

**Utilização de Jogos Como Método de
Avaliação**
**Um estudo exploratório de uma área a ser
desvendada**

Versão Final Após Defesa

Daniel Augusto da Silva Mendes Filho

Dissertação para obtenção do Grau de Mestre em
Design e Desenvolvimento de Jogos Digitais
(2º ciclo de estudos)

Orientador: Prof. Doutor Frutuoso G. M. Silva

Dezembro de 2025

Declaração de Integridade

Eu, Daniel Augusto da Silva Mendes Filho, que abaixo assino, estudante com o número de inscrição M9726 do 2º ciclo em Design e Desenvolvimento de Jogos Digitais da Faculdade de Artes e Letras, declaro ter desenvolvido o presente trabalho e elaborado o presente texto em total consonância com o Código de Integridades da Universidade da Beira Interior.

Mais concretamente afirmo não ter incorrido em qualquer das variedades de Fraude Académica, e que aqui declaro conhecer, que em particular atendi à exigida referenciação de frases, extratos, imagens e outras formas de trabalho intelectual, e assumindo assim na íntegra as responsabilidades da autoria.

Universidade da Beira Interior, Covilhã 04 / 12 / 2025

A handwritten signature in black ink, appearing to read 'DASMF', with a long horizontal stroke extending to the right.

(assinatura conforme Cartão de Cidadão ou preferencialmente
assinatura digital no documento original se naquele mesmo formato)

Dedicatória

Dedico este trabalho a Deus e à minha família, que nunca desistiu e sempre acreditou em mim, e a meus professores e amigos, que viram em mim a perseverança e vontade de sempre ir além.

Agradecimentos

Agradeço primeiramente a meu orientador, Frutuoso Silva, sem a confiança dele eu não teria conseguido chegar à conclusão desse trabalho. Também quero agradecer aos Professores Bruno Saraiva e Ricardo Wagner do Centro Universitário Farias Brito, que fizeram tudo que foi possível para auxiliar com a realização desse projeto.

À minha família, agradeço a minha mãe, Isabel, e pai, Daniel, e ao Guido por estarem torcendo e sempre me apoiando por mais difícil ou inusitada que seja a aventura. Também agradeço à minha avó Dulce e tia-avó Suzana, que sem o apoio delas, mesmo depois de partirem, eu nunca teria tido as oportunidades de chegar aonde cheguei. Agradeço de todo o coração a minha tia-avó Maria Augusta e meus avós Manoel e Armênia e tia Patrícia, que sem o apoio eu não teria como chegar aonde cheguei.

Agradeço também a meus amigos Gabriel Galam, Joaquim Cado e Adélia Cristina, sem o auxílio deles durante todos esses anos, não teria chegado tão longe como estou agora.

Esta dissertação foi desenvolvida no âmbito do “Instituto de Telecomunicações” sob orientação do Prof. Frutuoso Silva.

Resumo

Introdução

O uso de jogos como ferramenta para auxiliar o processo de aprendizagem já está em estudo há anos, e a existência de dezenas de exemplos práticos eficazes de jogos sendo usados para a educação infantil e até mesmo para o treinamento profissional não pode ser questionada. Mas há uma pequena quantidade de estudos que observem a capacidade de jogos avaliarem o conhecimento dos alunos, com os estudos mais expressivos muitas vezes encontrando soluções que, por mais efetivas que sejam, encontram conflitos na implementação, como no desenvolvimento prático e na análise dos dados obtidos. Este estudo se propõe a explorar uma linha mais simples de abordagem, que se adequa com maior facilidade a sistemas de avaliação já presentes em escolas, ao aplicar uma roupagem digital lúdica a provas mais tradicionais, mas também permitindo coletas de dados simples e não invasivos que podem ser interpretados pelo professor para maior compreensão do conhecimento e dificuldades do aluno.

Neste estudo, são utilizados e analisados a Teoria de Resposta ao Item de três parâmetros (TRI 3PL) para garantir a confiabilidade dos resultados dos testes e determinar quais dados poderiam ser usados junto com ela para reforçar sua eficácia, os impactos que esse método de teste poderia ter nos alunos, incluindo, mas não se limitando ao aumento ou diminuição da ansiedade de teste, e os ensaios de implementação de tal sistema nos sistemas escolares.

Metodologia

Na primeira fase, foi realizada uma pesquisa de artigos acadêmicos nas áreas de Jogos de Vídeo na Educação, para encontrar as maneiras como os jogos já são comumente usados nas escolas; TRI, para garantir a confiança nos dados dos testes; Ansiedade Escolar, para encontrar uma conexão entre ela e o desempenho dos alunos nos testes; a Eficácia dos Jogos de Vídeo no tratamento da ansiedade.

Na segunda fase, jogos desenvolvidos em 2019 pelo NPC – Núcleo de Produção Criativa, do Centro Universitário Farias Brito, sob a liderança do Professor Bruno Saraiva e supervisionados pelo Coordenador dos cursos de Ciência da Computação Ricardo

Wagner, foram modificados e usados para testes presenciais, com a devida autorização, em alunos da turma de Design de Jogos, dos quais fizeram um teste escrito de múltipla escolha, jogaram o jogo e preencheram um formulário online onde compartilharam suas experiências com o sistema tradicional de testes de papel e caneta e com os jogos, e casos comparando ambos diretamente. Um teste remoto também foi realizado com um grupo separado de sujeitos, onde o teste anterior de papel e caneta foi adaptado para um teste online através de um formulário online, mas devido à sua natureza, os dados coletados dessa forma são menos seguros em comparação com os testes presenciais.

Os dados coletados foram então analisados considerando os resultados no jogo e no teste de papel e caneta e comparando-os com suas experiências compartilhadas através dos formulários, mantendo os dados e análises dos testes presenciais e remotos separados devido às diferenças de como foram aplicados.

Resultados e Análise

Os testes presenciais apresentaram uma taxa de participantes que relataram frequentes casos de Ansiedade de Provas (AP) maior que aqueles que não tinham ou acontecia com pouca frequência. No geral, viu-se uma melhora expressiva quando comparando a avaliação escrita com a avaliação por jogo, com Geometria Analítica alcançando uma melhora de 60%. Sobre suas opiniões os voluntários apresentaram boa aceitação à proposta, tal como acreditam que teria uma influência positiva nos seus resultados caso aplicado quando fossem crianças. Porém foi notado que para os voluntários os jogos ainda deram a sensação de fazer uma avaliação, dessa forma levantando a possibilidade de uma efetividade reduzida na prática.

Os testes remotos apresentaram, diferente do presencial, uma taxa de participante os quais relataram casos de AP menor que aqueles que não tinham ou que aconteciam com pouca frequência, sendo este um caso inverso ao teste presencial. Novamente percebeu-se uma melhora nos resultados quando comparados exceto por Geometria Analítica, que teve uma queda nos resultados, esses podendo ter causas não pertinentes à competência, mas sim ao estado técnico da plataforma utilizada para os testes. Dos voluntários percebeu-se também uma maior rejeição a plataforma utilizada tal como a ideia de influenciar positivamente seus resultados se aplicados durante seu período de infância. Porém igualmente aos voluntários do teste presencial, mesmo sem concordarem em todos os pontos igualmente, há uma opinião favorável à utilização de jogos como provas, com a maioria preferindo jogar a fazer uma prova caso permitidos escolherem.

Conclusão e Trabalhos Futuros

Foi notada uma melhora significativa de performance dos voluntários em sua maioria, mas notando problemas presentes na plataforma, na inadequação do meio remoto para os testes, tal como baixa taxa de amostragem em ambos torna inconclusivas quaisquer teorias e questionamentos levantados nesse estudo, mas ainda abrindo margem a debates como precedentes a futuros estudos, onde ao considerar a melhora nos resultados dentre a prova escrita para o jogo pode-se levantar a possibilidade de uma efetividade do método proposto. São necessários estudos mais extensos tal como uma taxa de amostragem exponencialmente maior, com preferência ao público infantil, de tal maneira necessitando a parceria com uma instituição de ensino disposta a aplicar este método alternadamente com provas tradicionais de tal maneira que seus resultados possam ser comparados ao longo de um ou mais anos. Tal como o desenvolvimento de uma nova plataforma, esta por sua vez, completamente desenhada para estes testes, não a adaptação de uma solução sucateada de outro projeto similar.

O uso de TRI 3PL provou-se essencial para garantia da validade dos dados, tal como aliar a dados extras coletados dos voluntários, sendo esse o tempo de conclusão e resposta de itens, permitiu uma maior compreensão dos resultados assim como de seus comportamentos, como já é por padrão compatível com sistemas de notas já utilizados em escolas atualmente. Para estudos e testes futuros é também notada a necessidade de uma integração direta com a API de plataformas de análise de dados, como o ChatGPT, para facilitar e agilizar o output dos resultados tratados, porém isso acarreta em um impacto direto ao custo de manutenção após sua implementação.

Palavras-chave

jogos na educação, teoria de resposta ao item, ansiedade no ambiente escolar, ansiedade no ambiente de teste, jogos digitais, jogos no tratamento de ansiedade

Abstract

Introduction

The use of games as tools to support the learning process has been under study for many years, and the existence of numerous effective examples of games applied in early childhood education and even professional training is unquestionable. However, only a limited number of studies have examined the potential of games to assess students' knowledge. Even the most promising studies often face implementation challenges, such as practical development or the analysis of the data provided by such games. This study seeks to explore a simpler approach—one that can be more easily integrated into existing school assessment systems—by applying a playful digital layer to traditional tests while also enabling the collection of simple, non-invasive data that can be interpreted by teachers to better understand students' knowledge and difficulties.

In this research, the Three-Parameter Logistic Item Response Theory (IRT 3PL) model is applied to ensure the reliability of test results and to examine which types of additional data could strengthen its effectiveness. The study also investigates the potential impacts of this testing method on students, including (but not limited to) increases or decreases in test anxiety, as well as the feasibility of implementing such a system within school environments.

Methodology

In the first phase, a literature review was conducted in four areas: Video Games in Education, to identify how games are already commonly used in schools; Item Response Theory, to ensure the reliability of test data; Test Anxiety, to explore its relationship with student performance; and the Effectiveness of Video Games in addressing anxiety.

In the second phase, games developed in 2019 by the Núcleo de Produção Criativa (NPC) at Centro Universitário Farias Brito, under the leadership of Professor Bruno Saraiva and supervision of Computer Science Program Coordinator Ricardo Wagner, were modified and used in in-person trials. With proper authorization, students from the Game Design program participated by completing a multiple-choice written exam, playing the game, and filling out an online form where they shared their experiences with both the traditional paper-and-pencil test and the game-based assessment, including direct comparisons.

A remote test was also carried out with a separate group, in which the previous written exam was adapted into an online form. However, due to its nature, the remotely collected data were considered less secure than those gathered in person. Data was then analyzed by comparing the outcomes of the written and game-based tests with participants' reported experiences, keeping in-person and remote analyses separate due to differences in administration.

Results and Analysis

The in-person trials revealed that a higher proportion of participants reported experiencing frequent Test Anxiety (TA) compared to those who reported little or no anxiety. Overall, a significant improvement was observed when comparing written assessments with game-based assessments, with Analytic Geometry showing up to a 60% improvement. Volunteers expressed strong acceptance of the proposal, many believing it would have positively influenced their performance had it been applied during childhood. However, participants also noted that the games still felt like an assessment, raising the possibility of reduced effectiveness in real-world application.

By contrast, the remote trials showed the opposite trend: participants reported fewer cases of TA compared to those with little or no anxiety in the in-person group. Performance improvements were observed in most areas, except for Analytic Geometry, which showed a decline likely attributable not to content mastery but to technical issues with the online platform. Remote participants also expressed greater dissatisfaction with the platform and skepticism about its potential positive impact during their school years. Nevertheless, as with the in-person group, most participants maintained a favorable opinion toward game-based assessments, with the majority preferring to play rather than take a traditional test when given the choice.

Conclusion and Future Works

While a significant improvement in performance was observed among most volunteers, issues related to platform limitations, the unsuitability of remote testing, and the small sample size render the findings inconclusive. Nonetheless, the study raises important questions and provides precedents for future research. Considering the observed improvements from written to game-based assessments, the method shows potential effectiveness.

Further research is needed with substantially larger sample sizes, ideally focusing on children, which would require collaboration with educational institutions willing to alternate between game-based and traditional assessments over one or more academic years to allow longitudinal comparison. In addition, the development of a new, purpose-built platform—designed specifically for assessment rather than adapted from existing projects—is necessary to ensure proper implementation and data reliability.

The use of TRI3PL proved essential in ensuring data validity. When combined with additional data collected from volunteers—specifically item response and game completion times, it enabled a deeper understanding of both the results and participant behavior. This approach is inherently compatible with grading systems currently employed in schools. For future studies and testing, there is a noted need for direct integration with APIs from platforms capable of data analysis, such as ChatGPT, to streamline and accelerate the output of processed results. However, this integration would have a direct impact on post-implementation maintenance costs.

Keywords

games in education, item response theory, school-related anxiety, test anxiety, digital games, games for anxiety management

Índice

1	Capítulo 1 Introdução	1
1.1	Recursos Educacionais Digitais (RED).....	1
1.2	Ansiedade de Provas (AP).....	2
1.3	Projeto Núcleo de Produção Criativa (NPC).....	3
1.4	Stealth Assessments (SA)	4
1.5	Teoria e Questões.....	4
1.6	Organização do documento	5
2	Capítulo 2 Estado da Arte	7
2.1	Stealth Assessments (SA)	7
2.2	Learning Analytics (LA).....	8
2.3	Mixed Reality Mobile Games.....	9
2.4	Ansiedade de Provas (AP).....	10
2.5	Núcleo de Produção Criativa (NPC)	11
3	Capítulo 3 Metodologia.....	13
4	Capítulo 4 Ferramentas Utilizadas	23
4.1	Unity	23
4.2	Visual Studio Community 2022 (VCS).....	23
4.3	Plataforma NPC	24
4.3.1	Book Chase (BC)	26
4.3.2	Fighting Against Anxiety (FAA).....	29
4.3.3	The Cube: Cubex (TC).....	30
4.4	Teoria de Resposta ao Item (TRI)	34
5	Capítulo 5 Resultados e Análise.....	39
5.1	Resultados Teste Presencial	39
5.2	Resultados Teste Remoto	50
6	Capítulo 6 Conclusão e Trabalhos Futuros.....	61
6.1	Conclusão.....	61
6.2	Soluções e Trabalhos Futuros.....	62
7	Bibliografia	65

Lista de Figuras

Figura 1 Plataforma MEC RED.....	2
Figura 2 Tela de registro dos testes.....	16
Figura 3 Tela de seleção de jogo dos testes.....	16
Figura 4 Instruções de download e envio de resultados para os testes remotos.....	20
Figura 5 Foto da Agenda 2003 do Colégio Farias Brito (Neto, s.d.).....	25
Figura 6 Tela inicial Book Chase, com instruções.....	26
Figura 7 Tela de jogo BC.....	27
Figura 8 Cena principal e Arquivos do jogo BC no projeto Unity.....	28
Figura 9 Tela de início jogo FAA com instruções.....	29
Figura 10 Tela de jogo FAA.....	30
Figura 11 Tela de início do jogo TC com instruções.....	31
Figura 12 Tela de jogo TC.....	31
Figura 13 Cena principal TC no projeto Unity.....	32

Lista de Tabelas

Tabela 1 Pontuação jogos para inclusão nos testes	15
Tabela 2 Resultados da prova escrita de História do Brasil, sem TRI	39
Tabela 3 Resultados da prova escrita de História do Brasil, com TRI	40
Tabela 4 Resultados da prova escrita de Geometria Analítica, sem TRI	40
Tabela 5 Resultados da prova escrita de Geometria Analítica, com TRI	40
Tabela 6 Respostas da prova escrita de História do Brasil	41
Tabela 7 Parâmetros da prova escrita de História do Brasil.....	41
Tabela 8 Resultados do jogo de História do Brasil, sem TRI	43
Tabela 9 Resultados do jogo de História do Brasil, com TRI	43
Tabela 10 Resultados do jogo de Geometria Analítica, sem TRI	43
Tabela 11 Resultados do jogo de Geometria Analítica, com TRI	44
Tabela 12 Resultados jogo de Lógica, sem TRI.....	44
Tabela 13 Resultados jogo de Lógica, com TRI	44
Tabela 14 Tempo de conclusão do jogo TC	45
Tabela 15 Resultados da prova remota de História do Brasil, sem TRI	51
Tabela 16 Resultados da prova remota de História do Brasil, com TRI.....	51
Tabela 17 Resultados da prova remota de Geometria Analítica, sem TRI	51
Tabela 18 Resultados da prova remota de Geometria Analítica, com TRI	52
Tabela 19 Resultados do jogo remoto de História do Brasil, sem TRI	52
Tabela 20 Resultados do jogo remoto de História do Brasil, com TRI	52
Tabela 21 Resultados do jogo remoto de Geometria Analítica, sem TRI	53
Tabela 22 Resultados do jogo remoto de Geometria Analítica, com TRI	53
Tabela 23 Resultados do jogo remoto de Lógica, sem TRI	54
Tabela 24 Resultados do jogo remoto de Lógica, com TRI.....	54
Tabela 25 Tempo de conclusão do jogo remoto TC.....	55

Lista de Gráficos

Gráfico 1 Fluxo formulário Google Forms, para testes remotos	19
Gráfico 2 Estrutura Arquivo de questões do jogo BC	27
Gráfico 3 Preferência entre prova escrita e jogo	46
Gráfico 4 Opiniões da efetividade dos jogos	46
Gráfico 5 Avaliação da compreensão dos comandos	46
Gráfico 6 Preferência entre prova escrita e jogo caso possível escolher	47
Gráfico 7 Avaliação da efetividade do sistema atual de provas nas escolas	47
Gráfico 8 Quantificação dos resultados do Gráfico 7	48
Gráfico 9 Presença de AP nos voluntários	48
Gráfico 10 Taxa de redução de estresse ao jogar	49
Gráfico 11 Avaliação da efetividade em mascarar as provas através dos jogos	50
Gráfico 12 Opiniões da efetividade dos jogos no teste remoto	56
Gráfico 13 Avaliação do teste remoto da efetividade do sistema atual de provas nas escolas	57
Gráfico 14 Quantificação dos resultados do Gráfico 13	57
Gráfico 15 Preferência no teste remoto entre prova escrita e jogo se possível escolher	58
Gráfico 16 Opiniões no teste remoto sobre performance com provas como jogo quando criança	58
Gráfico 17 Presença de AP nos voluntários do teste remoto	59
Gráfico 18 Taxa de redução de estresse ao jogar no teste remoto	59

Lista de Equações

Equação 1 Cálculo ICC.....	35
Equação 2 Cálculo de verossimilhança $L(\theta)$	35
Equação 3 Incorporação do prior	35
Equação 4 Cálculo EAP(θ)	36
Equação 5 Cálculo de erro padrão da estimativa, $SE(\theta)$	36
Equação 6 Aproximar Integral por Soma	36
Equação 7 Cálculo Expected Score IRT(θ)	36
Equação 8 Formula de cálculo da nota final, após aplicação do TRI3 PL	42

Lista de Acrónimos

UBI	Universidade da Beira Interior
AP	Ansiedade de Provas
EAP	Estimativa a Posteriori
ECD	Evidence-Centered Design
e-ECD	Expanded Evidence-Centered Design
FBUni	Centro Universitário Farias Brito
GRP	Gabinete de Relações Públicas
IA	Inteligência Artificial
ICC	Curva Característica do Item
ID	Identifier
IRT	Item Response Theory
JDC	Jogos Digitais Casuais
LA	Learning Analytics
MCMC	Monte Carlo via Cadeiras de Markov
MAP	Máxima a Posteriori
MEC	Ministério da Educação do Brasil
MLE	Maximum Likelihood Estimation (Máxima Verossimilhança)
MML	Marginal Maximum Likelihood (Verossimilhança Marginal)
NaN	Not a Number (Não é um número)
NPC	Núcleo de Produção Criativa
RED	Recursos Educacionais Digitais
SA	Stealth Assessments (Avaliações Escondidas)
SE(θ)	Standard Error de θ (Erro Padrão da proficiência estimada)
TA	Test Anxiety (Ansiedade de Provas)
TMPro	Text Mesh Pro
TRI	Teoria de Resposta ao Item
TRI 3PL	Teoria de Resposta ao Item de Três Parâmetros
UBI	Universidade da Beira Interior
U19	Unity 2019

1 Capítulo 1

Introdução

A digitalização dos ambientes de estudo e trabalho não são algo de hoje, salas de aula em vários locais do mundo, incluindo várias regiões do Brasil, dispõem de ferramentas como Tablets, Projetores, Computadores, para uso de Recursos Educacionais Digitais (RED), (Conselho Nacional de Secretários de Educação, 2022). Mas muito desse foco fica na utilização de tais ferramentas para auxílio do ensino, mas outros estudos apontam a possibilidade de utilização de jogos, como no trabalho “*Stealth Assessment: Measuring and Supporting Learning in Video Games*”, aponta que a utilização de Stealth Assessments como uma forma de medir o conhecimento e o que o estudante é capaz de fazer com tal conhecimento (Shute & Ventura, 2013). Também levantando a possibilidade de que a utilização de tal poderia alterar a noção de “sucesso acadêmico” além das notas escolares.

Mas para começar a abordar seus impactos de forma aprofundada, precisamos passar por alguns pontos, apresentados nas seções seguintes.

1.1 Recursos Educacionais Digitais (RED)

O uso de ferramentas para auxílio do ensino em escolas já é algo comum de ser visto e debatido, com o tempo é natural um termo para referenciar esse tipo de ferramenta é criado, alguns são utilizados, mas um que vem ganhando destaque ao longo dos anos é Recursos Educacionais Digitais, ou RED.

O destaque ganho pelo termo RED é evidenciado pela criação da plataforma “MEC RED” pelo Ministério da Educação (MEC) do Brasil, que visa de catalogar e disponibilizar acesso gratuito a diversas RED. No entanto, a definição de RED carece de consenso. Segundo Hadjerrouit em seu artigo, descreve RED como uma ferramenta de 4 dimensões, onde deve ser entregue pela Web, deve abordar o conteúdo alinhado com o currículo, tem base em uma teoria de aprendizado, e contém elementos reutilizáveis (Hadjerrouit, 2010).

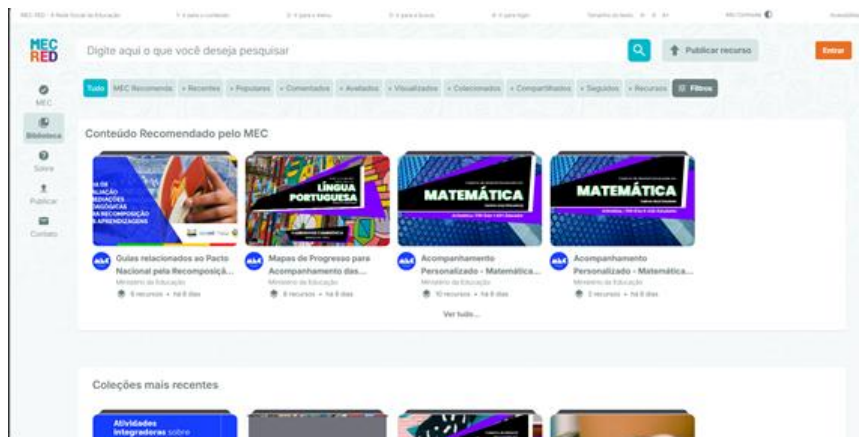


Figura 1 Plataforma MEC RED

1.2 Ansiedade de Provas (AP)

Na área de psicologia, diversos tipos de ansiedade são estudados, incluindo a ansiedade de provas (AP). Em seu artigo, “*The Prevalence and Effects of Test Anxiety in School Children*”, McDonald (2001), Descreve AP como “Ansiedade de provas é uma forma específica de um grupo mais abrangente, de problemas caracterizados por sentimentos de ‘ansiedade’” (McDonald, A., 2001, p. 3, tradução própria) e explica que podem ocorrer “Se em qualquer estágio de uma avaliação, nos sentimos despreparados, incertos de nossas habilidades, ou sentimento que não obtivemos nossa melhor performance.” (McDonald, A., 2001, p. 3, tradução própria). Esse é um tema que, desconectado de RED, vem sendo estudado há anos, com tentativas de identificar fatores causadores e medidas para redução da AP.

A ansiedade de provas é estudada tanto em seu escopo específico, como parte de um quadro mais amplo de ansiedade. O uso de jogos como uma ferramenta para tratamento de ansiedade já foi e ainda é objeto de estudo. Em “*The Effects of Casual Videogames on Anxiety, Depression, Stress, and Low Mood: A Systematic Review*” apesar não chegarem a um resultado definitivo, dados apontam um grau de efetividade do uso de Jogos Digitais Casuais (JDC) pode ser suplemento à medicação para ansiedade (Pine, Fleming, McCallum, & Sutcliffe, 2020). Isso levanta a necessidade de profissionais de saúde considerarem a utilização de JDC em tratamentos, inclusive especialmente devido aos diversos relatos de pessoas que utilizam JDC para se acalmar no dia a dia (Pine, Fleming, McCallum, & Sutcliffe, 2020).

Em “*Reducing Math Anxiety in School Children: A Systematic Review of Intervention Research*”, temos a afirmação que Ansiedade de Matemática (AM), pode ser causada por diversos fatores, dentre eles, de seu ambiente de convívio e pré-disposição genérica (Balt, Börnert-Ringleb, & Orbach, 2022). Isso pode se estender a AP quando a consideramos uma forma mais generalizada de ansiedade ligada à estudos.

Em “*A randomized controlled trial to test the effectiveness of an immersive 3D video game for anxiety prevention among adolescents.*”, testes foram feitos com o intuito de descobrir a efetividade da redução e prevenção de ansiedade em adolescentes com jogo “Dojo”, esse desenvolvido para o tratamento de ansiedade, e o comparando com “Rayman 2: The Great Escape”, que não possui esse objetivo. Os resultados mostraram que ambos os jogos têm eficácia semelhante na redução e prevenção da ansiedade, sugerindo que jogos em geral podem ter um efeito positivo na redução dos sintomas de ansiedade (Scholten, Malmberg, Lobel, Engels, & Granic, 2016). Isso apoia a ideia de que provas aplicadas como jogos podem ajudar na redução ou prevenção da ansiedade.

Fica evidente que AP é um fenômeno multifacetado, que requer uma abordagem abrangente para sua compreensão e tratamento. A integração de métodos tradicionais e inovadores, como JDC, pode oferecer novas perspectivas e soluções para aqueles que sofrem com essa condição. Um potencial sendo o projeto Núcleo de Produção Criativa, de 2019.

1.3 Projeto Núcleo de Produção Criativa (NPC)

No ano 2019, um projeto foi iniciado no Centro Universitário Farias Brito (FBUni), em Fortaleza, no estado do Ceará, comandado pelo Prof. Bruno Saraiva e supervisionado pelo Coordenador dos cursos de Ciência da Computação e Prof. Ricardo Wagner. Participei ativamente como Game Designer (GD) e Programador pelo período de aproximadamente 6 meses. O projeto utilizava jogos eletrônicos como ferramenta de avaliação do processo seletivo do Ensino Fundamental do Colégio Farias Brito (FB), testando seus conhecimentos em diversas competências, como História, Português, Matemática, Geografia, raciocínio lógico e interpretação de texto.

Para a realização foi formado um grupo voluntário de alunos e ex-alunos dos cursos de Design e Desenvolvimento de Jogos Digitais, Ciências da Computação e adjacentes, com habilidades em design de jogos digitais, arte e programação, desenvolveu um jogo utilizando o motor gráfico Unity 2019 (U19) como prova de conceito. Foram criados uma plataforma, que serve de central para registro de dados cadastrais, ajustes de perguntas por competência ensinada e acender aos jogos disponíveis. Na versão mais atualizada do projeto, encontra-se um total de 6 jogos: Book Chase(BC), Math Invaders (MI), Sky Spy (SkS), Fighting Against Anxiety (FAA), PC Adventure (PCA) e The Cube (TC). Esses jogos testavam competências em Matemática, Português, raciocínio lógico, dentre outras. Muitas das sensibilidades e design da plataforma e jogos foram feitos com base nas percepções dos participantes, algumas que em trabalhos pesquisados podem ter sido provadas ou não. Como por exemplo, os jogos feitos focavam em esconder os moldes de uma prova tradicional por meio de um jogo, diferente de outras abordagens como Stealth Assessments (SA).

1.4 Stealth Assessments (SA)

Como antes dito, estudos já foram feitos tentando observar a possibilidade de utilização de jogos, em seu livro “*Stealth Assessment: Measuring and Supporting Learning in Video Games*”, Shute & Ventura (2013) desenvolvem uma abordagem chamada de Stealth Assessments, que refere a aplicar avaliações dentro de jogos para que os estudantes sejam constantemente avaliados, sem uso de pontos específicos que interrompam a jogatina, preferindo sim coletar informações de ação dos jogadores, suas escolhas e estratégias para solução (Shute & Ventura, 2013). Também levantando uma redução no estresse aplicado no aluno por não estar sentindo que está sendo testado.

Apesar de inovador, Shute & Ventura apontam que tal método possui uma barreira grande de criação e implementação, requerendo um Game Design robusto, modelagem de psicometria e análise de dados. O que para esse trabalho o torna inviável de se aplicar (Shute & Ventura, 2013).

1.5 Teoria e Questões

Com todos esses conceitos introduzidos, o uso de jogos como método de avaliação tem seus desafios tanto de abordagem como implementação. O uso de SA prova-se demasiado abrangente para uma implementação nesse estudo exploratório, de tal forma, o foco para este define-se em uma abordagem mais tradicional, seguindo os moldes do NPC.

De tal maneira levantam-se as seguintes teorias e questões:

A utilização de jogos sem o uso de SA pode ter sua efetividade mesmo seguindo um sistema mais próximo ao tradicional das escolas, porém ainda permitindo uma coleta de dados maior que as provas tradicionais escritas, assim dando ao professor uma maior compreensão de cada aluno permitindo assim dar maior autonomia para com sua abordagem com o ensino.

Mesmo não utilizando SA, os usos desses jogos mais simples podem servir para a redução de AP em alunos mais jovens, de tal modo aumentando sua performance acadêmica.

Pelas dificuldades de implementação dos SA sua aplicação torna-se um desafio não somente monetário, mas logístico, pois da mesma forma de aplicação abre pouca possibilidade para mudanças no conteúdo. Para isso, questiona-se quais os custos de desenvolvimento, manutenção e equipamento seria necessário para a abordagem desse estudo.

1.6 Organização do documento

Esse documento foi estruturado em seis capítulos, os quais são separados em três partes que compõem o corpo, a Introdução, seguida pelo Desenvolvimento que abrange os capítulos de Estado da Arte, Ferramentas e “Resultados e Análise”; finalizando com a Conclusão com o capítulo de “Conclusão e Trabalhos Futuros”. Sendo descritos da seguinte maneira:

Introdução: Uma breve explicação da ideia principal. Apresentando, de forma resumida, por sua origem como ideia, seus possíveis impactos e pontos sensíveis, e levantando as teorias e questões que serão abordadas a frente. Também contém apresentação de informações essenciais para a compreensão desse estudo, como um relato do projeto que foi utilizado como base de todo esse estudo.

Estado da Arte: Uma visão ampla sobre outros estudos adjacentes, o que eles falam e como isso afeta este.

Ferramentas: Uma explicação sobre as ferramentas utilizadas nesse projeto, desde o que são, para o que servem, como para as mais importantes a esse estudo, como foram desenvolvidas, suas sensibilidades no processo, fraquezas e como eles foram remediadas.

Metodologia: Explica de forma detalhada o processo completo de coleta de dados e o tratamento deles. Descrevendo o uso das ferramentas descritas no capítulo anterior, o processo de coleta de dados, e descrevendo os dados coletados, seguido pela análise dos mesmos, para ser apresentada na Conclusão.

Resultados e Análise: Apresenta e de forma ordenada comenta os resultados dos testes feitos, com a presença de todas as tabelas e gráficos pertinentes à explicação e análise dos resultados.

Conclusão e Trabalhos Futuros: Explica a significância desses resultados e aponta quais os próximos passos para esse estudo, levando em conta todos os acertos e erros no processo desse estudo para que possam futuramente ser mantidos, alterados ou corrigidos caso necessários em estudos subsequentes.

2 Capítulo 2

Estado da Arte

2.1 Stealth Assessments (SA)

O estudo sobre a utilização de jogos como método de avaliação escolar vem se desenvolvendo ao longo dos anos, um dos mais notórios sendo uma série de trabalhos sobre SA criado por Shute & Ventura em “*Stealth Assessment: Measuring and Supporting Learning in Video Games*”. Neste livro são criados e desenvolvidos os o conceito dos SA, abordam sua implementação, impactos e desafios. SA são descritos como meios de continuamente através de um jogo coletar e analisar informações do jogador para avaliar seu conhecimento e lógica de solução (Shute & Ventura, 2013). Este estudo pega o uso de Evidence-centered Design (ECD), que em “*The Expanded Evidence-Centered Design (e-ECD) for Learning and Assessment Systems: A Framework for Incorporating Learning Goals and Processes Within Assessment Design*” ECD é descrito como uma estrutura sistemática para desenvolvimento de avaliações educacionais, que relaciona o conhecimento e as habilidades do aluno, garantindo validade ao definir explicitamente o conhecimento, perícias e habilidades para serem medidas (Arieli-Attali, Ward, Thomas, Deonovic, & von Davier, 2019).

Necessário para sua implementação é a constante atualização dos modelos para prover um retorno personalizado de informações. Aliado a interfaces utilizadas pelos professores, através dos SA, informações poderiam ser visualizadas em tempo real.

Uma barreira reconhecida de tal método é o custo inicial de implementação desse sistema, como visto em “*Games, Learning, and Assessment*” por Shute & Ke (2012), notando que para tal, é necessária uma modelagem de competência cuidadosa, aplicação de ECD e integração entre motores gráficos e pipelines de análise (Shute & Ke, Games, Learning, and Assessment, 2012).

Em “*The History of Stealth Assessment and a Peek Into Its Future*”, Shute (2023), analisa os anos de estudos e testes sobre SA, partindo de sua origem com sistemas de tutoria inteligente nos anos 80 e 90, para sua definição, aplicação, evoluções e debatendo os futuros rumos desse trabalho. Apontando seu futuro com a aplicação de inteligência Artificial (IA) para prover um retorno em escala, expansão para além de jogos, desafios quanto a privacidade de informação, interpretação dos modelos e custo de design. Para com a privacidade, é notado que SA coletam um volume grande de informação do aluno, levantando preocupações quanto a privacidade do mesmo, e apontando os riscos de perfilamento, mal uso de atributos pessoais e expor informações sensíveis dos alunos. Necessitando de tal forma uma governança responsável, incluindo comunicação clara com os alunos sobre o que é coletado, consentimento, segurança dos dados e políticas que garantam que esses dados serão utilizados apenas para avaliações e não formas não adequadas de

avaliação e monitoramento (Shute V. J., *The History of Stealth Assessment and a Peek Into Its Future*, 2023).

Os trabalhos de Shute servem como uma grande força nessa área, mas como notada em vários desses, seus custos de implementação são demasiados, fora a necessidade de uma colaboração multifacetada para seu desenvolvimento correto. Isso implica em uma ferramenta efetiva, mas difícil de ser implementada em todo o mundo, como no Brasil, onde a informatização do ambiente escolar público ainda é limitada, tornando ferramentas e sistemas que necessitam de custo de implementação e manutenção direta constante irreal para esses tipos de ambientes.

Outros estudos abordam as possibilidades além de SA, que poderiam se adequar de forma mais universal ao ambiente escolar, sem necessidade de um alto investimento.

2.2 Learning Analytics (LA)

Em “*Utilizing Learning Analytics to Support Study Success in Higher Education: A Systematic Review*.”, Ifenthaler & Yau. fazem uma análise sistemática de como Learning Analytics (LA), traduzido livremente para Análise de Aprendizado, pode ser aplicada como suporte ao aprendizado, engajamento e conquistas do estudante. Com uma forte ligação a ambientes digitais e jogos onde ações podem ser coletadas e analisadas (Ifenthaler & Yau, 2020).

Nele, é debatido que ao invés de provas discretas, o monitoramento contínuo da atividade de alunos dentro de plataformas de aprendizado pode ser utilizado. Descrevendo jogos como ambientes ricos em informações, com escolhas, tempo de solução, estratégias, etc. as quais podem ser passadas e processadas para prover testes em tempo real. Também afirmando que avaliações em jogos não precisam ser escondidas, podem ser efetuadas de forma explícita, com métricas de performance por tarefas em jogo, sequências de tomadas de decisão e uma interface para análise pelos professores (Ifenthaler & Yau, 2020).

Para com ansiedade de testes, Ifenthaler e Yau, afirmam igualmente a SA, há uma redução na AP por dissociação do ato de fazer uma prova, através das tarefas digitais ao invés de um exame momentâneo.

Porém igualmente ao sistema de SA, é notado que tanto seu custo de implementação como manutenção são uma barreira para escola e instituições com poucos fundos, devido a necessidade de infraestrutura robusta para a análise de dados, licenciamento de ferramentas de análise ou custo ainda maior de desenvolvimento de uma solução própria, treinamento para uso de professores e administradores para usar as ferramentas e interpretar os dados apresentados. Como também considerações sobre proteção de dados dos usuários.

Uma solução poderia ser aplicada com jogos aplicados a celulares, poderiam ser uma forma de expandir a acessibilidade para essas ferramentas pelos alunos e até professores, desde a ascensão dos smartphones, a acessibilidade e poder de processamento mínimo necessário pode estar presente neles.

2.3 Mixed Reality Mobile Games

Em “*A Mixed Methods Assessment of Students’ Flow Experiences During a Mobile Augmented Reality Science Game.*”, vemos o estudo da capacidade de jogos de realidade aumentada (RA), conseguem aumentar o engajamento do estudante, suportam experiências de imersão pela teoria flow, e se funcionam como ferramentas de avaliação, testando colaboração, investigação e resolução de problemas (Bressler & Bodzin, 2013). Para isso utilizam um jogo próprio chamado “*School Scene Investigators: The Case of the Stolen Score Sheets*”, que deve ser jogado dentro da escola e interagindo como códigos QR e elementos de RA. Os resultados apontam que estudantes ficaram engajados com sucesso na atividade, apontando surtos de foco e motivação a medida que jogavam e encontravam mais pistas, com melhores resultados presentes naqueles com interesse ou conforto prévio jogando jogos eletrônicos, de tal maneira, sugerindo que jogos como avaliações podem interagir de formas diferentes de um teste tradicional nas áreas afetivas e motivacionais (Bressler & Bodzin, 2013).

Diferente de avaliações normais, nesse estudo os alunos não agiam como em um teste tradicional, nem eram avaliados com notas e pontuações. Ele focava em permitir professores observarem como seus alunos interagem, colaboravam, persistiam entre desafios e utilizavam lógica com base nas evidências do jogo, dando uma visão maior em habilidades que não podem ser vistas em provas tradicionais de papel e caneta. Concluindo que jogos podem ser uma ferramenta complementar a avaliações, entretanto sua implementação requer que todos os alunos possuam acesso a esses dispositivos, treinamento dos professores e tempo para preparar os códigos QR. Sendo ainda sim, possivelmente o de menor custo de implementação dentre os trabalhos aqui abordados (Bressler & Bodzin, 2013).

Em contraste com SA, nesse os jogadores possuem ciência de seu progresso e capacidade de interação livre com seus semelhantes. SA focam na “invisibilidade” de sua operação, onde o estudante sequer tem ciência de que está a ser avaliado, com o objetivo de minimizar ao máximo AP e manter a experiência do jogador sem interferências automatizando ao máximo o processo. Quando comparado a abordagem de análise de aprendizado de Ifenthaler & Yau, também se nota contrastes, na qual os alunos sabem que estão sendo testados, e são testados de maneira contínua, mas o que é coletado são rastros deixados em jogos, como conclusão de objetivos, tempos de jogos, etc. tentando reduzir ansiedade de testes não pela omissão, mas tornando o ato em um processo formativo e não um evento único de alto impacto.

As três abordagens por si apontam filosofias diferentes, onde uma tenta omitir a existência de um teste, a outra busca diluir o impacto do teste ao tornar a avaliação algo constante, mantendo a transparência aos alunos e professores, e o outro estudo foca em complementar o sistema atual dando informações aos educadores sobre competências que não podem ser facilmente avaliadas com papel e caneta, também mantendo a transparência. Para tal, notou-se uma reincidência de custos de implementação altos, principalmente aqueles com altos volumes de coleta de dados, o que torna proibitivo a escolas e instituições com pouco poder financeiro para investimento em equipamentos desse calibre, tal como problemas quanto a privacidade dos alunos devido aos dados coletados. Fora, ambientes e escala de tempo necessárias de aplicação.

Um ponto em comum entre todos esses estudos foi a redução de AP, nos alunos testados. Mas prova-se necessária uma maior explicação de porquê é algo tão problemático.

2.4 Ansiedade de Provas (AP)

A ansiedade de provas é um fenômeno psicológico bem documentado e caracterizado por preocupação excessiva, tensão e reações físicas que afetam na performance de avaliação de situações. É por si, considerada uma forma específica de ansiedade de performance e tem sido estudada exaustivamente na psicologia educacional.

No livro *“International Handbook of Emotions in Education”*, Zeidner descreve AP como não meramente um nervosismo antes da prova, ele combina componentes cognitivos e psicológicos, que podem prejudicar memória e concentração, levando a uma performance reduzida em avaliações (Zeidner, 2014).

Os aspetos cognitivos incluem pensamentos intrusivos, que divergem recursos mentais da tarefa do momento. Afetivamente, estudantes podem sentir medo, temor, impotência, dentre outros, assim como sintomas físicos podem aparecer, como aceleração de batimentos cardíacos, suor e náusea (Zeidner, 2014).

Um dos aspetos mais problemáticos de AP é a constante relação negativa com a performance acadêmica. Estudantes com alta ansiedade geralmente obtêm notas mais baixas, e são mais propensos a abandonar seus estudos (Von der Embse, Barterian, & Segool, 2013).

“Effective interventions on test anxiety reduction: A meta-analysis.” aponta a existência de intervenções para tratamento, mais específico, Cognitive-behavioral therapy (CBT), que constantemente demonstrou efeitos positivos por ajudar estudantes a reenquadrar pensamentos mal-adaptados e reduzir reações físicas (Ergene, 2003). Outros estudos como em *“A Randomized Controlled Trial to Test the Effectiveness of an Immersive 3D Video Game for Anxiety Prevention among Adolescents”*, apontam que o uso de jogos pode ser utilizado como ferramenta complementar a tratamentos tradicionais, mas é necessário um maior interesse dos médicos para

poder testar sua eficácia de maneira mais ampla. Porém aliado a trabalhos mencionados anteriormente e seus resultados, pode-se considerar a conversão do ambiente de avaliação para jogos como um forte fator para a diminuição de AP (Scholten, Malmberg, Lobel, Engels, & Granic, 2016).

2.5 Núcleo de Produção Criativa (NPC)

Nesse sentido viu-se uma oportunidade de encontrar um mínimo denominador comum, ou um “minimum value product” (MVP), que seja capaz de se adequar a sistemas já presentes em escolas atualmente, seja facilmente acessível para escolas e instituições com menor poder financeiro, e permita uma coleta de dados relevante o suficiente para que o professor possa por si possa ter uma maior compreensão do progresso de seus alunos.

Para isso podemos observar o projeto de 2019, do Centro Universitário Farias Brito (FBUi), o Núcleo de Produção Criativa, ou NPC. Que buscava gamificar o processo de avaliação para candidatos a novos alunos do ensino fundamental I do Colégio Farias Brito. Para isso foi desenvolvida uma plataforma na qual o aluno deveria jogar jogos, que o avaliaria nas competências necessárias. Diferente dos estudos abordados até então, os jogos foram desenhados como uma nova roupagem de uma prova, não tentando escondê-la do estudante, mas dando uma apresentação mais amigável e descontraída. Os dados coletados originalmente não seriam mais que o padrão já coletado por provas de papel em caneta, o que diminui a necessidade de tal, mas abre outras oportunidades mais simples para se implementar tal sistema, tal como garantindo uma maior compatibilidade com sistemas escolares já presentes e fácil edição de seu conteúdo por professores, assim permitindo que um mesmo jogo teste várias competências. Mas podendo de mesmo modo, testar competências que não seriam facilmente avaliadas sem um acompanhamento direto, como raciocínio lógico e capacidade interpretativa.

Com o congelamento do projeto em 2020 em decorrência da pandemia COVID-19, a plataforma ficou em estado de abandono, mas servindo de pontapé para o desenvolvimento desse estudo.

3 Capítulo 3

Metodologia

Para este estudo, após um período de pesquisas de outros trabalhos acadêmicos de diversas competências ligadas a educação e jogos, mostrou-se intrigante ao ponderar as possibilidades e impactos do uso de jogos eletrônicos de tal maneira foi a ansiedade de provas. Como definido previamente, ansiedade de provas é vista uma generalização de um grupo de diferentes tipos de ansiedade ligadas ao ambiente escolar e que pode acontecer por diversos motivos. Com essa ideia em mente, foram pesquisados trabalhos os quais buscavam a utilização de diversos jogos como ferramenta para redução e ou prevenção de ansiedade.

Em "*Video game play and anxiety during late adolescence: The moderating effects of gender and social context*" Ohannessian, buscou por investigar as relações presentes entre sintomas de ansiedades em adolescentes, o papel do gênero e contexto social. Encontrou uma correlação inversa dependendo do gênero onde garotas que jogavam mais jogos eletrônicos possuíam um nível maior de ansiedade enquanto para garotos o nível de ansiedade foi reduzido, principalmente quando jogando com outras pessoas. Seus resultados levantam a possibilidade de que jogos possuem uma efetividade maior em garotos que garotas adolescentes, mas também abre questões quanto os tipos de jogos usados, e suas possíveis inadequações ao público feminino (Ohannessian, 2018). Com esse trabalho, levanta-se a possibilidade de que o efeito varie de acordo com o gênero, mas com as devidas sensibilidades, é possível que se tenha um resultado similar em ambos os gêneros. Vale ressaltar que esse artigo não leva em consideração identidade de gênero e somente os gêneros biológicos, que na sociedade atual e seus debates pode apresentar uma mudança nos resultados quando levados em consideração.

Este e outros trabalhos acadêmicos, como os previamente mencionados, serviram de base para a definição de AP como o principal foco desse estudo, assim centrando as perguntas mais abertas do teste nessa competência. Originalmente os testes eram planejados para serem feitos com alunos do fundamental I, com o auxílio da Organização Educacional Farias Brito, mas por problemas de cronograma por parte da organização, os testes foram revisados e alterados para que se adequassem a alunos universitários e ou profissionais formados.

Com o intuito de alcançar uma conclusão foi ponderada a necessidade de testes, e através de uma cuidadosa consideração de todas as possibilidades foi definida a seguinte estrutura. Primeiramente testes presenciais utilizando jogos digitais desenvolvidos para esse fim, seguido de uma prova escrita de múltipla escolha com as mesmas competências. Após os testes, foram pensadas duas possibilidades, uma entrevista com cada voluntário ou um questionário digital a ser preenchido onde os voluntários quantificariam de 1 a 10 sua experiência, apontariam suas

preferências de teste em cada cenário, e depois em alguns casos elaborariam abertamente suas opiniões.

Para a execução desses testes, como previamente dito, utilizando a ferramenta Unity foi modificado um projeto já existente no qual jogos estavam sendo desenvolvidos, pelo NPC, para o processo avaliativo de candidatos a novos alunos do ensino fundamental I, seguindo sistema escolar brasileiro, em que o autor deste estudo participou ativamente de seu desenvolvimento como programador, game designer e designer de UX, dos meses de março até setembro, do ano de 2019.

Durante seu desenvolvimento, noções e conceitos sobre as sensibilidades do design dos jogos foram definidos através de reuniões entre os integrantes e resultados de testes. Os testes efetuados nesse projeto, com alunos dentre o 5º e 6º ano do ensino fundamental II, e após os jogos havia o preenchimento de um formulário simples que tinha a seguinte estrutura:

- Qual a sua série? A) 5ª série ou B) 6ª Série;
- Qual tipo de jogo lhe agrada? A) Jogos de matemática ou B) Jogos de português;
- O Que achou da plataforma? A) Excelente; B) Bom; C) Regular ou D) Ruim;
- Sobre a sua experiência na plataforma, escolha uma nota de 1 a 5;
- Qual jogo você mais gostou? Pergunta de múltipla escolha, a qual o jogador poderia marcar os jogos que mais gostou;
- Qual tipo de jogo você gosta? Pergunta aberta;
- Descreva aqui seu elogio ou crítica da plataforma, gostaria que sua prova ou tarefas escolares fossem jogos? Pergunta aberta;

Originalmente os resultados desse formulário seriam utilizados para definir quais jogos seriam usados para os testes desse estudo, mas com a mudança de público para estudantes universitários e ou profissionais, os resultados foram desconsiderados.

Para determinar quais jogos seriam mais adequados, foi feita uma tabela, com pontos relevantes ao desenvolvimento, os jogos que receberam as pontuações mais altas foram escolhidos, esses pontos foram definidos por julgamento próprio. A tabela era composta da seguinte maneira:

- Familiaridade: é conhecimento do código, de estrutura a funcionamento;
- Simplicidade: é o quão fácil é para sentar e jogar em pouco tempo;
- Adaptabilidade: Quanto que esse jogo pode ser adaptado para outras competências

Tabela 1 Pontuação jogos para inclusão nos testes

Nome	Familiaridade	Simplicidade	Adaptabilidade	Total
Sky Spy	6	8	3	17
Fighting Against	9	9	9	27*
PC Adventure	3	1	0	4
The Cube	10	10	2	22*
Math Invadesr	4	6	6	16
Book Chase	7	8	4	19*

Com os jogos definidos, o desenvolvimento partiu na seguinte ordem:

1. Dar informações fictícias ao sistema de usuário antigo.
2. Isolamento das interfaces originais para com as novas.
3. Criação de um novo sistema de armazenamento de dados, na qual ele registaria o progresso do usuário e as seguintes informações com o intuito de criar um perfil. Esse perfil serve para garantir a existência do jogador e para organização de dados.
 - a. Nome;
 - b. Idade;
 - c. E-mail;
 - d. Gênero;
 - e. Área de estudo ou trabalho.
4. Adaptação do sistema de gravação de resultados dos jogos escolhidos para o novo sistema.
5. Criação de uma interface para acesso aos jogos pelo usuário.
6. Criação de métodos responsáveis por criptografia e descriptografia dos dados salvos.
7. Criação de uma build funcional.

Para evitar conflito com sistemas de gravação de dados da plataforma anterior, foi criado um método a ser executado na inicialização do jogo o qual injeta um perfil fictício para evitar problemas que pudessem vir da falta, tais como não carregamento das questões nos jogos. E um novo sistema de registo foi feito para os testes desse estudo, nele, o usuário cria um perfil que armazena os dados pessoais requeridos e em uma “List<>” de objetos da classe “ScrOb_SaveFile” estende de “ScriptableObject”. Nessa classe são guardados um objeto, “Script_User” responsável por guardar somente os dados pessoais, e uma “List<>” da classe gameSaves, que guarda os dados do progresso em cada jogo.

Foi escolhido por ser uma lista, para em caso de testes que precisassem ser realizados na mesma máquina em sequência, não fosse necessária fechar o jogo e retirar o arquivo “.enc” antes de começar o próximo. Durante uma sessão, o jogo gravava ao usuário um ID, que serviria para isolar os resultados em casos como o descrito. Isso se provou desnecessário para os testes, pois as vezes que foram testadas já havia máquinas o suficiente para todos os voluntários. Mas vale ressaltar

que após a bateria de testes foi encontrado um erro no carregamento das perguntas do jogo TC, nele ele só lia e armazenava os dados das perguntas para a primeira pessoa, nas seguintes, ele não recriava a List<Questions>. Esse bug não foi corrigido pois só foi encontrado após a bateria de testes.

A plataforma original possuía um sistema incompleto de gravação dos dados, e que se provava após análise pessoal, pouco prático para modificação. Para tal, foi criado o sistema de gravação que foi explicado em maiores detalhes no capítulo 2.

Com o processo de isolamento do sistema antigo já funcional, foi criada uma tela de registo, para qual o voluntário iria criar seu perfil, e outra para acesso aos jogos, o que pode ser visto nas imagens a seguir.



The image shows a registration form titled "Cadastro" on a green background. It includes the following fields: "Nome Completo:" with a placeholder "Nome*", "Idade:" with a placeholder "Idade (Somente Números)", "Gênero:" with a dropdown menu showing "Masculino", "E-Mail:" with a placeholder "E-Mail*", and "Área de Estudo / Trabalho:" with a dropdown menu showing "Computação" and a sub-field for "Ano (Somente Números)" with a placeholder "Exemplo: coloque '1' sem aspas para 1º ano*". Below the fields is a "Confirmar" button. There are also several lines of small text providing legal notices and terms of use.

Figura 2 Tela de registo dos testes



The image shows a game selection screen titled "Seleção da matéria:" on a green background. It features four buttons: "Book Chase (Matemática)", "F.A.A. (História)", "O Cubo (Lógica)", and "Encerrar e Salvar". There are also several lines of small text providing legal notices and terms of use.

Figura 3 Tela de seleção de jogo dos testes

Na tela de seleção de jogos, eles foram distribuídos como na Figura 5, seguindo pelos temas: Geometria Analítica (Matemática), História do Brasil e Lógica.

As perguntas feitas que compunham o jogo, principalmente da competência de Geometria Analítica, foram feitas levando em conta uma limitação de tempo dos testes e lembranças dos voluntários, uma vez que essa competência é uma das primeiras ensinadas aos alunos, e que nem todos poderiam se lembrar de seu conteúdo mais complexo após sua conclusão.

Para “Book Chase” (BC), foram feitas as cinco perguntas de múltipla escolha a seguir:

- Qual é o centro da circunferência $(x - 1)^2 + (y - 2)^2 = 4$? (1,2); (2,1); (-1,-2) ou (4,2);
- Qual é o raio da circunferência $x^2 + y^2 = 9$? 3; 4; 6 ou 8;
- Qual é a inclinação de uma reta horizontal? (0,0); (1,1); (-1,-1) ou (0,1);
- Qual é a equação da reta que passa pelo ponto (0,2) com inclinação -1 ? $Y = -X + 2$; $Y = -XY$; $Y = 2X - 1$ ou $Y = 2X - 1$;
- Qual é o valor da distância entre os pontos A(1,1) e B(4,5)? 5; 4; 7 ou 8.

Para “Fighting Against Anxiety” (FAA), foram feitas as 10 perguntas de múltipla escolha a seguir:

- Qual foi a principal atividade econômica do Brasil durante o período colonial? Cana-de-açúcar; Especiarias; Mineração ou Escravagista;
- Quem proclamou a independência do Brasil? Dom Pedro I; Getúlio Vargas; Dom Pedro II ou José Bonifácio;
- Em que ano ocorreu a Proclamação da República no Brasil? 1889; 1988; 1898 ou 1822;
- Quem foi o líder da Inconfidência Mineira? Tiradentes; Dom Pedro II; Marechal Deodoro ou Juscelino Kubitschek;
- Quando foi abolida a escravidão no Brasil? 1888; 1822; 1890 ou 1850;
- Durante qual período ocorreu o chamado "ciclo da borracha"? Final Sec. XIX; Sec. XVII; Sec. XVIII ou Sec. XVI;
- Quem assinou a Lei Áurea? Princesa Isabel; Dom Pedro II; Floriano Peixoto ou José Bonifácio;
- Quem foi Tiradentes? Líder Revolucionário; Líder Indígena; Imperador ou Cientista

Foram feitas 2 provas, que diferente dos jogos não houve teste de lógica, pois sua aplicação escrita se tornava menos prática. Seu conteúdo buscava ter a mesma área de conhecimento das questões dos jogos. Perguntas que porventura vieram a se repetir foram escolhidas por sua repetição não ser considerada algo prejudicial à aplicação dos testes.

Para a prova de Geometria Analítica, foram feitas as seguintes questões:

- Qual é a equação da reta que passa pelos pontos (2, 3) e (4, 7)? $Y = 2X - 1$; $Y = X + 1$; $Y = 2X + 1$ ou $Y = X - 1$;

- Qual é a distância entre os pontos (1,2) e (4,6)? 5; 7; 3 ou 6;
- Qual é a equação da circunferência de centro no ponto (3, -2) e raio 5? $(X - 3)^2 + (Y + 2)^2 = 25$; $(X + 3)^2 + (Y - 2)^2 = 25$; $(X - 3)^2 + (Y - 2)^2 = 25$ ou $(X + 3)^2 + (Y + 2)^2 = 25$;
- Qual é a equação da parábola de vértice em (0, 0) e foco em (0, 2)? $X^2 = 8Y$; $X^2 = 4Y$; $Y^2 = 4X$ ou $Y^2 = 8X$;
- Qual é a inclinação da reta que passa pelos pontos (-1, 4) e (3, 10)? 3/2; 2; 5-4 ou 3;

Para a prova de história do Brasil foram feitas as oito questões seguintes. Originalmente eram dez, mas devido a um erro na formulação da quinta e décima perguntas, estas foram cortadas para não afetar os resultados.

Em que ano ocorreu a independência do Brasil? 1808; 1822; 1889 ou 1914;

- Quem foi o primeiro imperador do Brasil? Pedro I; Pedro II; João IV ou Dom Luís;
- Qual foi a capital do Brasil antes de Brasília? São Paulo; Belo Horizonte; Rio de Janeiro ou Salvador;
- Em que ano ocorreu a abolição da escravidão no Brasil? 1865; 1877; 1888 ou 1900;
- Quem foi o principal líder da Inconfidência Mineira? Zumbi dos Palmares; Tiradentes; Dom Pedro I ou Dom Pedro II;
- Em que período ocorreu a ditadura militar no Brasil? 1930 – 1945; 1945 – 1964; 1964 – 1985 ou 1985 – 2000;
- Qual presidente brasileiro foi eleito após o fim da ditadura militar? Fernando Henrique Cardoso; José Sarney; Itamar Franco ou Fernando Collor;
- Qual foi o tratado que definiu as fronteiras entre Brasil e Portugal durante o período colonial? Tratado de Tordesilhas; Tratado de Madrid; Tratado de Utrecht ou Tratado de Paris;

Os testes presenciais foram aplicados no dia 9 de Abril de 2025, no Centro Universitário Farias Brito, com os alunos da aula de Design de Jogos Digitais, sob a supervisão e participação do professor Bruno Saraiva. O teste contou com 9 candidatos. Quais foram introduzidos aos conceitos que originaram esse estudo como uma visão aberta do objetivo dos estudos, com devidos cuidados para não influenciar expectativas nem os guiar à um resultado específico.

Lhes foram entregues as provas escritas junto a um temo de consentimento, que também fora explicado aos participantes para garantir completa ciência do que dados seriam guardados, como seriam utilizados e o que seria divulgado. Cada participante dispunha de um computador, com seu uso cedido pela organização e pertencente ao laboratório de informática, com Sistema Operacional Microsoft Windows 10 com o jogo instalado. No quadro se encontrava um link que ao introduzir no navegador ascenderia ao sítio do Google Forms com o formulário sobre sua experiência e envio dos resultados de suas seções no jogo.

O questionário, possui 16 itens com uma mistura de perguntas abertas e quantitativas, algumas as quais eram conjuntas, para que os candidatos pudessem melhor se expressar sobre o assunto. As perguntas abordaram desde a preferência do candidato a fazer uma prova pelo meio tradicional da caneta e papel ou como um jogo eletrônico, sobre a facilidade de entendimento das regras e comandos dos jogos, se estava mais calmo ao jogar, como relatos de se o candidato já teve algum episódio de ansiedade de prova. Anexo a este formulário, como explicado anteriormente, foi enviado o arquivo “.csv” com suas respostas dos jogos.

Com esse questionário buscou-se uma compreensão pessoal da experiência de cada participante. Vale então destacar que em nenhum momento dos testes fora requerido aos candidatos, de qualquer versão, qualquer nível de elaboração de suas respostas fosse acima do que os mesmos estivessem confortáveis a compartilhar. Tal como em manteve-se uma distância dos candidatos durante a execução do teste, só me aproximando caso chamado para responder a alguma dúvida, para garantir que minha presença não fosse um elemento causador de pressão a eles.

Nos testes remotos, tal qual a 2ª parte dos testes presenciais, foi feita através da plataforma Google Forms, com a prova escrita sendo convertida por completo a um formulário, e o acesso ao jogo sendo permitido através de um “hyperlink” que permitia a descarga, assim como instruções de como o iniciar e como encontrar e enviar seus resultados após a conclusão. Por necessidade de um termo de consentimento, ele foi aplicado ao início do formulário, com o requerimento de concordar com todos os pontos para iniciar o teste, assim como um termo de garantia da autenticidade de seus dados, confirmando que não trapaceou de qualquer maneira possível para obter melhores resultados nas provas.

Para melhor compreensão, o formulário foi estruturado na seguinte ordem:

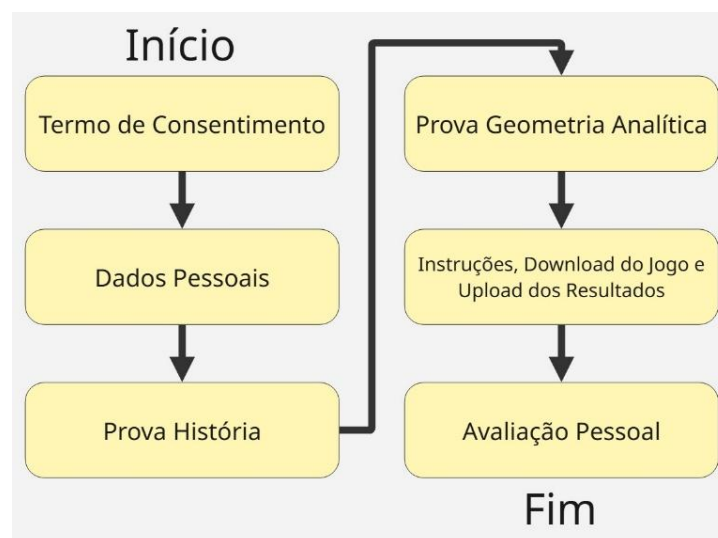


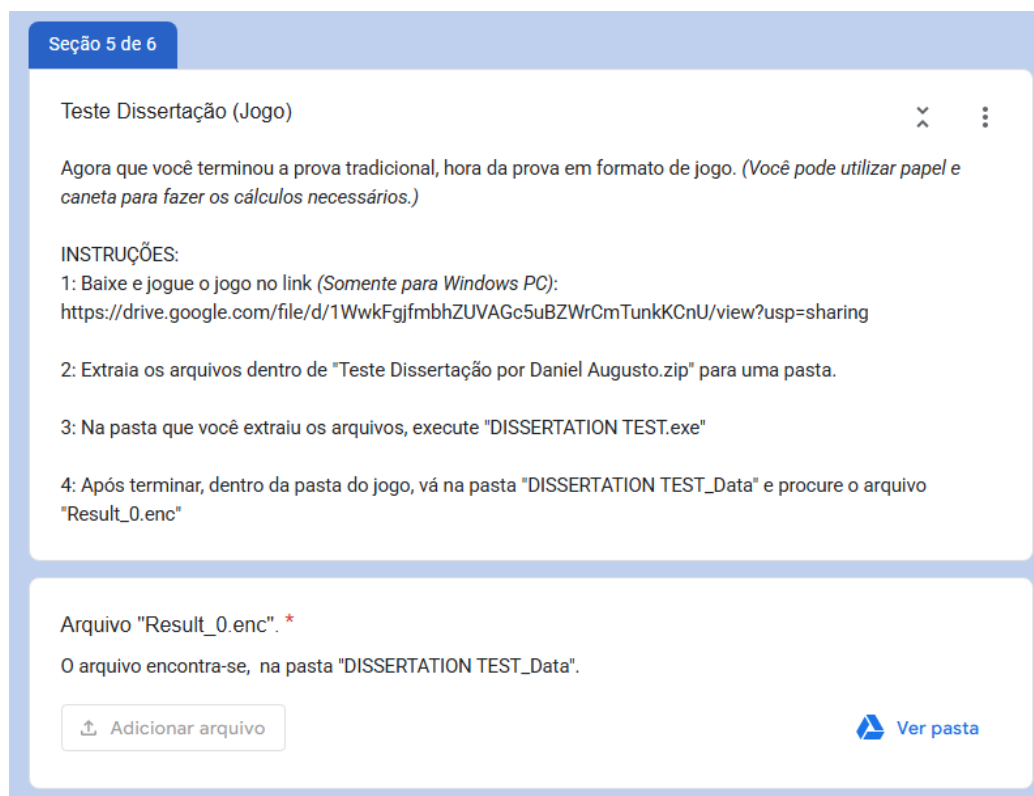
Gráfico 1 Fluxo formulário Google Forms, para testes remotos

Pela impossibilidade de os voluntários assinarem tradicionalmente o termo de consentimento nos testes remotos, foi elaborado um novo com pontos os quais o voluntário só poderia prosseguir com o teste após a confirmação de todos eles. Vale destacar que esse teste foi limitado somente a residentes legais no Brasil durante sua aplicação.

Foram coletados, Nome completo, Idade, Gênero e e-mail de contacto. Essas informações foram utilizadas para confirmar a autenticidade e para cruzar os dados com os arquivos dos jogos.

Diferente dos testes presenciais, os quais os voluntários possuíam a capacidade de os fazer na ordem desejada, no remoto, pelo funcionamento da plataforma Google Forms, foi necessária definir uma ordem específica. Levantou-se em consideração a hipótese de que, mesmo que minimamente, essa mudança possa afetar a experiência do voluntário. Porém para confirmar ou não tal hipótese são necessários mais estudos. As questões de cada uma das provas foram mantidas iguais as do teste presencial, porém a avaliação pessoal do estudo teve uma reorganização para o deixar esta parte mais sucinta.

O formulário foi distribuído através de link para grupos ligados a cursos de desenvolvimento de jogos, como de desenvolvedores profissionais, nas plataformas Discord, Telegram e WhatsApp e ficou disponível para ser respondido por um total de 5 dias. Como previamente explicado, os resultados salvos dos jogos para os testes remotos foram criptografados, assegurando sua autenticidade.



Seção 5 de 6

Teste Dissertação (Jogo)

Agora que você terminou a prova tradicional, hora da prova em formato de jogo. (Você pode utilizar papel e caneta para fazer os cálculos necessários.)

INSTRUÇÕES:

- 1: Baixe e jogue o jogo no link (Somente para Windows PC):
<https://drive.google.com/file/d/1WwkFgjfmhbhZUVAGc5uBZWrcmTunkKCnU/view?usp=sharing>
- 2: Extraia os arquivos dentro de "Teste Dissertação por Daniel Augusto.zip" para uma pasta.
- 3: Na pasta que você extraiu os arquivos, execute "DISSERTATION TEST.exe"
- 4: Após terminar, dentro da pasta do jogo, vá na pasta "DISSERTATION TEST_Data" e procure o arquivo "Result_0.enc"

Arquivo "Result_0.enc". *

O arquivo encontra-se, na pasta "DISSERTATION TEST_Data".

Adicionar arquivo

Ver pasta

Figura 4 Instruções de download e envio de resultados para os testes remotos

Como a estrutura dos arquivos com os resultados foi originalmente pensada para um sistema próprio, não era possível garantir sua compatibilidade com a ferramenta ChatGPT, por isso, os resultados foram manualmente transferidos a uma nova planilha que foi estruturada da com linhas representando os alunos e colunas as questões. O mesmo foi feito para guardar os resultados de cada prova física. Elas foram avaliadas, e seus resultados foram passados para uma planilha separada organizada da mesma maneira.

“student_ID; H01; H02; ... ; Hn”

Na primeira linha encontra-se a descrição das colunas, essa por sua vez, sendo disposta da seguinte maneira:

Servindo de guia para a ferramenta conhecer o que cada coluna representa. O student_ID é a identificação dos candidatos, esses “ID”, escritos como “SXXX” onde “S” é a representação de candidato e XXX é o número, por exemplo “S007” seria o ID do candidato número 7, para os testes remotos foram dados como “SRXXX”, seguindo a mesma lógica dos Ids dos testes presenciais, mas com o “R” como um indicador “Remote”. Esses Ids são consistentes entre todos os testes e foram determinados por ordem alfabética das planilhas originais.

As colunas subsequentes são os “ID” de cada questão. Nelas o “H” é um caractere variável, foi utilizado aqui para via de exemplificação, seu valor representa o tipo de prova, tal qual nas planilhas de Geometria Analítica, fora utilizada “GA”. O “n” é o número da questão. Isso se torna necessário para comparação com a planilha de “metadados” que será abordada a seguir. As respostas são representadas de forma binária, sendo “1” para acertos e “0” para erros.

Em conjunto é enviado uma segunda planilha com “metadados”, podendo também ser chamado “metadata”, incluem outras informações pertinentes ao cálculo do TRI. Essa planilha pode conter as informações de mais de uma prova, mas por organização esse trabalho fez uma para cada prova. Nela, divididas em colunas, essas são:

“test_ID; item_ID; n_alternativas; conteudo; conhecimento;”

O “test_ID”, por mais que seja redundante quando criando uma planilha separada de “metadados” para cada prova, é necessário quando se utilizando de uma única planilha para todas as provas.

O “item_ID”, é o mesmo da primeira planilha, serve de indicador para qual questão pertence a linha da planilha.

O “n_alternativas” é o número de opções de resposta que a questão possui, isso determina o valor de “c” no cálculo do TRI e pode variar caso questões não tenham um número padrão de opções,

mas nesses testes por via de simplicidade, foram padronizados como 4 opções por questão, tornando o valor de “c” mais comum como “0.25”. Exceto pela prova de lógica, a qual não era de múltipla escolha em si, por isso seu valor “c” diferiu das outras provas.

As colunas “conteúdo” e “conhecimento” são identificadores para caso a prova tenha mais de 1 matéria qual matéria ela pertence, e competência caso considere-se necessário a inclusão de “competência” serve como uma subcategoria para diferentes conhecimentos na mesma matéria. Por exemplo, temos as questões “HB; HO3; 4; História do Brasil; Regime Militar;” e “HB; HO4; 4; História do Brasil; Período Colonial;”. Assim para o cálculo do TRI, o valor de teta “ θ ” recebe maior influência de questões sobre o mesmo assunto. Para o primeiro teste por via de simplificação, foram aplicados “conteúdos” e “conhecimento” como o mesmo em todas as questões.

Nota-se também o uso de ponto e vírgula já que em alguns formatos “.csv”, são aplicadas vírgulas simples por padrão. Mas para a visualização correta pela plataforma Microsoft 365 Excel, a divisão entre células de uma linha é feita através de ponto e vírgula.

Por possuírem diferenças no acompanhamento entre os testes presenciais e remotos, tal como na organização de algumas perguntas, seus resultados foram tratados separadamente de acordo também e serão apresentados separadamente. Para os testes remotos, as respostas de cada candidato foram manualmente passadas para planilhas com a mesma formatação e organização das utilizadas nos testes presenciais.

4 Capítulo 4

Ferramentas Utilizadas

Para a execução dos testes se viu necessário o desenvolvimento de jogos com essa finalidade, o que, se desenvolvido desde seu início, levaria um tempo maior que o disponível. Assim, foi decidido ainda no início da fase de planejamento, que com a devida autorização, a adaptação e utilização dos jogos já desenvolvidos para o Projeto NPC da FBUni viabilizou a aplicação de testes para este estudo.

Para esse fim foi utilizada uma cópia dos arquivos do projeto que se encontram em um repositório privado da plataforma GitHub, da Microsoft. Assim, portado o projeto de sua versão original para a versão LTS 2019.4.40f1 do motor gráfico Unity, por falta de ferramentas que ao longo do desenvolvimento se provaram necessárias para sua conclusão, junto com a ferramenta Visual Studio Community 2022 (VCS), que oferece durante a instalação uma ferramenta, que possibilita a integração com a Unity.

4.1 Unity

A Unity, é um software denominado como “Motor Gráfico”, “Engine” ou “Game Engine”; para este estudo vamos utilizar o termo motor gráfico (MG). Ela é a ferramenta principal de desenvolvimento de jogos, e responsável por seu funcionamento. Nela o desenvolvedor tem disponível uma série de outras ferramentas para montar seu jogo utilizando-se de cenas, para montar telas, fases, dentre outras coisas, a capacidade de criar e colocar em cena objetos bidimensionais (2D) e ou tridimensionais (3D) que podem ser atribuídos módulos e scripts escritos pelo desenvolvedor para determinar seu funcionamento e comportamentos. Essa ferramenta também é responsável em pegar todos os arquivos do projeto e compilá-lo para se tornar o jogo em um programa próprio que possa ser executado para a plataforma desejada. Note-se que as exportações para diferentes plataformas possuem condições diferentes para ser possível.

4.2 Visual Studio Community 2022 (VCS)

Visual Studio Community, é uma plataforma criada pela Microsoft, e que é utilizada para a programação de aplicativos, jogos e páginas web. É a plataforma recomendada pela Unity, possuindo um módulo que permite uma melhor integração com o motor gráfico. Possui diversas ferramentas de conclusão inteligente e formatação automática, e “debug” para auxiliar e acelerar o desenvolvimento. Outra ferramenta, a qual foi recentemente adicionada, é integração com a plataforma de IA GitHub Copilot, também da Microsoft, que permite analisar, avaliar e corrigir problemas no código escrito. Também podendo sugerir e gerar código novo do zero, a partir de comandos e descrições dados pelo desenvolvedor.

4.3 Plataforma NPC

Originalmente desenvolvida em 2019, foi feita seguindo uma série de conceitos e sensibilidades, definidos em reuniões presenciais, as quais ditava como os jogos deveriam e poderiam se comportar, quais visuais deveriam ser utilizados, e quanto de informação deveria ser minimamente adquirida do professor e de seus alunos para a aplicação das provas, e como os resultados deveriam ser armazenados e organizados.

Para evitar que o aluno cometesse erros acidentais, como selecionar a resposta errada devido a um clique acidental do mouse, o jogo tem de ter a opção ou de regressar à pergunta anterior ou para recomeçar a prova. Para os testes feitos foi escolhida a opção de refazer a prova, pela maior facilidade de aplicação nos testes, mesmo que a opção de retroceder era considerada pelos membros uma mais efetiva. Todos os jogos desenvolvidos poderiam ser jogados quantas vezes o jogador acreditasse necessária, mas somente os resultados da última partida seriam gravados.

Um dos principais desafios do desenvolvimento se provou o balanceamento entre ação e retorno para com o usuário. Jogos para manter o interesse e chamar a atenção, utilizam-se de elementos de design de jogo, visuais e auditivos para indicar progresso e performance do usuário; como pontuação, efeitos visuais e sons. Foi levantado ainda no início do projeto como esses elementos poderiam causar confusão ao aluno, na qual o jogo poderia induzir a ideia de que ele estava acertando, mesmo errando, e em casos mais severos, o induzir a repetir o jogo até obter o retorno esperado. Mas com a necessidade de poder corrigir respostas anteriores aumenta significativamente a necessidade de ser neutro qualquer retorno após a seleção da resposta.

Para o visual, os membros, observaram que uma série de personagens da Organização Educacional Farias Brito estavam em desuso a anos, a “Turminha FB”, essas personagens foram criadas nos anos 2000 pela Organização Educacional Farias Brito junto ao artista Damasio Neto, para serem utilizados em materiais escolares, e promocionais, como agendas escolares, materiais de ensino dos professores, dentre outros (ver Figura 1). Em sua página da plataforma “artstation”, Damasio afirma que a ideia, vinha do desejo de ligar mais os alunos com a escola, onde essas personagens com personalidades variadas poderiam trazer uma identificação com os alunos e junto passar ensinar eles ideais como, mas não limitados há, diversidade, ética, cuidados com o meio ambiente (Neto, s.d.)

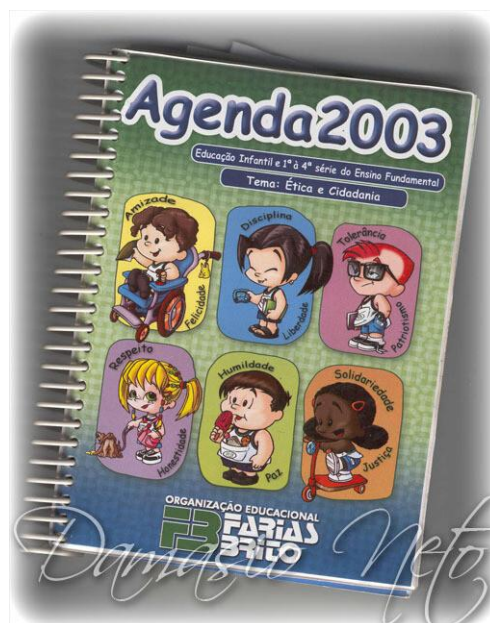


Figura 5 Foto da Agenda 2003 do Colégio Farias Brito (Neto, s.d.)

Com o alvo no público infantil, o uso dessas personagens levantou uma possibilidade de não só aumentar o interesse da instituição na procedência desse projeto, mas também a de uma revitalização das personagens com participação do NPC. Dentro desse jogo, não só foram feitos “assets” que utilizavam as personagens em diferentes situações para cada jogo, mas também foram criados a personagem do professor, que era a personagem responsável por explicar os comandos aos alunos.

Com a mudança da faixa etária dos voluntários aos testes para este estudo, mudanças necessitavam ser feitas tanto na plataforma, como nos jogos em si. Originalmente pensados para serem estudantes do Ensino Fundamental, seguindo o sistema escolar brasileiro, foram alterados para estudantes universitários por problemas com cronogramas da instituição que impediram a seleção e notificação dos pais, para que pudessem autorizar a participação de seus filhos nos testes.

Primeiro foi ponderada necessidade de tempo para adaptação dos jogos, incluindo formulação de novas questões e programação para adaptar os jogos existentes ao novo sistema de gravação de dados, foram escolhidos 3 jogos, levando em conta sua adaptabilidade para diferentes competências, onde o jogo deveria ser capaz de ser utilizado para mais de uma competência se necessário, facilidade de acesso aos jogadores, principalmente a facilidade de seus comandos, e conhecimento prévio do design do jogo e de seu código, este tendo o maior peso na hora da decisão, destacando que os jogos não possuíram necessidade de cumprir todos os três requisitos para serem escolhidos, como, caso um dos pontos se sobressaísse acima dos outros jogos. Levando-se em conta as necessidades, os jogos presentes foram reduzidos a três. Esses sendo:

4.3.1 Book Chase (BC)

Escolhido principalmente por sua familiaridade com o código escrito. Este foi o primeiro jogo desenvolvido por minha parte no NPC, e muito de sua estrutura e decisões de design ainda eram conhecidos, também sua temática narrativa, ajudava a assegurar uma maior imersão do jogador. Esse jogo buscava mascarar ao máximo o aspeto de prova e apresentar ao máximo os aspetos lúdicos, dando uma personagem que pode ser controlada diretamente pelo jogador e uma maior contextualização. Nele o jogador deve controlar uma personagem em uma biblioteca a qual, após uma confusão, livros caíram no chão e para guardá-los nas prateleiras corretas, ele deve pegar o livro com a resposta correta à pergunta no cartão da biblioteca. Seus comandos foram feitos para serem simples e intuitivos, onde para movimentar a personagem, o jogador necessitava usar as teclas “W” para movimentar para cima; “A” para movimentar para esquerda; “S” para movimentar para baixo; e “D” para movimentar para direita, e para selecionar a resposta desejada, era somente ir até o livro com a personagem e colidir com o livro (ver Figura 6 e Figura 7).

Esse jogo foi feito originalmente para a competência de álgebra, onde as respostas apresentadas necessitariam poucos dígitos para encaixar nos livros. Sua competência para os testes foi de Geometria Analítica, devido a sua presença na grade curricular dos primeiros semestres do curso de Jogos Digitais da FBUi.

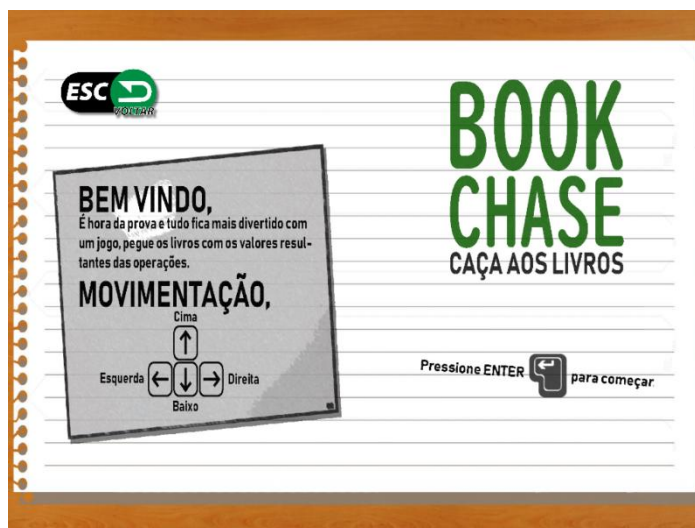


Figura 6 Tela inicial Book Chase, com instruções.

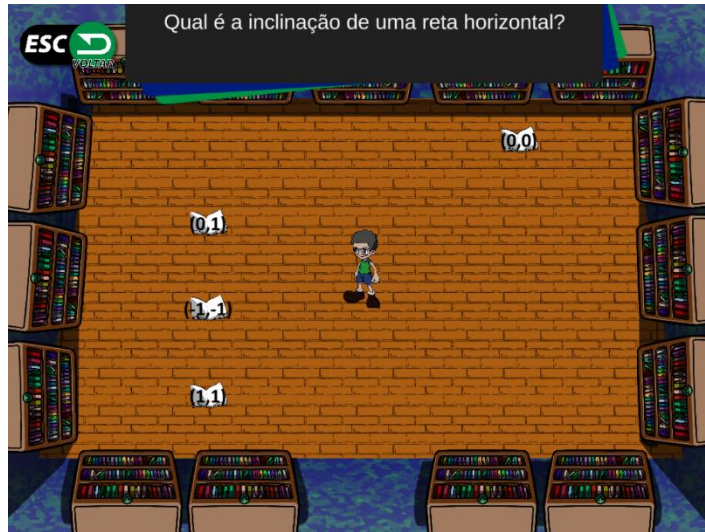


Figura 7 Tela de jogo BC

Para carregar as respostas, o jogo busca por um arquivo de texto o qual cada linha inclui uma pergunta e suas opções sob o seguinte formato:

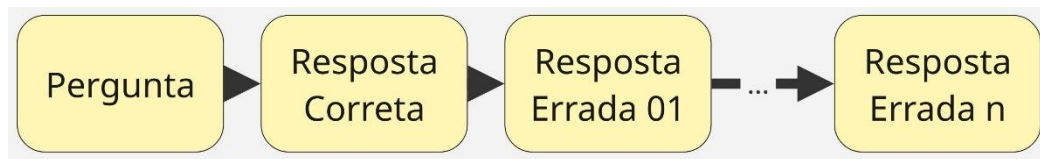


Gráfico 2 Estrutura Arquivo de questões do jogo BC

Sendo essas separadas pelo caractere “|” que foi escolhido por seu uso limitado na formulação de textos para provas.

Para carregar as perguntas foi decidido que, inicialmente, elas seriam lidas de um arquivo de texto “.txt”, porque durante o desenvolvimento dos jogos, o sistema para acesso do professor que o permitira adicionar, remover e editar as perguntas ainda estava em fase de estrutura, esse sistema nunca chegou a ser concluído, então para testes foi escolhido ler as informações de um arquivo de texto, por sua facilidade de acesso e edição.

Após o início do jogo as perguntas e respostas no arquivo de texto são colocadas em um “List<>” de uma classe “Question”, a qual armazena, uma “string” com a pergunta, e um “List<string>” com as opções de resposta, mantendo a ordem do arquivo de texto. Com o array pronto, ela escolhe aleatoriamente uma das perguntas, utilizando-se o método “UnityEngine.Random.Range(x,y);” (RandomRange), que retorna um número aleatório entre “x” e o valor anterior à “y” se trabalhando com “int” e um número entre “x” e “y” se trabalhar com “float”, números com pontos flutuantes, números que podem possuir decimais. A pergunta

selecionada é apresentada na tela e os objetos de livros são instanciados em tela com posições “parcialmente aleatórias”.

Para determinar a posição que o objeto terá, a cena do jogo possui uma série de objetos vazios que servem de marcadores para onde os livros podem ser instanciados. Quando uma opção é gerada, se for a primeira, um desses marcadores será escolhido através de um RandomRange, mas ignorando os marcadores dentro de um determinado raio da personagem do jogador, as demais levam em conta a personagem e outras opções já instanciadas em tela naquele dado momento.

Para determinar qual opção era a correta, como a primeira opção no array sempre era a correta, então dentro de um script no objeto do Livro, há uma booleana, onde os valores “true” representa resposta correta, e “false” representa uma resposta errada, a primeira opção a ser instanciada tem sua booleana dada o valor “true” e as demais recebem o valor “false”. Esse script também é o responsável por atualizar o texto da resposta no objeto do livro. Vale ressaltar que pela velocidade computacional atual, mesmo que haja uma ordem pela qual a primeira opção é sempre a correta, não é fisicamente possível um humano perceber sem que haja alguma falha na máquina, e por suas posições em tela serem sempre embaralhadas não é possível utilizar-se disso como uma trapaça.

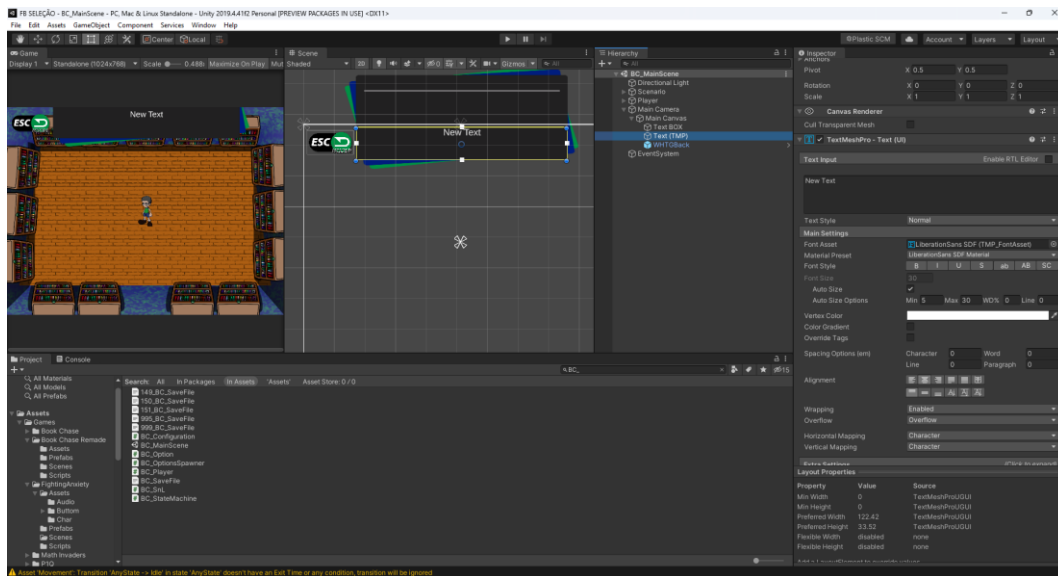


Figura 8 Cena principal e Arquivos do jogo BC no projeto Unity

Mudanças foram aplicadas ao projeto para a versão utilizada nos testes, estes por si incluem, a ajustes de elementos visuais como imagens, caixas de texto e fontes, estas que em ordem foram reescaloadas e tiveram suas âncoras ajustadas ao “canvas” da cena. Elementos de que previamente utilizavam o “UI Component” “Text” foram propriamente atualizados para “Text – Text Mesh Pro”, que possui melhor versatilidade como poder ajustar fontes, cores e contornos do texto sendo apresentado, como ajustar dinamicamente o tamanho da fonte, com valores máximos

e mínimos definidos pelo desenvolvedor. Porém isso também necessitou a atualização de arquivos como “BC_OptionSpawner.cs” que para acessar corretamente necessitava a inclusão da biblioteca “TMPro” como alterar todas as referências ao “Text” para equivalentes do novo componente. O arquivo “BC_Configurations.cs” também teve de ser alterado, onde foram adicionados uma classe extra “BC_SaveFile” que armazenaria os resultados, como na principal do arquivo fora adicionado o método “Diss_SaveFile”, que ao concluir o jogo grava os resultados no novo sistema de registro de progresso. No arquivo “BC_StateMachine.cs”, foram atualizadas as referências no estado “StateMachine.END” para gravar os dados salvos utilizando o novo sistema implementado.

4.3.2 Fighting Against Anxiety (FAA)

Escolhido principalmente por sua adaptabilidade para diferentes unidades curriculares, e pouco trabalho necessário para corrigir falhas presentes e adaptar o sistema de gravação de resultados criado para os testes desse estudo. Diferente do BC, esse diminuí significativamente o aspecto lúdico para, na época, servir de balança sob qual abordagem seria mais efetiva. Nele a personagem que representa o jogador, armado com seu lápis, está a enfrentar um monstro em formato de uma dobradura de origami feito de uma prova, e para o derrotar deve selecionar as respostas corretas, onde cada resposta representa uma ação em um combate de turnos de um jogo de RPG, esses sendo ataque, defesa, uso de item ou uso de magia. Originalmente a ideia era que a cada resposta do jogador, independente de acerto ou erro o monstro iria se desfazendo. Essa ideia não foi implementada por falta de tempo hábil dos artistas. Os comandos também foram feitos para serem os mais simples e intuitivos o possível, onde o jogador necessita somente mover o mouse e selecionar o item desejado. Para esse jogo foi escolhido a unidade curricular de História do Brasil, porque apesar de não ser algo presente na universidade, ainda é um assunto conhecido para todos que passaram pelo ensino médio.

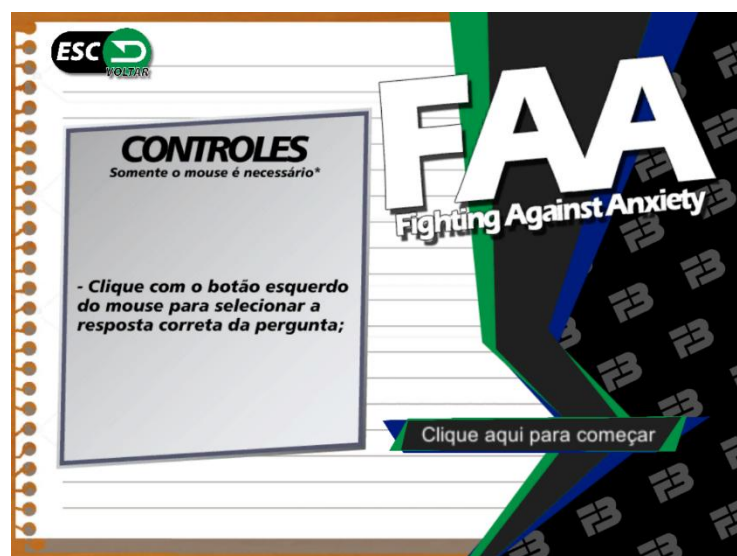


Figura 9 Tela de início jogo FAA com instruções



Figura 10 Tela de jogo FAA

Ao iniciar o jogo, igualmente ao BC, o arquivo de texto com as perguntas é lido e seus valores são guardados em uma “List<>” da classe “Question”. As perguntas são selecionadas de forma aleatória por um “RandomRange” e apresentada em tela. Mas diferentemente, suas opções são atribuídas aos botões em tela de forma aleatória.

Quando uma pergunta é carregada, uma “List<>” é gerada com referências aos objetos dos botões. Um desses botões é selecionado e aleatoriamente e os valores das opções de resposta são atribuídos a eles, botão que recebeu o valor é removido da lista. O primeiro botão tem a booleana definida como “true”, e as demais recebem o valor “false” (ver **Erro! Fonte de referência não encontrada.** e **Erro! Fonte de referência não encontrada.**). E igualmente ao BC, os mesmos princípios que garantem a segurança de integridade da prova e suas opções se aplicam aqui.

Mudanças feitas neste jogo limitam-se a similares aplicadas a BC. No qual para adaptar ao novo sistema de registro dos resultados, foram adicionadas ao método “Destroyer()” da classe “S_Player.cs”, que é responsável por remover os botões de opção e instanciar novos na cena, o que pode-se considerar contraintuitivo mas foi o método original aplicado para atualizar as opções de resposta e alterar poderia significar uma necessidade de reescrever muito do funcionamento do próprio jogo, neste método foi alterado então um condicional “if” que se não houverem mais questões a serem carregadas ao jogo, grava os resultados igualmente ao BC.

4.3.3 The Cube: Cubex (TC)

O último jogo, foi escolhido não por ser utilizado para alguma competência, mas sim por testar o raciocínio lógico do jogador e não possuir necessidade de um arquivo de texto para seu funcionamento. Nele os jogadores possuem em tela três objetos principais, primeiro, um cubo tridimensional que pode ser rotacionado pelo jogador e que em cada face possui uma letra ou

número, segundo, uma representação bidimensional de um cubo aberto com uma das faces com um dos caracteres ou números presentes no cubo tridimensional, e terceiro uma fila de quadrados com os caracteres presentes no cubo tridimensional e que não estão presentes no bidimensional. O jogador deve observar o cubo e através dele e das peças já aplicadas na imagem desconstruída do cubo, preencher as faces correspondentes com os quadrados na fila primando e carregando as faces para as posições corretas (ver Figura 11 e Figura 12).

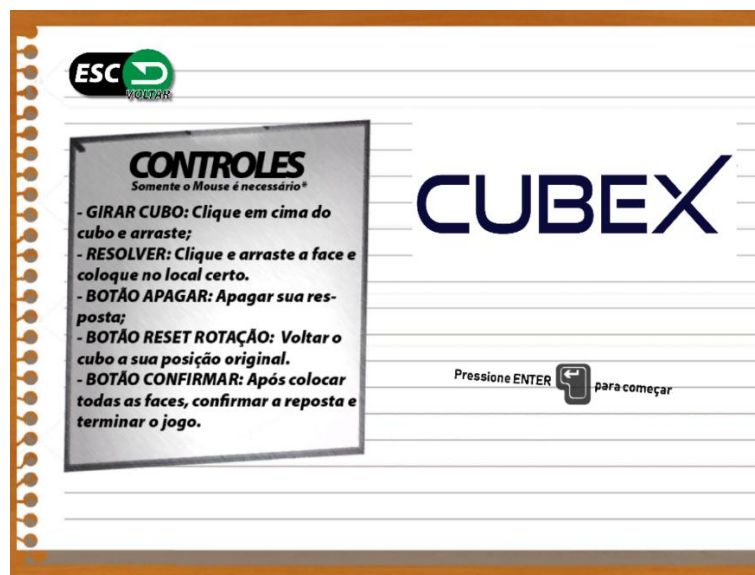


Figura 11 Tela de início do jogo TC com instruções

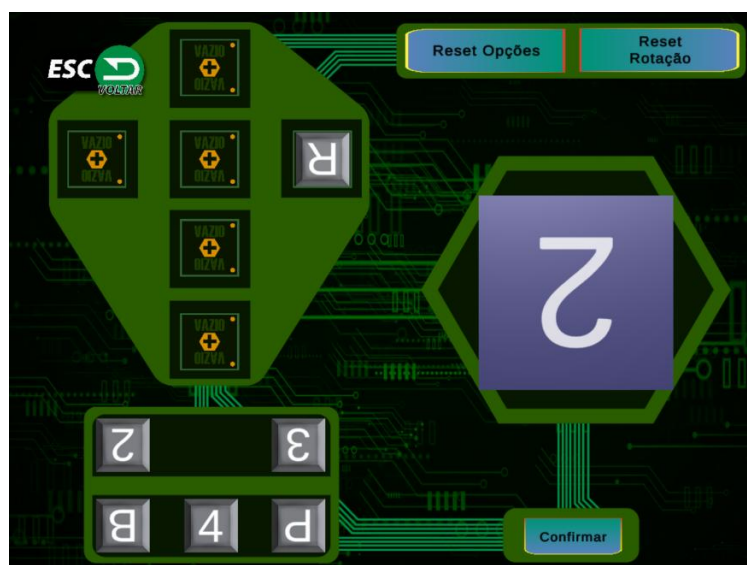


Figura 12 Tela de jogo TC

O uso mais importante para esse jogo não é comparar seus resultados com o de uma prova, tão pouco aumentar o número de testes para engordar os dados, mas sim, poder avaliar o nível de empenho de cada voluntário. Onde esse jogo, com sua simplicidade pode facilmente ser

respondido por um adulto, ou adolescente, assim resultados muito baixos, podem servir de indicador para a natureza de seu desempenho, mas não completamente.

Para esse jogo, os caracteres nas faces do cubo tridimensional são selecionados aleatoriamente, através de um RandomRange que permite pegar uma letra, maiúscula ou minúscula, ou um número e definir uma rotação também de forma aleatória. Para o cubo bidimensional, das faces do cubo tridimensional, são criados planos que possuem o mesmo caractere e rotação dele, uma é escolhida aleatoriamente e automaticamente colocada na posição correta como guia. Pois cada face do cubo desconstruído é referencia diretamente uma das faces do tridimensional e compara seu valor.

Se ao terminar de montar o cubo os valores de cada face são comparados com as respostas, para cada acerto, é guardado um “true” em uma “List<bool>” e para cada erro, na mesma lista é armazenado um “false”. Diferente dos outros, esse jogo não grava o tempo de resposta de cada face pois o jogador pode montar na ordem que desejar, então o tempo gravado é somente o tempo total para a conclusão do cubo. Vale notar que um erro essencialmente garante a existência de outro nesse teste, a pontuação não é para cada acerto, mas cada par, pois acertos e erros são em pares.

Este foi o que menos sofreu alterações, limitando-se atualização dos arquivos responsáveis por registro dos resultados, onde na classe “TC_TMControllers” foi atualizada o método “CompareAnswer()” para que fosse capaz de registrar as booleanas que representava cada acerto e erro de cada face do cubo, e na classe “TC_GameStateMachine”, que teve comentada a parte original de registro do progresso e foi adicionada as referências ao novo sistema implementado, igualmente aos jogos anteriores.

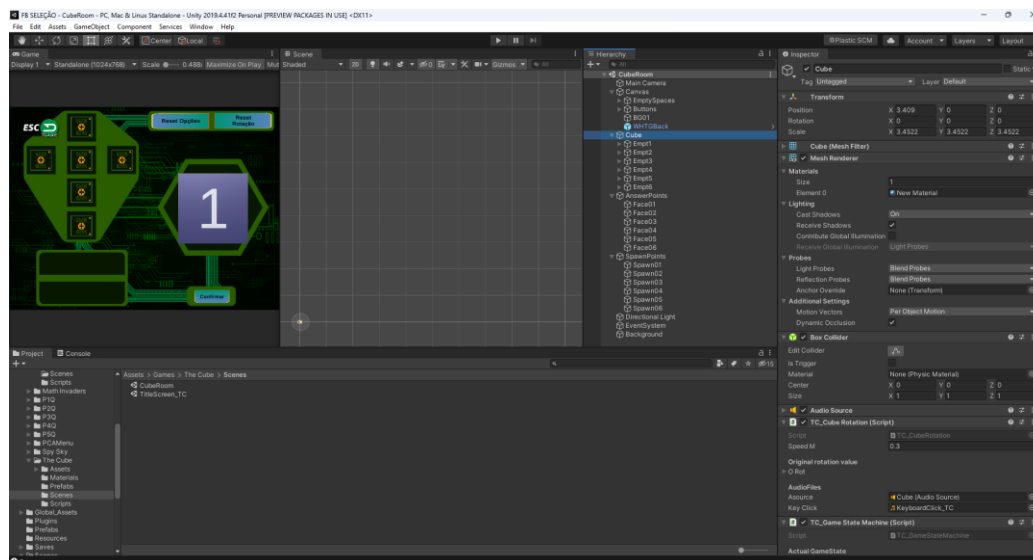


Figura 13 Cena principal TC no projeto Unity

Com cada jogo foi programado por uma pessoa diferente, em vários casos, a informação era gravada em arquivos e ficheiros diferentes, o que dificultava a análise. Para isso foi decidido que dos jogos a serem utilizados, seus sistemas para gravar resultados seriam sobrescritos com uma nova versão que seria capaz de salvar a todos os resultados de um aluno em um único arquivo “.csv”. Porém para os testes remotos, os resultados foram criptografados com a utilização do método “ExportToCSV” adicionado à classe “ScrOb_SaveFile”, o qual os salva em um arquivo “.csv” e depois, através da chamada do método estático “EncryptCsv” na classe “CsvEncryptor”, o criptografa e salva um novo arquivo sob a extensão “.enc”, apagando logo em sequência o arquivo original, a extensão “.enc” foi escolhida de forma arbitrária e não afeta em como os dados realmente foram guardados. Por não haver riscos imediatos de tentativa de trapaça, e facilidade de implementação, foi decidido que todos os arquivos teriam a mesma chave de criptografia.

Para fazer a descriptografia esses arquivos, foi criada a classe “Script_CsvDecryptor” no mesmo projeto da Unity, que em conjunto com a classe “Script_Editor_CsvDecryptor”, através do inspector, quando esse script está ligado ao objeto “BDConnection” que originalmente encontra-se na cena “InicializarBranco”, com o usuário pode aplicar o endereço onde se encontra o arquivo encriptado e onde encontra-se o arquivo com a chave de criptografia, ou abrir o “File Panel” da Unity para obter o endereço dos arquivos ao selecioná-los. E através de 1 clique no botão “Export to CSV”, fazer o processo de descriptografia ao chamar o método estático “DecryptCsv” que se encontra na classe “CsvDecryptor”, assim exportando um novo arquivo CSV.

As informações são divididas em 2 categorias, UserData, informações do usuário, e Game Data, os resultados dos jogos. As informações são guardadas com a seguinte estrutura

1. User Data:

- a. PLAYER ID: ID do jogador, para os testes foi determinado como zero por padrão, devida a necessidade de fazer o teste em múltiplas máquinas sem conexão com servidor e números aleatórios por menor a chance ainda podem causar conflitos por repetição, isso equivale a matrícula do aluno;
- b. NAME: Nome do jogador, não completamente necessário para o arquivo, mas foi mantido para organização dos dados e confirmação de identidade;
- c. AGE: A idade do jogador, utilizado organização dos dados;
- d. GENDER: Gênero do jogador, pode ser 1 de 3 valores, “Male” para masculino, “Female” para feminino e “Other”, para pessoas não binárias ou que preferem não divulgar seu gênero;
- e. “EMAIL”: endereço de correio eletrônico do jogador, serve para organização dos dados e para análise dos resultados e confirmação de identidade;
- f. COURSE: Curso ao qual está matriculado ou, para esse estudo, área de profissão, para organização dos dados;
- g. YEAR: Ano do curso, ou tempo de trabalho na área, utilizado para organização dos dados;

- h. MANDATORY CLASS: Aula obrigatória. Como os testes presenciais seriam feitos com alunos da FBUni, levou-se em consideração a política da organização para qual qualquer aluno de qualquer curso pode se registrar para aulas fora de sua unidade curricular como uma aula optativa. Essa informação provou-se redundante pois os testes foram feitos com alunos do curso de design e desenvolvimento de jogos digitais, qualquer outra opção marcada em course já confirmaria que essa é uma aula optativa, mas para evitar conflitos, ainda é registrado no arquivo “.csv”, apesar de não ser utilizado.
2. Game Data:
- a. GAME ID: Cada jogo é designado um ID, serve para identificação, todos os seus dados são guardados em um Lista, a qual, quando o progresso é gravado, busca por um ID compatível, se não o encontra, adiciona o resultado a uma nova posição na lista, se o encontrar, sobrescreve o resultado gravado previamente;
 - b. GAME RESULTS: Array booleano com os resultados de cada pergunta, essa é a informação principal para a análise dos dados;
 - c. TIME FOR ANSWER: Array float que guarda quantos segundo foram necessários para cada resposta ser respondida;
 - d. ANSWER VALUES: Array string que guarda os valores de cada opção escolhida, é guardado para maior feedback no erro do jogador.

Vale notar que a forma que as informações estão atualmente gravadas para o teste não são as ótimas para análise dos dados, foi mantida assim pela falta de tempo no desenvolvimento de um método mais robusto e mais bem estruturado.

Idealmente as informações do usuário e dos jogos devem ser organizadas em planilhas separadas, com os resultados de cada resposta sendo uma linha separada com as colunas organizadas como Booleano, Valor da Resposta e Tempo de Resposta.

Devido a limitação causada pela estrutura utilizada para criação dos arquivos “.csv”, os resultados de cada pessoa foram manualmente transferidos para um novo arquivo com uma organização similar a ideal que será descrito com maiores detalhes na metodologia.

4.4 Teoria de Resposta ao Item (TRI)

TRI é um conjunto de modelos estatísticos utilizados para analisar o desempenho de indivíduos em avaliações, buscando estimar a probabilidade de acerto com base em θ , que representa a habilidade.

“a” - Discriminação: Mede a capacidade do item de diferenciar entre alunos com diferentes níveis de habilidade próximos a curva logística de b (Dificuldade).

“b” - Dificuldade: indica o nível de θ (habilidade) necessário para ter 50% de chance de acerto acima de c (Acerto ao Acaso).

“c” -Acerto ao Acaso: representa a probabilidade mínima de acerto para alunos com baixa proficiência para $\theta \rightarrow -\infty$. Em provas de múltipla escolha é comum aproximar o valor de $1/k$, onde k = número de alternativas.

Existem diversos meios para obter os valores estimados de a , b , c e θ , como: Máxima Verossimilhança (MLE), Métodos Bayesianos (MAP / EAP) e Verossimilhança Marginal (MML). E podendo ser feita por algoritmo EM ou Montecarlo via Cadeias de Markov (MCMC), para o tratamento dos resultados utilizamos Estimativa a Posteriori (EAP), o qual pondera todos os valores possíveis de θ com base em quão verossímeis eles são dadas as respostas do aluno, e que para provas curtas como as aplicadas nos testes desse estudo é mais confiável, onde por exemplo, se aplicado MLE, caso um aluno ou voluntário acerte ou erre todas as questões, a verossimilhança excede-se para $+\infty$ ou $-\infty$. EAP resolve este problema ao impor um prior, o qual normalmente é dado como $N(0,1)$.

Primeiramente devemos calcular a curva característica de um item (ICC) para cada item presente no teste, que é dada pela seguinte equação:

Equação 1 Cálculo ICC

$$P_i(\theta) = Ci + \frac{1 - c}{1 + e^{-ai(\theta - bi)}}$$

Com $P_i(\theta)$ obtido, calculamos a verossimilhança $L(\theta)$. Esta mede o quão provável é observar o padrão de respostas de um aluno dado um valor de habilidade θ , e para isso utilizamos a seguinte equação:

Equação 2 Cálculo de verossimilhança $L(\theta)$

$$L(\theta) = \prod_{i=1}^n P_i(\theta)^{u_i} \cdot (1 - P_i(\theta))^{1-u_i}$$

Nela se o aluno acertou a questão ($u_i = 1$) utilizamos $P_i(\theta)$, e para erros ($u_i = 0$) utilizamos $1 - P_i(\theta)$, e seu resultado é uma função que mostra quais valores de θ tornam o padrão de respostas do aluno mais provável.

Com a $L(\theta)$ obtido, já que EAP é bayesiano, devemos incorporar o prior para garantir uma curva mais suave e que garante uma estimativa mais realista em provas curtas, para isso utilizamos a seguinte Equação, com $N(0,1)$ como dito anteriormente:

Equação 3 Incorporação do prior
 $Posteriori(\theta) \propto Prior(\theta) \cdot L(\theta)$

Com a curva suavizada podemos calcular o EAP, que é a média da posterior, para isso utilizamos a seguinte equação, na qual o numerador é a soma de todos os θ possíveis, ponderados pela chance de serem verdadeiros, e o denominador normaliza a distribuição garantindo que seja uma probabilidade.

Equação 4 Cálculo EAP(θ)

$$\hat{\theta}_{EAP} = \frac{\int \theta \cdot Prior(\theta)L(\theta)d\theta}{\int Prior(\theta)L(\theta)d\theta}$$

Com $\hat{\theta}_{EAP}$ definido agora devemos calcular o Erro Padrão da estimativa, o $SE(\hat{\theta})$, para tal utilizamos:

Equação 5 Cálculo de erro padrão da estimativa, SE(θ)

$$SE(\hat{\theta}) = \sqrt{\frac{\int (\theta - \hat{\theta}_{EAP})^2 \cdot Prior(\theta)L(\theta)d\theta}{\int Prior(\theta)L(\theta)d\theta}}$$

Com os valores estimados, integramos eles a grade $[-4, +4]$, e a divide em passo pequenos, para cada passo da grade se calcula o $P_i(\theta)$ de cada item, a verossimilhança $L(\theta)$, a $Prior(\theta)$ e $Posteriori(\theta)$ e aproximamos a integral por soma com:

Equação 6 Aproximar Integral por Soma

$$\int f(\theta) d\theta \approx \sum_i f(\theta_i)\Delta\theta$$

Feito isso, para reportar a escala de θ como nota esperada devemos integrar a posterior, para isso utilizamos o seguinte:

Equação 7 Cálculo Expected Score IRT(θ)

$$IRT(\theta) = \sum_i P_i(\hat{\theta}_{EAP})$$

Assim chegando a nota do TRI.

Recomenda-se a aplicação de TRI 3PL três vezes, uma com os resultados presentes, uma segunda com um aluno fantasma com 100% de acertos, e uma terceira com um aluno fantasma com 100% de erros, assim encontrando qual de facto é a maior e menor notas possíveis, pode-se então aplicar reescalonamento linear para melhor adequar ao sistema de notas escolar.

5 Capítulo 5

Resultados e Análise

Para uma melhor compreensão de suas experiências e relatos, é importante entender seus resultados. Os testes presencial e remoto serão abordados separadamente em ordem de aplicação, sendo o primeiro o teste presencial.

5.1 Resultados Teste Presencial

O teste presencial contou com uma amostra de 9 candidatos, esses de gêneros, idades e áreas de expertise diferentes. Para os cálculos e organização dos resultados, póstumo à aplicação foram atribuídos números de identificação que também foram aplicadas às planilhas com seus resultados, esses são os candidatos S001, S002, ..., S009, dessa forma ignorando a influência de gênero nos testes pois não é parte dos objetivos desse estudo.

Nas interações pessoais com os voluntários, percebeu-se uma inquietude vocal quanto a prova de geometria analítica, não por sua dificuldade, mas pelos voluntários não estarem mais tão familiarizados por desuso. Isso se provou importante para ilustrar o uso do TRI PL3 como garantia de precisão dos resultados. Vale ressaltar que TRI PL3 sempre considera uma chance de acerto como maior que zero e menor que dez, por isso um aluno tirar um zero ou dez absoluto é impossível após o tratamento das notas.

Tabela 2 Resultados da prova escrita de História do Brasil, sem TRI

student_ID	c_answers	w_answers	grade
S001	6	2	7.5
S002	6	2	7.5
S003	7	1	8.7
S004	5	3	6.2
S005	5	3	6.2
S006	4	4	5.0
S007	7	1	8.7
S008	6	2	7.5
S009	4	4	5.0

Tabela 3 Resultados da prova escrita de História do Brasil, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
S001	6	0.6014	0.6478	5.1657	6.46
S002	6	0.3603	0.7010	4.8776	6.10
S003	7	1.1718	0.7052	5.7369	7.17
S004	5	-0.2954	0.6797	4.0817	5.10
S005	5	0.3098	0.6168	4.8151	6.02
S006	4	-0.1521	0.6144	4.2456	5.31
S007	7	1.1718	0.7052	5.7369	7.17
S008	6	0.3603	0.7010	4.8776	6.10
S009	4	-0.8440	0.6967	3.5856	4.48

Tabela 4 Resultados da prova escrita de Geometria Analítica, sem TRI

student_ID	c_answers	w_answers	grade
S001	1	4	2.0
S002	1	4	2.0
S003	1	4	2.0
S004	3	2	6.0
S005	2	3	4.0
S006	0	5	0.0
S007	1	4	2.0
S008	0	5	0.0
S009	1	4	2.0

Tabela 5 Resultados da prova escrita de Geometria Analítica, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
S001	1	-0.402	0.835	0.864	1.73
S002	1	-0.402	0.835	0.864	1.73
S003	1	-0.352	0.700	0.808	1.62
S004	3	1.527	0.446	3.232	6.46
S005	2	0.163	0.722	1.224	2.45
S006	0	-0.561	0.755	0.738	1.48
S007	1	-0.352	0.700	0.808	1.62
S008	0	-0.561	0.755	0.738	1.48
S009	1	0.266	0.546	1.187	2.37

Com os resultados dispostos das provas escritas, se utilizarmos a média como 5, podemos observar que na prova de História do Brasil, todos acertaram ao menos 50% dela e atingiram a média. Porém ao aplicar a Teoria de Resposta ao Item nos mesmos resultados, a visão do real desempenho dos alunos se expande. Nessa temos o caso do aluno S009, que teve seus pontos reduzidos abaixo de cinco, mesmo tendo acertado quatro das oito questões. Isso ocorre ao aplicar os pesos de conhecimento a cada questão.

Tabela 6 Respostas da prova escrita de História do Brasil

student_ID	H01	H02	H03	H04	H05	H06	H07	H08
S001	0	1	1	1	1	1	0	1
S002	1	1	1	0	1	1	0	1
S003	1	1	1	1	1	1	0	1
S004	1	0	1	0	1	1	0	1
S005	0	1	1	1	1	0	0	1
S006	0	0	1	0	0	1	1	1
S007	1	1	1	1	1	1	0	1
S008	1	0	1	1	1	1	0	1
S009	0	0	1	0	1	0	1	1

Tabela 7 Parâmetros da prova escrita de História do Brasil

item_ID	A (discriminação)	B (Dificuldade)	C (Acerto ao Acaso)
H01	1.7068	0.8861	0.25
H02	1.8679	-0.0257	0.25
H03	1	0	0.25
H04	1.8679	-0.0257	0.25
H05	-1.3336	1.3972	0.25
H06	1.4590	1.3239	0.25
H07	-1.2754	-2.4509	0.25
H08	1	0	0.25

Observando os resultados do candidato S009, observam-se 4 acertos nas questões H03, H05, H07 e H08 (ver Tabela 6), onde notoriamente as questões três e oito, tem discriminação e dificuldade com valores de um e zero, respectivamente (ver Tabela 7). Questões com esses valores nos dá o nível médio de habilidade de uma questão, sendo mais informativo e influente para quem está ao redor de $\theta = 0$. Erros nessa área diminuem o valor de θ com maior intensidade do que os acertos podem elevar. Essa variação no valor de θ influencia o IRT, que é a probabilidade de acertos unindo todos os itens da prova utilizando o θ estimado (EAP).

De tal maneira, se aplicarmos a fórmula seguinte:

Equação 8 Formula de cálculo da nota final, após aplicação do TRI₃ PL

$$Nota = \frac{IRT(\theta)}{N} \times 10$$

Com N sendo o número de questões, temos como resultado o valor 4.48. Isso aponta que apesar de acertar 4 questões, o candidato possui um nível habilidade menor que outros, especialmente ao comparar com o aluno So06, que também acertou quatro questões, mas seu IRT quando aplicado a sua nota final, resultou em 5.31.

Para com a prova escrita de Geometria Analítica, vemos um caso interessante, no qual os alunos apresentaram-se com um conhecimento abaixo do esperado no assunto, onde apenas um dos nove candidatos apresentou uma nota acima da média cinco tanto com e sem TRI. Não somente isso, alguns apresentaram zero acertos. Como informado antes, TRI 3PL, se pressupõe sempre a chance de acerto ser maior que zero, por isso notas ao aplicar o TRI mesmo sem acertos, sua nota final não será zero absoluto. Para isso é necessário esclarecer que a nota mínima não é zero e para encontrar o menor valor possível se pode, após o cálculo normal, fazer outro com adicional de um “aluno fantasma”, aplicar um valor θ muito baixo, calcular as probabilidades de $Pi(\theta)$, calcular seu IRT e converter para a nota final. Assim, consegue-se definir um valor mínimo sem que esse afete as probabilidades dos resultados.

Os resultados da prova escrita indicam inicialmente uma familiaridade e facilidade maior dos candidatos com o assunto de História do Brasil do que Geometria Analítica. Podendo especular uma experiência mais agradável dos candidatos com a prova de História do Brasil comparada a Geometria Analítica.

Para as provas em jogos foram apresentados 3 jogos que, em seu total testam as mesmas competências da prova escrita, com o adicional da prova de Lógica. Durante os testes percebeu-se uma familiaridade com os controles utilizados e facilidade de entender os objetivos de cada um. O que se apresenta como uma possível indicação positiva quanto a sua viabilidade na aplicação desses testes como um padrão do dia-a-dia. Os resultados dos testes foram os seguintes.

Tabela 8 Resultados do jogo de História do Brasil, sem TRI

student_ID	c_answers	w_answers	grade
S001	4	4	5.0
S002	7	1	8.7
S003	7	1	8.7
S004	8	0	10.0
S005	7	1	8.7
S006	6	2	7.5
S007	6	2	7.5
S008	7	1	8.7
S009	5	3	6.2

Tabela 9 Resultados do jogo de História do Brasil, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
S001	4	-1.73	0.45	4.53	5.67
S002	7	-0.40	0.56	6.24	7.80
S003	7	0.16	0.70	6.63	8.29
S004	8	0.57	0.78	6.82	8.52
S005	7	0.57	0.78	6.81	8.52
S006	6	-0.44	0.55	6.21	7.77
S007	6	0.54	0.78	6.80	8.51
S008	7	0.57	0.78	6.81	8.52
S009	5	0.12	0.69	6.61	8.27

Tabela 10 Resultados do jogo de Geometria Analítica, sem TRI

student_ID	c_answers	w_answers	grade
S001	1	4	2.0
S002	1	4	2.0
S003	1	4	2.0
S004	2	3	4.0
S005	2	3	4.0
S006	1	4	2.0
S007	2	3	4.0
S008	2	3	4.0
S009	2	3	4.0

Tabela 11 Resultados do jogo de Geometria Analítica, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
S001	1	-0.60	0.75	1.53	3.07
S002	1	-1.20	0.80	1.40	2.80
S003	1	-1.20	0.80	1.40	2.80
S004	2	1.20	0.60	2.36	4.73
S005	2	1.10	0.60	2.30	4.61
S006	1	-1.20	0.80	1.40	2.80
S007	2	1.00	0.60	2.25	4.50
S008	2	0.80	0.62	2.14	4.28
S009	2	0.60	0.65	2.02	4.06

Tabela 12 Resultados jogo de Lógica, sem TRI

student_ID	c_answers	w_answers	grade
S001	4	2	6.6
S002	2	4	3.30
S003	6	0	10.0
S004	6	0	10.0
S005	6	0	10.0
S006	6	0	10.0
S007	6	0	10.0
S008	NaN	NaN	NaN
S009	2	4	3.33

Tabela 13 Resultados jogo de Lógica, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
S001	4	-0.14	0.35	3.99	6.65
S002	2	-1.83	0.50	1.95	3.25
S003	6	2.50	0.60	5.95	9.91
S004	6	2.50	0.60	5.95	9.91
S005	6	2.50	0.60	5.95	9.91
S006	6	2.50	0.60	5.95	9.91
S007	6	2.50	0.60	5.95	9.91
S008	NaN	NaN	NaN	NaN	NaN
S009	2	-1.83	0.50	1.95	3.25

Observando os resultados podemos fazer algumas afirmações e algumas presunções. Primeiramente, vemos que a média das notas da prova escrita de história foi de 6.92 sem o TRI e

5.99 com TRI aplicado. Já no jogo, a média das notas foram 7.88 sem o TRI e 7.925 com o TRI aplicado. Isso aplica incremento de 13.87% nas notas sem TRI e um aumento de 32.4% nas notas com TRI aplicado. Isso apresenta uma clara melhoria na performance entre os 2 meios para história, mas devida a baixa taxa de amostragem, ainda não é possível afirmar pois esse aumento pode ser causado por diversos fatores, como diferenças nas questões, desde sua formulação como itens; diferenças na dificuldade das questões, onde por mais balanceado que a percepção indicasse na hora da formulação das perguntas, a realidade pode diferir, e pela limitação de quantidade de testes, questionamentos sobre a paridade da dificuldade das questões foram ignorados na formulação das perguntas sobre a experiência dos voluntários. Também deve-se apontar a falta de respostas do voluntário S008, que sua planilha não apresentava resultados ao jogo TC, que aborda lógica. Originalmente considerou-se a possibilidade de um bug causando a perda de dados, mas repetidos testes tentando reproduzir esse bug sem sucesso e um caso com outro voluntário do teste remoto, foi confirmado como uma escolha deliberada do voluntário e não uma falha de software.

Para a prova escrita de geometria analítica, as médias da prova escrita foram de 2.22 sem o TRI e 2.32 com TRI aplicado, notavelmente balanceado. Já no jogo, a média das notas foram, 3.11 sem o TRI e 3.74 com o TRI aplicado. O que nos resultados sem TRI, o incremento foi de 40.09% e com TRI aplicado tivemos um incremento de 61.1%, mais uma vez apresentando uma melhora exponencial em comparação a prova escrita. Mas, novamente, esses resultados não podem ser utilizados em um vácuo para chegar a uma conclusão, pois diferenças na formulação de questões e itens podem influenciar a dificuldade final, mesmo que não intencionalmente.

Os resultados dos jogos de lógica servem tanto para exemplificar as possibilidades de conhecimento que podem ser avaliadas através dos jogos, pois, por sua simplicidade pode ser utilizado como um indicador de empenho do participante. E podemos assim destacar os participantes S002 e S009, que ambos obtiveram a nota mínima, podendo através dessa linha de pensamento indicar um menor empenho, porém seus resultados nas outras provas tornam impossível determinar a causa e para isso podemos utilizar o tempo de resposta para a conclusão.

Tabela 14 Tempo de conclusão do jogo TC

student_ID	c_answers	conclusion_time (seconds)
S001	4	56.63
S002	2	108.08
S003	6	138.29
S004	6	31.70
S005	6	25.03
S006	6	86.24
S007	6	69.58
S008	NaN	NaN
S009	2	106.05

Percebe-se apesar de ambos terem as notas mais baixas seus tempos são 2 dos maiores, só atrás do voluntário S003, que concluiu com 138.29 segundos, pouco mais de 2 minutos. Nenhum dos voluntários foi informado que seu tempo de resposta era gravado para garantir a integridade desses valores. De tal maneira pode-se considerar que os erros não foram causados por desinteresse, pelo tempo necessário para conclusão e não tendo ciência do registro do tempo de resposta, não há incentivo para atrasar suas respostas e assim trapacear os resultados.

Todos os nove participantes preencheram o formulário e foram obtidos os seguintes resultados.

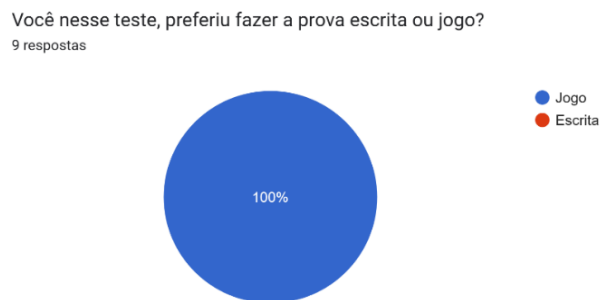


Gráfico 3 Preferência entre prova escrita e jogo

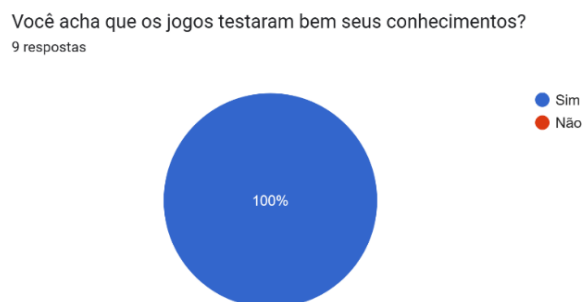


Gráfico 4 Opiniões da efetividade dos jogos

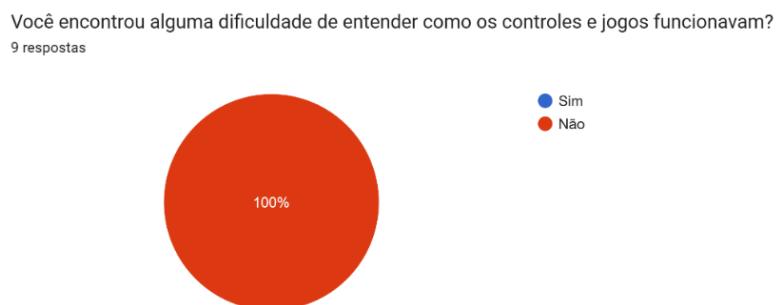


Gráfico 5 Avaliação da compreensão dos comandos

Com esses gráficos, inicialmente pode se observar uma satisfação com o teste aplicado, e pelo ponto de vista dos voluntários uma adequação às necessidades e facilidade de compreensão e uso para os testes. Também se percebe, no ambiente de teste, uma preferência aos jogos em comparação direta a prova escrita. Quando levantada a possibilidade de uma escolha para o uso de jogos ou provas tradicionais no ambiente real de estudo, 8 dos 9 voluntários escolheriam o jogo. Isso ajuda a demonstrar que uma solução mais lúdica pode ser apreciada, mas não é a preferência de todos. Isso também colabora com a melhora em performance presente nos resultados das provas quando comparadas a prova escrita. Os voluntários também descreveram de forma aberta suas experiências com os jogos, suas impressões foram positivas em sua maioria, mas também foi levantado por um dos voluntários um desconforto por não estar mais familiarizado com alguns dos tópicos.



Gráfico 6 Preferência entre prova escrita e jogo caso possível escolher

A ludificação do sistema escolar, como dito previamente, já é uma realidade e continua a se expandir com o uso de REDs. Com a constante busca pela digitalização do ambiente de estudo, com uso de computadores e tablets, a migração para um ambiente digital de avaliação já foi provada necessária, especialmente durante o período da COVID-19, com a obrigatoriedade de isolamento doméstico e migração, mesmo que temporária, para ensino a distância (EAD). Apesar da baixa taxa de amostragem para esse estudo, é possível iniciar um debate maior para a mudança do meio de avaliação, e argumentar que há um interesse nessa mudança.



Gráfico 7 Avaliação da efetividade do sistema atual de provas nas escolas.

Quando observando as opiniões dos voluntários para com o sistema de provas atual aplicado em escolas, por não possuir um meio absoluto, foi considerado cinco como indiferente pendendo a insatisfação e seis como indiferente pendendo a satisfação. Dessa forma se para cada posição se distanciando de o centro aumentarmos em um, sendo o centro zero, temos a seguinte imagem.

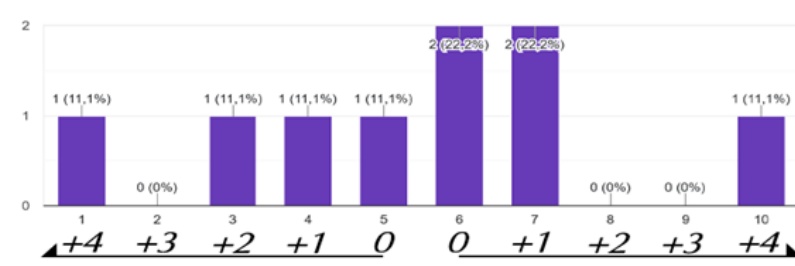


Gráfico 8 Quantificação dos resultados do Gráfico 7

Quando somamos os valores de cada lado e depois tiramos a média, temos 1.75 para os voluntários insatisfeitos, e 1 para os voluntários satisfeitos. Indicando um sentimento mais forte dentre os voluntários insatisfeitos. Quando pedido para que descrevam suas experiências com provas escritas, uma constante de pontos foi notada, onde cinco dos voluntários descreveram negativamente, apontando diretamente problemas com ansiedade, tensão e ou frustração. Também foram mencionados em três casos a presença de neurodivergência, em alguns, descritos como Autismo e TDAH. Apesar, um dos voluntários levantou um ponto importante, que nem todas as competências podem traduzir bem para um formato de jogo. Para isso, uma abordagem híbrida ainda sim completamente digital poderia ser uma solução, com jogos para matérias mais práticas, como química, ou que podem ser adaptadas de forma lúdica com uma facilidade, como língua portuguesa, história e geografia, e uma prova mais tradicional para matérias como matemática.

Mas com tais relatos, é visível um descontentamento com o formato atual, e com a maior conscientização de condições de neurodiversidade no dia a dia, a insistência em um formato arcaico é antagônico as necessidades desse grupo.

Você já teve ou tem episódio de Ansiedade com Provas?

9 respostas

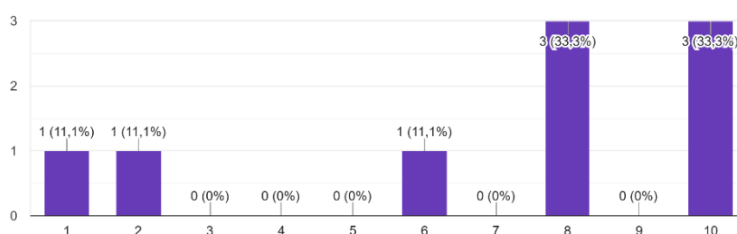


Gráfico 9 Presença de AP nos voluntários

A alta presença de reclamações envolvendo ansiedade, é reforçada quando se pediu para que quantificassem de zero a dez, o a frequência de episódios de AP, caso tivessem, e como demonstrado no gráfico 9, seis dos nove voluntários avaliaram como algo muito frequente. Em “*Worry and emotionality as components of test anxiety: replication and further data*” afirma um impacto maior de preocupação em comparação a emotividade (Doctor & Altman, 1969). Nisso torna-se válida a noção de que a performance acadêmica desses voluntários é afetada por AP.

Como abordado previamente, há estudos, como “*The effect of casual videogames on anxiety, depression, stress, and low mood: A systematic review*” que indicam que os usos de jogos podem não ser uma solução por si só, mas aliado com outros tratamentos para ansiedade, JDC podem auxiliar no tratamento e redução de ansiedade (Pine, Fleming, McCallum, & Sutcliffe, 2020). Isso traz a hipótese de jogos como prova ajudarem na redução de episódios de AP no ambiente de estudo.

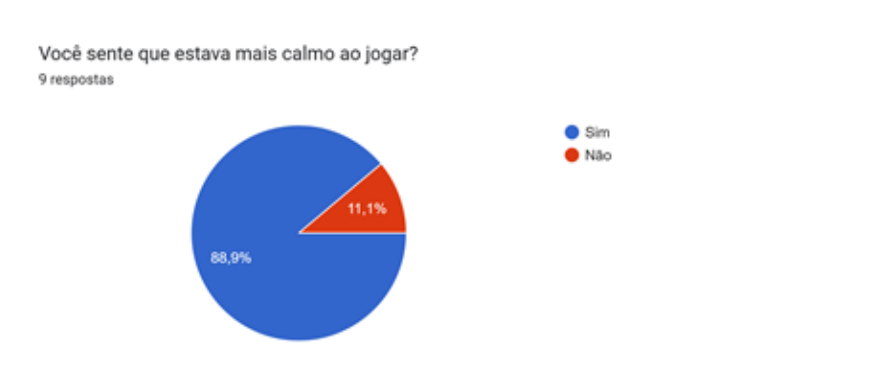


Gráfico 10 Taxa de redução de estresse ao jogar

Quando questionados sobre se sentiam-se mais calmos após jogar jogo, oito dos nove voluntários, afirmaram que sim, ao pedir para que elaborassem as opiniões ganharam mais detalhes, e mais questionamentos também foram levantados. Muitos mencionam que a apresentação como jogo, com as imagens e sons, ajudava a dissociar do ato de fazer uma prova, auxiliando na redução da ansiedade. Como descrito por um dos voluntários, a sensação dele não era da pressão de um bom resultado, mas sim empenhar-se ao máximo pela satisfação de jogar. Outro levanta o fato que se estivesse valendo pontuação real para seus estudos, teria se sentido mais tenso, estipulando que poderiam ser igualmente estressantes, mas por não ser algo que vai afetar ativamente seu futuro não é o mesmo, ainda sim ressalta que acredita que provas como jogos seriam mais divertidos.



Gráfico 11 Avaliação da efetividade em mascarar as provas através dos jogos

É necessário evidenciar, que quando requeridos a quantificar quanto eles sentiram que estavam fazendo uma prova ao jogar, percebe-se que apesar da ludificação, a grande maioria ainda se sentiu como fazendo uma prova, o que em um ambiente de prova real, os relatos anteriores podem divergir da realidade.

Após leituras do formulário, foi decidido que poderia ser refinado, tornando perguntas mais sucintas e reorganizando a ordem de algumas dessas perguntas. Essa mudança foi posteriormente aplicada ao teste remoto.

5.2 Resultados Teste Remoto

Para o teste remoto, contamos com uma amostra de 8 candidatos, esses possuindo um perfil de estudo e/ou profissional mais abrangente, o que pode afetar performance em alguns dos assuntos. Igualmente ao teste presencial, posterior aos testes, foi dado a cada candidato uma identificação a qual serão referenciados como neste trabalho.

Pela natureza da distribuição desse teste, contato direto com todos os candidatos era impossível, por isso não é possível afirmar através das interações quaisquer inquietudes com as competências aplicadas. Com exceção de um caso especial, que será abordado com maior profundidade a diante, o qual o padrão de suas respostas pode ser interpretado como uma falta de familiaridade ou rejeição a competência.

Como explicado na metodologia, por limitações da ferramenta Google Forms, foi definida uma ordem na qual a competência das provas escritas deve ser seguida, para via de organização, os resultados a seguir serão apresentados na mesma ordem.

Tabela 15 Resultados da prova remota de História do Brasil, sem TRI

student_ID	c_answers	w_answers	grade
SR001	4	4	5.0
SR002	3	5	3.7
SR003	8	0	10.0
SR004	3	5	3.7
SR005	4	4	5.0
SR006	5	3	6.25
SR007	5	3	6.25
SR008	3	5	3.7

Tabela 16 Resultados da prova remota de História do Brasil, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
SR001	4	-0.484	0.354	0.635	6.35
SR002	3	-1.283	0.536	0.408	4.08
SR003	8	1.158	0.621	0.958	9.58
SR004	3	-0.968	0.423	0.499	4.99
SR005	4	-1.194	0.424	0.431	4.31
SR006	5	-0.094	0.379	0.723	7.23
SR007	5	-0.436	0.342	0.647	6.47
SR008	3	-1.246	0.408	0.417	4.17

Tabela 17 Resultados da prova remota de Geometria Analítica, sem TRI

student_ID	c_answers	w_answers	grade
SR001	4	1	8.0
SR002	4	1	8.0
SR003	1	4	2.0
SR004	2	3	4.0
SR005	0	5	0.0
SR006	4	1	8.0
SR007	2	3	4.0
SR008	0	5	0.0

Tabela 18 Resultados da prova remota de Geometria Analítica, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
SR001	4	1.70	0.33	3.95	7.90
SR002	4	1.70	0.33	3.95	7.90
SR003	1	-3.00	2.59	1.26	2.52
SR004	2	-0.38	0.49	2.00	4.00
SR005	0	-3.50	3.95	1.25	2.50
SR006	4	1.70	0.33	3.95	7.90
SR007	2	-0.38	2.0	2.00	4.00
SR008	NaN	NaN	NaN	NaN	NaN

Para as provas escritas, percebe-se uma variação grande entre o nível de conhecimento entre alguns dos voluntários. Nisto pode destacar os voluntários SR005 e SR008, que tiveram as notas mais baixas nas duas provas, onde o SR008 não respondeu as questões de Geometria Analítica, o que de início indica uma rejeição a esse assunto, e o mesmo poderá ser observado adiante na apresentação e análise dos resultados dos jogos, onde o mesmo se repetiu com o jogo BC, que não apresentava resultados registrados. Podendo considerar como provável uma rejeição a competência.

Tabela 19 Resultados do jogo remoto de História do Brasil, sem TRI

student_ID	c_answers	w_answers	grade
SR001	7	1	8.75
SR002	5	3	6.25
SR003	6	2	7.50
SR004	7	1	8.75
SR005	7	1	8.75
SR006	5	3	6.25
SR007	6	2	7.50
SR008	7	1	8.75

Tabela 20 Resultados do jogo remoto de História do Brasil, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
SR001	7	0.35	0.59	6.83	8.54
SR002	5	-0.43	0.52	5.35	6.69
SR003	6	-0.17	0.56	5.91	7.39
SR004	7	0.44	0.57	6.95	8.69
SR005	7	0.35	0.59	6.83	8.54
SR006	5	-0.38	0.50	5.47	6.84
SR007	6	-0.054	0.54	6.14	7.69
SR008	7	0.32	0.60	6.79	8.49

Tabela 21 Resultados do jogo remoto de Geometria Analítica, sem TRI

student_ID	c_answers	w_answers	grade
SR001	3	2	6.0
SR002	3	2	6.0
SR003	3	4	6.0
SR004	3	2	6.0
SR005	1	2	2.0
SR006	3	2	6.0
SR007	1	4	2.0
SR008	NaN	NaN	NaN

Tabela 22 Resultados do jogo remoto de Geometria Analítica, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
SR001	3	0.16	0.83	2.24	4.48
SR002	3	0.29	0.86	2.35	4.69
SR003	3	0.18	0.90	2.26	4.51
SR004	3	0.29	0.86	2.35	4.69
SR005	1	-0.57	0.77	1.80	3.61
SR006	3	0.29	0.86	2.35	4.69
SR007	1	-0.43	0.75	1.87	3.74
SR008	NaN	NaN	NaN	NaN	NaN

Igualmente aos testes presenciais, ao observar as tabelas de História do Brasil, percebeu-se uma melhora entre os resultados da prova escrita e do jogo. Os voluntários obtiveram uma média de notas com TRI aplicado de 5.89 na escrita contra 7.85 no jogo, isso é um aumento de 33.28%, quando jogando. Isso vem de acordo com o primeiro teste, a qual tinha uma diferença de 32.4%, uma diferença menor que um por cento. Apesar, um ponto de discrepância foi observado quando observamos a diferença entre o número de acertos entre prova escrita e jogo entre os dois testes. No primeiro teste vimos em grande maioria um aumento no número de acertos, onde o maior acréscimo foi de três acertos pelo voluntário S004, mas a maior reincidência foi de um acerto a mais, isso valendo para três dos nove participantes com o participante s007 tendo uma redução em um no total de acertos. Para o teste remoto, vemos dois dos voluntários, esses sendo SR004 e SR008 tendo um acréscimo de quatro acertos. O que poderia indicar que a dificuldade das questões pode não ser equilibrada como antes imaginado. Porém para chegar a uma conclusão seria necessária uma taxa de amostragem maior para os testes, tal como uma bateria de testes mais longa e obter um feedback direto dos voluntários para suas opiniões quanto a dificuldade dentre a prova e o jogo.

Para o cálculo das médias os valores, que na tabela estão registados na tabela como NaN, do inglês “Not a Number”, foi considerado como uma nota zero, mesmo, o TRI 3PL sempre pressupõe uma

chance de acerto acima de zero, por menor que seja, nunca resultando em um zero absoluto a pontuação. Mas da mesma maneira que um aluno que não fez uma prova não pode ser avaliado, o voluntário também não pode receber uma nota. Mas pela possibilidade de uma prova de uma competência ser respondida, mas o jogo da mesma competência ser ignorado, e vice-versa.

Definido como esses dados serão tratados, podemos então ver que a média das notas da prova escrita com TRI aplicado é de 4.59, enquanto o jogo com TRI aplicado possui apenas, 4.34 de média. Sendo a primeira queda visível dos resultados entre a prova escrita e em jogo, com uma queda total de 5.4%. Quando comparados com os resultados do teste presencial, com o TRI aplicado, obteve-se uma melhora de 61% e sendo uma discrepância notória. Pela falta de mais testes, é impossível definir uma explicação para tal, mas pode-se teorizar quais fatores podem ter causado tal. Um dos principais pode ser o ambiente de prova, com a prova escrita, cada aluno possuía papel e caneta para solucionar as questões. Porém apesar de destacar a possibilidade do uso, não é possível garantir que todos os voluntários tenham acesso a isso no momento do teste. De tal forma afetando a capacidade de resolução em comparação aos voluntários do teste presencial. Também deve-se notar que essa questão levantada pode ser uma das ou principal causadora da rejeição da competência de geometria analítica pelo voluntário So08.

Tabela 23 Resultados do jogo remoto de Lógica, sem TRI

student_ID	c_answers	w_answers	grade
SR001	6	0	10.0
SR002	2	4	3.33
SR003	6	0	10.0
SR004	4	2	6.66
SR005	6	0	10.0
SR006	2	4	3.33
SR007	6	0	10.0
SR008	4	2	6.66

Tabela 24 Resultados do jogo remoto de Lógica, com TRI

student_ID	c_answers	θ (EAP)	SE(θ)	Acertos Esperados	grade
SR001	6	3.20	0.40	5.927	9.88
SR002	2	-2.30	0.65	2.235	3.72
SR003	6	3.20	0.40	5.927	9.88
SR004	4	-0.45	0.55	4.047	6.74
SR005	6	3.20	0.40	5.927	9.88
SR006	2	-2.30	0.65	2.235	3.72
SR007	6	3.20	0.40	5.927	9.88
SR008	4	-0.45	0.55	4.047	6.74

Para o jogo de lógica, sua presença se provou menos necessária, pois sua principal função foi explicada e demonstrada anteriormente e não ocorreu nenhuma situação que requeresse seus resultados para melhor compreensão.

Vemos nas tabelas que, novamente, dois alunos ficaram com o número mínimo de acertos, e para via de comparação com os resultados do primeiro teste, temos a seguir a seguinte tabela, com o tempo de conclusão de cada voluntário.

Tabela 25 Tempo de conclusão do jogo remoto TC

student_ID	c_answers	conclusion_time (seconds)
SR001	6	40.72
SR002	2	85.70
SR003	6	41.64
SR004	4	125.62
SR005	6	55.26
SR006	2	74.24
SR007	6	68.92
SR008	4	32.63

Com os tempos na Tabela 25, quando comparados com os registados na Tabela 14 vemos uma uniformidade entre os resultados, onde alguns resultados, como dos voluntários S001, S004 e S005 do teste presencial como os SR001, SR003 e SR005 obtiveram nota máxima com tempos abaixo de um minuto enquanto no outro lado do espectro os candidatos com menos acertos levaram um minuto ou mais para concluir. Podemos levar com quanto maior o tempo, maior foi a dificuldade de responder, mas não significa que sua nota será sempre baixa.

Para as avaliações pessoais, todos os voluntários preencheram e responderam cada uma das treze perguntas, os pontos mais importantes foram os seguintes. No geral, alguns pontos se repetiram, como a preferência geral na prova em jogo à prova escrita, mas diferentemente do primeiro teste houve ressalvas, algumas por falha na comunicação das intenções da pesquisa, outras sobre falhas na plataforma. Começando com as falhas na plataforma, ao menos três comentários foram feitos nas elaborações de perguntas quanto a um bug no jogo BC, que impedia o carregamento de questões e dessa forma, de prosseguimento do teste. Esse bug foi identificado e replicado após análise dos resultados. Por uma falha na lógica do carregamento das questões, este é feito uma única vez ao executável do jogo ser iniciado, porém para evitar que as perguntas aleatórias se repitam, as questões guardadas em um “array List” são removidas após serem escolhidas, assim quando o jogador tenta novamente executar o jogo, não há mais informações sendo carregadas na “array List” assim causando um estado que o jogo é incapaz de prosseguir. Outro bug também encontrado, mas não foi possível replicar após seu relato foi com a UI, na qual sua escala pode quebrar e impedir que elementos carreguem em suas posições corretas. Apesar de não ser possível replicar para afirmar, esse erro havia sido identificado no primeiro teste, durante seu

desenvolvimento a plataforma principal do jogo NPC, foi toda feita com foco no aspecto de tela 4:3 que garantia suporte em monitores mais antigos que ainda seguiam esse formato, tal como poderia ser aplicado a monitores mais recentes. Através disso, a plataforma possui um método que na inicialização do jogo ele força o jogo a rodar em “tela cheia”, mas mantendo a escala e resolução original. Por alguma falha não encontrada nesse método, pode acontecer de ele atualizar a resolução para a mesma do monitor, causando uma quebra no canvas da Unity. Porém apesar da seriedade de ambos os casos parecer alta, sua solução é simplesmente fechar o jogo teste e abrir seu executável novamente, que foi notado pelos próprios voluntários.

Você acha que os jogos testaram bem seus conhecimentos?
8 respostas

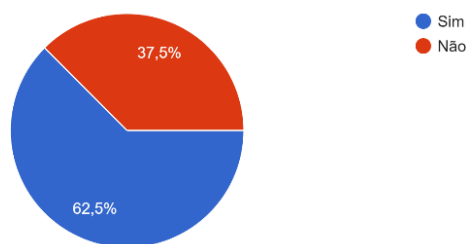


Gráfico 12 Opiniões da efetividade dos jogos no teste remoto

Outro problema notado pelos participantes foi a inadequação do meio remoto para a avaliação dos níveis de ansiedade. Onde diferente do primeiro teste, que foi feito presencialmente, em um laboratório do Centro Universitário Farias Brito, em Fortaleza, no estado Ceará, com o teste sendo aplicado poucos dias após uma semana de provas, que por si mantinha as memórias desse período mais presentes, o teste remoto foi feito em ambientes diferentes da escolha dos voluntários, podendo apenas contar com a memória deles para comparar como se sentiram, e mesmo, a sensação de fazer a prova como um jogo pode não ser precisa por não haver um real impacto de sua performance no seu dia a dia, diferente de uma prova acadêmica que pode afetar a todo seu futuro profissional. Isso por si, ao observar os resultados, indica um nível de influência sobre eles. E como pode se observar no gráfico 12, vemos uma presença de insatisfações sobre a efetividade dos jogos utilizados para testar o conhecimento dos participantes.

Quando questionados sobre suas opiniões ao sistema atual de provas aplicado em escolas, como se pode notar pelo gráfico 13 a seguir, houve uma divisão entre ambos como uma tendência maior a insatisfação comparada à satisfação.

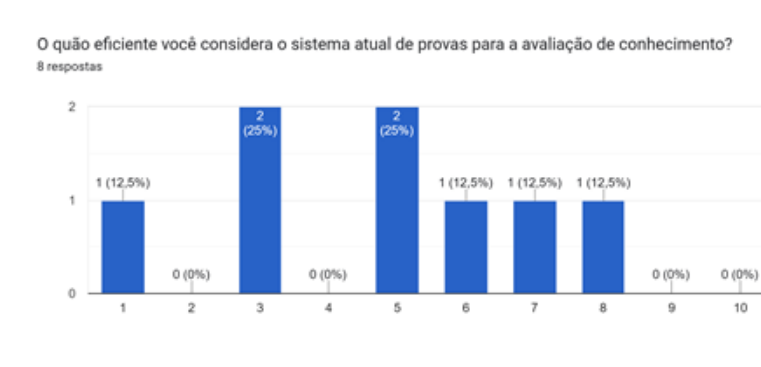


Gráfico 13 Avaliação do teste remoto da efetividade do sistema atual de provas nas escolas

Nele, utilizando cinco como mediano pendendo para insatisfação e seis como mediano pendendo a satisfação, podemos perceber que três dos voluntários sentiam-se indiferentes, mas com dois pendendo a estarem mais insatisfeitos, enquanto um ficou indiferente, pendendo à satisfação, e dois estão satisfeitos. Quando aplicamos a mesma lógica no tratamento dos resultados aplicada no primeiro teste temos a seguinte imagem:

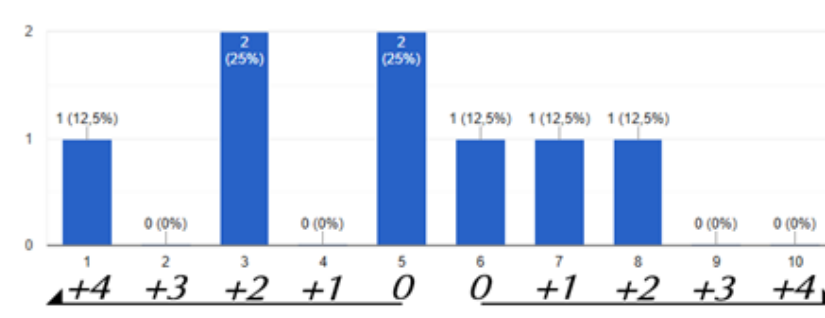


Gráfico 14 Quantificação dos resultados do Gráfico 13

A média dos voluntários insatisfeitos foi de 2.6 e enquanto dos satisfeitos foi 1, dessa forma podemos afirmar que a insatisfação dos voluntários que a sentem é maior que o mesmo pelos voluntários satisfeitos. Quando observamos os relatos vemos uma diferente série de motivos quando comparado ao primeiro teste, neste estiveram presentes reclamações a exaustão física, como o artigo “*Examining burnout profiles in university students: A person-centered approach. Motivation and Emotion*”, nota explicitamente que em seus testes alunos reportaram exaustão após provas, isso por sua vez é descrito no mesmo como um dos componentes chave de burnout em estudantes (Rentzios, Karagiannopoulou, & Ntritsos, 2025). Outras reclamações focam na incapacidade de provas tradicionais asertarem a retenção de conhecimento a longo prazo, de tal modo que desincentiva o real aprendizado e incentiva focar na memorização a curto prazo.

Esse sentimento de insatisfação com o sistema de provas tradicional também se reflete de forma consistente quando lhes é dado a opção entre os dois. Como vemos no gráfico 13, utilizando 5 e 6

como uma média, e atribuindo esses valores a indiferença temos a exata mesma taxa representada no gráfico 15 a seguir, no qual dois dos oito voluntários preferindo provas escritas.

Se você pudesse escolher entre jogar um jogo ou fazer uma prova escrita, qual você preferiria?
8 respostas

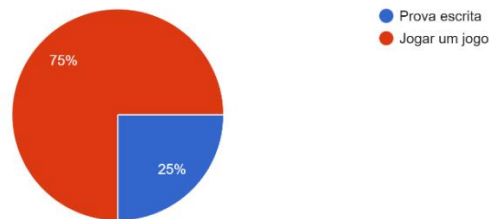


Gráfico 15 Preferência no teste remoto entre prova escrita e jogo se possível escolher

Porém quando questionados quanto a suas crenças de que suas performances acadêmicas teriam melhoras quando crianças se provas fossem aplicadas como jogos eletrônicos, como podemos observar no gráfico 16, somente 50% dos voluntários acreditam em alguma melhora.

Você acredita que quando criança sua performance nas provas seria melhor se elas fossem como jogos?
8 respostas

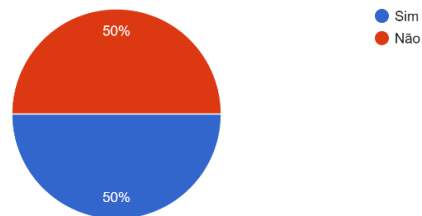


Gráfico 16 Opiniões no teste remoto sobre performance com provas como jogo quando criança

O que a princípio aparenta uma mudança grande em comparação ao primeiro teste, mas ao analisarmos o gráfico 15 a seguir podemos tirar algumas conclusões de porquê essa diferença é tão grande.

Em uma escala de 1 à 10. Você já teve ou tem episódio de Ansiedade com Provas?
8 respostas

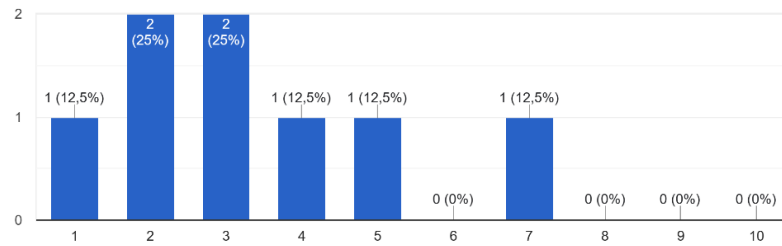


Gráfico 17 Presença de AP nos voluntários do teste remoto

Se compararmos o gráfico 17 ao gráfico 9, temos uma inversão nos casos, onde o teste presencial contou com uma maioria notória de alunos que relataram episódios constantes de AP, enquanto neste já temos uma quantidade menor. Se quantificarmos utilizando os próprios valores apresentados e tirando a média deles, temos uma média de 7.33 para o teste presencial e 3.37 para o remoto, isso mostra que AP era presente em mais que o dobro das vezes com os voluntários do teste presencial. Segundo os resultados presentes em “*Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. Journal of Affective Disorders*” percebe-se uma estabilidade através dos estágios formativos, com manifestações na infância mantendo-se na adolescência e idade adulta (von der Embse, Jester, Roy, & Post, 2018). Indicando que para os participantes do teste remoto, pela baixa presença de AP em seus relatos, sua performance acadêmica de fato pode não ser afetada de maneira significativa como para os participantes do teste presencial.

Você sente que estava mais calmo ao jogar em comparação com a prova tradicional?
8 respostas

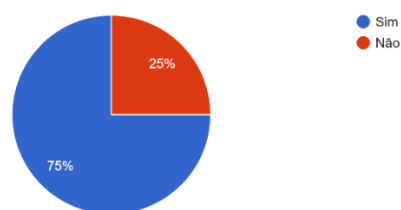


Gráfico 18 Taxa de redução de estresse ao jogar no teste remoto

Apesar do relato visto no gráfico 18, como dito previamente, pelas inadequações dos ambientes de teste remoto, aliados a falta de regência com o ambiente de provas acadêmicas, seus relatos apontam que apesar de se sentirem de fato mais calmos ao jogar, não há a pressão do ambiente de provas, o que torna essa comparação demasiada imprecisa.

Ao final, quando questionados sobre a crença na possibilidade de utilização de jogos digitais como método de avaliação, todos responderam positivamente. Seus relatos apontaram a

necessidade de uma boa execução e a criação de jogos mais focados para cada matéria, dessa maneira apontando um dos maiores empecilhos que foram o desenho de ferramentas dedicadas a esse estudo.

Para com os custos de implementação e manutenção, utilizando o escopo original pretendido da plataforma NPC, foram listados a necessidade do desenvolvimento dos jogos, que traz consigo, ferramentas, como motores gráficos, ferramentas de edição de imagem e som, pessoal como Game Designer, Programadores, Artistas Visuais e Sonoros, e um consultor pedagógico, esse custo variando com o escopo dos jogos desenvolvidos; tal como a necessidade de uma plataforma local ou online a qual os professores possam customizar os jogos podendo escolher quais serão utilizados e quais serão as questões usadas e apresentar de forma concisa os resultados para que possam ser interpretados por ele, apresentando a necessidade de um banco de dados, que já em muitas escolas já é algo presente e que pode ser integrado tendo assim uma redução de custos. O equipamento necessário para execução de uma prova pode variar dependendo do escopo dos jogos, o utilizado por esse estudo foi pensado de início para ser capaz de ser executado no máximo de máquinas possíveis. Mas necessita destacar que muito dos custos apresentados são especulativos pois não houve uma avaliação destes no desenvolvimento do projeto NPC, e por sua aplicação neste projeto não houve quaisquer gastos diretos com seu desenvolvimento e aplicação, apenas na impressão das provas escritas.

6 Capítulo 6

Conclusão e Trabalhos Futuros

6.1 Conclusão

A utilização de jogos no meio acadêmico para avaliações ainda é uma área pouco explorada, mas trabalhos nessa área focam se tornarem uma ferramenta para combate de AP, com cada abordagem tendo seus méritos, mas normalmente batendo de frente com custos e complexidades de desenvolvimento e implementação. Para isso uma solução simplificada que traz o ar lúdico dos jogos eletrônicos, mas mantendo estruturas já presentes em provas para uma melhor adaptabilidade a sistemas já presentes em escolas ao redor do mundo.

Através dos dois testes percebeu-se que mesmo com uma taxa de amostragem menor que a inicialmente desejada, na maior parte dos casos, houve melhoria no desempenho dos voluntários quando comparando seus resultados entre as provas escritas e em jogos. Porém por um desenvolvimento apressado e utilização de uma plataforma pré-existente que necessitou ser adaptada para este estudo, os resultados devem ser considerados com cautela, pois intuito e realidade não são sinônimas. O processo de desenvolvimento conturbado e necessidade de aderir a decisões de design arbitrárias do projeto inicial da plataforma, junto a escolhas de competências para os testes que se provaram equivocadas, podem ter afetado a qualidade dos dados coletados.

Foi notado que com o grupo que participou presencialmente, devido a seu ambiente e regência de avaliações a sério, e presença maior de participantes com frequentes episódios de AP, seus relatos apontam um otimismo e até possível efetividade do uso desses jogos para redução de AP, mas não se pode tirar nenhuma conclusão a definitivo considerando a baixa taxa de amostragem, mas abre caminho para um debate sobre essa abordagem e possibilidade de estudos futuros.

A aplicação dos testes remotos provou-se não adequada, com uma falta de controle sobre o ambiente de aplicação tal como de facilidade de atendimento para quaisquer dúvidas ou problemas, sendo caótica demais para ser considerada viável para testes futuros. Seus resultados mesmo encriptados mostraram, na secção de opiniões dos participantes, uma influência causada por equívocos na sua formulação, como a pouca explicação dos objetivos do estudo aos participantes como falta de informações sobre o funcionamento dos jogos. Quanto a redução nos resultados de Geometria Analítica, principalmente ao comparar com a melhora exponencial nos resultados, abrem a questão dos impactos causados pelos problemas técnicos da plataforma, como pode ser uma indicação de inadequação da competência a este método de teste, porém são necessários testes mais extensivos para chegar a uma conclusão.

Por sua complexidade a aplicação manual de TRI 3PL é inviável, para isso são necessárias ferramentas especializadas, para este estudo foi utilizada a ferramenta ChatGPT, a qual tratou os

dados da maneira explicada. Como a ferramenta descrita, pode tanto cometer erros como sua lógica pode mudar entre testes, os mesmos *prompts* foram utilizados entre os testes, ajustando valores pertinentes, e os mesmos dados foram tratados múltiplas vezes para assegurar os resultados apresentados. Diferente da solução aplicada nesse estudo, recomenda-se uma integração direta da plataforma dos jogos com a ferramenta ChatGPT por sua API ou outra ferramenta da análise de dados para tratamento e apresentação dos resultados sem necessidade de enviar *prompts* e arquivos manualmente, porém isso se junta aos custos para manter o sistema operacional.

O uso do TRI 3PL provou-se essencial para a garantia da autenticidade dos dados coletados, tal como melhor observar a real performance dos voluntários. Aliado a dados extras capturados, que para esse estudo, foi o tempo de resposta dos jogos e suas questões, os resultados apresentados após tratados TRI 3PL podem ser mais bem interpretados pelo professor, como se pode observar no capítulo 5, ao empregar os tempos de conclusão do jogo TC, vistos na Tabela 25, para chegar a uma interpretação dos resultados.

Seus custos não puderam ser avaliados além de especulações o que traz a necessidade de mais estudos, estes com uma nova plataforma desenvolvida do zero que não somente permitiria uma visão mais abrangente dos custos, como também permitiria uma adequação de design dos jogos para as competências escolhidas.

6.2 Soluções e Trabalhos Futuros

Para trabalhos futuros deve ter em consideração o processo que foi aplicado neste para poder evitar os problemas causados por falta de tempo e caminhos cortados em decorrência de tal. Incluindo abandono da aplicação de testes remotos.

Primeiramente é o desenvolvimento de uma plataforma e jogos novos de seu início com o intuito para uso nesses testes e acompanhamento de um pedagogo, não a reutilização de um projeto prévio que por similar em objetivo, não levou em conta estudos ou pedagogia nas decisões de design. E integrar na plataforma uma API capaz de análise de dados, seja essa ChatGPT ou outra, para maior facilidade e agilidade no tratamento dos resultados.

Também se percebe necessária a reformulação de dados coletados, para dar ao professor uma melhor atenção ao progresso do aluno, de tal forma que seja compatível com sistemas escolares já presentes e não seja invasivo à privacidade dos alunos.

Para a aplicação dos testes considera-se a comunicação direta com uma instituição de ensino que esteja disposta a aplicar esses testes com seus alunos mais jovens, e para garantir sua eficiência em situação de avaliação real, que a instituição aplique esses jogos como real meio de avaliação

durante um determinado período, podendo essa aplicação ser contínua ou alternada para ver as mudanças em tempo real com maior agilidade.

7 Bibliografia

- Arieli-Attali, M., Ward, S., Thomas, J., Deonovic, B., & von Davier, A. A. (2019, 04 26). The Expanded Evidence-Centered Design (e-ECD) for Learning and Assessment Systems: A Framework for Incorporating Learning Goals and Processes Within Assessment Design. *Frontiers in Psychology, 10*, p. 853. doi:10.3389/fpsyg.2019.00853
- Balt, M., Börnert-Ringleb, M., & Orbach, L. (2022). Reducing Math Anxiety in School Children: A Systematic Review of Intervention Research. *Frontiers in Education, 7*. doi:10.3389/educ.2022.798516
- Bressler, D. M., & Bodzin, A. M. (2013). A Mixed Methods Assessment of Students' Flow Experiences during a Mobile Augmented Reality Science Game. *Journal of Computer Assisted Learning, 29*, pp. 505-517. doi:10.1111/jcal.12008
- Conselho Nacional de Secretários de Educação. (2022, 03 14). *Uso de tablets na sala de aula é aliado no processo de ensino-aprendizagem*. Retrieved from Consed: <https://www.consed.org.br/noticia/saiba-como-o-uso-de-tablets-na-sala-de-aula-e-aliado-no-processo-de-ensino-aprendizagem>
- Doctor, R. M., & Altman, F. (04 de 1969). Worry and emotionality as components of test anxiety: replication and further data. *Psychological Reports, 24*(2), pp. 563-568. doi:10.2466/pro.1969.24.2.563
- Ergene, T. (2003). Effective interventions on test anxiety reduction: A meta-analysis. *School Psychology International, 24*(3), pp. 313-328. doi:10.1177/01430343030243004
- Hadjerrouit, S. (2010). A Theoretical Framework to Foster Digital Literacy: The Case of Digital Learning Resources. *Key Competencies in the Knowledge Society* (pp. 144-154). Brisbane: Springer. doi:https://doi.org/10.1007/978-3-642-15378-5_14
- Ifenthaler, D., & Yau, J. Y. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Education Tech Research and Development, 68*, pp. 1961-1990. doi:10.1007/s11423-020-09788-z
- McDonald, A. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology, 21*, pp. 89-101. doi:<https://doi.org/10.1080/01443410020019867>
- Neto, D. (s.d.). *Turminha FB*. Fonte: Damasio Neto: <https://damasioneto.artstation.com/projects/8lD45G>
- Ohannessian, C. M. (2018). Video game play and anxiety during late adolescence: The moderating effects of gender and social context. *Journal of Affective Disorders, 226*, pp. 216-219. doi:<https://doi.org/10.1016/j.jad.2017.10.009>
- Pine, R., Fleming, T., McCallum, S., & Sutcliffe, K. (2020, 08 13). The effect of casual videogames on anxiety, depression, stress, and low mood: A systematic review. *Games for Health Journal, 9*, pp. 255-264. doi:10.1089/g4h.2019.0132
- Rentzios, C., Karagiannopoulou, E., & Ntritsos, G. (2025). Academic Emotions, Emotion Regulation, Academic Motivation, and Approaches to Learning: A Person-Centered Approach. *Behavioral Sciences, 15*(7), p. 900. doi:<https://doi.org/10.3390/bs15070900>
- Scholten, H., Malmberg, M., Lobel, A., Engels, R. C., & Granic, I. (2016). A Randomized Controlled Trial to Test the Effectiveness of an Immersive 3D Video Game for Anxiety Prevention among Adolescents. *PLOS ONE, 11*(1), p. e0147763. doi: <https://doi.org/10.1371/journal.pone.0147763>
- Shute, V. J. (2023). The History of Stealth Assessment and a Peek Into Its Future. In *Games as Stealth Assessments*. IGI Global. doi:10.4018/979-8-3693-0568-9.ch001
- Shute, V. J., & Ke, F. (2012). Games, Learning, and Assessment. In D. Ifenthaler, D. Eseryel, X. Ge, D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives* (pp. 43-58). New York: Springer. doi:10.1007/978-1-4614-3546-4_4
- Shute, V., & Ventura, M. (2013). *Measuring and Supporting Learning in Video Games*. Cambridge, MA: The MIT Press. doi:<https://doi.org/10.7551/mitpress/9589.001.0001>
- Von der Embse, N., Barterian, J., & Segool, N. (2013). Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000-2010. *Psychology in the Schools, 50*, pp. 57-71. doi:10.1002/pits.21660
- von der Embse, N., Jester, D., Roy, D., & Post, J. (02 de 2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, pp. 483-493. doi:10.1016/j.jad.2017.11.048

Zeidner, M. (2014). Anxiety in Education. Em R. Pekrun, L. Linnenbrink-Garcia, R. Pekrun, & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 265-288). Boca Raton, FL: Taylor & Francis.