminava.

Figura 27 – Tempo jogado no Quiz Game.

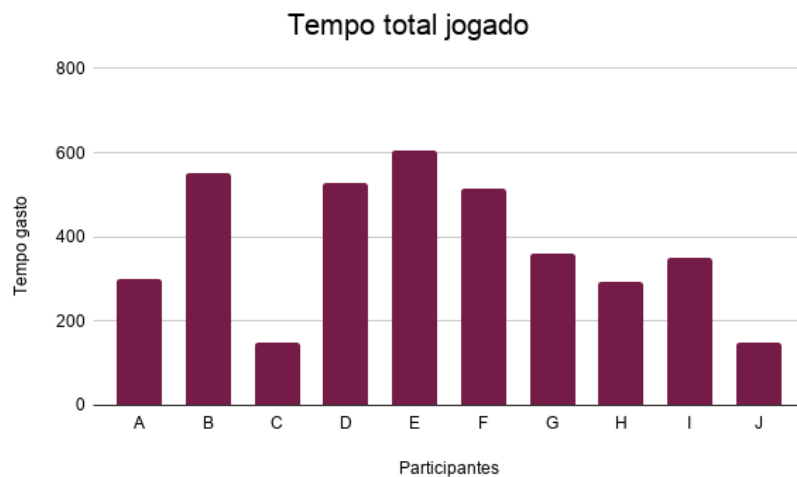
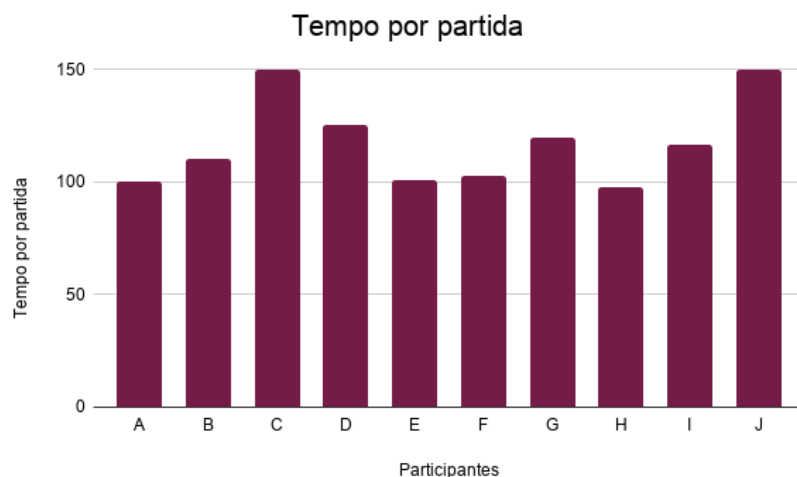
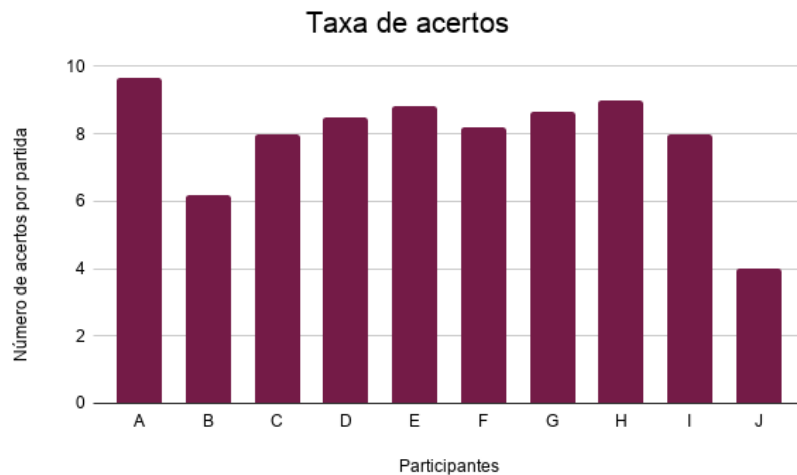


Figura 28 – Tempo por partida no Quiz Game.



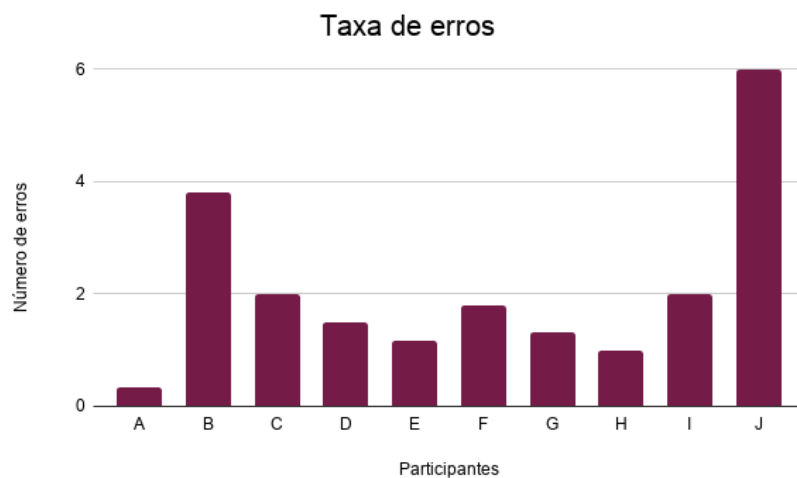
Foram calculadas as taxas de acertos dos participantes, isto é, a média de acertos por partida. A média dessa taxa entre os participantes foi de 7,91, com desvio padrão de 1,65. As taxas de acertos dos participantes podem ser observadas na Figura 29.

Figura 29 – Taxas de acertos por partida no Quiz Game.



Foram calculadas também as taxas de erros dos participantes, nesse caso, a média de erros por partida. A média dessa taxa entre os participantes foi de 2,09, com desvio padrão de 1,65. As taxas de erros dos participantes podem ser observadas na Figura 30.

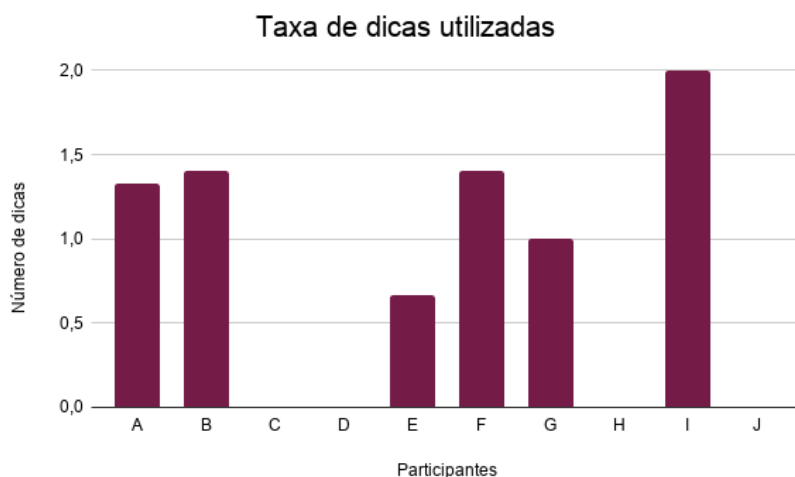
Figura 30 – Taxas de erros por partida no Quiz Game.



As taxas de dicas utilizadas pelos participantes por partida também foram calculadas. Essas dicas correspondiam às ajudas disponíveis no quiz. A média dessas taxas

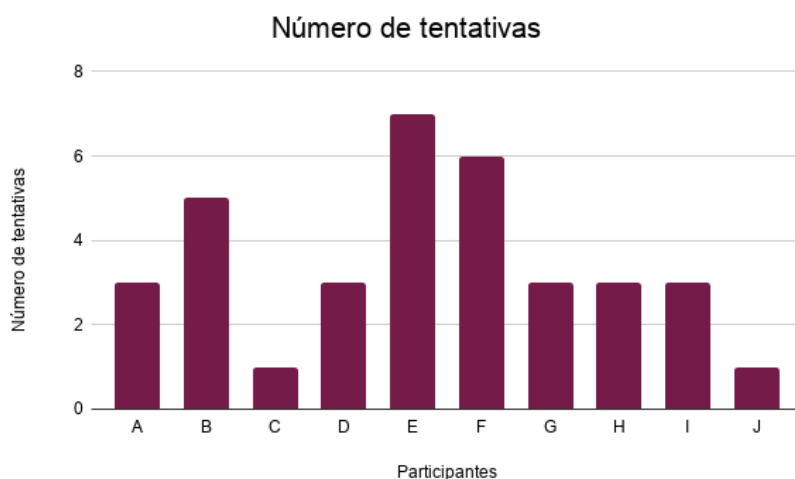
foi 0,78 com desvio padrão de 0,75. O gráfico com a taxa de dicas pode ser observado na Figura 31.

Figura 31 – Taxa de dicas utilizadas por partida no Quiz Game.



A média de tentativas por participante foi de 3,5 com desvio padrão de 1,96. O gráfico com o número de tentativas pode ser observado na Figura 32.

Figura 32 – Número de tentativas no Quiz Game.



6.2.1.1 Relacionamentos entre os dados extraídos

Após a visão geral dos dados, estes foram comparados para verificar o impacto das medidas de tempo, dicas e tentativas sobre o número de acertos dos participantes. Dessa forma, foi possível verificar, por exemplo, se o fato do participante jogar por mais

tempo ou utilizar mais dicas, ocasiona uma média de acertos maior e, portanto, um maior aprendizado. Para realizar cada análise, os participantes foram divididos em duas amostras diferentes de acordo com a mediana dos dados (tempo, dicas, tentativas). A relação entre o tempo total de jogo e o número de acertos pode ser observada na Figura 33. A média de acertos dos participantes que passaram menos tempo jogando foi menor que a dos demais (teste t, $p = 0,7635$). Ou seja, os participantes que jogaram o Quiz por mais tempo, conseguiram ter uma média de acertos maior.

Figura 33 – Relação entre número de acertos e o tempo total de jogo.

Mediana (Tempo total) = 354,5 segundos					
Tempo total de jogo < 354,5			Tempo total de jogo > 354,5		
Alunos	Tempo total de jogo	Número de acertos	Alunos	Tempo total de jogo	Número de acertos
C	150	8	G	360	8,67
J	150	4	F	513	8,2
H	293	9	D	527	8,5
A	301	9,67	B	553	6,2
I	349	8	E	606	8,83
Média:		7,73	Média:		8,08

A relação entre o número de acertos e o tempo médio por partida pode ser observada na Figura 34. Nota-se que a média de acertos dos participantes que levaram menos tempo para concluir a partida, é maior que a dos demais (teste t, $p = 0,3981$). Uma possível interpretação desse resultado é que como o participante sabia mais respostas, acabou respondendo mais rapidamente.

Figura 34 – Relação entre número de acertos e o tempo médio por partida.

Mediana (Tempo por jogada) = 113,6 segundos

Tempo por jogada < 113,6			Tempo por jogada > 113,6		
Alunos	Tempo por jogada	Número de acertos	Alunos	Tempo por jogada	Número de acertos
H	97,67	9,67	I	116,63	8
A	100,33	6,2	G	120	8,67
E	101	8,83	D	125,67	8,5
F	102,6	8,2	C	150	8
B	110,6	9	J	150	4
Média:		8,38	Média:		7,43

A relação entre o número de acertos e a média de dicas utilizadas por partida foi calculada conforme Figura 35. A média de acertos dos participantes que utilizaram mais dicas foi maior que a dos demais (teste t, $p = 0,67$).

Figura 35 – Relação entre número de acertos e dicas utilizadas.

Mediana (Dicas por jogada) = 0,835

Dicas < 0,835			Dicas > 0,835		
Alunos	Dicas	Número de acertos	Alunos	Dicas	Número de acertos
C	0	9,67	G	1	9,67
D	0	6,2	A	1,33	6,2
H	0	8,83	B	1,4	8,2
E	0	8,2	F	1,4	8,67
J	0,67	9	I	2	8
Média:		7,67	Média:		8,15

A relação entre o número de tentativas e o número de acertos foi calculada conforme Figura 36. A média de acertos dos participantes que jogaram mais vezes foi maior que a dos demais (teste t, $p = 0,3981$).

Figura 36 – Relação entre número de acertos e número de tentativas.

Mediana (Tentativas) = 3					
Tentativas <= 3			Tentativas >= 3		
Alunos	Tentativas	Número de acertos	Alunos	Tentativas	Número de acertos
C	1	8	H	3	9
J	1	4	A	3	9,67
G	3	8,67	B	5	6,2
D	3	8,5	F	6	8,2
I	3	8	E	7	8,83
Média:		7,43	Média:		8,38

6.2.2 Axio

Os dados extraídos do Axio foram: tempo jogado, níveis concluídos, número de erros/vidas perdidas e números de tentativas.

A média de tempo total jogado pelos nove participantes foi de 419 segundos, cerca de seis minutos e 59 segundos. O desvio padrão foi de 256,45 segundos. O gráfico do tempo jogado pode ser observado na Figura 37. O tempo médio por partida (Figura 38) foi de 177,92 segundos, com desvio padrão de 71,42 segundos.

Figura 37 – Tempo jogado no Axio.

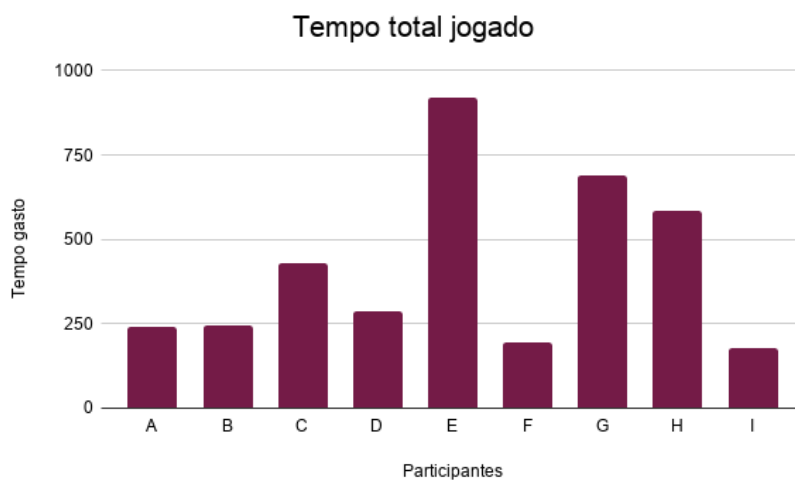
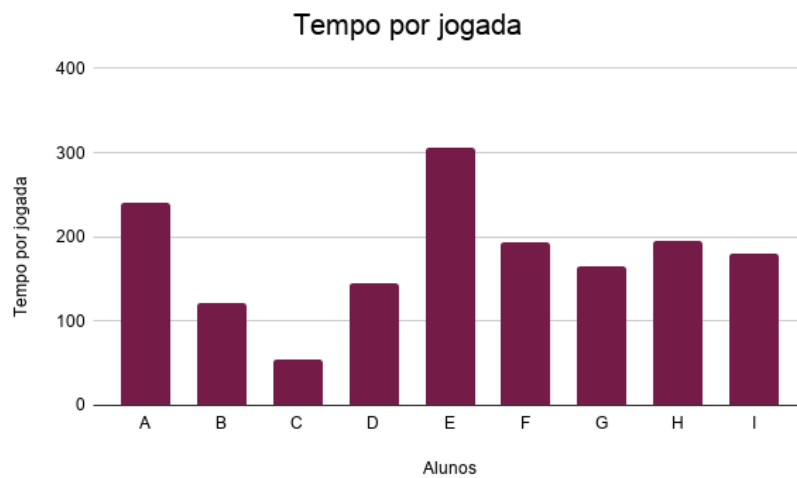
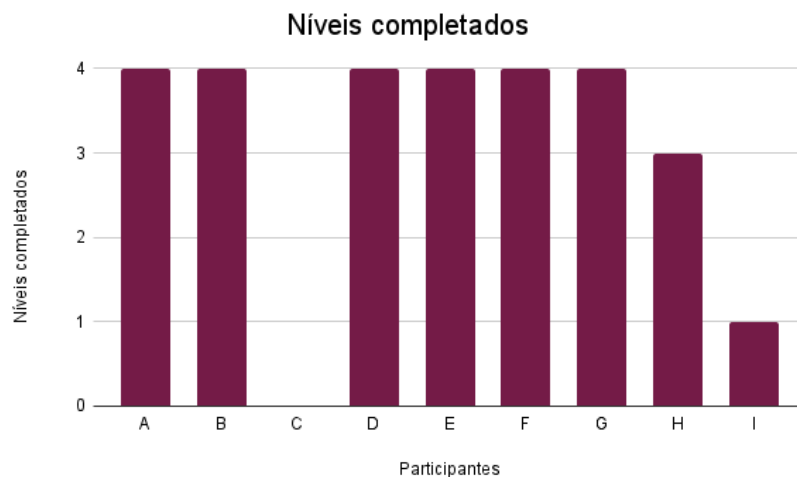


Figura 38 – Tempo médio por partida no Axio.



O Axio possui 4 níveis. Dos nove participantes, seis conseguiram concluir o último nível, finalizando o jogo. A média de níveis alcançados foi de 3,11, com desvio padrão de 1,54. O gráfico desse dado pode ser observado na Figura 39.

Figura 39 – Número de níveis concluídos no Axio.



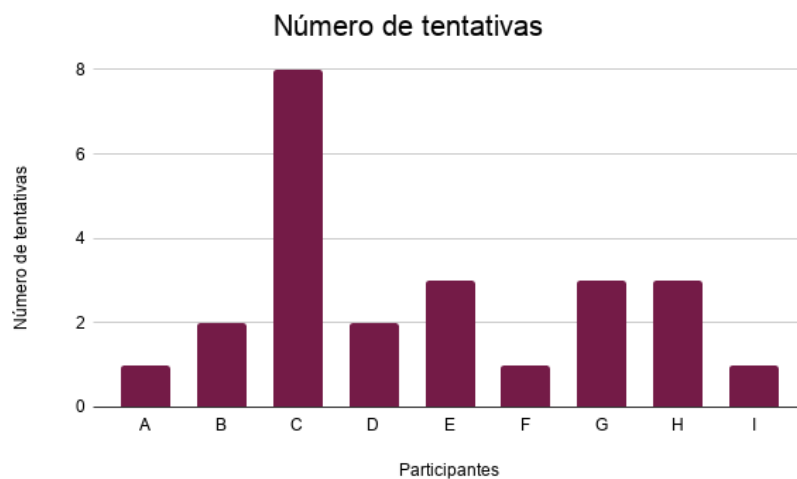
Os erros contabilizados dos jogadores foram as vidas perdidas durante o jogo. O participante poderia perder uma vida ao cair no chão ou caso o tempo esgotasse. A média de vidas perdidas por jogador cada rodada foi de 1,60, com desvio padrão de 1,20. O gráfico com a taxa de erros pode ser observado na Figura 40.

Figura 40 – Taxa de erros por partida no Axio.



A média de tentativas por participante foi de 2,67, com desvio padrão de 2,18. O gráfico com o número de tentativas pode ser observado na Figura 41.

Figura 41 – Número de tentativas no Axio.



6.2.2.1 Relacionamentos entre os dados extraídos

Para o Axio, foi analisado como o tempo jogado e o número de tentativas influenciaram no número de níveis alcançados e na quantidade de vidas perdidas pelo jogador. Assim, a partir dessas análises foi possível identificar se, por exemplo, mais tempo de jogo resultou em um melhor desempenho no jogo. Como três dos nove participantes não conseguiram completar os quatro níveis do jogo, os dados desses foram separados dos

demais participantes. A relação entre os níveis alcançados e o tempo total jogado pode ser observado na Figura 42. Os participantes que completaram o jogo tiveram uma média maior (teste t, $p = 0,8617$), ou seja, jogaram o jogo por mais tempo.

Figura 42 – Relação entre número de níveis alcançados e o tempo total de jogo.

Níveis alcançados < 4			Níveis alcançados = 4		
Alunos	Níveis alcançados	Tempo total	Alunos	Níveis alcançados	Tempo total
C	0	430	A	4	241
I	1	180	B	4	244
H	3	585	D	4	289
Média:		398,33	E	4	919
			F	4	193
			G	4	690
			Média:		429,33

A relação entre os níveis alcançados e o tempo médio por partida pode ser observada na Figura 43. É possível observar que os participantes que não conseguiram finalizar o jogo passaram menos tempo em cada partida (teste t, $p = 0,3812$). O que permite concluir que esse grupo acabou perdendo todas as vidas mais cedo que os demais.

Figura 43 – Relação entre número de níveis alcançados e o tempo médio por partida.

Níveis alcançados < 4			Níveis alcançados = 4		
Alunos	Níveis alcançados	Tempo por partida	Alunos	Níveis alcançados	Tempo por partida
C	0	53,75	A	4	241
I	1	180	B	4	122
H	3	195	D	4	144,5
Média:		142,92	E	4	306,33
			F	4	193
			G	4	165,67
			Média:		195,42

A relação entre os níveis alcançados e o número de tentativas foi analisada conforme Figura 44. A média de tentativas dos três participantes que não conseguiram alcançar o último nível do jogo foi maior que a dos demais participantes (teste t, $p = 0,4388$), ou seja, apesar de não terem conseguido finalizar o jogo, tentaram por mais vezes que os demais.

Figura 44 – Relação entre número de níveis alcançados e o número de tentativas.

Níveis alcançados < 4			Níveis alcançados = 4		
Alunos	Níveis alcançados	Tentativas	Alunos	Níveis alcançados	Tentativas
C	0	8	A	4	1
I	1	1	B	4	2
H	3	3	D	4	2
Média:		4	E	4	3
			F	4	1
			G	4	3
Média:		4	Média:		2

A quantidade média de vidas perdidas por participante também foi analisada com base no tempo total (Figura 45). A média de tempo dos participantes que perderam menos vidas foi menor em comparação aos demais (teste t, $p = 0,322$), ou seja, os participantes que erraram menos durante a partida conseguiram finalizar os níveis de forma mais rápida.

Figura 45 – Relação entre número de erros e o tempo total.

Vidas perdidas < 1.5			Vidas perdidas > 1.5		
Alunos	Vidas perdidas	Tempo total	Alunos	Vidas perdidas	Tempo total
A	0	241	H	2,25	585
F	0	244	E	2,67	919
G	0,67	289	C	3	430
B	1,33	193	I	3	180
D	1,5	690	Média:		528,5
Média:		331,4			

Mediana (Vidas perdidas) = 1.5

Em seguida, foi realizada a análise comparativa entre a quantidade média de vidas perdidas e o número de tentativas, conforme Figura 46. A média de tentativas dos participantes que erraram menos foi menor (teste t, $p = 0,285$), ou seja, os participantes que perderam mais vidas, jogaram o jogo por mais vezes.

Figura 46 – Relação entre número de erros e o número de tentativas.

Mediana (Vidas perdidas) = 1.5		
--------------------------------	--	--

Vidas perdidas < 1.5		
Alunos	Vidas perdidas	Tentativas
A	0	1
F	0	2
G	0,67	2
B	1,33	1
D	1,5	3
Média:		1,80

Vidas perdidas > 1.5		
Alunos	Vidas perdidas	Tentativas
H	2,25	8
E	2,67	3
C	3	3
I	3	1
Média:		3,75

6.3 Comparativo entre o questionário e os logs do sistema

Para análise de cada dado (acertos, tempo jogado, dicas utilizadas, tentativas), os participantes foram divididos em dois grupos para cada dimensão (usabilidade, diversão, atenção, entre outras), de acordo com suas respostas no MEEGA+. Os participantes que cumpriram pelo menos um dos seguintes critérios compuseram o primeiro grupo (G_x), os demais formaram o segundo grupo (G_y):

- Participantes que selecionaram a opção discordo ou discordo totalmente em mais de 20% das afirmações da dimensão.
- Participantes que selecionaram nem concordo, nem discordo em mais de 50% das afirmações da dimensão.

6.3.1 Quiz Game

Foram comparados os itens tempo jogado, taxa de acertos, dicas e tentativas com as respostas dos questionários. As dimensões do questionário consideradas foram usabilidade, desafio, diversão e atenção focada, visto que as demais possuíam respostas muito semelhantes entre todos os participantes.

Os resultados da análise do tempo de jogo total são apresentados na Tabela 9. As diferenças entre as médias dos grupos não foram significativas. A atenção focada teve uma diferença maior (teste t, $p = 0,2063$), o que significa que os participantes que melhor avaliaram esse dimensão, jogaram por mais tempo no total

Tabela 9 – Análise do tempo total em relação às dimensões do questionário

Dimensões	n G _x	n G _y	Média G _x	Média G _y	t	p
Usabilidade	5	5	373,2	387,2	-0.12759	0.9017
Desafio	5	5	379,8	380,6	-0.0072837	0.9944
Diversão	4	6	382,75	378,5	0.036479	0.9721
Atenção focada	3	7	286,33	420,43	-1.4265	0.2063

A análise do tempo por partida gerou os resultados da Tabela 10. É possível observar que a diferença entre as médias dos grupos não foi significativa, contudo, o tempo jogado por partida foi menor entre os participantes que mais discordaram das afirmações de usabilidade do jogo (G_x). Já nas dimensões desafio e diversão, os grupos que mais concordaram com as afirmações de cada dimensão (G_y) utilizaram menos tempo para concluir a rodada. A diferença mais expressiva (teste t, p = 0,2721) foi a dimensão atenção focada. Nesse caso, os participantes que afirmaram estarem mais concentrados no jogo levaram menos tempo em cada rodada.

Tabela 10 – Análise do tempo por partida em relação às dimensões do questionário

Dimensões	n G _x	n G _y	Média G _x	Média G _y	t	p
Usabilidade	5	5	116,03	118,87	-0,21764	0,8332
Desafio	5	5	122,98	111,92	0,88688	0,4023
Diversão	4	6	121,65	114,65	0,52207	0,6201
Atenção focada	3	7	128,88	112,55	1,273	0,2721

A relação entre o número de acertos e o questionário MEEGA+ pode ser observada na Tabela 11. Nota-se que a média de acertos foi maior entre os participantes que classificaram mais positivamente todas as dimensões analisadas (G_y), em destaque a usabilidade com a maior diferença entre os grupos (teste t, p = 0,2298).

Tabela 11 – Análise de número de acertos em relação às dimensões do questionário

Dimensões	n G _x	n G _y	Média G _x	Média G _y	t	p
Usabilidade	5	5	7,21	8,6	-1,4012	0,2298
Desafio	5	5	7,474	8,34	-0,81656	0,4409
Diversão	4	6	7,09	8,45	-1,0685	0,3609
Atenção focada	3	7	6,89	8,34	-0,95871	0,4263

Os resultados referentes a relação entre as dicas e às dimensões do questionário podem ser observados na Tabela 12. Nas dimensões usabilidade, desafio e atenção focada, os participantes que avaliaram mais positivamente o jogo (G_y), utilizaram menos dicas, com destaque para a usabilidade que apresentou a maior diferença entre as médias de seus dois grupos (teste t, $p = 0,07835$). Em contrapartida, o grupo que avaliou mais positivamente a diversão, utilizou mais dicas que o outro grupo.

Tabela 12 – Análise de número de dicas em relação às dimensões do questionário

Dimensões	n G_x	n G_y	Média G_x	Média G_y	t	p
Usabilidade	5	5	1.03	0.33	2.0333	0.07835
Desafio	5	5	0.88	0.68	0.40186	0.6989
Diversão	4	6	0.6825	0.8450	-0.31891	0.7597
Atenção focada	3	7	1	0.686	0.49624	0.6554

Na Tabela 13, são mostrados os resultados referentes à relação entre número de tentativas e as dimensões. As diferenças entre os grupos não foi significativa em nenhuma das dimensões, sendo a diferença mais expressiva (teste t, $p = 0,1504$) na dimensão de atenção focada. Nesse último caso, o grupo de participantes que avaliou mais positivamente a atenção focada (G_y) jogou mais vezes o jogo.

Tabela 13 – Análise de número de tentativas em relação às dimensões do questionário

Dimensões	n G_x	n G_y	Média G_x	Média G_y	t	p
Usabilidade	5	5	3,6	3,4	0.1525	0.8826
Desafio	5	5	3,2	3,8	-0.46291	0.6564
Diversão	4	6	3	3,83	-0.68168	0.5151
Atenção focada	3	7	2,33	4	-1.6161	0.1504

As tabelas completas que mostram a divisão dos grupos para cada dimensão podem ser observadas no Apêndice C.

6.3.2 Axio

Foram comparados os dados referentes ao tempo por partida, níveis alcançados, taxa de erros e número de tentativas. As dimensões consideradas foram usabilidade, desafio e atenção focada, já as demais dimensões apresentavam respostas muito semelhantes entre os participantes.

Os resultados referentes aos níveis alcançados são mostrados na Tabela 14. As diferenças entre as médias dos grupos também não foram significativas, principalmente a dimensão desafio que apresentou médias muito próximas entre os grupos (teste t, $p = 0,8234$). Na usabilidade, os participantes que avaliaram mais positivamente a dimensão (G_y) conseguiram alcançar mais níveis. Em contrapartida, os usuários que avaliaram a atenção focada mais negativamente (G_x) tiveram uma média de níveis alcançados maior, ou seja, apesar de não concordarem totalmente com a dimensão de atenção focada, conseguiram alcançar mais níveis do que os demais.

Tabela 14 – Análise do número de níveis alcançados em relação às dimensões do questionário

Dimensões	n G_x	n G_y	Média G_x	Média G_y	t	p
Usabilidade	2	7	2,5	3,29	-0.49014	0.6939
Desafio	4	5	3,25	3	0.23187	0.8234
Atenção focada	3	6	4	2,67	1.865	0.1212

A comparação entre o tempo total jogado e as dimensões pode ser observada na Tabela 15. Nota-se que os grupos que avaliaram mais positivamente as dimensões (G_y), jogaram o jogo por mais tempo. Na usabilidade, a diferença entre os dois grupos foi significativa (teste t, $p = 0.04064$).

Tabela 15 – Análise do tempo total em relação às dimensões do questionário

Dimensões	n G_x	n G_y	Média G_x	Média G_y	t	p
Usabilidade	2	7	216	478,14	-2.5178	0.04064
Desafio	4	5	326,75	492,8	-0.96185	0.3688
Atenção focada	3	6	167,72	183,01	-0.3945	0.7063

A comparação entre o tempo por partida e as dimensões pode ser observada na Tabela 16. Constata-se que o tempo por partida foi maior entre os participantes que avaliaram mais positivamente as três dimensões (G_y). A usabilidade foi a que apresentou a maior diferença entre os grupos (teste t, $p = 0,4573$).

Tabela 16 – Análise do tempo por jogada em relação às dimensões do questionário

Dimensões	n G _x	n G _y	Média G _x	Média G _y	t	p
Usabilidade	2	7	151	185,61	-0.83249	0.4573
Desafio	4	5	165,17	188,12	-0.50355	0.636
Atenção focada	3	6	167,72	183,01	-0.3945	0.7063

Conforme Tabela 17, a taxa média de erros dos participantes que avaliaram mais positivamente a usabilidade (G_y) foi menor, apesar de não ter sido uma diferença significativa (teste t, p = 0,5405). No desafio, o grupo que avaliou positivamente a dimensão (G_y), possuiu uma média de erros maior (teste t, p = 0,4755). Na atenção focada, a diferença foi mais expressiva (teste t, p = 0,0866), com os participantes que avaliaram mais positivamente a dimensão (G_y), tendo uma média de erros maior que os demais.

Tabela 17 – Análise da taxa de erros em relação às dimensões do questionário

Dimensões	n G _x	n G _y	Média G _x	Média G _y	t	p
Usabilidade	2	7	2,17	1,44	0.75448	0.5405
Desafio	4	5	1,25	1.884	-0.75849	0.4755
Atenção focada	3	6	0,723	2,04	-2.0345	0.0866

De acordo com a Tabela 18, os participantes que avaliaram mais positivamente as três dimensões (G_y) tiveram um número de tentativas maior em todos os grupos, com destaque para usabilidade que obteve a maior diferença entre os grupos (teste t, p = 0,1913). Assim, os participantes que avaliaram mais positivamente o jogo nessas dimensões, jogaram o Axio por mais vezes.

Tabela 18 – Análise do número de tentativas em relação às dimensões do questionário

Dimensões	n G _x	n G _y	Média G _x	Média G _y	t	p
Usabilidade	2	7	1,5	3	-1.4573	0.1913
Desafio	4	5	1,75	3,40	-1.2695	0.2583
Atenção focada	3	6	2	3	-0.82572	0.4366

As tabelas completas que mostram a divisão dos grupos para cada dimensão podem ser observadas no Apêndice D.

6.4 Discussão dos resultados

A partir dos resultados encontrados foi possível verificar o impacto de certos itens no desempenho do participante e nas suas respostas no questionário avaliativo. Um maior tempo total de jogo mostrou um impacto positivo no número de acertos do Quiz Game e no número de níveis alcançados no Axio. Com relação ao questionário, os alunos que jogaram mais tempo também avaliaram ambos os jogos de forma mais positiva. No Axio, por exemplo, o grupo que melhor avaliou a usabilidade jogou significativamente mais tempo que os demais. Já no Quiz Game, o maior impacto do tempo jogado foi na dimensão atenção focada.

A quantidade de vezes jogadas, ou seja, o número de tentativas foi maior entre os participantes que tiveram um melhor desempenho no Quiz Game. Em contrapartida, no Axio os participantes que não conseguiram completar o jogo tentaram mais vezes. Esse último fato mostra que os participantes jogavam repetidamente até conseguirem concluir o jogo. Como o Quiz é um jogo em que todos conseguem finalizar e o que difere é a pontuação, os participantes que jogaram mais vezes, possivelmente, buscavam melhorar sua pontuação. Em ambos os jogos, os participantes que tentaram mais, avaliaram o jogo de forma mais positiva no questionário. No Quiz Game, o maior impacto das tentativas foi na dimensão atenção focada e no Axio, nas dimensões usabilidade e desafio.

Com relação as dicas/ajudas disponíveis no Quiz Game, os participantes que mais as utilizaram tiveram um número de acertos maior. Já o questionário avaliativo mostrou que o grupo que avaliou mais positivamente a atenção focada utilizou menos dicas, o que sugere que os participantes que indicaram não estarem totalmente concentrados usaram mais dicas.

Vale destacar também a relação entre erros/vidas perdidas e a dimensão atenção focada no Axio. De acordo com os resultados do questionário avaliativo, os participantes que mais cometeram erros foram também os que melhor avaliaram a atenção focada. Isso pode se justificar pelo fato de que os participantes que erraram mais também jogaram por mais vezes.

Como citado anteriormente, nem todas as dimensões do questionário avaliativo foram comparadas, visto que, certas dimensões apresentavam respostas muito semelhantes entre os participantes, o que dificultaria dividi-los em grupos distintos. Essa questão de dados semelhantes entre os participantes foi notada também na avaliação dos logs. Uma possível solução para o problema seria a aplicação dos dados em uma população maior.

7 Conclusão

A partir da execução do Mapeamento Sistemático foram identificados 491 artigos relevantes e foram extraídos os seis métodos de avaliação mais recorrentes na avaliação de jogos sérios. Dentre estes, o que mais se destaca em números é a abordagem baseada na aplicação de questionários. Nessa metodologia, é possível avaliar vários fatores dos jogos. Contudo, essa avaliação é feita com base somente na perspectiva do usuário, atribuindo subjetividade aos resultados. Outros métodos identificados foram testes de desempenho e testes A/B. Estes conseguem ser mais eficientes na objetividade, porém o número de fatores que podem ser avaliados é mais limitado. As entrevistas e observações dos aplicadores também foram métodos identificados no mapeamento. Eles constituem-se como métodos qualitativos e por isso a sumarização e apresentação de seus dados são mais complexas. Por fim, os logs do sistema permitem avaliar o desempenho dos participantes de forma mais objetiva e sem a necessidade de aplicar questionários aos usuários.

No sentido de aplicar os resultados obtidos no mapeamento foi desenvolvida a Plataforma Euklides, uma plataforma web que permite que professores criem salas virtuais, nas quais é possível adicionar jogos e compartilhar o link para cadastro dos alunos. Para o professor, também é possível verificar o desempenho geral dos alunos bem como o desempenho detalhado, isto é, resultados de cada tentativa de cada aluno.

Foram também desenvolvidos os dois jogos que foram inseridos na plataforma Euklides: Quiz Game e Axio. O Quiz é um jogo baseado em questionários que permite flexibilidade no sentido de ser possível alterar as perguntas conforme a necessidade. Já o Axio é um jogo de plataforma que possibilita a revisão de expressões matemáticas e lógicas e que permite uma interação maior entre o jogo e o usuário.

A partir da aplicação da plataforma em um experimento com 10 voluntários foi possível avaliar o desempenho dos participantes através de dois métodos avaliativos: questionários e logs do sistema. O questionário aplicado (MEEGA+) permitiu avaliar nove dimensões. Os dados referentes ao desempenho do participante durante o jogo também foram coletados. Esses dados, que incluem tempo jogado, acertos, erros, dicas e outros, permitiram ter uma visão mais ampla e objetiva do desempenho do participante. Além de comparar os dados entre si, como o impacto do tempo jogado, dicas e tentativas no número de acertos do participante, foi possível comparar com as respostas do questionário avaliativo. A partir dessa análise foi possível destacar dados de desempenho que se relacionam com certas dimensões do questionário.

7.1 Contribuições

O Mapeamento Sistemático conduzido neste trabalho permitiu identificar as características dos métodos de avaliação de jogos sérios, destacando suas vantagens e desvantagens, bem como a forma como são aplicados e analisados. Espera-se que os achados do Mapeamento possam ser úteis tanto para desenvolvedores de modelos avaliativos como para pesquisadores que busquem um método ou modelo de avaliação para aplicar.

Com relação ao experimento realizado, como o mesmo teve todas as suas etapas documentadas, é possível reproduzir o mesmo modelo de avaliação utilizado em outros trabalhos relacionados a jogos sérios. A dinâmica da avaliação neste trabalho consistiu, principalmente, em uma análise dos logs coletados dos usuários durante os jogos e aplicação do modelo MEEGA+. Para a reprodução dessa avaliação, é possível seguir as etapas descritas, utilizando os mesmos cálculos realizados, tanto na análise dos logs, como na comparação entre logs e respostas do MEEGA+.

A pesquisa realizada neste trabalho gerou publicações em duas revistas. O primeiro trabalho (LIMA et al., 2020b), intitulado "*Desenvolvimento e Aplicação de Jogos Sérios para o Ensino de Cinética Química*", foi publicado na revista *Research, Society and Development*. O segundo trabalho (LIMA et al., 2020a), intitulado "Avaliando um Jogo Educacional para o Ensino de Inteligência Artificial - Qual Metodologia para Avaliação Escolher?", foi apresentado no XXVIII Workshop sobre Educação em Computação e consta, atualmente, nos anais do evento. Um artigo focado na condução do Mapeamento Sistemático está em fase de tradução.

7.2 Limitações

Além das limitações do Mapeamento Sistemático, já citadas, foram identificadas outras limitações da pesquisa como um todo. Primeiramente, não foi possível aplicar o experimento no público-alvo original, que seria alunos de uma turma de Fundamentos da Computação. Por conta da pandemia do coronavírus, o calendário acadêmico sofreu alterações significativas e não foi possível aplicar os jogos com a turma, na qual os alunos estariam, teoricamente, com o nível de conhecimento nivelado. Além disso, encontros presenciais foram suspensos, o que impediu a aplicação do experimento presencialmente. Como o experimento foi aplicado de forma remota, não foi possível controlar o ambiente dos participantes e observar seus comportamentos durante o experimento.

Ademais, é importante destacar que não foram comparadas todas as dimensões do questionário, tendo em vista a impossibilidade de dividir as respostas em grupos diferentes pelo fato de algumas serem muito semelhantes. Além disso, na comparação entre os logs de sistema e as respostas do questionário de avaliação percebeu-se que não houveram

diferenças significativas entre as dimensões. Um dos motivos para tal resultado pode se dar pelo número pequeno de participantes da amostra.

7.3 Trabalhos futuros

Diante dos resultados, alguns trabalhos futuros podem ser elencados. Inicialmente, a aplicação da plataforma em uma amostra maior de pessoas, de preferência, em sala de aula. Além disso, seria importante a aplicação não só dos questionários e dos logs, como também dos outros métodos de avaliação encontrados na pesquisa. Essa aplicação permitiria a realização de uma análise mais completa do comportamento dos métodos de avaliação. É também importante, a adoção de técnicas de inteligência artificial na análise dos logs de sistema, visto que, como mencionado no trabalho, mostram ter um potencial relevante na avaliação de jogos sérios. Por fim, um outro trabalho futuro seria a inserção dos cálculos realizados na análise dos logs de sistema diretamente na plataforma para facilitar o papel do avaliador do jogo.

APÊNDICE A – Manual do usuário - Plataforma Euklides

Neste apêndice é apresentado o manual do usuário da plataforma Euklides. A página inicial pode ser vista na Figura 47. Nessa página, o usuário pode fazer login e caso seja um professor, ele poderá se cadastrar. O aluno só poderá se cadastrar através de um link fornecido pelo professor.

Figura 47 – Login/Cadastro.

A imagem mostra duas telas da interface da plataforma Euklides. A primeira tela, no topo, apresenta o cabeçalho 'EUKLIDES' e dois botões: 'Entrar' e 'Criar Conta'. Um rótulo 'Fazer login' com uma seta vermelha aponta para o botão 'Entrar'. Abaixo dos botões, há um formulário de login com campos para 'E-mail' (contendo 'joão@gmail.com') e 'Senha' (com pontos para ocultar o texto), e um botão 'Entrar' na base. A segunda tela, na parte inferior, também possui o cabeçalho 'EUKLIDES' e os botões 'Entrar' e 'Criar Conta'. Um rótulo 'Fazer cadastro' com uma seta vermelha aponta para o botão 'Criar Conta'. Abaixo, há um formulário de cadastro com campos para 'Nome' (contendo 'João Oliveira'), 'E-mail' (contendo 'joão@gmail.com'), 'Idade' (contendo '25'), 'Sexo' (com opções 'Feminino' e 'Masculino', sendo 'Masculino' selecionado), 'Senha' e 'Confirmar a senha' (ambos com pontos para ocultar o texto), e um botão 'Cadastrar' na base.

Caso um professor tenha feito o login, é mostrada a tela inicial do professor, conforme Figura 48. Nessa página são mostradas as salas cadastradas pelo professor e opção para que ele cadastre uma nova sala. Caso o usuário selecione a opção de criar uma nova sala, será aberta uma *dialog* apresentada na Figura 49. Na *dialog*, o professor deverá informar o nome da sala, a disciplina relacionada, uma descrição (opcional) e quais objetos de aprendizagem, no caso, quais jogos ele irá adicionar.

Figura 48 – Página inicial do professor.

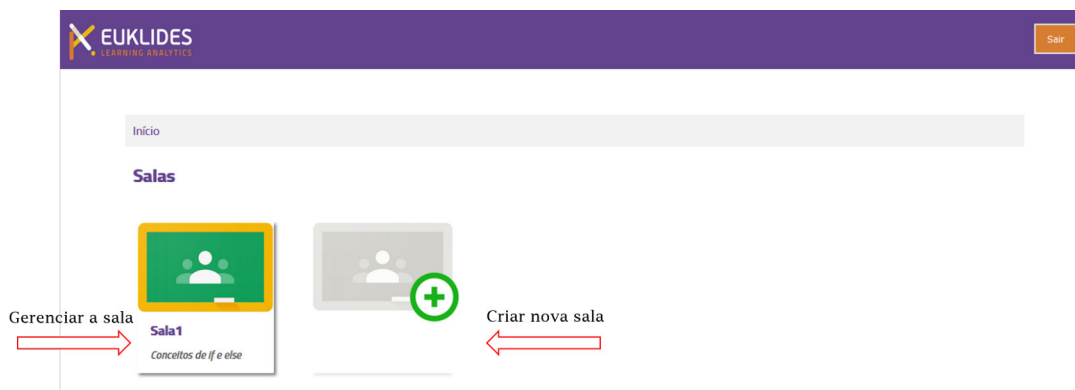
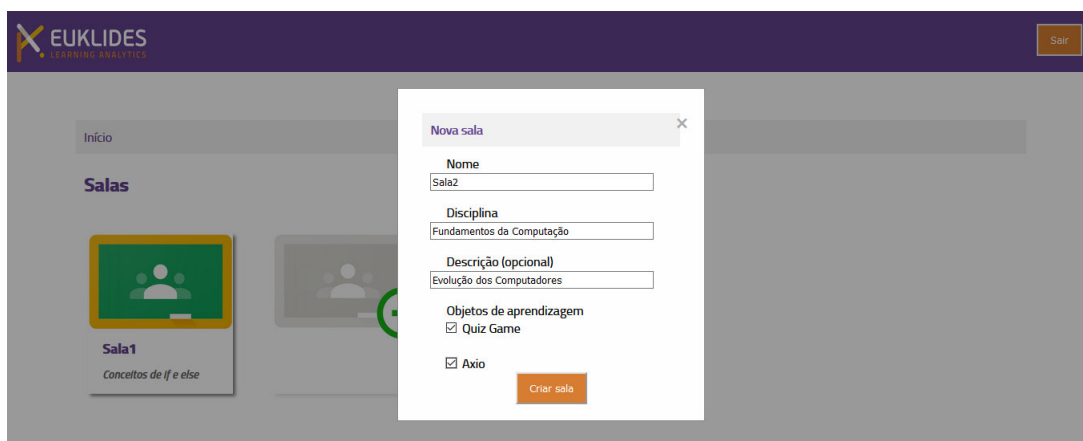


Figura 49 – Criação de uma nova sala.



Ao selecionar uma de suas salas, o professor visualiza a tela de gerenciamento da sala (Figura 50). Nessa página, é fornecido ao professor o link de cadastro da sala, que ao ser encaminhado aos alunos, permitirão que os mesmos sejam direcionados para a tela de cadastro e automaticamente inseridos na sala em questão. Há também três abas nessa tela: Jogos, Alunos e Gerenciar. Na aba de Jogos, é possível excluir os jogos existentes ou adicionar novos. Na aba Alunos (Figura 51), há uma visualização geral de todos os alunos presentes na sala, onde são mostrados os dados referentes a melhor partida do aluno, ou seja, a partida em que obteve a maior pontuação e em menor tempo. Ao clicar no nome do aluno, é mostrado o desempenho dele de forma detalhada (Figura 52). Por fim, na aba Gerenciar (Figura 53), o professor pode editar a sala, alterando o nome da sala, nome da disciplina ou a descrição. Além disso, também é possível excluir ou adicionar jogos à sala.

Figura 50 – Gerenciamento da sala - Aba Jogos.



Figura 51 – Gerenciamento da sala - Aba Alunos.

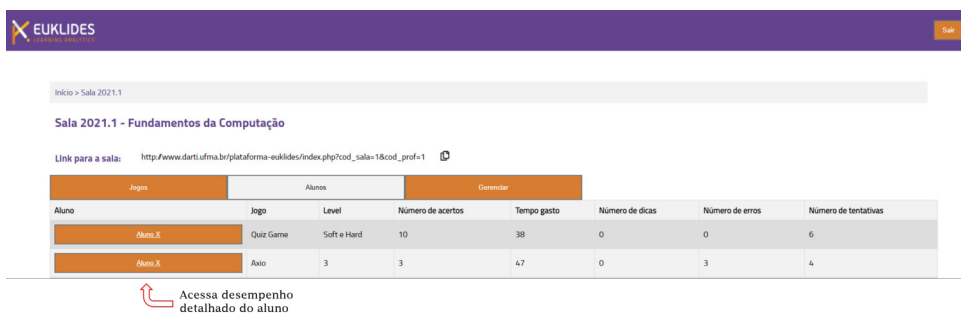


Figura 52 – Gerenciamento da sala - Aba Alunos - Desempenho detalhado do Aluno.

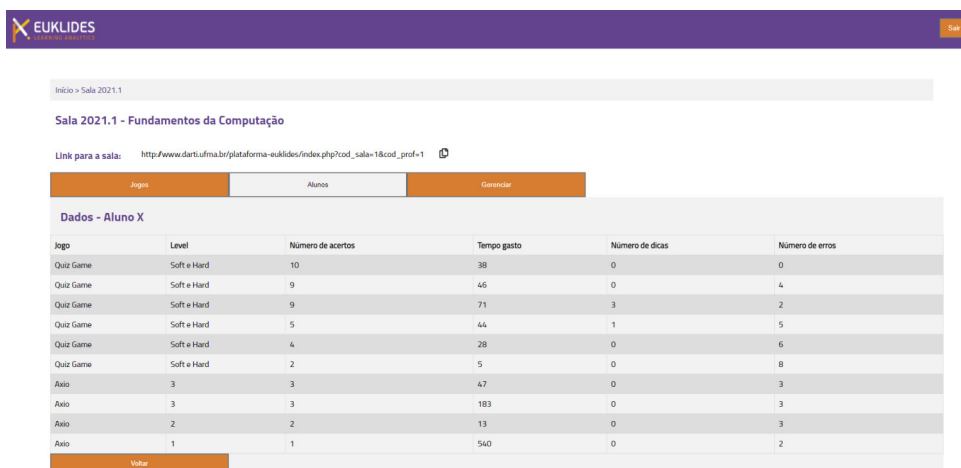
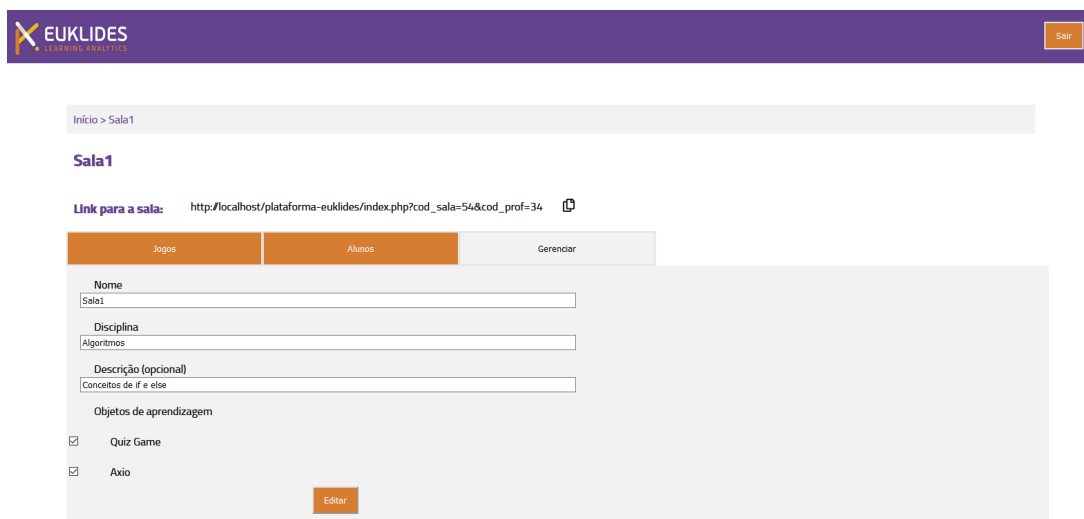


Figura 53 – Gerenciamento da sala - Aba Gerenciar.



A página inicial do aluno (Figura 54) é diferente, visto que o mesmo só pode fazer parte de uma sala, portanto, sua página apresenta os jogos da sala em que está cadastrado. É também disponibilizado um botão que permite ao aluno acessar um questionário para avaliação do jogo, ilustrado na Figura 55. O questionário utilizado foi o MEEGA+.

Figura 54 – Página inicial do aluno.

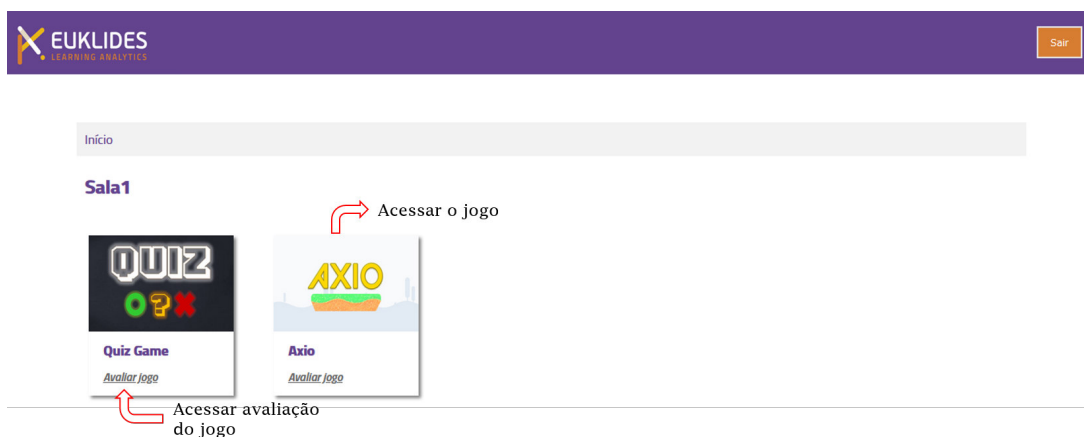


Figura 55 – Questionário para avaliação do jogo.

EUKLIDES
LEARNING ANALYTICS

Sair

Avaliação Quiz Game

Você está sendo convidado(a) a responder às perguntas deste questionário de forma voluntária. Ao responder este questionário você autoriza que as respostas sejam utilizadas em uma pesquisa acadêmica relacionada a avaliação de jogos sérios.

1 - Com que frequência você costuma jogar jogos digitais?

- Nunca: nunca jogo.
- Raramente: jogo de tempos em tempos.
- Mensalmente: jogo pelo menos uma vez por mês.
- Semanalmente: jogo pelo menos uma vez por semana.
- Diariamente: jogo todos os dias.

2 - Faixa etária:

Os dois jogos disponíveis na plataforma são o Quiz Game e o Axio.

A.1 Quiz Game

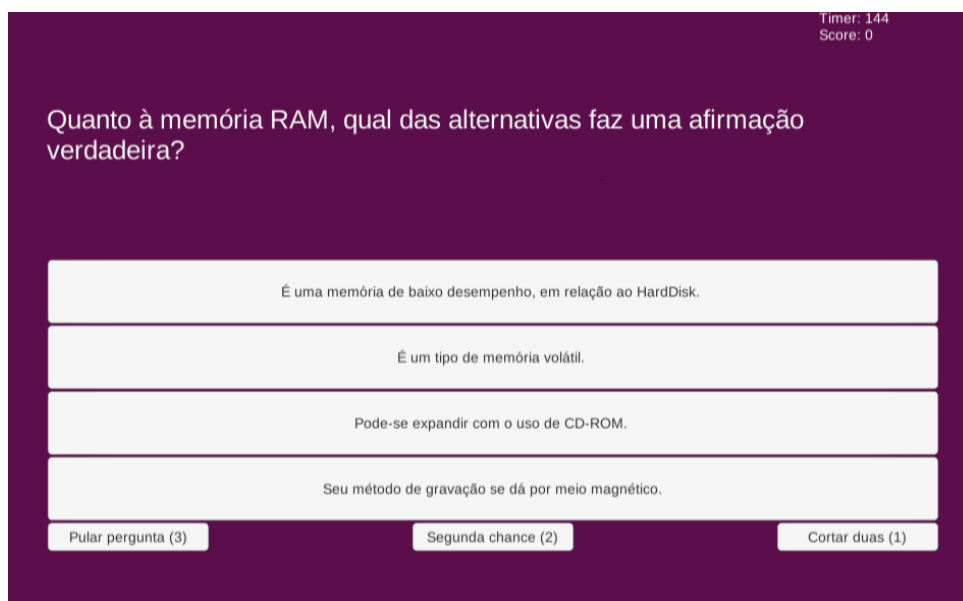
No Quiz Game, foram atribuídas perguntas de computação básica. É possível utilizar perguntas de outros temas alterando o arquivo JSON em que elas estão inseridas.

O menu principal do Quiz pode ser observado na Figura 56. O aluno pode escolher entre três níveis: Software e Hardware; História e Arquitetura dos computadores; Programação em python. Já na Figura 57, é mostrado como as perguntas são apresentadas aos alunos. Conforme a imagem, o aluno pode utilizar três opções: Pular pergunta, Segunda chance ou Cortar duas. No Pular a pergunta, que pode ser usada três vezes, o aluno será redirecionado a uma outra pergunta, sem contabilizar erro. Na Segunda chance, que pode ser usada duas vezes, o usuário pode selecionar a opção e, se estiver errada, ele receberá uma outra chance. No Cortar duas, serão excluídas duas das possíveis respostas, ou seja, o usuário poderá escolher entre apenas duas opções.

Figura 56 – Quiz Game - Menu principal.

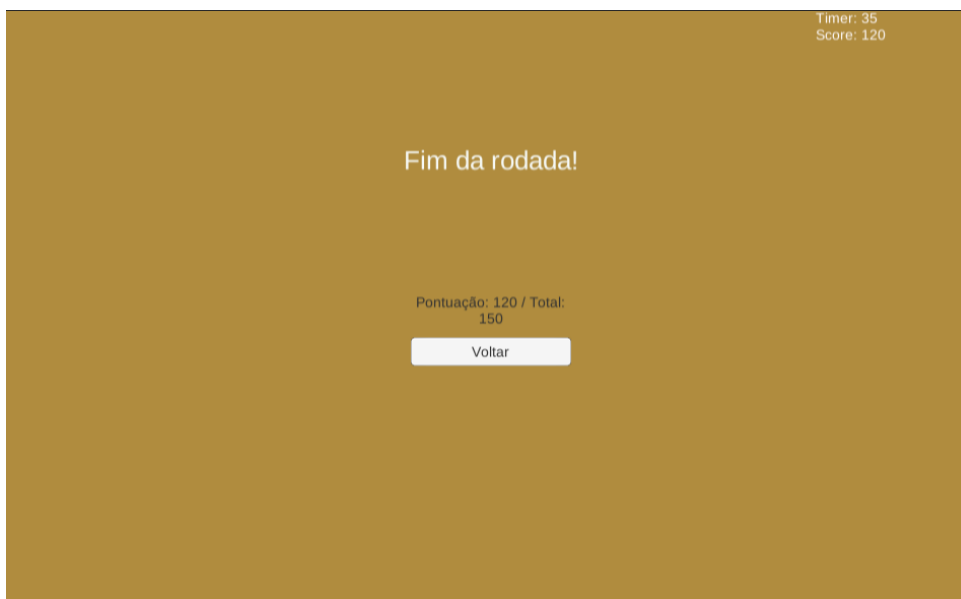


Figura 57 – Quiz Game - Tela de perguntas.



O quiz é finalizado quando o jogador responder 10 perguntas ou quando o tempo acabar. O tempo atribuído para cada partida do quiz foi de 150 segundos. No fim do jogo é apresentada a tela da Figura 58. Nela é mostrada a pontuação final do participante.

Figura 58 – Quiz Game - Tela de fim de jogo.



A.2 Axio

A tela inicial do Axio é mostrada na Figura 59. O jogador pode escolher entre duas opções: Novo Jogo e Como Jogar. A tela Como Jogar (Figura 60) apresenta o objetivo do jogo, os comandos para movimentar o personagem e informa que apenas as plataformas com axiomas verdadeiros são seguras para pular, ou seja, somente essas plataformas irão sustentar o personagem. Já na opção Novo Jogo, o jogo é iniciado.

Figura 59 – Axio - Menu principal.



Figura 60 – Axio - Como Jogar.



A Figura 61 mostra o início do jogo. O personagem inicia no chão e é mostrado seu número de vidas no canto superior esquerdo, o tempo para concluir a fase no canto superior direito e o nível atual no canto inferior direito. Nota-se ainda que cada plataforma possui uma expressão. O jogador deve identificar as plataformas com expressões verdadeiras e as com expressões falsas. As com expressões falsas não irão sustentar o personagem e o mesmo irá cair. Se o personagem tocar no chão, perde uma vida.

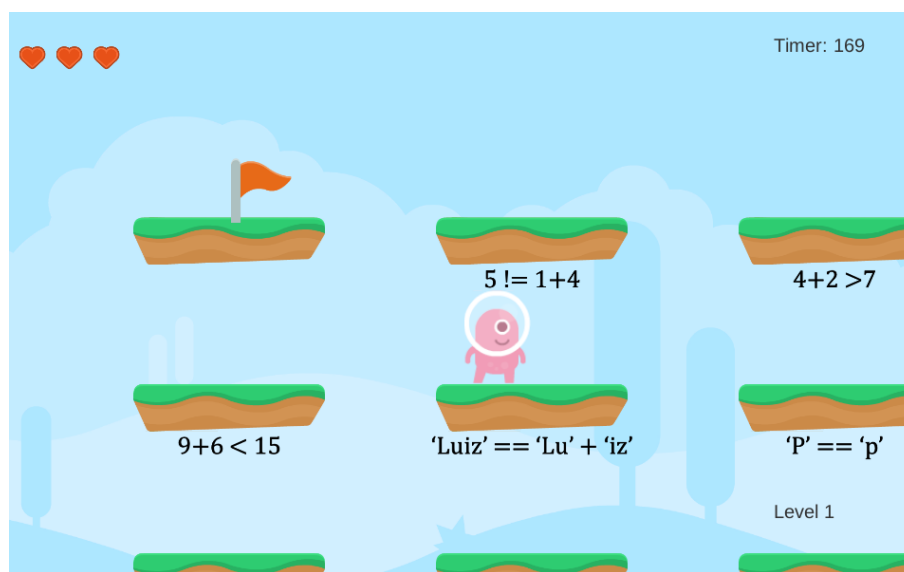
Figura 61 – Axio - Jogo iniciado.



O jogo apresenta quatro níveis que vão aumentando as dificuldades das expressões

matemáticas conforme o personagem avança. Quando o personagem alcança a bandeira (Figura 62), o mesmo é automaticamente redirecionado para o próximo nível. Caso o jogador atinja a bandeira e já esteja na última fase (nível 4), o jogo termina. O jogo também termina se o participante perder suas três vidas. O jogador também deve estar atento ao tempo. O tempo atribuído para concluir cada nível foi de 180 segundos, caso esgotasse esse tempo, o participante também perdia uma vida e o cronômetro era reiniciado.

Figura 62 – Axio - Objetivo do nível.



APÊNDICE B – Questionário do Experimento

Neste apêndice é apresentado o questionário do modelo MEEGA+. Para todas as perguntas, as opções de resposta eram: discordo completamente, discordo, neutro, concordo e concordo plenamente.

Tabela 19 – Questionário elaborado por Petri, von Wangenheim e Borgatto (2017).

Dimensão	Pergunta
Usabilidade	1. O design do jogo é atraente (tabuleiro, cartas, interface, gráficos, etc.). 2. Os textos, cores e fontes combinam e são consistentes. 3. Eu precisei aprender poucas coisas para poder começar a jogar o jogo. 4. Aprender a jogar este jogo foi fácil para mim. 5. Eu acho que a maioria das pessoas aprenderiam a jogar este jogo rapidamente. 6. Eu considero que o jogo é fácil de jogar. 7. As regras do jogo são claras e compreensíveis. 8. As fontes (tamanho e estilo) utilizadas no jogo são legíveis. 9. As cores utilizadas no jogo são compreensíveis.
Confiança	10. A organização do conteúdo me ajudou a estar confiante de que eu iria aprender com este jogo.
Desafio	11. Este jogo é adequadamente desafiador para mim. 12. O jogo oferece novos desafios (oferece novos obstáculos, situações ou variações) com um ritmo adequado. 13. O jogo não se torna monótono nas suas tarefas (repetitivo ou com tarefas chatas).
Satisfação	14. Completar as tarefas do jogo me deu um sentimento de realização. 15. É devido ao meu esforço pessoal que eu consigo avançar no jogo. 16. Me sinto satisfeito com as coisas que aprendi no jogo. 17. Eu recomendaria este jogo para meus colegas.

Tabela 20 – Continuação - Questionário elaborado por Petri, von Wangenheim e Borgatto (2017).

Dimensão	Pergunta
Interação social	18. Eu pude interagir com outras pessoas durante o jogo. 19. O jogo promove momentos de cooperação e/ou competição entre os jogadores. 20. Eu me senti bem interagindo com outras pessoas durante o jogo.
Diversão	21. Eu me diverti com o jogo. 22. Aconteceu alguma situação durante o jogo (elementos do jogo, competição, etc.) que me fez sorrir.
Atenção focada	23. Houve algo interessante no início do jogo que capturou minha atenção. 24. Eu estava tão envolvido no jogo que eu perdi a noção do tempo. 25. Eu esqueci sobre o ambiente ao meu redor enquanto jogava este jogo.
Relevância	26. O conteúdo do jogo é relevante para os meus interesses. 27. É claro para mim como o conteúdo do jogo está relacionado com a disciplina. 28. O jogo é um método de ensino adequado para esta disciplina. 29. Eu prefiro aprender com este jogo do que de outra forma (outro método de ensino).
Percepção de aprendiz.	30. O jogo contribuiu para a minha aprendizagem na disciplina. 31. O jogo foi eficiente para minha aprendizagem, em comparação com outras atividades da disciplina. 32. O jogo contribuiu para revisar alguns dos conceitos gerais da Computação (Quiz) / revisar expressões lógicas e seus operandos (Axio).

APÊNDICE C – Tabelas comparativas Quiz Game x MEEGA+

As tabelas a seguir mostram como os grupos foram divididos e como foram calculadas as médias de cada um. Foram analisados para o Quiz Game: tempo total jogado (Figura 63), tempo por partida (Figura 64), número de acertos (Figura 65), número de dicas (Figura 66) e número de tentativas (Figura 67).

Figura 63 – Tempo total x Dimensões MEEGA+

Tempo total x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
B	4 (44,4%)	0	9 (100%)	553	E	0	3 (33,3%)	9 (100%)	606
I	3 (33,3%)	0	9 (100%)	349	C	0	0	9 (100%)	150
A	2 (22,2%)	0	9 (100%)	301	D	0	0	9 (100%)	527
F	2 (22,2%)	0	9 (100%)	513	G	0	0	9 (100%)	360
J	0	5 (55,6%)	9 (100%)	150	H	0	0	9 (100%)	293
Média:				373,2	Média:				387,2

Tempo total x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
F	0	2 (66,7%)	3 (100%)	513	A	0	1 (33,3%)	3 (100%)	301
D	0	3 (100%)	3 (100%)	527	B	0	1 (33,3%)	3 (100%)	553
I	1 (33,3%)	0	3 (100%)	349	C	0	0	3 (100%)	150
G	1 (33,3%)	1 (33,3%)	3 (100%)	360	E	0	0	3 (100%)	606
J	2 (66,7%)	1 (33,3%)	3 (100%)	150	H	0	0	3 (100%)	293
Média:				379,8	Média:				380,6

Tempo total x Diversão

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
D	1 (50%)	1 (50%)	2 (100%)	527	G	0	1 (50%)	2 (100%)	360
J	1 (50%)	1 (50%)	2 (100%)	150	I	0	1 (50%)	2 (100%)	349
A	1 (50%)	0	2 (100%)	301	C	0	0	2 (100%)	150
B	1 (50%)	0	2 (100%)	553	E	0	0	2 (100%)	606
Média:				382,75	F	0	0	2 (100%)	513
					H	0	0	2 (100%)	293
					Média:				378,5

Tempo total x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
G	1 (33,3%)	1 (33,3%)	3 (100%)	360	A	0	1 (33,3%)	3 (100%)	301
I	1 (33,3%)	1 (33,3%)	3 (100%)	349	D	0	1 (33,3%)	3 (100%)	527
J	2 (66,7%)	1 (33,3%)	3 (100%)	150	E	0	1 (33,3%)	3 (100%)	606
Média:				286,33	F	0	1 (33,3%)	3 (100%)	513
					B	0	0	3 (100%)	553
					C	0	0	3 (100%)	150
					H	0	0	3 (100%)	293
					Média:				420,43

Figura 64 – Tempo por partida x Dimensões MEEGA+

Tempo por jogada x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
B	4 (44,4%)	0	9 (100%)	110,6	E	0	3 (33,3%)	9 (100%)	101
I	3 (33,3%)	0	9 (100%)	116,63	C	0	0	9 (100%)	150
A	2 (22,2%)	0	9 (100%)	100,33	D	0	0	9 (100%)	125,67
F	2 (22,2%)	0	9 (100%)	102,6	G	0	0	9 (100%)	120
J	0	5 (55,6%)	9 (100%)	150	H	0	0	9 (100%)	97,67
Média:				116,03	Média:				118,87

Tempo por jogada x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
F	0	2 (66,7%)	3 (100%)	102,6	A	0	1 (33,3%)	3 (100%)	100,33
D	0	3 (100%)	3 (100%)	125,67	B	0	1 (33,3%)	3 (100%)	110,6
I	1 (33,3%)	0	3 (100%)	116,63	C	0	0	3 (100%)	150
G	1 (33,3%)	1 (33,3%)	3 (100%)	120	E	0	0	3 (100%)	101
J	2 (66,7%)	1 (33,3%)	3 (100%)	150	H	0	0	3 (100%)	97,67
Média:				122,98	Média:				111,92

Tempo por jogada x Diversão

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
D	1 (50%)	1 (50%)	2 (100%)	125,67	G	0	1 (50%)	2 (100%)	120
J	1 (50%)	1 (50%)	2 (100%)	150	I	0	1 (50%)	2 (100%)	116,63
A	1 (50%)	0	2 (100%)	100,33	C	0	0	2 (100%)	150
B	1 (50%)	0	2 (100%)	110,6	E	0	0	2 (100%)	101
Média:				121,65	F	0	0	2 (100%)	102,6
					H	0	0	2 (100%)	97,67
					Média:				114,65

Tempo por jogada x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
G	1 (33,3%)	1 (33,3%)	3 (100%)	120	A	0	1 (33,3%)	3 (100%)	100,33
I	1 (33,3%)	1 (33,3%)	3 (100%)	116,63	D	0	1 (33,3%)	3 (100%)	125,67
J	2 (66,7%)	1 (33,3%)	3 (100%)	150	E	0	1 (33,3%)	3 (100%)	101
Média:				128,88	F	0	1 (33,3%)	3 (100%)	102,6
					B	0	0	3 (100%)	110,6
					C	0	0	3 (100%)	150
					H	0	0	3 (100%)	97,67
					Média:				112,55

Figura 65 – Número de acertos x Dimensões MEEGA+

Número de acertos x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Número de acertos
B	4 (44,4%)	0	9 (100%)	6,2
I	3 (33,3%)	0	9 (100%)	8
A	2 (22,2%)	0	9 (100%)	9,67
F	2 (22,2%)	0	9 (100%)	8,2
J	0	5 (55,6%)	9 (100%)	4
Média:				7,21

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Número de acertos
E	0	3 (33,3%)	9 (100%)	8,83
C	0	0	9 (100%)	8
D	0	0	9 (100%)	8,5
G	0	0	9 (100%)	8,67
H	0	0	9 (100%)	9
Média:				8,6

Número de acertos x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Número de acertos
F	0	2 (66,7%)	3 (100%)	8,2
D	0	3 (100%)	3 (100%)	8,5
I	1 (33,3%)	0	3 (100%)	8
G	1 (33,3%)	1 (33,3%)	3 (100%)	8,67
J	2 (66,7%)	1 (33,3%)	3 (100%)	4
Média:				7,47

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Número de acertos
A	0	1 (33,3%)	3 (100%)	9,67
B	0	1 (33,3%)	3 (100%)	6,2
C	0	0	3 (100%)	8
E	0	0	3 (100%)	8,83
H	0	0	3 (100%)	9
Média:				8,34

Número de acertos x Diversão

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
D	1 (50%)	1 (50%)	2 (100%)	8,5
J	1 (50%)	1 (50%)	2 (100%)	4
A	1 (50%)	0	2 (100%)	9,67
B	1 (50%)	0	2 (100%)	6,2
Média:				7,09

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
G	0	1 (50%)	2 (100%)	8,67
I	0	1 (50%)	2 (100%)	8
C	0	0	2 (100%)	8
E	0	0	2 (100%)	8,83
F	0	0	2 (100%)	8,2
H	0	0	2 (100%)	9
Média:				8,45

Número de acertos x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
G	1 (33,3%)	1 (33,3%)	3 (100%)	8,67
I	1 (33,3%)	1 (33,3%)	3 (100%)	8
J	2 (66,7%)	1 (33,3%)	3 (100%)	4
Média:				6,89

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo jogado
A	0	1 (33,3%)	3 (100%)	9,67
D	0	1 (33,3%)	3 (100%)	8,5
E	0	1 (33,3%)	3 (100%)	8,83
F	0	1 (33,3%)	3 (100%)	8,2
B	0	0	3 (100%)	6,2
C	0	0	3 (100%)	8
H	0	0	3 (100%)	9
Média:				8,34

Figura 66 – Número de dicas x Dimensões MEEGA+

Dicas utilizadas x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas
B	4 (44,4%)	0	9 (100%)	1,4	E	0	3 (33,3%)	9 (100%)	0,67
I	3 (33,3%)	0	9 (100%)	1	C	0	0	9 (100%)	0
A	2 (22,2%)	0	9 (100%)	1,33	D	0	0	9 (100%)	0
F	2 (22,2%)	0	9 (100%)	1,4	G	0	0	9 (100%)	1
J	0	5 (55,6%)	9 (100%)	0	H	0	0	9 (100%)	0
Média:				1,03	Média:				0,33

Dicas utilizadas x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas
F	0	2 (66,7%)	3 (100%)	1,4	A	0	1 (33,3%)	3 (100%)	1,33
D	0	3 (100%)	3 (100%)	0	B	0	1 (33,3%)	3 (100%)	1,4
I	1 (33,3%)	0	3 (100%)	2	C	0	0	3 (100%)	0
G	1 (33,3%)	1 (33,3%)	3 (100%)	1	E	0	0	3 (100%)	0,67
J	2 (66,7%)	1 (33,3%)	3 (100%)	0	H	0	0	3 (100%)	0
Média:				0,88	Média:				0,68

Dicas utilizadas x Diversão

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas
D	1 (50%)	1 (50%)	2 (100%)	0	G	0	1 (50%)	2 (100%)	1
J	1 (50%)	1 (50%)	2 (100%)	0	I	0	1 (50%)	2 (100%)	2
A	1 (50%)	0	2 (100%)	1,33	C	0	0	2 (100%)	0
B	1 (50%)	0	2 (100%)	1,4	E	0	0	2 (100%)	0,67
Média:				0,6825	F	0	0	2 (100%)	1,4
					H	0	0	2 (100%)	0
					Média:				0,845

Dicas utilizadas x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas	Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Dicas utilizadas
G	1 (33,3%)	1 (33,3%)	3 (100%)	1	A	0	1 (33,3%)	3 (100%)	1,33
I	1 (33,3%)	1 (33,3%)	3 (100%)	2	D	0	1 (33,3%)	3 (100%)	0
J	2 (66,7%)	1 (33,3%)	3 (100%)	0	E	0	1 (33,3%)	3 (100%)	0,67
Média:				1	F	0	1 (33,3%)	3 (100%)	1,4
					B	0	0	3 (100%)	1,4
					C	0	0	3 (100%)	0
					H	0	0	3 (100%)	0
					Média:				0,686

Figura 67 – Tentativas x Dimensões MEEGA+

Tentativas x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
B	4 (44,4%)	0	9 (100%)	5
I	3 (33,3%)	0	9 (100%)	3
A	2 (22,2%)	0	9 (100%)	3
F	2 (22,2%)	0	9 (100%)	6
J	0	5 (55,6%)	9 (100%)	1
Média:				3,6

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
E	0	3 (33,3%)	9 (100%)	7
C	0	0	9 (100%)	1
D	0	0	9 (100%)	3
G	0	0	9 (100%)	3
H	0	0	9 (100%)	3
Média:				3,4

Tentativas x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
F	0	2 (66,7%)	3 (100%)	6
D	0	3 (100%)	3 (100%)	3
I	1 (33,3%)	0	3 (100%)	3
G	1 (33,3%)	1 (33,3%)	3 (100%)	3
J	2 (66,7%)	1 (33,3%)	3 (100%)	1
Média:				3,2

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
A	0	1 (33,3%)	3 (100%)	3
B	0	1 (33,3%)	3 (100%)	5
C	0	0	3 (100%)	1
E	0	0	3 (100%)	7
H	0	0	3 (100%)	3
Média:				3,8

Tentativas x Diversão

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
D	1 (50%)	1 (50%)	2 (100%)	3
J	1 (50%)	1 (50%)	2 (100%)	1
A	1 (50%)	0	2 (100%)	3
B	1 (50%)	0	2 (100%)	5
Média:				3

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
G	0	1 (50%)	2 (100%)	3
I	0	1 (50%)	2 (100%)	3
C	0	0	2 (100%)	1
E	0	0	2 (100%)	7
F	0	0	2 (100%)	6
H	0	0	2 (100%)	3
Média:				3,83

Tentativas x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
G	1 (33,3%)	1 (33,3%)	3 (100%)	3
I	1 (33,3%)	1 (33,3%)	3 (100%)	3
J	2 (66,7%)	1 (33,3%)	3 (100%)	1
Média:				2,33

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
A	0	1 (33,3%)	3 (100%)	3
D	0	1 (33,3%)	3 (100%)	3
E	0	1 (33,3%)	3 (100%)	7
F	0	1 (33,3%)	3 (100%)	6
B	0	0	3 (100%)	5
C	0	0	3 (100%)	1
H	0	0	3 (100%)	3
Média:				4

APÊNDICE D – Tabelas comparativas Axio x MEEGA+

As tabelas a seguir mostram como os grupos foram divididos e como foram calculadas as médias de cada um. Foram analisado para o Axio: tempo por jogada (Figura 68), tempo total (Figura 69) níveis alcançados (Figura 70), número de erros/vidas perdidas (Figura 71), e número de tentativas (Figura 72).

Figura 68 – Tempo por jogada x Dimensões MEEGA+

Tempo por jogada x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo por jogada
B	3 (33,3%)	0	9 (100%)	122
I	2 (22,2%)	0	9 (100%)	180
Média:				151

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo por jogada
E	0	2 (22,2%)	9 (100%)	306,33
F	1 (11,1%)	0	9 (100%)	193
A	0	0	9 (100%)	241
C	0	0	9 (100%)	53,75
D	0	0	9 (100%)	144,5
G	0	0	9 (100%)	165,67
H	0	0	9 (100%)	195
Média:				185,61

Tempo por jogada x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo por jogada
B	1 (25%)	0	4 (100%)	122
F	1 (25%)	0	4 (100%)	193
G	1 (25%)	0	4 (100%)	165,67
I	1 (25%)	0	4 (100%)	180
Média:				165,17

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo por jogada
E	0	2 (22,2%)	4 (100%)	306,33
H	0	0	4 (100%)	195
A	0	0	4 (100%)	241
C	0	0	4 (100%)	53,75
D	0	0	4 (100%)	144,5
Média:				188,12

Tempo por jogada x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo por jogada
D	0	2 (66,7%)	3 (100%)	144,5
F	0	2 (66,7%)	3 (100%)	193
G	0	2 (66,7%)	3 (100%)	165,67
Média:				167,72

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Leveis
A	0	0	3 (100%)	241
B	0	1 (33,3%)	3 (100%)	122
C	0	0	3 (100%)	53,75
E	0	1 (33,3%)	3 (100%)	306,33
H	0	0	3 (100%)	195
I	0	1 (33,3%)	3 (100%)	180
Média:				183,01

Figura 69 – Tempo total x Dimensões MEEGA+

Tempo total x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
B	3 (33,3%)	0	9 (100%)	244
I	2 (22,2%)	0	9 (100%)	180
Média:				216

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
E	0	2 (22,2%)	9 (100%)	919
F	1 (11,1%)	0	9 (100%)	193
A	0	0	9 (100%)	241
C	0	0	9 (100%)	430
D	0	0	9 (100%)	289
G	0	0	9 (100%)	690
H	0	0	9 (100%)	585
Média:				478,14

Tempo total x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
B	1 (25%)	0	4 (100%)	244
F	1 (25%)	0	4 (100%)	193
G	1 (25%)	0	4 (100%)	690
I	1 (25%)	0	4 (100%)	180
Média:				326,75

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
E	0	2 (22,2%)	4 (100%)	919
H	0	0	4 (100%)	585
A	0	0	4 (100%)	241
C	0	0	4 (100%)	430
D	0	0	4 (100%)	289
Média:				492,8

Tempo total x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
D	0	2 (66,7%)	3 (100%)	289
F	0	2 (66,7%)	3 (100%)	193
G	0	2 (66,7%)	3 (100%)	690
Média:				390,67

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tempo total
A	0	0	3 (100%)	241
B	0	1 (33,3%)	3 (100%)	244
C	0	0	3 (100%)	430
E	0	1 (33,3%)	3 (100%)	919
H	0	0	3 (100%)	585
I	0	1 (33,3%)	3 (100%)	180
Média:				433,17

Figura 70 – Níveis alcançados x Dimensões MEEGA+

Níveis alcançados x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Níveis
B	3 (33,3%)	0	9 (100%)	4
I	2 (22,2%)	0	9 (100%)	1
Média:				2,5

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Níveis
E	0	2 (22,2%)	9 (100%)	4
F	1 (11,1%)	0	9 (100%)	4
A	0	0	9 (100%)	4
C	0	0	9 (100%)	0
D	0	0	9 (100%)	4
G	0	0	9 (100%)	4
H	0	0	9 (100%)	3
Média:				3,29

Níveis alcançados x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Níveis
B	1 (25%)	0	4 (100%)	4
F	1 (25%)	0	4 (100%)	4
G	1 (25%)	0	4 (100%)	4
I	1 (25%)	0	4 (100%)	1
Média:				3,25

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Níveis
E	0	2 (22,2%)	4 (100%)	4
H	0	0	4 (100%)	3
A	0	0	4 (100%)	4
C	0	0	4 (100%)	0
D	0	0	4 (100%)	4
Média:				3

Níveis alcançados x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Níveis
D	0	2 (66,7%)	3 (100%)	4
F	0	2 (66,7%)	3 (100%)	4
G	0	2 (66,7%)	3 (100%)	4
Média:				4

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Níveis
A	0	0	3 (100%)	4
B	0	1 (33,3%)	3 (100%)	4
C	0	0	3 (100%)	0
E	0	1 (33,3%)	3 (100%)	4
H	0	0	3 (100%)	3
I	0	1 (33,3%)	3 (100%)	1
Média:				2,67

Figura 71 – Número de erros x Dimensões MEEGA+

Erros x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Erros
B	3 (33,3%)	0	9 (100%)	1,33
I	2 (22,2%)	0	9 (100%)	3
Média:				2,165

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Erros
E	0	2 (22,2%)	9 (100%)	2,67
F	1 (11,1%)	0	9 (100%)	0
A	0	0	9 (100%)	0
C	0	0	9 (100%)	3
D	0	0	9 (100%)	1,5
G	0	0	9 (100%)	0,67
H	0	0	9 (100%)	2,25
Média:				1,44

Erros x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Erros
B	1 (25%)	0	4 (100%)	1,33
F	1 (25%)	0	4 (100%)	0
G	1 (25%)	0	4 (100%)	0,67
I	1 (25%)	0	4 (100%)	3
Média:				1,25

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Erros
E	0	2 (22,2%)	4 (100%)	2,67
H	0	0	4 (100%)	2,25
A	0	0	4 (100%)	0
C	0	0	4 (100%)	3
D	0	0	4 (100%)	1,5
Média:				1,884

Erros x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Erros
D	0	2 (66,7%)	3 (100%)	1,5
F	0	2 (66,7%)	3 (100%)	0
G	0	2 (66,7%)	3 (100%)	0,67
Média:				0,723

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Erros
A	0	0	3 (100%)	0
B	0	1 (33,3%)	3 (100%)	1,33
C	0	0	3 (100%)	3
E	0	1 (33,3%)	3 (100%)	2,67
H	0	0	3 (100%)	2,25
I	0	1 (33,3%)	3 (100%)	3
Média:				2,041

Figura 72 – Número de tentativas x Dimensões MEEGA+

Tentativas x Usabilidade

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
B	3 (33,3%)	0	9 (100%)	2
I	2 (22,2%)	0	9 (100%)	1
Média:				1,5

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
E	0	2 (22,2%)	9 (100%)	3
F	1 (11,1%)	0	9 (100%)	1
A	0	0	9 (100%)	1
C	0	0	9 (100%)	8
D	0	0	9 (100%)	2
G	0	0	9 (100%)	3
H	0	0	9 (100%)	3
Média:				3

Tentativas x Desafio

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
B	1 (25%)	0	4 (100%)	2
F	1 (25%)	0	4 (100%)	1
G	1 (25%)	0	4 (100%)	3
I	1 (25%)	0	4 (100%)	1
Média:				1,75

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
E	0	2 (22,2%)	4 (100%)	3
H	0	0	4 (100%)	3
A	0	0	4 (100%)	1
C	0	0	4 (100%)	8
D	0	0	4 (100%)	2
Média:				3,40

Tentativas x Atenção focada

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
D	0	2 (66,7%)	3 (100%)	2
F	0	2 (66,7%)	3 (100%)	1
G	0	2 (66,7%)	3 (100%)	3
Média:				2

Alunos	Respostas discordo / discordo totalmente	Respostas nem concordo, nem discordo	Número de perguntas da dimensão	Tentativas
A	0	0	3 (100%)	1
B	0	1 (33,3%)	3 (100%)	2
C	0	0	3 (100%)	8
E	0	1 (33,3%)	3 (100%)	3
H	0	0	3 (100%)	3
I	0	1 (33,3%)	3 (100%)	1
Média:				3

Referências

- AKL, E. A.; PRETORIUS, R. W.; SACKETT, K.; ERDLEY, W. S.; BHOOPATHI, P. S.; ALFARAH, Z.; SCHÜNEMANN, H. J. The effect of educational games on medical students' learning outcomes: A systematic review: Beme guide no 14. *Medical Teacher*, Taylor & Francis, v. 32, n. 1, p. 16–27, 2010. Citado 2 vezes nas páginas 19 e 20.
- ALFARAH, Z.; SCHÜNEMANN, H. J.; AKL, E. A. Educational games in geriatric medicine education: a systematic review. *BMC Geriatrics*, 2010. Citado 2 vezes nas páginas 19 e 20.
- ALVAREZ, J.; DJAOUTI, D. An introduction to serious game definitions and concepts. *Serious Games & Simulation for Risks Management*, p. 11–15, 12 2011. Citado na página 15.
- ALVAREZ, J.; PLANTEC, J.-Y.; VERMEULEN, M.; KOLSKI, C. Rdu model dedicated to evaluate needed counsels for serious game projects. *Computers Education*, Volume 114, p. Pages 38–56, 06 2017. Citado na página 24.
- ALVES, N. A.; ROZA, R. H. Gamificação no ensino de administração: o uso da plataforma kahoot em uma disciplina de gestão de projetos. *Tecnologias na Educação*, v. 25, p. 207–216, 2018. Citado na página 17.
- AMPATZOGLU, A.; BIBI, S.; AVGERIOU, P.; VERBEEK, M.; CHATZIGEORGIOU, A. Identifying, categorizing and mitigating threats to validity in software engineering secondary studies. *Information and Software Technology*, v. 106, p. 201 – 230, 2019. ISSN 0950-5849. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0950584918302106>>. Citado 2 vezes nas páginas 27 e 28.
- ASADIPOUR, A.; DEBATTISTA, K.; CHALMERS, A. Visuohaptic augmented feedback for enhancing motor skills acquisition. *The Visual Computer*, v. 33, p. 401–411, 2016. Citado na página 38.
- BROCKMYER, J. H.; FOX, C. M.; CURTISS, K. A.; MCBROOM, E.; BURKHART, K. M.; PIDRUZNY, J. N. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, v. 45, n. 4, p. 624–634, 2009. ISSN 0022-1031. Citado 2 vezes nas páginas 35 e 37.
- BROOKE, J. *SUS-A quick and dirty usability scale*. [S.l.]: Usability Evaluation in Industry, Taylor & Francis, 1996. 189–194 p. Citado 2 vezes nas páginas 35 e 36.
- BUSSAB, W. de O.; MORETTIN, P. A. *Estatística Básica*. 6. ed. [S.l.]: Saraiva, 2010. Citado na página 53.
- CABALLERO-HERNÁNDEZ, J. A.; PALOMO-DUARTE, M.; DODERO, J. M. Skill assessment in learning experiences based on serious games: A systematic mapping study. *Computers & Education*, v. 113, p. 42–60, 2017. ISSN 0360-1315. Citado na página 19.
- CAKIR, M. P.; ÇAKIR, N. A.; AYAZ, H.; LEE, F. J. An optical brain imaging study on the improvements in mathematical fluency from game-based learning. *Proceedings of the*

- 2015 Annual Symposium on Computer-Human Interaction in Play*, p. 209–219, 2015. Citado na página 38.
- CALDERÓN, A.; RUIZ, M.; O’CONNOR, R. V. A multivocal literature review on serious games for software process standards education. *Computers & Education*, v. 67, p. 156–167, 2013. Citado na página 19.
- Calvo-Morata, A.; Rotaru, D. C.; Alonso-Fernandez, C.; Freire, M.; Martinez-Ortiz, I.; Fernández-Manjón, B. Validation of a cyberbullying serious game using game analytics. *IEEE Transactions on Learning Technologies*, p. 1–1, 2018. ISSN 2372-0050. Citado na página 38.
- CASANO, J.; TEE, H. E. U.; AGAPITO, J.; ARROYO, I.; RODRIGO, M. M. T. Migration and evaluation of a framework for developing embodied cognition learning games. In: *VRCAI ’16*. [S.l.: s.n.], 2016. Citado na página 41.
- CHIN, J. P.; DIEHL, V. A.; NORMAN, K. L. Development of an instrument measuring user satisfaction of the human-computer interface. *CHI ’88 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 213–218, 1988. Citado na página 37.
- CODE.ORG. *Sobre nós*. 2020. <<https://code.org/international/about>>. Citado na página 17.
- COLLINS, A.; HALVERSON, R. The second educational revolution: rethinking education in the age of technology. *Journal of Computer Assisted Learning*, v. 26, p. 18–27, 2010. Citado na página 15.
- DAVIS, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, v. 13, n. 3, p. 319–340, 1989. Citado 2 vezes nas páginas 35 e 36.
- DUNWELL, I.; FREITAS, S. de; PETRIDIS, P.; HENDRIX, M.; ARNAB, S.; LAMERAS, P.; STEWART, C. A game-based learning approach to road safety: The code of everand. *Conference on Human Factors in Computing Systems - Proceedings*, 04 2014. Citado na página 31.
- EAGLE, M. Level up: a frame work for the design and evaluation of educational games. In: . [S.l.: s.n.], 2009. p. 339–341. Citado na página 15.
- EAGLE, M.; BARNES, T. Experimental evaluation of an educational game for improved learning in introductory computing. *SIGCSE ’09 Proceedings of the 40th ACM technical symposium on Computer science education*, p. 321–325, 2009. Citado na página 38.
- ECONOMOU, D.; DOUMANIS, I.; BOUKI, V.; PEDERSEN, F.; MENTZELOPOULOS, M.; GEORGALAS, N. Edu-simulation: A serious games platform designed to engage and motivate students. In: *2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*. [S.l.: s.n.], 2015. p. 244–248. Citado na página 18.
- ERHEL, S.; JAMET, E. Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, v. 67, p. 156–167, 2013. Citado 2 vezes nas páginas 15 e 32.

- ESCUDEIRO, P.; BIDARRA, J. Quantitative evaluation framework (qef). *Conselho Editorial/Consejo Editorial*, p. 16, 01 2008. Citado na página 37.
- FANG, X.; CHAN, S.; BRZEZINSKI, J.; NAIR, C. Development of an instrument to measure enjoyment of computer game play. *International Journal of Human-Computer Interaction*, v. 26, p. 868–886, 2010. Citado na página 37.
- FATTA, H.; MAKSOM, Z.; ZAKARIA, M. Systematic literature review on usability evaluation model of educational games: Playability, pedagogy, and mobility aspects. *Journal of Theoretical and Applied Information Technology*, v. 96, p. 4677–4689, 07 2018. Citado 2 vezes nas páginas 18 e 19.
- FIORI, M.; ROCHA, M.; MARQUES, A. Uma experiência de aprendizagem de lógica de programação com code.org no ensino médio: uma análise por gênero sobre a percepção dos estudantes. In: *Anais do XIII Women in Information Technology*. Porto Alegre, RS, Brasil: SBC, 2019. p. 124–128. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/wit/article/view/6722>>. Citado na página 17.
- FREIRE, M.; SERRANO-LAGUNA, ; MANERO, B.; MARTINEZ-ORTIZ, I.; GER, P. M.; FERNÁNDEZ-MANJÓN, B. Game learning analytics: Learning analytics for serious games. In: _____. [S.l.: s.n.], 2016. p. 1–29. ISBN 978-3-319-17727-4. Citado 2 vezes nas páginas 15 e 38.
- FU, F. L.; SU, R. C.; YU, S. C. Egameflow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, v. 52, p. 101–112, 2009. Citado 4 vezes nas páginas 24, 34, 35 e 36.
- GEORGE, C.; CARLOS, E.; DAVID, M. Serious game for the virtual practice of the emplantillado in the constructive system of adobe with reinforced cane. *ICETC 2017 Proceedings of the 2017 9th International Conference on Education Technology and Computers*, p. 99–103, 2017. Citado na página 38.
- GHALEB, E.; POPA, M.; HORTAL, E.; ASTERIADIS, S.; WEISS, G. Towards affect recognition through interactions with learning materials. In: . [S.l.: s.n.], 2018. p. 372–379. Citado 2 vezes nas páginas 41 e 43.
- GUAY, F.; VALLERAND, R.; BLANCHARD, C. On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (sims). *Motivation and Emotion*, v. 24, p. 175–213, 09 2000. Citado na página 37.
- GUILLÓ, A.; PERTEGAL-FELICES, M.; JIMENO-MORENILLA, A.; AZORIN-LOPEZ, J.; SOLIVERES, M. L. R.; RESTREPO-CALLE, F. Evaluating impact on motivation and academic performance of a game-based learning experience using kahoot. *Frontiers in Psychology*, v. 10, p. 2843, 12 2019. Citado na página 17.
- HOLZINGER, A. Construction and evaluation of a user experience questionnaire. *HCI and usability for education and work*, p. 63–76, 2008. Citado na página 37.
- 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*. New York, NY, USA: ACM, 2019. (ACSW 2019), p. 42:1–42:10. ISBN 978-1-4503-6603-8. Disponível em: <<http://doi.acm.org/10.1145/3290688.3290747>>. Citado na página 19.

- HUSSAAN, A.; SEHABA, K. Adaptive serious game for rehabilitation of persons with cognitive disabilities. In: . [S.l.: s.n.], 2013. p. 65–69. Citado na página 18.
- IJSSELSTEIJN, W.; HOOGEN, W.; KLIMMT, C.; KORT, Y. D.; LINDLEY, C.; MATHIAK, K.; POELS, K.; RAVAJA, N.; TURPEINEN, M.; VORDERER, P. Measuring the experience of digital game enjoyment. *Proceedings of Measuring Behavior*, 01 2008. Citado na página 37.
- JEFFRIES, P. R.; RIZZOLO, M. A. Designing and implementing models for the innovative use of simulation in teaching nursing care of ill adults and children: A national, multi-site, multi-method study. *National League for Nursing and Laerdal*, 2006. Citado na página 37.
- KAHOOT. *What is kahoot!* 2020. <<https://kahoot.com/what-is-kahoot/>>. Citado 2 vezes nas páginas 17 e 50.
- KANG, J.; MIN, L.; QU, W. Using gameplay data to examine learning behavior patterns in a serious game. *Elsevier*, v. 72, p. 757–770, 2016. Citado na página 15.
- KAZIMOGLU, C.; KIERNAN, M.; BACON, L.; MACKINNON, L. A serious game for developing computational thinking and learning introductory computer programming. *Procedia - Social and Behavioral Sciences*, v. 47, p. 1991–1999, 12 2012. Citado na página 24.
- KELLER, J. M. *Motivational Design for Learning and Performance: The Arcs Model Approach*. [S.l.]: Springer US, 2009. Citado na página 37.
- KENNEDY, R. S.; LANE, N. E.; BERBAUM, K. S.; LILIENTHAL, M. G. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, Taylor & Francis, v. 3, n. 3, p. 203–220, 1993. Citado na página 37.
- KHANE, A.; AHMAD, F. H.; MALIK, M. M. Use of digital game based learning and gamification in secondary school science: The effect on student engagement, learning and gender difference. *Education and Information Technologies*, p. 2767–2804, 2017. Citado na página 38.
- KIILI, K. Evaluations of an experiential gaming model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, v. 2, 10 2006. Citado na página 37.
- KIM, M. C.; PARK, T.; LIU, R.; FORTE, A. Understanding learning curves and trajectories in css layout. In: . [S.l.: s.n.], 2019. p. 504–510. Citado 2 vezes nas páginas 42 e 43.
- KIZONY, R.; KATZ, N.; WEISS, P. L. Virtual reality based intervention in rehabilitation: relationship between motor and cognitive abilities and performance within virtual environments for patients with stroke. In: . [S.l.: s.n.], 2004. Citado na página 37.
- KLAASSEN, R.; BUL, K.; AKKER, R.; BURG, G.; KATO, P.; BITONTO, P. Design and evaluation of a pervasive coaching and gamification platform for young diabetes patients. *Sensors*, v. 18, p. 402, 01 2018. Citado na página 18.

- LIMA, T.; FILHO, A. B.; BARROS, A. K.; VIANA, D.; JUNIOR, J. B. B.; RIVERO, L. Avaliando um jogo educacional para o ensino de inteligência artificial - qual metodologia para avaliação escolher? *Anais do XXVIII Workshop sobre Educação em Computação*, p. 66–70, 2020. Citado na página 75.
- LIMA, T. M. d. S.; MENEZES, R. d. F. d. S.; FILHO, A. O. B.; BARROS, A. K. D.; VIANA, D.; CABREJOS, L. J. E. R.; JUNIOR, J. B. B. Desenvolvimento e aplicação de jogos sérios para o ensino de cinética química. *Research, Society and Development*, v. 9, n. 7, p. e197973760, maio 2020. Disponível em: <<https://rsdjournal.org/index.php/rsd/article/view/3760>>. Citado na página 75.
- LOPES, X.; NUSSBAUM, M.; TSAI, C. Some recommendations for the reporting of quantitative studies. *Computers & Education*, v. 91, p. 106–110, 2015. Citado 2 vezes nas páginas 15 e 16.
- LUND, A. Measuring usability with the use questionnaire. *Usability and User Experience Newsletter of the STC Usability SIG*, v. 8, 01 2001. Citado na página 37.
- MARTINS, E.; GOUVEIA, L. B. Uso da ferramenta kahoot transformando a aula do ensino médio em um game de conhecimento. In: *ERI MT 2018*. [S.l.: s.n.], 2018. Citado na página 17.
- MARTINS, R.; REIS, R.; MARQUES, A. B. Inserção da programação no ensino fundamental uma análise do jogo labirinto clássico da code.org através de um modelo de avaliação de jogos educacionais. In: *Anais do WIE 2016*. [S.l.: s.n.], 2016. p. 121. Citado na página 17.
- MARTINS, T.; CARVALHO, V.; SOARES, F. Web platform for serious games' management. *Procedia Computer Science*, v. 64, p. 1115 – 1123, 2015. ISSN 1877-0509. Conference on ENTERprise Information Systems/International Conference on Project MANagement/Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN / HCist 2015 October 7-9, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1877050915027064>>. Citado na página 18.
- MCCOMBS, J. A path analysis of the behavioral intention of secondary teachers to integrate technology in private schools in florida. 03 2012. Citado na página 37.
- MOODY, D. L.; SINDRE, G. Evaluating the effectiveness of learning interventions: An information systems case study. *ECIS 2003 PROCEEDINGS*, 2003. Citado na página 37.
- MOUAHEB, H.; FAHLI, A.; MOUSSETAD, M.; ELJAMALI, S. The serious game: what educacional benefits? *Elsevier: Social and Behavioral Sciences*, v. 46, p. 5502–5508, 2012. Citado na página 15.
- NEWBERY, R.; LEAN, J.; MOIZER, J. Evaluating the impact of serious games: the effect of gaming on entrepreneurial intent. *Information Technology & People*, v. 29, p. 733–749, 11 2016. Citado na página 15.
- O'BRIEN, H. L.; TOMS, E. G. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, v. 61, p. 50–69, 2010. Citado na página 37.

- OKOLI, C.; SCHABRAM, K. A guide to conducting a systematic literature review of information systems research. *SSRN Electronic Journal*, v. 10, 05 2010. Citado na página 27.
- PARSIFAL. *About Parsifal*. 2018. <<https://parsif.al/about/>>. Citado na página 25.
- PETERSEN, K.; VAKKALANKA, S.; KUZNIARZ, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, Elsevier, v. 64, p. 1–18, 2015. Citado na página 27.
- PETRI, G.; VON WANGENHEIM, C. G. How games for computing education are evaluated? a systematic literature review. *Elsevier*, v. 107, p. 68–90, 2017. Citado 6 vezes nas páginas 15, 18, 24, 33, 34 e 43.
- PETRI, G.; VON WANGENHEIM, C. G.; BORGATTO, A. F. A large-scale evaluation of a model for the evaluation of games for teaching software engineering. *IEEE/ACM 39th International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET)*, p. 180–189, 2017. Citado 4 vezes nas páginas 11, 35, 86 e 87.
- RANKIN, Y. A.; MCNEAL, M.; SHUTE, M. W.; GOOCH, B. User centered game design: evaluating massive multiplayer online role playing games for second language acquisition. *Sandbox '08 Proceedings of the 2008 ACM SIGGRAPH symposium on Video games*, p. 43–49, 2008. Citado na página 37.
- Ribeiro, C.; Monteiro, M.; Hauge, J. B.; Pereira, J.; Antunes, T. Sepsis fast track: A simulation game for clinical education based on the sepsis fast track protocol. In: *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)*. [S.l.: s.n.], 2016. p. 1–8. ISSN null. Citado na página 40.
- RITTERFELD, U.; CODY, M.; VORDERER, P. *Serious Games: Mechanisms and Effects*. 1. ed. [S.l.]: Routledge, 2009. Citado na página 15.
- RODRIGUEZ, D. M.; TEESSON, M.; NEWTON, N. C. A systematic review of computerised serious educational games about alcohol and other drugs for adolescents. *Drug and Alcohol Review*, v. 33, p. 129–135, 2013. Citado 2 vezes nas páginas 19 e 20.
- RYAN, R.; DECI, E. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American psychologist*, v. 55, p. 68–78, 02 2000. Citado 2 vezes nas páginas 35 e 37.
- SAFITRI, A. G.; PRIHATMANTO, A. S.; RUSMIN, P. H. Design and implementation of educational game based on thematic curriculum using three layered thinking model (case study: Applying number and social arithmetic in the real life). *4th International Conference on Interactive Digital Media (ICIDM)*, 2015. Citado na página 39.
- SANDE, D.; SANDE, D. Uso do kahoot como ferramenta de avaliação e ensino aprendizagem no ensino de microbiologia industrial. *Holos*, v. 1, p. 170–179, 2018. Citado na página 17.
- SANTOS, A.; SOUZA, M.; FIGUEIREDO, E.; DAYRELL, M. Game elements for learning programming: A mapping study. In: . [S.l.: s.n.], 2018. Citado na página 19.

- SAVI, R.; VON WANGENHEIM, C. G.; BORGATTO, A. F. A model for the evaluation of educational games for teaching software engineering. *Proceedings of the 25th Brazilian Symposium on Software Engineering*, p. 194–203, 2011. Citado 3 vezes nas páginas 17, 35 e 49.
- SCHICKLER, M.; PRYSS, R.; REICHERT, M.; SCHOBEL, J.; LANGGUTH, B.; SCHLEE, W. Using mobile serious games in the context of chronic disorders: A mobile game concept for the treatment of tinnitus. *IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS)*, 2016. Citado na página 39.
- SILVA, J. B. da; ANDRADE, M. H.; OLIVEIRA, R. R. de; SALES, G. L.; ALVES, F. R. V. Tecnologias digitais e metodologias ativas na escola: o contributo do kahoot para gamificar a sala de aula. *Thema*, v. 15, p. 780–791, 2018. Citado na página 17.
- SLIMANI, A.; FATIHA, E.; LOTFI, E.; YEDRI, O. B.; SBERT, M. Learning analytics through serious games: Data mining algorithms for performance measurement and improvement purposes. *International Journal of Emerging Technologies in Learning (iJET)*, v. 31, p. 46, 01 2018. Citado 2 vezes nas páginas 41 e 43.
- SLINEY, A.; MURPHY, D. Jdoc: A serious game for medical learning. In: . [S.l.: s.n.], 2008. v. 39, p. 131–136. Citado na página 24.
- SU, C.; HSAIO, K. Developing and evaluating gamifying learning system by using flow-based model. *EURASIA Journal of Mathematics, Science & Technology Education*, v. 11, p. 1283–1306, 2015. Citado na página 34.
- TAHIR, R.; WANG, A. State of the art in game based learning: Dimensions for evaluating educational games. In: . [S.l.: s.n.], 2017. Citado 2 vezes nas páginas 15 e 19.
- TRUJILLO, K.; CHAMBERLIN, B.; WIBURG, K.; ARMSTRONG, A. Measurement in learning games evolution: Review of methodologies used in determining effectiveness of math snacks games and animations. *Technology, Knowledge and Learning*, v. 21, p. 155–174, 2016. Citado na página 38.
- Tseloudi, C.; Tsiatsos, T. Panic in the gallery: An online educational game for art history: Design and evaluation of a matching game. In: *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*. [S.l.: s.n.], 2015. p. 1–6. ISSN null. Citado na página 37.
- TWINING, P.; HELLER, R. S.; NUSSBAUM, M.; TSAI, C. Some guidance on conducting and reporting qualitative studies. *Computers & Education*, v. 106, p. A1–A9, 2017. Citado na página 16.
- UFSC, U. F. de S. C. MEEGA+ A model for evaluating educational games. 2021. <<http://www.gqs.ufsc.br/quality-evaluation/meega-plus/>>. Citado na página 49.
- ULRICH, F.; HELMS, N. H. Creating evaluation profiles for games designed to be fun. *Simulation & Gaming*, v. 105, p. 695–714, 2017. Citado na página 32.
- UNITY. *Unity*. 2021. <<https://unity.com/>>. Citado na página 50.
- VALLE, P. H. D.; TODA, A. M.; BARBOSA, E. F.; MALDONADO, J. C. Educational games: A contribution to software testing education. *IEEE Frontiers in Education Conference (FIE)*, p. 1–8, 2017. Citado na página 34.

- VIRVOU, M.; KATSIONIS, G. On the usability and likeability of virtual reality games for education: The case of vr-engage. *Computers Education*, v. 50, n. 1, p. 154 – 178, 2008. ISSN 0360-1315. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0360131506000789>>. Citado na página 40.
- WANG, R.; DEMARIAJR, S.; GOLDBERG, A.; KATZ, D. A systematic review of serious games in training health care professionals. *Society for Simulation in Healthcare*, v. 11, p. 41–51, 2016. Citado 2 vezes nas páginas 19 e 20.
- Weiss, P. L.; Gal, E.; Zancanaro, M.; Giusti, L.; Cobb, S.; Millen, L.; Hawkins, T.; Glover, T.; Sanassy, D.; Eden, S. Usability of technology supported social competence training for children on the autism spectrum. In: *2011 International Conference on Virtual Rehabilitation*. [S.l.: s.n.], 2011. p. 1–8. ISSN 2331-9569. Citado na página 37.
- WITMER, B. G.; SINGER, M. J. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, v. 7, n. 3, p. 225–240, 1998. Citado 2 vezes nas páginas 35 e 37.
- WRZESIEN, M.; RAYA, M. A. Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the e-junior project. *Computers Education*, v. 55, p. 178–187, 08 2010. Citado na página 24.
- Yuhana, U. L.; Mangowal, R. G.; Rochimah, S.; Yuniarno, E. M.; Purnomo, M. H. Predicting math performance of children with special needs based on serious game. In: *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)*. [S.l.: s.n.], 2017. p. 1–5. Citado na página 43.