



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

ARTHUR SILVA BASTOS

**O USO DE TÉCNICAS DE TESTES COM USUÁRIO COMBINADAS COM
VISUALIZAÇÃO DE DADOS PARA APERFEIÇOAR UM JOGO MASOCORE DE
PLATAFORMA 2D**

FORTALEZA

2021

ARTHUR SILVA BASTOS

O USO DE TÉCNICAS DE TESTES COM USUÁRIO COMBINADAS COM
VISUALIZAÇÃO DE DADOS PARA APERFEIÇOAR UM JOGO MASOCORE DE
PLATAFORMA 2D

Dissertação apresentada à coordenação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Ciência da Computação. Área de concentração: Computação Gráfica.

Orientadora: Profa. Dra. Emanuele Marques dos Santos.

Coorientador: Prof. Dr. George Allan Menezes Gomes.

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

B326u Bastos, Arthur Silva.

O uso de técnicas de testes com usuário combinadas com visualização de dados para aperfeiçoar um jogo masocore de plataforma 2D / Arthur Silva Bastos. – 2021.
124 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2021.

Orientação: Prof. Emanuele Marques dos Santos.

Coorientação: Profa. Dra. George Allan Menezes Gomes.

1. Level design. 2. Testes com usuários de jogos. 3. Visualização de dados. I. Título.

CDD 005

ARTHUR SILVA BASTOS

O USO DE TÉCNICAS DE TESTES COM USUÁRIO COMBINADAS COM
VISUALIZAÇÃO DE DADOS PARA APERFEIÇOAR UM JOGO MASOCORE DE
PLATAFORMA 2D

Dissertação apresentada à coordenação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Ciência da Computação. Área de concentração: Computação Gráfica.

Aprovada em: 10/02/2021

BANCA EXAMINADORA

Profa. Dra. Emanuele Marques dos Santos (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. George Allan Menezes Gomes (Coorientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Creto Augusto Vidal
Universidade Federal do Ceará (UFC)

Prof. Dr. Windson Viana de Carvalho
Universidade Federal do Ceará (UFC)

Prof. Dr. Esteban Walter Gonzalez Clua
Universidade Federal Fluminense (UFF)

AGRADECIMENTOS

Gostaria de agradecer à minha família e à minha futura esposa Lais por sempre me acompanharem, me auxiliarem e acreditarem em mim e nas minhas competências como desenvolvedor, pesquisador e, sobretudo, como ser humano.

Agradeço à minha orientadora, profa. Emanuele Santos, por ter me escolhido como orientando e ter me guiado durante esta jornada durante o curso de Mestrado. Foi graças a ela que eu descobri que o mundo de Visualização de Dados é, de fato, deveras apaixonante.

Agradeço também ao prof. George Gomes, que me acompanhou durante praticamente todo o meu tempo na UFC: foi meu professor e supervisor em projetos de docência e de pesquisa, esteve presente na banca do meu projeto na graduação e, por fim, atuou como coorientador desta pesquisa.

Gostaria de agradecer à banca revisora por dedicar seu tempo à revisão e correção desta pesquisa.

Gostaria de agradecer ao meu amigo Marcos Mourão por ter participado ativamente do desenvolvimento do jogo, especialmente na parte de concepção das mecânicas e projeto dos níveis.

Gostaria de agradecer a todos os voluntários que testaram o jogo, em especial àqueles que participaram dos testes com webcam. Vocês não sabem o quanto eu fico feliz quando eu vejo alguém jogando e sorrindo de empolgação por ter entendido como passar de um desafio.

Por fim, gostaria de agradecer à FUNCAP e à UFC, em especial aos departamentos de Computação e de Sistemas e Mídias Digitais. A primeira por ter me agraciado com uma bolsa de estudos por dois anos e a segunda por ter fornecido a infraestrutura e oportunidades necessárias para a realização deste projeto.

RESUMO

Conforme a indústria de videogames avança, criar jogos para jogadores mais diversos e exigentes torna-se uma tarefa cada vez mais complexa. Desenvolvedores modernos aplicam métodos de *Games User Research* (GUR) a fim de tomar decisões mais bem informadas baseadas no público-alvo. Métodos tradicionais incluem observação, entrevista e questionário. Para se obterem informações detalhadas a respeito dos usuários, métodos complementares podem ser necessários. Este trabalho analisa a inclusão de dois métodos complementares de baixo custo, o *eye-tracking* por meio de webcam e a telemetria, com auxílio de visualização de dados, em uma rotina de testes. Os dados adquiridos por estes métodos foram visualizados por meio de uma ferramenta customizada de acordo com as necessidades do projeto, que foi devidamente testada por profissionais e estudantes da área. O desenvolvimento de três versões de um jogo de plataforma hardcore 2D cujo diferencial reside em mecânicas de atenção multitarefa mostrou, por meio da comparação entre os protótipos, que os métodos complementares escolhidos identificaram uma parcela significativa dos problemas de *gameplay* encontrados. A análise da ferramenta mostrou que os dados advindos de métricas e *eye-tracking* proveram discernimento a respeito de *multitasking* e *level design*. O trabalho, por fim, discute os desafios em se avaliarem protótipos sob o critério da agradabilidade da experiência quando a frustração é um elemento de *gameplay* desejável.

Palavras-chave: Level design. Testes com usuários de jogos. Visualização de dados.

ABSTRACT

As the video game industry evolves, creating games for more diverse and demanding players becomes an increasingly complex task. Modern developers apply methods of Games User Research (GUR) to make better informed decisions based on the target audience. Traditional methods include observation, interview and questionnaire. To obtain detailed information about users, complementary methods may be required. This thesis analyzes the inclusion of two complementary low-cost methods, namely webcam-based eye-tracking and telemetry, combined with the data visualization, in a testing routine. The data acquired by these methods were visualized using a tool customized according to the needs of the project, which was duly tested by professionals and students in the field. The development of three versions of a 2D hardcore platform game whose differential lies in its multitasking mechanics showed, through the comparison of the prototypes, that the chosen complementary methods highlighted a significant part of the gameplay issues encountered. Analysis of the tool showed that the data derived from metrics and eye-tracking provided insight into multitasking and level design. Finally, the work discusses the challenges in evaluating prototypes under the criterion of the pleasantness of the experience when frustration is a desirable gameplay element.

Keywords: Level design. Games user research. Data visualization.

LISTA DE FIGURAS

| | |
|----------------------------------------------------------------------------------------------|----|
| Figura 1 – Exemplo de aparelho de miografia facial, usado em testes biométricos | 14 |
| Figura 2 – Painel da Unity Analytics | 15 |
| Figura 3 – Diagrama MDA | 19 |
| Figura 4 – Captura de tela de Super Meat Boy, jogo estilo <i>masocore</i> | 20 |
| Figura 5 – Gráfico que mostra a relação entre habilidade, desafio e <i>flow</i> | 21 |
| Figura 6 – Visualização de mapa de calor no jogo ParaLily | 24 |
| Figura 7 – Exemplo de <i>heatmap</i> | 25 |
| Figura 8 – <i>Screenshot</i> do jogo <i>Half-Life 2</i> | 27 |
| Figura 9 – <i>Screenshot</i> do jogo <i>Limbo</i> | 28 |
| Figura 10 – Combinação de <i>heatmap</i> e clusterização | 29 |
| Figura 11 – Um exemplo de Biometric Storyboard | 30 |
| Figura 12 – Visualização de dados biométricos no jogo SuperTux | 30 |
| Figura 13 – Material promocional do jogo <i>Saber's Edge</i> | 31 |
| Figura 14 – <i>Screenshot</i> do jogo <i>Portal</i> | 32 |
| Figura 15 – Técnica de clusterização utilizada por Wallner <i>et al.</i> (2019) | 34 |
| Figura 16 – Exemplo de visualização criada por meio do Vixen | 35 |
| Figura 17 – Técnica de clusterização utilizada por Kepplinger <i>et al.</i> (2020) | 35 |
| Figura 18 – Técnica de rotas quentes utilizada por Gomes <i>et al.</i> (2018) | 37 |
| Figura 19 – Diagrama que mostra a metodologia deste trabalho | 40 |
| Figura 20 – Captura de tela do desafio 1C do jogo <i>Downside Up</i> | 41 |
| Figura 21 – Captura de tela do jogo <i>Celeste</i> | 42 |
| Figura 22 – Captura de tela do jogo <i>I Wanna Be The Guy</i> | 42 |
| Figura 23 – Captura de tela do desafio 1D do jogo <i>Downside Up</i> | 43 |
| Figura 24 – Captura de tela do jogo <i>Chronos Twins</i> | 44 |
| Figura 25 – Captura de tela do jogo <i>Binaries</i> | 44 |
| Figura 26 – Captura de tela do desafio 1B do jogo <i>Downside Up</i> | 45 |
| Figura 27 – Estados do desafio 2B do jogo <i>Downside Up</i> | 46 |
| Figura 28 – Fluxo do processo de desenvolvimento dos protótipos | 47 |
| Figura 29 – Design inicial do desafio 1C conforme planejado no papel. | 47 |
| Figura 30 – Diagrama que mostra o procedimento de execução dos testes | 49 |
| Figura 31 – Exemplo de registro da telemetria | 51 |

| | |
|-------------------------------------------------------------------------------------------------|----|
| Figura 32 – Exemplo de visualização de <i>replay</i> | 53 |
| Figura 33 – Visualização agregada do desafio 3D | 54 |
| Figura 34 – Resumo do perfil dos entrevistados. | 56 |
| Figura 35 – Situações de uso da ferramenta de visualização | 58 |
| Figura 36 – Resumo das respostas dos entrevistados sobre a visualização | 59 |
| Figura 37 – Laboratório onde foram realizados os testes da Etapa 3 <i>Downside Up</i> | 62 |
| Figura 38 – Resultado do questionário de avaliação da dificuldade | 63 |
| Figura 39 – Exemplos de desafios com discrepância de dificuldade | 64 |
| Figura 40 – Captura de tela do desafio 1F | 66 |
| Figura 41 – Respostas aos módulos Core e Post-Game do GEQ | 68 |
| Figura 42 – Respostas ao PANAS (Teste 2), com intervalo de confiança de 95%. | 68 |
| Figura 43 – Captura de tela do desafio 3F | 74 |
| Figura 44 – Visualização agregada do desafio 3F de <i>Downside Up</i> | 75 |
| Figura 45 – Porções finais do desafio 3F em cada protótipo | 76 |
| Figura 46 – Desafio 1F implementado na versão v3 | 77 |
| Figura 47 – Visualizações agregadas dos desafios 1F e 3E | 79 |

LISTA DE TABELAS

| | |
|----------------------------------------------------------------------------------------|----|
| Tabela 1 – Resumo dos métodos tradicionais de testes com usuários de jogos | 22 |
| Tabela 2 – Resumo das medidas fisiológicas em GUR | 23 |
| Tabela 3 – Resumo das Metodologias de GUR adotadas neste trabalho | 36 |
| Tabela 4 – Comparação entre este trabalho e os trabalhos mais importantes | 37 |
| Tabela 5 – Resumo dos comentários dos entrevistados sobre a visualização | 60 |
| Tabela 6 – Resultados dos testes de soma de posto para os três questionários | 69 |
| Tabela 7 – Resultados dos testes <i>post hoc</i> para os três questionários. | 70 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|-------|-------------------------------------------|
| BioST | <i>Biometric Storyboards</i> |
| GEQ | <i>Game Experience Questionnaire</i> |
| GUR | <i>Games User Research</i> |
| MDA | <i>Mechanics, Dynamics, Aesthetics</i> |
| PANAS | <i>Positive and Negative Affect Scale</i> |

SUMÁRIO

| | | |
|--------------|----------------------------------------------------------------------------------|-----------|
| 1 | INTRODUÇÃO | 13 |
| 1.1 | Motivação | 13 |
| 1.2 | Objetivos | 16 |
| 1.3 | Etapas de Pesquisa | 17 |
| 1.4 | Contribuições | 17 |
| 1.5 | Publicação | 18 |
| 1.6 | Organização do Trabalho | 18 |
| 2 | CONCEITOS PRELIMINARES | 19 |
| 2.1 | Game Design | 19 |
| 2.2 | Experiência de Jogo | 20 |
| 2.3 | Testes com Usuários de Jogos | 21 |
| 2.4 | Visualização de Dados | 23 |
| 2.5 | Considerações Finais | 25 |
| 3 | TRABALHOS RELACIONADOS | 26 |
| 3.1 | Estratégia da Revisão Bibliográfica | 26 |
| 3.2 | Uso de GUR e Visualização de Dados para o Aperfeiçoamento de Level Design | 27 |
| 3.3 | Outros Usos de Visualização de Dados | 33 |
| 3.4 | Discussão | 34 |
| 3.5 | Considerações Finais | 38 |
| 4 | ABORDAGEM PROPOSTA | 39 |
| 4.1 | Metodologia | 39 |
| 4.2 | Downside Up | 40 |
| 4.2.1 | <i>Motivação</i> | 41 |
| 4.2.2 | <i>Game Design</i> | 43 |
| 4.3 | Dados Coletados | 48 |
| 4.3.1 | <i>Procedimento Padrão</i> | 48 |
| 4.3.2 | <i>Observação</i> | 48 |
| 4.3.3 | <i>Entrevista</i> | 49 |
| 4.3.4 | <i>Questionário de Avaliação de Dificuldade</i> | 49 |

| | | |
|-------|---------------------------------------------------------------------|-----|
| 4.3.5 | <i>Métricas de Jogo</i> | 50 |
| 4.3.6 | <i>Eye-Tracking</i> | 50 |
| 4.3.7 | <i>Estrutura de Dados</i> | 51 |
| 4.3.8 | <i>Questionários de Avaliação de Experiência</i> | 51 |
| 4.4 | Visualizações | 52 |
| 4.4.1 | <i>Replays</i> | 52 |
| 4.4.2 | <i>Visualização Agregada</i> | 53 |
| 4.4.3 | <i>Avaliação da Ferramenta</i> | 54 |
| 4.5 | Considerações Finais | 55 |
| 5 | RESULTADOS | 56 |
| 5.1 | Entrevista com Desenvolvedores | 56 |
| 5.2 | Testes com Usuários | 61 |
| 5.2.1 | <i>Etapa 2 - Desenvolvimento da Versão v2 (Teste 1.1)</i> | 62 |
| 5.2.2 | <i>Etapa 2 - Desenvolvimento da Versão v3 (Teste 1.2)</i> | 65 |
| 5.2.3 | <i>Etapa 3 - Comparação entre Protótipos (Teste 2)</i> | 67 |
| 5.3 | Considerações Finais | 70 |
| 6 | LIMITAÇÕES E LIÇÕES APRENDIDAS | 71 |
| 6.1 | Limitações | 71 |
| 6.2 | Lições aprendidas | 73 |
| 6.3 | Considerações Finais | 80 |
| 7 | CONCLUSÕES E TRABALHOS FUTUROS | 81 |
| 7.1 | Resultados Alcançados | 81 |
| 7.2 | Limitações | 81 |
| 7.3 | Trabalhos Futuros | 82 |
| | REFERÊNCIAS | 84 |
| | APÊNDICE A– QUESTIONÁRIO DE DIFICULDADE | 92 |
| | APÊNDICE B– QUESTIONÁRIOS DE EXPERIÊNCIA | 110 |
| | APÊNDICE C– TERMO DE CONSENTIMENTO | 117 |
| | ANEXO A– GAME EXPERIENCE QUESTIONNAIRE | 120 |
| | ANEXO B– POSITIVE AND NEGATIVE AFFECT SCALE | 124 |

1 INTRODUÇÃO

Esta dissertação apresenta uma metodologia que visa aperfeiçoar jogos digitais e que foi aplicada no desenvolvimento de um jogo intitulado Downside Up. O jogo foi aperfeiçoado seguindo técnicas de Testes com Usuários de Jogos (*Games User Research*, ou GUR), uma área emergente derivada da área de Experiência de Usuário (*User Experience*, ou UX) que investiga as particularidades dos jogos digitais. Os dados advindos desses testes foram visualizados através de uma ferramenta adaptada às necessidades do projeto. A ferramenta foi validada por profissionais e estudantes da área de jogos. Com base na interpretação dos dados, foram gerados múltiplos protótipos. A avaliação desses protótipos visou identificar as contribuições das metodologias de GUR. O critério utilizado foi o da diversão: o protótipo superior seria aquele que proporcionasse a experiência mais divertida.

Ao final desta dissertação são feitas reflexões sobre o processo, que contou com diversas conclusões surpreendentes e inesperadas pelo autor e demais participantes da pesquisa.

A motivação para este trabalho encontra-se na Seção 1.1, os objetivos na Seção 1.2. Um resumo da metodologia utilizada para alcançar estes objetivos está na Seção 1.3. As principais contribuições estão resumidas na Seção 1.4. Logo em seguida, é mostrada a publicação gerada por esta dissertação na Seção 1.5. Na Seção 1.6 há uma breve descrição do conteúdo a ser exposto nos demais capítulos.

1.1 Motivação

Jogos são fontes de inovações artísticas, movimentam uma indústria multimilionária e justificam evoluções tecnológicas (PAGULAYAN *et al.*, 2002), especialmente aquelas relacionadas a computação gráfica. Além disso, possuem fins que vão além do entretenimento, como a educação, a simulação, o treinamento e a saúde.

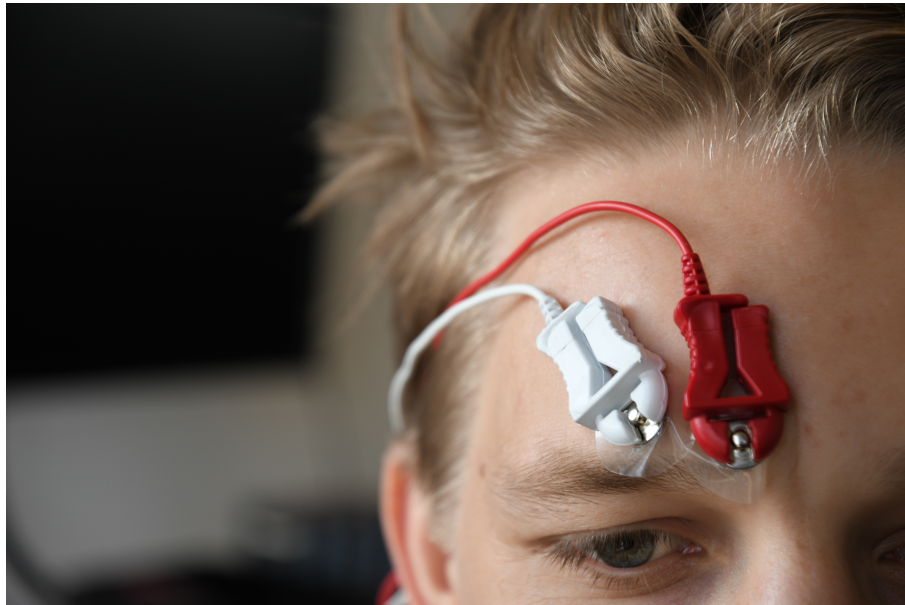
Denomina-se *game designer* o profissional responsável por assegurar que o jogo, enquanto produto, cumprirá sua finalidade (KOSTER, 2013). Guiar-se apenas pela intuição ou experiência do *game designer*, no entanto, pode levar a resultados insatisfatórios (ALBERT; TULLIS, 2013).

Desenvolvedores modernos aplicam técnicas de *Games User Research* (GUR) a fim de obter informações significativas sobre o jogo e seus jogadores. A partir delas é possível tomar decisões melhores que visem aperfeiçoar a experiência de jogo (MIRZA-BABAEI *et al.*, 2013b).

Isto é abordado em mais detalhes nos Capítulos 2 e 3.

Em desenvolvimento de jogos, costuma-se aplicar tradicionalmente observação, entrevista e questionário (RAJANEN; TAPANI, 2018). Esses métodos se destacam por responder bem a perguntas do tipo “*por quê*”, mas podem ser menos precisos para perguntas como “*o quê*” ou “*como*” para eventos de jogo, já que podem introduzir vieses advindos da subjetividade. Para coletar dados mais precisos sobre a experiência do jogador, podem ser necessários métodos complementares, como biometria (exemplo na Figura 1), *eye-tracking* e métricas de jogo.

Figura 1 – Exemplo de aparelho de miografia facial, usado em testes biométricos



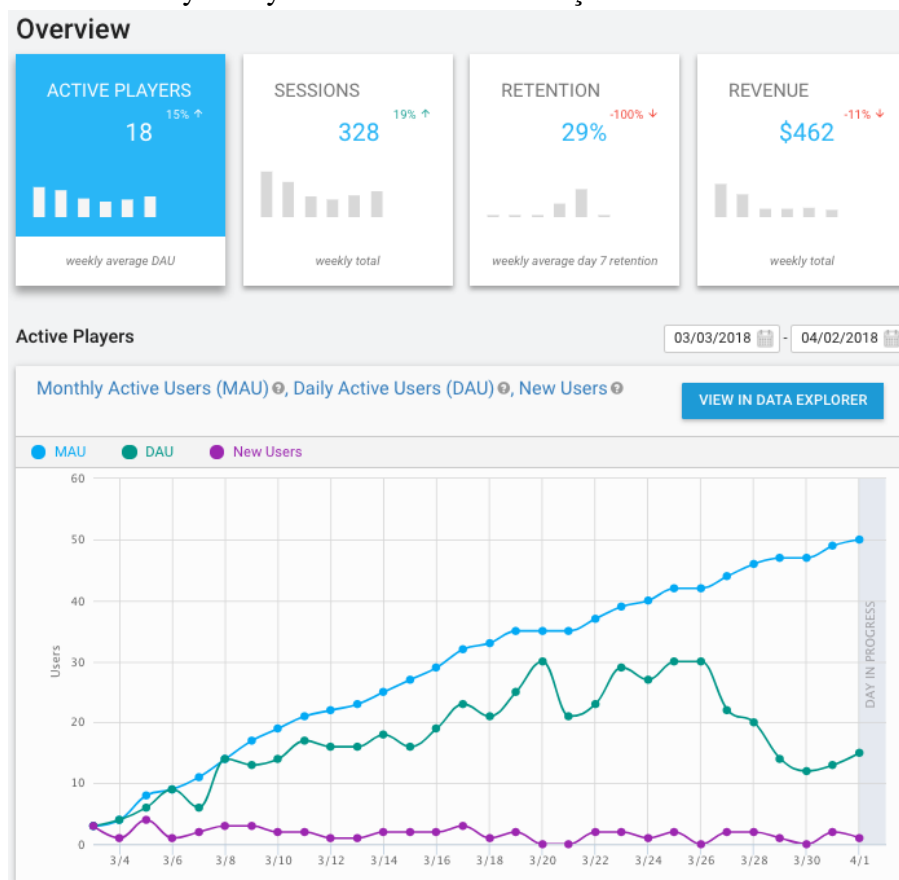
Fonte: (FARNSWORTH, 2018).

Há um crescente interesse por métodos biométricos, visto que eles proveem informação quantitativa sobre estados emocionais (DRACHEN *et al.*, 2018). Tais métodos, no entanto, normalmente requerem equipamentos caros, intrusivos, operados por profissionais especializados. Como consequência, a biometria costuma não ser economicamente viável para pequenos estúdios, estudantes e desenvolvedores independentes (MIRZA-BABAEI *et al.*, 2016). Por outro lado, os métodos que utilizam *eye-tracking* podem ser executados até com uma simples *webcam* (PAPOUTSAKI *et al.*, 2016; KRAFKA *et al.*, 2016; VALENTI *et al.*, 2009), mesmo que esta não garanta um resultado tão confiável quanto o de um aparelho dedicado a isto. O *eye-tracking* destaca-se por mensurar focos de atenção dos jogadores (EL-NASR; YAN, 2006). Além disso, costuma-se incluir rastreadores no código do jogo capazes de registrar a ocorrência de eventos e métricas. Este método é conhecido como telemetria (EL-NASR *et al.*, 2016). As métricas são definidas conforme a necessidade dos desenvolvedores e podem variar de condições genéricas

como medições do tempo de jogo até situações mais específicas como detecção de quais armas são mais utilizadas contra determinados inimigos em algum cenário.

Somando-se aos métodos citados, emprega-se também o uso de técnicas de visualização para tornar os dados adquiridos de diversas fontes mais compreensíveis (EL-NASR *et al.*, 2016). Plataformas de desenvolvimento como a Unity¹ oferecem ferramentas gratuitas para isto (Figura 2), tornando o uso de métricas e visualização de dados acessíveis para visualizações básicas como os gráficos de linhas.

Figura 2 – Painel da Unity Analytics mostrando informações sobre acesso de usuários a um jogo



Fonte: Manual da Unity (TECHNOLOGIES, 2020).

A despeito disso, o conhecimento acerca de visualização para jogos não é amplamente difundido. Historicamente, os jogos eletrônicos foram desenvolvidos sem o auxílio de visualização (EL-NASR *et al.*, 2016). Isto mudou em meados dos anos 2000, quando a conectividade online se tornou popular nos consoles, mas foi apenas na última década, com o advento dos jogos para dispositivos móveis, que a prática ganhou maior reconhecimento e o *status* de essencial para o desenvolvimento.

¹ <https://unity.com>

A aplicação de GUR costuma estar relacionada à maturidade dos estúdios em relação à área de Experiência de Usuário (DRACHEN *et al.*, 2018). Empresas maiores como Valve, EA Games e Ubisoft destinam recursos para GUR e costumam tratar seus métodos de testes como segredo estratégico. Estúdios menores, no entanto, aplicam GUR de forma mais restrita, seja pela falta de recursos, seja pela falta de conhecimento (NACKE *et al.*, 2016). Em estúdios informais sem fins lucrativos, geralmente formados por grupos de amigos que se empolgam com a ideia de desenvolver jogos, é bastante comum a ideia de que tomar decisões baseadas em dados torna o jogo “menos artístico” ou desmerece a criatividade do idealizador (DRACHEN *et al.*, 2018), o que cria mais barreiras à popularização dos testes de usuários e da visualização. No caso específico do Brasil, os métodos tradicionais de testes de usuários de jogos são os mais utilizados (ex.: observar um participante jogar e colher *feedback*), porém uma parcela significativa de estúdios relata não realizar nenhum tipo de teste (WANICK; BITELO, 2020).

É possível que a ignorância em relação à área de visualização também possa ser decorrente da formação acadêmica. Cursos de Ciências da Computação normalmente abordam visualização científica em suas grades curriculares, mas as ementas não tratam das particularidades da visualização para jogos (SUNDSTEDT, 2016).

Diante do exposto, surge a necessidade de ilustrar de forma prática as vantagens da aplicação de testes com usuários aliados à visualização de dados para aperfeiçoar um jogo.

1.2 Objetivos

Neste trabalho são discutidas e comparadas diversas metodologias de GUR aplicadas ao desenvolvimento de um jogo de plataforma 2D. Além disso, os dados de duas dessas metodologias, a telemetria e o *eye-tracking* usando *webcam*, são agregados e visualizados com uma técnica própria. Todo o processo é pensado para ser executado com baixo custo afim de mostrar que GUR e visualização podem ser utilizadas em estúdios de qualquer tamanho.

Esta pesquisa visa responder à seguinte pergunta: “como combinar de forma eficaz métodos de Testes com Usuários e Visualização de Dados para aperfeiçoar um jogo de plataforma 2D?”. Para tal, é necessário portanto:

- **Objetivo 1:** Realizar um estudo sobre GUR e visualização de dados para definir a abordagem metodológica;
- **Objetivo 2:** Aplicar a abordagem no desenvolvimento, aperfeiçoamento e prototipação de um jogo;

- **Objetivo 3:** Comparar protótipos segundo um critério objetivo;
- **Objetivo 4:** Analisar e discutir os resultados da comparação.

Espera-se também que este trabalho ilustre a importância do uso de visualização de dados no desenvolvimento de jogos tanto para profissionais da área quanto para estudantes e pesquisadores.

1.3 Etapas de Pesquisa

Inicialmente, houve a fundamentação teórica e o levantamento de trabalhos relacionados. Em seguida, foi executada uma metodologia de desenvolvimento e avaliação de múltiplos protótipos de um jogo. O aperfeiçoamento dos protótipos se deu com o auxílio de uma ferramenta de visualização validada por terceiros. Ao final, os protótipos foram comparados utilizando questionários de avaliação de experiência. Dessa avaliação foi possível extrair conclusões sobre todo o processo.

1.4 Contribuições

As principais contribuições deste trabalho são as seguintes:

- **Contribuição 1:** Uma metodologia de desenvolvimento, aperfeiçoamento e avaliação de jogos que emprega visualização de dados e técnicas de GUR de baixo custo;
- **Contribuição 2:** O desenvolvimento e a validação de uma nova solução de visualização de dados que agrega dados de métricas de jogo com dados de *eye-tracking*;
- **Contribuição 3:** A disponibilização dos protótipos do jogo desenvolvido, das visualizações produzidas, dos dados de questionários e outros materiais utilizados na pesquisa em um repositório público ²;
- **Contribuição 4:** Uma discussão sobre o uso de técnicas de testes com usuários aliadas a visualização de dados para o aperfeiçoamento de um jogo com base no experimento realizado aqui.

² <https://arthursb.github.io/Downside-Up/>

1.5 Publicação

Por meio da pesquisa relatada neste trabalho, foi publicado um artigo, intitulado *Evaluating the Use of Affordable User Testing and Visualization Techniques in Level Design of a Hardcore 2D Platform Game* (BASTOS *et al.*, 2019). O artigo foi publicado em 2019 e apresentado no Simpósio Brasileiro de Games e Entretenimento Digital (SBGames).

1.6 Organização do Trabalho

Este trabalho está organizado em oito capítulos. Neste capítulo foram introduzidos o contexto motivacional, os objetivos, as principais contribuições e uma breve descrição da metodologia. O restante organiza-se na seguinte estrutura:

- No Capítulo 2 são apresentados conceitos teóricos necessários para a compreensão total deste trabalho;
- No Capítulo 3 há uma revisão bibliográfica e a identificação dos principais trabalhos similares a este;
- No Capítulo 4, apresentam-se o detalhamento da metodologia e do jogo desenvolvido e a descrição e importância dos dados coletados, bem como suas visualizações, quando necessárias;
- No Capítulo 5, estão os resultados de cada teste;
- No Capítulo 6, discutem-se os resultados encontrados;
- No Capítulo 7 conclui-se o trabalho, mostrando direcionamentos para trabalhos futuros.

2 CONCEITOS PRELIMINARES

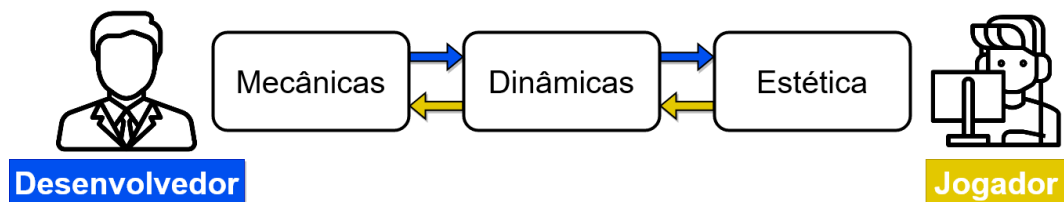
Neste capítulo, são mostrados os conceitos mais importantes para a compreensão do trabalho. A Seção 2.1 apresenta os conceitos adotados sobre a concepção de um jogo. A Seção 2.2 discute os fenômenos que ocorrem no ato de jogar. A Seção 2.3 resume métodos de GUR mais comuns. A Seção 2.4 introduz o conceito de visualização de dados e contextualiza sua aplicação na área de jogos. Na Seção 2.5, são feitas considerações finais.

2.1 Game Design

Game design refere-se ao processo de projetar um jogo. Segundo o modelo *Mechanics, Dynamics, Aesthetics* (MDA), proposto por Hunicke *et al.* (2004), um jogo é separado em três componentes: mecânicas, dinâmicas e estética. Mecânicas são os componentes mais básicos, como regras e ações. O ato de pular, por exemplo, pode ser uma mecânica. Dinâmicas são o resultado da combinação de mecânicas, formando sistemas controláveis pelo usuário. Uma tarefa no jogo, tal como pular para alcançar uma bandeira, pode ser vista como uma dinâmica, pois combina mecânicas como colisão, gravidade, pulo, missões etc. Estética¹ se refere à experiência do jogador e às suas sensações, emoções, percepções e fantasias em relação ao jogo.

O modelo MDA também ilustra as relações de produção e consumo. Um desenvolvedor costuma construir o jogo partindo das mecânicas, para depois combiná-las e formar as dinâmicas, que por sua vez possuem uma reação estética. O jogador participa do processo na ordem reversa: inicialmente ele vivencia uma experiência ao jogar, que o motiva a explorar e observar os sistemas para, por fim, compreender as mecânicas. A Figura 3 ilustra esse processo.

Figura 3 – Diagrama que mostra a percepção dos jogadores e desenvolvedores dos componentes do modelo MDA



Fonte: Adaptado de Hunicke *et al.* (2004)

Este trabalho foca em jogos de ação, mais especificamente no gênero plataforma 2D.

Os jogos de plataforma, além de se encontrarem em crescente popularidade, são relativamente

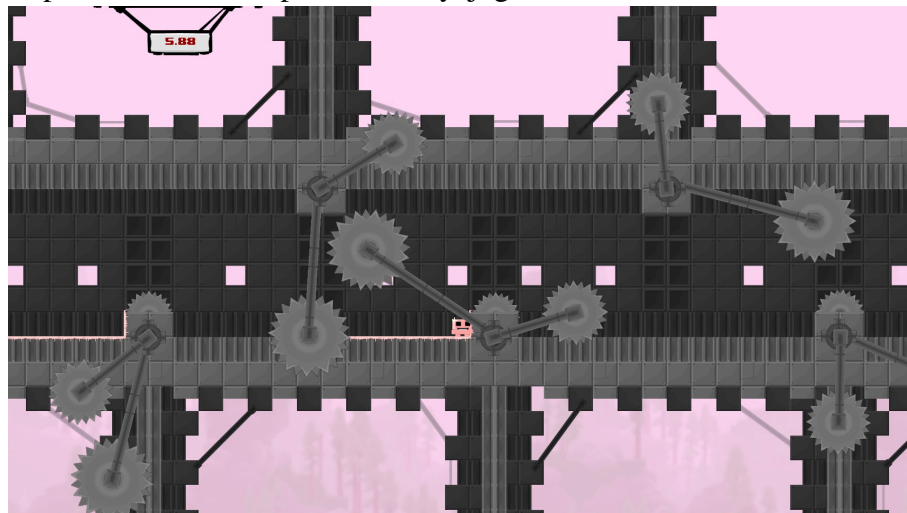
¹ O termo remete à Estética, área da filosofia que atua sobretudo no ramo da arte e trata de questões como o que é gostar ou apreciar, o que é belo, o que é bom e ruim, etc. (SANTAELLA, 1994)

simples de serem criados e compreendidos pelos jogadores. O gênero traz diversas praticidades que contribuem para alcançar os objetivos desta pesquisa de mestrado.

As principais mecânicas de jogos de plataforma incluem: um personagem controlável, plataformas fixas ou móveis, obstáculos, itens colecionáveis, recursos que modifiquem a movimentação (como cordas, esteiras e superfícies deslizantes), e gatilhos (como botões, chaves e cadeados). Um nível, ou fase, costuma ser formado pelo agrupamento de desafios. Desafios, por sua vez, são compostos por combinações de mecânicas (SMITH *et al.*, 2008).

O jogo desenvolvido neste trabalho enquadra-se no subgênero conhecido informalmente como *masocore* ou *platform hell* (um exemplo é mostrado na Figura 4). Este tipo de jogo apresenta alta dificuldade, requer reflexos rápidos e precisos, contém armadilhas ou truques para surpreender jogadores desatentos e, ocasionalmente, fazem comédia sobre o fracasso.

Figura 4 – Captura de tela de Super Meat Boy, jogo estilo *masocore*



Fonte: Comunidade Steam (SUP, 2013).

2.2 Experiência de Jogo

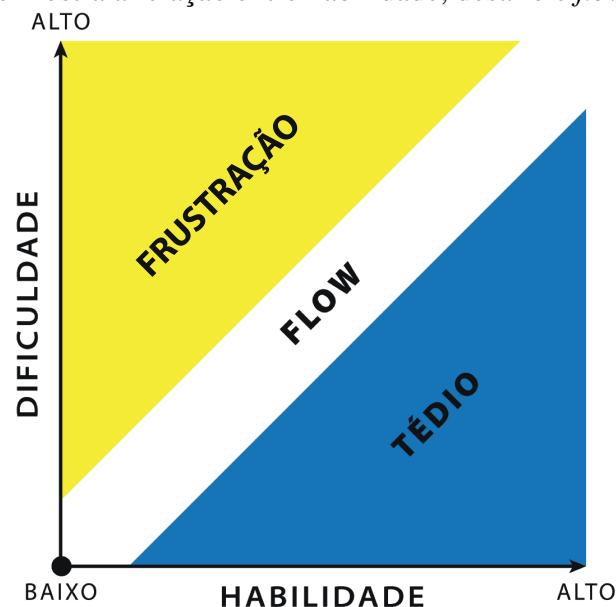
Frequentemente, a aprovação dos usuários se manifesta por meio de experiências positivas, que normalmente implicam no sentimento de diversão (FEDEROFF, 2002). Koster (2013) fala que diversão é a resposta emocional positiva ao processo cognitivo do aprendizado, da resolução de quebra-cabeças e da superação de desafios. Huizinga (2014) a reconhece como principal finalidade do entretenimento. Kim *et al.* (2015) complementam dizendo que também há diversão por meio da estética e das relações interpessoais.

É dito que o *flow*, ou estado de fluxo, proporciona a diversão (SHELDON; KING,

2001). O termo se refere a um estado de concentração no qual as ações são realizadas de forma semiautomática, sem intervenção consciente. Esse estado é acompanhado de percepções distorcidas de tempo, do ambiente externo e do ego (CSIKSZENTMIHALYI, 2013).

Csikszentmihalyi *et al.* (2014) citam que jogos são a ocasião perfeita para a ocorrência de *flow*, pois contêm elementos que justificam envolvimento emocional e regras que definem quais estímulos são relevantes. Os autores comentam que o *flow* acontece quando há um equilíbrio entre as habilidades do jogador e o nível de desafio exigido. Quando as capacidades do jogador excedem a cobrança do desafio, há o tédio. Quando o desafio excede a habilidade, há a frustração (CSIKSZENTMIHALYI *et al.*, 2014). A Figura 5 ilustra o conceito.

Figura 5 – Gráfico que mostra a relação entre habilidade, desafio e *flow*



Fonte: Adaptado de Csikszentmihalyi *et al.* (2014).

Propomos analisar a experiência por meio da proporção entre afetos positivos e negativos², partindo do princípio de que, no caso de jogos, a diversão é uma das principais consequências de uma experiência majoritariamente positiva (JACKSON, 2000).

2.3 Testes com Usuários de Jogos

É fundamental que o *game designer* ajuste os desafios de acordo com as habilidades do jogador para proporcionar o *flow*. Os principais métodos de GUR, que ajudam os designers a encontrar pontos que precisam de refinamento no jogo são mostrados aqui.

² Conceito usado na Psicologia para definir construtos que representam estados com repercussões físicas e mentais, ocasionados por reações a estímulos. Um exemplo de afeto é a raiva (MATSUMOTO, 2009).

Dadas a complexidade e a subjetividade em se avaliar a experiência, são aplicados diferentes testes na Academia e na Indústria com o intuito de obter dados sobre o produto ou sobre os usuários (DRACHEN *et al.*, 2018). Estes dados ajudam os *designers* a tomar decisões mais bem informadas que visam a melhorar o produto (AMBINDER, 2009; PAGULAYAN *et al.*, 2002).

A Tabela 1 resume os métodos aplicados tradicionalmente. Estes costumam depender apenas de lápis e papel, ou de outros meios de anotação. Normalmente, são subjetivos e contam com a presença de um pesquisador enquanto o testador joga. São métodos que, quando aplicados, fornecem informações pertinentes principalmente para o ajuste da dinâmica e da estética do jogo.

Tabela 1 – Resumo dos métodos tradicionais de testes com usuários de jogos

| Método | Descrição | Vantagens | Desvantagens |
|-------------------|-----------------------------------------------------------------------|--------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| Observação Direta | Observar o testador jogando e anotar comportamentos | Informa sobre comportamentos de que nem sempre o testador está ciente | Presença do observador pode enviesar os resultados |
| Verbalização | Pedir que o testador fale o que está pensando enquanto joga | Provê maior esclarecimento sobre os processos cognitivos | Processo distrativo que afeta a sessão de jogo |
| Questionários | Formular questões a serem respondidas com respostas padrão | Provê dados menos enviesados | Menor espaço para liberdade e variedade de respostas |
| Entrevistas | Conversar com os testadores sobre algum assunto pertinente à pesquisa | Responde a perguntas específicas de <i>design</i> | Potencial viés nas respostas, visto que há dependência na memória dos envolvidos |
| Gravação | Gravar a sessão de jogo e/ou o testador jogando | Registra com fidelidade sessões completas de jogo para análise posterior | Análise das gravações tende a ser demorada |

Fonte: Elaborada pelo autor com base nas obras de Ambinder (2009), Nacke *et al.* (2009), El-Nasr *et al.* (2016) e Drachen *et al.* (2018).

Há um crescente interesse por métodos biométricos (WALLNER *et al.*, 2019; MUSCAT *et al.*, 2019; HALABI, 2019). Nestes, são empregados aparelhos que realizam medições sobre dados fisiológicos dos jogadores. O maior destaque desses métodos é que eles são capazes de identificar emoções do jogador durante a sessão de jogo. Aplicar métodos biométricos, no entanto, requer equipamentos caros e geralmente intrusivos, além de ser necessário um especialista capaz de interpretar os dados obtidos pelos aparelhos. A Tabela 2 resume as medidas comumente observadas em testes.

Os métodos descritos anteriormente costumam detalhar ações ou sentimentos do jogador. Quanto a comportamentos do jogo, costuma-se utilizar métricas de jogo (telemetria). *Softwares* costumam implementar rastreadores que registram determinados parâmetros, tais como o tempo que um usuário utilizou a aplicação, quantas vezes por dia, em qual país etc.

Tabela 2 – Resumo das medidas fisiológicas pertinentes a testes com usuários de jogos

| Medida | Utilidade | Vantagens | Desvantagens |
|------------------------------------------|---------------------------------------------|--------------------------------------------------------|------------------------------------------------------------------------|
| Batimento Cardíaco | Identifica o nível de agitação | Simple de aferir | Não aponta a qual estímulo há uma reação |
| Condutividade da Pele | Identifica o nível de agitação | Fácil de identificar picos de estímulo | Suscetível a fatores externos (ex.: umidade) |
| Olhar | Mostra os focos de atenção na tela | Indica o que é mais chamativo em uma interface gráfica | Dados podem demorar a ser interpretados |
| Expressões Faciais | Sugerem o que os jogadores estavam sentindo | Afere com precisão emoções | Precisa ser combinada com outros métodos para evitar vieses |
| Eletroencefalografia (ondas cerebrais) | Mede relaxamento, engajamento e cansaço | Grande precisão | Dados costumam ser difíceis de serem tratados e interpretados |
| Eletromiografia facial (tensão muscular) | Determina emoções | Grande precisão | Pode enviesar a jogabilidade devido à necessidade de aparelhos na face |

Fonte: Elaborada pelo autor com base nas obras de Ambinder (2009), Nacke *et al.* (2009), El-Nasr *et al.* (2016) e Drachen *et al.* (2018).

Em jogos, podem ser aplicados rastreadores específicos, como locais onde há mais mortes no jogo, armas mais utilizadas, fases nas quais ocorrem a maior frequência de desistências, entre outros. As métricas fornecem informações pertinentes principalmente para o ajuste de mecânicas e dinâmicas do jogo (EL-NASR *et al.*, 2016).

2.4 Visualização de Dados

Visualização de dados se refere à apresentação e exploração de informações. Uma boa visualização melhora a comunicação ao facilitar a compreensão de dados complexos (CHEN *et al.*, 2007). Estes, no caso dos jogos, costumam advir da telemetria.

Conforme mostrado anteriormente, os dados pertinentes ao desenvolvimento de dados costumam ser numerosos e variados, podem representar informações subjetivas e objetivas, e se relacionam com inúmeras variáveis, inclusive com fatores externos ao jogo. Extrair informações significativas sobre esses dados frequentemente requer o uso de estratégias de visualização (DRACHEN; SCHUBERT, 2013).

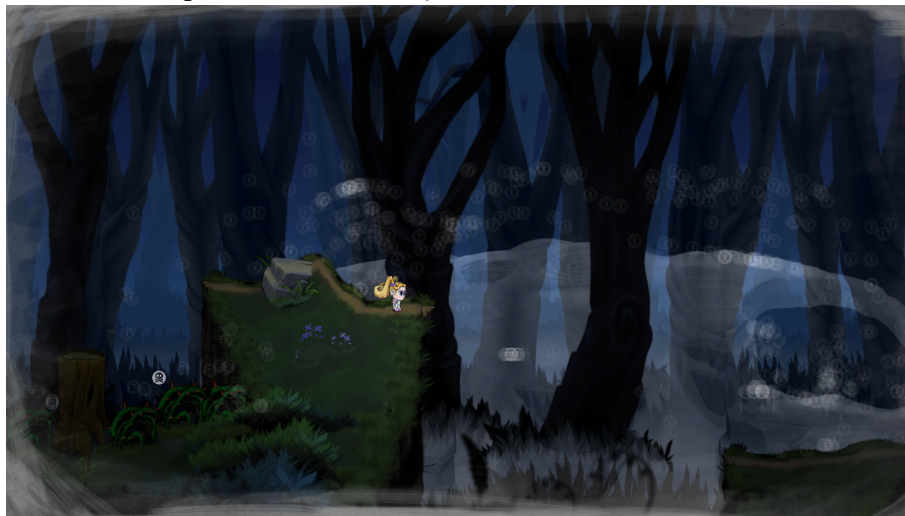
A utilização de visualização em jogos permeia toda a cadeia produtiva. Ela pode ser empregada em pesquisas para criação de provas de conceito, em avaliações de qualidade durante o desenvolvimento e em acompanhamento após o lançamento do jogo para averiguar o sucesso do produto.

As visualizações mais comuns para o desenvolvimento de jogos são aquelas que envolvem as dimensões de espaço e tempo, sendo a mais famosa delas o mapa de calor (ou

heatmap). O mapa de calor tem como finalidade mostrar a frequência ou magnitude de algum evento em um ponto específico através de propriedades das cores, como brilho, transparência e saturação (WILKINSON; FRIENDLY, 2009).

Grandes estúdios costumam não divulgar seus dados, visto que tendem a considerar este tipo de conhecimento como vantagem comercial estratégica (DRACHEN *et al.*, 2018). Estúdios menores costumam relatar de forma transparente o desenvolvimento dos jogos, o que inclui discussões sobre visualizações. Por exemplo, os desenvolvedores do jogo ParaLily, mostrado na Figura 6, utilizaram métricas para aperfeiçoar seções de um nível. Utilizaram a técnica de mapa de calor para descobrir onde os jogadores pulam e morrem com mais frequência. Dessa maneira, encontraram pontos problemáticos por serem muito difíceis. Similarmente, os desenvolvedores de Project Cargo mostraram em seu *devblog* (TEAM, 2016) como o uso de mapas de calor contribui para a criação de um cenário mais diverso.

Figura 6 – Visualização de mapa de calor. A transparência indica onde os jogadores pularam ou morreram com maior frequência em *ParaLily*

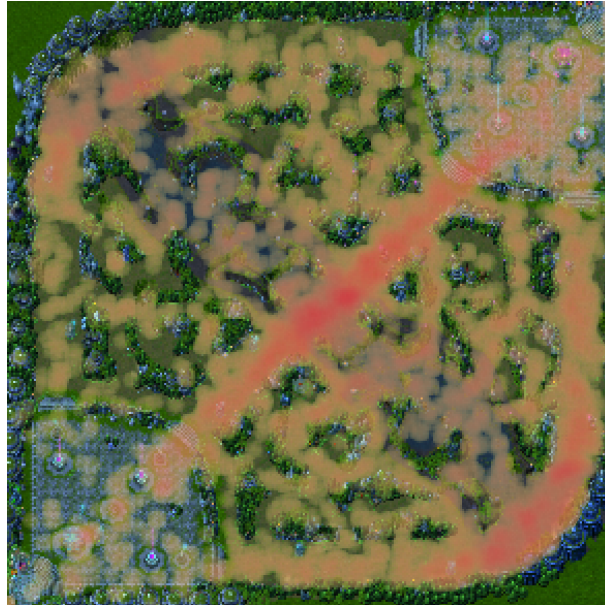


Fonte: *Devblog* do jogo ParaLily (PARALILY, 2018).

Os desenvolvedores do jogo Ultraflow ilustraram a importância das métricas no blog oficial da Unity (CHAN, 2015) ao mostrar que, por meio da medida da desistência dos usuários, é possível tornar a experiência mais agradável ao ajustar a dificuldade em níveis onde há abandonos frequentes. Já os desenvolvedores de Slyway descreveram quais métricas usaram para identificar desafios exageradamente complexos e por quais motivos deve-se balancear um jogo (VOURKOUS, 2018). Por fim, os desenvolvedores de Hero Sky: Epic Guild Wars descreveram que as métricas também são úteis para encontrar problemas de performance e otimizar o graficamente o jogo (LEE, 2015).

Além dos desenvolvedores, a visualização de dados é proveitosa para usuários, investidores e demais *stakeholders*. A Figura 7, por exemplo, ilustra um mapa de calor criada por um fã do jogo League of Legends para auxiliar outros fãs a compreender melhor quais áreas do mapa do jogo são mais perigosas, ou seja, onde ocorrem mortes com maior frequência.

Figura 7 – Exemplo de *heatmap* que mostra, em coloração vermelha, os locais no mapa do jogo League of Legends onde ocorrem mortes com maior frequência



Fonte: Postagem no fórum de discussão sobre o jogo League of Legends no Reddit (YERICH, 2014).

2.5 Considerações Finais

Neste capítulo, foram mostrados, de forma simplificada, conceitos necessários para a compreensão do trabalho. Foram apresentados os métodos de testes mais comuns utilizados na indústria. O desenvolvimento de jogos envolve diversos conceitos subjetivos, sendo os métodos de GUR uma maneira de tornar o processo mais pautado em dados e informação. Essa informação precisa ser visualizada de forma adequada. Foram apresentados o conceito de visualização, a contextualização da visualização na área de jogos e alguns exemplos. No próximo capítulo são apresentados trabalhos que refletem sobre a criação de jogos e os processos envolvidos.

3 TRABALHOS RELACIONADOS

Aqui são apresentados e discutidos trabalhos relacionados cujo tema principal é o uso de GUR para aperfeiçoar um jogo. O intuito da pesquisa realizada aqui é auxiliar na tomada de decisões metodológicas e de desenvolvimento com base no atual estado da arte. A estratégia para a identificação dos trabalhos relacionados está descrita na Seção 3.1. Na Seção 3.2, são mostrados trabalhos acadêmicos relacionados à aplicação de métodos de GUR e visualização de dados em jogos com o intuito de melhorar o *level design*. Em seguida, a Seção 3.3 apresenta outros trabalhos relacionados que utilizam visualização de dados. A Seção 3.4 discute as contribuições dos trabalhos relacionados, ressaltando-se as diferenças deste trabalho. Na Seção 3.5, são feitas as considerações finais.

3.1 Estratégia da Revisão Bibliográfica

A busca por trabalhos relacionados se deu pelo método Quick and Dirty Review, que consiste em identificar, de forma rápida, quais são os artigos mais relevantes para esta pesquisa (YI, 2014). Os artigos foram encontrados pela ferramenta Google Scholar¹, utilizando-se, em diversos idiomas, palavras-chaves como visualização, jogos, testes com usuários e GUR, bem como filtro por data.

Após encontrados os artigos, a leitura inicial foi realizada seguindo o método de Keshav (2007), que é dividido em três etapas. Na primeira, há a leitura na seguinte ordem: título do artigo, *abstract*, introdução, títulos de cada seção, conclusão e referências. Na próxima etapa analisam-se as figuras e marcam-se as referências mais importantes, que são inseridas no conjunto de artigos a serem analisados. Então, constrói-se uma planilha com as seguintes informações:

- Citação do artigo;
- Grau de excelência da publicação, segundo o sistema Qualis²;
- Nível de interesse pessoal do autor desta dissertação pelo artigo lido;
- Ideia central do artigo;
- Novas ideias que esse artigo desperta.

Em seguida é feita uma classificação dos artigos de acordo com o nível de interesse, para que, na última etapa, seja realizada uma leitura completa dos artigos melhor avaliados. Ao

¹ <https://scholar.google.com.br>

² <http://qualis.capes.gov.br/webqualis/>

longo desta pesquisa de mestrado, a busca por artigos foi repetida algumas vezes para detecção de artigos recém-publicados. Ao final do processo, foram separados os trabalhos que mais se relacionam a este, de acordo com as seguintes categorias: *game design*, *level design*, jogos de plataforma, testes com usuários de jogos, técnicas e ferramentas de visualização. O resultado é apresentado nas seções seguintes.

3.2 Uso de GUR e Visualização de Dados para o Aperfeiçoamento de Level Design

Um dos grandes desafios dos designers reside no balanceamento das mecânicas. Conforme dito no Capítulo 2, a correspondência entre experiência dos usuários e a intenção dos designers corrobora com o êxito do produto. Nesta seção, serão apresentados trabalhos que combinam técnicas de GUR e visualização de dados para, por meio destas, obter informações sobre a experiência de jogo dos usuários.

Nacke e Lindley (2010) desenvolveram três modificações do jogo *Half-Life 2* (Figura 8) para entender os efeitos de certas mecânicas. Utilizando métodos biométricos e um questionário, os autores descobriram relações de causa e efeito. Descobriram, por exemplo, que o tédio é causado pela lentidão, a repetição, a invariância, a facilidade e a linearidade. O *flow*, por sua vez, é alcançado por meio do aumento gradativo da dificuldade e alternância entre momentos de risco e segurança. Além de prover direcionamentos de level design, o trabalho ilustra a eficácia do *Game Experience Questionnaire* (GEQ), questionário de experiência específico para jogos posteriormente publicado por IJsselsteijn *et al.* (2013).

Figura 8 – *Screenshot* do jogo *Half-Life 2*



Fonte: Página oficial do jogo na loja Steam (VALVE, 2004)

Costuma-se associar diversão a níveis baixos de frustração, algo investigado por Boulton *et al.* (2017) em um estudo de caso com o jogo Limbo (Figura 9). Adotando a definição de que a frustração se refere ao resultado do impedimento do progresso ou ao fracasso do jogador, os autores descobriram indícios de que frustração pode, na verdade, ocasionar diversão. Esta hipótese é reforçada por Canossa *et al.* (2011) com sua análise de dados do jogo Kane & Lynch 2. Sabendo disso, pode-se considerar que convém inserir momentos frustrantes em um jogo para manter um nível de interesse tal que não seja grande o suficiente para justificar o abandono pelo jogador.

Figura 9 – *Screenshot* do jogo Limbo

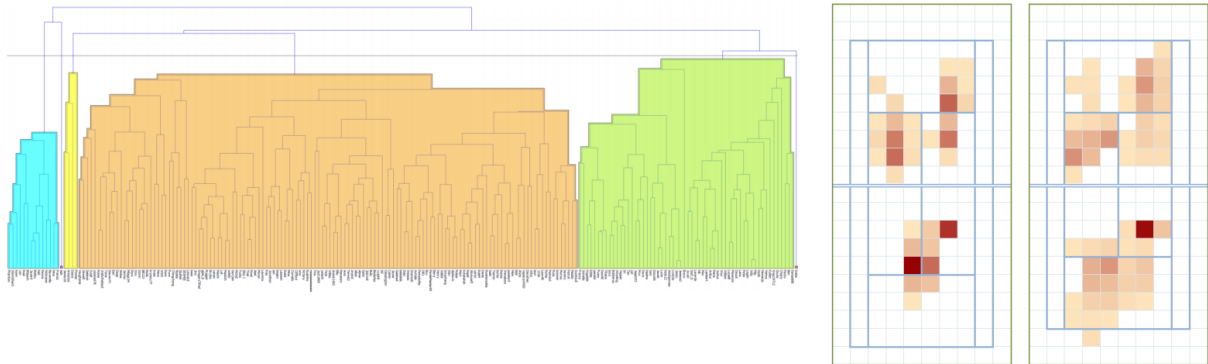


Fonte: Página oficial do jogo na loja Steam (PLAYDEAD, 2011)

Rodrigues e Brancher (2018) utilizaram a técnica de clusterização, que consiste no agrupamento de dados com características semelhantes, para identificar perfis de jogadores conforme uma série de métricas. Mozgovoy (2018) aliou a clusterização com mapas de calor, que consistem em representações da frequência de ocorrências de algum evento por meio da intensidade das cores (NELSON, 2011; DRACHEN; SCHUBERT, 2013). Os autores utilizam as duas técnicas para localizar em um jogo de tênis quais são as regiões que têm maior frequência de ações de movimento e de arremesso e entender os perfis de jogadores com base nelas (Figura 10). Em ambos os trabalhos, a compreensão sobre os jogadores é apontada como fator determinante para o balanceamento da dificuldade em jogos.

Argumentando que há uma falta de pesquisas que ilustram o uso de dados biométricos para aprimorar o desenvolvimento de jogos casuais, Gualeni *et al.* (2012) usaram aparelhos biométricos em conjunto com entrevistas, gravações e questionários para compreender a experiência de jogo no quebra-cabeças Gua-Le-Ni, de autoria própria. Os autores perceberam

Figura 10 – Combinação das técnicas de *heatmap* e clusterização utilizadas por Mozgovoy (2018). Os clusters à esquerda agrupam jogadores conforme comportamentos similares. O mapa à direita representa por meio da intensidade da cor vermelha os locais onde houve ações de movimento e arremesso frequentes num jogo de tênis



Fonte: Mozgovoy (2018).

uma relação entre a diversão e o estresse no jogo, que pôde ser explicitada graças aos aparelhos biométricos.

Mirza-Babaei *et al.* (2013a) investigaram o uso de biometria desenvolvendo uma nova técnica chamada *Biometric Storyboards* (BioST), que mostra dados biométricos em pontos-chave do jogo junto com a expectativa do *designer* sobre o quão empolgante esses momentos devem ser. Um exemplo de BioST pode ser visto na Figura 11. Os autores realizaram o seguinte experimento: recrutaram seis *game designers*, que desenvolveram iterativamente três versões de um jogo de plataforma: uma versão inicial sem modificações, uma aprimorada com técnicas tradicionais de GUR e visualização de dados, e outra aprimorada com BioST. Os resultados da comparação da experiência dos jogadores nos três protótipos mostraram que a biometria não trouxe melhora na usabilidade geral, porém trouxe melhora na qualidade da experiência.

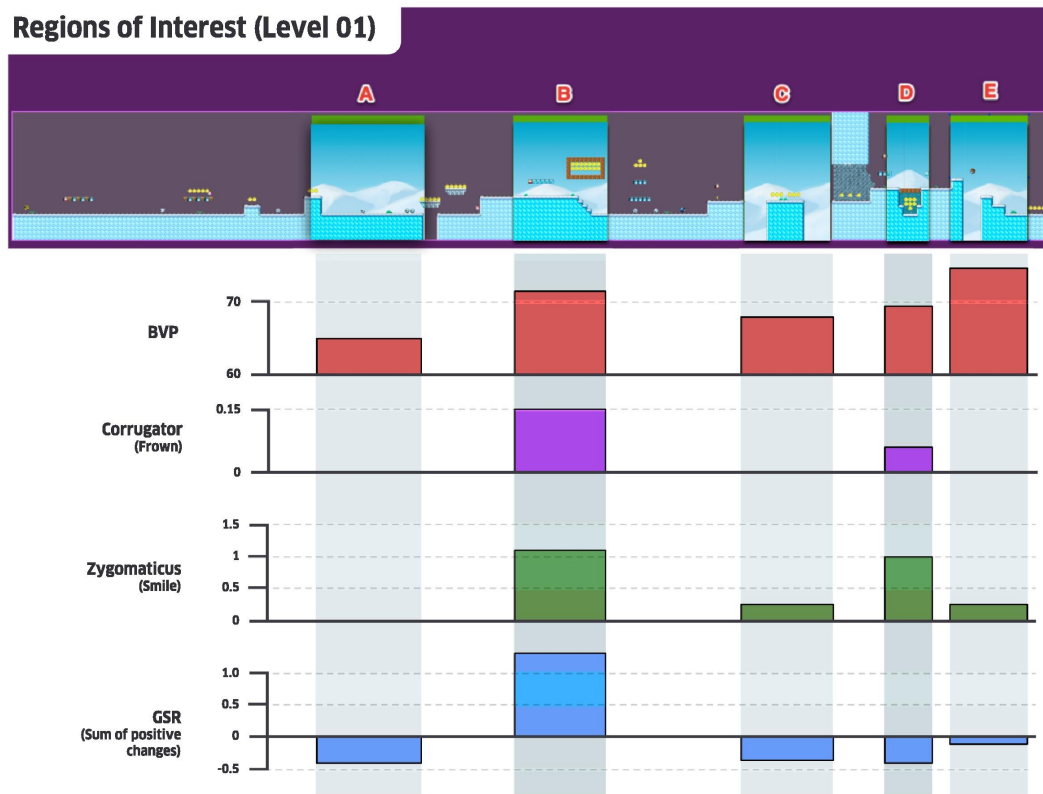
Gómez-Maureira *et al.* (2014) realizaram um estudo comparativo entre três metodologias de GUR (entrevistas, métricas e biometria) para averiguar de que forma elas contribuem na qualidade dos níveis de um jogo de plataforma e quais combinações propiciam resultados mais efetivos. Os autores modificaram o jogo Supertux, gerando três versões diferentes combinando pares de metodologias: uma modificada com base em dados advindos de entrevistas e métricas; outra via entrevista e biometria; e a última por meio de métricas e biometria. As análises foram feitas utilizando visualização de dados em cada caso, um exemplo delas está na Figura 12. Os autores concluíram que a entrevista é o método que provê o *feedback* mais construtivo e, portanto, deve ser incluído sempre que possível. Adicionalmente, a combinação de entrevista com métricas proporciona o acesso a dados quantitativos e qualitativos, além de responder satisfatoriamente a perguntas do tipo “como?”.

Figura 11 – Um exemplo de Biometric Storyboard. Há uma sequência de pontos-chave. Abaixo, têm-se as variações do grau de excitação do jogador, exibidas como uma linha azul. Ícones na linha marcam pontos de interesse. Esses pontos são descritos abaixo da linha



Fonte: Traduzido de HCT (2012).

Figura 12 – Visualização de dados biométricos em uma fase do jogo SuperTux, utilizada por Gómez-Maureira *et al.* (2014). A figura mostra regiões de interesse na fase bem como os níveis de batimentos cardíacos, franzimentos, sorrisos e condutividade da pele médios



Fonte: Gómez-Maureira *et al.* (2014)

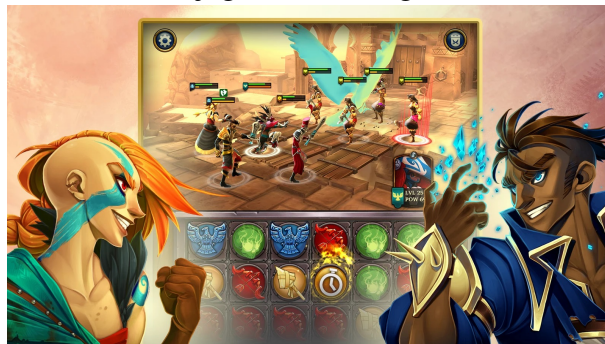
Os trabalhos de Mirza-Babaei *et al.* (2013a) e Gómez-Maureira *et al.* (2014) são especialmente valiosos para esta pesquisa, uma vez que não só apresentam conclusões acerca da experiência de jogo em seus respectivos objetos de estudo, como também refletem sobre os

benefícios da aplicação dos métodos de GUR utilizados e apresentam metodologias para o teste comparativo entre versões de jogos alteradas segundo critérios distintos.

Uma boa experiência de jogo também está relacionada ao aprendizado, que, no contexto dos jogos, se refere à compreensão das mecânicas (PLASS *et al.*, 2011). Uma consequência direta deste tipo de aprendizado é a aptidão para resolver desafios progressivamente mais difíceis. Se um jogador encontrar um desafio sem estar preparado para isso, pode gerar uma situação potencialmente frustrante, eventualmente ocasionando o abandono do jogo. Por outro lado, a combinação de conceitos pré-estabelecidos e a exploração criativa dos mesmos está associada à diversão, logo, proporcionam uma experiência prazerosa (JUUL, 2010). A seguir, são apresentados trabalhos cuja premissa é melhorar os processos de ensino e aprendizado em jogos.

Andersen *et al.* (2012) estudaram o impacto que tutoriais trazem em jogos de diferentes complexidades. Os autores desenvolveram oito tutoriais para três jogos diferentes de dificuldades e estilos variados. Métricas de mais de 40 mil jogadores mostraram que a presença de um tutorial é mais útil em jogos mais complexos. Morin *et al.* (2016) concluíram por meio de seu estudo de caso com o jogo Saber's Edge (Figura 13) que a presença de tutoriais explícitos compromete a retenção de usuários mais experientes, mas favorece a retenção de usuários inexperientes com jogos.

Figura 13 – Material promocional do jogo Saber's Edge



Fonte: Divulgação oficial no Unity Connect (DUPUIS, 2016)

Wang *et al.* (2020) utilizaram dados advindos de observação, entrevistas e *eye-tracking* em um jogo de estratégia para entender as diferenças entre o processo de aprendizado de jogadores novatos e experientes. Os autores descobriram que os novatos utilizam de tentativa e erro para aprender, enquanto que os experientes mapeiam todas as possíveis ações para a obtenção de um resultado. Segundo (HOFFMANN *et al.*, 2019), os jogos podem incentivar o perfeccionismo quando apresentam incentivos à realização de jogadas ótimas e avaliações de

desempenho. As conclusões desses trabalhos indicam que o que separa os jogadores novatos dos veteranos é a facilidade de compreensão de causa e efeito das mecânicas dos jogos, e que os veteranos podem se tornar perfeccionistas nos jogos que mais gostam.

Linehan *et al.* (2014) analisaram quatro jogos de sucesso, Portal (Figura 14), Portal 2, Braid e Lemmings, a fim de identificar padrões sobre como apresentar os desafios aos jogadores de acordo com o que eles já aprenderam. Os autores puderam identificar alguns padrões comuns aos jogos avaliados: as principais mecânicas são introduzidas separadamente em desafios que costumam requerer apenas a repetição de uma ação básica. Logo em seguida, quebra-cabeças reforçam o aprendizado daquela mecânica para, por fim, combiná-la com outras mecânicas prévias. A dificuldade desses desafios é crescente até a apresentação de uma nova mecânica. Scozzi *et al.* (2017) também analisaram o jogo Portal, seguindo outra metodologia: determinaram qual é o número mínimo de ações necessárias pra completar os níveis do jogo e compararam-no com o número de ações que os jogadores realizavam. Foi possível, assim, identificar níveis que apresentam curva de dificuldade elevada.

Figura 14 – Screenshot do jogo Portal



Fonte: Página oficial do jogo Portal (VALVE, 2007)

Iacovides *et al.* (2013) estudaram outros conceitos pertinentes, os de *breakdown* e *breakthrough*, referentes aos momentos de confusão causados por sucessivas falhas e a superação consciente de um obstáculo, respectivamente. Os autores argumentaram que dados psicofisiológicos não contribuíam para a identificação de *breakdowns* e *breakthroughs*. Estes puderam ser identificados apenas com os relatos dos participantes, reforçando ainda mais a importância de entrevistas.

Por serem artefatos inerentemente educativos, os jogos são frequentemente aplicados

em contextos educacionais. Nos jogos educativos, o aprendizado de uma mecânica é atrelado ao de algum assunto escolar (DÖRNER *et al.*, 2016). A avaliação do aprendizado escolar pode, nesse caso, ser utilizada como método de balanceamento do jogo (ALONSO-FERNÁNDEZ *et al.*, 2019).

Kohwalter *et al.* (2012) abordaram o aprendizado sob a perspectiva da compreensão dos próprios erros em um jogo. Os autores captaram dados da sessão de jogo e construíram visualizações que explicitam as relações de causa e efeito das ações dos jogadores. Os autores posteriormente expandiram sua abordagem para incluir o balanceamento de jogos no geral (KOHWALTER *et al.*, 2018) e até outros domínios, como mapeamento de tráfego (KOHWALTER *et al.*, 2016).

3.3 Outros Usos de Visualização de Dados

Nesta seção serão abordados trabalhos que utilizam visualização de dados mas não se enquadram na seção anterior por não terem como objetivo melhorar um jogo. São trabalhos que utilizam visualização para fornecer informações aos jogadores, ou que tenham como finalidade a descrição de uma técnica ou ferramenta de visualização.

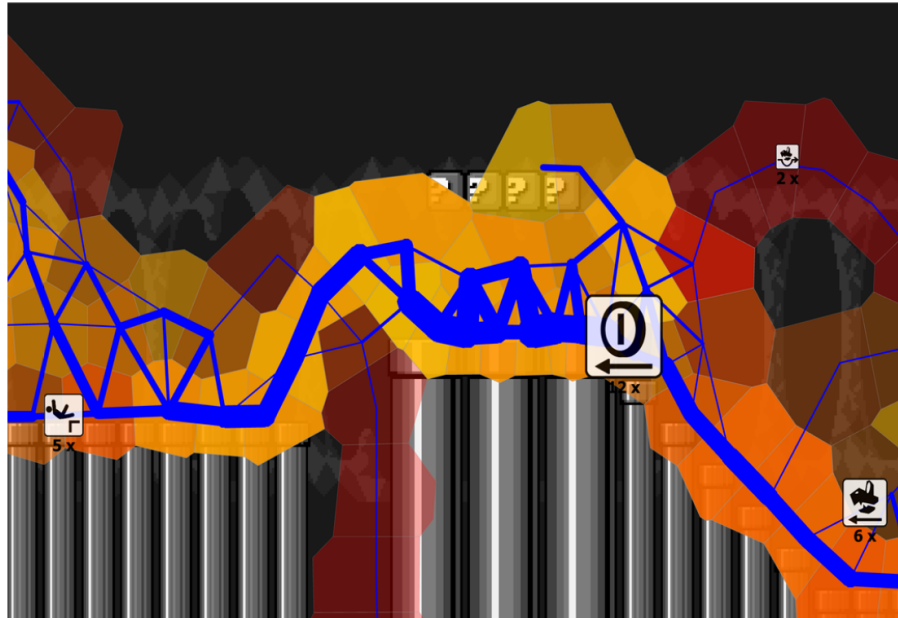
Siqueira *et al.* (2017) utilizaram análise de dados para prever quando aconteceriam desistências ou cancelamentos de assinaturas em jogos multijogador online. Feitosa *et al.* (2015) desenvolveram uma ferramenta que auxilia jogadores competitivos de Dota 2 ao mostrar detalhes sobre o *metagame*.³

Wallner *et al.* (2019) e Halabi *et al.* (2019) adaptaram a técnica de clusterização ao desenvolverem uma visualização que mistura dados de trajetória com dados coletados por aparelhos biométricos, relacionando locais no espaço do jogo com o grau de excitação dos jogadores. Em seu trabalho, os *clusters* são associados a uma codificação visual que os relacionam a algum fenômeno no jogo, tais como um pulo ou uma queda (ver Figura 15). Os símbolos são dimensionados conforme o número de ocorrências.

Diversos autores propuseram ferramentas de visualização. Entre elas destacam-se o PLATO (WALLNER; KRIGLSTEIN, 2014) e o Vixen (DRENIKOW; MIRZA-BABAEI, 2017), para visualização de dados de métricas de jogo no geral. A Figura 16 ilustra uma visualização espacial criada pelo Vixen. Algumas ferramentas são desenvolvidas com usos específicos em

³ Qualquer tipo de informação que transcenda o jogo em si mas que possa influenciar as partidas. Por exemplo, saber quais personagens são escolhidos com maior frequência pode influenciar os jogadores a treinar especificamente contra esses personagens durante as preparatórias para um campeonato.

Figura 15 – Técnica de clusterização utilizada por Wallner *et al.* (2019). As linhas azuis mostram trajetórias. As manchas alaranjadas mostram o nível de excitação dos jogadores naquela região. Ícones mostram ações. O ícone de moeda, por exemplo, indica que os jogadores voltaram para bater nas caixas e pegar moedas com frequência no trecho ilustrado



Fonte: Halabi *et al.* (2019)

mente. É o caso do ShowRunner (SANGHRAJKA *et al.*, 2019), capaz de visualizar fluxos de narrativas.

Kepplinger *et al.* (2020), por exemplo, propuseram uma visualização que agrega dados de emoção, olhar e emoções. A visualização aproveitou a estrutura de componentes da Unity para se estruturar como uma ferramenta, assim como o Vixen. Ela foi utilizada em um jogo 3D em primeira pessoa disponível como demonstração da Unity ⁴.

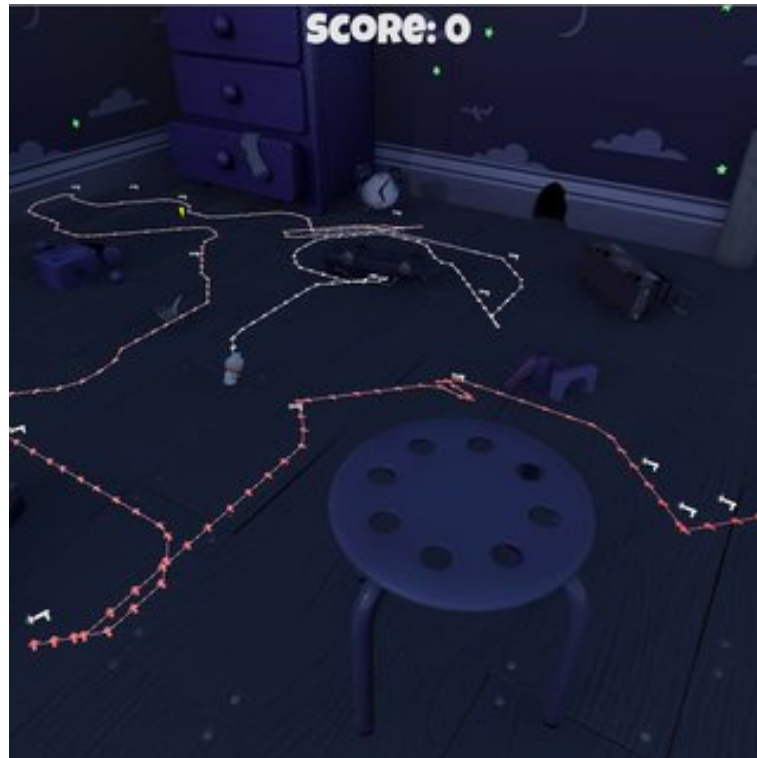
3.4 Discussão

A maioria dos trabalhos utiliza um ou mais dos métodos tradicionais como entrevista, questionário e observação devido a sua eficiência e facilidade de aplicação. São também os melhores métodos para obtenção de opiniões, críticas e sugestões dos jogadores. Este trabalho adotará estas metodologias, uma vez que as conclusões de Gómez-Maureira *et al.* (2014) e Iacovides *et al.* (2013) reforçam a necessidade do uso dessas metodologias.

Há também uma preferência pelas métricas. Conforme mostrado pelo trabalho de Bosnjak e Orehovacki (2018), a implementação de rastreadores é simples e confiável. Além disso, o método é essencial para a produção de visualizações espaciais, que fornecem informações

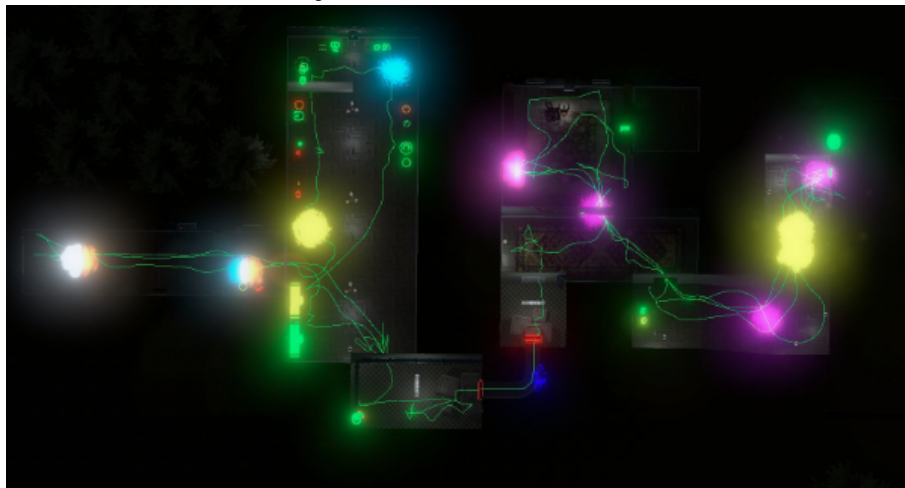
⁴ <https://unity3d.com/>

Figura 16 – Exemplo de visualização criada pelo Vixen em um projeto de demonstração fornecido pela Unity. As linhas mostram trajetórias, os ícones ações no jogo, como atirar, e as cores o nível de excitação dos jogadores.



Fonte: Drenikow e Mirza-Babaei (2017)

Figura 17 – Técnica de agregação de dados utilizada por Kepplinger *et al.* (2020). Linhas verdes mostram a trajetória do jogador, outlines coloridas nos objetos representam informações sobre o olhar e as partículas codificam emoções.



Fonte: Kepplinger *et al.* (2020)

imprescindíveis para a maioria dos jogos de ação. Para a aplicação das métricas, existem duas opções: modificar um jogo pré-existente, adicionando rastreadores ao código, ou criar um novo jogo e estruturá-lo tendo em mente as métricas que serão rastreadas. Esta última é preferível, uma vez que permite total controle criativo e funcional acerca do objeto de estudo.

Quanto à biometria, raramente os autores comentam sobre desvantagens relevantes como o custo dessas tecnologias, tempo de uso necessário para se obterem os dados e a necessidade de profissionais capacitados. Com o intuito de manter alguns dos ganhos dos métodos biométricos sem os ônus citados, decidiu-se optar pela utilização do *eye-tracking*. No contexto dos jogos, o *eye-tracking* é comumente aplicado em interfaces gráficas (ÍLHAN, 2018) e análise de cenário (ALMEIDA *et al.*, 2016; PETERSSON; HELGESSON, 2018). A inclusão dessa técnica como ferramenta de *level design* se dá pelas peculiaridades do jogo desenvolvido neste trabalho. Nele, a atenção por si só é uma grande agente na jogabilidade.

O uso dos métodos escolhidos aliados às técnicas de visualização é capaz de fornecer dados ricos e variados sobre múltiplos aspectos relevantes ao *game design*. A Tabela 3 resume quais métodos foram escolhidos e o tipo de dado que eles fornecem.

Tabela 3 – Resumo das Metodologias de GUR adotadas neste trabalho

| Método | Dado Fornecido |
|-------------------------|---------------------------------------------------------------------------------------------|
| Observação Participante | Comportamentos do testador enquanto ele joga, problemas de <i>level design</i> |
| Entrevista | Opiniões, comentários, críticas e sugestões |
| Questionário | Dados majoritariamente quantitativos acerca do <i>game design</i> ou da experiência de jogo |
| Métricas de jogo | Comportamentos do jogador e ocorrências de eventos dentro do jogo |
| <i>Eye-tracking</i> | Focos de atenção |

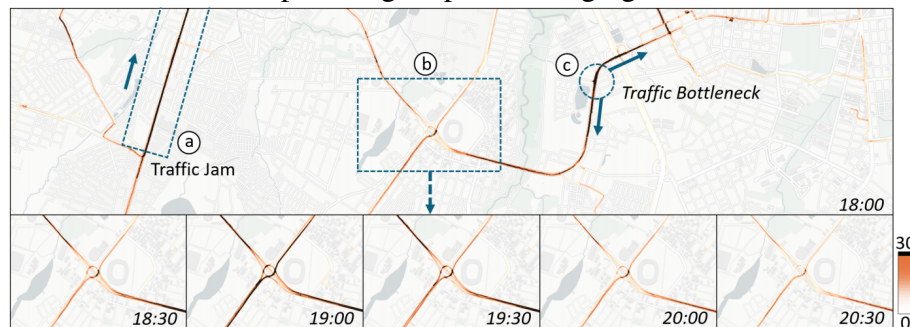
Fonte: Elaborada pelo autor com base no que foi discutido nos Capítulos 2 e 3.

A maior parte dos trabalhos utiliza algum tipo de técnica de visualização para analisar comportamentos dos usuários. Ao contrário de Wallner e Kriglstein (2014), Drenikow e Mirza-Babaei (2017) e Sanghrajka *et al.* (2019), que projetaram criadores de visualizações que funcionam em jogos quaisquer, este trabalho segue uma linha mais próxima à dos trabalhos citados como aplicações na indústria: as visualizações serão feitas mediante as necessidades específicas do jogo objeto de estudo.

No caso deste projeto, criou-se uma nova técnica de visualização capaz de aliar os diversos tipos de dados adquiridos pelas métricas de jogo e pelo *eye-tracking*. Conforme comentado anteriormente, as visualizações espaciais são essenciais para jogos de ação, como é o caso de Downside Up, o jogo desenvolvido aqui. Foi decidido adaptar a técnica de rotas quentes (Figura 18), proposta por Gomes *et al.* (2018).

Os autores aplicaram originalmente a técnica na representação de vias onde há tráfego mais intenso para auxiliar no controle de trânsito. Ao contrário das vias, que limitam o movimento a rotas pré-definidas, o avatar em Downside Up move-se livremente no plano 2D, tendo, assim, comportamento imprevisível. Embora haja uma perda de informação sobre direção

Figura 18 – Visualização de rotas quentes em um mapa. A intensidade do trânsito é ilustrada pela tonalidade vermelha. A cor preta sugere potenciais gargalos nas vias



Fonte: Gomes *et al.* (2018).

e sentido, a informação mais importante, que é a de presença ao longo do tempo em determinada região, se mantém. Ademais, diferente das vias, é possível tomar providências rápidas e alterar facilmente espaço do jogo caso a visualização aponte comportamentos indesejados pelos designers.

Tabela 4 – Comparação entre este trabalho e os trabalhos mais importantes

| | Mirza-Babaei <i>et al.</i> (2013a) | Gómez-Maureira <i>et al.</i> (2014) | Este trabalho |
|------------------------------------------------------|------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Técnicas de GUR | Observação, questionário, entrevista, biometria (eletromiografia, condutividade da pele) | Entrevistas, métricas, biometria (eletromiografia, condutividade da pele, pressão sanguínea) | Observação, questionário, entrevista, métricas, <i>eye-tracking</i> via webcam |
| Metodologia | Testar múltiplas versões de um protótipo | Testar múltiplas versões de um protótipo | Testar múltiplas versões de um protótipo |
| Abordagem | Adições sucessivas de metodologias de testes | Combinações de pares de metodologias distintas | Adições sucessivas de metodologias de testes |
| Validação das Mudanças de <i>level design</i> | Questionários de Experiência (SUS, PANAS, SAM) | Questionário de Experiência (GEQ) | Questionários de Experiência (GEQ, PANAS) |
| Tipo de jogo | Modificação de jogo produzido por terceiros | Modificação de jogo código-aberto | Criação de novo jogo autoral |
| Estilo de jogo | Plataforma 2D casual | Plataforma 2D casual | Plataforma 2D <i>hardcore</i> |
| Game designer externo à pesquisa | Sim | Não | Não |
| Utiliza nova técnica de visualização | Sim | Não | Sim |
| Visualização avaliada por profissionais? | Sim, porém em outro artigo (MIRZA-BABAEI <i>et al.</i> , 2011) | Não | Sim |

Fonte: Elaborada pelo autor.

Tendo em vista as considerações feitas neste capítulo, propõe-se um trabalho similar aos de Mirza-Babaei *et al.* (2013a) e Gómez-Maureira *et al.* (2014), adotando diferentes técnicas. A Tabela 4 mostra uma comparação dos trabalhos. Visa-se avaliar se as métricas de jogo e o *eye-tracking* complementam os métodos tradicionais em um jogo de plataforma 2D com

mecânicas não convencionais. Neste trabalho, o *eye-tracking* substitui a psicofisiologia, adotada pelos autores de referência, por poder ser utilizado de forma confiável e acessível via *web-cam* para o jogo estudado. Este trabalho difere-se também por introduzir dois artefatos devidamente testados com seus públicos-alvo, um jogo e uma técnica de visualização, tornando o trabalho interessante tanto para acadêmicos quanto para profissionais da área.

Por fim, este trabalho discute três tópicos principais: o aperfeiçoamento de jogos, a eficácia de diferentes métodos de GUR e a importância de visualização de dados para *game designers*.

3.5 Considerações Finais

Nas seções anteriores, foram descritas e discutidas outras abordagens que utilizam GUR e visualização de dados em jogos. Os trabalhos relacionados serviram de base para a definição dos métodos que serão aplicados. Diversas decisões sobre abordagens metodológicas foram tomadas com base no estado da arte da pesquisa com usuários de jogos. Estas decisões serão apresentadas no Capítulo 4.

4 ABORDAGEM PROPOSTA

Neste capítulo, está presente a abordagem metodológica proposta. A Seção 4.1 introduz e descreve as etapas da metodologia. A Seção 4.2 descreve o jogo desenvolvido durante a pesquisa. A Seção 4.3 discorre sobre os dados coletados dos usuários durante os testes para aperfeiçoamento dos protótipos ou para comparação da experiência de jogo. A Seção 4.4 mostra as estratégias de visualização empregadas e como elas são validadas. As considerações finais são feitas na Seção 4.5.

4.1 Metodologia

A metodologia visa analisar o quanto métodos de GUR melhoram o *level design* de jogos 2D. A metodologia foi aplicada durante o desenvolvimento do jogo *Downside Up*, que será detalhado adiante.

Este trabalho considera, no contexto desta pesquisa, que a qualidade geral do *design* está associada diretamente com o quão divertido um jogo é percebido pelos jogadores, conforme discutido no Capítulo 2. Sendo assim, o protótipo de maior sucesso seria aquele em que os usuários reportarem ter tido experiências mais positivas do que negativas. É importante salientar que a análise da qualidade gráfica está fora do escopo do projeto.

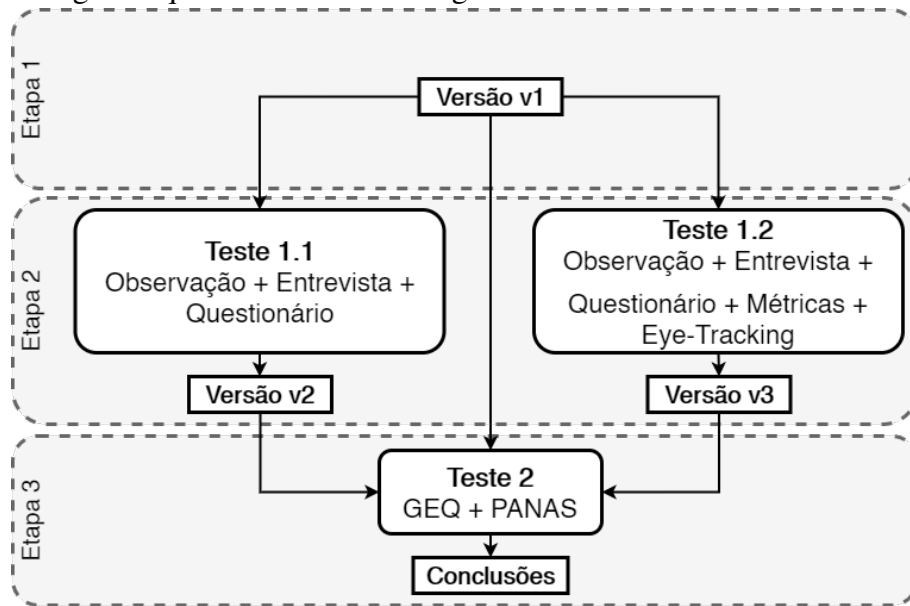
As pesquisas em GUR seguem um paradigma de estudos majoritariamente empíricos (CARTER *et al.*, 2014). Este trabalho segue a mesma linha.

Conforme dito no Capítulo 1, a metodologia procura oferecer meios pelos quais seja possível responder à seguinte pergunta: “como combinar de forma eficaz métodos de Testes com Usuários e Visualização de Dados para aperfeiçoar um jogo de plataforma 2D?”. A metodologia proposta foi dividida em três etapas, detalhadas a seguir. A Figura 19 apresenta um resumo das etapas.

Na Etapa 1, ocorreu o desenvolvimento do protótipo inicial de um jogo.

Em seguida, na Etapa 2, houve testes com usuários e desenvolvimento de duas versões aperfeiçoadas, v2 e v3. No primeiro teste com v1 (Teste 1.1) foram aplicados observação, entrevista e questionário de avaliação da dificuldade dos desafios. Os resultados desse teste determinaram as alterações nos desafios. A versão v2 foi o resultado destas alterações em v1. Similarmente, no segundo teste com v1 (Teste 1.2), foram aplicados observação, entrevista e questionário novamente, acrescentados às métricas de jogo e ao *eye-tracking*. As alterações

Figura 19 – Diagrama que mostra a metodologia deste trabalho



Fonte: Elaborada pelo autor.

sugeridas pelo resultado desse teste geraram a versão v3.

Por último, na Etapa 3, após jogarem alguma das três versões, os jogadores responderam a dois questionários de avaliação de experiência, GEQ (IJSSSELSTEIJN *et al.*, 2013) e *Positive and Negative Affect Scale* (PANAS) (WATSON *et al.*, 1988). Os resultados foram comparados para averiguar se as versões aperfeiçoadas por GUR são mais divertidas que o protótipo original e, em caso positivo, qual das duas tem o melhor desempenho.

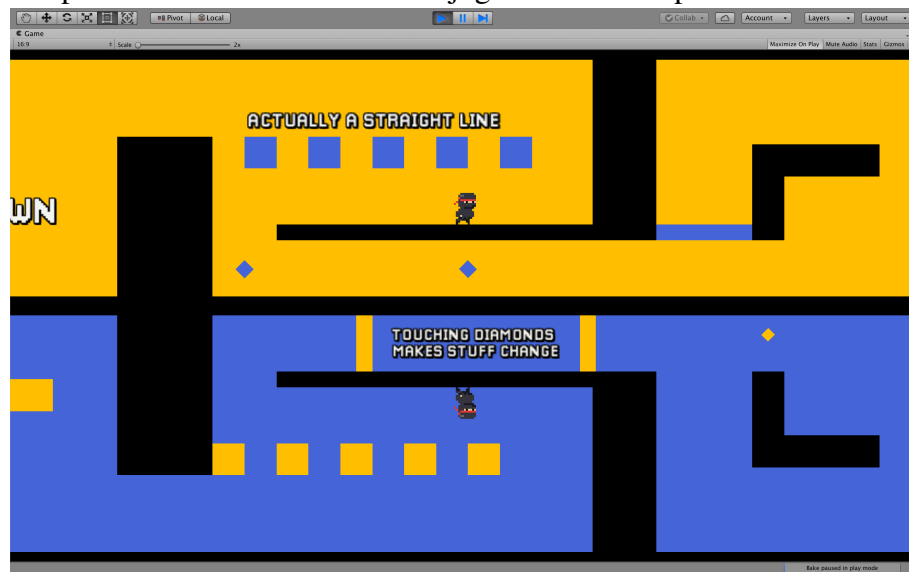
A seção seguinte descreve o jogo ao qual a metodologia foi aplicada.

4.2 Downside Up

Downside Up¹ é um jogo de plataforma 2D no qual jogadores devem guiar um avatar presente em duas telas espelhadas. Objetos vistos em uma tela podem não aparecer na outra tela, embora existam nas duas telas. Dessa forma, jogadores devem utilizar sua visão e memória para superar os desafios. A Figura 20 mostra uma captura de tela do jogo. O jogo é composto por três níveis, cada um deles composto por desafios menores. Ao todo, existem 18 desafios no jogo. Cada desafio foi rotulado com um número seguido de uma letra, ordenado alfabeticamente. Dessa forma, o desafio “3D”, por exemplo, representa o quarto desafio (“D”), da terceira fase (“3”). A duração esperada para completar o jogo é de 40 a 60 minutos. Optou-se pelo desenvolvimento de um jogo autoral por conta da liberdade criativa e do controle do código-fonte.

¹ O jogo e material suplementar, que inclui descrições ilustradas de todos os desafios e passo-a-passo de cada fase, estão disponíveis em: <https://arthursb.github.io/Downside-Up>

Figura 20 – Captura de tela do desafio 1C do jogo Downside Up



Fonte: Fornecida pelo autor.

4.2.1 Motivação

Dados de 2017 mostraram que jogos de ação estavam entre os gêneros de jogos mais populares dos Estados Unidos (GOUGH, 2019), cujo mercado de jogos é o maior do mundo. Os jogos de plataforma encontram-se em alta, conforme noticiado pela mídia (BRAUN, 2018; ACADEMY, 2014). Ganharam notoriedade novamente após o jogo Celeste (Figura 21) ter sido indicado a diversos prêmios em competições de “melhor jogo do ano”². Ademais, o *level design* de jogos de plataforma desperta interesse acadêmico até hoje (KHALIFA *et al.*, 2019). Baseado nisso, optou-se pelo desenvolvimento de um jogo de plataforma, subgênero dos jogos de ação. A simplicidade e as mecânicas básicas populares (andar e pular) trazem como vantagem a redução do tempo gasto pelos jogadores para dominar os controles. Os vários recursos já disponíveis para programação de jogos no estilo também reduzem do tempo de desenvolvimento. Isso tudo torna o processo de testes mais rápido e manejável.

Com base nos jogos Celeste³ e I Wanna be The Guy⁴, foi decidido seguir o princípio dos jogos *masocore*, denominação informal dada a jogos 2D com dificuldade excruciante, elementos que estimulam comportamentos de tentativa e erro. Elementos clássicos desse tipo de jogo incluem abundância de espinhos e mortes pouco punitivas no jogo.

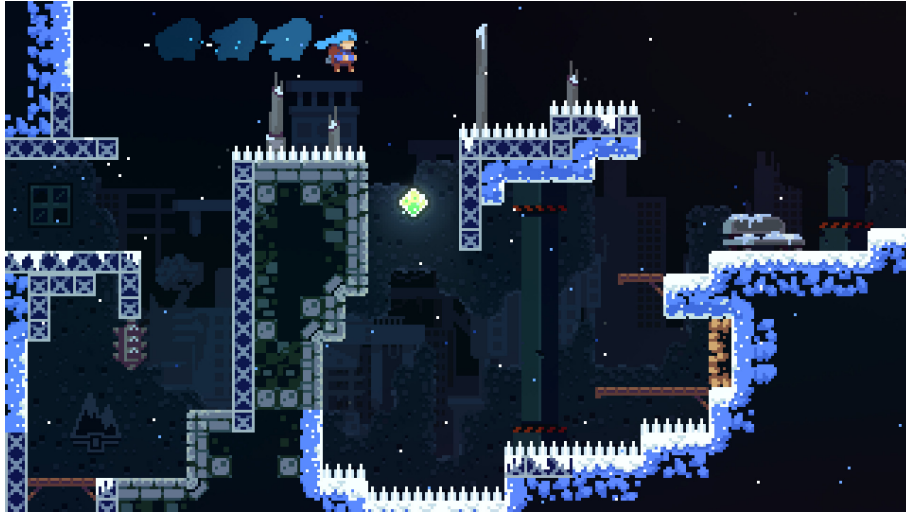
Os dois jogos selecionados, no entanto, apresentam diferenças fundamentais de *design*. Celeste preza pelo aprendizado do jogador, oferecendo desafios em dificuldade crescente

² <https://www.imdb.com/title/tt7903902/awards>

³ <http://www.celestegame.com/>

⁴ <http://kayin.moe/iwbgt/index.php>

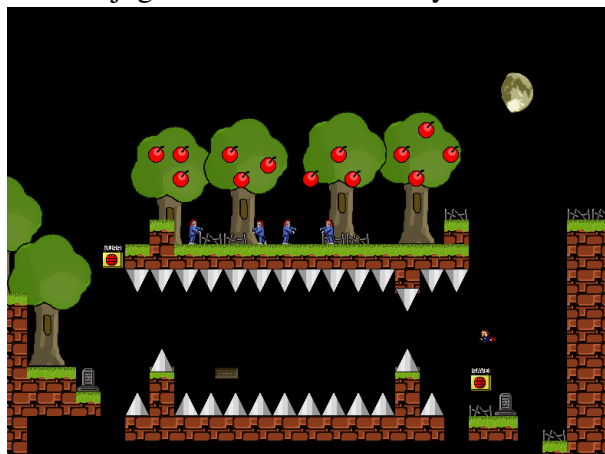
Figura 21 – Captura de tela do jogo Celeste.



Fonte: Página oficial do jogo na Steam (INC, 2018).

conforme novas mecânicas são apresentadas. O *design* estimula a criatividade e a combinação de conhecimentos prévios. O jogo, no entanto, requer bastante atenção e respostas rápidas do jogador. I Wanna Be The Guy, por sua vez, explora a dificuldade como forma de satirizar os jogos de plataforma. Diferentemente de Celeste, os desafios não são pensados na apresentação, combinação e exploração de mecânicas. Ao invés disso, o jogo prega peças nos jogadores, apresentando comportamentos imprevisíveis, tais como espinhos caindo do céu sem nenhum indicativo visível ou explosões aleatórias. No desafio mostrado na Figura 22, por exemplo, as maçãs podem cair da árvore e matar o jogador em momentos imprevisíveis à primeira vista.

Figura 22 – Captura de tela do jogo I Wanna Be The Guy.

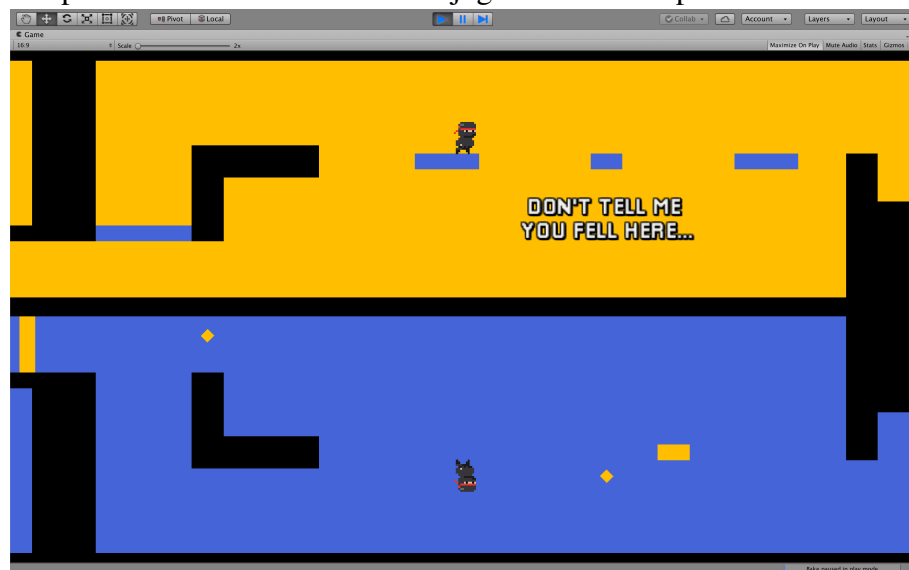


Fonte: Página oficial do jogo (KAYIN, 2007).

4.2.2 Game Design

O *design* de *Downside Up* encontra-se em um meio termo entre os dois jogos. Se por um lado há apreço pelo aprendizado do jogador e pela utilização criativa de mecânicas, por outro lado há também um elemento de surpresa. Ao contrário de *I Wanna Be The Guy*, no entanto, esses elementos não são imprevisíveis, mas podem surpreender os jogadores desatentos. No desafio 1D, mostrado na Figura 23, jogadores desatentos podem encostar em um losango que desativa uma plataforma logo acima do texto e cair no chão. Essas situações acontecem quando jogadores não percebem que há elementos distintos nas duas telas do jogo, mecânica que será explicada a seguir.

Figura 23 – Captura de tela do desafio 1D do jogo *Downside Up*



Fonte: Fornecida pelo autor.

Semelhante aos jogos *Chronos Twins*⁵ e *Binaries*⁶, usuários coordenam um personagem em duas telas simultaneamente. Obstáculos estão presentes em ambas as telas, porém possivelmente visíveis em apenas uma. Diferente de *Chronos Twins* (Figura 24), em *Downside Up* as telas são espelhadas. Também diferente de *Binaries*, no qual há dois avatares distintos presentes na mesma tela que compartilham os mesmos controles do jogador, em *Downside Up* o avatar é único, porém exibido em duas telas.

No desafio 1C, ilustrado anteriormente na Figura 20, no começo da seção, por exemplo, há uma linha reta na parte de cima do nível, porém esta é dividida em vários segmentos cuja visibilidade alterna entre as duas telas.

⁵ <http://www.enjoyup.com/>

⁶ <http://binaries.ant-workshop.com/>

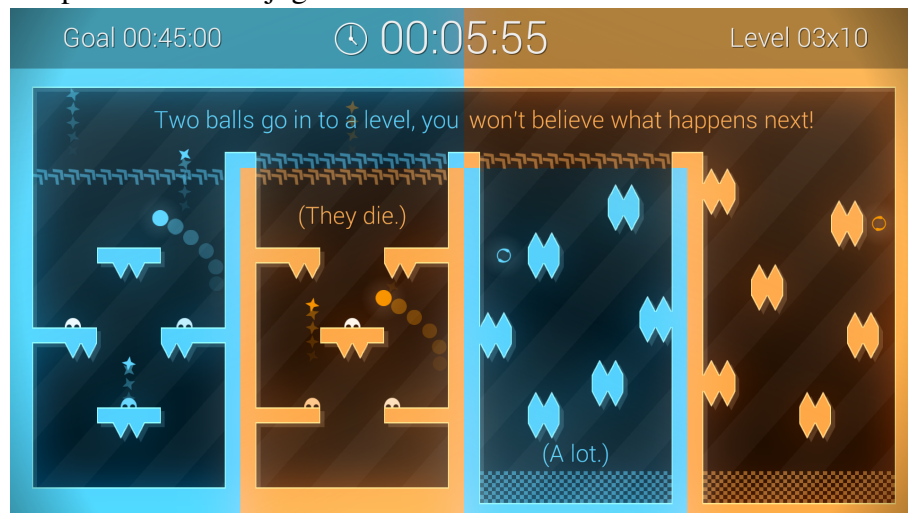
Figura 24 – Captura de tela do jogo Chronos Twins



Fonte: Página do jogo no banco de dados IGDB (IGDB, 2010).

O jogo é *offline e single player*, não há nenhum meio de interação entre jogadores, sendo assim a sensação é de que o jogador está jogando contra a máquina. Decidiu-se inserir um narrador para enriquecer esta interação, uma vez que ela pode ser potencialmente divertida. A Figura 23 mostra um caso de narração inspirada no narrador de Binaries (Figura 25).

Figura 25 – Captura de tela do jogo Binaries

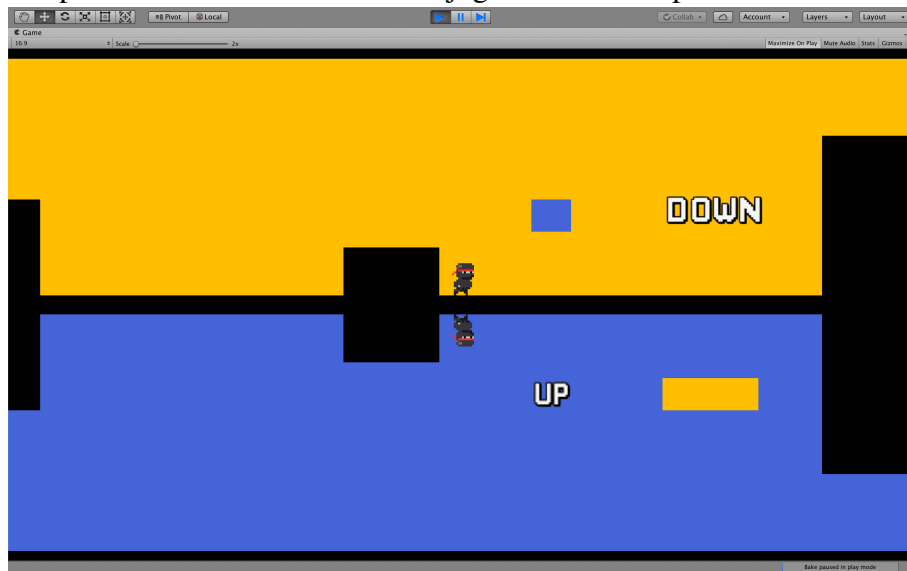


Fonte: Página oficial do jogo Binaries (WORKSHOP, 2016)

O narrador deixa mensagens irônicas ou informativas no plano de fundo desafiando ou provocando o jogador, bem como dando dicas sobre os desafios. Em alguns casos, como o do desafio mostrado na Figura 26, logo adiante, a existência do narrador possibilita tutoriais interessantes. O texto, no exemplo citado, está posicionado exatamente atrás de onde cada plataforma está, mas só aparece nas telas onde as respectivas plataformas estão invisíveis,

ilustrando a mecânica principal do jogo e dando uma dica de que o jogo requer atenção em duas telas simultaneamente. Esse recurso também é explorado no desafio 1C, que encontra-se na Figura 20, para concentrar a atenção do jogador em uma reta (parte superior do cenário) e para explicar o funcionamento de uma nova mecânica (parte inferior do cenário).

Figura 26 – Captura de tela do desafio 1B do jogo Downside Up



Fonte: Fornecida pelo autor.

Por meio do SteamSpy⁷, uma ferramenta de exploração de dados da loja Steam⁸, descobriu-se que jogadores de Celeste e Binaries descrevem esses jogos com rótulos como “difícil”, “quebra-cabeças de plataforma”, “minimalista”, “engraçado” e “agitado”. Estes rótulos sugerem não só que há um público em comum, como também explicitam características do design de jogos *masocore*.

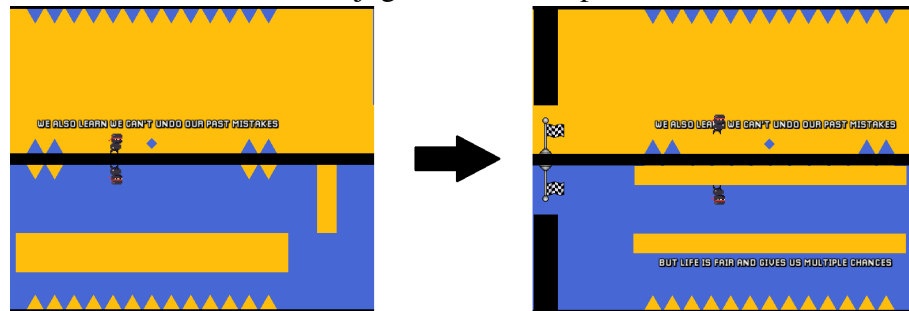
Inspirado pelo jogo Sometimes You Die⁹, há uma mecânica na qual as ações do jogador no mundo não são desfeitas quando ele morre. Dessa forma, é possível criar quebra-cabeças com resoluções inusitadas. Em Downside Up, por exemplo, se o personagem ativar uma armadilha que faz uma plataforma com espinhos cair e for esmagado pela plataforma, ao morrer e retornar àquela localização, a plataforma continuará no chão e a armadilha não precisará ser ativada novamente. Essa situação está ilustrada na Figura 27. A adição dessa mecânica foi pensada no fato de os jogadores morrerem constantemente em decorrência de espinhos. Ela torna o ato de morrer no jogo ligeiramente menos frustrante.

⁷ <https://steamspy.com/>

⁸ <https://store.steampowered.com/>

⁹ <http://www.kamibox.de/sometimesyoudie>

Figura 27 – Estados do desafio 2B do jogo Downside Up



Fonte: Fornecida pelo autor.

O jogo segue um estilo minimalista, que utiliza poucos elementos para compor o cenário, com o intuito de trazer o foco para os desafios e evitar sobrecarga de informação, uma vez que os jogadores devem lidar com duas telas ao mesmo tempo. Há pouca variedade de cores, cada uma representando um comportamento específico. Para evitar possíveis tonturas, os desafios foram planejados de forma que não sejam necessários movimentos constantes dos olhos entre as duas telas.

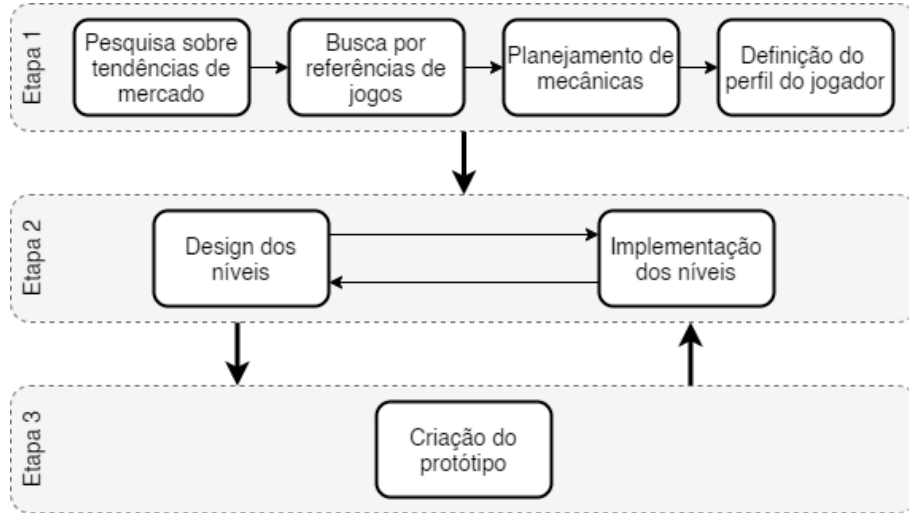
O time, composto por um *game designer* e um programador, que também é o autor desta pesquisa, desenvolveu os protótipos utilizando a plataforma Unity. A Unity, além de ser uma *game engine* renomada e com grande suporte da comunidade de desenvolvedores de jogos, é capaz de produzir versões utilizando tecnologia *WebGL*. Dessa forma, é possível criar protótipos testáveis por um navegador Web, o que facilita o processo de testes. A física e os controladores foram implementadas por terceiros na biblioteca disponível ao público “Unity 2D Platformer Controller”¹⁰. Ao todo foram desenvolvidos três níveis, cujo tempo de produção foi de aproximadamente um mês para cada. O material de arte foi gerado dentro da própria Unity ou utilizado dos exemplos fornecidos pela biblioteca utilizada.

O *design* dos níveis foi produzido por um *game designer* externo à pesquisa em conjunto com o autor da pesquisa. O *design* foi focado em um grupo específico de jogadores, adaptando a taxonomia que Mark Rosewater utilizou no jogo *Magic: the Gathering*: os jogadores *spike* (ROSEWATER, 2006). Este arquétipo de jogador é definido pela competitividade e grande habilidade em jogos e foi escolhido justamente por combinar com a proposta de jogo *masocore*. O processo de desenvolvimento encontra-se na Figura 28.

Após as etapas iniciais de pesquisa e planejamento, o processo era iterativo: o *game designer* produzia os níveis no papel, que em seguida eram implementados na Unity. O programador notificava o *game designer* quando ocorria alguma discrepância entre o planejamento no

¹⁰ <https://github.com/cjddmut/Unity-2D-Platformer-Controller>

Figura 28 – Fluxo do processo de desenvolvimento dos protótipos

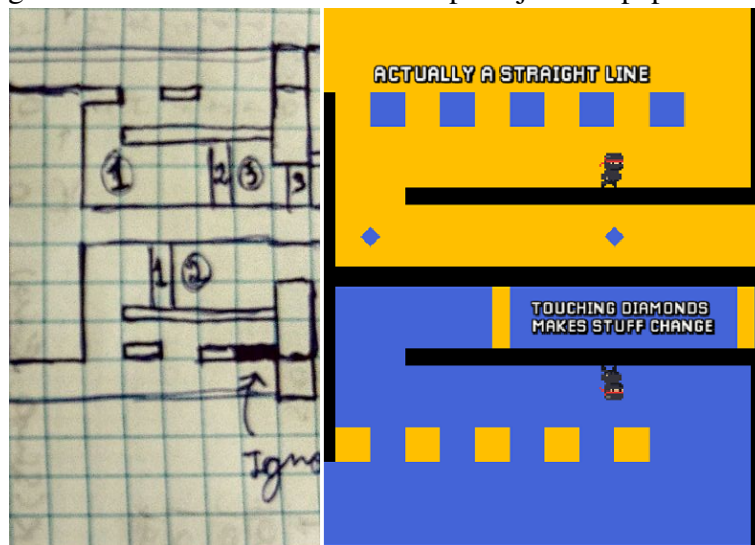


Fonte: Fornecida pelo autor.

papel e a prática no executável do jogo. O designer, em seguida, sugeria alterações com base no que foi reportado. O programador, por fim, implementava as alterações e procurava novamente por discrepâncias. O ciclo terminava quando nenhum ajuste precisava ser feito. O resultado era o protótipo de um nível. Ao todo foram feitos três níveis.

A Figura 29 mostra um trecho de um dos artefatos produzidos durante a etapa de planejamento do design do desafio 1C. A principal alteração quanto ao rascunho é o aumento da alternância de blocos na primeira plataforma. Essa maior segmentação tornou a mecânica das duas telas mais clara e chamativa, o que, combinada com o comentário do narrador, ajuda a causar *insights* nos jogadores.

Figura 29 – Design inicial do desafio 1C conforme planejado no papel.



Fonte: Fornecida pelo autor.

Por fim, criar um jogo próprio ao invés de modificar um pré-existente traz algumas vantagens. A primeira é a liberdade criativa, que facilita a produção de uma obra original e interessante. A segunda vantagem é a possibilidade de pôr em prática e documentar metodologias comuns de produção de jogos aliadas ao método científico. O produto final permite que modificações sejam feitas de forma rápida conforme as necessidades da pesquisa.

4.3 Dados Coletados

Nesta seção são descritos tipos de dados coletados e a abordagem da coleta.

4.3.1 Procedimento Padrão

O jogador foi convidado a participar do experimento por meio de listas de e-mail, postagens em redes sociais e convites pessoais. Após o participante ter chegado ao local do teste, a ele foi dada uma breve explicação do que se tratava a pesquisa, bem como quais métodos de coleta de dados seriam utilizados. O participante, então, preencheu um termo de consentimento (ver Apêndice C) atestando participar por vontade própria. Em seguida, houve uma breve explanação do jogo e dos controles. No caso do Teste 1.2, os participantes calibraram o WebGazer (PAPOUTSAKI *et al.*, 2016), necessário para a coleta de dados de *eye-tracking*, antes de cada nível. Após terem terminado de jogar, os participantes preencheram um formulário. Na Etapa 2 o formulário foi de avaliação da dificuldade dos níveis e, na Etapa 3, o formulário foi de avaliação da experiência. Nos Testes 1.1 e 1.2, os jogadores participaram de uma breve entrevista após terem preenchido o questionário. Este processo está resumido na Figura 30.

4.3.2 Observação

Nos testes da Fase 2, um pesquisador acompanhou os jogadores. Durante toda a sessão de jogo, o observador respondeu a dúvidas e anotou comportamentos que pudessem ajudar os *designers* a melhorar o jogo. Esses comportamentos iam desde as ações dentro do jogo, como repetir um pulo várias vezes no mesmo desafio, as ações do testador enquanto jogava, tais como inquietações ou mudanças posturais. Ao final da sessão, os jogadores podiam ver as anotações, caso desejassem.

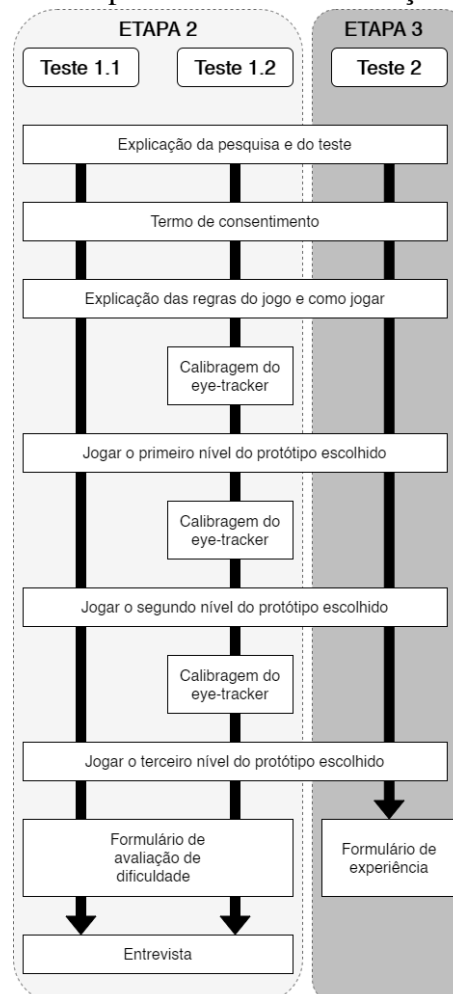
4.3.3 Entrevista

A entrevista após a sessão de jogo servia para reforçar as anotações. Além de perguntas padrão, por exemplo "o que você gostou do jogo?", "onde você sentiu maior dificuldade?" e "o que você melhoraria no jogo?", havia também perguntas estruturadas conforme o observado pelo pesquisador. Exemplos de perguntas incluíam "você se sentiu impaciente neste desafio específico?". Os testadores podiam, além disso, ser solicitados a explicar suas linhas de raciocínio para resolução de algum desafio. As novas informações adquiridas através da entrevista foram adicionadas às notas de observação.

4.3.4 Questionário de Avaliação de Dificuldade

O questionário da Fase 2 avaliou a dificuldade dos níveis. Essa avaliação foi necessária para estimar se a cobrança e a habilidade estavam adequadas para o surgimento do *flow*.

Figura 30 – Diagrama que mostra o procedimento de execução dos testes



Fonte: Elaborada pelo autor.

Nele eram apresentados todos os 18 desafios da versão inicial (v1). Os usuários marcavam numa escala de 1 a 5, para cada desafio, o grau de dificuldade que eles consideravam que aquele desafio possuía. O valor 1 se referia ao grau "muito fácil" e o 5 ao "muito difícil". Os resultados foram posteriormente comparados com a expectativa dos *designers*. O questionário está presente no Apêndice A.

4.3.5 Métricas de Jogo

As métricas de jogo e o *eye-tracking* foram exclusivos ao Teste 1.2. No caso de Downside Up, foram capturadas as seguintes métricas, que costumam ser pertinentes a todos os jogos de plataforma: posição do avatar e localização das mortes dentro do jogo, acompanhadas pela marcação de tempo exata do registro. Essas métricas ajudam os *designers* a entender a dificuldade do jogo e se os jogadores estavam demorando a passar de um desafio por conta de problemas de usabilidade ou de *level design*.

Além disso, quando o usuário morre, o local da morte é registrado junto com uma marcação de tempo. Ao contrário da posição, que era registrada por usuário, as mortes foram registradas por nível. Sendo assim, foram usados apenas dois arquivos, que representam os níveis 2 e 3, já que no primeiro nível não há espinhos.

4.3.6 Eye-Tracking

A solução escolhida para esta pesquisa foi o WebGazer¹¹, uma vez que este pode ser executado em paralelo ao jogo por meio de um navegador web.

Experimentos preliminares mostraram que, mesmo em condições de bom posicionamento e iluminação adequada, a precisão do rastreamento de olhar não era confiável utilizando a solução escolhida. O rastreamento ocular era inconstante e podia variar bastante se o usuário movimentasse o pescoço.

Para este projeto, no entanto, foi necessário rastrear apenas o eixo Y da tela, contornando o problema da imprecisão. Assim, o WebGazer foi alterado para registrar, a cada *frame*, se o testador estava olhando para cima ou para baixo. Este registro obedeceu a seguinte estrutura: o valor 1 indica que o usuário está olhando para cima, e o valor 2 indica para baixo. Esse dado é acompanhado de uma marcação de tempo.

¹¹ Disponível em: <https://webgazer.cs.brown.edu>

Soluções mais atuais e sofisticadas, por exemplo uso de *eye-trackers* dedicados, podem ser usadas para produzir resultados tão bons quanto ou superiores aos obtidos aqui.

4.3.7 Estrutura de Dados

Foram utilizados arquivos de texto simples gerados a partir da união entre os registros das métricas de jogo e do eye-tracking. Essa união foi feita combinando os dados usando como chave a marcação de tempo.

Após a união, foi feita uma limpeza dos dados. Esta limpeza se deu pelo fato de Downside Up e o WebGazer executarem a 60 quadros por segundo, frequentemente ocasionando em registros repetidos. Essa mesma lógica foi seguida com os dados dos mortes de usuários.

Ao final, foram utilizados dois tipos de estruturas. A primeira registra as posições do avatar e a a tela observada pelo usuário a cada instante. A segunda registra as posições das mortes e a tela observada pelo usuário naquele momento. Um exemplo de registro da estrutura de dados encontra-se na Figura 31. Os arquivos que registraram as mortes foram construídos de forma similar.

Figura 31 – Exemplo de registro da telemetria. O primeiro campo é uma marcação de tempo, o segundo as posições X e Y do avatar, o terceiro a tela que estava sendo observada pelo jogador naquele momento (1 é a tela de cima, 2 é a tela de baixo)

| | | | |
|----------|------------|----------|---|
| 17:34:20 | -0.5565398 | 3.879763 | 2 |
| 17:34:20 | -0.5039569 | 3.982254 | 2 |
| 17:34:20 | -0.4613739 | 4.076745 | 2 |
| 17:34:21 | -0.4287909 | 4.163236 | 2 |
| 17:34:21 | -0.406208 | 4.241727 | 2 |
| 17:34:21 | -0.393625 | 4.312219 | 2 |
| 17:34:21 | -0.3910421 | 4.37471 | 2 |
| 17:34:21 | -0.3910421 | 4.475692 | 2 |
| 17:34:21 | -0.3910421 | 4.514183 | 2 |
| 17:34:21 | -0.3910421 | 4.544674 | 2 |

Fonte: Elaborada pelo autor.

Todas as visualizações agregadas e os arquivos de texto com os dados utilizados para gerá-las estão disponíveis publicamente no repositório deste trabalho ¹².

4.3.8 Questionários de Avaliação de Experiência

Na Fase 3 foram aplicados somente dois questionários de avaliação de experiência, o PANAS (WATSON *et al.*, 1988) e o GEQ (IJSELSTEIJN *et al.*, 2013), disponíveis nos Anexos

¹² https://arthursb.github.io/Downside-Up/index_data.html

A e B. Nesses questionários, o usuário marcava numa escala de 1 a 5, o grau de concordância com afirmações relacionadas a afetos positivos e negativos. A escolha desses questionários se deve ao renome e rigor científico: o PANAS é recomendado para aplicações multimídia em geral, já o GEQ é focado em jogos.

Neste trabalho, o GEQ (módulos principal e pós-game) e o PANAS foram traduzidos para o português (ver Apêndice B). As implicações dessa tradução são discutidas no Capítulo 6. Conforme exposto no Capítulo 2, para esta pesquisa, considera-se que uma experiência positiva, e consequentemente divertida, ocorrerá com base na análise da incidência e intensidade dos afetos positivos e negativos. A hipótese é de que a experiência positiva é composta de afetos positivos mais frequentes e de maior intensidade (respostas de valor 4 ou 5).

4.4 Visualizações

Os dados captados advêm de fontes distintas, sendo assim, são necessárias técnicas de visualização para melhor compreendê-los. Para os dados das entrevistas, notas de observação e questionários, técnicas simples como gráficos de barras, gráficos de pizza, etc. são suficientes. Os dados relacionados aos softwares, no entanto, requerem abordagens mais elaboradas. As subseções a seguir detalham as soluções de visualização empregadas neste trabalho.

4.4.1 Replays

Combinando o que foi obtido via *eye-tracking* com as métricas de jogo, é possível formar *replays*. Os *replays* consistem em passos automatizados a serem replicados por um avatar customizado. É possível sincronizar a posição do avatar com a região para onde o jogador estava olhando em dado momento. Destaca-se apenas a região visada a cada instante. Por exemplo, se o jogador está olhando para a tela de cima, apenas ela estará colorida, enquanto que a tela de baixo ficará preta. A Figura 32 ilustra cinco *frames* de um *replay*¹³.

Uma grande vantagem desta visualização é que é possível observar a confusão visual. Em sessões do jogo como a ilustrada na Figura 32, ocorre com frequência a alternância de olhar entre as porções superior e inferior da tela, algo que não é desejável segundo os direcionamentos de *design* descritos na Seção 4.2. A desvantagem, no entanto, é que os *replays* podem ter longa duração, o que demanda bastante tempo para análise.

¹³ Um exemplo desta visualização em forma de vídeo pode ser encontrada em: <https://arthursb.github.io/Downside-Up/>

Figura 32 – Cinco *frames* de uma visualização de *replay*. Em um *replay* é possível observar com clareza para onde os jogadores estavam olhando em dado instante.



Fonte: Elaborada pelo autor.

4.4.2 Visualização Agregada

Pode ser necessária apenas uma compreensão do comportamento geral, algo que os *replays* não mostram satisfatoriamente. Foi desenvolvida uma ferramenta que produz uma visualização que combina todos os dados adquiridos pelo jogo em uma imagem só. Essa técnica de visualização agregada encontra-se na Figura 33.

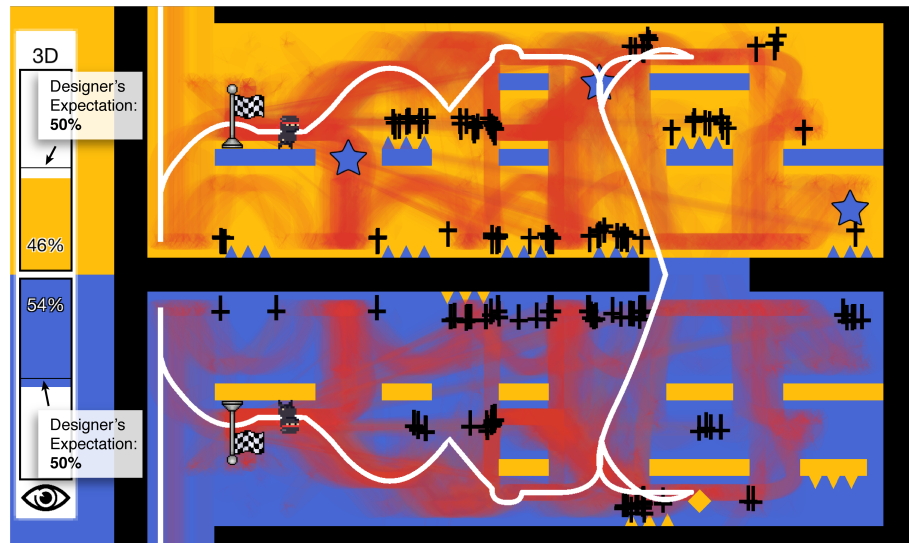
A técnica desenvolvida apresenta uma trajetória adaptando a técnica de rotas quentes (GOMES *et al.*, 2018). A rota branca representa a trajetória ideal, ou mais rápida, para a solução do desafio. Por sua vez, as rotas vermelhas representam trajetórias reais, realizadas pelos participantes dos testes. Cada rota vermelha possui um grau de transparência baixo. Quando uma trajetória é realizada repetidamente, as rotas se sobrepõem, o que faz a cor vermelha se tornar cada vez mais vibrante, definindo um padrão na movimentação dos jogadores.

As rotas vermelhas foram construídas usando o componente *LineRenderer* da Unity passando como entrada os pontos obtidos das estruturas apresentadas na Subseção 4.3.7. Não foram utilizadas técnicas de simplificação desses pontos, como clusterização. As rotas brancas foram geradas a partir de um registro gerado pelo próprio desenvolvedor.

Para representar em que lugares os jogadores morreram, são utilizadas marcações em forma de cruz. Combinando com os dados do *eye-tracking*, é possível determinar para qual porção da tela o participante estava olhando quando o evento ocorreu. Sendo assim, uma cruz posicionada na tela de cima indica que o jogador estava olhando para cima naquele instante. Este dado é especialmente interessante para averiguar locais onde ocorrem mortes acidentais com frequência.

Por fim, é possível também mostrar a distribuição do olhar em um desafio. Ela é representada por meio de duas barras na lateral esquerda. O resultado é comparado com a expectativa do *designer* em relação à predominância de uma tela em relação a outra, representada por uma linha preta. Com essa visualização, é possível ver de forma mais clara se uma das duas porções da tela está visualmente mais chamativa do que o planejado.

Figura 33 – Visualização agregada do desafio 3D. Foram construídos 154 caminhos neste cenário.



Fonte: Elaborada pelo autor.

4.4.3 Avaliação da Ferramenta

A visualização agregada foi avaliada por profissionais da área. Foi feita uma entrevista na qual os participantes responderam a perguntas sobre suas experiências prévias e realizaram tarefas utilizando a ferramenta de visualização.

A avaliação buscou mensurar, por meio da entrevista, a eficácia segundo os parâmetros definidos por Zhu (2007): acurácia, utilidade e eficiência. A acurácia pode ser medida por meio da quantidade de erros de interpretação, a utilidade pelo número de objetivos de pesquisa alcançados, e a eficiência pela curva de aprendizado. O roteiro da entrevista está listado a seguir:

- **Fase 1:** entrevistador se apresenta e fala sobre a pesquisa;
- **Fase 2:** participante se apresenta e fala sobre sua experiência com desenvolvimento de jogos;
- **Fase 3:** participante dá uma nota de 1 a 5, sendo 1 menor e 5 maior, sobre seu nível de expertise com desenvolvimento em Unity e visualização de dados, em seguida justifica as respostas;
- **Fase 4:** entrevistador mostra o jogo e explica as mecânicas, o participante joga por alguns minutos;
- **Fase 5:** entrevistador explica a ferramenta de visualização, o participante pode explorá-la livremente se quiser;
- **Fase 6:** entrevistador exhibe 6 situações as quais o participante deve responder

como ele utilizaria a ferramenta para identificar algo (mais detalhes no Capítulo 5);

- **Fase 7:** participante responde às seguintes perguntas sobre o *level design* na empresa onde trabalha: “você sabe como é feito o *level design*?”, “você já usou uma ferramenta customizada para *level design*?”, “você imagina que uma ferramenta customizada poderia ter lhe ajudado em algum projeto?”, “você acredita que a ferramenta apresentada é útil para o *level design*?” ;
- **Fase 8:** participante comenta livremente, dando críticas e sugestões;
- **Fase 9:** entrevistador encerra com agradecimentos;

A visualização foi empregada pelos próprios desenvolvedores e pesquisadores. Espera-se que o resultado da análise das entrevistas aponte problemas e limitações da técnica empregada graças aos olhares externos de participantes de diversos níveis de experiência. O resultado é apresentado no Capítulo 5 e discutido no Capítulo 6.

4.5 Considerações Finais

Neste capítulo foram apresentadas a metodologia utilizada neste trabalho, as técnicas de coleta e as de visualização de dados. Posteriormente, no Capítulo 5, há o relato da aplicação dos testes e a sua repercussão, bem como o resultado da entrevista com profissionais. A discussão sobre os resultados e limitações estão presentes no Capítulo 6.

5 RESULTADOS

Neste capítulo, serão apresentados os resultados da entrevista com desenvolvedores de jogos e dos testes com usuários, conforme descrito no Capítulo 4. A Seção 5.1 mostra o protocolo da entrevista com desenvolvedores de jogos e a análise das respostas. A Seção 5.2 mostra os resultados dos testes com usuários, especificando detalhes sobre a aplicação das técnicas descritas no Capítulo 4. São feitas considerações finais na Seção 5.3.

5.1 Entrevista com Desenvolvedores

Figura 34 – Resumo do perfil dos entrevistados.



Fonte: Elaborada pelo autor.

Ao todo, seis participantes com diferentes níveis de experiência realizaram a entrevista. A Figura 34 resume o perfil dos participantes. Entre os 6, só havia um estudante. Apenas um não trabalhava com jogos na sua profissão atual. Houve somente uma participante do sexo feminino. Os entrevistados com ensino superior completo eram formados em computação e/ou jogos digitais. Os participantes A, B, C e F trabalharam com jogos de forma independente além de seus empregos formais.

Foi pedido aos entrevistados para definirem, numa escala de 1 a 5, sendo 1 “nenhuma experiência” e 5 “domínio completo”, o quanto eles se consideram experientes com Unity e Visualização de Dados. A maioria dos entrevistados assumia o cargo de programador e se considerava experiente com Unity (respostas com valor 4 ou 5), normalmente por conta do tempo

que trabalharam com essa *engine*. Apenas o participante E marcou 1 neste item, visto que não era programador nem nunca precisou usar a *engine*. Os participantes consideraram ter conhecimento básico ou moderado sobre visualização de dados (respostas com valor 2 ou 3). A justificativa mais comum para as respostas foi a de ter aprendido a teoria em aulas na graduação mas não necessariamente aplicar o conteúdo na prática com frequência.

Os participantes jogaram Downside Up por 5 minutos para que entendessem melhor as regras e a filosofia de design dos níveis. Após isto, foram introduzidos ao projeto no editor da Unity e à ferramenta de visualização. Na sequência, foram apresentadas situações no jogo (Figura 35) sobre as quais o participante utilizou a ferramenta para responder perguntas. Os enunciados e respostas serão detalhados a seguir.

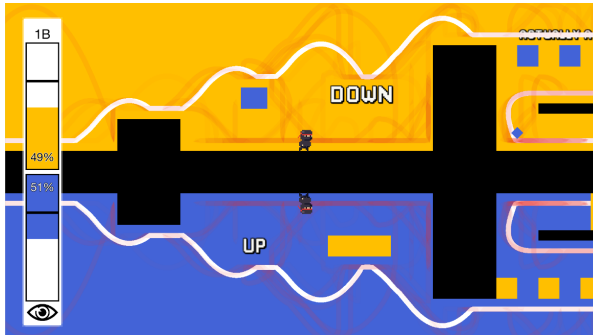
A situação 1 (Figura 35a) mostra um desafio introdutório. O jogador é devidamente apresentado à mecânica de duas telas pela primeira vez. O caminho ideal é representado pela linha branca, e mostra o simples pulo de uma plataforma a outra. Como nem todos os jogadores entendem rapidamente a mecânica, muitos caem e repetem o caminho, conforme evidenciado pelas linhas vermelhas. No teste, todos os participantes apontaram o mesmo caminho, porém três deles não mencionaram explicitamente a trajetória branca.

A situação 2 (Figura 35b) explora a interpretação do participante quanto às linhas vermelhas. No exemplo anterior, elas estavam com cores claras, pois as trajetórias não foram feitas repetidamente. Já neste exemplo, algumas regiões apresentam trajetórias com vermelho intenso. Muitos jogadores chegaram a esse desafio sem estarem acostumados a coordenar o olhar nas duas telas. O resultado é que muitos não enxergaram o losango na tela de baixo. Este losango, quando acionado, desativa uma plataforma na tela de cima, fazendo o jogador cair e repetir o trajeto até a parte superior do cenário. Três dos entrevistados apontaram para o chão e três para as plataformas, mas apenas dois mencionaram explicitamente as trajetórias vermelhas.

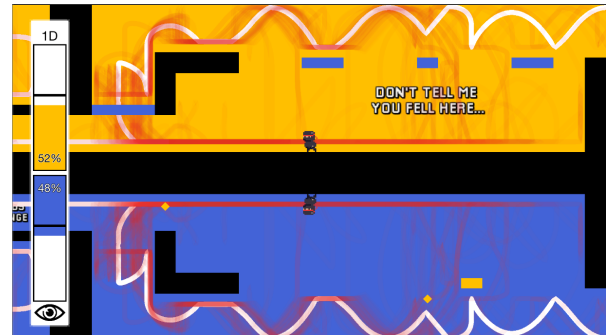
A situação 3 (Figura 35c) mostra um desafio onde há espinhos e um losango que ativa uma armadilha. Quando ativado, o losango revela que a barra larga laranja é composta na verdade por duas barras mais finas. Uma delas desaparece, fazendo com que a outra, repleta de espinhos, caia. Os jogadores são pegos de surpresa pela armadilha ou pelos espinhos ao redor. No teste, todos os participantes observaram as cruzes e mencionaram os espinhos. Além disso, um entrevistado mencionou que havia cruzes marcadas no ar e supôs que havia elementos dinâmicos no cenário sem precisar jogar a fase.

A situação 4 (Figura 35d) é a última que testa diretamente os recursos da visualização

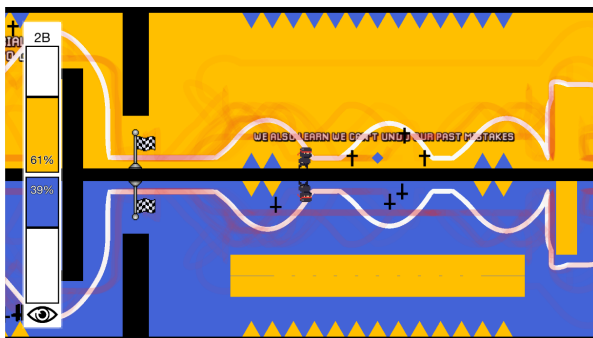
Figura 35 – Situações mostradas aos entrevistados durante a avaliação da ferramenta de visualização agregada



(a) Situação 1: identificar o caminho ideal no desafio 1B



(b) Situação 2: identificar um caminho comum realizado pelos jogadores no desafio 1D



(c) Situação 3: identificar o local de morte comum no desafio 2B



(d) Situação 4: identificar qual dos lados foi o mais visto pelos jogadores no desafio 3A



(e) Situação 5: identificar locais onde houve mortes acidentais no desafio 2C



(f) Situação 6 (desafio 1E): analisar onde os jogadores sentiram mais dificuldade

proposta. Nela foi apresentado um desafio no qual havia apenas dois focos de atenção: um texto na tela de baixo e um losango na tela de cima. A ação ocorreu totalmente na tela de cima, o que justifica a porcentagem mais alta na barra lateral. Todos os entrevistados citaram que foi o lado de cima. Apenas um não fez menção explícita à barra enquanto explicava o porquê.

As duas últimas situações exploram possibilidades de interpretação prática dos resultados. Na situação 5 (Figura 35e), é apresentado o desafio 1C. Foi pedido que o participante identificasse locais onde ocorreram possíveis mortes acidentais. Todos apontaram a região onde havia espinhos na tela de baixo invisíveis na de cima. Dois citaram cruzes em locais onde não

havia espinhos, causados por elementos dinâmicos na fase. Dois participantes focaram mais na explicação e não detalharam quais recursos da visualização foram usados para tirar sua conclusão. Um dos participantes disse que não houve indícios de mortes acidentais.

Na situação 6 foi pedida a identificação de um local onde os usuários sentiram dificuldade. Este local é uma discreta parede azul presente apenas na tela de cima (ver Figura 35f). No entorno dela há uma trajetória sólida vermelha indicando que houve quedas frequentes precedidas pela colisão com a parede. Todos os participantes citaram a coluna azul, justificando com a intensidade da trajetória vermelha perto dela. Apenas um justificou o comportamento com a barra de distribuição de olhar, indicando que os jogadores não viram a coluna. Dois adicionalmente mencionaram a região onde há um losango, que corresponde a uma armadilha que provoca uma queda.

As quatro primeiras situações testaram a compreensão de cada recurso da ferramenta, enquanto que as duas últimas simularam casos de uso reais. Um resumo das respostas encontra-se na Figura 36.

Figura 36 – Resumo das respostas dos participantes ao teste com situações de visualização. Células azuis indicam que o participante respondeu corretamente com apoio de algum recurso da ferramenta. Células amarelas indicam respostas corretas sem apoio da ferramenta. Células vermelhas indicam que o participante não respondeu corretamente

| Situação | | Resultados | | | | | |
|----------|-----------------------------------------------|------------|---------|----------|---------|---------|---------|
| Número | Objetivo | A | B | C | D | E | F |
| 1 | Avaliar compreensão da trajetória branca | Amarelo | Azul | Amarelo | Azul | Azul | Amarelo |
| 2 | Avaliar compreensão das trajetórias vermelhas | Amarelo | Amarelo | Azul | Amarelo | Amarelo | Azul |
| 3 | Avaliar compreensão das marcações das mortes | Azul | Azul | Azul | Azul | Azul | Azul |
| 4 | Avaliar compreensão da barra lateral | Azul | Amarelo | Azul | Azul | Azul | Azul |
| 5 | Avaliar exploração dos recursos | Azul | Amarelo | Vermelho | Azul | Amarelo | Azul |
| 6 | Avaliar exploração dos recursos | Azul | Azul | Azul | Azul | Azul | Azul |

Legenda:

- Respondeu corretamente com apoio da ferramenta
- Respondeu corretamente sem apoio da ferramenta
- Não respondeu corretamente

Fonte: Elaborada pelo autor.

Em seguida, os participantes responderam a perguntas sobre o uso da ferramenta. Somente o participante C não sabia como era feito o *level design* na empresa onde trabalhava. Apenas os participantes A e E nunca usaram uma ferramenta customizada para *level design*, mas todos acreditaram que esse tipo de ferramenta ajudaria em algum projeto passado ou atual.

O participante C citou que um projeto passado falhou por conta de não ter tido estruturas para incluir o *feedback* dos jogadores com facilidade.

Ao final, os participantes comentaram sobre a ferramenta. A recepção foi bastante positiva. A ferramenta foi elogiada por sua simplicidade a despeito da necessidade de um conhecimento mínimo em Unity para ser usada. Todos os participantes citaram que a ferramenta apresentada era útil para *level design* e que atendia com sucesso as necessidades específicas do *Downside Up*. O participante B comentou que não seria muito custoso adaptar os recursos implementados para utilização em outros jogos de ação. D comentou que a ferramenta era capaz de fornecer bons dados para a análise da experiência do jogo, principalmente no que se referia à localização de situações frustrantes nos desafios. Por fim, os participantes sugeriram diversas alterações ou recursos novos que poderiam ser feitos futuramente pra melhorar a ferramenta. Estas sugestões estão listadas na Tabela 5.

Tabela 5 – Resumo das sugestões de melhoria que os entrevistados apontaram para a ferramenta de visualização.

| Participante | Sugestão | Motivo |
|--------------|----------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| A | Mostrar direção e sentido nas trajetórias vermelhas | Compreender melhor se o jogador estava explorando o cenário ou resolvendo o desafio |
| A e C | Inclusão de técnicas de inteligência artificial | Acelerar o processo de testes e avaliação de modificações |
| B | Linha do tempo | Planificar os replays em uma figura |
| B e F | Registrar os inputs do jogador | Tornar os dados dos replays mais precisos |
| B e E | Remover a dependência da interface da Unity e adicionar atalhos de teclado | Melhorar usabilidade e tornar a ferramenta mais customizável e amigável para leigos |
| B e C | Reduzir a poluição visual (clusterização) | Melhorar o design da visualização |
| C e D | Adição de comentários | Justificar comportamentos não triviais, como os que ocorrem em desafios com elementos dinâmicos |
| E e F | Melhorar a paleta de cores do jogo | Tornar a visualização mais acessível e evitar que as cores do jogo dificultem a interpretação dos dados |
| F | Adição de rastreadores em outros objetos que não o jogador | Permitir uma melhor compreensão de eventos do jogo |

Fonte: Elaborada pelo autor com base na entrevista com desenvolvedores.

Conforme explicado no Capítulo 4, foram avaliados três elementos na entrevista com os desenvolvedores: acurácia, utilidade e eficiência da ferramenta de visualização.

A visualização se mostrou acurada, uma vez que os recursos da mesma foram interpretados corretamente nas seis situações apresentadas, com exceção da resposta do participante C à situação 5. A ferramenta, no entanto, apresenta alguns problemas quanto à clareza. As trajetórias não mostram direção e sentido e o significado das linhas pretas na barra de distribuição do olhar requerem explicação externa. Em alguns momentos, as cores do jogo se misturam com

a da visualização. Em outros, há um excesso de elementos na tela.

Embora não tenham sido coletados dados quantitativos sobre o uso da ferramenta, como o número de vezes que algum recurso é usado pelo participante, os testadores foram unânimes quanto à utilidade da ferramenta, citando que ela certamente colaboraria com a melhoria na qualidade do jogo testado. Houve ênfase para os recursos de rotas quentes e *eye-tracking*, citados como os principais provedores de informações relevantes acerca do comportamento dos usuários.

A ferramenta pode ser considerada parcialmente eficiente, uma vez que ela possibilita acesso simples a informações complexas de se interpretar sem elementos visuais, mas exige pré-requisitos. O usuário deve ter experiência com a interface da Unity e uma compreensão prévia do *level design* do jogo, principalmente por conta da quantidade de elementos dinâmicos. Isso foi evidenciado principalmente pelo fato de que as quatro primeiras tarefas, cuja intenção era avaliar a compreensão do participante acerca dos recursos isoladamente, foram realizadas rapidamente, ao contrário das duas últimas. O participante B, por exemplo, não entendeu o conceito de “morte acidental”, que foi definida como a ocorrência das seguintes situações: jogador morreu por conta de algo que ele não enxergou; jogador morreu e não entendeu o motivo; jogador morreu e sentiu que não por morreu como consequência de decisões erradas no jogo.

De forma resumida, a ferramenta é acurada e útil, mas nem sempre eficiente. Como esperado, a maioria das sugestões dadas pelos participantes tem como objetivo melhorar a eficiência.

5.2 Testes com Usuários

Aqui são descritos os resultados dos testes aplicados conforme a metodologia descrita na Seção 4.1 do Capítulo 4. Para os testes da Etapa 2, relativa à identificação de melhorias e balanceamento no jogo, os participantes jogaram *Downside Up* em um Macbook Pro Retina de 2014¹, situados em uma sala com apenas mesa, cadeiras e outros materiais necessários para a condução dos testes, como caneta e papel. Já na Etapa 3, os testes foram realizados em um laboratório (Figura 37) equipado com computadores iMac de 2015². Em ambos os casos, participantes foram recrutados por meio de listas de email, redes sociais ou pessoalmente. Ao todo, o processo de testes durou dois meses.

¹ Especificações técnicas: <https://support.apple.com/kb/SP704>

² Especificações técnicas: <https://support.apple.com/kb/SP733>

Figura 37 – Laboratório onde foram realizados os testes da Etapa 3 *Downside Up*.



Fonte: Fornecida pelo autor.

5.2.1 Etapa 2 - Desenvolvimento da Versão v2 (Teste 1.1)

Os testes foram realizados com 12 participantes. Destes, a maioria (8 jogadores) possuía idade entre 21 e 26 anos, tinham hábito de jogar por 8 ou mais horas por semana (7 jogadores) e preferiam jogar em um computador (9 jogadores).

Observou-se que os jogadores demoraram mais que o esperado para aprender as mecânicas básicas do jogo, alcançando a compreensão completa somente na segunda fase. Os participantes relataram passar mais tempo olhando para a parte de cima do jogo e percebiam a parte de baixo como uma “sombra”. Esse comportamento sugere que o tutorial não está desafiador o suficiente ou não apresenta as mecânicas de forma apropriada.

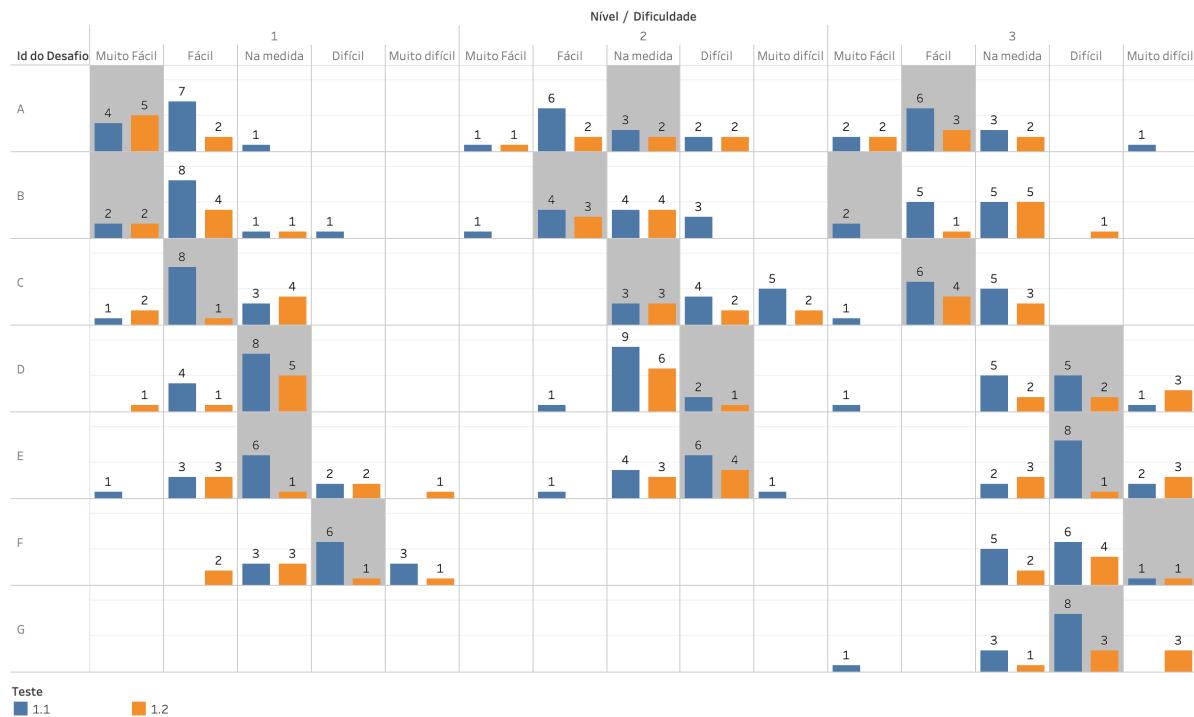
Foram observados também problemas de usabilidade. Jogadores sentiram que os controles estavam pouco responsivos em algumas situações. Comentaram também que havia muitos espinhos posicionados de maneira injusta e que serviam a nenhum propósito além de tornar o jogo frustrante. Um usuário precisou de ajuda para completar certos desafios de destreza. No desafio 2D, que consiste em uma sala repleta de espinhos e objetos que desencadeiam a aparição de mais espinhos, uma nova mecânica foi introduzida sem uma explicação propriamente dita, fazendo com que os participantes a compreendessem exclusivamente por meio de tentativa e erro.

Os participantes morreram mais frequentemente na última fase do jogo. Embora não

tenham explicitamente dito isso nas entrevistas, foi observado que os jogadores se frustraram constantemente com mortes acidentais. Dessa forma, conclui-se que um reposicionamento dos espinhos e ajuste da quantidade se faz necessário.

Ao final da sessão de jogo, os participantes puderam expor sua opinião quanto ao jogo. Elogiaram as mecânicas gerais, comentando de forma positiva sobre a filosofia de criar um *design* que prega peças. Muitos participantes jogaram sorrindo, ou esboçaram reações de alívio ao superar um desafio. Um participante comentou ter gostado do fato de haver um narrador que dava dicas sutis.

Figura 38 – Resultado do questionário de avaliação de dificuldade das três fases, realizado nos testes da Etapa 2. Células acinzentadas marcam a opção que os designers esperavam ser a mais marcada em cada desafio

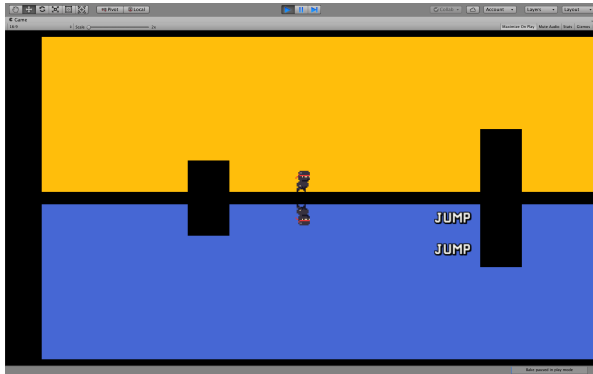


Fonte: Elaborada pelo autor.

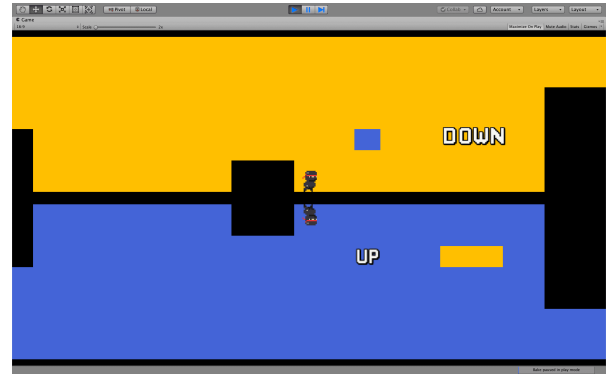
As respostas do questionário de avaliação de dificuldade encontram-se resumidas na Figura 38. Nas três fases, as respostas corresponderam à expectativa dos *designers*, com exceção dos desafios 1A, 1B, 2A, 2C, 2D, 3B e 3F, mostrados na Figura 39. Alguns deles possuem problemas de design, que serão discutidos adiante.

Os desafios 1A e 1B supostamente deveriam ensinar todas as mecânicas de pulo, mas não ensinaram apropriadamente como pular no ar. É possível que isto tenha causado mais marcações na opção *Fácil* que na *Muito Fácil*.

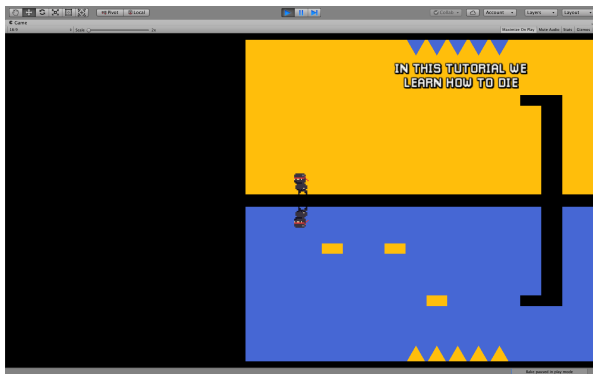
Figura 39 – Desafios nos quais a percepção de dificuldade dos usuários difere da expectativa dos designers, conforme indicado no formulário de avaliação de dificuldade



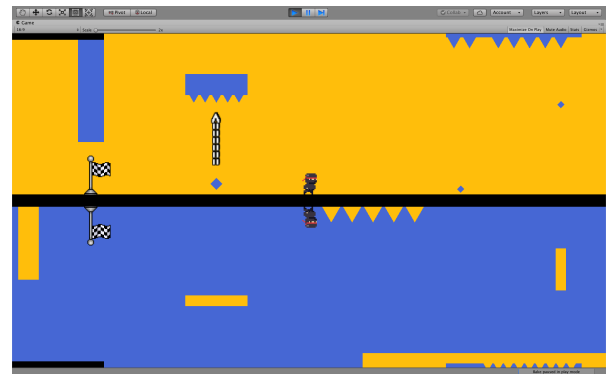
(a) Desafio 1A



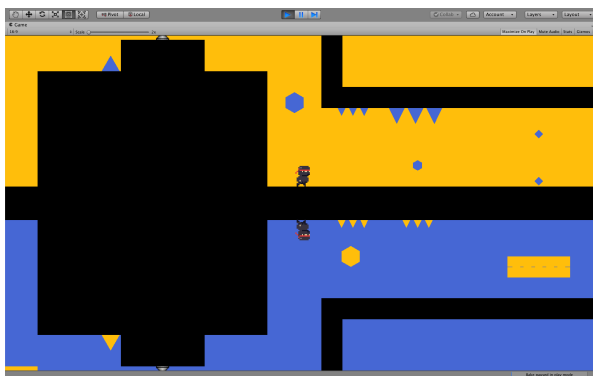
(b) Desafio 1B



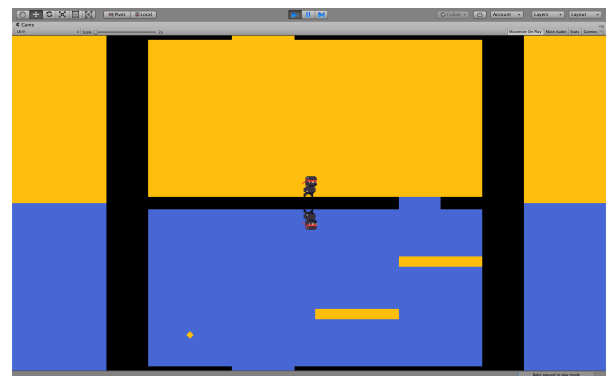
(c) Desafio 2A



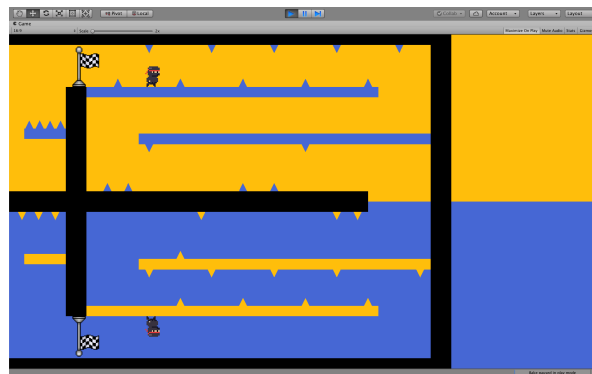
(d) Desafio 2C



(e) Desafio 2D



(f) Desafio 3B



(g) Desafio 3F

O desafio 2C foi categorizado como *Muito Difícil* provavelmente devido a um problema de câmera. Para resolver este desafio, os jogadores deviam chegar perto de uma plataforma para desencadear uma mudança de câmera que revelaria uma passagem secreta no teto. Essa era a primeira situação na qual a câmera se movimentava verticalmente a despeito de não ter havido nenhuma indicação prévia de que isso era possível no jogo, tornando essa mudança de câmera algo inesperado.

No desafio 2D foi apresentada uma nova mecânica sem explicações ou locais seguros para experimentação. Os jogadores, portanto, precisaram entender a mecânica por meio de tentativa e erro. Surpreendentemente, este desafio foi categorizado como *Na Medida*.

Em 3B, todos os objetos relevantes estavam presentes na tela de baixo, e como este era o único desafio a se passar inteiramente nessa tela, explicaria o fato desse desafio ter sido categorizado como *Muito Fácil*, conforme esperado.

O último desafio notável é o 3F. Nele o jogador devia manobrar o avatar com precisão e estratégia para atravessar um corredor repleto de espinhos. Provavelmente o senso de dificuldade dos jogadores quanto a este desafio não correspondeu aos dos *designers* por conta da sensação de maestria ocasionada pela repetição. Em outras palavras, quanto mais os jogadores morriam e repetiam o caminho, mais fácil eles percebiam o desafio.

5.2.2 *Etapa 2 - Desenvolvimento da Versão v3 (Teste 1.2)*

Estiveram presentes 7 participantes. A maioria com idade entre 18 e 20 anos (3 jogadores), com hábito de jogar de 2 a 8 horas por semana (6 jogadores) no computador (3 jogadores).

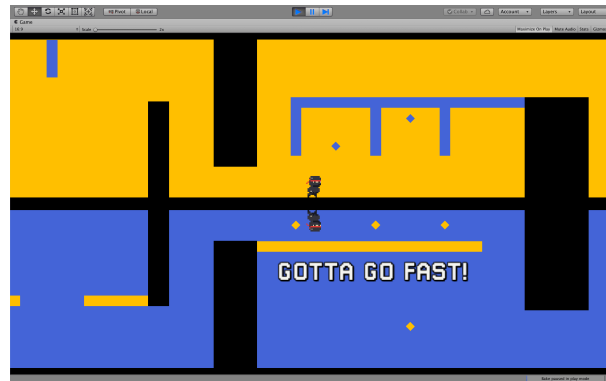
Os resultados das observações e entrevistas foi similar ao relatado na Subseção 5.2.1. Uma porcentagem mais alta de jogadores teve dificuldade em entender as mecânicas fundamentais. Os mesmos problemas de usabilidade identificados no Teste 1.1 foram relatados no Teste 1.2. Os testadores elogiaram a criatividade do jogo.

O questionário também apontou resultados similares, conforme pode ser visto na Figura 38. Sendo assim, ambos os grupos perceberam o jogo com um nível de dificuldade similar. Notavelmente, neste teste, nenhum jogador marcou a resposta esperada pelos *designers* nos desafios 2B e 3B, que deveriam ser os mais fáceis em suas respectivas fases.

Ao analisar os *replays*, foram detectadas ocorrências de cintilação, sugerindo movimentação constante do olhar em certos desafios. Conseqüentemente, houve confusão a respeito

de como manejar informação nas duas telas ou os jogadores estavam explorando o cenário antes de realizar alguma ação. Em casos como o do desafio 1F (Figura 40), isso é esperado, uma vez que é o último desafio da fase e foi projetado para utilizar bem o aspecto de atenção dividida. Em outros casos, como o do desafio 2C, não há muita necessidade de olhar para a tela de baixo, e essas cintilações não são esperadas, o que reforça o que foi observado e anotado das entrevistas quanto a problemas de usabilidade. As cintilações acontecem justamente quando os jogadores tentam realizar o pulo em direção à plataforma que leva à passagem secreta. A Figura 32 do Capítulo 4 mostra essa situação.

Figura 40 – Captura de tela do desafio 1F de *Downside Up*, que requer grande atenção tanto na parte de cima quanto de baixo



Fonte: Elaborada pelo autor.

No desafio 3D, também ilustrado no Capítulo 4 (Figura 33), os jogadores devem explorar um ambiente cheio de objetos que desencadeiam mudanças nos espinhos ou nas plataformas. Um desses objetos abre o caminho para o personagem cair e ir para o próximo desafio. Por meio da visualização agregada, foi possível encontrar diversos pontos de morte acidental no entorno de espinhos cuja posição muda de acordo com algum evento no jogo. Além disso, era esperado que ambas as telas apresentassem a mesma distribuição do olhar, uma vez que os elementos estão bem distribuídos entre as duas porções, mas os jogadores passaram mais tempo olhando para a parte de baixo. Isso indica que ou a parte de baixo era menos distrativa ou foi percebida como mais segura para se orientar. Uma solução para equilibrar essa distribuição seria a de simplificar a resolução do desafio e remover espinhos.

Em resumo, as conclusões obtidas com o auxílio das estratégias de visualização não apenas reforçaram aquilo que foi anotado da observação e entrevistas como também proveram maior *insight* a respeito da natureza multitarefa do jogo. Foi possível determinar com clareza espinhos que causaram frustração desnecessariamente, plataformas indesejadas e desafios com

um desbalanço nos elementos visuais entre as duas telas. Foi possível também observar diferenças entre manobrar o personagem na tela de cima e na de baixo.

5.2.3 Etapa 3 - Comparação entre Protótipos (Teste 2)

Oitenta e quatro participantes realizaram os testes nesta etapa. A maioria (47) possuía idade entre 18 e 20 anos, jogava por mais de 8 horas na semana (32 jogadores) e preferia jogar no computador (57 jogadores). Dentre os testadores, 39 jogaram v1, 22 jogaram v2 e os outros 23, v3.

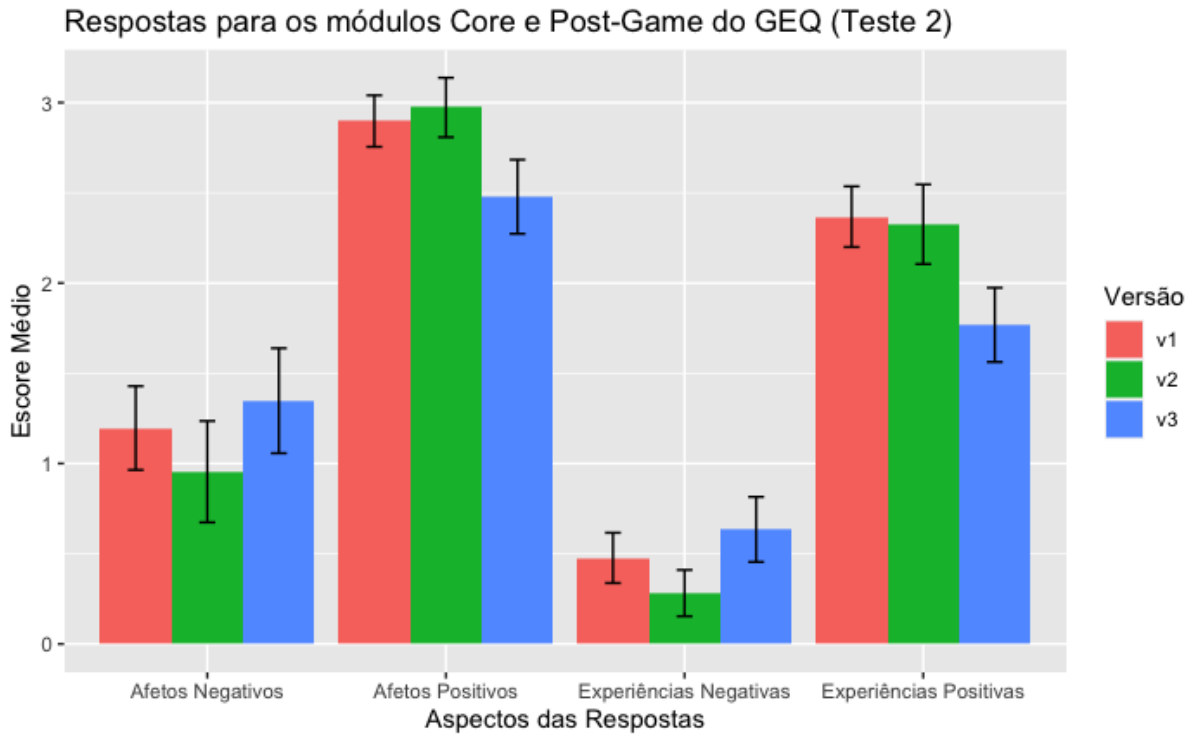
Os resultados do GEQ, mostrados na Figura 41, apontam que os jogadores experienciaram mais afeto positivo e menos afeto negativo jogando v2 ($M = 2.97$, 95% IC = 2.81, 3.13 e $M = 0.95$, 95% IC = 0.67, 1.23, respectivamente) que jogando v1 ($M = 2.89$, 95% IC = 2.75, 3.04 e $M = 1.19$, 95% IC = 0.96, 1.42, respectivamente) ou v3 ($M = 2.48$, 95% IC = 2.27, 2.68 e $M = 1.34$, 95% IC = 1.05, 1.63, respectivamente). Além disso, v2 manteve a experiência positiva equiparável à do protótipo original (v1: $M = 2.36$, 95% IC = 2.22, 2.53, v2: $M = 2.33$, 95% IC = 2.11, 2.55 e v3: $M = 1.77$, 95% IC = 1.56, 1.97), mas com o menor valor na experiência negativa entre as três versões (v2: $M = 0.48$, 95% IC = 0.34, 0.62, v2: $M = 0.28$, 95% IC = 0.15, 0.42 e v3: $M = 0.63$, 95% IC = 0.45, 0.81).

Os resultados do PANAS (Figura 42) mostram que as reações à versão v1 foram mais fortes, sendo esta versão a que apresentou os maiores valores tanto nos afetos positivos quanto nos negativos ($M = 2.63$, 95% IC = 2.51, 2.75 e $M = 1.39$, 95% IC = 1.25, 1.52, respectivamente). Entre as versões alteradas, os jogadores sentiram mais emoções positivas em v2 ($M = 2.52$, 95% IC = 2.35, 2.69) que em v3 ($M = 2.04$, 95% IC = 1.87, 2.21). Também sentiram ligeiramente menos emoções negativas em v2 ($M = 1.11$, 95% IC = 0.93, 1.29) que em v3 ($M = 1.19$, 95% IC = 1.02, 1.36).

Para verificar se esses resultados são significativos, foi realizado o teste de soma de postos de Kruskal-Wallis (MCKIGHT; NAJAB, 2010), um teste não paramétrico para amostras independentes, após usar o teste de Shapiro-Wilk para verificar que os dados não possuíam distribuição normal. Os resultados do teste de Kruskal-Wallis encontram-se na Tabela 6. Foi feito um teste *post hoc* para comparar as versões duas a duas. Os resultados encontram-se na Tabela 7.

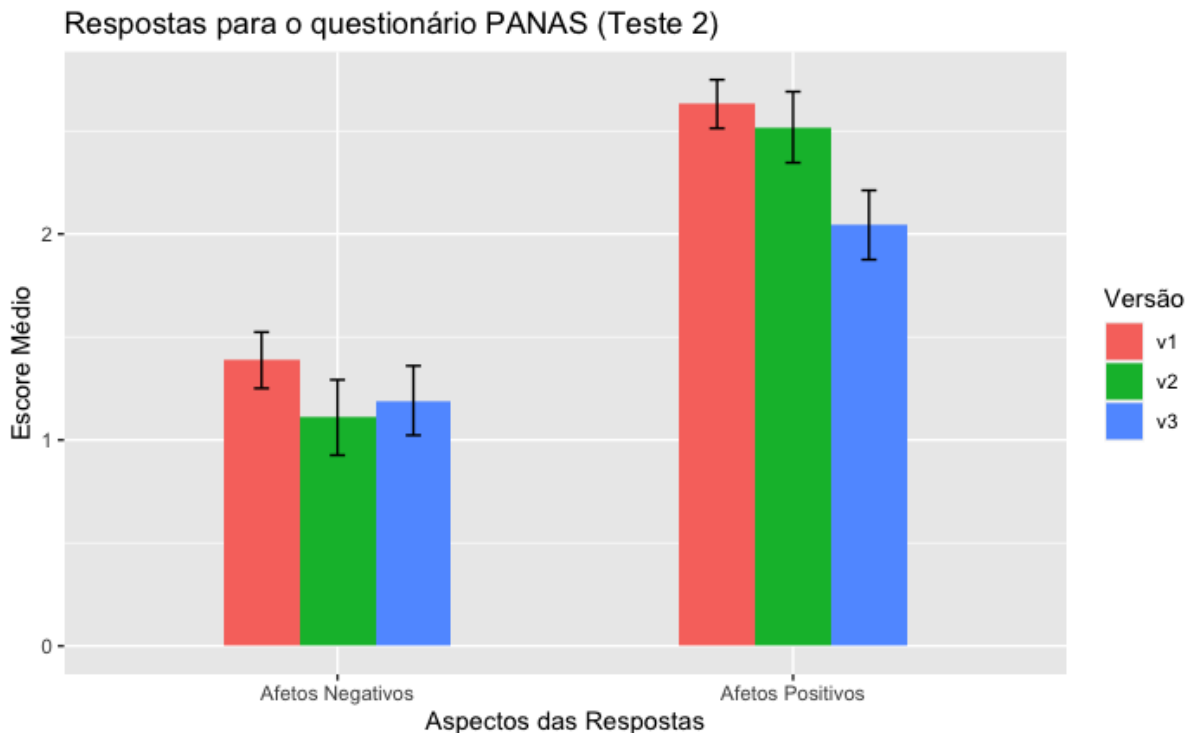
Os escores para “afeto positivo” do GEQ foram influenciados significativamente pelo grupo de v3 ($H(2) = 14.79$, $p = 0.0006$). A análise comparativa mostrou que v1 e v2

Figura 41 – Respostas aos módulos Core e Post-Game do GEQ (Teste 2), com intervalo de confiança de 95%.



Fonte: Elaborada pelo autor.

Figura 42 – Respostas ao PANAS (Teste 2), com intervalo de confiança de 95%.



Fonte: Elaborada pelo autor.

são semelhantes entre si ($dif = 5.88$), mas v3 apresenta escores significativamente menores que as outras duas versões ($dif_{v1v3} = 46.40, dif_{v2v3} = 52.28$). A versão v3 também afetou

Tabela 6 – Resultados dos testes de soma de posto de Kruskal-Wallis para os três questionários.

| Questionário | Aspecto | Resultado do teste de Kruskal-Wallis | Escore de cada versão |
|--------------|----------------------|------------------------------------------|---------------------------------------------|
| GEQ | Afeto Positivo | $\chi^2 = 14.785, gl = 2, p = 0.0006158$ | v1 = 221.66, v2 = 227.54, v3 = 175.26 |
| | Afeto Negativo | $\chi^2 = 4.2345, gl = 2, p = 0.1204$ | v1 = 126.81, v2 = 113.61, v3 = 138.30 |
| | Experiência Positiva | $\chi^2 = 21.411, gl = 2, p = 0.00002$ | v1 = 272.18, v2 = 267.37, v3 = 204.90 |
| | Experiência Negativa | $\chi^2 = 12.863, gl = 2, p = 0.00161$ | v1 = 207.07, v2 = 190.94, v3 = 235.02 |
| PANAS | Afeto Positivo | $\chi^2 = 31.367, gl = 2, p = 0.00001$ | v1 = 454.21, v2 = 437.48, v3 = 347.09 |
| | Afeto Negativo | $\chi^2 = 8.146, gl = 2, p = 0.01703$ | v1 = 443.40, v2 = 389.53, v3 = 411.28 |

Fonte: Elaborada pelo autor.

significativamente o “afeto negativo” ($H(2) = 4.23, p = .12$). Nesse item, todas as versões apresentam resultados diferentes.

Para o item “experiência positiva” do GEQ, v3 novamente influenciou significativamente os escores ($H(2) = 21.41, p = 0.00002$). Os escores de v1 e v2 são semelhantes entre si ($dif = 4.82$), mas novamente significativamente maiores quando comparados a v3 ($difv1v3 = 67.29, difv2v3 = 62.47$). A “experiência negativa” também foi impactada por v3 ($H(2) = 12.86, p = 0.0016$). Nesse item, no entanto, os escores de v3 foram significativamente maiores apenas que os de v2 ($dif = 44.08$). Não foram vistas diferenças significativas de v1 para as demais versões ($difv1v2 = 16.12, difv1v3 = 27.95$).

No caso do “afeto positivo” do PANAS, v3 impactou significativamente ($H(2) = 31.37, p = 0.00001$). As versões v1 e v2 foram similares entre si ($dif = 16.72$) e tiveram escores significativamente superiores aos de v3 ($difv1v3 = 107.11, difv2v3 = 90.40$). O “afeto negativo” também foi influenciado por v3 ($H(2) = 8.15, p = 0.017$). Os escores de v3 não foram substancialmente diferentes de v1 e v2 ($difv1v2 = 32.12, difv1v3 = 21.75$), mas v1 teve um escore significativamente maior que o de v2 ($dif = 53.87$).

Os resultados mostraram que v2 é estatisticamente semelhante a v1 em todos os aspectos, com exceções do item “afetos negativos” do questionário PANAS. Há diferenças significativas de v1 e v3 em todos os casos positivos, que são os itens “afeto positivo” e “experiência positiva” do GEQ e “afeto positivo” do PANAS. Por último, v2 e v3 são significativamente diferentes nos seguintes aspectos: “afeto positivo”, “experiência positiva” e “experiência negativa”

Tabela 7 – Resultados dos testes *post hoc* para os três questionários.

| Questionário | Aspecto | Comparação | Resultado do teste <i>post hoc</i> | Diferença? |
|--------------|----------------------|------------|-----------------------------------------------------|------------|
| GEQ | Afeto Positivo | v1 e v2 | <i>dif.obs.</i> = 5.88, <i>dif.crit.</i> = 34.65 | Não |
| | | v1 e v3 | <i>dif.obs.</i> = 49.40, <i>dif.crit.</i> = 34.16 | Sim |
| | | v2 e v3 | <i>dif.obs.</i> = 52.28, <i>dif.crit.</i> = 38.75 | Sim |
| | Afeto Negativo | v1 e v2 | <i>dif.obs.</i> = 13.20, <i>dif.crit.</i> = 26.86 | Não |
| | | v1 e v3 | <i>dif.obs.</i> = 11.49, <i>dif.crit.</i> = 26.48 | Não |
| | | v2 e v3 | <i>dif.obs.</i> = 24.68, <i>dif.crit.</i> = 30.04 | Não |
| | Experiência Positiva | v1 e v2 | <i>dif.obs.</i> = 4.82, <i>dif.crit.</i> = 37.95 | Não |
| | | v1 e v3 | <i>dif.obs.</i> = 67.28, <i>dif.crit.</i> = 37.42 | Sim |
| | | v2 e v3 | <i>dif.obs.</i> = 62.47, <i>dif.crit.</i> = 42.44 | Sim |
| | Experiência Negativa | v1 e v2 | <i>dif.obs.</i> = 16.12, <i>dif.crit.</i> = 34.65 | Não |
| | | v1 e v3 | <i>dif.obs.</i> = 27.95, <i>dif.crit.</i> = 34.16 | Não |
| | | v2 e v3 | <i>dif.obs.</i> = 44.07, <i>dif.crit.</i> = 38.75 | Sim |
| PANAS | Afeto Positivo | v1 e v2 | <i>dif.obs.</i> = 16.72, <i>dif.crit.</i> = 48.97 | Não |
| | | v1 e v3 | <i>dif.obs.</i> = 107.11, <i>dif.crit.</i> = 48.295 | Sim |
| | | v2 e v3 | <i>dif.obs.</i> = 90.39, <i>dif.crit.</i> = 54.78 | Sim |
| | Afeto Negativo | v1 e v2 | <i>dif.obs.</i> = 53.87, <i>dif.crit.</i> = 48.97 | Sim |
| | | v1 e v3 | <i>dif.obs.</i> = 32.12, <i>dif.crit.</i> = 48.29 | Não |
| | | v2 e v3 | <i>dif.obs.</i> = 21.74, <i>dif.crit.</i> = 54.77 | Não |

Fonte: Elaborada pelo autor.

do GEQ e “afeto positivo” do PANAS.

A maioria dos resultados aponta que, dentre as três versões, v2 é versão que os jogadores gostaram mais, conseqüentemente o protótipo mais bem sucedido. Além disso, em praticamente todos os casos, a versão v3 se mostrou inferior. Surpreendentemente, v3, segundo as métricas estabelecidas neste trabalho, é considerada pior que v1, o protótipo ao qual ela se baseia. A discussão sobre esse resultado será apresentada no Capítulo 6.

5.3 Considerações Finais

Neste capítulo foram apresentados os resultados dos diversos testes propostos na metodologia. O resultado da análise desses testes influenciou nas modificações da versão inicial do protótipo. Os protótipos foram, então, comparados. Este capítulo concentrou-se no relato da execução das técnicas de GUR. As limitações e lições aprendidas serão discutidas no capítulo seguinte.

6 LIMITAÇÕES E LIÇÕES APRENDIDAS

Neste capítulo, são discutidas as limitações do trabalho e as lições aprendidas durante a sua execução, bem como reflexões acerca do resultado surpreendente reportado no Capítulo 5. A Seção 6.1 discute as limitações do trabalho e as medidas utilizadas para contorná-las, quando possível. A Seção 6.2 mostra as lições aprendidas com todo o processo. Encerra-se o capítulo com as considerações finais na Seção 6.3.

6.1 Limitações

Este trabalho executou uma metodologia que combinou o método científico com os métodos empregados na produção de jogos digitais (CHANDLER, 2009) visando torná-lo relevante tanto para a academia e quanto para a indústria. Em trabalhos similares, há casos nos quais um ou mais autores também participam do desenvolvimento. Se por um lado, a participação de um pesquisador no desenvolvimento introduz vieses na pesquisa (GÓMEZ-MAUREIRA *et al.*, 2014), por outro, a completa terceirização desse processo não representa uma situação cotidiana na indústria (MIRZA-BABAEI *et al.*, 2013a). Neste trabalho, optou-se pela primeira opção, visto o caráter de contribuição para o campo prático. Algumas medidas para a redução da influência de vieses são a participação de um *designer* externo à pesquisa, o uso de testes com usuários para validar as alterações de *game design*, e as entrevistas com desenvolvedores nacionais de jogos para avaliação das ferramentas e técnicas utilizadas.

Este trabalho também possui escopo bastante reduzido: mesmo em estúdios *indie* costuma-se utilizar métricas de milhares de usuários. Pode-se argumentar que o escopo está atrelado, em parte, à ausência de um teste anônimo online. Uma desvantagem advinda disso é a ausência de estratificação dos usuários. Com uma amostra pequena de testadores, é difícil obter conclusões significativas isolando por características como idade, sexo e preferências de estilos de jogos.

O GEQ e PANAS originais foram escritos na língua inglesa. Até o momento da aplicação dos testes, não havia versão traduzida oficial do GEQ. Já o PANAS apresenta uma versão em português de Portugal (GALINHA; PAIS-RIBEIRO, 2005). Uma versão para português brasileiro foi testada e divulgada em 2019 (NUNES *et al.*, 2019), posteriormente à execução dos testes. Neste trabalho foi utilizada uma tradução informal de ambos os questionários em vez do formulário em inglês ou em português lusitano.

O projeto se sustenta na premissa do uso de metodologias de baixo custo, portanto excluiu o uso de biometria e *eye-trackers* dedicados. Para alguns desenvolvedores, no entanto, essa limitação pode simplesmente não ser um problema. Notavelmente, *eye-trackers* dedicados, como o Tobii¹, podem ser considerados financeiramente acessíveis em diversos países.

A escolha de utilizar a webcam para o WebGazer impossibilitou que a mesma fosse usada para sistemas de detecção de emoções via expressões faciais como os desenvolvidos por Talegaonkar *et al.* (2019), Dingli e Giordimaina (2017) e Magdin *et al.* (2016). Tais sistemas podem ser escolhidos como alternativas de baixo custo aos métodos biométricos de GUR.

Ademais, a metodologia do trabalho também assumiu que existe uma correlação simples entre diversão e dificuldade (ex.: quanto mais difícil for um desafio, mais divertido será superá-lo). Na Etapa 2 a dificuldade é utilizada como critério de avaliação de qualidade, porém, na Etapa 3, os protótipos são avaliados conforme o que o jogador sentiu enquanto jogava. Os resultados mostrados no Capítulo 5 indicam que a versão v2 foi vista como a mais divertida, porém os desenvolvedores consideram v3 a mais fácil. Isto novamente reforça a importância de realizar testes com usuários e de se ter métricas de avaliação de *level design*.

Downside Up, assim como os chamados jogos “*masocore*”, propositadamente provoca os jogadores e intercala desafios intensos e fáceis. Embora o jogo não apresente elementos de aleatoriedade, armadilhas foram posicionadas de maneiras tidas como injustas. Isto é, reconhecidamente, um princípio controverso de *game design* que pode forçar jogadores a resolver desafios usando tentativa e erro. Isso tornou complexa a avaliação de o quão adequada a dificuldade de um desafio está.

Jogos de plataforma “*masocore*” contêm mecânicas bastante específicas. Sendo assim, é mais difícil generalizar e afirmar que as descobertas discutidas neste trabalho sejam igualmente válidas para outros jogos de plataforma 2D, ou até a outros gêneros de jogos. O estilo de jogo também apresenta um perfil de usuário peculiar. Encontrar participantes cujos jogos favoritos são desse estilo foi, como esperado, problemático. Considerando todos os participantes, a maioria preferiu jogos de RPG ou MOBA (*Multiplayer Online Battle Arena*), que se distanciam consideravelmente do público-alvo do jogo desenvolvido. A especificidade do jogo, no entanto, trouxe consequências positivas, que serão discutidas na próxima seção.

Os participantes podem ter se sentido pressionados a terminar o jogo uma vez que possuíam o conhecimento de que estavam participando de um teste. Embora alguns tenham

¹ Especificações: <https://gaming.tobii.com/product/tobii-eye-tracker-4c/>

perguntado sobre uma data de lançamento de uma possível versão comercial do jogo, a abordagem executada neste trabalho não considerou a hipótese de desistências, portanto não há uma estimativa precisa de quantos testadores realmente terminariam o jogo.

Um resumo das principais limitações encontra-se listado a seguir:

- Autor da pesquisa assumiu os papéis de pesquisador e desenvolvedor do jogo;
- Quantidade de testadores pequena, especialmente nos testes que envolveram *eye-tracking*;
- Uso de *eye-tracking* de baixo-custo;
- Ausência de um sistema capaz de avaliar afetos de forma automática;
- Associação de diversão e dificuldade sem considerar as nuances das mesmas;
- Uso de um jogo desenvolvido com uma filosofia de *level design* incomum, dificultando avaliações posteriores;
- Maioria dos testadores não corresponde ao público-alvo planejado.
- Desistências dos usuários não foram contabilizadas.

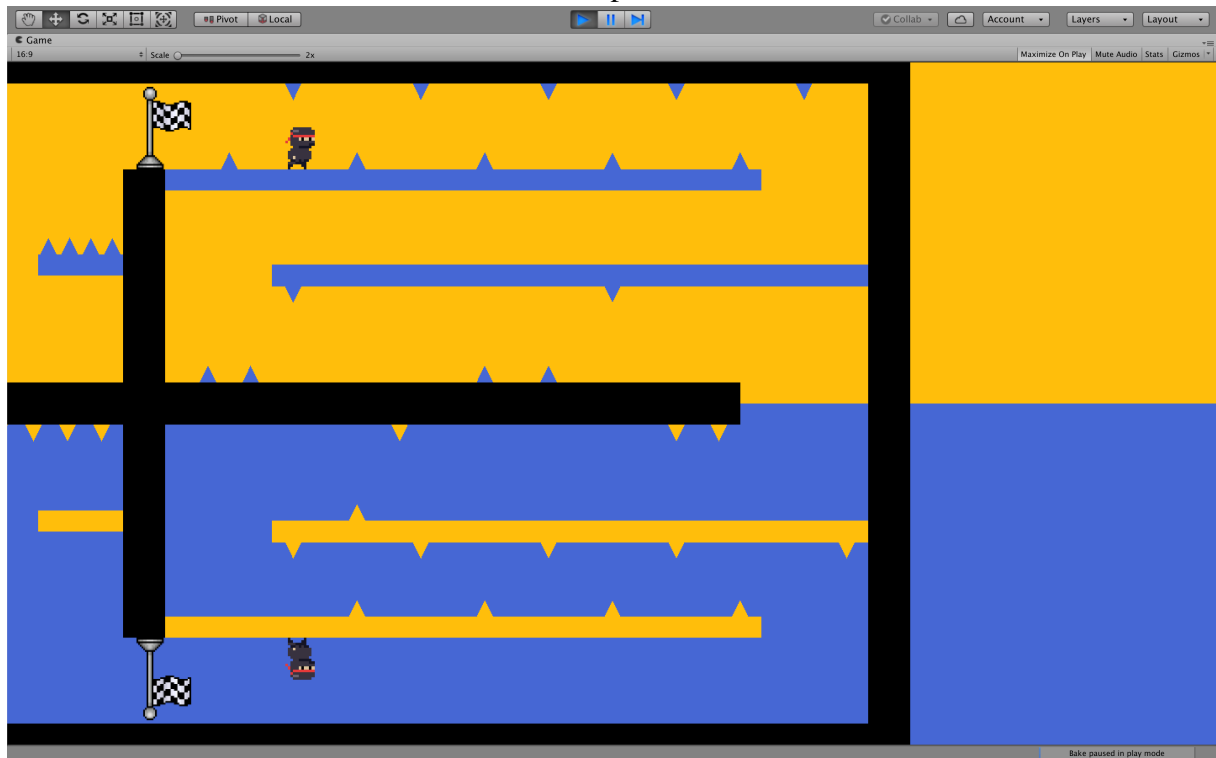
6.2 Lições aprendidas

Houve, de fato, grandes benefícios aos desenvolvedores ao aplicar diferentes métodos de GUR. A combinação de métricas com *eye-tracking* não apenas confirmou os dados obtidos pela observação e entrevistas, mas também proveu *insights* sobre a compreensão de *level design* em um jogo multitarefa. Em contrapartida, essa combinação não forneceu informações sobre o que os jogadores sentiram ou como se deu o processo de raciocínio para a resolução de um desafio. Para isto, a combinação de métricas, *eye-tracking* e questionário foi a mais proveitosa.

Downside Up requer que o jogador lide com duas telas simultaneamente. Foi observado que os jogadores tenderam a focar na tela superior, que se comportava como um jogo de plataforma tradicional. Curiosamente, no desafio 3F (Figura 43), que consiste em corredores em formato de S repletos de espinhos, os dois primeiros andares estão organizados da mesma forma. Os jogadores relataram que o andar do meio é consideravelmente mais difícil que o de cima. Isto pode ter acontecido devido ao esforço cognitivo necessário para mover-se na dimensão espelhada da tela de baixo.

Visualizar os dados de diferentes fontes facilitou a compreensão de como os participantes jogaram o jogo de uma forma mais simples e rápida que checar as notas de observação para cada participante. A visualização também conseguiu mostrar dados que podem não ser tão facilmente observáveis de outras maneiras. A Figura 44 mostra que no desafio 3F há uma grande

Figura 43 – Captura de tela do desafio 3F de Downside Up. Curiosamente, os jogadores tiveram muito mais dificuldade no andar do meio, a despeito de ele ser idêntico ao andar de cima



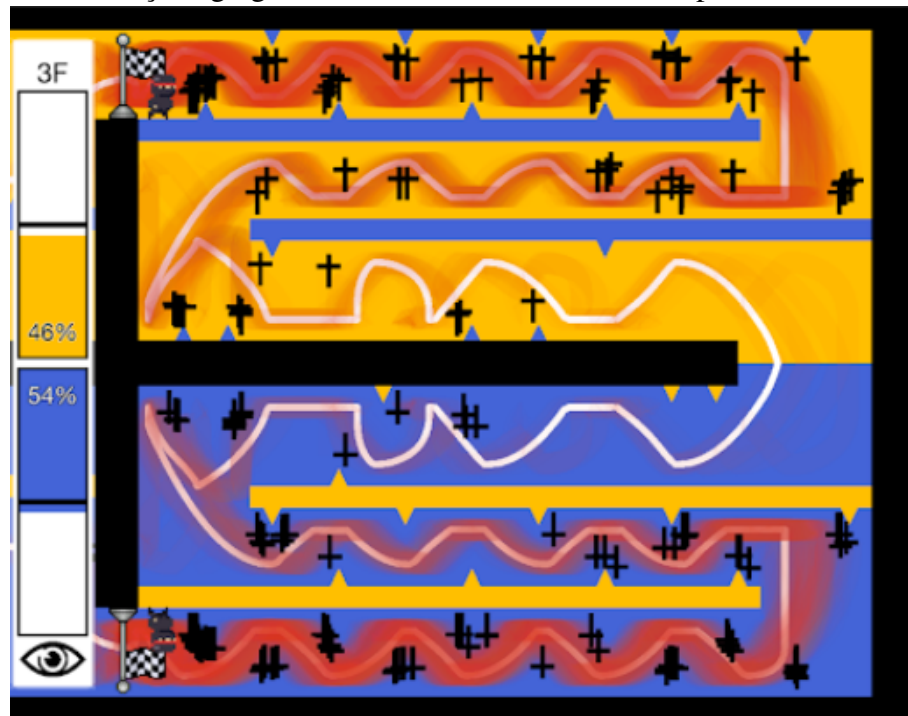
Fonte: Elaborada pelo autor.

concentração de mortes e repetição de movimentos no primeiro andar. Tanto a quantidade de marcações em forma de cruz quanto a opacidade da trajetória vermelha são reduzidas conforme o jogador desce. Isto indica que os jogadores aprenderam a manobrar no cenário no primeiro andar até ganharem confiança para repetir o movimento nos andares seguintes. Além disso, a concentração da cor vermelha ao redor da trajetória branca indica que os jogadores estavam focados em praticar a execução da solução ótima do desafio, o que vai ao encontro do design focado em destreza proposto para este desafio.

No caso específico de Downside Up, usar o WebGazer como ferramenta de avaliação de *level design* proporcionou informações valiosas sobre a atenção dos jogadores. Como mencionado anteriormente, as imprecisões do WebGazer podem torná-lo inadequado para certos tipos de jogos. No entanto, a ferramenta funcionou em Downside Up principalmente porque o jogo poderia ser dividido em duas áreas de interesse (cima e baixo). É possível que essa mesma solução funcione apropriadamente para outros jogos que apresentem poucas mudanças de câmera e cujo espaço possa ser discretizado em porções menores.

Os questionários escolhidos, GEQ e PANAS, tratam afetos negativos como estritamente indesejáveis, o que pode nem sempre estar de acordo com o *game design*. É possível que

Figura 44 – Visualização agregada do desafio 3F de Downside Up



Fonte: Fornecida pelo autor.

a adição de um espaço para comentários nesses questionários ajudasse a compreender melhor as experiências dos respondentes.

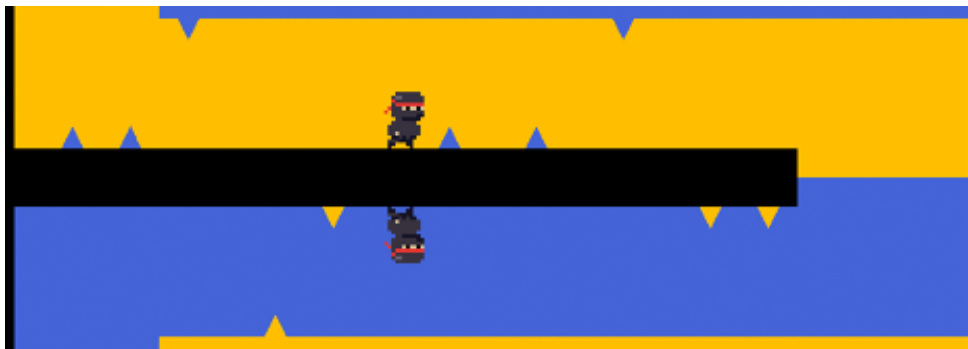
A Etapa 2 da metodologia consistiu em testar v1 e encontrar pontos a ser melhorados nos protótipos subsequentes. Conforme já apontado no experimento de Mirza-Babaei *et al.* (2013a), foi observado que, neste experimento, os protótipos desenvolvidos na Etapa 2 deveriam ter passado por um reteste antes do começo da Etapa 3 para assegurar a qualidade das mudanças. Mesmo que apoiado por dados, o processo de *game design* continua sendo bastante subjetivo e propenso a erro humano. Um desafio que pode ter sido motivado com a inclusão de frustração indesejada foi o desafio 3F (ver Figura 45). Os resultados dos testes levaram a crer que o mesmo deveria se tornar mais difícil na versão v3, pois os jogadores estavam realizando a movimentação dos andares do meio e de baixo com maior facilidade que no primeiro andar (rever Figura 44). Apesar de um ajuste mais refinado em diversos aspectos que tradicionalmente qualificam um bom *game design*, como progressão da dificuldade, clareza nos tutoriais e melhorias na usabilidade, é possível que o desafio 1F tenha sido modificado para tornar-se excessivamente difícil na versão v3. Especula-se que este desafio tenha sido um dos responsáveis pelos *scores* inferiores de v3 nos questionários de avaliação de experiência.

Os feedbacks para a versão v3 sugeriram que o narrador poderia ser mais presente. Dessa forma, foram incluídos textos em quase todos os desafios dessa versão. Alguns textos já

Figura 45 – Porções finais do desafio 3F conforme implementado nas versões v1 (a), v2 (b) e v3 (c). A adição e reposicionamento de espinhos pode ter tornado v3 mais difícil do que o previsto



(a) Parte do desafio 3F em v1, o qual v2 e v3 se basearam



(b) Esta parte do desafio manteve-se inalterada em v2

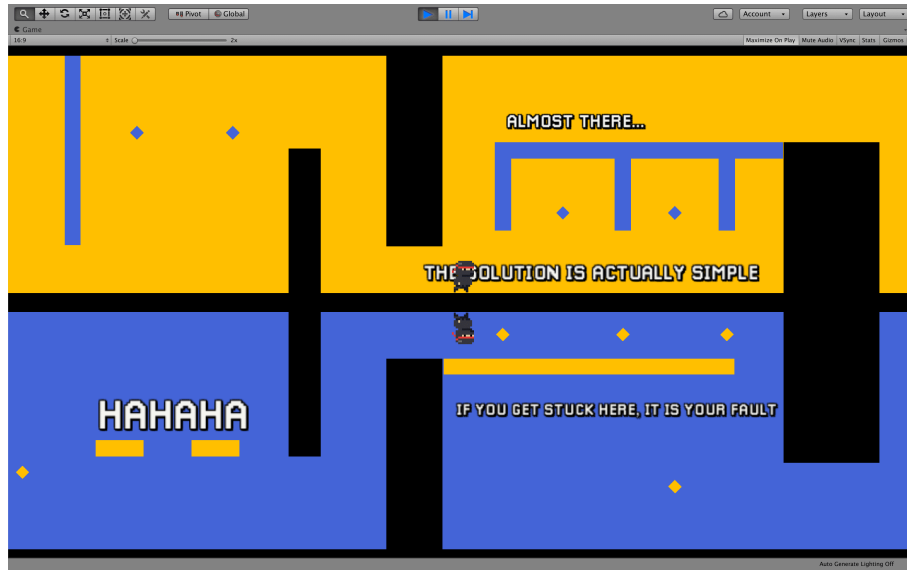


(c) Esta parte do desafio foi alterada em v3 para se tornar mais difícil

existentes foram alterados para tornar o narrador mais provocativo. Um exemplo encontra-se na Figura 46, que mostra o desafio 1F presente na versão v3. Na figura, o texto original “vai ligeiro” foi substituído por “a solução na verdade é simples” e “se você ficar preso aqui, foi por culpa sua”. O desafio 1F, no entanto, foi alterado para se tornar mais fácil de ser resolvido. É possível que o teor das provocações possa ter se tornado mais agressivo, intimidador ou desestimulante, influenciando negativamente a experiência dos jogadores menos habilidosos no gênero. Esta hipótese também poderia ser avaliada em caso de reteste.

Foi observado, com base nas divergências entre as avaliações dos usuários e a

Figura 46 – Desafio 1F presente na versão v3. É possível que mudanças nos textos do narrador tenham ocasionado experiências menos agradáveis a jogadores menos habilidosos em jogos “masocore”



Fonte: Fornecida pelo autor.

expectativa dos designers, que a dificuldade dos desafios também tem influências relativas. Em outras palavras, a percepção de dificuldade é influenciada pelas adjacências. Por exemplo, se um desafio fácil for seguido de um desafio difícil e em seguida por outro fácil, esse desafio do meio vai ser percebido como mais difícil, por consequência. Da mesma maneira, os desafios fáceis passam a ser percebidos como mais fáceis. Essa questão da dificuldade também poderia ter sido mais refinada caso houvesse uma rodada de testes de avaliação de level *level design* com v2 e v3.

Foram poucos casos em que a maioria dos usuários votou em uma dificuldade mediana, o que indica que a progressão de dificuldade talvez não estivesse suficientemente refinada. Uma possível consequência disso pode ter ocorrido pelas constantes reações do tipo “ame ou odeie” por parte dos jogadores. Embora uma parcela significativa dos testadores em todas as etapas tenha mostrado sinais de raiva, também houve uma grande quantidade de elogios ao fato de o jogo causar raiva, uma vez que a sensação de maestria causada pela superação dos desafios é prazerosa. Por outro lado, foi mencionado diversas vezes que o jogo estava excessivamente difícil, e que a dificuldade por si só já seria motivo suficiente para o abandono do jogo, caso os participantes estivessem jogando em um ambiente casual.

Embora não haja dados biométricos na Etapa 3 para confirmar esta afirmação, acredita-se que, baseado nas anotações de observação e entrevistas na Etapa 2, a “boa” frustração acontece quando o jogo provoca o jogador de forma astuta ou quando o jogador sente que é culpa dele não ter conseguido realizar alguma tarefa. A “má” frustração acontece quando os jogadores

sentem que o progresso está sendo impedido pelo próprio jogo. Exemplos disso incluem espinhos posicionados em locais onde mortes acidentais acontecem, seções em que manobrar o avatar é trabalhoso, plataformas que requerem pulos difíceis de executar, seções onde há elementos cruciais escondidos pela câmera e momentos em que o jogo requer tanto reflexos precisos e agilidade quanto memória e visão das duas telas ao mesmo tempo.

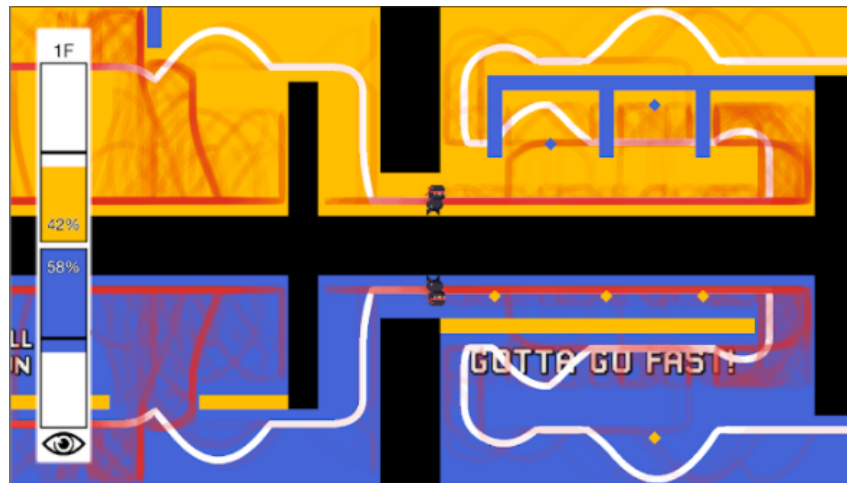
A aplicação da frustração no *level design* ainda hoje é um assunto controverso, uma vez que não é possível identificar precisamente que tipos de jogadores são capazes de resistir a ela sem abandonar o jogo. Embora seja possível listar as características de algo frustrante, os métodos para chegar a tal são, essencialmente, qualitativos e dependentes de subjetividade. Mesmo o desenvolvimento de um sistema capaz de prever situações frustrantes pode falhar em julgar se aquilo entretém o usuário ou não (CANOSSA *et al.*, 2011).

Dito isso, com o auxílio das técnicas de GUR, foi possível encontrar exemplos de desafios que empregam a “boa frustração”. É o caso do desafio 3E, cuja visualização agregada encontra-se na Figura 47b. O desafio consiste em utilizar quatro losangos para desativar quatro plataformas azuis com espinhos e pular rapidamente pelas plataformas amarelas até chegar em um local seguro. Como os espinhos voltam depois de um tempo, muitos jogadores precisaram repetir o desafio, considerado difícil pela maioria dos jogadores. O desafio 3E é um *redesign* do último desafio da primeira fase (Figura 47a), adicionando espinhos como punição pela falta de destreza. Tanto em 1F quanto em 3E, os jogadores esboçaram reações muito positivas ao conseguir atravessar os obstáculos. O comportamento nos dois casos também foi similar: a opacidade da trajetória vermelha é maior no começo, indicando que os jogadores repetiram o movimento algumas vezes até entenderem que a dificuldade do desafio reside na rapidez dos *inputs*.

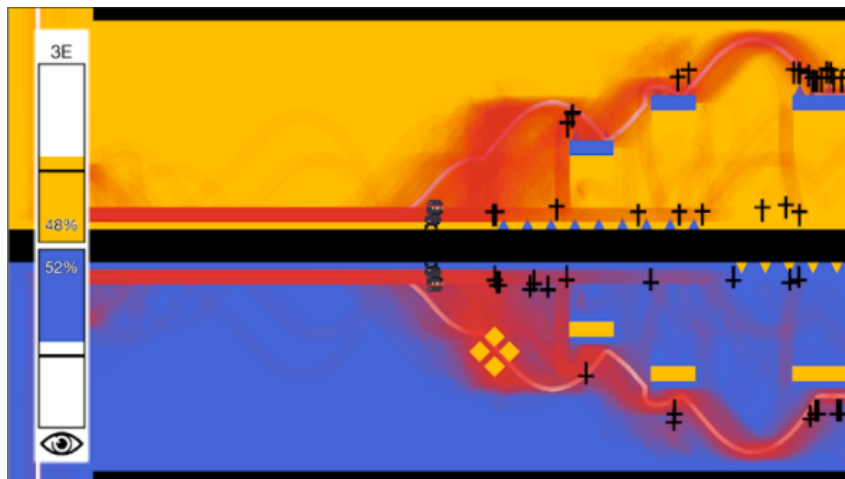
Apesar de o perfil dos testadores ter se distanciado do público-alvo do jogo, ainda foi possível extrair informações proveitosas sobre o *level design* com base nos testes. Jogadores inexperientes, por exemplo, providenciaram dados que melhor ajudaram a regular desafios introdutórios, como tutoriais, e apontaram mais claramente falhas de usabilidade. Jogadores mais experientes passaram pelos tutoriais com maior facilidade, provendo *feedback* menos rico nesse caso, mas sua experiência de jogo ajudou a refinar desafios complexos, que testam se o jogador realmente aprendeu uma mecânica.

Por fim, os resultados sugerem que observação, entrevistas e questionários sozinhos já são suficientes para melhorar de forma eficiente o *level design*. Embora v2 tenha tido *score*

Figura 47 – Visualizações agregadas dos desafios 1F e 3E, que apresentam similaridades de *level design*.



(a) Visualização agregada do desafio 1F de Downside Up.



(b) Visualização agregada do desafio 3E de Downside Up.

maior que v3, do ponto de vista dos desenvolvedores, a inclusão de métricas e *eye-tracking* continuam sendo valiosas. Neste trabalho, o uso de dados espaciais em conjunto com dados sobre atenção foi capaz de expor problemas de usabilidade. Além disso, esses métodos não dependem da memória de quem aplicou o teste nem de quem participou, tornando-os mais precisos e menos propensos a vieses.

O uso de visualização acelerou o processo de desenvolvimento ao simplificar informações advindas de diversas fontes. Não só isso, como também seu uso foi validado e aprovado por profissionais da área. Surpreendentemente, a maioria dos usuários não empregava visualização rotineiramente em seus trabalhos. Isto sugere que, a despeito das grandes vantagens, a comunidade de desenvolvedores de jogos ainda desconhece a área de visualização de dados.

Um resumo das principais lições aprendidas encontra-se listado a seguir:

- O *eye-tracking* proveu *insights* sobre o desenvolvimento de um jogo que requer habilidades multitarefa;
- O *eye-tracking* via webcam funcionou muito bem em Downside Up devido ao ambiente discretizado em apenas duas regiões de interesse;
- A combinação das metodologias empregadas ofereceu informações sobre atenção, processo de raciocínio, dificuldade dos desafios, entre outros comportamentos;
- Há um esforço cognitivo para coordenar o avatar na tela cuja câmera é espelhada. Esse esforço, por exemplo, pode dificultar outras habilidades, como a destreza;
- Foi possível descobrir novas informações graças à visualização agregada;
- Os questionários GEQ e PANAS tratam afetos negativos, como a frustração, como estritamente indesejáveis, algo que pode nem sempre estar de acordo com o *level design*;
- É necessário realizar testes para confirmar se as melhorias implementadas não causam regressões na experiência do usuário. Sendo assim, o mínimo de rodadas de testes visando o aperfeiçoamento de um jogo é 2;
- O teor das frases do narrador pode influenciar negativamente a experiência dos usuários;
- Os jogadores percebem a dificuldade dos desafios de forma relativa com base na lembrança da dificuldade de desafios adjacentes;
- Existem casos em que é desejável desafios frustrantes. Nesses casos, o prazer da vitória deve ser superior ao incômodo da frustração;
- A visualização agregada desenvolvida nesse trabalho consegue sugerir quais desafios apresentam a “boa frustração”;
- O *feedback* de jogadores inexperientes ou que não se enquadram no público-alvo de um jogo contribui na identificação de problemas de usabilidade;
- A prática de visualização de dados, apesar de bastante benéfica aos desenvolvedores, ainda é pouco difundida no meio.

6.3 Considerações Finais

Este capítulo apresentou a discussão extraída com base nos resultados da aplicação da metodologia. A conclusão encontra-se no próximo Capítulo.

7 CONCLUSÕES E TRABALHOS FUTUROS

Neste capítulo, são apresentadas as considerações finais a respeito da pesquisa realizada nesta dissertação de mestrado. Na Seção 7.1, encontram-se os resultados alcançados. Estes resultados foram obtidos dentro de limitações resumidas na Seção 7.2. Na Seção 7.3, estão as possibilidades de continuidade para a pesquisa, os trabalhos futuros.

7.1 Resultados Alcançados

Esta dissertação investigou como GUR é aplicada para aperfeiçoar um jogo de plataforma 2D. Foi feito um experimento em que entrevistas, questionários, telemetria e eye-tracking, aliados a uma ferramenta de visualização de dados que agrega os resultados das últimas duas, foram capazes de prover informações sobre o comportamento dos usuários e mostrar com maior precisão pontos de melhoria no jogo.

Ao longo do caminho, foram descobertas limitações e especificidades do gênero *masocore*. Evidenciaram-se a complexidade de se avaliar a dificuldade dos desafios, a relação que os jogadores possuem com o sentimento de frustração, e as questões acerca de usar afetos positivos e negativos como fatores relacionados a diversão. As principais contribuições da abordagem executada são listadas a seguir:

- Uma abordagem de testes que usa apenas metodologias de baixo custo para aperfeiçoar um jogo;
- A criação de uma solução de visualização de dados capaz de agregar dados de métricas e *eye-tracking*. Este último, normalmente empregado para avaliação de interfaces, é utilizado aqui como um meio de avaliação de *level design*;
- A publicação online de um repositório público contendo todo o material utilizado neste trabalho (jogo, dados, visualizações etc.);
- Uma discussão sobre o uso de técnicas de testes com usuários e visualização de dados para o aperfeiçoamento de um jogo.

7.2 Limitações

O trabalho atingiu seus objetivos dentro de uma série de limitações de desenvolvimento, metodologia e avaliação. Elas são explicitadas a seguir:

- Escopo reduzido devido a restrições de tempo e tamanho da equipe;

- Estudo envolvendo apenas um jogo;
- Necessidade de jogadores com perfil específico;
- Quantidade de testadores pequena;
- Utilização de traduções próprias de questionários renomados sem o suporte de estudo prévio quanto a equivalência dos resultados nos dois idiomas;
- Aplicação somente de metodologias capazes de operar sob restrições orçamentárias;
- Premissa de que há uma correlação simples entre diversão e dificuldade no jogo escolhido;
- Realização de apenas uma rodada de testes para o aperfeiçoamento do protótipo.

7.3 Trabalhos Futuros

A área de GUR ainda está em amadurecimento, portanto há diversos direcionamentos possíveis na mesma. Nesta Seção, são apresentadas sugestões de continuidade considerando, principalmente, as limitações deste experimento.

Neste trabalho, as soluções desenvolvidas foram pensadas com o intuito de resolver problemas específicos de *Downside Up*. A realização de testes em uma abordagem similar com outros jogos de plataforma ou com jogos de outros estilos, por exemplo, de quebra-cabeça, daria maior robustez ao processo. Os jogos de quebra-cabeça são ideais por costumarem ter câmera fixa e regiões de interesse bem delimitadas.

O experimento adotou um jogo que requer habilidades multitarefa, algo incomum em jogos comerciais, visto que há um esforço cognitivo para dividir a atenção em múltiplas interfaces (HU *et al.*, 2016). Trabalhos futuros podem se aprofundar no estudo desse esforço utilizando visualização de dados nos jogos citados como inspiração para *Downside Up*, por exemplo.

O trabalho apontou uma potencial falha metodológica relacionada a como avaliar diversão e dificuldade em jogos *masocore*. Futuramente, pode-se estudar isto por meio de pesquisas qualitativas com jogadores de obras nesse estilo, como *Celeste* e *Super Meat Boy*, o que melhoraria futuras metodologias.

Durante o processo de desenvolvimento e avaliação, foi notado que frustração não é necessariamente um indicador ruim no que tange à criação de fases para jogos de plataforma *masocore* 2D. É possível que o uso de biometria seja capaz de apontar melhor quais são as

características das frustrações “boa” e “ruim”, teorizadas aqui. Caso o orçamento continue sendo uma limitação, é possível experimentar técnicas de baixo custo, porém menos precisas, também com a *webcam* (DINGLI; GIORDIMAINA, 2017).

Neste trabalho, foi utilizada com sucesso a *webcam* para realizar o *eye-tracking*. Trabalhos posteriores podem comparar a solução empregada com o uso de *eye-trackers* dedicados a fim de determinar em quais situações cada caso é apropriado.

Por fim, o perfil de usuário de maior valor, segundo o público alvo do jogo, estava em um subconjunto de jogadores mais raro do que o esperado. A repetição da metodologia com algum jogo de plataforma mais fácil ou popular, como os da franquia Super Mario Bros, poderia ser mais inclusiva. A utilização de testes online com jogadores globais também é uma possibilidade para a aquisição de dados em maior quantidade e variedade.

REFERÊNCIAS

- ACADEMY, N. Y. F. **Genre Revival: 2d platformers**. Estados Unidos: [s. n.], 2014. Disponível em: <https://www.nyfa.edu/student-resources/genre-revival-2d-platformers/>. Acesso em: 28 fev. 2020.
- ALBERT, W.; TULLIS, T. **Measuring the user experience: collecting, analyzing, and presenting usability metrics**. Oxford, Reino Unido: Newnes, 2013.
- ALMEIDA, S.; MEALHA, Ó.; VELOSO, A. Video game scenery analysis with eye tracking. **Entertainment Computing**, Elsevier, Amsterdam, Holanda, v. 14, p. 1–13, 2016.
- ALONSO-FERNÁNDEZ, C.; CANO, A. R.; CALVO-MORATA, A.; FREIRE, M.; MARTÍNEZ-ORTIZ, I.; FERNÁNDEZ-MANJÓN, B. Lessons learned applying learning analytics to assess serious games. **Computers in Human Behavior**, Elsevier, Amsterdam, Holanda, v. 99, p. 301–309, 2019.
- AMBINDER, M. Valve's approach to playtesting: the application of empiricism. In: **Game Developers Conference 2009**. San Francisco, Estados Unidos: [s. n.], 2009.
- ANDERSEN, E.; O'ROURKE, E.; LIU, Y.-E.; SNIDER, R.; LOWDERMILK, J.; TRUONG, D.; COOPER, S.; POPOVIC, Z. The impact of tutorials on games of varying complexity. In: ACM. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. Austin, Estados Unidos, 2012. p. 59–68.
- BASTOS, A. S.; SANTOS, E.; GOMES, G. A. M.; MOURAO, M. A. Evaluating the use of affordable user testing and visualization techniques in level design of a hardcore 2d platform game. **Proceedings of SBGames 2019**, Rio de Janeiro, Brasil, p. 145–154, 2019.
- BOSNJAK, M.; OREHOVACKI, T. Measuring quality of an indie game developed using unity framework. In: IEEE. **2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)**. Opatija, Croácia, 2018. p. 1574–1579.
- BOULTON, A.; HOURIZI, R.; JEFFERIES, D.; GUY, A. A little bit of frustration can go a long way. In: SPRINGER. **Advances in Computer Games**. Leiden, Holanda, 2017. p. 188–200.
- BRAUN, E. **The Rise and Fall and Rise of The Platformer**. Estados Unidos: [s. n.], 2018. Disponível em: <https://cultureofgaming.com/the-rise-and-fall-and-rise-of-the-platformer/>. Acesso em: 28 fev. 2020.
- CANOSSA, A.; DRACHEN, A.; SØRENSEN, J. R. M. Arrrgghh!!!: blending quantitative and qualitative methods to detect player frustration. In: ACM. **Proceedings of the 6th international conference on foundations of digital games**. Bordeaux, França, 2011. p. 61–68.
- CARTER, M.; DOWNS, J.; NANSEN, B.; HARROP, M.; GIBBS, M. Paradigms of games research in hci: a review of 10 years of research at chi. In: ACM. **Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play**. Toronto, Canadá, 2014. p. 27–36.
- CHAN, S. **The Story of Ultraflow: Designing with unity analytics**. Estados Unidos: [s. n.], 2015. Disponível em: <https://blogs.unity3d.com/2015/08/18/the-story-of-ultraflow-designing-with-unity-analytics/>. Acesso em: 12 mai. 2019.

- CHANDLER, H. M. **Manual de produção de jogos digitais**. Porto Alegre: Bookman Editora, 2009.
- CHEN, C.-h.; HÄRDLE, W. K.; UNWIN, A. **Handbook of data visualization**. London, Reino Unido: Springer Science & Business Media, 2007.
- CSIKSZENTMIHALYI, M. **Flow: the psychology of happiness**. Sydney, Austrália: Random House, 2013.
- CSIKSZENTMIHALYI, M.; ABUHAMDEH, S.; NAKAMURA, J. Flow. In: **Flow and the foundations of positive psychology**. McLean, Estados Unidos: Springer, 2014. p. 227–238.
- DINGLI, A.; GIORDIMAINA, A. Webcam-based detection of emotional states. **The Visual Computer**, Springer, London, Reino Unido, v. 33, n. 4, p. 459–469, 2017.
- DÖRNER, R.; GÖBEL, S.; KICKMEIER-RUST, M.; MASUCH, M.; ZWEIG, K. **Entertainment Computing and Serious Games: International gi-dagstuhl seminar 15283, dagstuhl castle, germany, july 5-10, 2015, revised selected papers**. London, Reino Unido: Springer, 2016. v. 9970.
- DRACHEN, A.; MIRZA-BABAEI, P.; NACKE, L. E. **Games user research**. Oxford, Reino Unido: Oxford University Press, 2018.
- DRACHEN, A.; SCHUBERT, M. Spatial game analytics and visualization. In: IEEE. **2013 IEEE Conference on Computational Intelligence in Games (CIG)**. Niagara Falls, Canadá, 2013. p. 1–8.
- DRENIKOW, B.; MIRZA-BABAEI, P. Vixen: interactive visualization of gameplay experiences. In: **Proceedings of the 12th International Conference on the Foundations of Digital Games**. Hyannis, Estados Unidos: [s. n.], 2017. p. 1–10.
- DUPUIS, B. **Saber's Edge**. Estados Unidos: [s. n.], 2016. Disponível em: <https://connect.unity.com/p/games-sabers-edge>. Acesso em: 13 mai. 2020.
- EL-NASR, M. S.; DRACHEN, A.; CANOSSA, A. **Game analytics**. London, Reino Unido: Springer, 2016.
- EL-NASR, M. S.; YAN, S. Visual attention in 3d video games. In: ACM. **Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology**. Hollywood, Estados Unidos, 2006. p. 22.
- FARNSWORTH, B. **What Is EMG (Electromyography) and How Does It Work?** Estados Unidos: [s. n.], 2018. Disponível em: <https://imotions.com/blog/electromyography-101/>. Acesso em: 7 fev. 2020.
- FEDEROFF, M. A. **Heuristics and usability guidelines for the creation and evaluation of fun in video games**. Tese (Doutorado) – Citeseer, Pennsylvania, Estados Unidos, 2002.
- FEITOSA, V. R.; MAIA, J. G.; MOREIRA, L. O.; GOMES, G. A. Gamevis: game data visualization for the web. In: IEEE. **2015 14th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)**. Piauí, Brasil, 2015. p. 70–79.

- GALINHA, I. C.; PAIS-RIBEIRO, J. L. Contribuição para o estudo da versão portuguesa da positive and negative affect schedule (panas): Ii-estudo psicométrico. **Análise psicológica**, Instituto Superior de Psicologia Aplicada, Lisboa, Portugal, v. 23, n. 2, p. 219–227, 2005.
- GOMES, G. A.; SANTOS, E.; VIDAL, C. A.; SILVA, T. L. C. da; MACEDO, J. A. F. Real-time discovery of hot routes on trajectory data streams using interactive visualization based on gpu. **Computers & Graphics**, Elsevier, Amsterdam, Holanda, v. 76, p. 129–141, 2018.
- GÓMEZ-MAUREIRA, M. A.; WESTERLAKEN, M.; JANSSEN, D. P.; GUALENI, S.; CALVI, L. Improving level design through game user research: a comparison of methodologies. **Entertainment Computing**, Elsevier, Amsterdam, Holanda, v. 5, n. 4, p. 463–473, 2014.
- GOUGH, C. **Genre breakdown of video game sales in the United States in 2017**. Estados Unidos: [s. n.], 2019. Disponível em: <https://www.statista.com/statistics/189592/breakdown-of-us-video-game-sales-2009-by-genre/>. Acesso em: 8 mai. 2019.
- GUALENI, S.; JANSSEN, D.; CALVI, L. How psychophysiology can aid the design process of casual games: a tale of stress, facial muscles, and paper beasts. In: ACM. **Proceedings of the international conference on the foundations of digital games**. Raleigh, Estados Unidos, 2012. p. 149–155.
- HALABI, N. **Moving towards enhancing playtesting reports through data visualization**. Tese (Doutorado) – University of Ontario Institute of Technology, Oshawa, Canadá, 2019.
- HALABI, N.; WALLNER, G.; MIRZA-BABAEI, P. Assessing the impact of visual design on the interpretation of aggregated playtesting data visualization. In: **Proceedings of the Annual Symposium on Computer-Human Interaction in Play**. Barcelona, Espanha: [s. n.], 2019. p. 639–650.
- HCT, S. **Biometric Storyboards**. Canadá: [s. n.], 2012. Disponível em: <https://sussexhct.wordpress.com/2012/10/23/lab-meeting-biometric-storyboards/>. Acesso em: 8 mai. 2020.
- HOFFMANN, G.; MARTIN, R.; WEINHARDT, C. Perfectionism in games-analyzing playing behaviors in an educational game. In: IEEE. **2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)**. Vienna, Áustria, 2019. p. 1–4.
- HU, G.; HANNAN, N. B.; TEARO, K.; BASTOS, A.; REILLY, D. Doing while thinking: physical and cognitive engagement and immersion in mixed reality games. In: **Proceedings of the 2016 ACM Conference on Designing Interactive Systems**. Brisbane, Australia: [s. n.], 2016. p. 947–958.
- HUIZINGA, J. **Homo Ludens IIs 86**. Londres, Reino Unido: Routledge, 2014.
- HUNICKE, R.; LEBLANC, M.; ZUBEK, R. Mda: a formal approach to game design and game research. In: **Proceedings of the AAAI Workshop on Challenges in Game AI**. Palo Alto, Estados Unidos: [s. n.], 2004. v. 4, n. 1, p. 1722.
- IACOVIDES, I.; ACZEL, J.; SCANLON, E.; WOODS, W. Making sense of game-play: how can we examine learning and involvement? **Transactions of the Digital Games Research Association**, Pittsburgh, Estados Unidos, v. 1, n. 1, 2013.

IGDB. **Chronos Twins DX (2010)**. Estados Unidos: [s. n.], 2010. Disponível em: <https://www.igdb.com/games/chronos-twins-dx>. Acesso em: 8 ago. 2020.

IJSSELSTEIJN, W.; KORT, Y. D.; POELS, K. The game experience questionnaire. Technische Universiteit Eindhoven, Eindhoven, Holanda, 2013.

İLHAN, A. E. Eye-tracking to enhance usability: A race game. In: SPRINGER. **Proceedings of SAI Intelligent Systems Conference**. London, Reino Unido, 2018. p. 201–214.

INC, M. M. G. **Celeste**. Canadá: [s. n.], 2018. Disponível em: <https://store.steampowered.com/app/504230/Celeste/>. Acesso em: 8 ago. 2020.

JACKSON, S. A. Joy, fun, and flow state in sport. **Emotions in sport**, Washington, Estados Unidos, p. 135–155, 2000.

JUUL, J. The game, the player, the world: looking for a heart of gameness. **PLURAIIS-Revista Multidisciplinar**, Salvador, Brasil, v. 1, n. 2, 2010.

KAYIN. **IWBTG! - Screen Shots!** Estados Unidos: [s. n.], 2007. Disponível em: <https://kayin.moe/iwbtg/screenshots.php>. Acesso em: 8 ago. 2020.

KEPPLINGER, D.; WALLNER, G.; KRIGLSTEIN, S.; LANKES, M. See, feel, move: player behaviour analysis through combined visualization of gaze, emotions, and movement. In: **Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems**. Honolulu, Estados Unidos: [s. n.], 2020. p. 1–14.

KESHAV, S. How to read a paper. **ACM SIGCOMM Computer Communication Review**, New York, Estados Unidos, v. 37, n. 3, p. 83–84, 2007.

KHALIFA, A.; SILVA, F. de M.; TOGELIUS, J. Level design patterns in 2d games. In: IEEE. **2019 IEEE Conference on Games (CoG)**. London, Reino Unido, 2019. p. 1–8.

KIM, J.; JUNG, J.; KIM, S. The relationship of game elements, fun and flow. **Indian Journal of Science and Technology**, Chennai, Índia, v. 8, n. S8, p. 405–411, 2015.

KOHWALTER, T.; CLUA, E.; MURTA, L. Provenance in games. **XI SBGames**, Brasília, Brasil, p. 162–171, 2012.

KOHWALTER, T.; OLIVEIRA, T.; FREIRE, J.; CLUA, E.; MURTA, L. Prov viewer: a graph-based visualization tool for interactive exploration of provenance data. In: SPRINGER. **International Provenance and Annotation Workshop**. McLean, Estados Unidos, 2016. p. 71–82.

KOHWALTER, T. C.; FIGUEIRA, F. M. de A.; SERDEIRO, E. A. de L.; JUNIOR, J. R. da S.; MURTA, L. G. P.; CLUA, E. W. G. Understanding game sessions through provenance. **Entertainment Computing**, Elsevier, Amsterdam, Holanda, v. 27, p. 110–127, 2018.

KOSTER, R. **Theory of fun for game design**. Newton, Estados Unidos: "O'Reilly Media, Inc.", 2013.

KRAFKA, K.; KHOSLA, A.; KELLNHOFER, P.; KANNAN, H.; BHANDARKAR, S.; MATUSIK, W.; TORRALBA, A. Eye tracking for everyone. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. Las Vegas, Estados Unidos: [s. n.], 2016. p. 2176–2184.

- LEE, J. **Optimizing Unity games on Android**: A case study. Estados Unidos: [s. n.], 2015. Disponível em: https://www.gamasutra.com/view/news/255007/Sponsored/Optimizing_Unity_games_on_Android_A_case_study.php. Acesso em: 12 mai. 2019.
- LINEHAN, C.; BELLORD, G.; KIRMAN, B.; MORFORD, Z. H.; ROCHE, B. Learning curves: analysing pace and challenge in four successful puzzle games. In: **ACM. Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play**. New York, Estados Unidos, 2014. p. 181–190.
- MAGDIN, M.; TURČÁNI, M.; HUDEC, L. Evaluating the emotional state of a user using a webcam. **International Journal of Interactive Multimedia & Artificial Intelligence**, La Rioja, Espanha, v. 4, n. 1, 2016.
- MATSUMOTO, D. E. **The Cambridge dictionary of psychology**. [S. l.]: Cambridge University Press, 2009.
- MCKIGHT, P. E.; NAJAB, J. Kruskal-wallis test. **The corsini encyclopedia of psychology**, Wiley Online Library, Manhattan, Estados Unidos, p. 1–1, 2010.
- MIRZA-BABAEI, P.; LONG, S.; FOLEY, E.; MCALLISTER, G. Understanding the contribution of biometrics to games user research. In: **CITeseer. DiGRA conference**. Hilversum, Holanda, 2011.
- MIRZA-BABAEI, P.; MOOSAJEE, N.; DRENKOW, B. Playtesting for indie studios. In: **ACM. Proceedings of the 20th International Academic Mindtrek Conference**. Tampere, Finlândia, 2016. p. 366–374.
- MIRZA-BABAEI, P.; NACKE, L. E.; GREGORY, J.; COLLINS, N.; FITZPATRICK, G. How does it play better?: exploring user testing and biometric storyboards in games user research. In: **ACM. Proceedings of the SIGCHI conference on human factors in computing systems**. Paris, França, 2013. p. 1499–1508.
- MIRZA-BABAEI, P.; ZAMMITTO, V.; NIESENHAUS, J.; SANGIN, M.; NACKE, L. **Games user research: practice, methods, and applications**. New York, Estados Unidos: ACM, 2013.
- MORIN, R.; LÉGER, P.-M.; SENEAL, S.; BASTARACHE-ROBERGE, M.-C.; LEFÈBRVE, M.; FREDETTE, M. The effect of game tutorial: comparison between casual and hardcore gamers. In: **Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts**. Austin, Estados Unidos: [s. n.], 2016. p. 229–237.
- MOZGOVOY, M. Analyzing user behavior data in a mobile tennis game. In: **IEEE. 2018 IEEE Games, Entertainment, Media Conference (GEM)**. Galway, Irlanda, 2018. p. 1–9.
- MUSCAT, A.; DUCKWORTH, J.; WILSON, D. Methods beyond the screen: conducting remote player studies for game design research. In: **DIGITAL GAMES RESEARCH ASSOCIATION DIGRA. DiGRA'19**. Kyoto, Japão, 2019. p. 1–16.
- NACKE, L.; AMBINDER, M.; CANOSSA, A.; MANDRYK, R.; STACH, T. Game metrics and biometrics: the future of player experience research. In: **ALGOMA UNIVERSITY. Conference on Future Play**. Vancouver, Canadá, 2009.
- NACKE, L. E.; LINDLEY, C. A. Affective ludology, flow and immersion in a first-person shooter: measurement of player experience. **arXiv preprint arXiv:1004.0248**, Ithaca, Estados Unidos, 2010.

NACKE, L. E.; MOSER, C.; DRACHEN, A.; MIRZA-BABAEI, P.; ABNEY, A.; ZHENYU, Z. Lightweight games user research for indies and non-profit organizations. In: **Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems**. San Jose, Estados Unidos: [s. n.], 2016. p. 3597–3603.

NELSON, M. J. Game metrics without players: strategies for understanding game artifacts. In: **Workshops at the Seventh Artificial Intelligence and Interactive Digital Entertainment Conference**. Palo Alto, Estados Unidos: [s. n.], 2011.

NUNES, L. Y. O.; LEMOS, D. C. L.; JÚNIOR, R. de C. R.; BEHAR, C. B.; SANTOS, P. P. Pires dos. Análisis psicométrico de la panas en brasil. **Ciencias Psicológicas**, Facultad de Psicología. Universidad Católica del Uruguay, Montevideú, Uruguai, v. 13, n. 1, p. 45–55, 2019.

PAGULAYAN, R. J.; KEEKER, K.; WIXON, D.; ROMERO, R. L.; FULLER, T. User-centered design in games. In: **The human-computer interaction handbook**. Boca Raton, Estados Unidos: CRC Press, 2002. p. 915–938.

PAPOUTSAKI, A.; SANGKLOY, P.; LASKEY, J.; DASKALOVA, N.; HUANG, J.; HAYS, J. Webgazer: scalable webcam eye tracking using user interactions. In: AAI. **Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)**. Palo Alto, Estados Unidos, 2016. p. 3839–3845.

PARALILY, T. **ParaLily Update #4: Improving level design with player analytics and heatmaps**. Estados Unidos: [s. n.], 2018. Disponível em: <http://paralilygame.com/updates/paralily-update-4-improving-level-design-with-player-analytics-and-heatmaps/>. Acesso em: 12 mai. 2019.

PETERSSON, E.; HELGESSON, F. Visual attention in level design for a 3d adventure platform game-analyzing visual cues in a 3d environment. Malmö universitet/Teknik och samhälle, Malmö, Suécia, 2018.

PLASS, J. L.; HOMER, B. D.; KINZER, C.; FRYE, J.; PERLIN, K. Learning mechanics and assessment mechanics for games for learning. **G4LI White Paper**, New York, Estados Unidos, v. 1, p. 2011, 2011.

PLAYDEAD. **LIMBO**. Estados Unidos: [s. n.], 2011. Disponível em: <https://store.steampowered.com/app/48000/LIMBO/>. Acesso em: 7 fev. 2020.

RAJANEN, M.; TAPANI, J. A survey of game usability practices in north american game companies. In: **Designing Digitalization (ISD2018 Proceedings)**. Lund, Suécia: [s. n.], 2018.

Rodrigues, L. A. L.; Brancher, J. D. Improving players' profiles clustering from game data through feature extraction. In: **2018 17th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)**. Foz do Iguaçu, Brasil: [s. n.], 2018. p. 177–186. ISSN 2159-6662.

ROSEWATER, M. Timmy, johnny, and spike revisited. **Making Magic**, Renton, Estados Unidos, v. 18, 2006.

SANGHRAJKA, R.; YOUNG, R. M.; SALISBURY, B.; LANG, E. W. Showrunner: a tool for storyline execution/visualization in 3d game environments. In: SPRINGER. **International Conference on Interactive Digital Storytelling**. Lake City, Estados Unidos, 2019. p. 323–327.

- SANTAELLA, L. **Estética: de Platão a Pierce**. São Paulo, Brasil: Experimento São Paulo, 1994.
- SCOZZI, M. V.; IACOVIDES, I.; LINEHAN, C. A mixed method approach for evaluating and improving the design of learning in puzzle games. In: ACM. **Proceedings of the Annual Symposium on Computer-Human Interaction in Play**. New York, Estados Unidos, 2017. p. 217–228.
- SHELDON, K. M.; KING, L. Why positive psychology is necessary. **American psychologist**, American Psychological Association, Massachusetts, Estados Unidos, v. 56, n. 3, p. 216, 2001.
- SIQUEIRA, E. S.; CASTANHO, C. D.; RODRIGUES, G. N.; JACOBI, R. P. A data analysis of player in world of warcraft using game data mining. In: IEEE. **2017 16th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)**. Curitiba, Brasil, 2017. p. 1–9.
- SMITH, G.; CHA, M.; WHITEHEAD, J. A framework for analysis of 2d platformer levels. In: ACM. **Proceedings of the 2008 ACM SIGGRAPH symposium on Video games**. Los Angeles, Estados Unidos, 2008. p. 75–80.
- SUNDSTEDT, V. A visualisation course in a game development curriculum. In: EUROGRAPHICS ASSOCIATION. **Proceedings of the 37th Annual Conference of the European Association for Computer Graphics: Education Papers**. Goslar, Alemanha, 2016. p. 9–16.
- SUP. **Hardcore Achievement Guide**. Estados Unidos: [s. n.], 2013. Disponível em: <https://steamcommunity.com/sharedfiles/filedetails/?id=152292067>. Acesso em: 6 fev. 2020.
- TALEGAONKAR, I.; JOSHI, K.; VALUNJ, S.; KOHOK, R.; KULKARNI, A. Real time facial expression recognition using deep learning. **ICCP2019**, Elsevier, Tokyo, Japão, 2019.
- TEAM, L. D. **Level Design 101**. Estados Unidos: [s. n.], 2016. Disponível em: <https://www.projectargo.net/blog/dev-blogs/level-design-101>. Acesso em: 8 mai. 2020.
- TECHNOLOGIES, U. **Unity Analytics Overview**. Estados Unidos: [s. n.], 2020. Disponível em: <https://docs.unity3d.com/Manual/UnityAnalyticsOverview.html>. Acesso em: 7 fev. 2020.
- VALENTI, R.; STAIANO, J.; SEBE, N.; GEVERS, T. Webcam-based visual gaze estimation. In: SPRINGER. **International Conference on Image Analysis and Processing**. Vietri sul Mare, Itália, 2009. p. 662–671.
- VALVE. **Half-Life 2**. Estados Unidos: [s. n.], 2004. Disponível em: https://store.steampowered.com/app/220/HalfLife_2/. Acesso em: 7 fev. 2020.
- VALVE. **Portal**. Estados Unidos: [s. n.], 2007. Disponível em: <https://store.steampowered.com/app/400/Portal/>. Acesso em: 13 mai. 2020.
- VOURKOUS, A. **Mobile Indie Spotlight Episode #2: Slyway**. Estados Unidos: [s. n.], 2018. Disponível em: <https://medium.com/@vourkosa/mobile-indie-spotlight-episode-2-slyway-f21d059e352a>. Acesso em: 12 mai. 2019.
- WALLNER, G.; HALABI, N.; MIRZA-BABAEI, P. Aggregated visualization of playtesting data. In: ACM. **Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems**. Glasgow, Reino Unido, 2019. p. 363.

WALLNER, G.; KRIGLSTEIN, S. Plato: a visual analytics system for gameplay data. **Computers & Graphics**, Elsevier, Amsterdam, Holanda, v. 38, p. 341–356, 2014.

WANG, H.; CHEN, W.-W.; SUN, C.-T. Play teaches learning?: a pilot study on how gaming experience influences new game learning. In: **Interactivity and the Future of the Human-Computer Interface**. Pennsylvania, Estados Unidos: IGI Global, 2020. p. 147–168.

WANICK, V.; BITELO, C. Exploring the use of participatory design in game design: a brazilian perspective. **International Journal of Serious Games**, Genova, Itália, v. 7, n. 3, p. 3–20, 2020.

WATSON, D.; CLARK, L. A.; TELLEGEN, A. Development and validation of brief measures of positive and negative affect: the panas scales. **Journal of personality and social psychology**, American Psychological Association, Massachusetts, Estados Unidos, v. 54, n. 6, p. 1063, 1988.

WILKINSON, L.; FRIENDLY, M. The history of the cluster heat map. **The American Statistician**, Taylor & Francis, Riverside, Estados Unidos, v. 63, n. 2, p. 179–184, 2009.

WORKSHOP, A. **Binaries - A Controller-Smashingly Tough Platform Game**. Estados Unidos: [s. n.], 2016. Disponível em: <https://binaries.ant-workshop.com/>. Acesso em: 12 mai. 2019.

YERICH. **What 5000 deaths looks like (Heatmap)**. Estados Unidos: [s. n.], 2014. Disponível em: https://www.reddit.com/r/leagueoflegends/comments/1wfmm9/what_5000_deaths_looks_like_heatmap/. Acesso em: 6 fev. 2020.

YI, J. S. Qndreview: read 100 chi papers in 7 hours. In: ACM. **CHI'14 Extended Abstracts on Human Factors in Computing Systems**. Toronto, Canadá, 2014. p. 805–814.

ZHU, Y. Measuring effective data visualization. In: SPRINGER. **International Symposium on Visual Computing**. Lake Tahoe, Estados Unidos, 2007. p. 652–661.

APÊNDICE A – QUESTIONÁRIO DE DIFICULDADE

Este foi o formulário o qual os participantes preencheram dando notas quanto à dificuldade de cada desafio da primeira versão do jogo *Downside Up*.

Formulário de Avaliação de Level Design

Você foi convidado a participar da sessão de testes de um protótipo de jogo 2D denominado Downside Up.

Este formulário tem a seguinte estrutura:

- 1 - Dados Demográficos
- 2 - Avaliação de dificuldade dos desafios do nível 1
- 3 - Avaliação de dificuldade dos desafios do nível 2
- 4 - Avaliação de dificuldade dos desafios do nível 3
- 5 - Avaliação de dificuldade geral e considerações finais

A duração esperada é de 5 a 10 minutos para o preenchimento. A qualquer momento você poderá cancelar sua participação nesta pesquisa. Os dados coletados serão anônimos e confidenciais.

* Required

Dados Demográficos

Por favor, conte-nos mais sobre você e sua experiência prévia com jogos eletrônicos.

Sexo *

Masculino

Feminino

Other: _____

Faixa etária *

Até 17 anos

De 18 a 20 anos

De 21 a 25 anos

De 26 a 30 anos

A partir de 31 anos

Tempo médio gasto com jogos digitais por semana *

- Até 2 horas por semana
- Mais que 2 horas mas menos que 4 horas
- Mais que 4 horas mas menos que 8 horas
- 8 horas ou mais por semana

Em que aparelho você prefere jogar? *

- Console de mesa (controle)
- Console portátil (controle e/ou touchscreen)
- Computador (teclado e mouse)
- Computador (controle)
- Smartphone (touchscreen)
- Other: _____

Gênero de jogo eletrônico favorito *

- Ação ou Aventura (luta, plataforma, tiro, hack-n'-slash, battle royale)
- Narrativa (escape room, narrativas interativas, visual novel)
- RPG ou Estratégia (por turno, de ação, roguelike, RTS, Moba, tactics)
- Simulação (construção, vida, gerenciamento, veículos, esportes)
- MMO (MMORPG, MMOFPS, MMORTS)
- Other: _____

Quais são seus jogos favoritos? (opcional)

Your answer

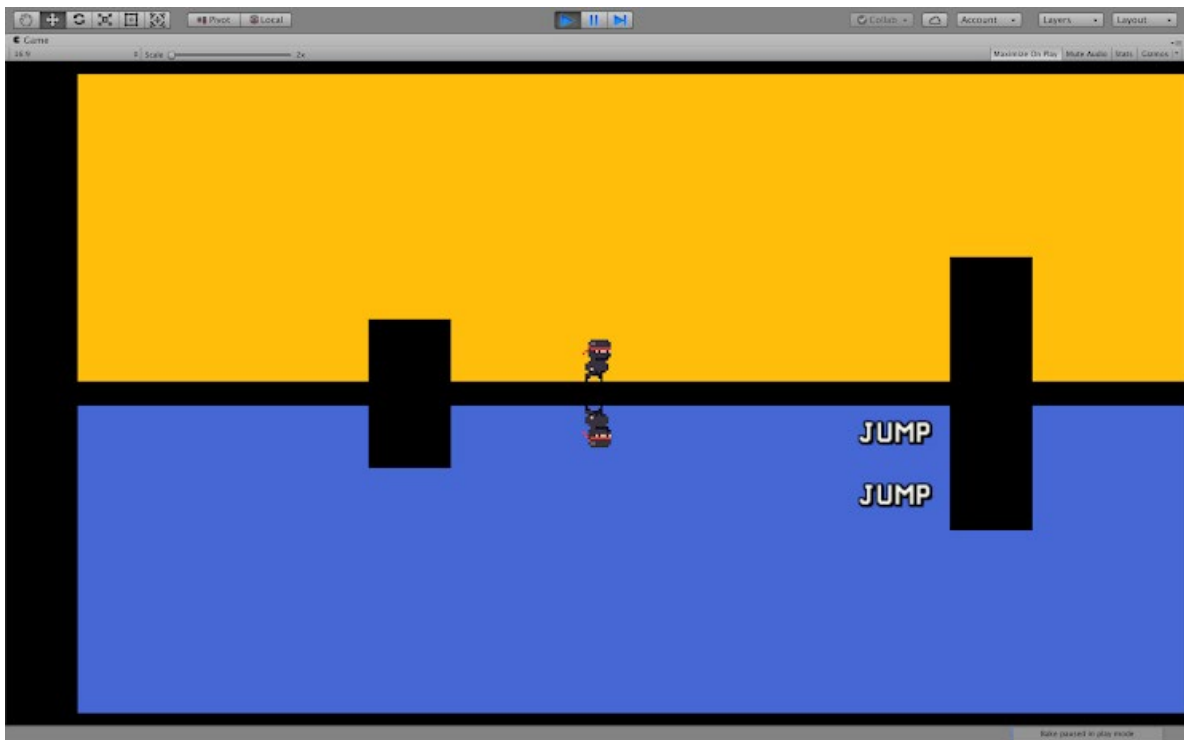
Dificuldade do nível 1

Por favor, julgue a dificuldade numa escala de 1 a 5 de cada desafio presente no nível 1. Com a seguinte padronização:

- 1 - Fácil demais: o desafio ficou aquém das expectativas, não me senti desafiado
- 2 - Fácil: não me senti desafiado, porém já esperava que este desafio seria fácil
- 3 - Na medida certa: o desafio e sua dificuldade parecem adequados à minha expectativa
- 4 - Difícil: me senti desafiado, porém já esperava que este desafio seria difícil
- 5 - Difícil demais: o desafio superou as expectativas, me senti demasiadamente desafiado

Ao final, há espaço para críticas, sugestões e comentários a respeito do nível 1.

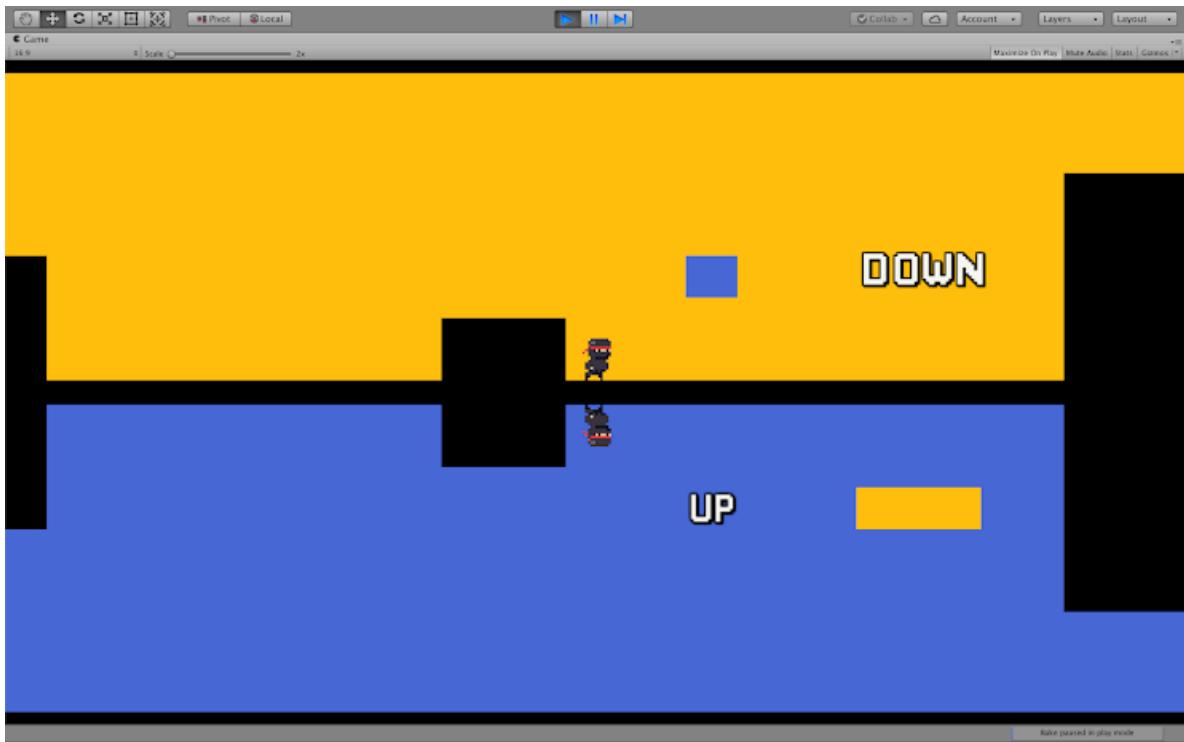
Desafio 1A



Dificuldade do desafio 1A *

Fácil demais 1 2 3 4 5 Difícil demais

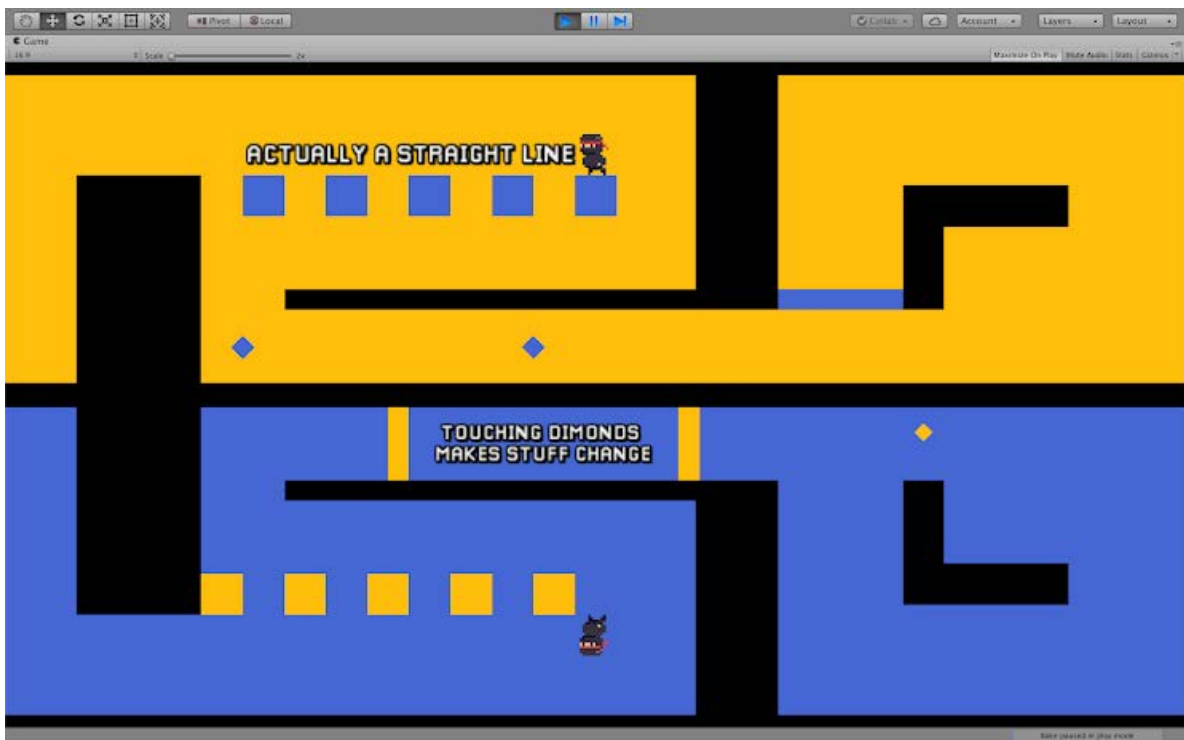
Desafio 1B



Dificuldade do desafio 1B *

Fácil demais 1 2 3 4 5 Dificil demais

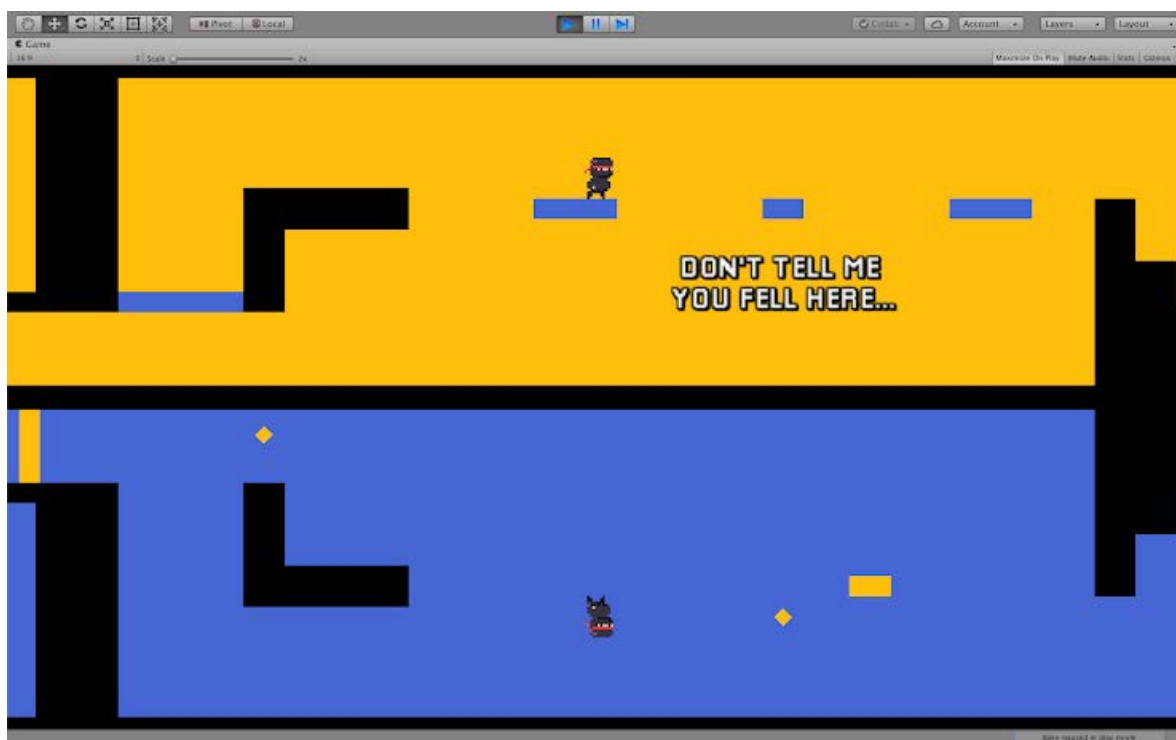
Desafio 1C



Dificuldade do desafio 1C *

Fácil demais 1 2 3 4 5 Dificil demais

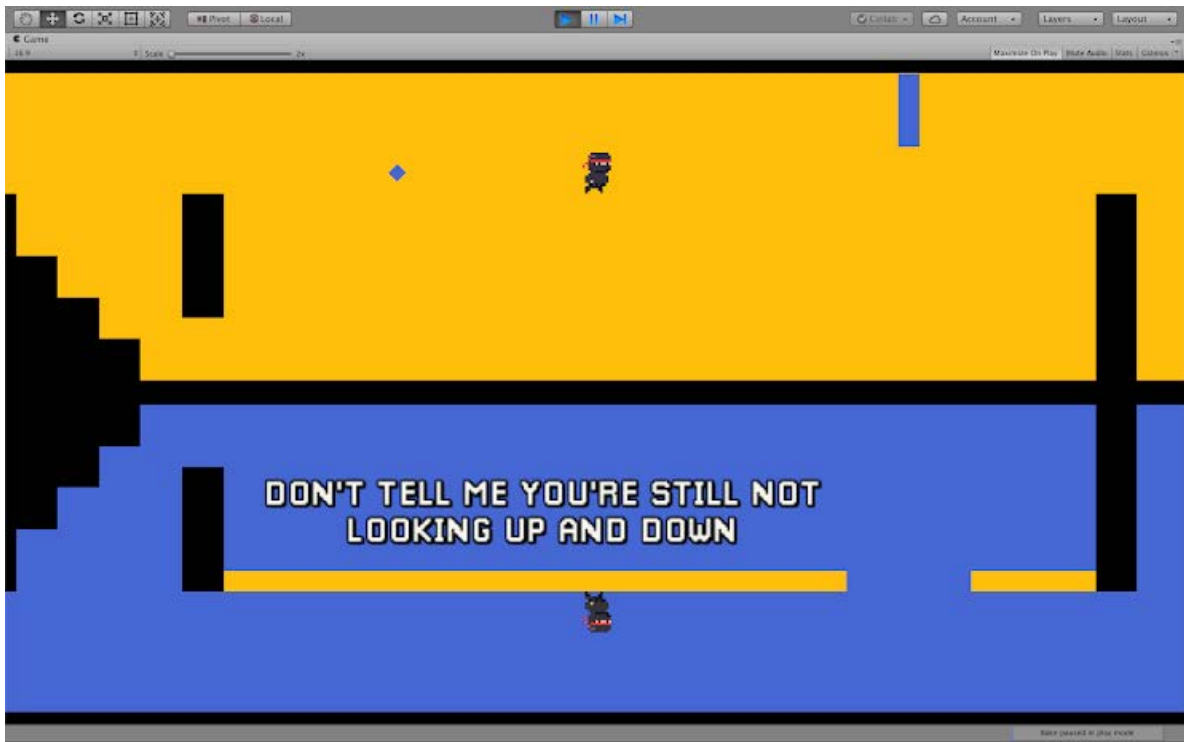
Desafio 1D



Dificuldade do desafio 1D *

Fácil demais 1 2 3 4 5 Dificil demais

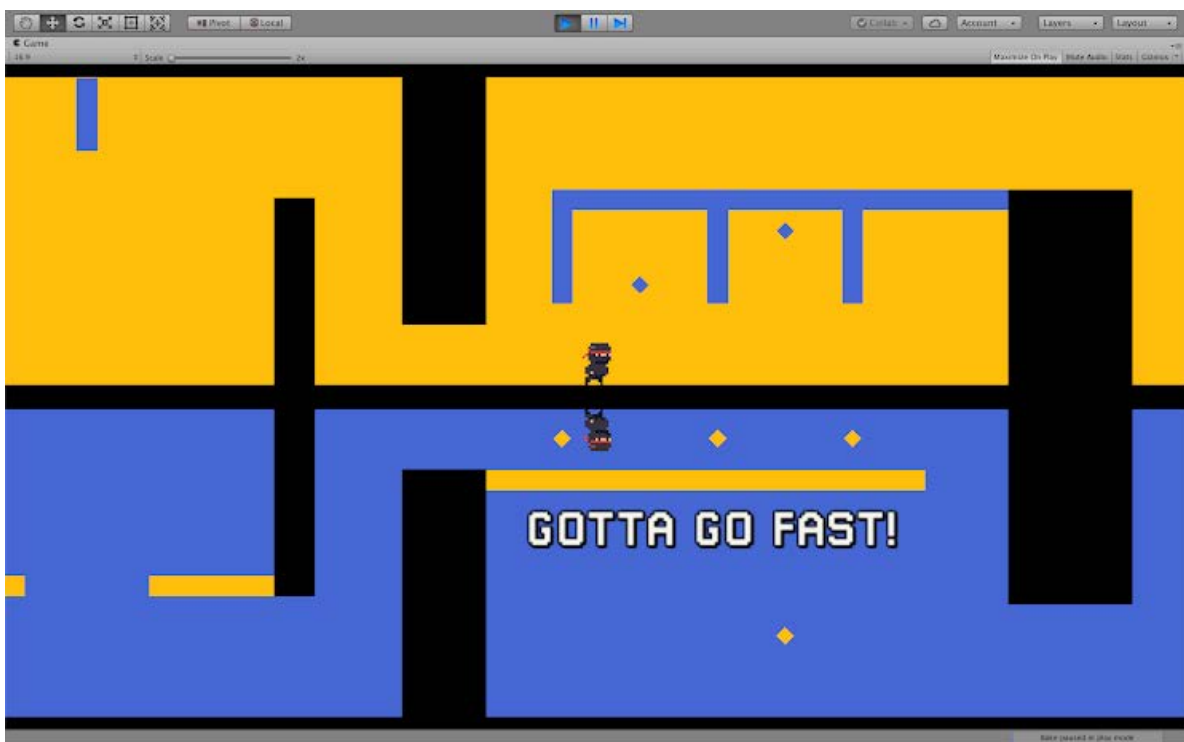
Desafio 1E



Dificuldade do desafio 1E *

Fácil demais 1 2 3 4 5 Dificil demais

Desafio 1F



Dificuldade do desafio 1F *

Fácil demais 1 2 3 4 5 Difícil demais

Comentário final sobre a fase 1 (opcional)

Your answer

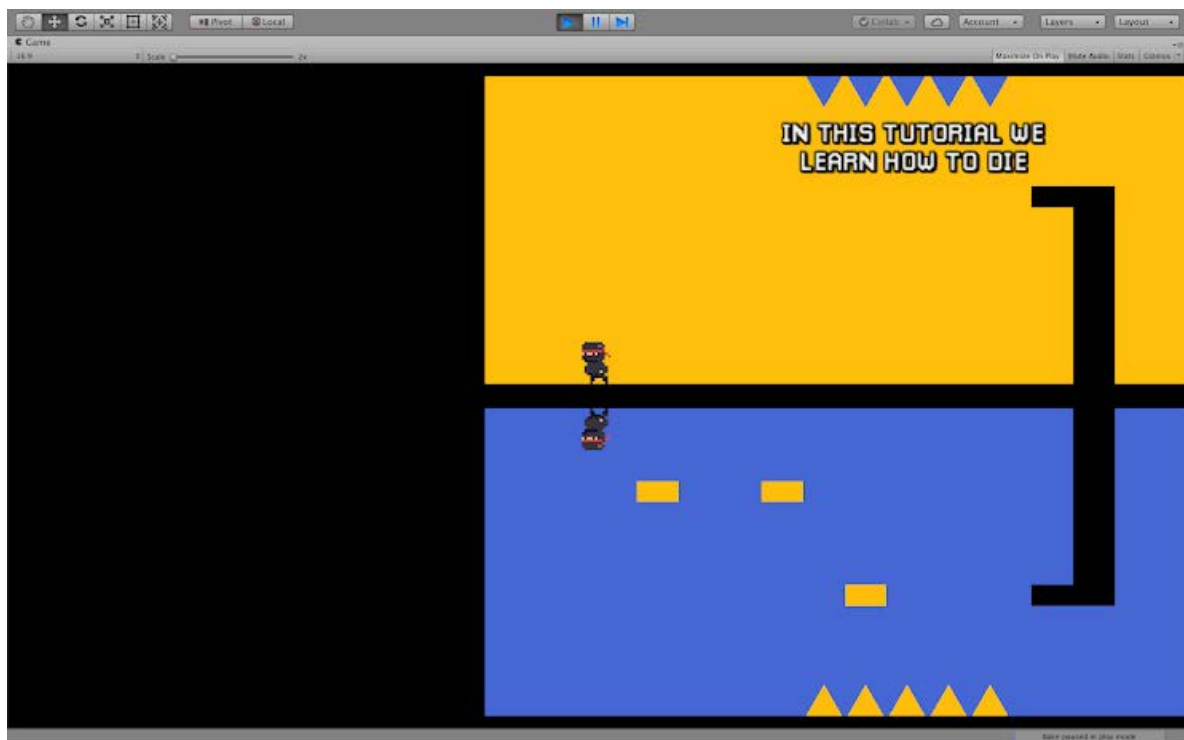
Dificuldade do nível 2

Por favor, julgue a dificuldade numa escala de 1 a 5 de cada desafio presente no nível 2. Com a seguinte padronização:

- 1 - Fácil demais: o desafio ficou aquém das expectativas, não me senti desafiado
- 2 - Fácil: não me senti desafiado, porém já esperava que este desafio seria fácil
- 3 - Na medida certa: o desafio e sua dificuldade parecem adequados à minha expectativa
- 4 - Difícil: me senti desafiado, porém já esperava que este desafio seria difícil
- 5 - Difícil demais: o desafio superou as expectativas, me senti demasiadamente desafiado

Ao final, há espaço para críticas, sugestões e comentários a respeito do nível 2.

Desafio 2A



Dificuldade do desafio 2A *

Fácil demais 1 2 3 4 5 Difícil demais

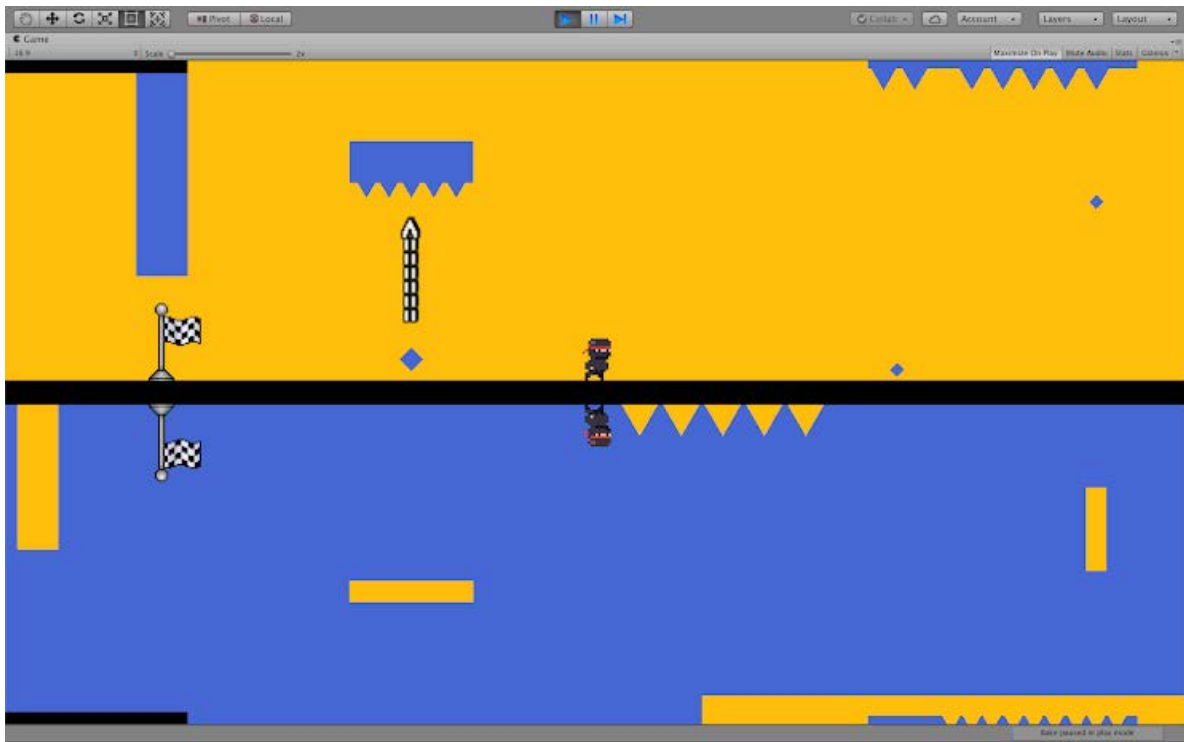
Desafio 2B



Dificuldade do desafio 2B *

Fácil demais 1 2 3 4 5 Difícil demais

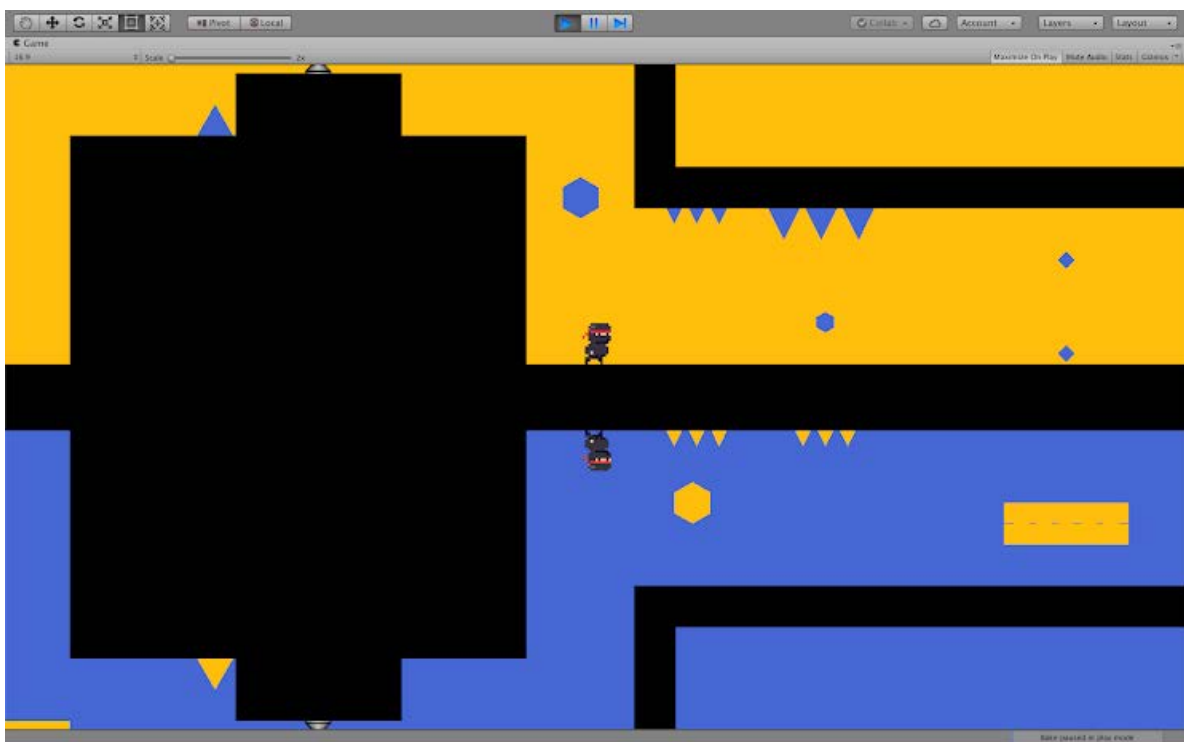
Desafio 2C



Dificuldade do desafio 2C *

Fácil demais 1 2 3 4 5 Difícil demais

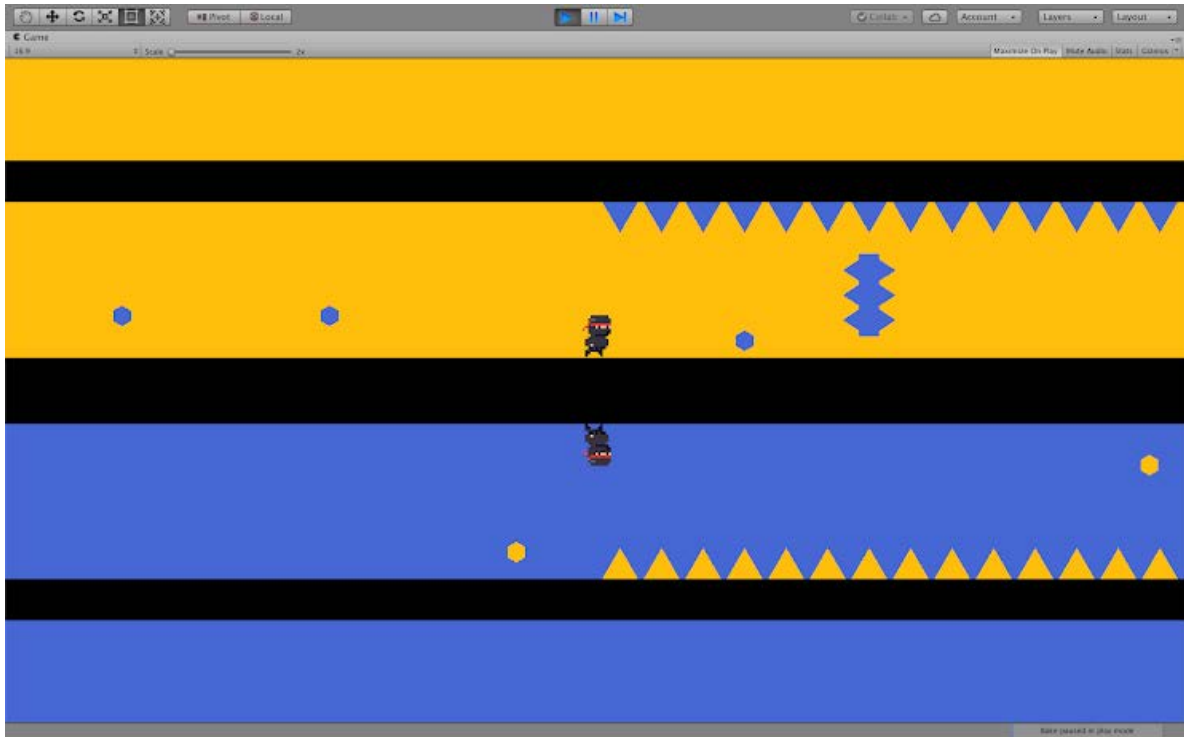
Desafio 2D



Dificuldade do desafio 2D *

Fácil demais 1 2 3 4 5 Difícil demais

Desafio 2E



Dificuldade do desafio 2E *

Fácil demais 1 2 3 4 5 Difícil demais

Comentário final sobre a fase 2 (opcional)

Your answer

Dificuldade do nível 3

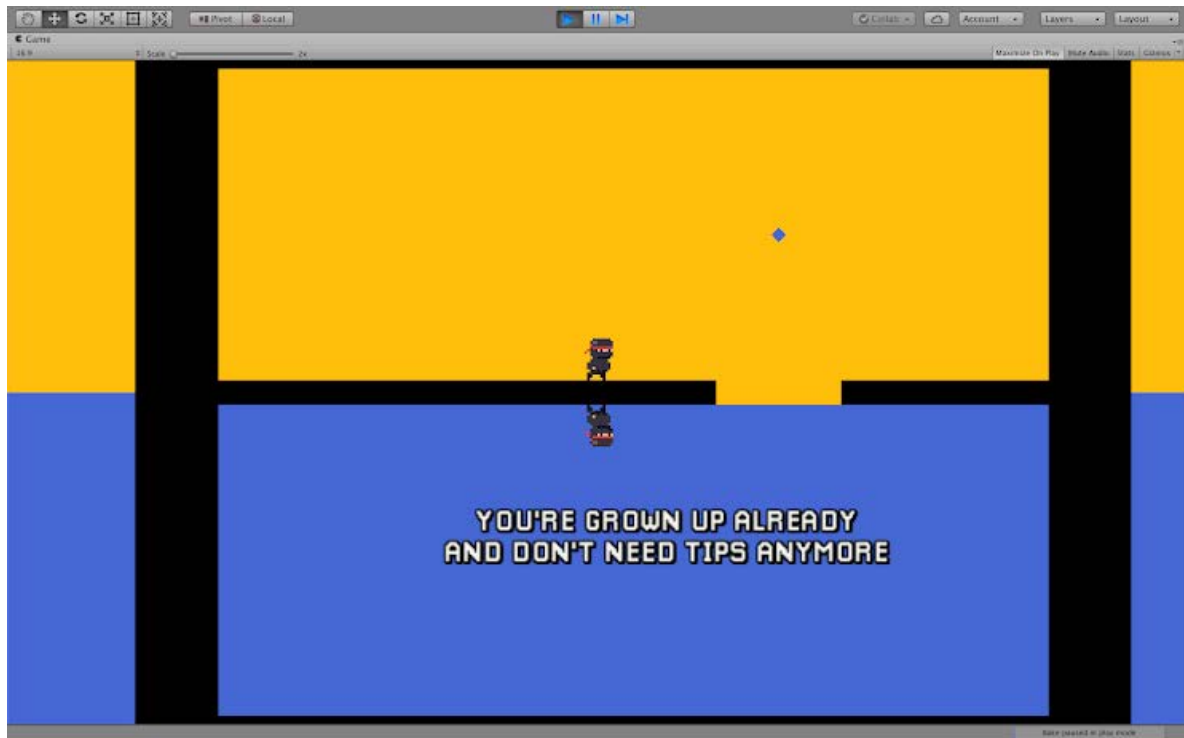
Por favor, julgue a dificuldade numa escala de 1 a 5 de cada desafio presente no nível 3. Com a seguinte

padronização:

- 1 - Fácil demais: o desafio ficou aquém das expectativas, não me senti desafiado
- 2 - Fácil: não me senti desafiado, porém já esperava que este desafio seria fácil
- 3 - Na medida certa: o desafio e sua dificuldade parecem adequados à minha expectativa
- 4 - Difícil: me senti desafiado, porém já esperava que este desafio seria difícil
- 5 - Difícil demais: o desafio superou as expectativas, me senti demasiadamente desafiado

Ao final, há espaço para críticas, sugestões e comentários a respeito do nível 3.

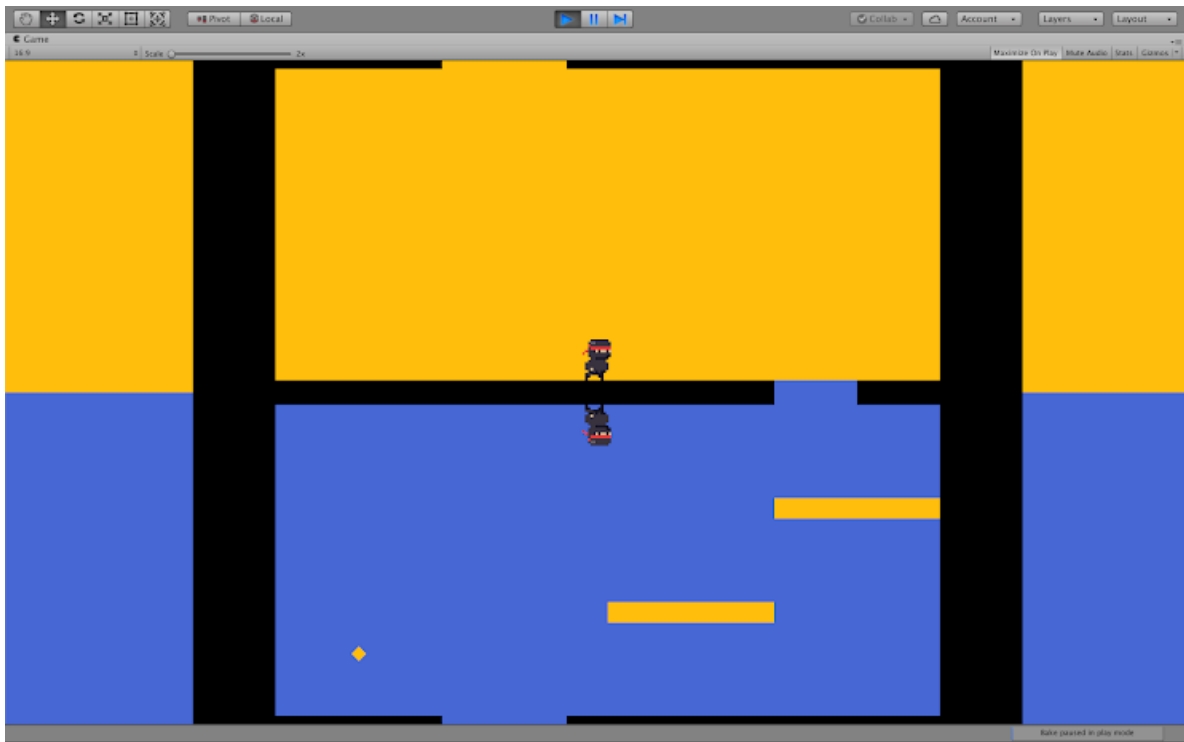
Desafio 3A



Dificuldade do desafio 3A *

Fácil demais 1 2 3 4 5 Difícil demais

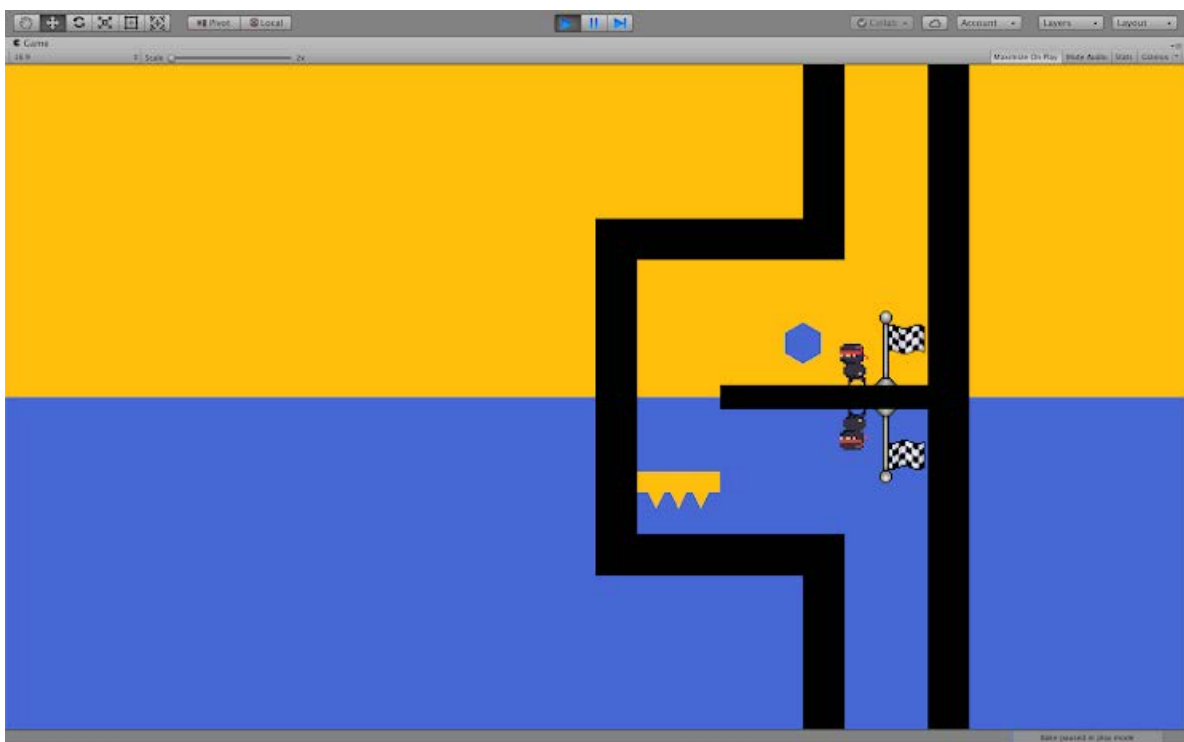
Desafio 3B



Dificuldade do desafio 3B *

Fácil demais 1 2 3 4 5 Difícil demais

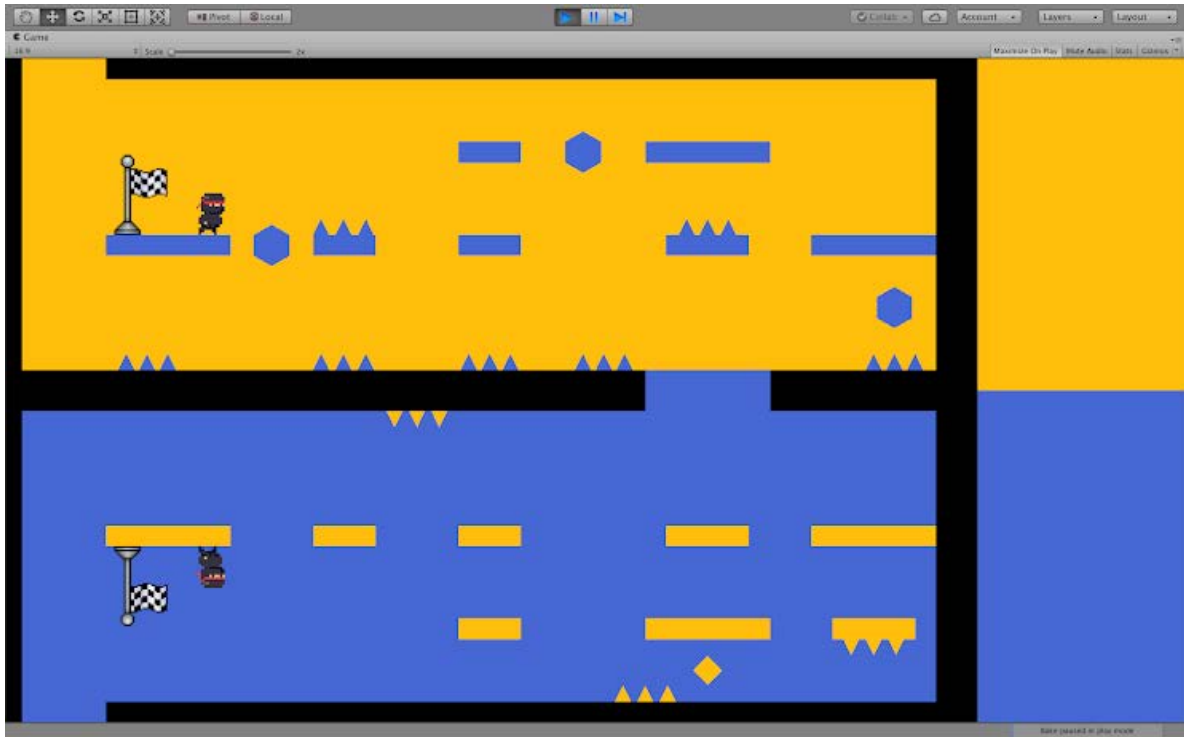
Desafio 3C



Dificuldade do desafio 3C *

Fácil demais 1 2 3 4 5 Difícil demais

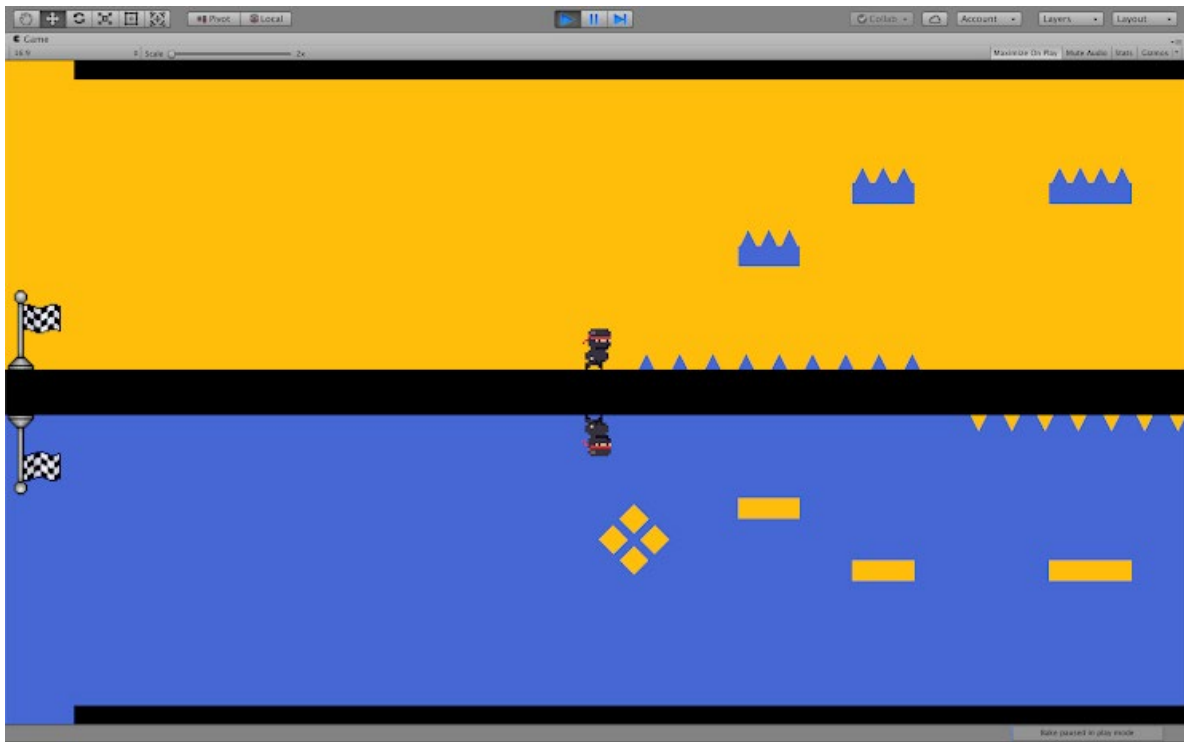
Desafio 3D



Dificuldade do desafio 3D *

Fácil demais 1 2 3 4 5 Difícil demais

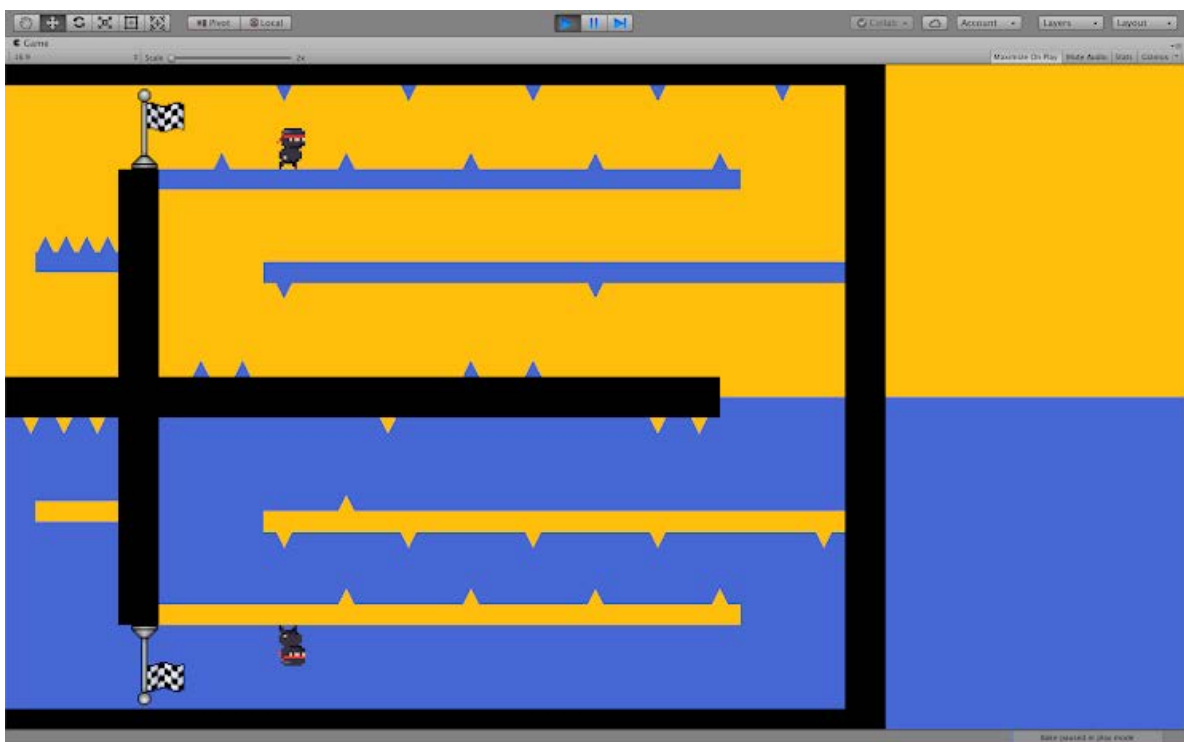
Desafio 3E



Dificuldade do desafio 3E *

Fácil demais 1 2 3 4 5 Difícil demais

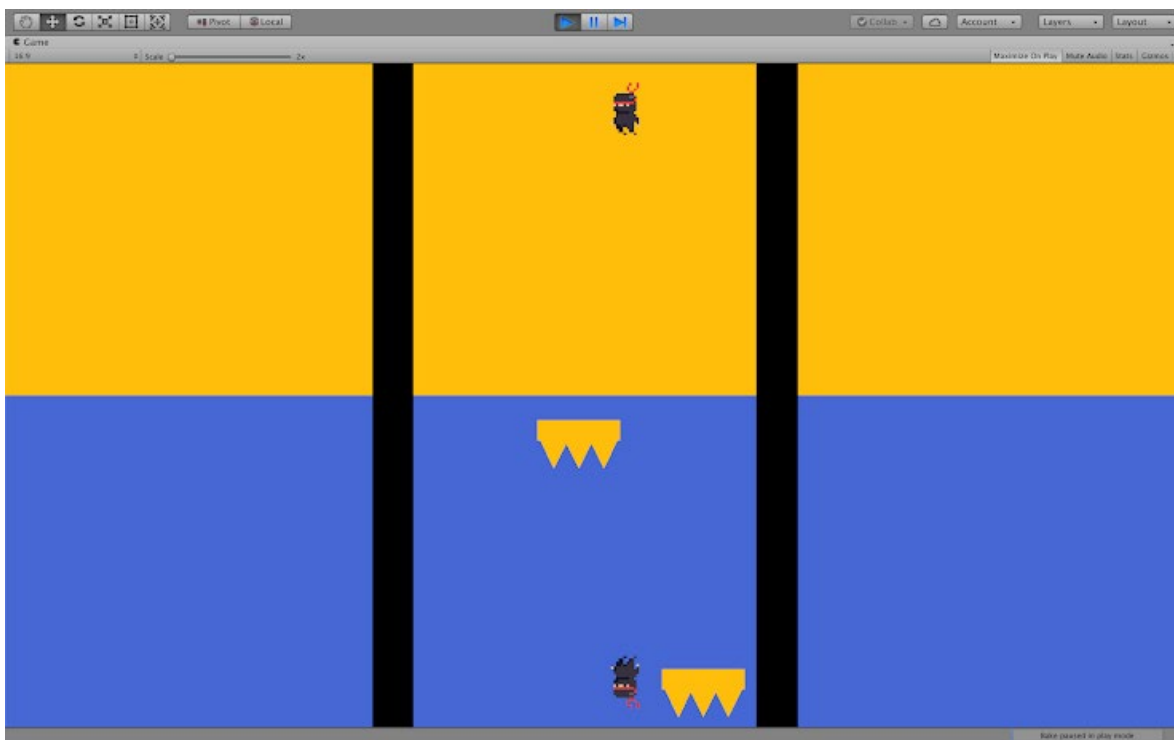
Desafio 3F



Dificuldade do desafio 3F *

Fácil demais 1 2 3 4 5 Difícil demais

Desafio 3G



Dificuldade do desafio 3G *

Fácil demais 1 2 3 4 5 Difícil demais

Comentário final sobre a fase 3 (opcional)

Your answer

Encerramento

Por favor, conte-nos mais sobre sua experiência geral testando este jogo.

Nível de satisfação com o protótipo *

Nada satisfeito 1 2 3 4 5 Completamente satisfeito

Justificativa (opcional)

Your answer

Nível de dificuldade do protótipo *

Difícil demais 1 2 3 4 5 Fácil demais

Justificativa (opcional)

Your answer

Você recomendaria este jogo (ou jogo similar) para seus amigos? *

Sim

Não

Talvez

Justificativa (opcional)

Your answer

Avalie sua experiência com o jogo *

- Não me diverti nem um pouco jogando, a maioria dos desafios não me agradou
- Não me diverti na maioria dos momentos, mas gostei de alguns desafios
- Os momentos divertidos e não divertidos foram aproximadamente iguais para mim
- Me diverti na maioria dos momentos, mas não gostei de alguns desafios
- Me diverti muito jogando, a maioria dos desafios me agradou

Justificativa (opcional)

Your answer

Comentários finais sobre o jogo como um todo (opcional)

Your answer

APÊNDICE B – QUESTIONÁRIOS DE EXPERIÊNCIA

Este foi o formulário o qual os participantes preencheram dando notas quanto à dificuldade de cada desafio da primeira versão do jogo *Downside Up*.

Testes de Avaliação de Experiência

Você foi convidado a participar da sessão de testes de um protótipo de jogo 2D denominado Downside Up.

Este formulário tem a seguinte estrutura:

- 1 - Dados Demográficos
- 2 - Game Experience Questionnaire (GEQ)
- 3 - Positive And Negative Affect Scale (PANAS)

A duração esperada é de 5 a 10 minutos para o preenchimento. A qualquer momento você poderá cancelar sua participação nesta pesquisa. Os dados coletados serão anônimos e confidenciais.

* Required

1 - Dados Demográficos

Por favor, conte-nos mais sobre você e sua experiência prévia com jogos eletrônicos.

Sexo *

Masculino

Feminino

Other: _____

Faixa etária *

Até 17 anos

De 18 a 20 anos

De 21 a 25 anos

De 26 a 30 anos

A partir de 31 anos

Tempo médio gasto com jogos digitais por semana *

- Até 2 horas por semana
- Mais que 2 horas mas menos que 4 horas
- Mais que 4 horas mas menos que 8 horas
- 8 horas ou mais por semana

Em que aparelho você prefere jogar? *

- Console de mesa (controle)
- Console portátil (controle e/ou touchscreen)
- Computador (teclado e mouse)
- Computador (controle)
- Smartphone (touchscreen)
- Other: _____

Gênero de jogo eletrônico favorito *

- Ação ou Aventura (luta, plataforma, tiro, hack-n'-slash, battle royale)
- Narrativa (escape room, narrativas interativas, visual novel)
- RPG ou Estratégia (por turno, de ação, roguelike, RTS, Moba, tactics)
- Simulação (construção, vida, gerenciamento, veículos, esportes)
- MMO (MMORPG, MMOFPS, MMORTS)
- Other: _____

Você conseguiu passar de todas as fases do protótipo? *

- Sim

2 - GEQ (Filtrado)

Este teste visa avaliar como foi sua experiência jogando o protótipo do jogo Downside Up.

Módulo Principal

Seja o mais sincero possível e marque as opções conforme você concorda ou discorda das 9 afirmações a seguir

Indique o quanto você concorda com as afirmações a seguir sobre a sua experiência enquanto jogava: *

| | 0 - Muito pouco | 1 - Um pouco | 2 - Moderadamente | 3 - Bastante | 4 - Extremamente |
|----------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Me senti satisfeito | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Achei divertido | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Estive feliz | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Fiquei mal humorado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Pensei em outras coisas | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Fiquei cansado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. Me senti bem | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. Fiquei entediado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 20. Gostei do jogo | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Módulo Pós-Game

Seja o mais sincero possível e marque as opções conforme você concorda ou discorda das 12 afirmações a seguir

Indique o quanto você concorda com as afirmações a seguir sobre a sua experiência depois de jogar: *

| | 0 - Muito pouco | 1 - Um pouco | 2 - Moderadamente | 3 - Bastante | 4 - Extremamente |
|------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Me senti renovado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Me senti mal | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Me senti culpado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Me senti vitorioso | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Me senti perdendo tempo | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Me senti estimulado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Me senti satisfeito | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. Senti que poderia ter feito algo mais útil | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. Me senti poderoso | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. Me senti arrependido | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. Me senti envergonhado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. Me senti orgulhoso | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

4 - PANAS

Este teste visa avaliar seus sentimentos e emoções enquanto jogava. Seja o mais sincero possível e atribua um valor de intensidade a cada um dos 20 sentimentos ou emoções listados a seguir. Os itens devem ser marcados segundo esta lógica:

- 0 - Muito pouco ou nada
- 1 - Um pouco
- 2 - Moderadamente
- 3 - Bastante
- 4 - Extremamente

*

| | 0 | 1 | 2 | 3 | 4 |
|--------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Interessado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Angustiado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Empolgado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Aborrecido | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Forte | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Culpado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Assustado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Hostil | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Entusiasmado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Orgulhoso | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Irritado | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Alerta | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Envergonhado



Inspirado



Nervoso



Determinado



Atento



Agitado



Ativo



Apreensivo



APÊNDICE C – TERMO DE CONSENTIMENTO

Este foi o formulário que os testadores dos protótipos do jogo *Downside Up* assinaram para atestar que desejaram participar por livre e espontânea vontade.

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Você está sendo convidado para participar da pesquisa intitulada “Uma Abordagem para o Aperfeiçoamento de Level Design de Jogo de Plataforma 2D Usando Visualização e Técnicas de Baixo Custo”. Você não deve participar contra a sua vontade. Leia atentamente as informações abaixo e faça qualquer pergunta que desejar, para que todos os procedimentos desta pesquisa sejam esclarecidos.

O estudo visa coletar dados sobre qualidade de experiência com o objetivo de aperfeiçoar um jogo eletrônico. O projeto está sendo coordenado por Arthur Silva Bastos e faz parte da pesquisa da dissertação do curso de Mestrado em Ciências da Computação na Universidade Federal do Ceará. A sua participação nesta pesquisa consistirá em quatro etapas sequenciais. O pesquisador estará presente observando e tomando notas, além de acompanhá-lo(a) ou orientá-lo(a). A duração varia entre 30 e 60 minutos, conforme a sua performance no jogo. As etapas serão as seguintes:

1. Preencher um formulário demográfico e de perfil de jogador;
2. Jogar, em um computador fornecido pelo pesquisador, um protótipo com 3 fases de um jogo. Os movimentos na tela (posição X,Y) do avatar no jogo serão rastreados, bem como em que locais do ambiente de jogo ocorreram mortes. O pesquisador poderá solicitar captura de olhar por meio da webcam do computador a fim de identificar regiões de foco de atenção. Nenhuma imagem ou vídeo será capturada nesta etapa;
3. Preencher um formulário de experiência de jogo ("game experience questionnaire", ou GEQ), com 31 questões;
4. Participar de uma entrevista informal de até 2 minutos sobre sua opinião como jogador e o que você acha que é necessário melhorar no jogo;

Todos os dados coletados serão confidenciais. Em caso de publicação, os resultados serão apresentados de forma que seu anonimato seja garantido. Você poderá solicitar os dados a qualquer momento. Caso você autorize, fotos serão tiradas a fim de ilustrar o processo de pesquisa. A participação é não remunerada e voluntária. O benefício da sua participação será jogar um jogo potencialmente divertido.

A participação nesta pesquisa oferece riscos mínimos ao participante, sendo estes os mesmos decorrentes do ato de jogar: possível má-postura por conta da distração com o jogo, incômodo visual ou nas mãos caso a sessão de jogo ultrapasse a duração de uma hora. Participantes tímidos poderão se sentir incomodados com o aplicador do teste presente observando o jogo. A qualquer momento você pode desistir da sua participação desse projeto e retirar o seu consentimento sem qualquer tipo de prejuízo. Para aceitar participar da pesquisa, o TCLE precisa ser assinado por você e pelo pesquisador responsável.

Nome: Arthur Silva Bastos
Instituição: Universidade Federal do Ceará
Endereço: Rua Rangel Pestana, 950, Casa 3
Telefones para contato: (85) 992075231

ATENÇÃO: Se você tiver alguma consideração ou dúvida, sobre a sua participação na pesquisa, entre em contato com o Comitê de Ética em Pesquisa da UFC/PROPESQ – Rua Coronel Nunes de Melo, 1000 - Rodolfo Teófilo, fone: 3366-8346/44. (Horário: 08:00-12:00 horas de segunda a sexta-feira).
O CEP/UFC/PROPESQ é a instância da Universidade Federal do Ceará responsável pela avaliação e acompanhamento dos aspectos éticos de todas as pesquisas envolvendo seres humanos.

Eu, _____, ____ anos, RG: _____, declaro que é de livre e espontânea vontade que participarei desta pesquisa. Declaro que li cuidadosamente este Termo de Consentimento Livre e Esclarecido e que, após sua leitura, tive a oportunidade de fazer perguntas sobre o seu conteúdo, como também sobre a pesquisa, e recebi explicações que responderam por completo minhas dúvidas. E declaro, ainda, estar recebendo uma via assinada deste termo.

Fortaleza, ____/____/____

| | | |
|--------------------------------------------------------|------|------------|
| Nome do participante da pesquisa | Data | Assinatura |
| Nome do pesquisador principal | Data | Assinatura |
| Nome do Responsável legal/testemunha (se aplicável) | Data | Assinatura |
| Nome do profissional que aplicou o TCLE | Data | Assinatura |

ANEXO A – GAME EXPERIENCE QUESTIONNAIRE

Versões originais dos módulos do GEQ adaptados neste trabalho.

2. Game Experience Questionnaire – Core Module

Please indicate how you felt while playing the game for each of the items, on the following scale:

| not at all | slightly | moderately | fairly | extremely |
|------------|----------|------------|--------|-----------|
| 0 | 1 | 2 | 3 | 4 |
| < > | < > | < > | < > | < > |

- 1 I felt content
- 2 I felt skilful
- 3 I was interested in the game's story
- 4 I thought it was fun
- 5 I was fully occupied with the game
- 6 I felt happy
- 7 It gave me a bad mood
- 8 I thought about other things
- 9 I found it tiresome
- 10 I felt competent
- 11 I thought it was hard
- 12 It was aesthetically pleasing
- 13 I forgot everything around me
- 14 I felt good
- 15 I was good at it
- 16 I felt bored
- 17 I felt successful
- 18 I felt imaginative
- 19 I felt that I could explore things
- 20 I enjoyed it
- 21 I was fast at reaching the game's targets
- 22 I felt annoyed
- 23 I felt pressured
- 24 I felt irritable
- 25 I lost track of time
- 26 I felt challenged
- 27 I found it impressive
- 28 I was deeply concentrated in the game
- 29 I felt frustrated
- 30 It felt like a rich experience
- 31 I lost connection with the outside world
- 32 I felt time pressure

33 I had to put a lot of effort into it

5. GEQ – post-game module

Please indicate how you felt after you finished playing the game for each of the items, on the following scale:

| not at all | slightly | moderately | fairly | Extremely |
|------------|----------|------------|--------|-----------|
| 0 | 1 | 2 | 3 | 4 |
| < > | < > | < > | < > | < > |

- 1 I felt revived
- 2 I felt bad
- 3 I found it hard to get back to reality
- 4 I felt guilty
- 5 It felt like a victory
- 6 I found it a waste of time
- 7 I felt energised
- 8 I felt satisfied
- 9 I felt disoriented
- 10 I felt exhausted
- 11 I felt that I could have done more useful things
- 12 I felt powerful
- 13 I felt weary
- 14 I felt regret
- 15 I felt ashamed
- 16 I felt proud
- 17 I had a sense that I had returned from a journey

ANEXO B – POSITIVE AND NEGATIVE AFFECT SCALE

Versão original do PANAS adaptada neste trabalho.

- "Blue Monday" phenomenon. *Journal of Personality and Social Psychology*, 49, 129-134.
- Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681-706). Hillsdale, NJ: Erlbaum.
- Tessler, R., & Mechanic, D. (1978). Psychological distress and perceived health status. *Journal of Health and Social Behavior*, 19, 254-262.
- Warr, P. (1978). A study of psychological well-being. *British Journal of Psychology*, 69, 111-121.
- Warr, P., Barter, J., & Brownbridge, G. (1983). On the independence of positive and negative affect. *Journal of Personality and Social Psychology*, 44, 644-651.
- Watson, D. (in press). The vicissitudes of mood measurement: Effects of varying descriptors, time frames, and response formats on measures of Positive and Negative Affect. *Journal of Personality and Social Psychology*.
- Watson, D. (1988). Intraindividual and interindividual analyses of Positive and Negative Affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology*, 54, 1020-1030.
- Watson, D., & Clark, L. A. (1984). Negative Affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96, 465-490.
- Watson, D., Clark, L. A., & Carey, G. (in press). Positive and Negative Affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*.
- Watson, D., Clark, L. A., & Tellegen, A. (1984). Cross-cultural convergence in the structure of mood: A Japanese replication and a comparison with U.S. findings. *Journal of Personality and Social Psychology*, 47, 127-144.
- Watson, D., & Pennebaker, J. W. (in press). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review*.
- Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98, 219-235.
- Wills, T. A. (1986). Stress and coping in early adolescence: Relationships to substance use in urban school samples. *Health Psychology*, 5, 503-529.
- Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology*, 43, 111-122.

Appendix

The PANAS

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent [INSERT APPROPRIATE TIME INSTRUCTIONS HERE]. Use the following scale to record your answers.

| 1 | 2 | 3 | 4 | 5 |
|--------------------------------|--------------------|------------|------------------|-----------|
| very slightly or not at all | a little | moderately | quite a bit | extremely |
| | _____ interested | | _____ irritable | |
| | _____ distressed | | _____ alert | |
| | _____ excited | | _____ ashamed | |
| | _____ upset | | _____ inspired | |
| | _____ strong | | _____ nervous | |
| | _____ guilty | | _____ determined | |
| | _____ scared | | _____ attentive | |
| | _____ hostile | | _____ jittery | |
| | _____ enthusiastic | | _____ active | |
| | _____ proud | | _____ afraid | |

We have used PANAS with the following time instructions:

| | |
|----------------|---------------------------------------------------------------------|
| Moment | (you feel this way right now, that is, at the present moment) |
| Today | (you have felt this way today) |
| Past few days | (you have felt this way during the past few days) |
| Week | (you have felt this way during the past week) |
| Past few weeks | (you have felt this way during the past few weeks) |
| Year | (you have felt this way during the past year) |
| General | (you generally feel this way, that is, how you feel on the average) |

Received May 10, 1987
Revision received September 14, 1987
Accepted November 11, 1987 ■