



**PODER EXECUTIVO
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**



AVALIAÇÃO DA PLAYABILITY DE JOGOS DIGITAIS COM AVALIADORES NÃO ESPECIALISTAS ATRAVÉS DE CONJUNTOS HEURÍSTICOS CUSTOMIZÁVEIS

FELIPE SONNTAG MANZONI

Manaus
2021

FELIPE SONNTAG MANZONI

**AVALIAÇÃO DA PLAYABILITY DE JOGOS
DIGITAIS COM AVALIADORES NÃO
ESPECIALISTAS ATRAVÉS DE
CONJUNTOS HEURÍSTICOS
CUSTOMIZÁVEIS**

Dissertação de Mestrado submetida ao corpo docente do Programa de Pós-Graduação em Informática (PPGI) da Universidade Federal do Amazonas (UFAM) como um dos requisitos para alcançar o grau de Mestre em Informática.

**ORIENTADORA:
PROF^A. DRA. TAYANA UCHÔA CONTE**

Manaus
2021

FELIPE SONNTAG MANZONI

**DIGITAL GAMES PLAYABILITY
ASSESSMENT THROUGH CUSTOMIZABLE
HEURISTIC SETS WITH NON-EXPERT
EVALUATORS**

Master's thesis presented to the faculty members of the Informatics Postgraduate Program (*Programa de Pós-Graduação em Informática*) - PPGI, at the Federal University of Amazonas (*Universidade Federal do Amazonas*) - UFAM as one of the requirements for achieving the Master's degree in Computer Science.

ADVISOR:
PhD. PROFESSOR TAYANA UCHÔA CONTE

Manaus
2021

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

M296a Manzoni, Felipe Sonntag
Avaliação da playability de jogos digitais com avaliadores não
especialistas através de conjuntos heurísticos customizáveis /
Felipe Sonntag Manzoni . 2021
186 f.: il. color; 31 cm.

Orientadora: Tayana Uchôa Conte
Dissertação (Mestrado em Informática) - Universidade Federal do
Amazonas.

1. Playability. 2. Avaliação Heurística. 3. Conjuntos Heurísticos. 4.
Avaliação da Jogabilidade. 5. Jogos Digitais. I. Conte, Tayana
Uchôa. II. Universidade Federal do Amazonas III. Título



PODER EXECUTIVO
MINISTÉRIO DA EDUCAÇÃO
INSTITUTO DE COMPUTAÇÃO



PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

FOLHA DE APROVAÇÃO

"Avaliação da Playability de Jogos Digitais com Avaliadores Não Especialistas através de Conjuntos Heurísticos Customizáveis"

FELIPE SONNTAG MANZONI

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

Profa. Tayana Uchôa Conte - PRESIDENTE

Profa. Carla Ilane Moreira Bezerra - MEMBRO EXTERNO

Prof. Breno Bernard Nicolau de França - MEMBRO EXTERNO

Manaus, 26 de Agosto de 2021

Para toda a minha família, agregados e a pessoa de minha vida,
Mesmo para todos aqueles com quem eu não posso estar em todos os momentos,
Por todo o amor, suporte incondicional e incentivo que me forneceram,
Sem o qual não seria capaz de chegar a este momento.

Para todos os meus professores,
Por todo o tempo gasto me ensinando qual era o caminho certo a seguir e como chegar aqui,
Nós não desistiremos, pois ensinar é uma dádiva que é capaz de mudar as pessoas e o mundo.

AGRADECIMENTOS

A minha mãe, Miriam Sonntag Manzoni, meu pai, Eduardo Manzoni, meu irmão, Fernando S. Manzoni, minha cunhada, Anne Kellen R. G. Manzoni, e minha irmã, Tábita S. Manzoni, por todo o suporte e encorajamento para seguir em frente durante os dois processos, minha graduação e por todo o processo do meu Mestrado. Todos vocês acreditaram em mim até mesmo quando não acreditei e me encorajaram a ir atrás dos meus objetivos, mesmo sabendo que o caminho a minha frente não seria nem um pouco fácil de percorrer. Obrigado por sempre me incentivarem a seguir em frente, compartilharem as minhas alegrias e conquistas pessoais, assim como por ainda estarem ao meu lado. Devo esse grau de mestre a todos vocês, sem vocês nunca estaria aqui. Da mesma forma, agradeço a todos os meus familiares, principalmente as minhas Avós: Iris e Eneiva; assim como aos meus Tios: Luiz, Rosa, Fábio, Christie, Osmar, Sônia, Xande, Ester, Dedé e Jaques; também aos meus primos e amigos deste lugar que tanto amo: Rodrigo, Letícia, Matheus, Bianca, Júlia, Natália, Rafael, Natanael e Walter; que durante este trajeto me apoiaram nas minhas viagens, estadias, um café com brincadeiras depois de um dia cansativo ou apenas com uma palavra amiga. Amo todos vocês do fundo do meu coração!

A minha namorada, Caroline Amâncio de Oliveira, por todo o amor que você me ensinou a ter por você e pelos outros. Obrigado por sempre me compreender, suportar-me em qualquer decisão e por sua companhia durante este processo. Cada palavra de incentivo, dias de choro, raiva e tristeza que você passou ao meu lado me fizeram uma pessoa melhor, obrigado por sempre estar ali quando precisei; eu te amo.

A minha querida orientadora, Tayana Conte, que suportou todas as minhas brincadeiras através dos anos, e que viu em mim alguém disposto a trabalhar duro e disposto a aprender a como ser um pesquisador de verdade. Obrigado por sempre ter acreditado em mim e no meu trabalho mesmo quando não acreditei em mim mesmo. É por causa de você Tay que me tornei o pesquisador que hoje eu sou. Obrigado por sempre me encorajar, me ensinar, por ter paciência, e por me guiar no caminho correto. Obrigado por todo o valioso conhecimento que você me ensinou e por todas as oportunidades que você me deu para que eu crescesse profissionalmente e pessoalmente. Eu te agradeço e devo a você tudo que sou hoje, eu sempre serei grato por tudo isso, e terei orgulho de dizer que fui seu aluno.

Para Simone D. J. Barbosa e Milene S. Silveira, por terem concordado em participar de alguns processos desenvolvidos durante este mestrado. Obrigado por todo o tempo investido revisando artigos, participando nas discussões do processo metodológico e por terem aberto as portas de suas universidades no processo de execução dos experimentos. Vocês duas são modelos para mim, com as quais um dia espero poder ter a chance de trabalhar novamente.

Para Bruna M. Ferreira, por ter aceitado o desafio de ajudar a Tay a me guiar por alguns dos processos em meus primeiros passos na pesquisa enquanto eu estava na graduação. Sem a sua co-orientação, não teria alcançado alguns dos meus maiores objetivos e realizações.

Para Altigran Soares, por sempre ter acreditado em mim e ter me ajudado a entender o que eu enfrentaria neste processo assim como por ter me suportado e torcido por mim. Admiro o professor que o senhor é, e sempre serei grato por ter tido a chance de trabalharmos juntos, e espero que nos encontremos novamente no futuro e possamos colaborar novamente.

Para Bruno Freitas Gadelha, por todos os valiosos conselhos em cada reunião de grupo e por sempre deixar o clima mais leve com suas brincadeiras. Obrigado por esse conhecimento compartilhado e por sempre nos suportar em cada passo!

Especialmente para Leandro S. G. de Carvalho e José Luiz de Souza Pio, por terem me guiado em meus primeiros passos neste trajeto para me tornar um pesquisador. Obrigado Pio, por ter me aceitado e acreditado em mim para desenvolver meu primeiro projeto tecnológico na UFAM. Obrigado Leandro, por ter me mostrado o potencial que existia em mim desde o meu primeiro dia de aula, e por todo o conhecimento que o senhor me passou através de todos esses anos. O senhor tem meu maior respeito pelo professor que o senhor é e espero um dia ser tão bom quanto o senhor nesta nobre arte que é ser um professor.

Agradeço especialmente a minha banca de defesa de Mestrado, Prof. Breno Bernard Nicolau de França e Prof(a). Carla Ilane Moreira Bezerra, e suplentes, Prof. Bruno Gadelha e Prof. Edson César, que tão gentilmente aceitaram e ofereceram seu tempo para participar da banca, ler o texto e fazer parte desta avaliação com seus comentários e críticas.

A minha primeira banca de defesa de graduação e banca de qualificação de mestrado: Leandro S. G. de Carvalho, Natasha M. C. Valentim, Edson César C. de Oliveira e Luciana A. M. Zaina; por terem aceitado participar, reservado tempo de participar e ler meu texto, por todas as suas críticas, discussões e sugestões que com certeza aprimoraram este trabalho.

Para todos os pesquisadores do USES que, de alguma forma, receberam-me neste grupo, suportaram-me, torceram por mim, aconselharam-me e de toda forma possível me ajudaram a chegar aqui: Bia, Natasha, Bruna, Carol, Drica, Mary, Úrsula, Márcia, Léo, Will, Walter, Randerson, Flávio, Lucas, Felipinho, Paty, Cezar, Edson, Gretchen, Letícia, Atacílio, Carlos, Fabian, Leonardo, Milka, Romualdo, Maria, Genildo, Myrian e Sérgio. Obrigado por todo o compartilhamento de conhecimento, alegrias, tristezas, brincadeiras e piadas, apoio, horas, dias e noites, semanas, meses e anos que passamos juntos assim como as experiências que tivemos. Tenho um carinho especial por esse grupo de pesquisa e por todos que dele já fizeram parte.

Para todos os amigos que, de alguma forma ou de outra, ajudaram-me, torceram, tomaram um café comigo ou que somente me deram uma palavra amiga, assim como para todos aqueles que de forma direta ou indireta colaboraram nessa pesquisa: Thomaz, Laura, Gilvan, Anny, Matheus, Kiong, Marcela, Vitória, Larissa, Felipe, Vicky, Giovana e Brandel.

Para Ariel Antony Afonso e Paulo Martins, por terem sempre me empurrado em frente e sempre estarem ali por mim, por sempre terem me ajudado quando precisei, por terem participado da minha vida e continuado ali todos os dias e noites acordados resolvendo problemas. Sempre serei grato por ter vocês dois em minha vida e que nós conseguimos alcançar nossos objetivos e sonhos.

A UFAM e ao IComp juntamente com todos os professores e pessoal administrativo, por sempre terem dado o suporte necessário para que eu conseguisse concluir esta fase de minha vida e sempre serem capazes de resolver os problemas que surgiam. Ao Prof. Eduardo Luzeiro Feitosa e Prof. Edleno, um agradecimento especial pela devoção e tempo investido coordenando o PPGI desta universidade.

Finalmente, a FAPEAM, CAPES e CNPq pelos apoios financeiros dados durante este mestrado, assim como por terem aceitado este projeto de mestrado e disponibilizar uma bolsa de estudos e apoio financeiro.

“You, me, or nobody is gonna hit as hard as life. But it ain't about how hard you hit. It's about how hard you can get hit and keep moving forward. How much you can take and keep moving forward. That's how winning is done!”

Rocky Balboa

Resumo da Dissertação apresentada ao PPGI/UFAM como um dos requisitos para alcançar o grau de Mestre em Informática.

AVALIAÇÃO DA PLAYABILITY DE JOGOS DIGITAIS COM AVALIADORES NÃO ESPECIALISTAS ATRAVÉS DE CONJUNTOS HEURÍSTICOS CUSTOMIZÁVEIS

Felipe Sonntag Manzoni
2021

Orientadora: Ph.D. Tayana Uchôa Conte

Programa: Programa de Pós-Graduação em Informática (*PPGI*) da Universidade Federal do Amazonas (*UFAM*).

A Playability é definida como o grau pelo qual um jogador é capaz de entender, controlar e compreender um jogo através de todas as suas interações. As avaliações da Playability de jogos digitais são essenciais no desenvolvimento e na criação de jogos com melhores interações, interfaces e mecânicas. Startups e pequenas empresas que desenvolvem seus próprios jogos necessitam de métodos de avaliação bem projetados, flexíveis e fáceis de aplicar. Além disso, atualmente, as empresas de desenvolvimento de jogos buscam soluções de baixo custo e com resultados imediatos para a empresa. Mas, as soluções definidas pela literatura não levam em consideração as diferenças entre tipos e gêneros de jogos, além de os inspetores necessitarem de um maior suporte para realizar o processo de avaliação. Dessa forma, este trabalho propõe, avalia e desenvolve tecnologias capazes de suprir as necessidades das empresas de desenvolvimento de jogos atuais. Este trabalho apresenta a proposta e avaliação de duas técnicas para a avaliação da Playability em Jogos Digitais. A NExPlay é a primeira técnica desenvolvida para a avaliação da Playability e tem por objetivo realizar as avaliações através de avaliadores não especialistas com o objetivo de diminuir custos e tempo com avaliadores especialistas. A segunda proposta é a técnica CustomCheck4Play que tem por objetivo principal ser uma técnica customizável de acordo com o tipo e gênero de jogo sendo avaliado. Adicionalmente foi proposta a ferramenta de apoio PEACH para facilitar no processo de customização do conjunto heurístico utilizado pela CustomCheck4Play, habilitando a capacidade de realizar avaliações de modo mais fácil, rápido e organizado. A NExPlay, através de sua avaliação, mostrou-se uma técnica viável na avaliação da Playability de jogos digitais, no entanto, os inspetores identificaram dificuldades no entendimento de algumas heurísticas, solicitaram uma categorização mais completa e identificaram a existência de diversas heurísticas em apenas uma, o que dificultava entender qual aspecto da heurística deveria ser avaliado ou não. Por essas razões foi proposta a técnica CustomCheck4Play. A CustomCheck4Play mostrou estatisticamente ter uma grande diferença, sendo mais eficaz e eficiente na detecção de problemas da Playability que outro conjunto heurístico da literatura quando avaliamos o mesmo jogo ou jogos diferentes propostos dentro do espectro dos tipos e gêneros de jogos desejados pelas técnicas.

Palavras-chave: Playability, Jogabilidade, Avaliação Heurística, Conjuntos Heurísticos, Avaliação da Jogabilidade, Jogos Digitais, Avaliação de Jogos.

Abstract of the Thesis presented to PPGI/UFAM as a partial fulfillment of the requirements for the Degree of Master in Informatics.

DIGITAL GAMES PLAYABILITY ASSESSMENT THROUGH CUSTOMIZABLE HEURISTIC SETS WITH NON-EXPERT EVALUATORS

Felipe Sonntag Manzoni
2021

Advisor: Ph.D. Tayana Uchôa Conte

Program: Programa de Pós-Graduação em Informática (*PPGI*) from the Universidade Federal do Amazonas (*UFAM*).

Playability is defined as the degree by which a player is able to learn, control, and understand a game through all its interactions with the game. Playability assessments for digital games are essential in game development for creating games with better interactions, interfaces, and mechanics. Startups and small businesses that develop games need well-designed, flexible, and easy-to-apply evaluation methods. Also, nowadays, game development companies are looking for low-cost solutions with immediate results for the company. Besides, solutions defined in the literature do not consider the differences between types and genres of games and inspectors need more support to carry out the evaluation process. In this way, this work proposes, evaluate, and develops technologies that can supply the needs of the current game development companies. This work presents the proposal and evaluation of two assessment techniques for Playability in Digital Games. The first developed technique for the assessment of Playability was NExPlay and it aims to carry out evaluations through non-expert evaluators with the objective of reducing costs and time with expert evaluators. The second proposal is the CustomCheck4Play technique, whose primary objective is to be a customizable technique according to the types and genres of games being evaluated by it. Additionally, the PEACH Support Toll was proposed in order to facilitate the customization process of the heuristic set used by CustomCheck4Play, enabling the ability to perform assessments in an easier, faster and more organized way. NExPlay, through its evaluation, proved to be a viable solution for the evaluation of Playability in digital games, however, inspectors identified difficulties in the understanding of some heuristics, requested a more complete categorization, and identified the existence of several heuristics in only one, which made it difficult to understand which aspect of the heuristic should be evaluated or not. For these reasons, we proposed CustomCheck4Play. CustomCheck4Play greatly statistically proved to be more effective and efficient in detecting Playability problems than another heuristic set from literature when we evaluated the same game or different games that are from the same subset of types and genres proposed for the techniques.

Keywords: Playability, Heuristic Evaluation, Heuristic Sets, Playability Assessment, Digital Games, Game Evaluation.

Lista de Figuras

Figura 1.1. Metodologia de pesquisa definida.....	32
Figura 2.1. Características da qualidade de software gerais e aspectos da usabilidade segundo a NBR ISO/IEC 9126-1.	38
Figura 2.2. String de busca utilizada na biblioteca selecionada.....	49
Figura 2.3. Tela inicial do jogo The Cannon Man Adventures	61
Figura 2.4. Tela inicial do jogo Green the Planet 2	61
Figura 2.4. Tela inicial do jogo Green the Planet 2	68
Figura 2.3. Tela inicial do jogo The Cannon Man Adventures	68
Figura 4.1. Tela inicial do jogo Leap of Cat.....	82
Figura 4.2. Boxplot para os valores de Eficiência.....	88
Figura 4.3. Boxplot para os valores de Eficácia	88
Figura 6.1. Figura da tela inicial do jogo <i>Leap of Cat</i>	111
Figura 6.1. Boxplot para a eficiência do estudo CustomCheck4Play.	117
Figura 6.2. Boxplot para a Eficácia do estudo CustomCheck4Play.	118
Figura 7.1. Tela inicial do site da ferramenta PEACH	124
Figura 8.1. Boxplot para os valores da Eficiência e Eficácia deste estudo.	142

Lista de Tabelas

Tabela 2.1. Características da usabilidade avaliadas na Playability baseada nas heurísticas avaliadas por Pinelle, Wong e Stach (2008).	44
Tabela 2.2. Objetivo de pesquisa do Mapeamento Sistemático de acordo com o paradigma GQM.	47
Tabela 2.3. Termos utilizados para os parâmetros do PICO	49
Tabela 2.4. Critérios de inclusão e exclusão para o mapeamento desenvolvido.	50
Tabela 2.5. Número de publicações encontradas a partir da biblioteca digital e snowballing.	53
Tabela 2.6. Artigos selecionados após o segundo filtro do processo de MS desenvolvido divididos entre publicações da pesquisa automática e por snowballing.	53
Tabela 2.7. Feature analysis dos conjuntos heurísticos avaliados no mapeamento sistemático.	63
Tabela 3.1. Heurísticas da categoria Game Play	73
Tabela 3.2. Heurísticas da categoria Usabilidade	74
Tabela 3.3. Heurísticas da categoria Mecânicas	75
Tabela 4.1. Níveis de conhecimento dos avaliadores por conjunto utilizado	80
Tabela 4.2. Valores absolutos quantitativos por participante do estudo	85
Tabela 4.3. Valores médios por solução utilizada	86
Tabela 4.4. Valores de significância dos testes de normalidade para os grupos	87
Tabela 5.1. Categorias da técnica CustomCheck4Play de acordo com as seções do jogo.	102
Tabela 5.2. Heurísticas da Técnica CustomCheck4Play.	103
Tabela 6.1. Valores absolutos para o estudo da técnica CustomCheck4Play.	114
Tabela 6.2. Valores das médias para o conjunto CustomCheck4Play.	116
Tabela 6.3. Valores para o teste de normalidade do estudo da CustomCheck4Play.	117
Tabela 8.1. Valores quantitativos absolutos por avaliador.	138
Tabela 8.2. Médias calculadas de acordo com os resultados quantitativos.	140
Tabela 8.3. Valores do teste de normalidade para a eficiência e eficácia.	141

Sumário

CAPÍTULO 01 – INTRODUÇÃO	28
1.1 Contexto	28
1.2 Definição do problema/motivação	29
1.3 Objetivo de pesquisa	31
1.3.1 Objetivos Específicos	31
1.4 Metodologia	32
1.5 Organização do texto	36
CAPÍTULO 02 – PLAYABILITY E MÉTODOS DE AVALIAÇÃO DA PLAYABILITY	38
2.1 Jogos Digitais	38
2.1.1 Tipos de Jogos	39
2.1.2 Gêneros de Jogos	42
2.2 Playability	42
2.2.1 Usabilidade	44
2.2.2 Mecânicas	45
2.2.3 Game Play	45
2.3 Mapeamento Sistemático	46
2.3.1 Metodologia de pesquisa	46
2.3.2 Protocolo do mapeamento sistemático	46
2.3.2.1 Definição de objetivo e questão de pesquisa	47
2.3.2.2 Estratégia de busca das publicações	48
2.3.2.3 Critérios de seleção de artigos	50
2.3.2.4 Procedimentos para seleção de artigos	51
2.3.3 Condução do mapeamento sistemático	52
2.3.4 Métodos para avaliação da Playability	54
2.3.4.1 Questionários e Entrevistas	55
2.3.4.2 Avaliação da Playability através de Conjuntos Heurísticos	56
2.3.4.2.1 Conjuntos Heurísticos para Avaliação da Jogabilidade identificados através do MSL 57	
2.3.4.2.2 Estudos de avaliação e definições da Playability	60
2.3.5 Discussão de resultados	63
2.4 Hands-On conjuntos heurísticos	67
2.4.1 Objetivo primário e conjuntos selecionados	67
2.4.2 Jogos selecionados para o processo de avaliação	68
2.4.3 Processo de avaliação	68

2.4.4	Resultados avaliados.....	69
2.5	Conclusão.....	70
CAPÍTULO 03 – NExPlay: NON-EXPERT PLAYABILITY HEURISTIC SET		71
3.1	Introdução	71
3.2	Desenvolvimento	71
3.3	Conjunto Heurístico.....	73
3.4	Discussão.....	75
CAPÍTULO 04 – ESTUDO DE VIABILIDADE DA NEXPLAY		77
4.1	Introdução	77
4.2	Estudo Experimental	77
4.2.1	Metodologia.....	78
4.2.2	Jogo Selecionado	81
4.2.3	Processo Experimental	82
4.2.4	Seleção de variáveis	84
4.2.5	Hipóteses.....	84
4.3	Resultados.....	85
4.3.1	Resultados Quantitativos	85
4.3.2	Resultados Qualitativos.....	89
4.3.2.1	Maneiras de Condução da Avaliação Utilizando Conjuntos Heurísticos	90
4.3.2.2	Especificidade das Heurísticas.....	90
4.3.2.3	Categorização das Heurísticas	91
4.3.2.4	Quantidade de Heurísticas	92
4.3.2.5	Entendimento das Heurística e Direcionamento por Gênero e Tipo de Jogo	93
4.3.2.6	Entrevistas	94
4.3.2.7	Descrição dos Problemas Identificados.....	97
4.3.3	Ameaças a Validade.....	97
4.4	Discussão.....	98
4.5	Conclusão.....	99
CAPÍTULO 05 – CustomCheck4Play: UMA TÉCNICA CUSTOMIZÁVEL PARA A AVALIAÇÃO DA PLAYABILITY		101
5.1	Introdução	101
5.2	Heurísticas da Técnica CustomCheck4Play.....	102
5.3	Processo de Customização.....	105
CAPÍTULO 06 – ESTUDO DE VIABILIDADE DA CUSTOMCHECK4PLAY		109
6.1	Estudo Experimental	109
6.2	Processo Experimental	110
6.2.1	Primeira rodada.....	111

6.2.2	Segunda Rodada	112
6.2.3	Seleção de variáveis	112
6.2.4	Hipóteses.....	112
6.2.5	Execução do Estudo (1a e 2a rodadas).....	113
6.3	Resultados.....	115
6.3.1	Resultados quantitativos	116
6.3.1.1	Eficiência.....	117
6.3.1.2	Eficácia.....	118
6.3.2	Resultados qualitativos.....	118
6.3.2.1	Entendimento e especificidade das heurísticas.....	119
6.3.2.2	Viabilidade das sentenças de ajuda	120
6.3.2.3	Número de heurísticas e Heurísticas faltantes	120
6.3.2.4	Categorização da técnica	120
6.3.3	Ameaças à validade	121
6.4	Discussão.....	122
CAPÍTULO 07 – ASSISTENTE DE CUSTOMIZAÇÃO PEACH		123
7.1	Introdução	123
7.2	Processo de customização.....	124
7.3	Estudo de viabilidade da PEACH e primeiras impressões	126
7.4	Resultados.....	128
7.5	Conclusão.....	131
CAPÍTULO 08 – ESTUDO COMPARATIVO CUSTOMCHECK4PLAY		132
8.1	Estudo Experimental	132
8.2	Processo Experimental	134
8.2.1	Primeira rodada.....	135
8.2.2	Segunda Rodada	135
8.2.3	Seleção de variáveis	136
8.2.4	Hipóteses.....	136
8.2.5	Execução do Estudo.....	136
8.3	Resultados.....	139
8.3.1	Resultados quantitativos	140
8.3.1.1	Eficiência.....	141
8.3.1.2	Eficácia.....	142
8.3.2	Resultados qualitativos.....	142
8.3.2.1	Entendimento das heurísticas e Especificidade das heurísticas.....	144
8.3.2.2	Viabilidade das Sentenças de Ajuda.....	144

8.3.2.3	Tamanho do Conjunto de Heurísticas e Heurísticas Faltantes	145
8.3.2.4	Categorização do Conjunto Heurístico	146
8.3.3	Ameaças à validade	147
8.4	Discussão	148
CAPÍTULO 09 – CONCLUSÕES E TRABALHOS FUTUROS.....		149
9.1	Conclusões	149
9.2	Contribuições.....	150
9.3	Trabalhos Futuros	152
REFERÊNCIAS BIBLIOGRÁFICAS		154
APÊNDICE A – QUESTIONÁRIO DE CARACTERIZAÇÃO.....		159
APÊNDICE B – TABELA DE IDENTIFICAÇÃO E CLASSIFICAÇÃO DOS DEFEITOS.....		160
APÊNDICE C – QUESTIONÁRIO DE AVALIAÇÃO E OPINIÃO PÓS-USO DOS CONJUNTOS HEURÍSTICOS.....		161
APÊNDICE D – QUESTIONÁRIO DE MODULARIZAÇÃO DA TÉCNICA CUSTOMCHECK4PLAY		162
APÊNDICE E – QUESTIONÁRIO DE CARACTERIZAÇÃO DO ESTUDO REALIZADO PARA A CUSTOMCHECK4PLAY		166
APÊNDICE F – QUESTIONÁRIO DE FEEDBACK DO ESTUDO REALIZADO PARA A PEACH		167
APÊNDICE G – QUESTIONÁRIO DE FEEDBACK DO ESTUDO REALIZADO PARA A CUSTOMCHECK4PLAY		168
APÊNDICE H – CODEBOOK DESENVOLVIDO PARA A ANÁLISE QUALITATIVA DO ESTUDO COMPARATIVO DA CUSTOMCHECK4PLAY.....		169

CAPÍTULO 01– INTRODUÇÃO

Este capítulo apresenta a contextualização desta pesquisa de mestrado, o problema e a motivação, a questão e os objetivos da pesquisa, além da metodologia aplicada.

1.1 Contexto

Playability, termo em inglês derivado da palavra Jogabilidade em português, é definida como um meio de avaliação das interações entre jogadores e o próprio jogo (Desurvire, Caplan e Toth, 2004; Desurvire e Wiberg, 2008, 2009). Esta pesquisa utiliza o termo inglês, uma vez que a palavra em português Jogabilidade, pode representar diversos significados. Além disso, a palavra em inglês é mais utilizada e referenciada. Sendo assim, a Playability avalia diretamente o design de interação com a interface e impacta diretamente em como os usuários interagem com os jogos (Desurvire e Wiberg, 2009; Korhonen, Paavilainen e Saarenpää, 2009; Nacke et al., 2009). Conseqüentemente, Playability é considerada um aspecto indispensável na avaliação de jogos digitais, de modo que, através dela é possível diretamente aprimorar a qualidade dos jogos (Nacke et al., 2009).

Embora exista o conceito de usabilidade para a avaliação de interfaces e interação dos usuários com elas (Barbosa e Silva, 2010), não se pode avaliar jogos digitais utilizando esse mesmo conceito. A usabilidade se preocupa primeiramente, e fundamentalmente, com a facilidade de interação dos usuários e com a facilidade de finalizar uma ação pretendida (Bevan, 2001). Porém, em jogos digitais, existe a necessidade de que as tarefas a serem completadas possuam certa dificuldade e apresentem desafios que não devem ser facilitados (Rogers, 2014). Portanto, certos aspectos da usabilidade devem ser considerados na avaliação da Playability, porém, nem todos, uma vez que jogos não podem ser extremamente fáceis, mas podem, por exemplo, ser extremamente intuitivos.

Embora a Playability seja importante e apresente muitos benefícios, diversas empresas de desenvolvimento de jogos não realizam a avaliação e correção dos defeitos identificados por ela (Politowski *et al.*, 2016; Rajanen & Tapani, 2018). De acordo com o avanço das tecnologias do mundo digital, além da necessidade de fases de desenvolvimento cada vez mais curtas para lançamentos rápidos, existe a pressão da comunidade que espera maiores e melhores jogos da indústria com tecnologia mais avançada. Além disso, desenvolvedoras de jogos digitais diminuem cada vez mais seus orçamentos de desenvolvimento para que haja um maior retorno.

Disso resultam a falta de avaliações e fases de teste apropriadas as quais a indústria deveria desenvolver e realizar (Duarte e Veloso, 2016).

Diversos autores descrevem na literatura métodos para a avaliação da Playability em jogos digitais que focam especificamente em diversos aspectos utilizando conjuntos heurísticos (Desurvire, Caplan e Toth, 2004; Korhonen e Koivisto, 2006; Barcelos et al., 2011). Conjuntos heurísticos são uma coleção, um agrupamento ou uma união de heurísticas, que podem ser definidas como defeitos comumente identificados ou boas práticas para um determinado objetivo de um determinado software ou objeto de avaliação. Por exemplo, M. A. Federoff (2002) desenvolveu heurísticas para a criação e avaliação de ambientes que sejam mais divertidos aos jogadores a partir de uma *guideline* de boas práticas para a promoção da diversão no jogo. De outro modo, Pinelle et al. (2008a) desenvolveu um conjunto de heurísticas a partir da análise de diversos comentários online sobre jogos previamente desenvolvidos, listando uma série de defeitos encontrados nos jogos após seu lançamento. Ou seja, um conjunto heurístico pode ser uma coleção de defeitos descritos para a avaliação de um software, onde o avaliador busca tais defeitos no software em questão. Além disso, a utilização de conjuntos heurísticos provém da técnica de avaliação heurística proposta por Nielsen e Molich (1990), uma técnica muito conhecida e que pode ser facilmente adaptada para diversos contextos de avaliação de acordo com a quantidade de estudos e técnicas desenvolvidas a partir dela (Desurvire, Caplan e Toth, 2004; Korhonen e Koivisto, 2006; Desurvire e Wiberg, 2008; Barcelos et al., 2011; Duarte e Veloso, 2016).

1.2 Definição do problema/motivação

Embora existam diversas propostas de técnicas utilizando conjuntos heurísticos para a avaliação da Playability em jogos digitais, essas propostas possuem diversos defeitos e lacunas que precisam ser avaliadas: os conjuntos utilizados por tais técnicas são muito longos, dificultando a avaliação e atrasando a avaliação dos inspetores (Korhonen e Koivisto, 2006; Barcelos et al., 2011); os conjuntos apresentados nessas técnicas utilizam linguagem extremamente técnica, dificultando o entendimento por um inspetor inexperiente (Barcelos et al., 2011; Desurvire e Wiberg, 2008); os conjuntos utilizados são generalizados, aumentando demais o escopo de avaliação criando situações em que heurísticas não se relacionam com o jogo avaliado (Desurvire e Wiberg, 2008; Barcelos et al., 2011). Além desses problemas identificados na literatura, também existem os problemas reportados pelas empresas de desenvolvimento de jogos para que as mesmas não realizem avaliações durante o desenvolvimento dos seus jogos

(Aleem, Capretz e Ahmed, 2016; Politowski et al., 2016). Na maioria dos casos, as empresas de desenvolvimento de jogos reportam o alto custo para realização dessas avaliações, a necessidade de especialistas, alto custo de tempo para sua realização e a falta de um sistema de apoio para realizá-las com uma metodologia não especificada.

De acordo com o *survey* desenvolvido por Rajanen e Tapani (2018), empresas de desenvolvimento de jogos preferem utilizar entrevistas semiestruturadas e seções de jogo com avaliação assistida. Um dos principais defeitos encontrados na forma como as avaliações de Playability são propostas é a necessidade de especialistas para realizá-las, aumentando, assim, o custo para realização dessas avaliações também. A necessidade de especialistas pode criar ainda um custo maior de tempo de acordo com a sua disponibilidade, principalmente se a empresa de desenvolvimento em questão não possui nenhum. Por fim, esta pesquisa não foi capaz de identificar na literatura disponível, no decorrer de sua execução, um sistema ou ferramenta que realize o papel de apoio durante as avaliações. Esse sistema ou ferramenta poderia facilitar a interação e possivelmente diminuir o tempo e custo para realização dessas avaliações.

A motivação para o desenvolvimento deste trabalho surge da necessidade de técnicas de avaliação da Playability que permitam a avaliação de jogos digitais através de inspetores “não especialistas”, além de uma redução de custo (baseado na diminuição de tempo e necessidade de especialistas) e tempo para a realização desse tipo de avaliação. A diminuição de custo prevista para esta pesquisa está relacionada com a previsão de diminuição de acordo com a diminuição de tempo para identificação de defeitos (aumento da eficiência) e o aumento do número de defeitos identificados neste mesmo tempo sem o aumento do número de falso-positivos (aumento da eficácia). O emprego de avaliadores “não especialistas”, ou pelo menos a permissão e habilitação de uma técnica a ser utilizada por eles, pode significar uma diminuição de custos no processo de avaliação além de facilitar o processo de modo que todo tipo de avaliador possa utilizá-la. Avaliadores “não especialistas” podem ser mais facilmente encontrados e contratados/contatados para a realização desse tipo de atividade se considerado que: a maioria das pessoas com pequena base de conhecimento de processos de engenharia de software, e conceitos de interação humano-computador, pode ser considerada um avaliador em potencial. A disponibilidade cria uma diminuição de custo na contratação para a realização do estudo além de criar uma diminuição no tempo necessário para a avaliação, uma vez que não exige o tempo necessário para a pesquisa de avaliadores especialistas. Além disso, existe a necessidade do desenvolvimento de uma ferramenta ou sistema de apoio para os inspetores, de modo que ela possa apoiá-los durante a avaliação de jogos. Além de ser um fator de apoio no

processo de avaliação, esse é um dos motivos indicados pela pesquisa desenvolvida por Rajanen e Tapani (2018) para que empresas de desenvolvimento de jogos não realizem esse tipo de avaliação dentro de seus processos de desenvolvimento. De outro modo, a ferramenta ou sistema de apoio pode ainda facilitar o entendimento dos avaliadores no processo de avaliação, uma vez que, mesmo sem a presença de um especialista, existe uma ferramenta que pode guiá-los durante a sua extensão.

Portanto, a questão de pesquisa deste trabalho consiste em “Como reduzir custo, tempo e diminuir a necessidade de especialistas na execução e uso das técnicas de avaliação da Playability em jogos digitais?”. Para responder a essas questões, as técnicas de avaliação existentes foram mapeadas de forma sistemática de modo que serviram de base para a técnica apresentada neste trabalho com o objetivo de tentar corrigir os defeitos apresentados pela literatura.

1.3 Objetivo de pesquisa

O objetivo principal deste trabalho consiste em desenvolver e evoluir uma técnica de avaliação da Playability em jogos digitais de forma a apoiar a avaliação de artefatos por avaliadores não especialistas em todas as fases de desenvolvimento dos jogos digitais. Para esta pesquisa, utilizaremos os conceitos apresentados por Falbo, R. (1998), onde o mesmo define uma técnica como um procedimento para a realização de uma atividade geral, que não descreve como fazê-la em termos de sub-atividades. De forma semelhante, entendeu-se método como um procedimento sistemático, definindo passos e heurísticas para a realização de uma ou mais atividades. Desta forma, pretendemos desenvolver uma técnica, já que não serão definidas cada uma das sub-atividades específicas a serem realizadas para o processo de avaliação (Severino, 2007; Barbosa, 1989).

Além disso, foi desenvolvida uma ferramenta de apoio à avaliação da Playability por tais avaliadores. Dessa forma, foi desenvolvida uma técnica eficiente, capaz de identificar um maior número de defeitos no menor tempo possível, e eficaz, identificando um menor número de falso-positivos possível em relação ao total de defeitos identificados, apoiando a melhoria da qualidade dos jogos desenvolvidos e lançados.

1.3.1 Objetivos Específicos

Os objetivos específicos desta pesquisa são:

- Criar uma base de conhecimento sobre Playability e técnicas de avaliação da Playability para integração da sua avaliação durante o processo de desenvolvimento de jogos.
- Desenvolver um conjunto de técnicas para a avaliação da Playability contendo aspectos específicos para tipos e gêneros de jogos existentes, sendo possível customizar a avaliação de acordo com o jogo a ser avaliado.
- Verificar características que influenciam na avaliação da Playability em jogos digitais para aprimoramento da técnica de avaliação.
- Desenvolver uma ferramenta de apoio à avaliação da Playability de jogos digitais através das técnicas desenvolvidas.

1.4 Metodologia

A metodologia utilizada para atingir os objetivos desta pesquisa é descrita a seguir e apresentada na Figura 1.1. Divide-se em quatro etapas: definição do tema de pesquisa; investigação do problema; definição da solução; e avaliação e refinamento da solução. As atividades pertinentes a cada etapa são resumidamente descritas a seguir:

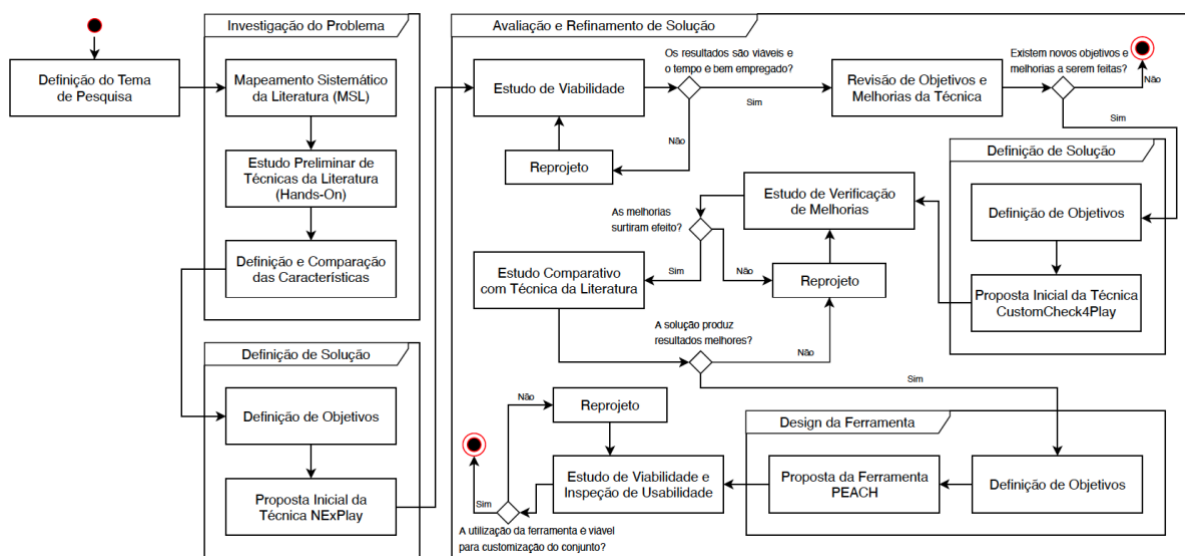


Figura 1.1. Metodologia de pesquisa definida.

A primeira etapa representada na Figura 1.1, de **Definição do tema de pesquisa**, é apresentada no CAPÍTULO 01 desta pesquisa, referente à introdução da pesquisa, definição do problema e seus objetivos. Nessa etapa, uma clara definição do tema e área de pesquisa em que o trabalho será desenvolvido é definida e um problema específico dessa área é identificado para ser explorado e, posteriormente, resolvido.

Investigação do problema: objetiva compreender melhor o problema a ser investigado, identificando soluções existentes na literatura. Um estudo da literatura foi realizado em busca de conceitos e soluções previamente propostas para a solução do problema investigado. Essa etapa possui três atividades descritas abaixo:

- Mapeamento Sistemático da Literatura (MSL): foi realizado um mapeamento sistemático da literatura com o objetivo de encontrar trabalhos que apresentassem conjuntos heurísticos, metodologias de avaliação e estudos sobre a avaliação da Playability em jogos digitais. Os resultados desse mapeamento foram utilizados como a base literária desta pesquisa e apoio no processo de entendimento e pesquisa do problema investigado, além da identificação das soluções pré-existentes.

- Estudo preliminar de técnicas da literatura (Hands-On): um estudo preliminar foi realizado com a intenção de utilizar as técnicas e metodologias identificadas na literatura através do mapeamento sistemático. O objetivo desse estudo foi melhor compreender e investigar não só o processo de avaliação utilizado por essas técnicas como também os defeitos e lacunas existentes no uso delas dentro dos processos de desenvolvimento de jogos.

- Definição e comparação das características: após o uso das técnicas identificadas na literatura, as principais características que definem o processo de uso, população e necessidades / lacunas que podem aprimorar o processo de avaliação de jogos foram listadas e comparadas. Dessa forma, as características identificadas e comparadas foram utilizadas para classificar os objetivos que poderiam ter o maior impacto positivo quando trabalhados dentro da técnica proposta.

Definição da solução: de acordo com os resultados do mapeamento sistemático e dos estudos experimentais, juntamente com a análise das características identificadas nos trabalhos da literatura, foi observada a necessidade de uma técnica reduzida, voltada para avaliadores não especialistas com foco nos defeitos mais comuns identificados. Dessa forma, os objetivos iniciais desta pesquisa foram definidos e uma proposta inicial da técnica foi desenvolvida através da NExPlay.

- Definição de objetivos: a partir da análise e definição das características identificadas nos trabalhos da literatura, foram definidos alguns objetivos a serem alcançados com a proposta e desenvolvimento da técnica inicial. Essa técnica inicial destaca-se em poder ser utilizada por avaliadores não especialistas, além de apoiá-los pela facilidade de aplicação, entendimento das heurísticas e do processo. Além disso, um segundo foco da técnica inicial é

desenvolver uma técnica mais curta e simples com o objetivo de diminuir o tempo e esforço necessário para sua aplicação durante o processo de desenvolvimento de jogos nas empresas.

- Proposta inicial da técnica NExPlay: foi desenvolvida uma técnica de avaliação com o objetivo de diminuir o número total de heurísticas avaliadas pela técnica em relação aos conjuntos apresentados na literatura, encontrados através do mapeamento sistemático. Também desejou-se que a linguagem utilizada na descrição de tais heurísticas fosse simples o suficiente para o entendimento por avaliadores não especialistas. Essa proposta inicial da técnica de avaliação teve como base para sua formulação os conjuntos encontrados durante o mapeamento preliminar da literatura. A proposta inicial da técnica de avaliação pode ser encontrada na descrição do artigo '*NExPlay: Playability Assessment for Non-Experts Evaluators*', aceito na conferência *International Conference on Enterprise and Information Systems (ICEIS) 2018* (Manzoni, Ferreira e Conte, 2018).

Avaliação e refinamento da solução: essa etapa da metodologia é direcionada para que, após a definição inicial da técnica, possamos observar os efeitos gerados pela proposta como a proposta pode ser incrementada para assim apoiar mais ainda os avaliadores não especialistas. Para esse processo ser realizado, alguns estudos experimentais e desenvolvimentos foram executados, conforme segue:

- Estudo de viabilidade da técnica NExPlay: foi realizado um estudo preliminar utilizando a técnica proposta e um conjunto heurístico da literatura de forma comparativa. Os resultados desse estudo serviram para verificar a aceitação da técnica NExPlay em termos de sua eficiência, eficácia, utilidade e facilidade de compreensão das heurísticas apresentadas pela técnica de avaliação da Playability por avaliadores não especialistas. O resultado desse estudo preliminar também é descrito no artigo '*NExPlay: Playability Assessment for Non-Experts Evaluators*', aceito na conferência *International Conference on Enterprise and Information Systems (ICEIS)*, 2018 (Manzoni, Ferreira e Conte, 2018). A partir dos resultados coletados foi observado que eram viáveis de acordo com os objetivos definidos e que o tempo era mais bem empregado em relação à técnica comparada. Assim, pudemos passar para o próximo passo com a identificação de pontos de melhoria e a definição de novos objetivos.

- Revisão de objetivos e melhorias da técnica: de acordo com os resultados obtidos e pela análise qualitativa dos participantes do estudo conduzido, foi observado existirem pontos de melhoria que poderiam ser mais bem trabalhados em uma nova versão da técnica. De acordo com os pontos observados, foi definido que a técnica deveria ter um foco maior no jogo sendo avaliado naquele momento ao invés de uma técnica geral que poderia ser aplicada a todos os tipos de jogos. Dessa forma, o objetivo principal definido para essa fase foi de criar uma técnica

customizável para os diferentes tipos e gêneros de jogos avaliados, criando uma técnica mais focada, que demandasse menos conhecimento técnico e menor tempo para avaliação.

Definição de solução: de acordo com os novos objetivos definidos, um novo processo de definição da solução foi realizado, desenvolvendo uma segunda solução focada nos defeitos identificados na primeira fase de definição de solução de acordo com o que segue:

- Definição de objetivos e proposta da técnica customizável CustomCheck4Play: utilizando-se das características observadas durante o estudo preliminar e da evolução da técnica NExPlay, foi proposta uma nova técnica customizável de forma a adaptá-la aos tipos e gêneros de jogos a serem avaliados. A proposta dessa técnica é descrita no artigo '*Straight to the Point - Evaluating What Matters for You: A Comparative Study on Playability Heuristic Sets*', aceito na conferência *International Conference on Enterprise and Information Systems (ICEIS)*, 2020 (Manzoni, Barbosa, Silveira e Conte, 2020).

Após a definição da nova técnica de avaliação da Playability CustomCheck4Play, o ciclo de **avaliação e refinamento da solução** pôde ser continuado de forma a validar as melhorias desenvolvidas e identificar o quanto é capaz de apoiar os objetivos definidos. A seguir os passos para validação, verificação e melhoria da técnica são descritos:

- Estudo de verificação de melhorias propostas com a técnica CustomCheck4Play e estudo comparativo com técnica da literatura: essas duas etapas consistem em caracterizar a técnica proposta e verificar sua possibilidade de uso. Para isso, foi conduzido um experimento controlado de modo a verificar a aceitação da técnica para a avaliação da Playability em relação a outro conjunto heurístico existente na literatura. Foram avaliadas eficiência e eficácia da técnica proposta (CustomCheck4Play) para verificar se haveria uma melhora significativa em relação a outro conjunto da literatura. Os resultados desse estudo podem ser encontrados no artigo '*Straight to the Point - Evaluating What Matters for You: A Comparative Study on Playability Heuristic Sets*', aceito na conferência *International Conference on Enterprise and Information Systems (ICEIS)*, 2020 (Manzoni, Barbosa, Silveira e Conte, 2020). A partir dos resultados obtidos nesse estudo, identificamos que a técnica cumpria os objetivos definidos para ela, sendo possível passar para o próximo passo da pesquisa com um experimento comparativo de técnicas.

Design da ferramenta de apoio: nessa etapa pretendeu-se desenvolver uma ferramenta que pudesse apoiar os avaliadores durante todo o processo de avaliação. Essa ferramenta sugere as heurísticas e categorias adequadas para avaliação, contidas na técnica de avaliação proposta, dependendo do tipo e gênero de jogo, além de apoiar todo o processo de avaliação.

- *Definição de objetivos e desenvolvimento da ferramenta PEACH:* a partir dos resultados observados nos estudos experimentais realizados, foram definidos objetivos a serem atingidos em relação ao processo de customização da técnica desenvolvida através da ferramenta PEACH. A ferramenta foi proposta de modo que qualquer desenvolvedor, avaliador ou jogador pudesse fazer a customização do conjunto heurístico a ser utilizado no processo de avaliação da Playability de jogos digitais. A ferramenta é direcionada a ser um apoio aos avaliadores, de modo a customizar o conjunto e possivelmente apresentar um conjunto de heurísticas mais voltadas especificamente ao jogo sendo avaliado.

- *Estudo de viabilidade e inspeção de usabilidade:* a partir da definição dos objetivos específicos para a ferramenta de apoio e seu desenvolvimento, um estudo experimental e inspeção de usabilidade foram desenvolvidos com a ferramenta em conjunto com a técnica de avaliação CustomCheck4Play. O primeiro objetivo desse estudo é direcionado a entender o quanto a ferramenta de apoio pode melhor direcionar o conjunto ao jogo específico sendo avaliado. O segundo objetivo é de inicialmente compreender o quanto a ferramenta é de fácil compreensão e utilização durante o processo de avaliação. No estudo desenvolvido, foi verificada uma melhora significativa em relação à eficiência e eficácia do conjunto da literatura comparado. De acordo com os resultados identificados a técnica atende a todos os objetivos propostos e demonstrou que pode ser eficiente e eficaz no processo de avaliação da Playability de jogos digitais. Os resultados desse estudo serão apresentados em um artigo a ser submetido para um Journal.

1.5 Organização do texto

Esse capítulo introdutório apresentou a contextualização, a definição do problema e da motivação, os objetivos da pesquisa e a metodologia de pesquisa. O conteúdo será detalhado ao longo dos próximos capítulos. Este trabalho está organizado segundo a estrutura descrita abaixo:

Capítulo 2 – Playability e Métodos para Avaliação da Playability: Apresenta os conceitos da Playability, assim como alguns métodos existentes para avaliação da Playability os quais foram selecionados através de um Mapeamento Sistemático da Literatura.

Capítulo 3 – NExPlay: Non-Expert Playability Heuristic Set: apresenta a técnica inicial proposta para a avaliação da Playability em jogos digitais por avaliadores não especialistas.

Capítulo 4 – Estudo de Viabilidade com a NExPlay: apresenta o estudo de viabilidade conduzido para a avaliação inicial do conjunto inicial NExPlay.

Capítulo 5 – CustomCheck4Play: Uma Técnica Customizável Para a Avaliação da Playability: apresenta a proposta da técnica desenvolvida CustomCheck4Play com o propósito de caracterizá-la na pesquisa.

Capítulo 6 – Avaliação de viabilidade da técnica CustomCheck4Play: Apresenta a avaliação realizada com o conjunto com o propósito de validar sua utilização em um contexto específico.

Capítulo 7 – Assistente de Customização PEACH: apresenta a proposta e avaliação inicial da usabilidade e funcionalidades do assistente de customização que apoia a CustomCheck4Play.

Capítulo 8 – Estudo Comparativo CustomCheck4Play: apresenta o segundo estudo comparativo desenvolvido para avaliar a CustomCheck4Play em comparação a um conjunto heurístico da literatura diferente da proposta anterior de modo a possuir resultados mais abrangentes em tipos de jogos diferentes fazendo uso de um conjunto customizado produzido pela PEACH.

Capítulo 9 – Conclusões e Trabalhos Futuros: apresenta as conclusões deste texto e os próximos passos a serem realizados nesta pesquisa.

CAPÍTULO 02 – PLAYABILITY E MÉTODOS DE AVALIAÇÃO DA PLAYABILITY

Este capítulo apresenta os conceitos e definições da Playability, discutindo sobre o objetivo de cada atributo da Playability que são utilizados para a avaliação de jogos digitais. São apresentados métodos e exemplos de técnicas aplicadas na avaliação da Playability de jogos digitais e seus impactos para a sociedade e indústria. Além disso, apresenta o processo de Mapeamento Sistemático da Literatura realizado e seus resultados.

2.1 Jogos Digitais

Os jogos digitais consistem em um conjunto de regras e atividades as quais devem ser desenvolvidas e seguidas de modo a fazer avanços em uma história, sendo ela cronológica ou não. Esse conjunto de regras, atividades, objetivos e história são entendidos e traduzidos no conceito de *gameplay* nos jogos digitais, em que esta *gameplay* pode modificar e definir este conjunto de regras, atividades, objetivos e história de acordo com o tipo de jogo sendo desenvolvido ou jogado (Orlando, 2007). De acordo com Arsenault e Perron (2009), a *gameplay* é definida como sendo a seção de um jogo a qual interage diretamente com o design e mecânicas do jogo, representando as atividades e ações possíveis de serem executadas no jogo e o quão facilmente essas ações são executadas. Além disso, jogos digitais têm como o objetivo principal serem objetos de entretenimento interativos (Prensky, 2007). Por isso, eles também definem mecânicas de interação com esse mundo virtualmente gerado, de modo a modificar seu conteúdo (Kapp, 2012). Com o avanço tecnológico, diversas mudanças aconteceram no modo como os usuários interagem com esses jogos assim como nas diversas formas que existem para contar a história atrelada a eles, além de constantemente evoluírem as mecânicas utilizadas para esta interação (Orlando, 2007).

Fica claro o avanço tecnológico ocorrido nos jogos digitais quando são observados os gráficos e mecânicas utilizadas pelo “*Tennis for Two*”¹, primeiro jogo desenvolvido na história. O jogo utilizava um osciloscópio² conectado a dois reles de energia variáveis, cada um destes, conectado a um potenciômetro e um botão. O potenciômetro registrava a quantidade de energia

¹ Acessível em: https://pt.wikipedia.org/wiki/Tennis_for_Two. Acessado em: 31/01/2020

² Acessível em: <https://pt.wikipedia.org/wiki/Oscilosc%C3%B3pio>. Acessado em: 31/01/2020

enviada ao sistema, a qual, por sua vez, determinava a altura (ou força) com que a bola seria rebatida, e o botão acionava a passagem de energia ao sistema. O jogo se baseava na simulação digital de um jogo de Tênis o qual seria disputado por duas pessoas, considerando as regras normais de um jogo de Tênis oficial.

Com o passar do tempo, os computadores passaram a possuir capacidades de processamento muito maiores em relação aos primeiros computadores desenvolvidos. Assim, novos jogos poderiam ser desenvolvidos, considerando diversas e variadas mecânicas do mundo real assim como ações e histórias mais imersivas utilizando diversos e diferentes hardwares (Orlando, 2007). Com o aumento da complexidade de mecânicas, atividades e na própria *gameplay*, diversos problemas de interação com os usuários surgiram. Diversos usuários não eram capazes de entender as diversas e infinitas informações que agora faziam parte da tela, não eram capazes de realizar as ações por causa da grande quantidade de botões e micro atividades necessárias para isso. Por isso, uma metodologia de avaliação e guias de boas práticas para o desenvolvimento desses jogos deveriam ser criados, para suporte aos desenvolvedores nesse processo (Rogers, 2014).

Com essa evolução nas mecânicas e *gameplay* dos jogos digitais, foi necessário também que uma classificação fosse criada de forma a identificar e caracterizar os jogos de acordo com seu conteúdo. Assim, surge o termo: ‘Tipo de jogo’; em que as mecânicas, o ritmo de jogo, estratégias e outras atividades poderiam ser identificadas simplesmente por um jogo pertencer a um certo tipo de jogo (Rogers, 2014).

2.1.1 Tipos de Jogos

A classificação inicial dos jogos se baseou nas categorizações criadas por Roger Caillois, em seu livro ‘*Man, Play and Games*’ de 1958 (Tradução: Caillois and Barash, 2001), e por Chris Crawford, em seu livro ‘*The Art of Computer Game Design*’ de 1984 (Crawford, 1984), os quais definiam, de forma muito preliminar, uma classificação de jogos de mesa através dos níveis de agitação e calma pelos quais seus jogadores passavam. Nos dias atuais, as classificações de tipos e gêneros de jogos foram modificadas de acordo com os avanços e criação de diferentes tipos. Por exemplo, a ESA (*Entertainment Software Association*) publica, anualmente, uma pesquisa de opinião e análise do último ano no mercado de jogos e entretenimento mundial. Nesses relatórios, os jogos apresentados são classificados de acordo com algumas características definidas por eles. Sendo assim, os tipos de jogos definem especificamente os aspectos da *gameplay* utilizada no jogo, através de diferentes modos de

input, desafios, modos de se contar a história e muitos outros. Uma classificação e exemplificação de alguns tipos de jogos definidos hoje em dia estão disponíveis abaixo (*ESA Report on Games Usage*, 2009 & 2019; Pinelle et al., 2008b):

- **Ação:** Este tipo de jogo começou classicamente com os jogos de tipo “*twitch*” – do inglês: contração muscular, pois os jogadores realizavam intensos movimentos físicos em máquinas de fliperama para realizar as ações solicitadas nos jogos. Jogos de ação requerem do jogador reflexos rápidos, lógica para enfrentar desafios, além de coordenação para lidar com diferentes mecânicas. Possuem a característica de apresentar diferentes elementos no jogo em um passo rápido o qual o jogador deve acompanhar para continuar jogando. Por exemplo, temos jogos clássicos como “*Super Mário Bros.*”³, “*Sonic the Hedgehog*”⁴, e outros. Esse tipo inclui também jogos de *side-scrolling* e jogos com labirintos como “*Pac-Man*”⁵, além de jogos de plataforma como “*Gex*”⁶ e jogos de objetos caindo que necessitam ser destruídos como “*Missile Command*”⁷. Além destes, temos jogos como “*Doom*”⁸, “*Quake*”⁹ e “*Duke Nukem*”¹⁰ com seu estilo de jogo de tiro em primeira pessoa contra inimigos pré-definidos e níveis progressivos.
- **Aventura:** Esse é o tipo de jogo em que se deve normalmente explorar um mundo virtual desconhecido em busca de objetos específicos, solucionando quebra-cabeças e problemas lógicos de modo a ser capaz de progredir em uma história específica. Jogos de aventura possuem a característica de possuir um personagem principal (protagonista) o qual se envolve em uma história interativa ao longo do jogo. Alguns exemplos clássicos são, por exemplo: “*Myst*”¹¹, “*The Legend of Zelda: Ocarina of Time*”¹² e “*Zork*”¹³.
- **Jogos de Luta:** Este tipo de jogo é definido por uma batalha corpo a corpo entre dois personagens até que um deles caia ou seja derrotado. Aquele que ficar em pé, ganha. Da mesma forma, é possível considerar jogos nos quais um ou mais personagens batalham

³ Informações adicionais sobre este jogo podem ser encontradas em: [Super Mario Bros.](#) Acessado em: 04 Fevereiro de 2020.

⁴ Informações adicionais sobre este jogo podem ser encontradas em: [Sonic the Hedgehog.](#) Acessado em: 04 Fevereiro de 2020.

⁵ Informações adicionais sobre este jogo podem ser encontradas em: [Pac-Man.](#) Acessado em: 04 Fevereiro de 2020.

⁶ Informações adicionais sobre este jogo podem ser encontradas em: [Gex.](#) Acessado em 4 Fevereiro de 2020.

⁷ Informações adicionais sobre este jogo podem ser encontradas em: [Missile Command.](#) Acessado em: 04 Fevereiro de 2020.

⁸ Informações adicionais sobre este jogo podem ser encontradas em: [Doom.](#) Acessado em: 04 Fevereiro de 2020.

⁹ Informações adicionais sobre este jogo podem ser encontradas em: [Quake.](#) Acessado em: 04 Fevereiro de 2020.

¹⁰ Informações adicionais sobre este jogo podem ser encontradas em: [Duke Nukem.](#) Acessado em: 04 Fevereiro de 2020.

¹¹ Informações adicionais sobre este jogo podem ser encontradas em: [Myst.](#) Acessado em: 04 Fevereiro de 2020.

¹² Informações adicionais sobre este jogo podem ser encontradas em: [The Legend of Zelda: Ocarina of Time.](#) Acessado em: 04 Fevereiro de 2020.

¹³ Informações adicionais sobre este jogo podem ser encontradas em: [Zork.](#) Acessado em: 04 Fevereiro de 2020.

em lutas corpo-a-corpo com diversos outros oponentes. A luta continua até que todos os inimigos sejam derrotados. Esses jogos são definidos por ritmos frenéticos, muita atenção dos seus jogadores em vida e ataques inimigos, diferentes e variadas maneiras de combinar ataques e botões, e pouquíssima história envolvida. São exemplos clássicos desse tipo de jogo “*Mortal Kombat*¹⁴” e “*Street Fighter*¹⁵”.

- **Role Playing Game (RPG):** Jogos de RPG têm a principal característica de que seus usuários possuem o controle de um personagem principal o qual se desenvolve e evolui ao longo da história. Além disso, esses jogos possuem uma característica de serem jogados de acordo com uma história pré-definida por um jogador que define as consequências das ações de outros jogadores. Como exemplos as séries de jogos ‘*Final Fantasy*’¹⁶ ou ‘*Diablo*’¹⁷.
- **Multiplayer online battle arena (MOBA):** é um estilo de jogo no qual os jogadores controlam um personagem principal de um dos dois times competidores, com habilidades e itens, com o objetivo principal de derrotar a base principal inimiga. Esse estilo de jogo foi principalmente derivado do estilo ‘*Action real-time strategy*’ (ARTS). Temos como exemplo o jogo ‘*Defense of the Ancients*’ (DotA)¹⁸.
- **First Player Shooter (FPS):** é um estilo de jogo no qual os jogadores controlam personagens utilizando diversos tipos de armas de fogo e armas brancas ou explosivos, normalmente tendo como objetivo a eliminação do time adversário em rodadas. Além disso, o nome deste tipo de jogo vem do fato de que os jogadores possuem uma câmera fixa na visão em primeira pessoa. Temos ainda alguns jogos que modificam este ângulo de visão para uma visão em terceira pessoa, chamados de *Third Person Shooter* (TPS). Um exemplo muito famoso de jogo FPS é o jogo *Counter-Strike* (CS)¹⁹, e um exemplo de jogo que utiliza as mecânicas de tipos TPS é o jogo *Grand Theft Auto* (GTA)²⁰.
- **Real Time Strategy (RTS):** jogos RTS normalmente são baseados em jogos voltados para a coleta de recursos e o desenvolvimento de bases e unidades através do uso dos recursos coletados de forma online e em tempo real sem turnos. Neste tipo de jogo

¹⁴ Informações adicionais sobre este jogo podem ser encontradas em: [Mortal Kombat](#). Acessado em: 04 Fevereiro de 2020.

¹⁵ Informações adicionais sobre este jogo podem ser encontradas em: [Street Fighter](#). Acessado em: 04 Fevereiro de 2020.

¹⁶ Informações básicas e adicionais sobre o jogo podem ser encontradas em: [Final Fantasy](#). Acessado em: 21 Out. 2018.

¹⁷ Informações adicionais sobre o jogo podem ser encontradas em: [Diablo](#). Acessado em: 21 Out. 2018.

¹⁸ Informações adicionais podem ser encontradas no site do jogo: [DotA 2](#). Acessado em: 20 Out. de 2018.

¹⁹ Informações adicionais sobre o jogo podem ser encontradas em: [Counter-Strike](#). Acessado em: 08 Out. de 2021.

²⁰ Informações adicionais sobre o jogo podem ser encontradas em: [Grand Theft Auto](#). Acessado em: 08 Out. de 2021.

normalmente o objetivo é centrado no jogador que desenvolver a maior base e destruir as bases inimigas. Um exemplo deste tipo de jogo é o jogo *Age of Empires*²¹.

2.1.2 Gêneros de Jogos

Gêneros de jogos são, essencialmente, a característica principal que define como a história do jogo será contada e qual a ambientação específica dela. Gêneros de jogos são derivados basicamente das situações reais do mundo, no qual tais histórias acontecem. Porém, é importante saber que a combinação de gêneros e tipos de jogos é importante, uma vez que certas combinações não conseguiriam chamar a atenção de seus jogadores ou poderiam não ser totalmente aceitos na comunidade. Por exemplo, um jogador pode achar estranho um jogo de dança ambientado num ambiente de construção. Sendo assim, existem muitas dezenas de gêneros de jogos e combinações de gêneros que podem ser criados. Todo e qualquer tipo de situação que defina um ambiente e tipo de história que pode ser contada se transforma em um gênero. Dessa forma, é possível ter jogos definidos dentro de um gênero espacial, com astronautas, naves, mundos diferentes e coisas do espaço não conhecidos. Juntamente com esse gênero pode ser incluído o tema futurístico, definindo objetos e situações ainda irreais para a atualidade. Ainda mais, é possível ter jogos do tipo FPS que podem ser dos gêneros de guerra, zumbis, futurístico, espacial, sobrevivência e muitos outros. Em certas situações, a combinação de algum tipo com algum gênero pode dar surgimento a um novo tipo de jogo especificamente ambientado para situações similares à original que definiu esse tipo. Um exemplo desse acontecimento foi com o jogo '*Defense of the Ancients*' (DotA²²), em que ele foi definido como um jogo RTS (*real time strategy*) e de gênero de defesa da torre em combinação com o gênero surrealista e gênero de campo de batalha de times. A partir disso, foi definido o tipo de jogo que hoje conhecemos como MOBA's, os quais são jogos de RTS especificamente voltados para o objetivo de defesa da torre com a mecânica de batalha em equipe.

2.2 Playability

De acordo com Nacke *et al* (2009), Playability pode ser definida como a análise das interações e relações entre o próprio jogo e seu design. Ou seja, a Playability analisa o design geral dos jogos e como ele influencia os jogadores e suas ações. A Playability se preocupa principalmente

²¹ Informações adicionais sobre o jogo podem ser encontradas no site: [Age of Empires](#). Acessado em: 08 Out. de 2021.

²² Informações adicionais podem ser encontradas no site do jogo: [DotA 2](#). Acessado em: 20 Out. de 2018.

em analisar se as informações desejadas para aquele artefato estão em conformidade com seu design geral e como tais informações se apresentam para o jogo em si.

A avaliação da Playability é importante para que possam ser identificados problemas de interação e design dos jogos durante seu desenvolvimento. Esses problemas, quando tratados, podem aumentar a qualidade geral dos jogos fazendo com que sejam mais bem aceitos no mercado pelo seu público-alvo. Além disso, essas avaliações podem ajudar a identificar problemas de usabilidade na interface dos jogos que podem afetar o entretenimento e a experiência dos usuários com os jogos (Sánchez *et al.*, 2012).

Deve ser observado ainda que a Playability é um conceito de qualidade do software e não apenas um método de avaliação de jogos digitais. Neste sentido, a Playability deve ser considerada como um fator de qualidade do software que identifica a qualidade do design de um jogo em relação as mecânicas e gameplay deste jogo. E sendo assim, os métodos de avaliação existentes apoiam e são utilizados para a medição deste fator de qualidade dos softwares. Ou seja, a Playability é o conceito pelo qual é possível avaliar a qualidade do design de um jogo e como os jogadores entendem e usam este design dentro do jogo.

É possível apontar também como um dos fatores positivos de impacto da avaliação da Playability, a redução de custos com atualizações e correções futuras dos jogos. Isso seria válido para avaliações realizadas nos estágios iniciais de produção do jogo, uma vez que erros implementados em softwares se tornam mais custosos à medida que o projeto avança (Pressman, 2016). Além disso, quando estes defeitos são identificados na fase de implementação do projeto e necessitam de alterações no projeto e design do software para sua correção, este processo pode acarretar em esforço e orçamento até seis vezes maiores para serem modificados (Pressman, 2016). Assim, de fato, o conceito pode ser expandido para uma melhora da qualidade de softwares em geral, os quais se beneficiam dessa característica (Melo, Shull e Travassos, 2001). No caso específico de jogos digitais, é possível supor que o esforço pode ser possivelmente maior ainda, uma vez que um jogo possui elementos gráficos que devem acompanhar código e ações apresentadas a seus usuários.

De acordo com Desurvire e Wiberg (2009), a avaliação da Playability pode ser subdividida em diversas categorias que constituem as características desejadas de um dado jogo. Porém, tais categorias não podem ser aplicadas a qualquer tipo de jogo, uma vez que os jogos podem não possuir as características necessárias para ser avaliado pelas heurísticas definidas em tais categorias. Dessa forma, uma categorização mais simples é desejável, já que é possível abranger mais jogos de diferentes tipos e gêneros.

As principais categorias definidas na literatura são: Usabilidade, Mecânicas, *Game Play*, *Game Story*. A seguir, são apresentadas as definições de cada uma dessas categorias e sua utilização na avaliação da Playability de jogos digitais.

2.2.1 Usabilidade

Segundo Nielsen (1994b), usabilidade é definida como: “um atributo de qualidade que avalia quão fácil uma interface é de se usar” ou “é uma medida de qualidade da experiência de um usuário ao interagir com um produto ou sistema”. Desta forma, Nielsen (1994a) define fatores que influenciam na avaliação da usabilidade de interfaces assim como alguns aspectos que, quando aplicados, ajudam a melhorar a usabilidade dos sistemas.

Além disto, Nielsen (1994a), ainda define que a usabilidade é definida como a medida provinda da relação entre eficácia, eficiência e satisfação subjetiva que um usuário possui durante o manuseio de um sistema.

Segundo os critérios de qualidade de software, norma ISO/IEC 25010:2011 (2011), a usabilidade é entendida como um conjunto de atributos de software que se relacionam com o esforço necessário para seu uso por determinado conjunto de usuários, o qual é definido por suas diferentes características, seus respectivos aspectos, como é mostrado na **Error! Reference source not found.** com os aspectos específicos para a usabilidade de um software.

Porém, para a Playability, a usabilidade é uma característica de avaliação que foca em aspectos específicos de seu conceito completo. Isso acontece porque, em jogos digitais, não se deve aplicar o conceito de nível de facilidade de uso, uma vez que a dificuldade e os desafios são partes integrantes de um jogo. De acordo com as heurísticas de usabilidade de interfaces avaliadas por Pinelle, Wong e Stach (2008) em seu estudo, as características da usabilidade aplicadas na Playability são apresentadas na Tabela 2.1.

Tabela 2.1. Características da usabilidade avaliadas na Playability baseada nas heurísticas avaliadas por Pinelle, Wong e Stach (2008).

Característica	Definida como
Inteligibilidade	Atributos do software que evidenciam o esforço do usuário para reconhecer o conceito lógico e sua aplicabilidade.
Operacionalidade	Atributos do software que evidenciam o esforço do usuário para sua operação e controle de sua operação.
Conformidade	Atributos do software que fazem com que o software esteja de acordo com as normas, convenções ou regulamentações previstas em leis e descrições similares encontradas em sistemas semelhantes, relacionadas à aplicação.

Encontro das informações	Relacionada com o esforço necessário de um usuário na tarefa de se encontrar uma informação específica dentro do sistema, seja ela em suas configurações ou durante seu uso (auxiliado por arquitetura de informação, categoria de nomes, navegação, links, rotulação).
Entendimento das informações	Relacionado especificamente com o tipo de linguagem utilizada na descrição das informações, seja ela técnica ou não, de modo que qualquer tipo de usuário possa entender e assimilar tais informações.

Nielsen (1994b) e a ISO/IEC 25010:2011 (2011) definem características de qualidade de software relacionadas a Usabilidade de um software qualquer através das características a seguir: Adequação funcional, Eficiência de desempenho, Compatibilidade, Usabilidade, Confiabilidade, Segurança, Manutenção e Portabilidade. Sendo assim, Pinelle, Wong e Stach (2008), em seu estudo, se baseiam nestas características de qualidade de softwares de modo que possam avaliar diretamente o design dos jogos e as atividades realizadas para interação com o sistema.

2.2.2 Mecânicas

Quando associamos a mecânica a jogos digitais estamos falando de um conceito que vai além dos controles físicos utilizados por seus jogadores. No contexto de Playability, a mecânica dos jogos pode ser entendida como elementos escolhidos para explorar motivações, interesses e desejos dos jogadores. Além disso, a mecânica de um jogo digital está associada ao tema e pode ser alterada de acordo com a temática do jogo no qual está inserida. Ou seja, existe um padrão comum para os controles físicos utilizados em diferentes temáticas e gêneros de jogos (Bjork e Holopainen, 2004).

Além disso, a mecânica avaliada pela Playability vem associada com o conceito de aprimorar a interação do jogador com as características essenciais do jogo, da mesma forma que a mecânica está ligada com a temática geral do jogo. Por exemplo, as configurações de um jogo podem afetar o modo como um jogador interage com ele e, dessa forma, até mesmo os controles pré-definidos necessitam de um meio para que possam ser adaptados para cada jogador (Korhonen e Koivisto, 2006).

2.2.3 Game Play

Game Play é a estrutura de interações, informações e ações desenvolvidas durante um jogo por seus jogadores além das interações existentes entre eles (Bjork e Holopainen, 2004). Dessa forma, a *Game Play* se dedica à avaliação do design projetado para os jogos e de como consegue

interagir com seus usuários, além de como eles interagem entre si. Além disso, de acordo com seu conceito, a *Game Play* avalia como os objetivos, primários e secundários, relacionam-se e como são projetados para dar sentido às ações realizadas pelos jogadores ou para a história do jogo avaliado.

Na Playability, geralmente a avaliação da *Game Play* engloba também a avaliação da *Game Story*. Isso pode ser observado em diversas pesquisas realizadas nos últimos anos (Korhonen e Koivisto, 2006; Barcelos *et al.*, 2011). Para a avaliação da *Game Story*, temos a avaliação da história do jogo e seus personagens, de como eles interagem com os usuários para que seja construído um ambiente o mais realista possível. Além disso, a *Game Story* avalia como os objetivos, primários e secundários, relacionam-se com a história e como suportam as ações dos jogadores em relação à história apresentada. No entanto, geralmente essa avaliação é integrada com a avaliação da *Game Play*, de forma a simplesmente expandir sua área de avaliação geral.

2.3 Mapeamento Sistemático

2.3.1 Metodologia de pesquisa

Para que fosse possível propor uma nova técnica de avaliação da Playability em jogos digitais, foi necessário entender e caracterizar as soluções existentes na literatura, identificando potenciais lacunas, defeitos e qualidade, além de criar o devido embasamento do trabalho a ser desenvolvido sem criar soluções similares às pré-existentes. Com esse objetivo foi desenvolvido um mapeamento sistemático da literatura.

Segundo Kitchenham e Charters (2007) uma revisão sistemática da literatura, ou apenas revisão sistemática, é indicado como um método para a avaliação, caracterização, identificação e interpretação das pesquisas relevantes e disponíveis com relação a um tópico ou questão de pesquisa específica. Porém, no escopo desta pesquisa, foi realizado um Mapeamento Sistemático (MS). Um Mapeamento Sistemático (MS), segundo Kitchenham e Charters (2007), é um tipo de revisão sistemática da literatura, porém, apresentando uma ampla visão da área de pesquisa ou questão aberta de pesquisa, assim estabelecendo a existência de evidências a uma questão de pesquisa.

2.3.2 Protocolo do mapeamento sistemático

Nas próximas seções podem ser encontradas as informações condizentes com o protocolo adotado para realização do mapeamento sistemático desta pesquisa, sendo ele baseado no processo definido por Kitchenham e Charters (2007).

2.3.2.1 Definição de objetivo e questão de pesquisa

Primeiramente foi definido o objetivo do Mapeamento Sistemático formalmente através do paradigma GQM (*Goal-Question-Metric*) definido por Basili e Rombach (1988). O objetivo definido para esta pesquisa é apresentado conforme ilustra a Tabela 2.2.

Tabela 2.2. Objetivo de pesquisa do Mapeamento Sistemático de acordo com o paradigma GQM.

Analisar	Trabalhos e pesquisas publicadas em meios científicos através de um estudo baseado em mapeamento sistemático
Com o propósito de	Identificar conjuntos heurísticos, <i>guidelines</i> , métodos, técnicas e ferramentas
Em relação a	A avaliação da Playability
Do ponto de vista de	Pesquisadores e Desenvolvedores
No contexto de	Jogos digitais

De acordo com objetivo de pesquisa, foi também definida a questão de pesquisa para o Mapeamento Sistemático desenvolvido. A questão principal de pesquisa foi definida como:

RQ1. Quais os conjuntos heurísticos, *guidelines*, metodologias, técnicas e ferramentas foram desenvolvidas e utilizadas para a avaliação da Playability em jogos digitais?

Conjuntos heurísticos são, neste escopo, definidos como todo e qualquer conjunto de heurísticas que definem os principais defeitos ou diretrizes de avaliação que podem ser encontrados em jogos digitais em relação às características da sua Playability. Da mesma forma, *guidelines* nesse contexto podem ser consideradas as pesquisas que definem conjuntos de regras e diretrizes para o **desenvolvimento** de jogos digitais em relação a sua usabilidade e Playability considerando a aplicação de boas práticas e seu aprimoramento.

O termo “**métodos**”, ou método de avaliação, foi utilizado nesta pesquisa para identificar trabalhos da literatura que apresentem um conjunto de tarefas, ou procedimentos, que funcionam como guias aos usuários para identificar e obter resultados de uma medida específica a um produto ou parte de um produto (ISO/IEC/IEEE 24765:2017, 2017, p. 168). Ou seja, indo de acordo com a definição de Falbo, R. (1998), temos que um método é um processo sistemático que define passos e heurísticas para a execução de atividades específicas através de um conjunto de sub-atividades que satisfaçam uma, ou mais, atividades.

Para esse escopo, é definida uma **técnica** como qualquer conjunto de passos que apresente uma forma de realizar uma atividade específica sem que exista a definição de sub-atividades que satisfaçam a conclusão da primeira, por exemplo, uma entrevista estruturada

pode ser considerada uma técnica ou uma avaliação estruturada utilizando conjuntos heurísticos (Falbo, 1998).

Adicionalmente, uma **ferramenta** é um software que representa um apoio significativo no processo de aplicar um método específico para tarefas definidas pelo uso de um conjunto de técnicas tendo o objetivo de avaliar um produto específico.

Por fim, o mapeamento sistemático desenvolvido buscou responder a três questões secundárias com o objetivo de avaliar mais profundamente os aspectos de cada uma das soluções apresentadas por diferentes autores:

SQ1. Por qual público alvo de aplicadores/avaliadores que as técnicas propostas são intencionadas para serem utilizadas?

SQ2. As soluções propostas são direcionadas para qual escopo (em relação a tipos, gêneros e contextos gerais dos jogos)?

SQ3. A solução proposta pode ser configurada para a avaliação de diferentes jogos digitais, possui contexto generalizado, ou é focada na avaliação de um específico subtipo de jogos, tipo e gênero específicos?

2.3.2.2 Estratégia de busca das publicações

Escopo da pesquisa: A pesquisa foi realizada na biblioteca digital *Scopus*, uma vez que essa biblioteca digital permite a utilização de expressões lógicas na *string* de busca automática. Além disso, ela permite a busca através apenas de *abstract* e título além de possuir uma grande gama de áreas e subáreas de pesquisa, disponibilizando a recuperação de textos completos. Kitchenham e Charters (2007) afirmaram que é a maior base de dados de indexação de resumos e citações científicas. Além disso, a biblioteca digital *Scopus* abrange uma grande parte de outras bibliotecas digitais além de conferências e periódicos os quais publicam em mais de uma base digital. Para o contexto desta pesquisa a biblioteca digital *Scopus* foi a única biblioteca digital utilizada pois todos os artigos base desejados foram retornados pela busca na biblioteca digital. Com o objetivo de mitigar o número de artigos relacionados não retornados pela biblioteca, um processo de *snowballing* foi realizado nos artigos retornados pela biblioteca.

Idiomas dos artigos: Os idiomas adotados pela pesquisa foram o Inglês, por ser o maior idioma utilizado para a publicação de artigos científicos, e o português, por se tratar do idioma nativo utilizado nesta pesquisa e pelo interesse de identificar a situação atual da área em conferências nacionais.

Definição da *String* de busca: Para a definição dos termos de busca da *string* de busca, foi utilizada a definição pelos termos do PICO (*Population, Intervention, Comparison e Output*) e os termos relacionados a eles, quando aplicável (veja Tabela 2.3).

Tabela 2.3. Termos utilizados para os parâmetros do PICO

Parâmetro	Termos de busca utilizados
(P) <i>Population</i> : Trabalhos publicados em conferências ou periódicos (<i>peer-reviewed</i>) relatando conjuntos heurísticos, <i>guidelines</i> , metodologias, técnicas e ferramentas para a avaliação da Playability de jogos digitais.	<i>“playability”</i> OR <i>“usability”</i> OR <i>“player experience”</i> OR <i>“game usability”</i> OR <i>“play testing”</i>
(I) <i>Intervention</i> : tecnologias utilizadas para a avaliação e teste de jogos digitais e processos de desenvolvimento destes softwares.	<i>“measurement”</i> OR <i>“evaluat*”</i> OR <i>“heuristic”</i> OR <i>“sets”</i> OR <i>“method”</i> OR <i>“assess”</i> OR <i>“assessment”</i> OR <i>“technique”</i> OR <i>“framework”</i> OR <i>“model”</i> OR <i>“inspection”</i>
(C) <i>Comparison</i> : Não se aplica, pois este é um mapeamento sistemático.	-
(O) <i>Output</i> : Soluções de avaliação e desenvolvimento de jogos digitais.	<i>“computer games”</i> OR <i>“digital games”</i> OR <i>“game design*”</i> OR <i>“game evaluation”</i> OR <i>“games”</i> OR <i>“evaluation method”</i> OR <i>“evaluation technique”</i> OR <i>“evaluation model”</i>

Dessa forma, a *string* de busca é definida através dos termos descritos na Tabela 2.3 utilizando os seus termos em inglês uma vez que a biblioteca digital selecionada utiliza essa linguagem para realizar a busca nos termos das publicações. A *string* de busca é apresentada na figura abaixo:

“playability” OR *“usability”* OR *“player experience”* OR *“game usability”* OR *“play testing”*) AND (*“measurement”* OR *“evaluat*”* OR *“heuristic”* OR *“sets”* OR *“method”* OR *“assess”* OR *“assessment”* OR *“technique”* OR *“framework”* OR *“model”* OR *“inspection”*) AND (*“computer games”* OR *“digital games”* OR *“game design*”* OR *“game evaluation”* OR *“games”* OR *“evaluation method”* OR *“evaluation technique”* OR *“evaluation model”*)

Figura 2.1. String de busca utilizada na biblioteca selecionada.

A string de busca desenvolvida foi executada em dois momentos diferentes. A primeira vez que a string de busca foi executada foi no período de Setembro e Outubro de 2016, onde a biblioteca Scopus retornou 97 artigos no total, apresentando todos os artigos de referência desejados. Posteriormente, durante a pesquisa, uma atualização deste mapeamento foi realizado onde a string de busca foi executada no período de Outubro à Novembro de 2018, onde a biblioteca Scopus retornou 48 artigos, completando os 145 artigos apresentados em sua totalidade neste mapeamento sistemático.

Além da realização de uma busca através de biblioteca digital, foi realizada uma busca de publicações através da realização de um *snowballing* (Wohlin, 2014). A técnica de *snowballing* é uma técnica com a qual, a partir de artigos base, um pesquisador analisa as publicações referenciadas por tais artigos, em busca de publicações que possam ter relação com o objetivo da pesquisa ou com o assunto abordado pelo mapeamento desenvolvido. Nesse caso, apenas um nível de *backward snowballing* foi executado, de forma a pesquisar novas informações nas referências dos artigos retornados pela busca na biblioteca digital. Os trabalhos foram recolhidos para estudo a partir do Google Scholar. Os artigos identificados a partir do processo de *snowballing* foram avaliados e passaram por todos os filtros do mapeamento sistemático, da mesma forma que os artigos coletados na fase de busca pela biblioteca digital. O processo de *snowballing* foi realizado após a segunda execução do processo de pesquisa de artigos na biblioteca digital *Scopus* com string de busca.

2.3.2.3 Critérios de seleção de artigos

Durante o processo de mapeamento sistemático, é desejado que sejam selecionadas, identificadas e avaliadas as publicações que sejam importantes ou relevantes à questão de pesquisa definida para o processo realizado. Kitchenham e Charters (2007) sugerem a definição de critérios para inclusão e exclusão de publicações de acordo com as questões de pesquisa definidas para o mapeamento em questão. Tais critérios, de acordo com as questões de pesquisa definidas, devem ser suficientes para excluir a maior parte dos trabalhos que, a partir de uma leitura inicial, não sejam compatíveis com o assunto abordado pelas questões definidas. Para este mapeamento em questão, os critérios definidos são apresentados na Tabela 2.4 a seguir.

Tabela 2.4. Critérios de inclusão e exclusão para o mapeamento desenvolvido.

Tipo	Identificador	Descrição
Critérios de Inclusão	[Inc1]	O artigo descreve métodos de avaliação ou medição da Playability.
	[Inc2]	O artigo descreve características para avaliação da Playability em jogos digitais.
	[Inc3]	O artigo descreve fatores de influência para a medição ou avaliação da Playability.
	[Inc4]	O artigo fundamenta/explica/apresenta diferenças entre a usabilidade e a Playability no processo de desenvolvimento de jogos digitais.
	[Inc5]	O artigo descreve estudos experimentais que avaliem métodos de avaliação da Playability ou de desenvolvimento de jogos inclusivos com Playability.

Tipo	Identificador	Descrição
	[Inc6]	O artigo compara métodos de avaliação ou medição da Playability em jogos digitais.
Critério de Exclusão	[Exc1]	O artigo não está em Inglês ou Português.
	[Exc2]	O artigo não é gratuitamente disponível on-line para download.
	[Exc3]	O artigo não está limitado ao escopo de jogos digitais.
	[Exc4]	O artigo não apresenta definições, aplicações ou métodos de medição e/ou avaliação da Playability.
	[Exc5]	Não é um artigo ou periódico “ <i>peer reviewd</i> ”.
	[Exc6]	O artigo não atende a nenhum dos critérios de inclusão.

2.3.2.4 Procedimentos para seleção de artigos

De forma a sistematizar o processo de seleção de artigos a partir da *string* de busca definida, foi executado um processo de quatro etapas: (1) *execução da busca*, (2) *primeiro filtro* (3) *segundo filtro*, (4) *execução de snowballing*. Para isso na execução da busca por publicações, o pesquisador responsável pelo mapeamento utilizou a *string* de busca definida anteriormente na biblioteca digital selecionada e armazenou o conjunto de referências retornadas pela biblioteca na ferramenta Start²³. A ferramenta utilizada indica automaticamente as entradas duplicadas e o pesquisador responsável pela execução do mapeamento sistemático pode marcar ou excluir os artigos duplicados, para não serem avaliados no primeiro filtro de avaliação.

Após a realização da busca por publicações através da *string* de busca definida, os artigos foram avaliados apenas através do título de suas referências para identificação de possíveis publicações que não seriam relacionadas pela busca na biblioteca digital. Os artigos os quais os pesquisadores julgassem relacionados com a pesquisa em questão seriam selecionados e incluídos manualmente na ferramenta Start.

A partir da lista de artigos retornados pela biblioteca digital, foi realizado o primeiro filtro em que o pesquisador responsável pelo mapeamento realizou a leitura de título e abstract de cada publicação, classificando-as de acordo com os critérios de inclusão e exclusão definidos. A decisão de inclusão ou exclusão era registrada pelo avaliador na ferramenta Start. Quando existia algum tipo de incerteza quanto ao conteúdo do artigo ser relevante ou não, a decisão sempre seria de incluir para maior avaliação durante o segundo filtro. Após a finalização desse primeiro filtro, o pesquisador responsável pelo mapeamento sistemático pesquisou nas referências de todos os artigos que foram aceitos inicialmente trabalhos que não

²³ Start: ferramenta de apoio ao planejamento e execução de revisões sistemáticas. Disponível em: http://lapes.dc.ufscar.br/tools/start_tool

havia sido identificados pela busca na biblioteca digital. Durante essa pesquisa, apenas o título das referências foi lido para decidir quais artigos deveriam ser melhor explorados. Quando em dúvida sobre a inclusão ou exclusão, o pesquisador sempre optou pela sua inclusão.

Durante o segundo filtro, o pesquisador responsável pela realização do mapeamento realizaria a leitura completa de todos os artigos incluídos a partir do primeiro filtro novamente, além dos artigos incluídos a partir da técnica de *snowballing*, analisando os critérios de inclusão e exclusão deles após leitura completa. Para cada exclusão, o pesquisador deveria indicar os critérios que justificavam essa decisão na ferramenta, indicando a parte do conteúdo que não condizia com a pesquisa. Ao final desse segundo filtro, uma base de conhecimento em técnicas, metodologias e ferramentas de avaliação da Playability foi formada.

2.3.3 Condução do mapeamento sistemático

Os resultados quantitativos obtidos durante a execução da busca automática na biblioteca selecionada são apresentados na Tabela 2.5 abaixo, juntamente com os resultados quantitativos da realização do *snowballing* a partir dos artigos encontrados na busca pela biblioteca digital, sendo dividido de acordo com os filtros realizados.

Como resultado do processo de mapeamento sistemático, a busca automatizada por string de busca na biblioteca digital *Scopus*, retornou 145 artigos, dos quais 2 eram duplicados e foram retirados da avaliação do primeiro filtro. No primeiro filtro, todas as 143 publicações foram analisadas lendo seus títulos e abstracts e, ao final desse filtro, 31 publicações foram selecionadas para o segundo filtro. Nesse momento, os títulos das referências das 31 publicações foram lidas de modo a identificar estudos que não foram identificados pela busca automática de acordo com o processo de *backward snowballing*. Após essa leitura de títulos, 25 publicações foram selecionadas para passarem por um processo de primeiro filtro, avaliando, além de seus títulos, os abstracts das publicações. Após o primeiro filtro dessas publicações, 9 artigos foram selecionados para serem avaliados no segundo filtro. No segundo filtro, 40 publicações foram analisadas, lidos seus textos completos de modo a identificar a base de conhecimento desta pesquisa. Após o segundo filtro, 16 artigos foram selecionados dos quais 11 eram provenientes da análise automática com *string* de busca; e 5 eram provenientes do *snowballing* realizado.

Tabela 2.5. Número de publicações encontradas a partir da biblioteca digital e *snowballing*.

Método de busca	Total	Duplicados	Publicações analisadas 1º filtro	Publicações analisadas 2º filtro	Publicações selecionadas
Busca por string (Biblioteca digital - <i>Scopus</i>)	145	2	143	31	11
<i>Snowballing</i>	25	0	25	9	5

Os artigos selecionados ao final do segundo filtro formam a nossa base de conhecimento em técnicas, métodos e ferramentas para a avaliação da Playability de jogos digitais. Todos os artigos selecionados e suas referências estão disponíveis na Tabela 2.6 abaixo:

Tabela 2.6. Artigos selecionados após o segundo filtro do processo de MS desenvolvido divididos entre publicações da pesquisa automática e por *snowballing*.

Título da publicação	Autores	Ano
Publicações provenientes da pesquisa automática por String		
NExPlay: Playability assessment for non-experts evaluators	Manzoni, F. S.; Ferreira, B. M. and Conte, T. U.	2018
Developing playability heuristics for computer games from online reviews	Zhu, M. and Fang, X.	2014
Methods for game user research: Studying player behavior to enhance game design	Desurvire, H. and El-Nasr, M.S.	2013
The explanatory power of playability heuristics	Korhonen, H.	2011
Comparison of playtesting and expert review methods in mobile game evaluation	Korhonen, H.	2010
Expert review method in game evaluations: Comparison of two playability heuristic sets	Korhonen, H.; Paavilainen, J. and Saarenpää, H.	2009
Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration	Desurvire, H. and Wiberg, C.	2009
Usability, playability, and long-term engagement in computer games	Febretti, A. and Garzotto, F.	2009
Playability heuristics for mobile multi-player games	Korhonen, H. and Koivisto, E. M. I.	2007

Título da publicação	Autores	Ano
Playability heuristics for mobile games	Korhonen, H. and Koivisto, E. M. I.	2006
Using Heuristics to Evaluate the Playability of games	Desurvire, H.; Caplan, M. and Toth, J. A.	2004
Publicações provenientes do <i>Snowballing</i>		
Análise comparativa de heurísticas para avaliação de jogos digitais.	Barcelos, T. S.; Carvalho, T.; Schimiguel, J. and Silveira, I. F.	2011
Game development software engineering process life cycle: A systematic review	Aleem, S.; Capretz, L. F. and Ahmed, F.	2016
Software engineering processes in game development: A survey about Brazilian developers	Politowski, C.; De Vargas, D.; Fontoura, L. and Foletto, A.	2016
Informal HCI: Fixing Playability Issues as a Strategy to Improve the Skills of Novice Programmers	Barcelos, T. S.; Costa, G. C.; Muñoz, R. and Silveira, I. F.	2014
Expert vs novice evaluators - Comparison of heuristic evaluation assessment	Borys, M. and Laskowski, M.	2014

Na seção a seguir, serão apresentados os métodos para avaliação da Playability identificados, estudados e analisados de acordo com os resultados identificados no mapeamento sistemático desenvolvido. Além disso, outros trabalhos, que foram identificados na literatura durante a execução desta pesquisa, são descritos, demonstrados e utilizados para descrever os conceitos necessários a ela, incrementando o corpo de conhecimento utilizado nela.

2.3.4 Métodos para avaliação da Playability

A avaliação da Playability é um tipo de avaliação de jogos que segue regras definidas, da mesma forma que podem ser encontradas na avaliação da usabilidade, uma vez que, esses dois tipos de avaliação são relacionados com a qualidade do software, ou jogo, em questão. Sendo assim, da mesma forma como na usabilidade, é possível avaliar a Playability através de diferentes métodos. Em usabilidade, é comum classificarmos os métodos de avaliação em: métodos de investigação, métodos de inspeção e métodos de observação de uso (Barbosa e Silva, 2010). Esses métodos são descritos abaixo:

- **Métodos de investigação** permitem ao avaliador ter acesso, interpretar e analisar concepções, opiniões perspectivas e comportamentos dos usuários de um sistema

específico. Isso pode ser feito com a utilização de questionários, entrevistas, grupos focais e estudos de campo.

- **Métodos de inspeção** permitem ao avaliador examinar soluções de IHC para tentar antever as possíveis consequências que certas decisões de Design podem causar à experiência do usuário. De forma mais direta, pretende identificar problemas que possam existir durante a utilização de um sistema por qualquer usuário. Contudo, esse tipo de método não envolve usuários diretamente, o que o torna uma avaliação de possíveis experiências e não situações reais de uso (Barbosa e Silva, 2010). Entre os métodos de inspeção estão a avaliação heurística e percurso cognitivo.
- **Métodos de observação** fornecem informações sobre situações em que os usuários realizam atividades com ou sem o apoio de sistemas interativos. Nesse caso, observamos problemas de uso reais os quais são identificados através da observação das ações realizadas pelos usuários durante suas atividades (Barbosa e Silva, 2010). Dentre os métodos de observação temos testes de usabilidade, *user testing* (quando o jogador é observado e suas atividades avaliadas) e prototipação em papel.

Entre todos os trabalhos da literatura pesquisados e revisados durante esta pesquisa, foram identificados a utilização de métodos de observação, através de estudos de *user testing*, métodos de investigação, em que entrevistas e questionários foram utilizados e métodos de inspeção avaliações através da utilização de conjuntos heurísticos para a avaliação da Playability. A utilização de métodos de inspeção podem ser capazes de apoiar o objetivo geral desta pesquisa de diminuição de custos para o desenvolvimento de jogos, uma vez que esse tipo de técnica pode ser aplicado mais cedo no desenvolvimento. Essa capacidade de aplicação da técnica em fases iniciais deve-se à capacidade de inspecionar o jogo de forma a antever as possíveis consequências de interações realizadas através da geração de experiências potenciais de uso (Barbosa e Silva, 2010). A seguir são apresentadas algumas técnicas para a avaliação da Playability em jogos digitais identificadas durante o MSL e, na seção 2.3.4.2.2, são apresentados diferentes estudos que pesquisaram alguns diferentes pontos no processo de avaliação da Playability ou que identificaram pontos de melhorias e defeitos existentes atualmente.

2.3.4.1 Questionários e Entrevistas

O Questionário é uma técnica simples e de fácil aplicação a qual pode descrever opiniões dos usuários por meio impresso ou on-line. Questionários podem descrever perguntas de forma

aberta ou fechada, porém, o mais comum são as perguntas fechadas de modo a agilizar as respostas dadas pelos usuários podendo-se contabilizá-las de forma simples também. Por isso, pode ser utilizado na avaliação da Playability de maneira a complementar as opiniões dos jogadores e dos avaliadores em como os jogos se comportam e em quais pontos necessitam de maior atenção.

Tanto para os Questionários quanto para as Entrevistas, as perguntas devem ser formuladas de forma cautelosa, evitando ambiguidades ou mal entendidos entre as partes, assim como não influenciar respostas. Questionários, assim como Entrevistas, podem ser utilizados de forma simples para a coleta de informações sobre necessidades dos usuários, dificuldades com a técnica aplicada, mensuração da qualidade e satisfação de uso, entre outros. Porém, questionários, por geralmente utilizarem questões fechadas, podem não fornecer informações suficientes para a análise de resultados desejada. Dentro dos resultados analisados do MSL, nenhum resultado específico foi identificado o qual definisse ou propusesse diferentes maneiras de execução dessa técnica. Porém, em vários artigos foi identificado que os autores utilizavam a técnica de questionários e entrevistas para avaliação final das técnicas aplicadas nas avaliações dos jogos (Korhonen e Koivisto, 2016; Korhonen e Koivisto, 2017; Barcelos et al., 2011; Desurvire e El-Nasr, 2013; Desurvire e Wiberg, 2009). Além disso, não foram identificados trabalhos que utilizassem as técnicas de questionários e entrevistas para a avaliação de jogos e avaliação da percepção dos jogadores em si.

2.3.4.2 Avaliação da Playability através de Conjuntos Heurísticos

A Avaliação Heurística é um método de inspeção utilizado para identificação de problemas em um determinado software através do uso de diretrizes as quais descrevem características e comportamentos desejáveis da interação e interface. A técnica foi inicialmente proposta por Nielsen e Molich (1990) que chamam essas diretrizes de heurísticas, as quais podem ser definidas para qualquer tipo de avaliação.

De acordo com Nielsen e Molich (1990), as interfaces devem ser percorridas pelo menos duas vezes, de modo que, na primeira vez, os avaliadores ganhe uma visão geral do conjunto, e, na segunda vez, os avaliadores possam analisar cautelosamente cada elemento presente na tela. Durante essa fase de avaliação da interface, os avaliadores devem anotar as discrepâncias identificadas, descrevendo a heurística violada, o local em que a discrepância se encontra e uma justificativa para ter sido considerada discrepância, indicando também sua severidade.

Porém, ao longo dos anos essa técnica foi modificada para ser utilizada de diferentes formas por diferentes grupos de pessoas e tipos de avaliação que não a própria avaliação da usabilidade de interfaces. Assim como certas fases e atividades definidas por Nielsen e Molich (1990) foram modificadas e adaptadas para melhor avaliar diferentes situações e softwares. Nesse sentido, a avaliação realizada para a Playability utilizando heurísticas é utilizada de maneira modificada, de modo a apoiar a avaliação feita em jogos digitais e focando nos aspectos necessários e desejados pela Playability ao invés da usabilidade.

2.3.4.2.1 Conjuntos Heurísticos para Avaliação da Jogabilidade identificados através do MSL

Barcelos *et al.* (2011) definem em seu trabalho um conjunto heurístico para a avaliação da Playability de jogos digitais com escopo amplo de tipos e gêneros. Na definição de seu conjunto, os autores não definiram um escopo específico para sua utilização. Sendo assim, o conjunto pode ser utilizado para a avaliação de uma grande variedade de tipos e gêneros. Os autores ainda definiram que o principal objetivo com o desenvolvimento de seu conjunto heurístico seria o de desenvolver um conjunto heurístico reduzido, uma vez que o número de heurísticas, comumente identificado na literatura é de 30 a 40 heurísticas por conjunto. Os autores desejavam identificar a existência de uma diferença estatística entre a utilização de conjuntos reduzidos e os conjuntos extensos existentes na literatura em relação à eficiência e eficácia. O estudo experimental desenvolvido por eles foi feito de forma comparativa com o conjunto desenvolvido por Federoff (2002), além de avaliação e identificação das heurísticas existentes na literatura. Além disso, os resultados apresentados pelos autores demonstraram não existir uma diferença estatística significativa entre a utilização de conjuntos reduzidos ou não em termos de eficiência e eficácia (defeitos por tempo e defeitos reais, excluindo falso-positivos). Por fim, os resultados mostraram que o conjunto heurístico definido por eles era viável para a avaliação de jogos digitais em contexto geral, uma vez que os resultados comparativos eram semelhantes.

Um conjunto heurístico para a avaliação generalizada de jogos digitais, sem especificação de tipos e gêneros chamado de *Heuristic Evaluation for Playability (HEP)* é definido no trabalho desenvolvido por Desurvire, Caplan e Toth (2004). O conjunto definido pelos autores foi desenvolvido de forma a analisar e identificar as heurísticas existentes na literatura para a definição de um conjunto heurístico mais completo e com categorias que pudessem efetivamente identificar as diferenças entre tais categorias durante a avaliação. Os autores realizaram um estudo experimental em que o conjunto heurístico desenvolvido foi

comparado com a metodologia de *user testing*, uma técnica mais simples e comumente utilizada para a avaliação de softwares em geral. De acordo com os resultados apresentados por eles, o conjunto definido obteve um melhor desempenho em relação a número de defeitos e avaliação qualitativa dos seus usuários.

Desurvire e Wiberg (2008) definiram em seu trabalho um *guideline* com um conjunto de heurísticas voltadas à boa prática de desenvolvimento de jogos digitais chamado de *Game Approachability Principles (GAP)*. Este *guideline* foi desenvolvido com o objetivo de apoiar os desenvolvedores de jogos nas fases iniciais de desenvolvimento com um conjunto de boas práticas. De modo mais específico, o *guideline* desenvolvido por eles é voltado para a melhoria de fases iniciais dos jogos, nas quais são apresentadas as mecânicas de jogo por exemplo. A *GAP* foi definida através da avaliação por questionários aplicados a jogadores comuns durante uma feira de jogos eletrônicos. As questões variavam sobre os tipos de jogos mais utilizados e suas principais características necessárias. Além disso, a *GAP* foi avaliada durante o desenvolvimento de um jogo para um maior entendimento de sua cobertura no desenvolvimento de jogos e de sua aceitação pelos usuários.

No trabalho de Desurvire e Wiberg (2009), o conjunto de heurísticas para a avaliação de jogos digitais chamado *PLAY (Heuristics of Playability)* foi definido. De acordo com os resultados apresentados pelo seu conjunto heurístico anterior (HEP; Desurvire, Caplan e Toth, 2004) e os avanços nos conjuntos heurísticos para avaliação da Playability, assim como a partir dos resultados identificados em seu trabalho anterior que definiu a *GAP* (Desurvire e Wiberg, 2008), os autores desenvolveram e propuseram um novo conjunto heurístico baseado na *HEP*. A principal alteração realizada foi a inclusão de diversas heurísticas que realizavam a avaliação de aspectos relacionados à *Player Experience (PX)* de jogos digitais assim como da usabilidade das interfaces de jogos. Neste contexto, a *Player Experience* é relacionada diretamente com a identificação e avaliação dos níveis de diversão dos jogadores durante as seções além da avaliação de quanto os jogadores mantem os seus níveis de interesse e foco durante as seções (Aker *et al.*, 2020). E, para se adaptar ao grande número de heurísticas adicionadas ao conjunto, os autores também definiram uma série de categorias e subcategorias para a separação das heurísticas de modo a facilitar sua localização. A *PLAY*, de acordo com os resultados quantitativos e qualitativos apresentados pelo trabalho, é capaz de eficientemente avaliar jogos digitais de escopo geral, identificando a maioria dos defeitos da Playability presente neles. Porém, o conjunto foi identificado como extenso e de difícil compreensão por seus avaliadores por conta da quantidade de categorias que subdividiam as heurísticas sem apresentar inicialmente uma explicação destas categorias e como utilizá-las.

Korhonen e Koivisto (2006) definiram e apresentaram um conjunto heurístico para a avaliação de jogos digitais de escopo específico para jogos móveis com o nome de *Mobile Game Playability Heuristics*. Os autores especificaram que, para o contexto móvel de jogos digitais, existem diversas características únicas que os conjuntos heurísticos definidos pela literatura não são capazes de identificar. Dessa forma, o conjunto heurístico definido pelos autores é a união da análise e identificação das heurísticas definidas pela literatura, juntamente com as heurísticas especificamente definidas para a avaliação das características móveis de jogos digitais. De acordo com os resultados apresentados por eles, em seu estudo experimental para verificar a viabilidade do conjunto heurístico, o conjunto definido foi uma opção viável para a avaliação de jogos digitais móveis de gênero e tipo com escopo aberto, em que a eficiência do conjunto era comparável com os conjuntos definidos na literatura.

Em uma continuação ao conjunto heurístico apresentado anteriormente por Korhonen e Koivisto (2006) foram incluídos aspectos *multiplayer* no conjunto de heurísticas foi apresentado no trabalho de Korhonen e Koivisto (2007). De acordo com os autores, existe uma lacuna nos conjuntos heurísticos da literatura que não possuem a capacidade de identificar defeitos no escopo de jogos digitais móveis com funções multijogadores. De acordo com os defeitos apontados pelos autores, o conjunto heurístico definido é baseado no conjunto definido em seu trabalho anterior (Korhonen e Koivisto, 2006) no qual realizam a inclusão de heurísticas para a identificação e avaliação das características multijogador em jogos digitais móveis de tipo e gênero com escopo amplo.

Pinelle, Wong e Stach (2008a) apresentam em seu trabalho um conjunto heurístico para a avaliação da Playability o qual é definido através da avaliação da usabilidade de interface de diversos jogos digitais. Com ele, os autores identificaram os principais defeitos da usabilidade de interfaces dos jogos, relacionando-os com os conceitos da Playability de jogos digitais. A partir da definição das heurísticas, os autores realizaram um estudo de viabilidade no qual puderam identificar a eficiência do conjunto heurístico para a avaliação de defeitos da Playability em jogos digitais.

Um conjunto heurístico desenvolvido a partir de seu trabalho anterior (Pinelle, Wong e Stach, 2008a) é apresentado em Pinelle *et al.* (2009). Nele, os autores identificaram a necessidade de desenvolver um conjunto heurístico para a avaliação de jogos digitais baseados em conexão contínua com a rede. De acordo com os autores, existia uma limitação nos conjuntos heurísticos da literatura de forma que tais conjuntos não poderiam identificar os defeitos específicos de jogos com conexão constante a rede de dados. Os autores identificaram as heurísticas que, de acordo com os trabalhos da literatura, cobriam tais necessidades e, sendo

assim, incluíram tais heurísticas em seu conjunto heurístico anteriormente apresentado. Esse novo conjunto heurístico apresentado nesse trabalho é chamado de *Network Game Heuristics (NGH)*. Os resultados de viabilidade apresentados identificam o conjunto NGH como um conjunto viável para a avaliação de jogos nesse escopo específico, de acordo com o número de defeitos identificados quando o conjunto NGH é utilizado em avaliações da Playability.

2.3.4.2.2 Estudos de avaliação e definições da Playability

Aleem *et al.* (2016) desenvolveram uma revisão sistemática da literatura com foco na identificação, análise e apresentação dos processos e técnicas de desenvolvimento de jogos nas empresas de desenvolvimento de jogos focado nos processos de ciclo de vida utilizados e como os mesmos são adaptados para o uso durante o processo de desenvolvimento de jogos. Além do foco nos ciclos de vida e as adaptações feitas nos mesmos, a pesquisa desenvolvida pelos autores verificou os padrões de qualidade verificados e identificados nas empresas e como as mesmas aplicam e validam tais padrões. O artigo apresenta conceitos sobre a Playability como padrão de qualidade para avaliação dos jogos, porém, não apresenta técnicas para a avaliação da mesma. Neste contexto, os autores apresentam análises sobre o estado atual das técnicas utilizadas pela indústria e como os mesmos definem a qualidade de jogos e os processos realizados durante o ciclo de vida de desenvolvimento de tais empresas.

Duarte e Veloso (2016) descrevem em seu trabalho um estudo experimental utilizando um conjunto heurístico da literatura definido por Pinelle *et al.* (2009) para a avaliação de um jogo em desenvolvimento. O jogo em questão estava em sua segunda versão e, como objetivo para a realização dessa avaliação, os autores discutem a importância da análise da Playability de forma a avaliar os defeitos existentes para o desenvolvimento de uma terceira versão de seu jogo. Os autores demonstram, de acordo com os resultados apresentados, a importância da realização de avaliações da Playability de forma que são apresentados os principais defeitos encontrados quando utilizada uma técnica de avaliação durante o processo de desenvolvimento. O trabalho apresentado pode ser utilizado como base de conhecimento dos principais problemas encontrados através da utilização de uma técnica de avaliação da Playability como conjuntos heurísticos durante o desenvolvimento de um tipo de jogo específico.

Pinelle, Wong e Stach (2008) descrevem em seu trabalho um estudo experimental utilizando-se de uma avaliação da Usabilidade de interfaces de diversos jogos digitais de diferentes tipos e gêneros. Os autores realizaram-no com o objetivo de identificar a existência de similaridades entre diferentes jogos de um mesmo tipo e gênero. De acordo com os resultados apresentados, foi identificada a existência de uma similaridade entre as interfaces de

diferentes jogos digitais de um mesmo tipo e gênero, mesmo que cada uma destas interfaces seja proposta de diferentes formas. Com esses resultados, os autores desenvolveram um conjunto de *guidelines* para o desenvolvimento de interfaces melhores de acordo com o tipo e gênero do jogo em questão. No contexto dessa pesquisa, o trabalho apresentado por eles pode ser usado como uma base de conhecimento dos problemas e padrões das interfaces de jogos digitais de diversos tipos e gêneros.

Sánchez *et al.* (2012) desenvolveram em seu trabalho um estudo experimental para identificação das características da PX em jogos digitais. O estudo experimental desenvolvido foi realizado através do uso de um conjunto heurístico para avaliação da Playability. Além de utilizar um conjunto heurístico pré-existente na literatura para avaliação da Playability, os autores identificaram, através de estudos de *User Experience* (UX) existentes, as principais características relacionadas à experiência para a avaliação da PX em tais jogos. Os resultados do trabalho apresentado pelos autores são a definição de uma base de dados sobre conceitos necessários para a avaliação da PX em jogos digitais, além de uma base de conhecimento sobre os problemas comumente identificados da PX durante avaliações de jogos.

Por fim, a partir dos resumos das publicações analisadas através do mapeamento sistemático executado e de outros trabalhos identificados durante esta pesquisa, foi desenvolvida uma *feature analysis* dos conjuntos heurísticos identificados na literatura como meio de identificação dos defeitos e lacunas neles existentes. A *feature analysis* focou na identificação, e classificação, de três fatores considerados principais para o alcance dos objetivos deste trabalho. Os conjuntos identificados na literatura foram analisados em termos de sua categorização, tipos de jogos que o conjunto é capaz de avaliar e quais tipos de avaliadores poderiam interagir e utilizar os conjuntos propostos. Abaixo, uma definição de cada fator é apresentada:

- A **categorização** dos conjuntos da literatura foi analisada e classificada em três níveis diferentes: inexistente (IN) – quando o conjunto analisado não apresentava nenhuma classificação de suas heurísticas ou separação clara; existente (EX) – quando o conjunto analisado possuía qualquer tipo de categorização ou separação e classificação das heurísticas apresentadas no conjunto analisado; customizável (C) – quando o conjunto analisado pode ser utilizado com uma quantidade X qualquer de categorias definidas no conjunto dependendo do jogo sendo avaliado no momento. A **categorização** do conjunto foi considerada importante para que o conjunto utilizado na avaliação dos jogos fosse capaz de analisar as características de cada jogo de forma independente.

- O **tipo de jogo** que o conjunto é capaz de avaliar foi analisado e classificado em dois tipos diferentes: generalizado (G) – quando os conjuntos analisados eram definidos e poderiam ser utilizados para a avaliação de qualquer jogo, de qualquer tipo e gênero, independentemente de seu contexto; específico ou de escopo limitado (ES) – quando o conjunto analisado era definido apenas para ser utilizado em contextos específicos e tipos/gêneros de jogos específicos, como jogos online, educativos, jogos de tiro e outros. O **tipo de jogo** que o conjunto é capaz de avaliar foi considerado importante para que fosse possível melhor compreender se os conjuntos da literatura estavam sendo desenvolvidos com jogos gerais ou específicos em mente e por quê.
- O **tipo de avaliador** - para utilizar o conjunto e realizar avaliações da playability de jogos digitais há dois tipos diferentes de avaliador: especialistas (E) – classificado para os conjuntos que são definidos e desenvolvidos apenas para serem utilizados por avaliadores especialistas em playability e avaliações utilizando conjuntos heurísticos; não especialistas ou qualquer avaliador (NE) – classificado para os conjuntos que são definidos e desenvolvidos para serem utilizados por qualquer avaliador que possua nenhum ou quase nenhum conhecimento sobre playability e avaliações utilizando conjuntos heurísticos assim bem como por especialistas e pessoas experientes. O **tipo de avaliador** que pode utilizar o conjunto heurístico desenvolvido foi considerado importante uma vez que o objetivo desta pesquisa é de desenvolver um conjunto heurístico capaz de ser utilizado por qualquer avaliador e aumentar a disponibilidade e número de possíveis avaliadores. Para a identificação do tipo de avaliador ao qual o conjunto é direcionado, foram considerados, a explicação dos autores dos artigos sobre seu público-alvo, a população estudada, a estrutura do conjunto apresentado, e as heurísticas descritas em cada um dos mesmos.

A tabela contendo a *feature analysis* das publicações analisadas é apresentada a seguir (Tabela 2.7).

Tabela 2.7. Feature analysis dos conjuntos heurísticos avaliados no mapeamento sistemático.

Referência	NT	Categorização			Tipo de jogos		Tipo de Avaliador	
		IN	EX	C	G	ES	E	NE
Barcelos, T. S. <i>et al.</i> (2011)	NE	X			X		X	
Desurvire, H., Caplan, M. e Toth, J. A. (2004)	HEP - Heuristic Evaluation for Playability		X		X		X	
Desurvire, H. e Wiberg, C. (2008)	GAP - Game Approachability Principles	X			X		X	
Desurvire, H. e Wiberg, C. (2009)	PLAY - Heuristics of Playability		X		X		X	
Korhonen, H. e Koivisto, E. M. I. (2006)	Mobile Game Playability Heuristics		X			X	X	
Korhonen, H. e Koivisto, E. M. I. (2007)	Playability Heuristics for Mobile Multi-Player Games		X			X	X	
Pinelle, D., <i>et al.</i> (2008)	NE (Nome Generalizado: <i>Game Heuristics</i>)	X			X		X	
Pinelle, D., <i>et al.</i> (2009)	NGH - Network Game Heuristics		X			X	X	
Sánchez, J. L. G. <i>et al.</i> (2012)	NE (<i>PX - Player Experience: PHET</i>)		X			X	X	

Legenda: NT - Nome da técnica; IN - Inexistente; EX - Existente; C - Customizável; G - Generalizado; ES - Específico (Escopo Limitado); E - Especialista; NE - Não Especialista (Qualquer avaliador).

2.3.5 Discussão de resultados

De acordo com os resultados apresentados acima, provenientes do processo de avaliação e mapeamento da literatura atual dentro do contexto de métodos e definições para a avaliação da playability de jogos digitais, serão respondidas abaixo as três questões de pesquisa que foram levantadas para o desenvolvimento deste mapeamento:

SQ1 – Por qual público alvo de aplicadores/avaliadores que as técnicas propostas são intencionadas para serem utilizadas?

Como demonstrado na Tabela 2.7 de *feature analysis* dos conjuntos heurísticos encontrados na literatura, o foco principal dos conjuntos heurísticos desenvolvidos tem sido na definição de conjuntos que devem ser utilizados e avaliados através de especialistas e avaliadores com experiência prévia na utilização desse tipo de técnicas e conceitos. Dentre todos os conjuntos avaliados, nenhum deles se propôs a desenvolver um conjunto ou técnica que pudesse ser utilizada e aplicada por avaliadores com pouca ou nenhuma experiência nesse contexto. Assim, foi identificada uma lacuna, e uma oportunidade de pesquisa, dentro dos conjuntos heurísticos atualmente propostos para a avaliação da playability de jogos digitais. Um conjunto de heurísticas capaz de ser utilizado por avaliadores com pouca ou nenhuma experiência, bem como por especialistas, poderia facilitar o processo de seleção de avaliadores e agilizar a avaliação dos jogos em si.

SQ2 – As soluções propostas são direcionadas para qual escopo (em relação a tipos, gêneros e contextos gerais dos jogos)?

Em geral, dentre os conjuntos analisados no mapeamento sistemático, a divisão entre conjuntos utilizados para avaliação de jogos de maneira geral ou específica é muito balanceada. Cinco conjuntos analisados foram desenvolvidos de forma a serem utilizados para a avaliação geral de jogos digitais, não levando em consideração nenhum tipo ou gênero específicos de jogos digitais. Porém, na maioria dos trabalhos analisados que desenvolveram seus conjuntos de forma generalista, sem levar em conta um tipo e gênero específico em mente, houve problemas com heurísticas que não poderiam ser avaliadas em alguns jogos durante as seções de sua avaliação. Os outros quatro conjuntos analisados da literatura foram desenvolvidos com contextos específicos em mente, contendo heurísticas próprias para tipos, gêneros ou contextos definidos. Korhonen e Koivisto (2006) desenvolveram um conjunto com aspectos específicos para jogos mobile, e na continuação desse trabalho, Korhonen e Koivisto (2007) incluíram aspectos e heurísticas para a avaliação de jogos que possuíam características *multiplayer online*. Pinelle et al. (2009) desenvolveram um conjunto e focaram na cobertura de casos de jogos digitais que contavam com a conexão contínua com a internet e demandavam transmissão contínua de dados. Sánchez et al. (2012) desenvolveram e pesquisaram os conceitos e atributos da PX que deveriam ser avaliados nos jogos digitais e os principais defeitos e características que deveriam ser avaliados neste tipo de avaliação.

SQ3 – A solução proposta pode ser configurada para a avaliação de diferentes jogos digitais, possui contexto generalizado, ou é focada na avaliação de um específico subtipo de jogos, tipo e gênero específicos?

Como descrito nas seções 2.3.4.2.1 e 2.3.4.2.2 e demonstrado pela Tabela 2.7, nenhum conjunto de heurísticas proposto, dentre os que foram analisados por este mapeamento, tem a possibilidade de ser utilizado de forma customizada ou configurada de acordo com o jogo sendo avaliado. Dentre os nove conjuntos avaliados, cinco são focados em conjuntos generalistas, que tem por objetivo avaliar de forma geral jogos digitais sem um tipo, gênero ou contexto específico de jogos. Outros quatro conjuntos analisados têm contextos, ou tipos e gêneros de jogos, específicos a serem avaliados, sendo que dois deles definem heurísticas específicas para jogos mobile, dois outros possuem heurísticas específicas para contextos *multiplayer* ou online com conexão constante com a rede, e um deles analisou e definiu cenários e contextos específicos de PX que devem ser analisados nos jogos digitais. Fica evidenciada a falta de conjuntos e técnicas que sejam capazes de serem customizadas e modificadas de acordo com os jogos avaliados. De forma análoga, os quatro conjuntos da literatura analisados que possuem contextos específicos de avaliação demonstram a necessidade e a oportunidade de heurísticas que façam a avaliação de contextos específicos dependendo do jogo avaliado.

Através desta análise e mapeamento, foi possível também identificar algumas lacunas (“*gaps*”) iniciais nos conjuntos e trabalhos apresentados que suportam esta pesquisa: Desurvire & Wiberg (2009) demonstraram de forma qualitativa e quantitativa que o uso de conjuntos heurísticos contendo uma divisão, ou caracterização de categorias, para as heurísticas pode influenciar na qualidade e eficiência das avaliações de jogos digitais; Barcelos *et al.* (2011) demonstrou de forma qualitativa e quantitativa que a utilização de conjuntos reduzidos de heurísticas não altera a eficácia de identificação de defeitos mas pode influenciar na eficiência e no tempo total da avaliação, criando avaliações mais curtas com menos heurísticas a serem avaliadas; Korhonen & Koivisto (2006 & 2007) demonstram de forma qualitativa e quantitativa em seus trabalhos que é necessário avaliar contextos específicos através de heurísticas específicas para tais contextos, uma vez que as diferentes características dos tipos de jogos avaliados necessitam de heurísticas que possam identificar defeitos específicos. De forma semelhante, Chang *et al.* (2019) e De Souza Filho *et al.* (2019), desenvolveram trabalhos com o desenvolvimento de conjuntos heurísticos e avaliações da Playability em contextos específicos pois os mesmos possuem características específicas que necessitam ser analisadas de forma independente. Chang *et al.* (2019) desenvolve um trabalho com foco na avaliação de heurísticas para a Playability e para a Usabilidade de jogos de realidade aumentada para

dispositivos móveis. De Souza Filho *et al.* (2019) desenvolve um trabalho com foco na adaptação de um conjunto de heurísticas para a avaliação da gamificação de softwares incluindo os conceitos de Playability nas heurísticas, demonstrando que softwares gamificados podem utilizar os conceitos de qualidade da Playability, mas necessitando de pontos específicos.

Dentre todos os resultados obtidos durante o mapeamento da literatura, nenhuma ferramenta de apoio a identificação de defeitos da Playability foi encontrada. Além disso, os conjuntos de heurísticas encontrados na literatura não são desenvolvidos de forma a habilitar a avaliação de gêneros e tipos de jogos específicos. Portanto, mesmo que alguma ferramenta de apoio fosse definida na literatura, ainda seria necessário o desenvolvimento de uma ferramenta que apoiasse a customização do conjunto dependendo do tipo e gênero de jogo sendo avaliado.

Um ponto importante para a discussão e desenvolvimento do conjunto de heurísticas é que dentre os conjuntos analisados durante o mapeamento sistemático, todos os conjuntos possuíam, pelo menos, uma heurística relacionada aos seguintes pontos: conhecimento sobre as habilidades necessárias para a conclusão de uma tarefa atual ou de um futuro próximo; apresentação de feedbacks coerentes para cada uma das ações realizadas pelos jogadores; as terminologias utilizadas para cada um dos menus, ações ou atividades na tela fazem sentido para os jogadores e podem ser entendidas, ou facilmente identificadas, sem a necessidade de ajudas e traduções. Desta forma, é pretendido que estas heurísticas estejam presentes nos conjuntos desenvolvidos para esta pesquisa, indo de encontro com o conhecimento gerado pela pesquisa dos conjuntos da literatura.

Por último, foi realizada uma avaliação nos conjuntos da literatura em busca de categorias e métodos de classificação dos conjuntos heurísticos. Ou seja, cada conjunto heurístico desenvolve certas características que necessitam ser classificadas de acordo com elas. Cada categoria descreve uma parte de um jogo ou conjunto heurístico, contendo características que devem ser avaliadas em jogos digitais para sua completude e qualidade. Através da feature analysis apresentada na Tabela 2.7, foram identificadas quatro categorias gerais mais comuns que são utilizadas para a categorização e classificação dos conjuntos heurísticos, como, por exemplo, em Desurvire *et al.*, 2004 e em Korhonen & Koivisto, 2006. Dentre os seis conjuntos avaliados que possuíam algum tipo de classificação de suas heurísticas, cinco possuíam explicitamente (ou implicitamente utilizando nomes análogos) tais categorias identificadas. Essas categorias são: Usabilidade, *Game Play*, Mecânicas de Jogo e História. Porém, de acordo com a análise da literatura, foi possível identificar que a categoria relacionada à História poderia, em diversas situações, ser usada como um complemento dentro da categoria de *Game Play*.

2.4 Hands-On conjuntos heurísticos

Para um melhor entendimento dos conjuntos heurísticos e técnicas de avaliação da Playability selecionadas a partir do mapeamento sistemático, foi realizada uma série de avaliações ‘*Hands-On*’, em que o pesquisador pôde avaliar individualmente conjuntos e técnicas de avaliação na prática. Durante esse processo, o pesquisador utilizou quatro dos conjuntos heurísticos provenientes da *feature analysis* (Desurvire, Caplan e Toth, 2004; Korhonen e Koivisto, 2006; Desurvire e Wiberg, 2009; Barcelos *et al.*, 2011) em dois jogos diferentes randomicamente selecionados da loja de aplicativos *Android* (*Google Play Store*)²⁴.

2.4.1 Objetivo primário e conjuntos selecionados

O objetivo primário com a realização desta atividade foi a definição de uma questão exploratória sobre alguns conjuntos identificados no Mapeamento Sistemático. Desejava-se identificar inicialmente quais seriam os defeitos e limitações para cada um dos conjuntos assim como suas qualidades durante a avaliação de um jogo qualquer. O pesquisador responsável poderia melhor compreender como cada conjunto e heurística seria aplicado para a avaliação da Playability de um jogo. Para isso, foram selecionados quatro conjuntos heurísticos que apresentavam diferentes características de acordo com a *feature analysis* apresentada anteriormente.

O conjunto proposto por Barcelos *et al.* (2011) é um conjunto sem a presença de classificação das heurísticas e apresenta o menor conjunto de heurísticas existente com o desenvolvimento a partir da literatura. O conjunto proposto por Desurvire, Caplan e Toth (2004) possui uma categorização simples de três categorias utilizando-se de um conjunto de heurísticas para a avaliação generalizada de tipos e gêneros de jogos digitais. O conjunto proposto por Desurvire e Wiberg (2009) é um conjunto heurístico com sistema de categorização complexo e com um maior número de heurísticas que, intencionalmente, são desenvolvidas para cobrir um maior número de tipos e gêneros utilizados por ele para avaliação. O conjunto proposto por Korhonen e Koivisto (2006) é um conjunto com categorização simples e heurísticas definidas de forma similar às outras através da literatura, porém, ele é definido mais especificamente para um contexto específico de jogos digitais móveis.

²⁴ Google Play Store, disponível em: <https://play.google.com/store?hl=pt>.

2.4.2 Jogos selecionados para o processo de avaliação

Foram selecionados dois jogos exclusivos para a plataforma Android para a realização da avaliação “*Hands-On*”. Os jogos foram utilizados em um dispositivo novo e compatível com os dois jogos em questão para diminuição da interferência de hardware fraco ou insuficiente em sua avaliação. O primeiro jogo selecionado foi o jogo *Green the Planet 2*²⁵, um jogo de tipo Arcade gratuitamente disponível com história intensiva e poucos objetivos. O segundo foi o jogo *The Cannon Man Adventures*²⁶, um jogo disponível tanto para a plataforma *Android* quanto para a plataforma *iOS (Apple)*. Esse jogo tem temática simples, poucas mecânicas, mundo 2D e pouca história baseado em fases únicas.

Os jogos acima foram selecionados por estarem gratuitamente disponíveis em suas lojas de aplicativos, possuem boas notas gerais de avaliação nas mesmas ao mesmo tempo que apresentam diversos comentários com diferentes defeitos de jogabilidade, bugs e propostas de melhoria. Além disso, os dois jogos selecionados possuíam uma campanha de jogo pequena que possibilitou a investigação de quase todas as mecânicas de jogo em um curto espaço de tempo e o avanço rápido na história/progressão do jogo.



Figura 2.2. Tela inicial do jogo Green the Planet 2



Figura 2.3. Tela inicial do jogo The Cannon Man Adventures

2.4.3 Processo de avaliação

O processo de avaliação foi realizado pelo pesquisador em forma de duplas: cada conjunto heurístico selecionado era testado nos dois jogos, primeiro em um e depois no outro. A cada par (conjunto/jogos) realizado, o pesquisador invertia a sequência dos jogos utilizados como uma tentativa de não criar um ciclo contínuo de fadiga. Assim, para cada descrição de problema

²⁵ Green the Planet 2, disponível em: <https://play.google.com/store/apps/developer?id=Kikaku%20Damashii%2C%20Inc.&hl=pt>

²⁶ The Cannon Man Adventures, disponível em: <http://cannonman.trdreams.com/>

existente em cada um dos jogos, o pesquisador faria a relação problema/heurísticas as quais melhor definissem a combinação entre as mesmas.

Ao fim de cada par (conjunto/jogo) o pesquisador anotaria os problemas, lacunas e qualidades de cada conjunto, caracterizando a utilização deles. Ao final, são descritos os principais defeitos e lacunas apresentadas pelos conjuntos avaliados nesse processo de *Hands-On*.

2.4.4 Resultados avaliados

A partir da avaliação de *feature analysis* dos conjuntos heurísticos avaliados, juntamente com os resultados obtidos do processo de *hands-on* dos conjuntos citados, temos os seguintes resultados/observações sobre eles:

- **Linguagem técnica:** de acordo com o processo de desenvolvimento dos conjuntos, todos os identificados são voltados para avaliações realizadas por especialistas, desta forma, a linguagem utilizada para definir as heurísticas de avaliação são extremamente técnicas, podendo causar confusão de termos em alguns avaliadores.
- **Apoio na localização das heurísticas:** fica clara a necessidade de utilização de categorização das heurísticas quando avaliados os resultados apresentados por cada um dos artigos citados no mapeamento realizado. Da mesma forma, durante o processo de *hands-on*, foi possível identificar uma maior facilidade quando os conjuntos avaliados apresentavam categorias de classificação. Porém, em alguns casos, as heurísticas estavam mal localizadas, criando uma busca mais prolongada por certas heurísticas.
- **Direcionamento a tipos e gêneros:** analisando a tabela de *feature analysis* dos conjuntos selecionados do mapeamento sistemático, é possível identificar que nenhum dos conjuntos avaliados leva em consideração a diferença existente entre tipos e gêneros de jogos. Essa característica se fez presente durante a avaliação dos conjuntos no processo de *hands-on*, em que certas heurísticas não poderiam ser analisadas, por conta do tipo ou gênero do jogo em questão; ou determinado defeito não poderia ser classificado uma vez que o conjunto não apresentava uma heurística que o definisse.
- **Apoio à indústria:** identificamos uma grande lacuna onde os conjuntos heurísticos existentes não possuem direcionamento à indústria de jogos especificamente, uma vez que, de acordo com Politowski *et al.* (2016), a indústria procura metodologias e técnicas de avaliação com baixo custo de aplicação e com resultados imediatos. A necessidade de utilização de especialistas para a realização da avaliação torna esse processo

dispendioso, uma vez que especialistas tem custos elevados, e, por conta de sua escassez, há um aumento no tempo para realização da avaliação, já que deve ser feito todo o processo de contrato e avaliação separadamente.

De acordo com os resultados observados, e a comparação das técnicas utilizadas, foi definido que o conjunto heurístico proposto por Barcelos *et al.* (2011) seria utilizado como objeto de comparação nos estudos desenvolvidos para validação da técnica inicialmente proposta. O conjunto desenvolvido por Barcelos *et al.* (2011) foi escolhido por ser um dos poucos conjuntos que não apresentam uma classificação de heurísticas além de ser um conjunto proposto com o objetivo de ser menor e mais genérico em relação aos conjuntos da literatura. Por fim, esse foi o conjunto heurístico capaz de identificar o maior número de defeitos nos jogos em que foi utilizado, sendo capaz de identificar, em média, 5 defeitos da Playability a mais que os outros conjuntos.

2.5 Conclusão

Nesse capítulo foram apresentados os conceitos e principais métodos de avaliação da Playability para avaliação de jogos digitais. Além de embasamento teórico para esta pesquisa, foram introduzidos os conceitos que serão utilizados ao longo dela. Portanto, analisando juntamente com os conceitos apresentados no Capítulo 1, foram apresentados a contextualização, problema, método de pesquisa, e embasamento teórico utilizado para a proposta de pesquisa apresentada neste texto. Também foram apontados diversos defeitos e lacunas existentes nas pesquisas atuais da literatura em conjuntos heurísticos os quais poderiam ser solucionados de acordo com os objetivos definidos para esta pesquisa. Por fim, a partir dos resultados observados, foi selecionado um conjunto heurístico específico para a realização das avaliações comparativas das versões da técnica desenvolvida neste texto. Nesse caso, o conjunto heurístico proposto por Barcelos *et al.* (2011) foi selecionado, primeiramente por ser um dos poucos conjuntos propostos que não possui categorização; depois, por ser o único conjunto proposto que define um conjunto reduzido com o objetivo principal de identificar uma diferença estatística através dessa abordagem comparado aos conjuntos propostos na literatura. Por fim, durante o processo de avaliação dos conjuntos heurísticos (veja seção 2.4) o conjunto proposto por Barcelos *et al.* (2011) foi o conjunto que conseguiu identificar o maior número de defeitos em um menor tempo de avaliação.

CAPÍTULO 03 – NExPlay: NON-EXPERT PLAYABILITY HEURISTIC SET

Este capítulo apresenta a proposta da primeira versão da técnica de avaliação da Playability em jogos digitais chamada de NExPlay (Non Expert Playability heuristic set). Também são apresentados os objetivos desejados com esta proposta e a metodologia de desenvolvimento aplicada.

3.1 Introdução

De acordo com os problemas identificados nos conjuntos heurísticos existentes na literatura apresentados na seção 2.4, e com as necessidades atuais da indústria de desenvolvimento de jogos para avaliações de baixo custo e com resultados imediatos (Politowski *et al.*, 2016), este capítulo apresenta um novo conjunto de heurísticas para avaliação da *Playability* chamado NExPlay (*Non-Expert Playability Heuristic Set*). A NExPlay tem por objetivo principal propor uma técnica de avaliação generalizada com a qual todo tipo de inspetor pode realizar a avaliação da *Playability*, tendo ele conhecimento sobre processos de avaliação utilizando heurísticas, ou de *Playability*, ou não. Além disso, essa técnica pretende diminuir o número geral de heurísticas apresentadas aos usuários com o objetivo de reduzir o tempo necessário para a realização de uma avaliação da *Playability* em jogos digitais.

Neste contexto, o NExPlay é desenvolvido como um conjunto de heurísticas reduzido a partir dos conjuntos heurísticos presentes na literatura (Desurvire *et al.*, 2004; Korhonen & Koivisto, 2006; Desurvire & Wiberg, 2009; Barcelos *et al.*, 2011). O conjunto foi desenvolvido como um conjunto reduzido da literatura uma vez que os conjuntos apresentados pela literatura já tiveram suas heurísticas validadas através de estudos experimentais não sendo necessário identificar novas métricas para o desenvolvimento dessas heurísticas. Os objetivos desejados com o desenvolvimento dessa técnica são:

1. Minimizar a quantidade de heurísticas necessárias para a avaliação da *Playability* de jogos digitais.
2. Viabilizar avaliação da *Playability* por avaliadores não especialistas.

3.2 Desenvolvimento

Para o desenvolvimento da NExPlay, foram considerados os conjuntos heurísticos da literatura bem conhecidos e com os conjuntos que, de acordo com o entendimento dos autores,

apresentaram os melhores resultados e melhor processo de desenvolvimento das heurísticas. Inicialmente esses conjuntos foram selecionados e, individualmente, analisados em diferentes cenários de diferentes jogos, como explicado anteriormente na seção 2.5. A partir dos resultados dessas análises individuais, selecionamos 4 conjuntos heurísticos da literatura (Desurvire *et al.*, 2004; Korhonen & Koivisto, 2006; Desurvire & Wiberg, 2009; Barcelos *et al.*, 2011) os quais utilizaríamos para o desenvolvimento da NExPlay. Durante a análise individual dos conjuntos, foi realizada, de forma conjunta, a avaliação individual das características apresentadas por eles e também dos métodos aplicados para seu desenvolvimento.

A partir desses resultados foi realizada a seleção das heurísticas que se mostraram úteis durante a avaliação da *Playability* em jogos digitais diversos. Também foram incluídas heurísticas que, de acordo com a análise dos pesquisadores, poderiam ser utilizadas em outras situações e jogos distintos. A partir dessa seleção, uma lista de heurísticas contendo a união das heurísticas provenientes dos trabalhos da literatura selecionados foi obtida. Foi avaliada a similaridade entre as heurísticas buscando realizar uma unificação das semelhantes. Este processo foi realizado com o objetivo de minimizar o tamanho do conjunto heurístico final, um dos principais objetivos para o desenvolvimento dessa técnica.

Com a lista de heurísticas similares unificadas, foi realizada uma análise da terminologia e linguagem utilizada para sua descrição. Heurísticas que apresentavam terminologias ou linguagens incompatíveis com inspetores não especialistas em jogos ou em avaliação foram reavaliadas e reescritas, de maneira a ser compatíveis com esse ‘cenário’. O objetivo dessa fase foi simplificar as frases que definiam cada uma das heurísticas da técnica para atingir um maior entendimento da parte dos inspetores durante as avaliações. Esse também é um objetivo central definido para o desenvolvimento dessa técnica. Esperava-se que, com sentenças mais simples e com terminologias bem conhecidas, qualquer inspetor, seja ele especialista ou não, poderia utilizar e facilmente entender as heurísticas contidas nesse conjunto.

De acordo com os resultados do MSL apresentados na seção 2.3.5, existem quatro categorias principais na literatura utilizadas na definição de conjuntos heurísticos. Além disso, assim como já descrito nestes resultados, a categoria relacionada à História pode ser considerada como um complemento dentro da categoria de *Game Play*. Sendo assim, com o objetivo de diminuir a quantidade de heurísticas e, possivelmente, tornar mais fácil a avaliação dos jogos pelos inspetores, utilizamos a categoria de *Game Play* em conjunto com as heurísticas de História.

A partir do conjunto heurístico totalmente desenvolvido e categorizado em *Game Play*, Mecânicas e Usabilidade, é possível avaliar sua completude e concordância com o mundo real. Para isso, foram realizadas entrevistas com 5 jogadores especialistas para que avaliassem o conjunto de heurísticas e apontassem, de acordo com suas experiências, se havia a necessidade de inclusão ou não de outras heurísticas para a completude da técnica de avaliação. Os jogadores identificaram a necessidade de inclusão de algumas heurísticas assim como a modificação da sentença que definia outras para um melhor entendimento delas durante a avaliação. O conjunto completo das heurísticas é apresentado na seção 3.3.

Por fim, foi realizada uma avaliação de viabilidade da técnica proposta em comparação a um conjunto heurístico da literatura que fosse semelhante à abordagem utilizada para exclusão de maiores vieses. A metodologia e resultados desse estudo são mostrados no próximo capítulo.

3.3 Conjunto Heurístico

A técnica NExPlay apresenta um total de 19 heurísticas categorizadas em 3 seções: *Game Play*, Usabilidade e Mecânicas. Cada uma das três tabelas abaixo (Tabela 3.1, Tabela 3.2, Tabela 3.3) apresenta uma das 3 categorias selecionadas para a técnica de avaliação NExPlay e suas heurísticas. Para identificação da origem de cada heurística iremos abreviar a pesquisa de onde cada heurística foi proveniente, tendo: HEP – para as heurísticas definidas por Desurvire *et al.* (2004); *Mobile* – para as heurísticas desenvolvidas por Korhonen & Koivisto (2006); *Play* – para as heurísticas desenvolvidas por Desurvire & Wiberg (2009); e Barcelos – para o conjunto desenvolvido por Barcelos *et al.* (2011). Dentre as heurísticas de *Game Play* apresentadas na Tabela 3.1, as heurísticas G1 e G6 estão relacionadas a avaliação da história do jogo de forma geral.

Tabela 3.1. Heurísticas da categoria *Game Play*

ID	Heurísticas – <i>Game Play</i>	Origem
G1	Atividades desenvolvidas durante o jogo são variadas de modo a diminuir fadiga além de serem balanceadas com a dificuldade do jogo.	<i>Play</i> , HEP
G2	O jogo é balanceado de modo a aplicar pressão no jogador sem frustrá-lo, desta forma, desafios são experiências positivas ao jogador e o mantém interessado.	Barcelos, <i>Mobile</i>
G3	O jogador não deve ser penalizado repetitivamente pelo mesmo erro ou perder qualquer objeto ganho através de grande esforço.	<i>Play</i> , <i>Mobile</i> , HEP

G4	A inteligência artificial deve apresentar desafios e surpresas inesperadas ao jogador independente de seu nível assim como acompanhar seu aprendizado.	<i>Play</i>
G5	As habilidades necessárias para alcançar um objetivo atual ou futuro são conhecidas e ensinadas e, para tais habilidades aprendidas, o jogo oferece uma recompensa clara e imediata.	<i>Play, Mobile, HEP, Barcelos</i>
G6	O jogo apresenta diversos objetivos secundários e opcionais os quais complementam um (ou mais) objetivo (s) principal (ais) sendo que para cada possível conquistá-los de diversas maneiras.	<i>Barcelos, Mobile</i>

Tabela 3.2. Heurísticas da categoria Usabilidade

ID	Heurísticas - Usabilidade	Origem
U1	Todas as vezes que o jogador iniciar o jogo o mesmo possui informações suficientes para começar a jogar, ou seja, não há necessidade de acessar manuais, documentações ou o tutorial constantemente.	<i>Play, HEP, Barcelos</i>
U2	O jogador consegue obter facilmente informações sobre tudo a sua volta além de informações sobre estado atual do jogo, pontuações e outras informações de maneira óbvia e simples.	<i>Play, Mobile</i>
U3	O jogo apresenta feedback ao jogador de forma apropriada (música/efeitos sonoros/vibrações de controle/avisos visuais) e de modo consistente (possui um padrão sendo seguido), imediato e desafiador em relação às ações do jogador.	<i>Play, Mobile, HEP, Barcelos</i>
U4	Layouts e menus devem ser intuitivos e organizados de forma que o jogador possa manter o foco no jogo além de ser capaz de evitar erros não intencionais dos jogadores.	<i>Mobile, Play</i>
U5	O jogador deve ser capaz de interromper facilmente a partida a qualquer momento, além de ser possível salvar o estado atual para retomar o jogo posteriormente (ou seja, interrupções são esperadas e planejadas).	<i>Mobile</i>
U6	O jogo deve apresentar um tutorial inicial bem estruturado para familiarização do jogador além de o desafio do jogo poder ser ajustado de acordo com as habilidades dos mesmos.	<i>Play, HEP, Barcelos</i>
U7	Todas as representações visuais do jogo devem ser de fácil compreensão pelo jogador naturalmente.	<i>Barcelos, Mobile</i>
U8	O tutorial deve ensinar ao jogador as funcionalidades e mecânicas básicas sem ser intrusivo ou criando uma espécie de Manual do jogo durante sua extensão.	<i>Mobile, Play</i>

Tabela 3.3. Heurísticas da categoria Mecânicas

ID	Heurísticas - Mecânicas	Origem
M1	Controles devem ser claros, customizáveis e fisicamente confortáveis (se aplicável), além de suas respectivas ações de resposta serem imediatas.	<i>Mobile</i> , Barcelos
M2	Os controles devem ser suficientemente simples para que o aprendizado seja rápido, porém, estes devem ser expansíveis também para jogadores mais habilidosos.	<i>Play</i> , HEP
M3	Jogadores devem receber ajuda de contexto sensível de modo que não se encontrem presos ao ponto de necessitar de um manual, porém, estas não devem facilitar demais as tarefas necessárias a uma fase.	<i>Play</i> , <i>Mobile</i> , HEP, Barcelos
M4	O jogador deve possuir a habilidade de modificar as configurações do jogo (som, gráficos e outros) de maneira simples e de modo que o satisfaça.	<i>Mobile</i> , Barcelos
M5	As terminologias utilizadas para todos os objetos/funções/outras são entendidas pelo usuário.	<i>Play</i> , <i>Mobile</i> , HEP, Barcelos

3.4 Discussão

Nesse capítulo é apresentada uma primeira versão da técnica proposta NExPlay, que tem por objetivo principal viabilizar a avaliação da *Playability* por avaliadores não especialistas. Para alcançar esse objetivo, uma avaliação dos conjuntos existentes na literatura para identificar os problemas enfrentados por eles foi realizada, tanto através da *feature analysis* dos conjuntos da literatura quanto pelo *hands-on* desses conjuntos. De forma similar, foram utilizadas para entender e identificar quais heurísticas melhor apoiariam, e como apoiariam, os avaliadores durante uma inspeção de jogo. Esse processo de avaliação dos conjuntos existentes, e suas dificuldades para avaliação da *Playability* são descritos na seção 2.4.

De acordo com os passos apresentados no desenvolvimento da técnica NExPlay, formou-se um conjunto de heurísticas composto por três seções, ou categorias, as quais apresentam, no total, 19 heurísticas. Comparado aos conjuntos heurísticos apresentados pela literatura, esse é um número pequeno, uma vez que os maiores conjuntos da literatura apresentam aproximadamente 50 heurísticas. Além disso, a partir da simplificação das sentenças nas heurísticas, foi possível viabilizar seu entendimento por qualquer pessoa, sem a necessidade de um conhecimento específico em *Playability* ou jogos digitais.

Da forma como a técnica NExPlay foi proposta, é esperado que tenha um desempenho melhor ou similar ao desempenho de um conjunto heurístico qualquer da literatura. Também que possa ser utilizada por avaliadores não especialistas na avaliação de jogos digitais e

Playability. Para a verificação e avaliação da solução proposta um estudo experimental comparativo com outro conjunto heurístico da literatura foi desenvolvido. Esse estudo e seus respectivos resultados são apresentados no próximo capítulo. Além disso, ao fim do capítulo, discutiremos os benefícios de se usar tal técnica assim como seus defeitos e como aprimorá-la para ser uma técnica mais bem adaptada às situações desejadas.

CAPÍTULO 04 – ESTUDO DE VIABILIDADE DA NEXPLAY

Este capítulo apresenta um estudo experimental comparativo para verificação da viabilidade da técnica proposta NExPlay. Tem como objetivo principal avaliar a viabilidade de utilizar a técnica NExPlay para a avaliação da Playability em jogos digitais por avaliadores “não especialistas”. Ele apresenta resultados qualitativos e quantitativos de acordo com as métricas de eficiência e eficácia da avaliação.

4.1 Introdução

Para a avaliação da técnica proposta um estudo experimental comparativo foi realizado para identificação de sua utilidade, além de avaliar se os objetivos desejados foram alcançados com seu desenvolvimento. O objeto de comparação selecionado foi o conjunto heurístico proposto por Barcelos *et al.* (2011). O conjunto proposto por Barcelos *et al.* (2011) foi selecionado pois foi desenvolvido com uma metodologia semelhante à apresentada no capítulo anterior para a NExPlay além de ter apresentado o melhor desempenho entre todos os conjuntos comparados no processo de Hands-On feito inicialmente. O objetivo dos autores era apresentar um conjunto de heurísticas reduzido da literatura para avaliar se a quantidade de heurísticas afetaria o tempo de avaliação pelos usuários de forma significativa. Além disso, o conjunto proposto pelos autores não apresenta uma categorização, dessa forma, será possível avaliar se existe alguma diferença significativa entre avaliações utilizando conjuntos categorizados ou não, assim como obter a percepção dos avaliadores que utilizaram o conjunto.

4.2 Estudo Experimental

O objetivo desse estudo foi :

- Verificar a viabilidade de uso da técnica por avaliadores especialistas e não especialistas em jogos e avaliações utilizando heurísticas.

De acordo com esse objetivo geral também foram definidos objetivos secundários que deveriam ser alcançados através da avaliação do objetivo anteriormente definido. Os objetivos secundários são apresentados a seguir:

- Avaliar quantitativamente a técnica para identificação da eficiência e eficácia produzidas pelo mesmo.

- Avaliar qualitativamente a técnica para identificação de facilidade de uso e interação com o mesmo.
- Identificar a facilidade de entendimento das heurísticas da técnica.
- Identificar a necessidade de categorização da técnica durante um ciclo de uso.
- Identificar a abrangência das heurísticas em relação ao jogo utilizado para a avaliação.
- Identificar a diferença da qualidade de avaliação entre usuários especialistas ou não.

4.2.1 Metodologia

Para a avaliação de viabilidade da NExPlay, as métricas de eficiência e eficácia foram selecionadas para identificação de sua performance em relação ao conjunto heurístico proposto por Barcelos *et al.* (2011). O estudo foi projetado para uma avaliação de um fator e dois tratamentos, em que o fator é a solução proposta para a avaliação da Playability e os tratamentos são a NExPlay e o conjunto proposto por Barcelos *et al.* (2011).

Para realização desse estudo foram selecionados 24 alunos entre graduandos de Ciência da Computação e integrantes do Programa de Pós-Graduação em Informática da Universidade Federal do Amazonas (UFAM) durante todo o mês de Junho de 2017. Todos os participantes possuíam conhecimento prévio dos conceitos de Interação Humano-Computador (IHC) e Engenharia de Software (ES). Os participantes possuíam diferentes níveis de conhecimento em jogos digitais e inspeção de software. Como o estudo objetivava medir a viabilidade do conjunto em relação a outro conjunto da literatura, foi definido um experimento de apenas uma via. Não foi realizado um processo de avaliação cruzada (*Cross-Evaluation*) uma vez que o conhecimento adquirido pelo uso de um dos conjuntos poderia influenciar o uso e avaliação utilizando o segundo conjunto. Sendo assim, metade dos avaliadores utilizou a NExPlay, enquanto a outra metade utilizou o conjunto proposto por Barcelos *et al.* (2011).

A principal diferença observada entre as duas soluções utilizadas para esse estudo foi que o conjunto heurístico proposto por Barcelos *et al.* (2011) não possuía uma classificação das heurísticas de acordo com as 3 categorias previamente definidas e discutidas. Dessa forma, um dos objetivos desse estudo foi o de avaliar o efeito causado por essas categorias na avaliação dos jogos. Além disso, o conjunto heurístico proposto por Barcelos *et al.* (2011) foi desenvolvido de forma semelhante ao processo de desenvolvimento da NExPlay além de apresentar uma quantidade semelhante de heurísticas. Além disso, Barcelos *et al.* (2011) apresentou o melhor desempenho entre todos os conjuntos comparados no processo de Hands-

On feito inicialmente nesta pesquisa. Sendo assim, um dos motivos para tal escolha era a de excluir possíveis vieses de avaliação entre os dois conjuntos, uma vez que eram semelhantes em quantidade de heurísticas e método de desenvolvimento para elas.

Para que não houvesse diferenças significativas de conhecimento entre os dois grupos que realizaram as avaliações, foi realizado um balanceamento dos grupos de acordo com seus conhecimentos em inspeção e jogos. Para este balanceamento, todos os avaliadores responderam ao questionário de caracterização previamente à realização do estudo. O questionário incluía questões abordando o conhecimento geral dos avaliadores em inspeção de software e em experiência no uso de jogos digitais sem definição de escopo em tipo ou gênero de jogo específico. Para a questão sobre o conhecimento em inspeções de software foram definidos quatro níveis de conhecimento baseados nas experiências: **sem experiência, baixa experiência, média experiência e alta experiência**. Esses níveis de experiência foram definidos de acordo com a experiência dos avaliadores em projetos de inspeção de software na indústria e academia:

- **Sem experiência:** não possui conhecimento de inspeções de software e não possui nenhuma experiência anterior na avaliação de qualidade de softwares em geral.
- **Baixa experiência:** possui pequeno conhecimento em inspeções e avaliações de qualidade de software adquiridos através de experiências anteriores em projetos de avaliação na academia.
- **Média experiência:** possui conhecimento dos conceitos gerais para a inspeção de software e já obteve experiências anteriores em pelo menos um projeto de avaliação de software na indústria além da academia.
- **Alta experiência:** possui conhecimento de processos de inspeção de software e já obteve experiências na avaliação de diversos projetos na indústria com softwares gerais.

Além da avaliação do conhecimento em inspeção e avaliação de qualidade de software, também foi questionado o conhecimento dos participantes em relação ao seu conhecimento dos jogos digitais. Para esta questão foram definidos quatro níveis de conhecimento. O critério para a avaliação do conhecimento em jogos foi considerar a frequência do uso de jogos digitais durante uma semana:

- **Sem experiência:** não joga nenhum tipo de jogo digital.
- **Baixa experiência:** joga casualmente ou, pelo menos, com frequência constante de uma a duas vezes na semana.

- **Média experiência:** joga, pelo menos, com frequência constante de três a quatro vezes na semana.
- **Alta experiência:** joga, pelo menos, com frequência constante de cinco vezes na semana ou todos os dias da semana.

Deve ser observado que, para essa questão, não existiram restrições quanto a tipos ou gêneros conhecidos e utilizados pelos avaliadores durante essas frequências. Além disso, também não existiram restrições quanto ao tempo mínimo utilizado durante as seções do avaliador, qualquer quantidade de tempo poderia ser considerada como um dia completo. O questionário completo, suas perguntas e possíveis respostas, está disponível no APÊNDICE A. A seguir, na Tabela 4.1, são apresentados os níveis de conhecimento de cada avaliador que participou desse estudo experimental estando separados pela solução utilizada.

Antes da realização do balanceamento entre os dois grupos, foi realizada a distribuição dos avaliadores partindo do princípio do design randômico, aleatoriamente distribuindo-os entre os grupos. Após essa distribuição, os participantes foram alternados entre os dois grupos de modo a obedecer ao balanceamento das equipes em relação ao conhecimento nos dois quesitos. Ao final dessa fase, os dois grupos possuíam 12 participantes balanceados de acordo com seus níveis de conhecimento.

Para garantir que possuíssem o mesmo nível de conhecimento nos processos de avaliação de software/jogos utilizando heurísticas, assim como nos conceitos e objetivos da Playability, um treinamento foi realizado com todos os participantes. O treinamento teve duração média de pouco menos de uma hora com o objetivo de apresentar os conceitos básicos da Playability e de avaliações utilizando heurísticas. Durante o treinamento, foram apresentados exemplos práticos utilizando um conjunto heurístico diferente dos dois apresentados para a avaliação, assim como jogos diferentes do utilizado na inspeção. Além disso, o tipo de jogo apresentado no exemplo e suas mecânicas eram similares ao do jogo selecionado para o estudo de modo a familiarizar ao máximo os participantes com o estudo a ser realizado. Por fim, exemplos de como os participantes deveriam preencher as tabelas de discrepância e como assinalar as heurísticas afetadas foram apresentadas a eles, além de como deveriam ser preenchidos os campos de tempo de início e fim para a inspeção.

Tabela 4.1. Níveis de conhecimento dos avaliadores por conjunto utilizado

AV	CI	CJ	CU
AV - 01	Baixo	Baixo	NExPlay
AV - 02	Baixo	Baixo	
AV - 03	Médio	Nenhum	

AV	CI	CJ	CU
AV - 04	Alto	Baixo	Barcelos <i>et al.</i>
AV - 05	Baixo	Médio	
AV - 06	Baixo	Baixo	
AV - 07	Alto	Nenhum	
AV - 08	Baixo	Baixo	
AV - 09	Médio	Baixo	
AV - 10	Baixo	Alto	
AV - 11	Médio	Baixo	
AV - 12	Médio	Alto	
AV - 13	Alto	Alto	
AV - 14	Médio	Baixo	
AV - 15	Alto	Médio	
AV - 16	Baixo	Baixo	
AV - 17	Baixo	Nenhum	
AV - 18	Nenhum	Nenhum	
AV - 19	Baixo	Baixo	
AV - 20	Baixo	Baixo	
AV - 21	Médio	Baixo	
AV - 22	Nenhum	Baixo	
AV - 23	Baixo	Alto	
AV - 24	Alto	Nenhum	

Legenda: AV – Avaliador; CI – Conhecimento em Inspeções de Software; CJ – Conhecimento em Jogos; CU – Conjunto Utilizado

4.2.2 Jogo Selecionado

Para a seleção do jogo a ser avaliado, inicialmente, um especialista em jogos e Playability selecionou quatro jogos de diferentes gêneros e tamanhos (tempo de jogo). Esse especialista realizou uma avaliação dos jogos utilizando a técnica NExPlay para identificação de possíveis defeitos. O objetivo dessa avaliação era identificar um jogo que apresentasse uma quantidade significativa de defeitos para maior facilidade de identificação deles pelos participantes. Além disso, o jogo deveria ser de tipo e gênero **casual** para que os participantes não necessitassem de treinamento ou grandes tutoriais e esforço para utilizar o jogo. A partir dos quatro jogos selecionados, dois foram descartados por apresentarem *game play* muito monótona ou poucos defeitos. Os outros dois jogos selecionados foram avaliados por dois inspetores externos e cada inspetor utilizou apenas um jogo, juntamente com a NExPlay. Esta avaliação preliminar teve dois objetivos principais:

- Identificar se existia a necessidade de alteração preliminar da técnica NExPlay anteriormente à realização do estudo e;
- Identificar qual dos dois jogos deveria ser utilizado de acordo com os defeitos e dificuldades identificadas pelos dois inspetores.

Após essa avaliação preliminar, foi selecionado o jogo *Leap of Cat*²⁷ de gênero casual com história e mecânicas simplificadas, de fácil compreensão por todos os participantes. O mesmo tem distribuição gratuita e disponível apenas para a plataforma Android através da *Google Play Store* em qualquer dispositivo compatível. Esse jogo foi selecionado por apresentar uma grande quantidade de problemas referentes à *Playability* mas, mesmo assim, ter uma boa avaliação na loja de aplicativos, de acordo com seus usuários. Além disso, o jogo em questão é desenvolvido localmente e a pesquisa possui apoio dos desenvolvedores do jogo de forma que os mesmos poderiam tirar dúvidas sobre seu uso e poderiam dar apoio na utilização do jogo nos dispositivos caso o mesmo porventura não viesse a estar disponível para download.



Figura 4.1. Tela inicial do jogo Leap of Cat

4.2.3 Processo Experimental

Primeiramente, um treinamento com os participantes foi realizado para a normalização de seus conhecimentos em relação aos conceitos fundamentais da *Playability* e de avaliações utilizando heurísticas. Esse treinamento ocorreu em dois dias para que todos os participantes pudessem selecionar o melhor dia para a sua participação. O mesmo treinamento foi realizado nos dois dias utilizando-se dos mesmos exemplos. Após sua realização, os participantes foram instruídos a realizar a avaliação do jogo e identificar seus defeitos, assinalando-os na tabela de discrepâncias entregue a eles, apresentada no APÊNDICE B. O processo de avaliação e identificação dos defeitos foi mantido aberto para que os participantes o realizassem no momento mais adequado, respeitando o prazo estipulado para entrega.

²⁷ https://play.google.com/store/apps/details?id=com.fpftech.leapofcat&hl=en_US&gl=US

Para que esse processo fosse realizado, foram enviados para os e-mails individuais de cada participante os documentos necessários para realizar a inspeção da *Playability* do jogo, os quais eram:

- Tabela de descrição e associação das discrepâncias e heurísticas afetadas individualmente.
- Solução a ser utilizada.
- Curta descrição das atividades a serem executadas juntamente com as instruções para preencher o questionário de avaliação de uso da NExPlay.

Os participantes deveriam enviar suas inspeções para o mesmo e-mail apenas com a tabela de discrepâncias preenchida e responder digitalmente ao questionário de avaliação da técnica utilizada. Além disso, o critério e processo para a realização da inspeção do jogo ficaram em aberto para que os participantes pudessem decidir o melhor modo de realizá-la. Para aqueles participantes que não possuíam dispositivos *Android* compatíveis com o jogo a ser avaliado, foram disponibilizados *tablets* para realização da inspeção em dia e local específicos com o jogo previamente instalado.

Todos os participantes realizaram avaliações individuais, anotando seus tempos de início e fim respectivamente. Para os participantes que realizassem pausas entre fases diferentes de avaliação do jogo, foi instruído que os mesmos notassem os tempos de início e fim respectivamente a cada fase desenvolvida. Quando as tabelas retornaram com todas as respostas, tempos de intervalo foram descartados e apenas os tempos totais de avaliação do jogo foram considerados.

Durante a fase de avaliação do jogo, os organizadores do experimento não poderiam ajudar os participantes quanto a possíveis problemas ou sua classificação dentre qualquer uma das heurísticas. Eles poderiam apenas sanar dúvidas quanto ao entendimento das heurísticas e às definições dadas a elas ou à *Playability* em geral. Após essa fase de identificação dos possíveis problemas no jogo, as tabelas de discrepâncias foram lidas e consolidadas em uma única tabela contendo todas as discrepâncias encontradas no jogo juntamente com a técnica com que fora identificada.

Após a consolidação dessa tabela, dois especialistas avaliaram as discrepâncias identificadas em busca de duplicatas (discrepâncias semelhantes identificadas por avaliadores distintos). Tais duplicatas eram consolidadas em apenas uma discrepância geral a qual teria sua descrição, contemplando o contexto geral do problema em questão. Para realizar a fase de discriminação, essa tabela foi entregue a dois especialistas em Engenharia de Software e Interação Humano-Computador, os quais deveriam identificar se tais discrepâncias eram

defeitos reais ou falso-positivos, além de identificar se existiam outras discrepâncias duplicadas ou não. A tabela contendo o número de discrepâncias identificadas por avaliador, número de falso-positivos, defeitos consolidados e número total de defeitos por técnica é apresentada na Tabela 4.2 na seção 4.3 (Resultados) desse capítulo.

4.2.4 Seleção de variáveis

As variáveis independentes desse estudo são os conjuntos heurísticos empregados para a avaliação do jogo em questão e, as variáveis dependentes são a eficácia e eficiência produzida pelos conjuntos heurísticos durante o estudo. A eficiência foi calculada como a razão entre o número de defeitos únicos e o tempo gasto, por pessoa, para avaliar o jogo selecionado. A eficácia foi calculada como a razão entre o número de defeitos encontrados por um avaliador e o número total de defeitos conhecidos do jogo selecionado.

Discrepâncias são definidas como problemas encontrados pelos avaliadores antes da classificação, retiradas de duplicatas e avaliação, por especialistas, de falso-positivos. Após a avaliação por especialistas temos os defeitos identificados no jogo, sem duplicatas. As discrepâncias identificadas pelos avaliadores como falso-positivos indicam que tais discrepâncias não representam um defeito real da aplicação e que devem ser desconsideradas.

4.2.5 Hipóteses

O experimento foi realizado com o intuito de testar as seguintes hipóteses (nulas e alternativas respectivamente) de acordo com os fatores de tratamentos selecionados para o mesmo:

H01: Não existe diferença em termos de eficiência entre a técnica NExPlay e o conjunto heurístico proposto por Barcelos *et al.* (2011).

HA1: Existe diferença em termos de eficiência entre a técnica NExPlay e o conjunto heurístico proposto por Barcelos *et al.* (2011).

H02: Não existe diferença em termos de eficácia a técnica NExPlay e o conjunto heurístico proposto por Barcelos *et al.* (2011).

HA2: Existe diferença em termos de eficácia a técnica NExPlay e o conjunto heurístico proposto por Barcelos *et al.* (2011).

4.3 Resultados

Para a apresentação dos resultados inferiremos o nome dos conjuntos heurísticos e dos grupos que utilizaram tais conjuntos da seguinte forma:

- **Grupo 1**, para os participantes **AV 1** ao participante **AV 12**, os quais utilizaram a técnica NExPlay.
- **Grupo 2**, para os participantes **AV 13** ao participante **AV 24**, os quais utilizaram o conjunto heurístico proposto por Barcelos *et al.* (2011) utilizado como objeto de comparação nesse estudo.
- Além disso, a sigla **AV**, se refere a um participante qualquer chamado neste caso de **Avaliador**, cuja numeração corresponde ao avaliador de mesma numeração nas tabelas: Tabela 4.1 e Tabela 4.2.

Dessa forma, os resultados quantitativos e qualitativos desse estudo são apresentados a seguir.

4.3.1 Resultados Quantitativos

Em números absolutos, a Tabela 4.2, a seguir, mostra a quantidade de discrepâncias, falso-positivos, defeitos por participante, eficiência individual, tempo gasto para cada um dos avaliadores de acordo com a solução utilizada por eles e o total de defeitos por técnica.

Tabela 4.2. Valores absolutos quantitativos por participante do estudo

AV	ND	FP	D	T(h)	E (D/T)	TD
AV - 01	4	0	4	0,67	5,970	NExPlay = 74
AV - 02	4	1	3	0,83	3,614	
AV - 03	9	3	6	1,97	3,045	
AV - 04	11	5	6	1,25	4,800	
AV - 05	6	0	6	0,53	11,320	
AV - 06	7	1	6	1,17	5,128	
AV - 07	3	1	2	1,05	1,904	
AV - 08	8	1	7	0,63	11,111	
AV - 09	14	4	10	1,00	10,000	
AV - 10	18	3	15	0,92	16,304	
AV - 11	5	0	5	1,00	5,000	
AV - 12	6	2	4	0,42	9,523	
AV - 13	20	5	15	1,50	10,000	Barcelos <i>et al.</i> = 101
AV - 14	5	1	4	1,18	3,389	
AV - 15	18	3	15	0,92	16,304	
AV - 16	13	1	12	1,85	6,486	

AV	ND	FP	D	T(h)	E (D/T)	TD
AV - 17	9	1	8	0,60	13,333	
AV - 18	4	1	3	0,63	4,761	
AV - 19	3	0	3	0,57	5,263	
AV - 20	10	0	10	0,78	12,820	
AV - 21	9	0	9	0,92	9,782	
AV - 22	5	1	4	1,35	2,962	
AV - 23	7	1	6	1,23	4,878	
AV - 24	12	0	12	1,42	8,450	

Legenda: AV – Avaliador; ND – Número de Discrepâncias; FP – Falso-Positivos; D – Defeitos; T(h) – Tempo total (em horas); E (D/T) – Eficiência individual (Defeitos / Tempo); TD – Total de Defeitos (por técnica contando duplicatas)

Ao todo, foram encontrados 49 defeitos únicos (defeitos sem duplicatas e contados apenas uma vez e não o total encontrado por todos os participantes) para a aplicação em questão sendo avaliada, considerando os defeitos encontrados pela NExPlay e pelo conjunto de Barcelos *et al.* Dentre os 49 defeitos únicos, a NExPlay identificou 21 defeitos e o conjunto de Barcelos *et al.* identificou 38 defeitos únicos. Na Tabela 4.3, a seguir, apresentamos os valores de eficiência e eficácia contabilizados em relação aos valores totais de cada solução.

Tabela 4.3. Valores médios por solução utilizada

Conjunto	TDU	MD	EM (%)	TTA (h)	E (TDU / TTA)
NExPlay	21	6,16	42,8	9,26	2,2
Barcelos et al.	38	8,41	77,5	10,17	3,7

Legenda: TDU – Total de Defeitos Únicos; MD – Média de Defeitos (TDU / número de avaliadores); EM – Eficácia Média; TTA (h) – Tempo Total de Avaliação (em horas); E (TDU / TTA) – Eficiência do conjunto

Para a avaliação de diferença significativa na avaliação utilizando os conjuntos heurísticos, testes estatísticos foram conduzidos com os valores de eficiência e eficácia coletados durante o estudo. Nesses testes, a eficiência foi considerada como o número de defeitos por avaliador em razão do tempo gasto para realizar a avaliação. De forma análoga, a eficácia foi considerada como o número de defeitos encontrados pelos avaliadores em razão do número total de defeitos identificados durante o experimento.

Dessa forma, para a seleção de qual teste estatístico deveria ser conduzido para a comparação das duas amostras, foram realizados testes de normalidade para cada um dos grupos

e em relação a cada uma das variáveis independentes. Os valores para o teste de normalidade das amostras são apresentados na Tabela 4.4 a seguir:

Tabela 4.4. Valores de significância dos testes de normalidade para os grupos

Eficiência	Shapiro-Wilk	
	Grupo	Significância
	Grupo 1	0,283
	Grupo 2	0,407
Eficácia	Shapiro-Wilk	
	Grupo	Significância
	Grupo 1	0,033
	Grupo 2	0,208

De acordo com os valores apresentados pelos testes de normalidade Shapiro-Wilk, em relação à eficiência dos conjuntos, as variáveis são normalmente distribuídas, uma vez que todos os valores para a significância são acima de 0,05. Sendo assim, de acordo com o tamanho da amostra (24 participantes), devemos utilizar o teste estatístico “*t-Student*” em que o intervalo de confiança desejado é de 0,05.

Analogamente, de acordo com os valores apresentados pelo teste de normalidade Shapiro-Wilk para os valores de eficácia dos conjuntos, a variável não é normalmente distribuída entre os grupos. Dessa forma, de acordo com o tamanho da amostra, devemos utilizar o teste estatístico *Mann-Whitney* em que o intervalo de confiança desejado é de 0,05.

Os testes estatísticos, de normalidade e plotagem de gráficos auxiliares foram realizados através da ferramenta estatística *SPSS (Statistical Package for the Social Sciences)*²⁸. Essa ferramenta foi selecionada por sua simplicidade, familiaridade dos pesquisadores com ela, por sua precisão nos cálculos realizados, além da presença de uma interface gráfica para interação com ela. Além dos testes estatísticos, foi utilizada a análise dos diagramas de caixa (*Boxplots*) para melhor visualização das possíveis diferenças existentes entre os grupos e variáveis. A seguir, a Figura 4.3 mostra o diagrama de caixa para os valores da eficácia e, a Figura 4.2, mostra o diagrama de caixa para os valores da eficiência.

Quando comparados os dois conjuntos em relação à eficiência obtida pelo estudo experimental utilizando o teste estatístico “*t-Student*”, nenhuma diferença estatística foi encontrada entre os grupos (*p-value* = 0,618). Esse resultado sugere que a eficiência do Grupo

²⁸ <https://www.ibm.com/br-pt/products/spss-statistics>

1 (NExPlay) e do Grupo 2 (Barcelos et al.) foi similar quando utilizadas para avaliar jogos digitais de gênero e tipos específicos como o utilizado. Da mesma forma como nos testes estatísticos, os diagramas de caixa demonstraram esse comportamento indicando que os conjuntos teriam eficiências similares. Dessa forma, não é possível anular ou rejeitar a hipótese nula e nem confirmar ou verificar a hipótese alternativa (H01 e HA1) relacionada à eficiência dos conjuntos heurísticos nesse estudo experimental.

De forma análoga, quando comparados os dois conjuntos em relação à eficácia obtida pelo estudo experimental utilizando o teste estatístico “*Mann-Whitney*”, nenhuma diferença estatística foi encontrada entre os grupos ($p\text{-value} = 0,266$). Esse resultado sugere que a eficácia do Grupo 1 (NExPlay) e Grupo 2 (Barcelos *et al.*) são similares quando utilizados para avaliar jogos digitais de gêneros e tipos específicos como os do jogo utilizado. Da mesma forma, os diagramas de caixa demonstraram que, em relação à eficácia, os grupos teriam comportamentos e valores semelhantes. Assim não é possível anular ou rejeitar a hipótese nula e nem confirmar ou verificar a hipótese alternativa (H02 e HA2) relacionada à eficácia dos conjuntos heurísticos nesse estudo experimental.

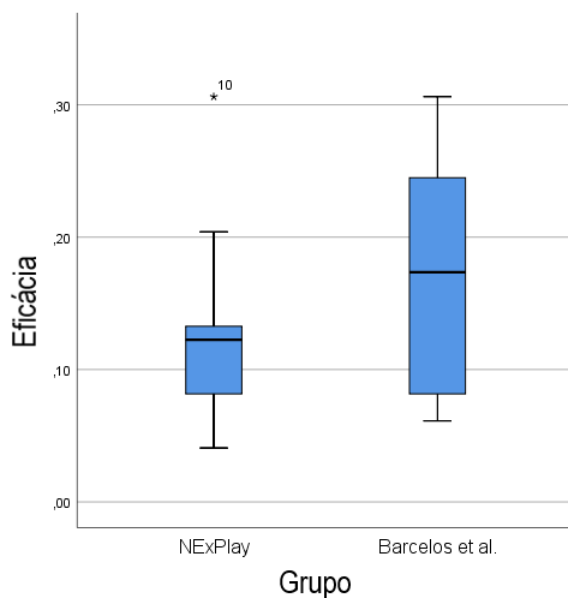


Figura 4.3. Boxplot para os valores de Eficácia

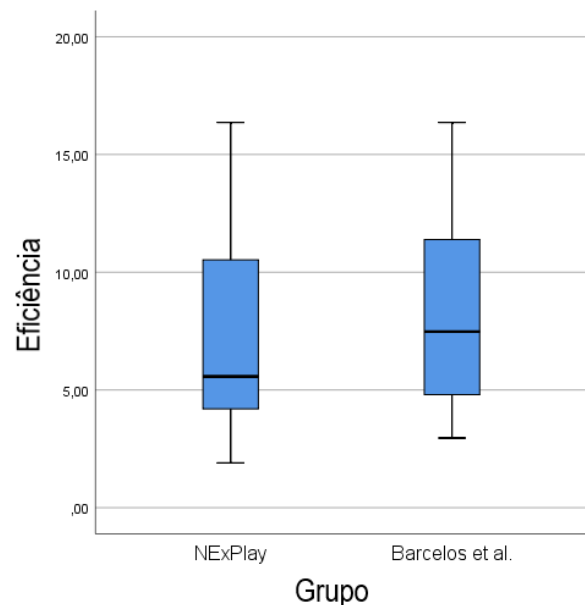


Figura 4.2. Boxplot para os valores de Eficiência

Dessa forma, analisando valores brutos e absolutos obtidos com esse estudo, é possível inferir que a NExPlay não teve um melhor desempenho em relação ao conjunto heurístico proposto por Barcelos *et al.* (2011). A técnica NExPlay teve um menor número de defeitos únicos encontrados na aplicação (21 defeitos únicos) em relação ao conjunto proposto por Barcelos *et al.* (38 defeitos únicos), porém, obteve um tempo total de avaliação menor em

relação ao conjunto proposto por Barcelos *et al.* (9,26 horas contra 10,17 horas). Além disso, existiu um número maior de falso-positivos ao se utilizar a NExPlay (21 falso-positivos) em relação ao conjunto proposto por Barcelos *et al.* (14 falso-positivos). Porém, não é possível deixar de observar que o conjunto NExPlay não teve nenhum de seus avaliadores com conhecimento alto nos dois objetos de avaliação desse estudo como teve o conjunto proposto por Barcelos *et al.* (2011) com o avaliador **AV-13**. Isso se deve ao fato de o balanceamento ter sido realizado de modo a favorecer o conjunto proposto por Barcelos *et al.* (2011). Portanto, em termos quantitativos, o conjunto NExPlay não obteve melhores resultados que o conjunto proposto por Barcelos *et al.* (2011). Porém, mesmo com esses resultados, o conjunto NExPlay ainda se apresenta como uma opção viável na avaliação da Playability de jogos digitais uma vez que a diferença em termos quantitativos não é significativa, sendo, portanto, apenas indícios.

4.3.2 Resultados Qualitativos

Para entender a percepção dos avaliadores em relação ao uso dos conjuntos heurísticos aplicados durante o experimento, foi solicitado que os mesmos preenchessem questionários após a realização do experimento. O questionário foi disponibilizado online aos avaliadores e está disponível no APÊNDICE C. Além desses questionários, entrevistas foram conduzidas com alguns dos avaliadores para que pudessem ser identificadas, com mais detalhes, as dificuldades e percepções em relação aos conjuntos heurísticos. Para a análise dos resultados qualitativos apresentados a seguir, cada uma das entrevistas e questionários de opinião foram analisados individualmente onde o autor marcou e identificou sentenças relevantes em cada uma das mesmas. Após a demarcação das sentenças significantes, as mesmas foram classificadas em termos de categorias que deveriam ser observadas ou em pontos de melhoria que deveriam ser implementados em uma futura melhoria. A partir dessa categorização das sentenças dos usuários, foi feita uma análise geral dos pontos observados e transcritos em cada uma das categorias desenvolvidas. Para prevenir a omissão de informações, as sentenças não foram sintetizadas e sempre que possível, pelo menos duas sentenças foram destacadas, uma positiva e outra negativa. Na ausência de uma sentença negativa, pontos de melhoria foram levantados a partir das frases positivas.

Os resultados a seguir são descritos com relação a cada tipo de problema ou às diferentes visões apresentadas pelos avaliadores.

4.3.2.1 Maneiras de Condução da Avaliação Utilizando Conjuntos Heurísticos

De acordo com as respostas fornecidas pelos avaliadores, cada um deles utilizou os conjuntos de formas diferentes durante a avaliação do jogo. Dentre os doze avaliadores que utilizaram a NExPlay, quatro deles (Avaliador 1, 6, 11 e 12) descreveram que utilizaram o conjunto de heurísticas como um guia para a realização do experimento, verificando se as heurísticas definidas eram respeitadas dentro do jogo ou se eram violadas em algum ponto. Outros quatro avaliadores (Avaliador 2, 4, 8 e 9) jogaram o jogo disponibilizado observando possíveis defeitos e, somente após essa fase, associaram os problemas identificados com uma heurística do conjunto a qual satisfazia suas condições, além de pesquisar as próprias heurísticas dentro do jogo. Esse processo foi explicado pelo avaliador 4 em seu questionário:

“Eu utilizei o jogo e explorei todas as suas possibilidades, então li as heurísticas e analisei cada heurística de acordo com o que eu joguei anteriormente. Para algumas heurísticas, eu retornei ao jogo para verificar se o mesmo possuía tal defeito ou não” – Avaliador 4.

Outros três avaliadores (Avaliador 3, 7 e 10) avaliariam o jogo apenas escrevendo os possíveis problemas durante a seção de avaliação do experimento e então associando tais problemas encontrados com as suas respectivas heurísticas.

Quando analisados os resultados individuais dos avaliadores juntamente com essa informação (veja a Tabela 4.3), é possível observar que, na maioria dos casos, o modo como a avaliação é realizada não é definido através somente do conhecimento prévio dos avaliadores em inspeção ou jogos. Porém, pode ser observado que a maioria dos avaliadores com um grau maior de conhecimento em jogos decidiu realizar uma avaliação do jogo em busca dos problemas e então correlacioná-los com as suas respectivas heurísticas afetadas. Por outro lado, a maioria dos avaliadores que possuía um maior conhecimento em inspeção, escolheu avaliar o jogo inicialmente utilizando as heurísticas como um guia para a sua avaliação.

4.3.2.2 Especificidade das Heurísticas

Com relação a NExPlay, alguns avaliadores apontaram que, devido a decisões de projeto e design das heurísticas utilizadas, a união de um ou mais problemas na mesma heurística afetara seu julgamento no momento de identificação das heurísticas que seriam afetadas pelo problema. Isso pode ser observado através do comentário feito pelo avaliador 11:

“(…), as heurísticas possuem muitos aspectos diferentes sendo considerados em uma mesma heurística. Por exemplo, em algumas heurísticas eu concordava com a primeira parte da

mesma, mas com a segunda parte a mesma não se encaixava no problema que eu avaliava no momento. Logo, eu não assinalava essa heurística para o problema.” – Avaliador 11.

Esse mesmo problema é descrito pelo avaliador 09 em seu questionário além de descrever que gostaria de ver mais heurísticas, dispersas de forma a tornar únicas para seu contexto de avaliação, além de serem mais objetivas nos quesitos de sua avaliação:

“(…), existem heurísticas que nos fazem ficar confusos à medida que as respondemos ou lemos. Isso se deve porque uma parte das heurísticas é verdade e a outra não, quando avaliamos um problema. E quando nós respondemos os problemas, nós não sabemos se devemos descrever de uma forma geral ou responder apenas em relação à parte da heurística que realmente se encaixa (...); as heurísticas deveriam ser divididas, ...minha sugestão seria de revisar as heurísticas e tornar as mesmas mais únicas em seus contextos” – Avaliador 9.

Ou seja, de acordo com os avaliadores, o uso de mais de uma sentença ou problemas semelhantes em uma heurística sendo descrita afetou o entendimento do problema geral descrito por essas heurísticas. Da mesma forma, isso impedia que os avaliadores pudessem assinalar tais heurísticas para certos problemas identificados no jogo. Esse problema seria solucionado quebrando-se as heurísticas que possuíssem tais características em diversas outras heurísticas menores. Porém, fazer essa quebra de heurísticas, causaria o aumento do conjunto no seu número, ferindo um dos princípios listados para esse conjunto de heurísticas.

Sendo assim, uma das soluções para esse problema listado na NExPlay, seria a revisão das heurísticas de modo a construir frases mais objetivas que englobassem os problemas desejados. Sendo assim, apenas uma sentença seria avaliada e seu tamanho não seria afetado.

O mesmo problema foi observado no conjunto de Barcelos et al através dos comentários feitos por seus avaliadores, por exemplo, o avaliador 16 escreveu:

“(…), as heurísticas possuem duas sentenças que precisam ser verificadas, como na heurística A10 – ‘Os gráficos e a trilha sonora devem ajudar a aumentar o interesse dos jogadores’, esta poderia ser separada para mais facilmente ser avaliada e verificada no jogo...” – Avaliador 16.

4.3.2.3 Categorização das Heurísticas

O avaliador 8 (o qual utilizou a NExPlay) apontou que a categorização das heurísticas em *Game Play*, Usabilidade e Mecânicas não seria necessária. Além disso, o mesmo avaliador apontou que dentro dessas categorias existiam heurísticas semelhantes. O mesmo problema pode ser observado no conjunto de Barcelos *et al* em relação à repetitividade e à similaridade entre as heurísticas. Como observado também pelo Avaliador 24:

“Eu encontrei alguns itens similares e ficou confuso de entender a diferença entre eles. Por exemplo, os itens A14 e A18 pareciam falar sobre a mesma coisa e semelhantes” – Avaliador 24.

Em contrapartida, em relação à categorização, o avaliador 02 apontou que, dentro da NExPlay, a categorização das heurísticas nesses três aspectos, ajudou durante a execução da avaliação. Na maior parte, os avaliadores não apontaram exatamente quais pontos, ou heurísticas, o conjunto era parecido, assim como demonstrado pelo avaliador 11:

“A técnica possui muitas heurísticas, algumas muito similares umas às outras (...)” – Avaliador 11.

Analisando os resultados apresentados nessa seção, é possível observar que o uso desse tipo de categorização não afetou a performance durante o processo de avaliação do jogo. Porém, o avaliador 10 descreveu como ponto positivo a utilização de uma separação em categorias:

“Como ponto positivo, eu destaco o pequeno número de heurísticas e a categorização das heurísticas dentro do conjunto (...)” – Avaliador 10.

4.3.2.4 Quantidade de Heurísticas

Como mencionado anteriormente, alguns avaliadores indicaram que o conjunto estaria extenso demais. Uma possível solução seria a criação de conjuntos heurísticos específicos para cada tipo e gênero de jogo, assim, apenas heurísticas significativas para aquela avaliação estariam disponíveis, diminuindo significativamente o tamanho dos conjuntos. Por outro lado, alguns avaliadores especificaram que o conjunto estaria pequeno e abrangente:

“Eu achei o conjunto de heurísticas com boa completude e com tamanho razoavelmente pequeno, não tornando a avaliação cansativa. Além de as mesmas não serem repetitivas” – Avaliador 10.

O Avaliador 6 apontou que as heurísticas apresentadas poderiam não cobrir todos os problemas encontrados no jogo avaliado, como descrito em seu questionário:

“...algumas heurísticas podem não cobrir alguns problemas apresentados pelos jogos avaliados” – Avaliador 6.

Contudo, não ficou claro quais problemas o conjunto de heurísticas não conseguiria avaliar. Alguns avaliadores explicaram que as heurísticas poderiam ser mais objetivas, facilitando entendimento em relação às heurísticas. Isso pode ser observado especificamente na sentença descrita pelo Avaliador 3:

“...os conceitos apresentados são muito extensos, talvez frases mais objetivas sobre as heurísticas deveriam ser formuladas, assim facilitando a interpretação das mesmas pelos avaliadores” – Avaliador 3.

Baseado nesses resultados, estudos futuros deveriam verificar se as heurísticas devem ser reformuladas, reduzindo sua extensão e aumentando sua especificidade.

4.3.2.5 Entendimento das Heurística e Direcionamento por Gênero e Tipo de Jogo

Alguns avaliadores apontaram que a NExPlay pode ser entendido facilmente da forma como as heurísticas estão descritas:

“Eu não encontrei nenhuma dificuldade no entendimento das heurísticas” – Avaliador 2.

“A linguagem utilizada para descrição das heurísticas assim como a técnica são simples e de fácil compreensão, sendo assim, não tive nenhuma dificuldade no entendimento das mesmas, mesmo que eu não possuísse conhecimento prévio em avaliação de jogos” – Avaliador 1.

Em relação à compreensão do conjunto de Barcelos *et al*, algumas dificuldades foram observadas e reportadas pelos avaliadores em relação ao entendimento das heurísticas:

“Eu não consegui associar/avaliar esta heurística: ‘O ritmo do jogo deve levar em conta a fadiga e a manutenção dos níveis de atenção dos jogadores.’ Eu não tinha certeza se o que eu deveria avaliar era o nível de stress ao utilizar o jogo. Como seu objetivo não era claro, eu não avalei a heurística” – Avaliador 20.

“A heurística ‘Os atores digitais e o mundo do jogo devem parecer realistas e consistentes’. Eu não sei se o termo ‘realista’ seria ideal para esta descrição (...). Eu tive esta dúvida sobre o que deveria ser realista no contexto da avaliação realizada” – Avaliador 20.

De acordo com alguns avaliadores, a dificuldade no entendimento ocorreu tanto por conta de algumas palavras utilizadas serem muito específicas a jogos, assim como outras palavras eram confusas para a avaliação do jogo selecionado. Contudo, em alguns casos os avaliadores não especificaram por que encontraram tais dificuldades com essas heurísticas ou palavras. Isso pode ser observado pelos comentários feitos pelos avaliadores 16, 22 e 24:

“A heurística A15 não está clara (...). Eu fiquei hesitante ao utilizar as heurísticas A03 e A05 para o problema no jogo onde o mesmo não apresentava ajudas de contexto” – Avaliador 16.

“Alguns termos pareciam confusos como: fácil, facilidade, rico, desafios e estratégias” – Avaliador 22.

“Eu encontrei alguns itens similares e foi confuso conseguir entender a diferença entre eles. Por exemplo, os itens A14 e A18 pareceram se referir a mesma coisa. Em alguns itens, existiam alguns termos que deixavam o avaliador confuso. Por exemplo, inteligência artificial. Não é claro se a inteligência artificial que o mesmo se refere é a do jogo” – Avaliador 24.

É possível observar durante a avaliação dos problemas reportados pelos dois grupos que os avaliadores que utilizaram o conjunto de Barcelos et al tiveram dificuldades no entendimento das heurísticas. Entretanto, não foram reportados problemas quanto ao entendimento das heurísticas apresentadas pela NExPlay. Estas são indicações de que a NExPlay pode ser mais facilmente entendida, tornando-se uma possível opção para utilização por avaliadores sem experiências prévias em tal contexto.

Alguns avaliadores ainda apontaram que as heurísticas deveriam ser direcionadas para cada tipo de jogo e gênero especificamente. Isto é, deveriam ser criadas diferentes heurísticas as quais poderiam avaliar diferentes tipos e gêneros de jogos existentes, desde que as heurísticas apresentadas são focadas em um contexto geral. Isso pode ser observado na descrição do problema feita pelo avaliador 12:

“...algumas heurísticas não são compatíveis com todos os tipos de jogos. O problema de customização dos controles (M1 e M2) e inteligência artificial (G4), por exemplo, não podem ser adaptadas para o jogo avaliado. Isto me faz imaginar se eu deveria reportar que o jogo não permite esta customização ou se eu deveria considerar que este jogo é simples e não possui a possibilidade de incorporar estes aspectos...” – Avaliador 12.

Além do mais, a falta de categorização dessas heurísticas, com relação ao gênero de jogo avaliado, afetou a avaliação feita pelos avaliadores. A mesma observação foi feita pelo avaliador 23 no conjunto de Barcelos et al como descrito abaixo:

“As heurísticas cobrem relevantes características de jogos em geral, mas tipos de jogos específicos podem não necessariamente possuir todas estas características, o que faz com que os avaliadores reportem problemas que não se aplicam a certos tipos de jogos” – Avaliador 23.

4.3.2.6 Entrevistas

Para um melhor entendimento da aplicação das heurísticas e para realizar as melhorias no conjunto de heurísticas NExPlay, 4 avaliadores foram selecionados para uma entrevista. Esses avaliadores foram escolhidos por apresentarem a menor taxa de problemas reportados, quando

comparados aos outros avaliadores. Dessa forma, entrevistas foram realizadas com os avaliadores 1, 2, 7 e 11.

As entrevistas foram conduzidas de forma a conhecer as razões que levaram os avaliadores a possuírem tais taxas assim como suas dificuldades gerais em relação ao conjunto utilizado. Além disso, era desejável compreender alguns aspectos específicos do conjunto assim como conhecer como eles foram aplicados.

O Avaliador 1 identificou que algumas heurísticas poderiam não ser aplicadas no contexto do jogo avaliado por seu tipo e gênero. Isto é, o jogo selecionado não possuía as características necessárias para ser avaliado pelas heurísticas definidas. Porém, foi considerado que esse é um problema em relação à utilização das heurísticas não dirigidas a tipos e gêneros específicos de jogos, mas heurísticas formuladas de forma a cobrir aspectos presentes em contextos gerais. Ainda relacionado a este aspecto, o avaliador identificou que o gênero e tipo do jogo avaliado durante o experimento era simples demais para realização da inspeção e avaliação completa de todos os aspectos apresentados pelas heurísticas.

O avaliador também explicou que a utilização das heurísticas como um guia para sua avaliação ajudou-o a avaliar todos os aspectos necessários daquele jogo. Isto é, o avaliador leu cada uma das heurísticas e então encontrou uma metodologia para testar tais heurísticas no contexto do jogo avaliado, assim identificando os problemas presentes no jogo. Essa decisão também pode ter sido influenciada pelo fato de que o avaliador não possuía conhecimento prévio de processos de inspeção de software assim como de jogos. O próprio avaliador indicou que sua decisão na condução do experimento se deu pelo seu nível de conhecimento prévio.

Quando questionado sobre suas dificuldades iniciais com o jogo, o Avaliador 2 indicou que o mecanismo de salto do personagem era difícil e que necessitava de tempo consideravelmente grande para aprendê-lo. Porém, o avaliador não indicou esse problema durante sua avaliação do jogo em meio aos problemas descritos. Sendo assim, deve ser verificado se as heurísticas apresentadas aos avaliadores descrevem o problema apresentado por ele ou se a heurística que descreve tal problema pode por ele ser compreendida. O avaliador também relatou que um dos problemas encontrados durante sua avaliação fora identificado diretamente pela leitura prévia das heurísticas de avaliação. Tal problema foi identificado pela heurística G4, a qual descreve a falta de modificações do cenário, dificuldades e surpresas no decorrer do jogo.

Quando questionado sobre a utilização de categorização das heurísticas no conjunto, o Avaliador 2 indicou que tal classificação em grupos distintos ajudara durante sua avaliação, que à utilizara para identificação do que deveria avaliar no jogo. Isto é, o avaliador analisou o jogo

durante sua avaliação separadamente em termos de *Game Play*, Usabilidade e Mecânicas de jogo. Ainda mais, quando teve dificuldades com a classificação dos problemas utilizando as heurísticas, ele se baseou na classificação das heurísticas para realizar a descrição e associação dos problemas encontrados. Considerando tal informação, deve ser considerado uma maior investigação sobre tais categorizações, principalmente em tipos de avaliação semelhantes para identificação de possíveis categorias e sua utilização fora desse contexto.

Assim como apontado em seu questionário, o Avaliador 2 reforçou a necessidade da criação de heurísticas relacionadas a diferentes tipos e gêneros de jogos específicos, uma vez que, em sua maioria, as heurísticas apresentadas foram descritas de forma a generalizar o jogo avaliado. Dessa forma, algumas heurísticas não podem ser adaptadas para a avaliação do jogo proposto, uma vez que não apresentam as características desejadas para isso.

O Avaliador 7 apontou que decidira conduzir o experimento de forma a inicialmente utilizar o jogo e então utilizar as heurísticas para classificar os problemas identificados por conta de sua inexperiência com jogos. O avaliador indicou que não possuiu dificuldades com a linguagem utilizada para descrição das heurísticas. Porém, as heurísticas poderiam não possuir o correto direcionamento uma vez que, durante a entrevista, o avaliador indicou ter enfrentado problemas com as mecânicas de salto do personagem principal, mas não indicou tal problema em sua descrição na tabela de problemas identificados. O avaliador também apontou que as heurísticas da categoria de Mecânicas possuíam descrições muito longas. Tais descrições afetaram as decisões tomadas pelo avaliador no momento da identificação do problema com a sua devida heurística, uma vez que a mesma estava descrita de forma diferente da desejada.

O Avaliador 11 indicou que o conjunto possuía informações repetidas nas heurísticas U9, U10 e U15. Ele indicou que todas essas heurísticas avaliavam o mesmo tipo de objeto referente a informações visuais do jogo. Dessa forma, o avaliador indicou que durante sua avaliação teve que avaliar o mesmo aspecto diversas vezes retirando sua concentração e desenvolvimento. O avaliador indicou que, por o jogo a ser avaliado ser pequeno, simples e repetitivo, o conjunto apresentado era extenso demais para avaliação afetando sua performance. Mesmo com esses problemas identificados, o avaliador indicou que o jogo cumprira seu objetivo de ser simples e interessante por curtos períodos de tempo. A categorização das heurísticas não afetou as avaliações do participante de modo algum. A sua experiência em avaliação e jogos influenciou a forma como ele realizou tal experimento, primeiramente utilizando o jogo e posteriormente avaliando os problemas através das heurísticas apresentadas.

4.3.2.7 Descrição dos Problemas Identificados

Além de coletar a percepção dos avaliadores em relação ao uso da técnica, uma análise das descrições dos problemas apresentados pelos avaliadores foi realizada. Todos os defeitos encontrados pelos dois conjuntos foram analisados por três especialistas de forma a analisar se tais descrições poderiam ser facilmente entendidas. A análise individual indicou que as descrições dos defeitos apresentados pela técnica NExPlay possuíam descrições mais objetivas e simples para o entendimento geral quanto ao problema descrito e situação ocorrida para tal problema.

Em alguns casos de defeitos encontrados pela NExPlay, foi observado que a descrição do problema seguia a descrição da heurística afetada em si, enquanto tal descrição também localizava o problema no cenário do jogo e momento em que ele podia ocorrer. Esse fato não foi observado nos defeitos descritos pelos avaliadores que utilizaram o conjunto descrito por Barcelos et al (2011).

4.3.3 Ameaças a Validade

Ameaças à validade existem em todo tipo de estudo experimental, e podem afetar o resultado final. Dessa forma, avaliá-las e preveni-las se torna parte essencial em todo tipo de estudo experimental. Através da avaliação dessas ameaças é possível levá-las em conta durante a avaliação dos resultados obtidos e assim não criarmos ‘desvios’ tendenciosos neles. Dessa forma, os quatro tipos possíveis de ameaças à validade (Wohlin *et al.*, 2012) são apresentados a seguir em relação ao experimento conduzido nesta pesquisa:

Validade Interna: Foram considerados três problemas principais em relação à incorreta interpretação dos resultados: (1) efeitos de treinamento, (2) classificação da experiência dos avaliadores e (3) medição do tempo. Em relação à primeira ameaça, poderia existir um efeito causado pelo treinamento dado aos participantes da NExPlay se esse fosse de menor qualidade do que o treinamento dado aos participantes do Conjunto 2 e vice-versa. Essa ameaça pôde ser controlada garantindo que o mesmo treinamento fosse ministrado aos participantes dos dois conjuntos de heurísticas de forma a utilizar exemplos genéricos e aplicados a jogos não relacionados em gênero ou tipo com o jogo avaliado por eles.

Em relação à classificação do conhecimento dos participantes, foi uma avaliação realizada pelos próprios participantes de acordo com seus conhecimentos prévios (Jogos e Inspeção), não havendo interferência ou interpretação por parte dos organizadores.

Em relação à medição do tempo, esse pode ter sido um fator de impacto relevante, uma vez que não havia um controle de como os avaliadores anotaram seus tempos assim como não é possível obter certeza se os mesmos anotariam os tempos corretos e reais de avaliação. Em uma tentativa de controlar os efeitos dessa ameaça, os formulários entregues aos avaliadores, em que eles deveriam anotar tais tempos, possuíam campos distintos para início e fim da fase de jogo e para a fase de inspeção.

Validade Externa: O jogo “*Leap of Cat*” utilizado para o experimento não representa todos os tipos e gêneros de jogos existentes. Os avaliadores nesse estudo não eram participantes da indústria de jogos ou especialistas em *Playability*. Porém, isso não pode ser considerado uma ameaça real uma vez que o público-alvo era de avaliadores não especialistas. Com isso, a amostra representa a população desejada, mas não sendo possível generalizar para o artefato avaliado.

Validade de Conclusão: A ameaça seria o tamanho da amostra utilizada, uma vez que o número de participantes não é ideal de um ponto de vista estatístico. Diante deste fato, existe uma limitação nos resultados, sendo eles considerados sinais e não evidências conclusivas.

Validade de Constructo: Para esse tipo de ameaça, as definições para os indicadores de eficiência e eficácia foram consideradas. Tais indicadores são comumente usados em estudos experimentais investigando a detecção de defeitos. Ainda mais, tais indicadores foram utilizados da mesma forma como apresentado em outros estudos da literatura (Korhonen, Paavilainen e Saarenpää, 2009; Valentim, Conte e Maldonado, 2015).

4.4 Discussão

De acordo com os resultados e metodologia de desenvolvimento da NExPlay apresentados nesse capítulo, foi possível identificar diversos pontos que necessitam de melhorias ou resultados que apoiam as decisões de desenvolvimento da técnica proposta. A técnica NExPlay obteve resultados que não podem evidenciar uma melhora em relação aos conjuntos presentes na literatura, contudo, foi possível validar a idealização e desenvolvimento da técnica.

Em termos quantitativos a técnica NExPlay obteve um menor desempenho em relação ao conjunto heurístico proposto por Barcelos *et al.* (2011), de forma que a técnica NExPlay não obteve uma diferença estatística relevante quando comparado com o mesmo. Em termos absolutos, a NExPlay obteve uma quantidade menor de defeitos em relação ao conjunto proposto por Barcelos *et al.* (2011) realizando tal avaliação em um menor tempo. Contudo,

mesmo obtendo um menor tempo de avaliação, essa diferença não se reflete na eficiência do conjunto quando comparado com Barcelos *et al.* (2011). Todavia, deve ser levado em consideração o fato de que o balanceamento realizado para as equipes favorece o Grupo 2 (Barcelos *et al.*) de modo que isso pode ter influenciado a diferença quantitativa observada.

Em relação à avaliação qualitativa, é possível observar diversos pontos que necessitam de melhoria bem como alguns pontos em que pode haver, de acordo com os participantes, melhorias em relação aos conjuntos da literatura. Para que seja possível melhorar o conjunto proposto, será desenvolvida uma nova interação de desenvolvimento do mesmo, de modo a observar os pontos positivos e negativos apontados pelos avaliadores. Um dos principais problemas encontrados com a NExPlay é o fato de que diversas heurísticas necessitam ser direcionadas a tipos e gêneros específicos de jogos, uma vez que cada um dos tipos e gêneros específicos possuem características diferenciadas que necessitam de atenção especial. Além disso, foi reportada uma grande redundância entre algumas heurísticas, que pode ter sido causada durante a minimização dos conjuntos da literatura. Também foi identificado que a classificação das heurísticas em diferentes características ajudou consideravelmente durante a avaliação, sendo necessária a sua melhoria para melhor atender as características dos jogos digitais.

É possível identificar também que a utilização de avaliações utilizando conjuntos heurísticos apoiou diversas decisões dos participantes, assim como guiou os mesmos para a correta identificação dos defeitos nos jogos. Porém, seria necessário definir um processo para a realização da avaliação de modo a incluir outros métodos qualitativos, de forma a completar a avaliação de todas as características existentes nos jogos.

4.5 Conclusão

Nesse capítulo foi apresentada a proposta, metodologia de desenvolvimento e resultados do estudo experimental para a avaliação da técnica NExPlay. O conjunto tem por objetivo central apoiar a avaliação da Playability em jogos digitais realizadas por avaliadores não especialistas ou com pouco conhecimento. O estudo experimental conduzido tinha por objetivo validar a utilização da NExPlay para a avaliação de um jogo aleatório selecionado para dispositivos digitais móveis. Os resultados quantitativos apresentados pelo estudo não podem apoiar uma conclusão em relação ao desempenho do conjunto comparado a outro da literatura. Porém, a partir da análise dos resultados qualitativos é possível identificar que, de acordo com a

percepção dos avaliadores que utilizaram o conjunto, a NExPlay pode ser considerada uma opção viável na avaliação de jogos digitais. O conjunto proposto permitiu a identificação e classificação de defeitos da Playability em um jogo digital aleatoriamente selecionado de forma simples e sem criar barreiras para os avaliadores.

Porém, uma das limitações identificadas para a NExPlay é a de que ela não tem a capacidade de avaliar qualquer tipo e gênero de jogo disponível. Consequentemente, essa limitação entrará nos defeitos identificados para que possa ser corrigido em uma nova versão da técnica. No próximo capítulo será apresentada uma nova versão a qual é desenvolvida especificamente para sua customização de acordo com o tipo e gênero de jogo selecionado para avaliação. Além disso, novos estudos devem ser conduzidos com o objetivo de identificar o impacto causado por tais defeitos da Playability nos usuários finais do jogo digital. A partir dos resultados de novos estudos, é possível identificar como apoiar os desenvolvedores de jogos e melhorar a satisfação dos usuários com tais jogos desenvolvidos.

CAPÍTULO 05 – CustomCheck4Play: UMA TÉCNICA CUSTOMIZÁVEL PARA A AVALIAÇÃO DA PLAYABILITY

Este capítulo apresenta o processo de desenvolvimento e proposta de uma técnica especialmente desenvolvida para ser customizável de acordo com tipo e gênero de jogo sendo avaliado. Da mesma forma como na técnica NExPlay, a técnica CustomCheck4Play apoia a avaliação da Playability através de avaliadores “não especialistas”.

5.1 Introdução

Inicialmente, uma técnica é proposta para minimizar os defeitos identificados durante a avaliação do conjunto NExPlay. Por exemplo, os avaliadores identificaram a necessidade de uma melhor categorização das heurísticas assim como a necessidade de desenvolver heurísticas únicas, sem a dupla identificação de problemas, quando cada heurística podia definir mais de uma situação/defeito. Um dos objetivos principais propostos nesta pesquisa refere-se ao desenvolvimento de uma técnica que possa ser modificada de acordo com o tipo e gênero do jogo sendo avaliado. É desejável desenvolver essa característica em um conjunto heurístico para que ele possa se adaptar ao jogo sendo avaliado e, desta forma, ter a capacidade de avaliar todos os aspectos apresentados por um jogo específico. Com isso, foram definidos os seguintes objetivos na proposta dessa nova técnica:

- **Aprimorar a categorização das heurísticas:** de acordo com os avaliadores durante o estudo de avaliação da NExPlay, a categorização das heurísticas foi útil, porém algumas heurísticas estavam mal classificadas e poderiam existir mais categorias para melhor classificar as heurísticas. Portanto, na proposta desse conjunto, é desejado criar uma categorização que possa melhor classificar as heurísticas e apoiar a criação de categorias que possam ser customizáveis de acordo com tipo e gênero de jogo sendo avaliado.
- **Categorias customizáveis:** a técnica proposta deve ter uma mecânica que possa torná-la customizável de acordo com o tipo e gênero do jogo sendo avaliado. Por isso, de acordo com o tipo e gênero do jogo, as categorias que classificam o conjunto poderão ser incluídas ou não, customizando o conjunto de heurísticas de acordo com suas características.

- **Inclusão e divisão de heurísticas:** de acordo com os avaliadores, durante a avaliação da técnica NExPlay, a técnica se apresentou completa em relação aos aspectos do jogo avaliado. Porém, alguns aspectos do jogo não puderam ser alcançados ou avaliados por falta ou má descrição deles. Além disso, a união de diversas heurísticas em apenas uma, como a técnica NExPlay foi projetada, criou o problema de identificar corretamente o problema existente para uma heurística específica. Dessa forma, desejamos dividir essas heurísticas e melhor classificá-las de acordo com seu contexto para melhor apoiar os avaliadores.
- **Inclusão de sentenças de ajuda para heurísticas:** foi identificado que os avaliadores tiveram dificuldades em diversas heurísticas e, com isso, é desejado incluir uma seção de sentenças de ajuda para as heurísticas em que os avaliadores indicaram essas dificuldades.

5.2 Heurísticas da Técnica CustomCheck4Play

De acordo com os objetivos definidos, foi desenvolvida a técnica CustomCheck4Play, a qual é apresentada a seguir juntamente com sua categorização. Na Tabela 5.1 é apresentada cada categoria definida para o conjunto desenvolvido para a CustomCheck4Play, explicando cada uma das categorias como uma seção de jogo e o que esta seção de jogo apresenta/possui. Na Tabela 5.2, são apresentadas as heurísticas desenvolvidas para a CustomCheck4Play divididas nas categorias apresentadas anteriormente na Tabela 5.1 como as seções de jogo.

Tabela 5.1. Categorias da técnica CustomCheck4Play de acordo com as seções do jogo.

Seções dos jogos	Definição da seção do jogo
Introdução	Cada jogo tem uma motivação inicial para realizar atividades e história, é aqui onde o objetivo principal geralmente será apresentado ao jogador e o tutorial inicial também acontecerá, apresentando a mecânica mais básica do jogo como: movimentos, visão e menus. / informações da tela / status do player.
Apresentação de personagens	Geralmente conectado diretamente com a apresentação inicial da história, os personagens mais importantes que nos acompanham durante a história são apresentados. Além disso, se o jogo se concentrar em um único personagem principal, teremos a personalização de nosso personagem que controlaremos e usaremos ao longo da história.

Seções dos jogos	Definição da seção do jogo
Introdução ao Gameplay	Logo após a história inicial do jogo, é normal ter uma primeira fase / obstáculo em que são apresentadas as mecânicas do jogo. É aqui que podemos ter nossa primeira interação com os controles do jogo. Além disso, a partir daqui, começaremos a perceber se temos todas as informações que queremos e precisamos para continuar jogando, como pontuação, vida, munição e tudo mais na tela.
Desenvolvimento da Gameplay	Ao longo do jogo, várias atividades devem ser realizadas para que metas sejam alcançadas. Nesse momento teremos que avaliar se o jogo pode prender nossa atenção, modificando as formas que temos para atingir esses objetivos. Além disso, é aqui que podemos avaliar se o jogo pode se adaptar ao nosso conhecimento, apresentando novos desafios e obstáculos.
Progresso da História	Cada vez que atingimos um objetivo secundário ou principal, uma nova parte da história será apresentada ao jogador para criar uma ordem cronológica que pode ser modificada pelas ações dos jogadores.
Evolução da Gameplay	Quanto mais avançamos na história, mais o jogo deve evoluir para incluir novos obstáculos e novos desafios, como novos inimigos ou quebra-cabeças.
Pausas de Jogo	Os jogos devem fornecer meios para os jogadores poderem parar a sessão e retomá-la mais tarde.
Ajudas de Contexto e Recuperação de erros	Em vários momentos, alguns jogadores podem ficar presos porque não sabem o que fazer com uma informação / caminho, etc. Nesses momentos, o jogo deve ser capaz de apresentar ajuda ao jogador.
Dificuldade e Níveis Progressivos	Os jogos precisam de diferentes níveis de dificuldade, já que jogadores diferentes, com habilidades diferentes, podem jogá-lo. Além disso, o jogo deve saber como progredir ou regredir a dificuldade de acordo com o tempo jogado.
Configurabilidade e Menus	Os jogos devem ter meios para modificar certas configurações de acordo com cada situação.

Tabela 5.2. Heurísticas da Técnica CustomCheck4Play.

Nº	Heurísticas
Introdução	
H1	O jogo apresenta uma história inicial que justifique as ações do jogador apresentando um objetivo principal.
H2	O jogo deve apresentar um tutorial para familiarização do jogador com as mecânicas e gameplay.

N°	Heurísticas
H3	As terminologias utilizadas para todos os objetos/funções são entendidas pelo jogador.
Apresentação de personagens	
H4	É possível customizar o personagem principal de modo a criarmos uma identidade única a partir do mesmo.
H5	O personagem principal consegue ter uma personalidade e características que prendem nossa atenção acreditando que o mesmo é real.
H6	Os personagens do jogo criam uma relação pessoal com o jogador caracterizando-os como indivíduos reais.
Introdução ao Gameplay	
H7	Todas os símbolos, botões, ícones do jogo devem ser de fácil compreensão pelo jogador.
H8	O jogador consegue obter facilmente informações sobre o estado atual do jogo como pontuações, vida e munições restantes de maneira simples.
H9	Os controles devem ser simples para que o aprendizado seja rápido.
Desenvolvimento da Gameplay	
H10	O jogo apresenta feedback para o jogador de forma adequada.
	Sentença de ajuda: Ou seja, o jogo faz uso correto de músicas / efeitos sonoros / controle de vibrações / avisos visuais quando necessário ao jogador, assim como tem um padrão sendo seguido para essas ações.
H11	Atividades desenvolvidas durante o jogo são variadas de modo a diminuir fadiga.
H12	As habilidades necessárias para se alcançar um objetivo atual ou futuro são conhecidas e ensinadas.
H13	O jogo é balanceado de modo a aplicar pressão no jogador sem frustrá-lo.
	Sentença de ajuda: Ou seja, desafios são experiências positivas ao jogador e o mantém interessado.
Progresso da História	
H14	É possível conquistar um objetivo principal de diversas maneiras e existem objetivos secundários e opcionais.
	Sentença de ajuda: O jogo apresenta diversos objetivos secundários e opcionais os quais complementam um (ou mais) objetivo(s) principal(ais).
H15	É possível entender a cronologia da história de forma quase natural e as ações do jogador interferem na mesma.
H16	O jogo apresenta uma história imersiva que prenda a atenção do jogador.
Evolução da Gameplay	
H17	O jogador percebe alterações no mundo do jogo através de suas ações e progresso
H18	Controles devem ser claros e fisicamente confortáveis, além de suas ações de resposta serem imediatas.
H19	O jogador não deve perder qualquer objeto ganho através de grande esforço.
H20	Os inimigos representam desafios e surpresas inesperadas ao jogador independente do nível selecionado.
H21	O jogador consegue aprender como não errar em atividades semelhantes.

Nº	Heurísticas
	Sentença de ajuda: O jogo não penaliza de forma imediata nos primeiros erros dando oportunidade para o jogador aprender com estes erros.
H22	O jogador consegue se divertir com o jogo.
H23	Os gráficos do jogo são satisfatórios/atrativos. Sentença de ajuda: Mesmo que feito de forma simples, os gráficos devem ser apresentados de forma atrativa para os jogadores.
H24	O Design do jogo é atrativo o suficiente para manter a atenção do jogador. Sentença de ajuda: O design do jogo não causa fadiga no jogador.
Pausas de Jogo	
H25	O jogador deve ser capaz de interromper facilmente a partida a qualquer momento. Sentença de ajuda: Ou seja, interrupções são esperadas e planejadas.
H26	É possível salvar o estado atual do jogo para retomar o mesmo posteriormente do mesmo ponto salvo.
H27	Todas as vezes que o jogador iniciar/recomeçar/continuar o jogo, o mesmo possui informações suficientes para começar a jogar. Sentença de ajuda: Não existe a necessidade de consultar manuais, documentações ou tutoriais constantemente para continuar jogando.
Ajudas de Contexto e Recuperação de erros	
H28	Jogadores devem receber ajudas de contexto durante as fases para que não se encontrem presos nas mesmas.
H29	As ajudas de contexto não devem facilitar demais as fases de modo que não existam desafios ao jogador.
H30	O jogo deve ser capaz de evitar erros não intencionais dos jogadores. Sentença de ajuda: Por exemplo, não permitir que o jogador perca seu slot de salvamento do jogo quando acidentalmente o mesmo apertar um botão errado.
Dificuldade e Níveis Progressivos	
H31	O desafio do jogo pode ser ajustado de acordo com as habilidades dos jogadores.
H32	A dificuldade do jogo é gradual e acompanha o desenvolvimento do jogador. Sentença de ajuda: O jogo segue a dificuldade escolhida pelo jogador, porém, se ajusta ao longo das fases, quanto mais fases o mesmo consegue passar, mais difícil e desafiador o jogo se torna.
Configurabilidade e Menus	
H33	Layouts e menus devem ser intuitivos e organizados de forma que o jogador possa manter o foco no jogo.
H34	O jogo deve permitir customização das configurações e controles.
H35	Os controles devem ser expansíveis para jogadores mais habilidosos.

5.3 Processo de Customização

Um dos principais problemas identificados nas avaliações da Playability é que os conjuntos heurísticos existentes não consideram o tipo de jogo que está sendo avaliado. Cada tipo de jogo

existente tem diferenças muito claras entre eles. Por exemplo, a personalização de personagens e a história não devem ser avaliadas se um jogo não os apresentar. Portanto, é necessário ter um conjunto de heurísticas que possa avaliar apenas características essenciais para cada jogo.

Este problema foi tratado desenvolvendo um novo procedimento através do qual é possível selecionar heurísticas para a avaliação do jogo. A técnica foi desenvolvida usando como base as heurísticas da técnica NExPlay. Como a NExPlay é uma técnica desenvolvida de modo aberto para qualquer tipo de jogo, é possível supor que a CustomCheck4Play é igualmente aberta a qualquer tipo de jogo. Para cada categoria em nossa técnica, é avaliado se há algum tipo de jogo que não exija uma avaliação completa. Assim, um questionário que faz três perguntas principais sobre o jogo pretendido para avaliação é desenvolvido. Primeiro, o tipo principal do jogo, possuindo 9 opções: RPG, MMORPG, Aventura, Emulação, Simulação, Estratégia, Ação, Puzzle e Casual. Através dessa questão, já é possível restringir as categorias que serão necessárias para a avaliação. Para esta questão, cada opção gera saídas diferentes de conjuntos:

- RPG (*Role Playing Game*) – as heurísticas H28 e H29 não são apresentadas; a seção de “Dificuldade e níveis progressivos” não é apresentada – em jogos RPG ajudas de contexto não são necessárias pois o usuário pode seguir o fluxo de história que desejar (H28, H29);
- MMORPG (*Massively Multiplayer Online Role Playing Game*) – as heurísticas H25, H26, H28 e H29 não são apresentadas; a seção de “Dificuldade e níveis progressivos” não é apresentada – por se tratar de um jogo online, pausas de jogo e salvamento de missões não existem (H25, H26); em jogos MMORPG ajudas de contexto não são necessárias pois o usuário pode seguir o fluxo de história que desejar (H28, H29);
- Aventura – as heurísticas H4, H16, H28 e H35 não são apresentadas – jogos de Aventura possuem personagens principais mas que não podem ser customizados (H4); normalmente apresentam histórias simples e lineares com um objetivo simples (H16); este tipo de jogos não necessitam de ajudas de contexto por possuir linearidade e ações repetitivas (H28); com ações repetitivas e poucas opções, controles não necessitam de customizações extensas (H35);
- Emulação – as heurísticas H1, H5, H11, H13, H15, H16, H17, H20, H28, H29 e H35 não são apresentadas – jogos de Emulação não possuem história, por serem emulações de ações do mundo real assim como não ter um personagem principal

(H1, H5); o jogo objetiva Emular situações reais, logo não necessita balancear ou modificar as ações realizadas (H11, H13); normalmente estes jogos não apresentam histórias, mas necessitam de tutoriais extensos (H15, H16); ações podem não surtir efeitos no mundo e pode não existir inimigos em emulações (H17, H20); o jogo não precisa dar ajudas de contexto para não modificar a imersão da emulação (H28, H29);

- Simulação – as heurísticas H1, H5, H9, H11, H17, H28 e H29 não são apresentadas; a seção de “Progresso da História” não é apresentada – normalmente não apresentam personagens principais e não apresentam histórias por simular o mundo real (H1, H5); jogos de simulação podem conter milhares de controles que podem não ser compreendidos rapidamente (H9); não apresenta uma dificuldade que possa ser balanceada (H11); ações não modificam o mundo real (H17); o jogo não precisa dar ajudas de contexto para não modificar a imersão da simulação (H28, H29);
- Estratégia – as heurísticas H1, H11, H28 e H29 não são apresentadas; as seções de “Apresentação de Personagens” e “Progresso da História” não são apresentadas – jogos de Estratégia normalmente não possuem personagens principais (H1); normalmente jogos de estratégia são definidas em níveis antes da seção (H11); o jogo não precisa dar ajudas de contexto uma vez que o jogador deve possuir sua estratégia por si mesmo (H28, H29);
- Ação – as heurísticas H5, H11, H14, H16, H19, H21, H28 e H29 não são apresentadas – jogos de ação nem sempre tem foco em história, apenas nos acontecimentos importantes (H5); normalmente a dificuldade é definida por níveis (H11); a história é linear com objetivos únicos e superficial na ação (H14, H16); o jogador pode perder tudo e os adversários não são lineares, logo o jogador pode errar consecutivamente (H19, H21); o jogo não precisa dar ajudas de contexto por ser linear e com ações repetitivas (H28, H29);
- Puzzles – as heurísticas H19, H20 e H35 não são apresentadas; as seções de “Apresentação de Personagens” e “Progresso da História” não são apresentadas – jogos de puzzle são feitos para o jogador perder tudo e não possuem inimigos (H19, H20); com ações repetitivas e poucas opções, controles não necessitam de customizações extensas (H35);

- Casuais – as seções de “Apresentação de Personagens”, “Progresso da História” e “Ajuda de Contexto e Recuperação de Erros” não são apresentadas;

A segunda questão pergunta se o jogo que está sendo avaliado está em fase de desenvolvimento ou se já está finalizado. Esta questão não modifica o conjunto final apresentado para o avaliador, nem gera nenhum aviso ou alteração visível, mas é aconselhável que jogos em desenvolvimento sejam avaliados de forma mais leve ou de forma menos rigorosa para cada quesito selecionado. Esse processo e conhecimento depende da experiência do pesquisador que estiver conduzindo a avaliação do jogo, além de depender do time de desenvolvimento considerar quais pontos devem ser observados e avaliados ou não. Finalmente, a terceira questão pergunta se o jogo tem essas três características específicas: história progressiva, personalização de personagem e ajudas de contexto durante o jogo. Estas três opções fazem com que cada uma dessas seções em específico sejam incluídas ou não, dependendo se a o avaliador responder que as mesmas estão presentes ou não no jogo sendo avaliado.

O questionário é um formulário on-line que adapta as heurísticas apresentadas pela técnica CustomCheck4Play e fornece ao usuário um conjunto mais adequado para o jogo em avaliação. Cada resposta dada faz com que o questionário online escolha qual categoria de heurística precisa ou não estar no conjunto para a avaliação. O questionário completo pode ser encontrado no APÊNDICE D.

CAPÍTULO 06 – ESTUDO DE VIABILIDADE DA CUSTOMCHECK4PLAY

Este capítulo apresenta um estudo experimental comparativo para verificação da viabilidade da técnica proposta CustomCheck4Play. Este estudo tem como objetivo principal avaliar a viabilidade de utilizar a técnica CustomCheck4Play para a avaliação customizada da Playability em um jogo digital específico por avaliadores “não especialistas”. O estudo apresenta resultados qualitativos e quantitativos de acordo com as métricas de eficiência e eficácia da avaliação.

6.1 Estudo Experimental

Para a avaliação da técnica proposta foi realizado um estudo experimental comparativo para identificar a utilidade e avaliar se os objetivos desejados foram alcançados com o desenvolvimento dela. O objeto de comparação selecionado foi o conjunto heurístico proposto por Barcelos *et al.* (2011). O conjunto proposto por Barcelos *et al.* (2011) foi selecionado por ter sido desenvolvido com uma metodologia semelhante à apresentada no capítulo anterior para a NExPlay além de ter apresentado o melhor desempenho entre todos os conjuntos comparados no processo de Hands-On feito inicialmente nesta pesquisa. O objetivo do autor era apresentar um conjunto de heurísticas reduzido da literatura para avaliar se a quantidade de heurísticas afetaria o tempo de avaliação pelos usuários de forma significativa. Além disso, o conjunto proposto pelo autor não apresenta uma categorização, dessa forma, é desejado avaliar se existe alguma diferença significativa entre avaliações utilizando conjunto categorizado ou não, assim como observar a percepção dos avaliadores que utilizaram o conjunto.

Para avaliar o desempenho da técnica, foi realizado um estudo comparativo com o conjunto proposto por Barcelos *et al.* (2011). O principal objetivo com esse estudo foi comparar a eficiência e eficácia da técnica CustomCheck4Play ao conjunto heurístico proposto por Barcelos *et al.* (2011). Além disso, era desejado identificar se o CustomCheck4Play pode avaliar um jogo de maneira simples e se o conjunto pode oferecer suporte a avaliadores novatos e especialistas a qualquer momento para uma avaliação de Playability. Além disso, o objetivo foi de avaliar se o conjunto poderia cobrir todos os aspectos necessários para uma avaliação de Playability com o jogo selecionado, bem como as percepções dos inspetores.

O estudo experimental foi realizado em duas rodadas em duas universidades diferentes, para que fosse possível cobrir um número maior de pessoas e uma maior diversidade de experiências e conhecimentos entre os inspetores.

Para a realização do estudo, foram utilizados os dois conjuntos distintos em que metade dos participantes utilizou o conjunto CustomCheck4Play e a outra metade, o conjunto heurístico proposto por Barcelos *et al.* (2011). Uma das principais diferenças entre os dois conjuntos é que Barcelos *et al.* (2011) não tem nenhuma separação nas categorias, enquanto o CustomCheck4Play tem a categorização separada por seções cronológicas dos jogos. Outra grande diferença entre os dois conjuntos é que o conjunto proposto nesta pesquisa possui uma seção de sentença de ajuda. Os grupos de avaliadores foram equilibrados para garantir que ambos os grupos tenham participantes com o mesmo nível de conhecimento.

Para equilibrar os dois grupos, em ambas as rodadas, todos os participantes responderam a um questionário de caracterização antes do estudo. O questionário conteve duas questões, primeiro sobre os níveis de conhecimento em jogos em termos gerais. Essa questão foi subdividida em três possíveis respostas: *Heavy Player* (Pessoas que sabem muito sobre muitos tipos diferentes de jogos e dedicam um tempo especial de suas vidas para estudar e jogar esses jogos); *Casual Player* (Pessoas que gostam de algum tipo de jogo específico, mas não dedicam um tempo de suas vidas a ele); Não gosta de jogos.

Para a segunda questão, era desejado identificar se os participantes tinham algum conhecimento prévio / experiência com desenvolvimento de jogos. Essa pergunta teve quatro diferentes respostas possíveis: *Sim, eu já trabalhei em projetos de desenvolvimento de jogos na indústria*; *Sim, já trabalhei em projetos de desenvolvimento de jogos na academia*; *Sim, mas apenas por mim mesmo como curiosidade ou passatempo*; *Não, eu não tenho nenhum conhecimento / experiência anterior com desenvolvimento de jogos*. O questionário de caracterização pode ser encontrado no APÊNDICE E.

6.2 Processo Experimental

Cada processo de avaliação da Playability começa com a seleção do jogo que é desejado ser avaliado. Para o propósito deste estudo, um tipo de jogo casual de um desenvolvedor local foi selecionado, chamado *Leap of Cat* (Figura 6.1), mesmo jogo que foi utilizado no estudo anterior para a avaliação e validação do conjunto heurístico proposto NExPlay. Dessa forma, é possível

realizar comparações entre os dois estudos e as duas técnicas propostas, indicando se houve uma melhoria real na definição da nova técnica CustomCheck4Play. O jogo apresenta um tema e uma história simples, sem grandes mudanças ao longo de seu gameplay ou mudanças súbitas em mecânicas e narrativas. O jogo *Leap of Cat* foi selecionado porque está disponível gratuitamente na loja da Google Play para qualquer dispositivo *Android* compatível. Além disso, o jogo apresenta uma lista extensa de problemas de *Playability*, apesar de ser bem avaliado na loja de aplicativos. Além disso, o jogo em questão foi utilizado no estudo anterior da técnica NExPlay, criando uma base de conhecimento sobre os defeitos do jogo e os falso-positivos previamente identificados. Desta forma é possível realizar a comparação dos estudos para entender a melhoria desenvolvida entre as duas técnicas já que o mesmo objeto de avaliação está sendo utilizado. Por fim, como mencionado no primeiro estudo da NExPlay, o jogo é desenvolvido localmente, tendo apoio dos desenvolvedores do jogo em questão.



Figura 6.1. Figura da tela inicial do jogo *Leap of Cat*

Esse experimento foi projetado como um estudo unidirecional. O design cruzado não foi selecionado por causa da propagação do conhecimento entre a avaliação dos conjuntos. Duas rodadas em diferentes instituições com o objetivo de ter uma população mais diversificada e abrangente foram realizadas.

6.2.1 Primeira rodada

Uma primeira rodada do estudo experimental foi conduzida na Universidade Federal do Amazonas, UFAM (Manaus, Amazonas), com alunos de graduação em Ciência da Computação. Todos os participantes tinham conhecimento prévio em Interação Humano-Computador e Engenharia de Software. 28 estudantes participaram dessa primeira rodada do estudo empírico, sendo selecionados durante o mês de Junho de 2018. Todos os participantes deste estudo não participaram de nenhum estudo anterior relacionado a esta pesquisa ou outra pesquisa que tivesse relação com o assunto em questão. Cada participante preenchia o questionário de caracterização e era designado a um dos grupos que usaria um dos dois conjuntos heurísticos disponíveis. Os participantes foram aleatoriamente designados para ambos os grupos (princípio do desenho aleatório), respeitando seu nível de conhecimento

conforme informado por eles. No final do balanceamento, cada grupo teve 14 participantes e todos os 28 participantes completaram com sucesso o estudo empírico pretendido.

6.2.2 Segunda Rodada

A segunda rodada do estudo empírico foi conduzida na Pontifícia Universidade Católica do Rio Grande do Sul, PUCRS (Porto Alegre, Rio Grande do Sul), com alunos de graduação em Engenharia de Software. Todos os participantes tinham conhecimento prévio em Engenharia de Software e estavam cursando a Interação Humano-Computador. 26 alunos participaram da segunda rodada do estudo empírico durante o mês de Setembro de 2018. Todos os participantes deste estudo não participaram de nenhum estudo anterior relacionado a esta pesquisa ou outra pesquisa que tivesse relação com o assunto em questão. Cada participante preencheu o questionário de caracterização e foi atribuído a um dos grupos que usaria um dos dois conjuntos heurísticos disponíveis. Os participantes foram aleatoriamente designados para ambos os grupos (princípio do desenho aleatório), respeitando seu nível de conhecimento conforme informado por eles. No final do balanceamento, o primeiro grupo teve 11 participantes e o segundo grupo teve 15 participantes e todos os 26 participantes completaram com sucesso o estudo empírico pretendido.

6.2.3 Seleção de variáveis

As variáveis independentes desse estudo são os conjuntos heurísticos empregados para a avaliação do jogo em questão e, as variáveis dependentes, são a **eficácia** e **eficiência** produzida pelos conjuntos heurísticos durante o estudo. A eficiência foi calculada como a razão entre o número de defeitos únicos (conhecidos) para o artefato e o tempo gasto, por pessoa, para avaliar o jogo selecionado. A eficácia foi calculada como a razão entre o número de defeitos encontrados por um avaliador e o número total de defeitos conhecidos do jogo selecionado.

A definição para discrepâncias, defeitos, falso-positivos e o processo para confirmação de defeitos é o mesmo utilizado no estudo anterior e apresentado na seção 4.2.3.

6.2.4 Hipóteses

O experimento foi realizado com o intuito de testar as seguintes hipóteses (nulas e alternativas respectivamente) de acordo com os fatores de tratamentos selecionados para ele:

H01: Não existe diferença em termos de eficiência entre a técnica CustomCheck4Play e o conjunto heurístico proposto por Barcelos *et al.* (2011).

HA1: Existe uma diferença em termos de eficiência ao usar a técnica CustomCheck4Play e o conjunto definido por Barcelos et al (2011).

H02: Não existe diferença em termos de eficácia entre a técnica CustomCheck4Play e o conjunto heurístico proposto por Barcelos *et al.* (2011).

HA2: Existe uma diferença em termos de efetividade ao usar a técnica CustomCheck4Play e o conjunto definido por Barcelos et al (2011).

6.2.5 Execução do Estudo (1a e 2a rodadas)

Antes que qualquer experimento pudesse ser executado, os participantes tinham que ser introduzidos aos conceitos fundamentais da Playability e ao processo de sua avaliação. Foi realizada uma sessão de treinamento sobre os conceitos básicos da Playability e o processo de sua avaliação com todos os participantes em ambas as rodadas. Após o treinamento, foi considerado que todos os participantes tinham níveis semelhantes de conhecimento em Playability e avaliações utilizando heurísticas.

Após o treinamento da Playability, os participantes foram instruídos a jogar e usar a solução designada ao jogo para identificar problemas nele. Além disso, os participantes foram instruídos a escrever todos os problemas encontrados em uma tabela fornecida para eles, bem como a heurística violada para cada problema (APÊNDICE B). Na primeira rodada, a solução designada, a tabela de especificação de problemas e algumas instruções breves foram enviadas para o e-mail pessoal de cada participante. Eles foram instruídos a retornar a tabela de especificação do problema, preenchida no máximo dois dias depois. Além disso, cada participante deveria enviar individualmente sua tabela de especificação de problemas para os organizadores da inspeção.

Na segunda rodada, toda a inspeção foi feita *in loco*, com a presença de dois organizadores de inspeção. As mesmas instruções dadas na primeira rodada foram dadas aos participantes da segunda, limitando a sessão de avaliação do jogo a duas horas. Organizadores de inspeção só podiam interferir se os participantes não tivessem entendido completamente o sentido de algumas palavras em instruções. Se a dúvida se referisse à solução designada ou ao próprio jogo, os organizadores não interfeririam. Por exemplo, se os participantes não compreendessem o processo de instrução de avaliação, os organizadores poderiam interferir para explicar como conduzir o processo de avaliação. Mas se os inspetores tivessem alguma dúvida sobre regras do jogo avaliado ou qualquer coisa relacionada ao jogo, os organizadores não interfeririam.

Cada participante realizou a avaliação do jogo individualmente e anotou os tempos inicial e final na tabela de especificação do problema. Todas as tabelas com discrepâncias identificadas foram enviadas aos pesquisadores ou devolvidas a eles no final da avaliação. Depois de receber todas as tabelas de discrepâncias, os autores criaram uma lista de discrepâncias sem duplicatas das duas rodadas. Essa lista foi analisada por dois especialistas que classificaram cada discrepância como um problema ou um falso-positivo. A Tabela 6.1 mostra a lista final contendo o número de problemas únicos encontrados por cada inspetor.

Tabela 6.1. Valores absolutos para o estudo da técnica CustomCheck4Play.

AV	ND	FP	D	T(h)	E (D/T)	TD
AV - 01	10	0	10	0,57	17,65	CustomCheck4Play = 221
AV - 02	12	2	10	0,65	15,38	
AV - 03	6	1	5	0,6	8,33	
AV - 04	8	1	7	0,35	20	
AV - 05	6	1	5	0,42	12	
AV - 06	8	1	7	0,72	9,77	
AV - 07	14	3	11	1,33	8,25	
AV - 08	12	1	11	0,77	14,35	
AV - 09	9	1	8	0,9	8,89	
AV - 10	8	1	7	0,65	10,77	
AV - 11	7	1	6	0,88	6,79	
AV - 12	9	1	8	0,67	12	
AV - 13	15	3	12	0,42	28,8	
AV - 14	5	0	5	0,67	7,5	
AV - 15	8	0	8	0,25	32	
AV - 16	11	0	11	0,45	24,44	
AV - 17	9	2	7	0,33	21	
AV - 18	11	0	11	0,22	50,77	
AV - 19	17	1	16	0,42	38,4	
AV - 20	10	1	9	0,3	30	
AV - 21	15	2	13	0,28	45,88	
AV - 22	6	1	5	0,35	14,29	
AV - 23	4	0	4	0,27	15	
AV - 24	13	1	12	0,25	48	
AV - 25	15	2	13	0,48	26,9	
AV - 26	8	0	8	0,58	13,71	Barcelos et al. (2011) = 166
AV - 27	8	1	7	0,55	12,73	
AV - 28	9	2	7	0,37	19,09	
AV - 29	7	1	6	0,8	7,5	
AV - 30	6	0	6	0,45	13,33	

AV	ND	FP	D	T(h)	E (D/T)	TD
AV – 31	7	1	6	0,32	18,95	
AV – 32	8	4	4	0,4	10	
AV – 33	6	1	5	0,4	12,5	
AV – 34	9	0	9	0,63	14,21	
AV – 35	8	3	5	0,25	20	
AV – 36	6	3	3	0,23	12,86	
AV – 37	8	0	8	1,12	7,16	
AV – 38	9	2	7	1,08	6,46	
AV – 39	6	0	6	0,47	12,86	
AV – 40	2	1	1	0,25	4	
AV – 41	7	1	6	0,52	11,61	
AV – 42	7	2	5	0,42	12	
AV – 43	7	0	7	0,33	21	
AV – 44	6	0	6	0,5	12	
AV – 45	2	0	2	0,25	8	
AV – 46	6	0	6	0,43	13,85	
AV – 47	7	0	7	0,42	16,8	
AV – 48	5	0	5	0,43	11,54	
AV – 49	4	0	4	0,42	9,6	
AV – 50	9	0	9	0,38	23,48	
AV – 51	5	0	5	0,43	11,54	
AV – 52	6	1	5	0,38	13,04	
AV – 53	5	1	4	0,38	10,43	
AV – 54	7	0	7	0,48	14,48	

Legenda: AV – Avaliador; ND – Número de Discrepâncias; FP – Falso-Positivos; D – Defeitos; T(h) – Tempo total (em horas); E (D/T) – Eficiência individual (Defeitos / Tempo); TD – Total de Defeitos (por técnica contando duplicatas)

6.3 Resultados

Para facilitar o relato dos resultados encontrados, o primeiro grupo, que usou a CustomCheck4Play, será referido como Grupo 1 e, da mesma forma, segundo grupo, que usou o conjunto de Barcelos *et al.*, será referido como Grupo 2. Em cada grupo, as duas rodadas são consideradas e unificadas pela mesma abreviação. Os resultados são apresentados a seguir, separados em resultados quantitativos e qualitativos, agregando as duas rodadas.

6.3.1 Resultados quantitativos

No total, foram encontrados 49 problemas únicos considerando a técnica CustomCheck4Play e o conjunto de Barcelos et al. Na primeira rodada, a técnica CustomCheck4Play encontrou 33 problemas únicos e o conjunto de Barcelos et al encontrou 21. Na segunda rodada do estudo, a técnica CustomCheck4Play encontrou 38 problemas únicos e o conjunto de Barcelos et al encontrou 28. Além disso, foi formulada uma lista unificada de todos os problemas conhecidos no jogo. Essa lista contém todos os problemas únicos identificados em todos os estudos experimentais feitos até hoje com o jogo selecionado, tanto na avaliação do conjunto NExPlay, quanto nessa avaliação e, na Tabela 6.1, são apresentados os resultados individuais de cada participante.

A Tabela 6.2 apresenta as médias calculadas para eficiência e eficácia. Além disso, toda informação apresentada é dada como um total de ambas as rodadas do estudo experimental. Para avaliar estatisticamente os resultados obtidos, foram realizados testes para avaliar a diferença estatística e a significância entre os conjuntos com relação à eficiência e eficácia.

Tabela 6.2. Valores das médias para o conjunto CustomCheck4Play.

Conjunto	TDU	MD	EM (%)	TTA (h)	E (TDU / TTA)
CustomCheck4Play	71	8.84	14.98	13.18	5.39
Barcelos <i>et al.</i>	49	5.72	9.70	13.68	3.59

Legenda: TDU – Total de Defeitos Únicos; MD – Média de Defeitos (TDU / número de avaliadores); EM – Eficácia Média; TTA (h) – Tempo Total de Avaliação (em horas); E (TDU / TTA) – Eficiência do conjunto

Para os procedimentos dos testes estatísticos, foi realizado um teste de normalidade de Shapiro-Wilk para escolher o teste estatístico adequado para comparar as amostras. Como mostra a Tabela 6.3, a variável de eficácia passou no teste de normalidade. Para a variável de eficiência, por não estar distribuída normalmente, foi realizado um teste estatístico de Mann-Whitney com nível de confiança de 0,05.

Para a variável de eficácia, como é normalmente distribuída, foi realizado um teste estatístico T-Student com nível de confiança de 0,05. As análises estatísticas foram realizadas através da ferramenta SPSS, uma ferramenta amplamente utilizada e muito precisa tanto em âmbito acadêmico como em âmbito industrial.

Tabela 6.3. Valores para o teste de normalidade do estudo da CustomCheck4Play.

Eficiência	Shapiro-Wilk	
	Grupo	Significância
	CustomCheck4Play	0.005
Barcelos et al	0.337	

Eficácia	Shapiro-Wilk	
	Grupo	Significância
	CustomCheck4Play	0.252
Barcelos et al	0.212	

6.3.1.1 Eficiência

Primeiramente, para entender a distribuição de dados entre os dois conjuntos, foi realizada uma análise de *boxplot* (diagrama de caixa). O *boxplot* para eficiência é mostrado na Figura 6.2. É possível observar claramente que o CustomCheck4Play teve um desempenho melhor em termos de eficiência. Os resultados do teste estatístico de Mann-Whitney confirmam essa observação, mostrando uma diferença estatisticamente significativa entre a técnica CustomCheck4Play e o conjunto de Barcelos et al ($p\text{-value} = 0,030$), confirmando que a CustomCheck4Play é mais eficiente que o conjunto de Barcelos et al quando usado para avaliar problemas de Playability no jogo especificamente selecionado. De acordo com esses resultados, é possível rejeitar a hipótese nula (H_0) e aceitar a hipótese alternativa (H_A) levantada nesse estudo experimental.

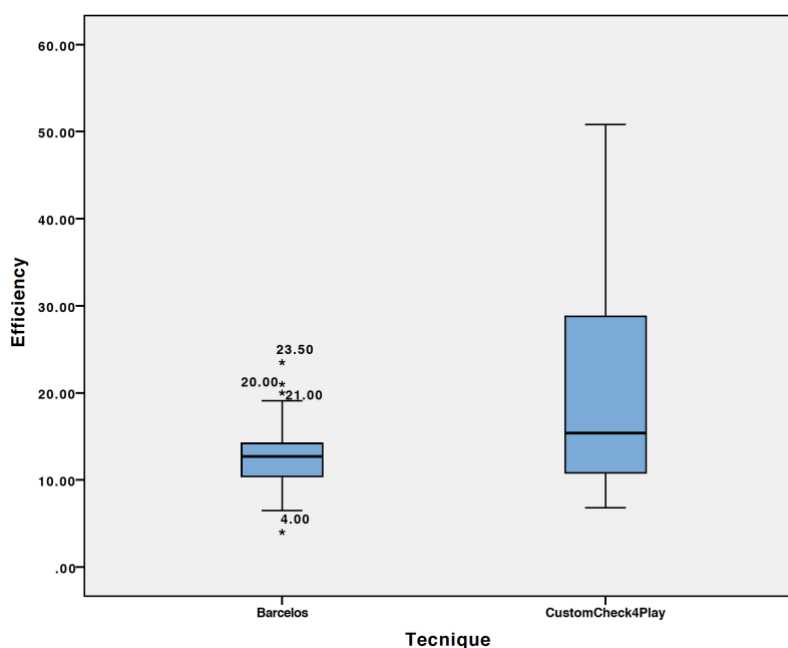


Figura 6.2. Boxplot para a eficiência do estudo CustomCheck4Play.

6.3.1.2 Eficácia

O *boxplot* para eficácia é mostrado na Figura 6.3. É possível observar claramente que o CustomCheck4Play teve um desempenho melhor em termos de Eficácia. Os resultados do teste estatístico T-Student mostram uma diferença estatística significativa entre a técnica CustomCheck4Play e o conjunto de Barcelos et al ($p\text{-value} = 0,004$). Esse resultado confirma que a CustomCheck4Play é mais eficaz que o conjunto de Barcelos et al quando usado para avaliar problemas da Playability no jogo especificamente selecionado. De acordo com esses resultados, é possível rejeitar a hipótese nula (H_0) e aceitar a hipótese alternativa (H_A) formulada nesse estudo experimental.

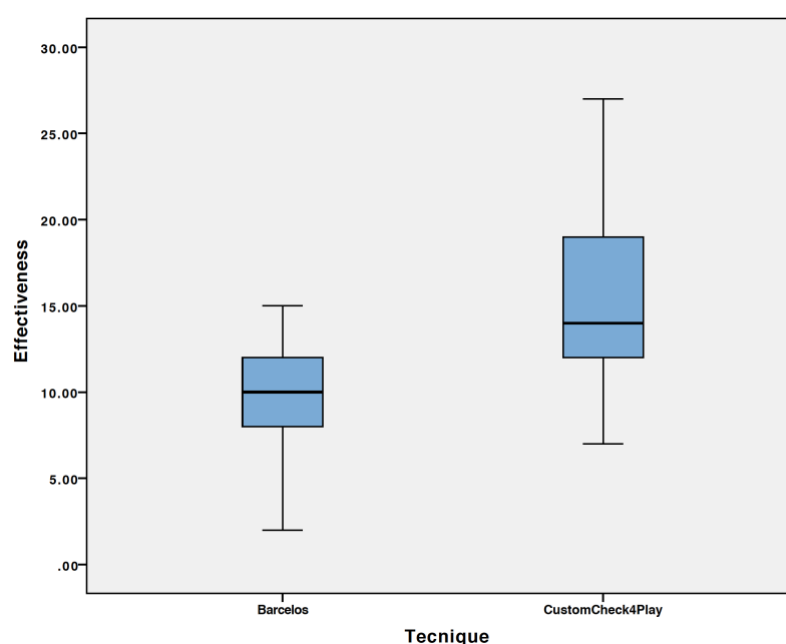


Figura 6.3. Boxplot para a Eficácia do estudo CustomCheck4Play.

6.3.2 Resultados qualitativos

Em cada rodada, ao final do estudo, os participantes preencheram um questionário para analisar sua percepção sobre o uso do conjunto CustomCheck4Play. Na primeira rodada, o questionário foi fornecido on-line e recebido de inspetores juntamente com os relatórios de problemas. Na segunda rodada, o questionário foi fornecido no final do estudo, individualmente, para cada participante que o preencheu no final da sessão. Ele está disponibilizado no APÊNDICE G.

Para a análise dos resultados qualitativos apresentados a seguir, cada uma das entrevistas e questionários de opinião foram analisados individualmente onde o autor marcou e identificou sentenças relevantes em cada uma das mesmas. Após a demarcação das sentenças

significantes, as mesmas foram classificadas em termos de categorias que deveriam ser observadas ou em pontos de melhoria que deveriam ser implementados em uma futura melhoria. A partir dessa categorização das sentenças dos usuários, uma análise geral foi realizada e transcritas em cada uma das categorias desenvolvidas. De modo a prevenir a omissão de informações, as sentenças não foram sintetizadas e sempre que possível, pelo menos duas sentenças foram destacadas, uma positiva e outra negativa. Na ausência de uma sentença negativa, pontos de melhoria foram levantados a partir das frases positivas. Os resultados qualitativos analisados a partir desses questionários são apresentados abaixo.

6.3.2.1 Entendimento e especificidade das heurísticas

Em relação à técnica CustomCheck4Play, a maioria dos inspetores declarou que não tinha dificuldades em entender o que cada heurística significava e que as heurísticas eram claramente escritas:

“Eu não tive dificuldades em entender as heurísticas da solução dada; todas as sentenças foram claramente escritas e fáceis de associar a problemas no jogo.” – Avaliador 08.

O avaliador 06 teve uma dificuldade específica em entender a heurística H25, que avalia a customização de controles e configurações. Analisando o H25, supomos que a dificuldade do avaliador ocorreu porque o jogo selecionado não apresenta esse tipo de personalização:

“Encontrei alguma dificuldade em entender e avaliar a heurística H25 para o jogo selecionado. Além disso, o resto das heurísticas para o conjunto estão bem escritas e não tive dificuldade em compreendê-las.” - Avaliador 06.

Dois avaliadores, 07 e 14, tiveram dificuldades com as heurísticas H3, H8 e H26. Na heurística H3, o avaliador 07 alegou que a terminologia de uma das palavras poderia induzir a erro os inspetores durante a avaliação. Para H8 e H26, os avaliadores não especificaram a dificuldade que tinham. Em relação ao H26, que avalia a mecânica dos controles utilizados no jogo, supusemos que a dificuldade ocorrera porque o jogo selecionado possuía apenas uma mecânica, que não se diferenciava durante o jogo.

6.3.2.2 Viabilidade das sentenças de ajuda

Alguns avaliadores indicaram que a existência de uma sentença de apoio para algumas heurísticas foi fundamental para compreendê-las e que ela não tornava as heurísticas maçantes ou longas demais. Além disso, um dos inspetores sugeriu que toda heurística deveria ter uma sentença de ajuda, com a qual entenderia rápida e facilmente.

“Todas as frases são muito claras; além disso, o conjunto contém uma frase de "ajuda" que me ajudou muito em alguns casos, para que eu pudesse entender rapidamente frases heurísticas.” - Avaliador 11.

“Eu não tive problemas em entender as heurísticas do set. Quando eu tinha alguma, a sentença de apoio sempre foi muito útil.” - Avaliador 23.

6.3.2.3 Número de heurísticas e Heurísticas faltantes

O Avaliador 03 reclamou que sentira a falta de mais heurísticas que discutissem o personagem principal e suas mecânicas. Outro avaliador, o avaliador 04, queixou-se da falta de heurísticas baseadas no conceito da usabilidade, que poderia avaliar problemas de consistência e navegação. Além disso, o avaliador 18 solicitou a inclusão de heurísticas sobre redes sociais e compartilhamento de conteúdo.

“Eu gostaria de mais algumas heurísticas sobre o personagem principal e sua mecânica. No jogo avaliado eu encontrei dois problemas que eu combinei com mais de uma heurística para que eles se complementem.” - Avaliador 03.

“Eu perdi heurística que avaliou coisas como: padrão de navegação, consistência, o jogo não tinha publicidade, mas ainda tem um botão para remoção de anúncios, eu teria assinado como um problema, mas não havia nada para associá-lo.” - Avaliador 04.

“Eu perdi uma heurística para a avaliação de conexão de rede social e compartilhamento de conteúdo do jogo em si.” - Avaliador 18.

6.3.2.4 Categorização da técnica

Por fim, era desejado saber se as alterações feitas na categorização do conjunto, a partir da separação em 10 características baseadas em jogos digitais, foram aceitas pelos avaliadores e poderiam ser utilizadas de forma mais eficiente e eficaz. Avaliando o ponto de vista dos

avaliadores, nenhum dos participantes se queixou da categorização das heurísticas. As categorias fornecidas pela técnica CustomCheck4Play foram suficientes para distribuir corretamente as heurísticas apresentadas. Além disso, a categorização desenvolvida para o conjunto permitiu formular e utilizar a técnica de customização, em que as categorias do conjunto poderiam ou não estar no conjunto usado para a avaliação.

“O conjunto heurístico me apoiou durante a avaliação, principalmente porque era fácil encontrar todas as heurísticas apenas avaliando qual seção eu queria avaliar. Além disso, cobre as partes mais importantes de um jogo através da classificação de seções que qualquer jogo poderia apresentar.” - Avaliador 07.

“... a subdivisão das heurísticas apresentadas apoia de maneira que qualquer heurística fosse facilmente encontrada, basta procurarmos a seção que estamos avaliando.” - Avaliador 20.

6.3.3 Ameaças à validade

Todo estudo experimental apresenta algumas ameaças à sua validade que precisam ser identificadas para que possam ser tratadas durante todo o processo experimental. As ameaças associadas a este trabalho são separadas em quatro categorias e apresentadas abaixo: validade interna, validade externa, validade de conclusão e validade de construto (Wohlin et al., 2012).

Validade Interna: Com relação à medição do tempo, pode haver um erro de medição feito pelos inspetores durante a avaliação. Essa ameaça pode ter um impacto significativo em nossos resultados da primeira rodada, já que não temos nenhum controle sobre como os inspetores registraram o tempo ou se o tempo registrado estava correto. Na tentativa de controlar essa ameaça, a tabela fornecida tinha lugares distintos nos quais os inspetores precisavam registrar seus tempos, e essa etapa específica foi mencionada durante a fase de treinamento. Com relação ao treinamento dado, pode haver efeitos colaterais durante a inspeção. Isso pode acontecer se o treinamento dado aos inspetores da técnica CustomCheck4Play fosse melhor ou pior do que o treinamento dado para os inspetores do conjunto definidos por Barcelos *et al.* (2011), ou entre a primeira e a segunda seção. Para evitar essa ameaça, o mesmo treinamento foi dado para os inspetores de ambos os conjuntos, bem como para as rodadas um e dois. Além disso, o treinamento dado usava apenas exemplos simples que não eram relacionados a nenhum dos dois conjuntos, além de não estarem relacionados ao jogo avaliado. Por fim, em relação aos níveis de conhecimento, esse estudo foi realizado por meio do conhecimento autodeclarado. Os próprios inspetores preencheram o questionário sem qualquer confirmação, independentemente de sua avaliação.

Validade Externa: Para fins de generalização, o jogo usado para a avaliação, “*Leap of Cat*”, não representa todos os tipos de jogos existentes. Além disso, os participantes desse estudo não eram da indústria do jogo, nem eram especialistas em jogabilidade. Portanto, os resultados são limitados até que seja realizada uma investigação mais aprofundada. No entanto, essa não é uma grande ameaça, já que o público-alvo do nosso conjunto são avaliadores não especialistas e jogadores em geral. Portanto, a amostra usada representa nossa população.

Validade de Conclusão: Essa ameaça seria válida se, por exemplo, o tamanho da amostra utilizado não fosse grande o suficiente para conter representatividade estatística. Porém, do ponto de vista estatístico, uma amostra acima de 45 participantes pode ser considerada relevante. Porém, para a população desejada deste trabalho, essa ainda é uma representatividade pequena da população alvo, sendo assim, ainda é possível considerar o tamanho da amostra como uma ameaça se comparada à população total almejada.

Validade de construção: Para esse tipo de ameaça, consideramos a definição de eficácia e eficiência usada para esse estudo. Tais indicadores são amplamente utilizados na literatura para este tipo de estudo e adotamos a mesma definição que Korhonen, Paavilainen e Saarenpää (2009) e Valentim, Conte e Maldonado (2015).

6.4 Discussão

De acordo com os resultados e metodologia de desenvolvimento da CustomCheck4Play apresentados nesse capítulo, é possível concluir que a técnica definida é uma opção viável para a avaliação da Playability de Jogos Digitais. Mas, de acordo com os resultados quantitativos e qualitativos obtidos a técnica CustomCheck4Play é mais eficiente e eficaz que o conjunto definido por Barcelos *et al.* (2011) para a avaliação de um jogo específico em um cenário específico, mas maiores avaliações seriam necessárias para generalização desses resultados. Além disso, é possível definir e viabilizar a avaliação de uma técnica definida especificamente para a realização de avaliações customizadas de jogos digitais de acordo com o tipo e gênero do jogo em questão. E, de acordo com os avaliadores que utilizaram a CustomCheck4Play, todos aplicariam esse conjunto em alguma avaliação futura.

Por fim, nos próximos passos dessa pesquisa, a CustomCheck4Play será utilizada para o desenvolvimento de uma ferramenta de avaliação da Playability de jogos digitais. Essa ferramenta contará com um meio facilitado para a utilização tanto do questionário de customização quanto da técnica para a avaliação de jogos digitais.

CAPÍTULO 07 – ASSISTENTE DE CUSTOMIZAÇÃO PEACH

Este capítulo apresenta o desenvolvimento de um estudo de viabilidade do assistente de customização desenvolvido para o conjunto CustomCheck4Play nomeado PEACH (Playability Evaluation Assistant for Customizing Heuristics). Esse assistente tem por objetivo apresentar o conjunto heurístico CustomCheck4Play, o processo de customização do conjunto, os conceitos básicos para que seja utilizado corretamente na avaliação da playability e efetivamente guiar os seus usuários no processo de customizar e aplicar o conjunto na avaliação da Playability desses jogos. O estudo de viabilidade da PEACH apresenta resultados qualitativos em relação à experiência de uso dos participantes do estudo com o objetivo de observar a capacidade de apoio e o design da ferramenta.

7.1 Introdução

Um dos principais problemas identificados na literatura para a avaliação da Playability de jogos digitais, previamente discutida no CAPÍTULO 01 e CAPÍTULO 02, é de que os conjuntos heurísticos propostos e a maioria dos métodos de avaliação não consideram as diferenças existentes entre os diversos tipos e gêneros de jogos do mercado atual (Desurvire et al., 2004; Desurvire e Wiberg, 2015; *ESA Report on Games Usage*, 2009 & 2019). Cada tipo e gênero de jogo existente possuem claras diferenças entre si. Por exemplo, história e personalização de personagens não deveriam ser avaliados caso o jogo em questão não possuísse esse aspecto. Ao avaliar esse aspecto em um jogo que não apresenta esse tipo de interação com o mundo, seriam gerados problemas que não podem ser relacionados ao jogo que está sendo avaliado. Assim, existe a necessidade de um método de avaliação capaz de avaliar apenas as características essenciais de cada jogo, baseada no seu tipo e gênero. Sendo assim, foi desenvolvido um processo de customização do conjunto heurístico CustomCheck4Play e um website que serve como um assistente no processo de customização desse conjunto heurístico e um assistente para o entendimento do uso e aplicação da avaliação da Playability em jogos digitais. O método de avaliação, e o website de customização devem ser utilizados em conjunto para que a avaliação da playability desses jogos seja feita corretamente. Por fim, todo esse processo pode ser acompanhado e executado a partir da ferramenta web PEACH (*Playability Evaluation Assistant for Customizing Heuristics*).



Figura 7.1. Tela inicial do site da ferramenta PEACH

7.2 Processo de customização

De forma a criar um processo de customização da CustomCheck4Play, para que possa ser utilizada para avaliar especificamente os componentes necessários do jogo em questão, um questionário de customização foi desenvolvido, baseando-se no questionário inicialmente desenvolvido para a CustomCheck4Play e apresentado no CAPÍTULO 05 e mostrado no APÊNDICE D.

Para desenvolver esse questionário de customização das heurísticas, cada categoria de heurísticas pertencente a CustomCheck4Play foi analisada para verificar a existência de algum tipo ou gênero de jogo que não atenderia ou que não necessitasse das heurísticas descritas naquela categoria avaliada. Como as categorias da CustomCheck4Play foram desenvolvidas de forma a serem compatíveis com as características dos oito principais tipos de jogos hoje presentes no mercado, eles foram utilizados para realizar a classificação das heurísticas entre os tipos e definir quais heurísticas seriam ou não necessárias para a avaliação deles. Essa análise resultou em um mapa de conceitos que definem os componentes mais influentes em cada um dos oito tipos de jogos definidos assim como os componentes que não geram nenhum tipo de influência ou não estão presentes neles.

A partir dos resultados dessa análise e do questionário definido anteriormente, um novo questionário, contemplando três seções diferentes, foi desenvolvido para a customização do conjunto da CustomCheck4Play. A primeira seção questiona o tipo e gênero principal do

jogo a ser avaliado, verificando apenas o tipo e gênero principal. Para esta primeira pergunta, 9 opções de respostas estão disponíveis: RPG, MMORPG, Aventura, Emulação, Simulação, Estratégia, Ação, Puzzle e Casual. Se o jogo possuir tipos ou gêneros secundários, os mesmos não são considerados aqui, uma vez que o maior peso nesse momento é do gênero primário do jogo. Essa seção irá servir de base para as seções seguintes do questionário, afinando as opções de resposta e as características que podem fazer parte do jogo em questão. Além disso, é nessa seção que identificamos o nome do jogo, a expertise do avaliador que está customizando o conjunto, além de outras informações pessoais. Essas informações iniciais serão utilizadas no futuro para classificar, identificar e melhorar a customização dos jogos de acordo com as respostas prévias feitas no sistema e o entendimento das pessoas que o utilizam. Os resultados da customização de acordo com a resposta dada para esta primeira questão não foram modificadas e as mesmas seguem o que é apresentado na seção 5.3 de customização da CustomCheck4Play.

Na segunda fase, o avaliador irá identificar características específicas que podem influenciar na game play do artefato de avaliação selecionado. O primeiro aspecto específico avaliado do jogo é se ele possui uma seção ou opção de “*multiplayer online*” parecidos com modos de jogo de time contra time em jogos FPS. O segundo aspecto avaliado é se o jogo apresenta um personagem principal que seja o centro das atenções na história principal do jogo. O terceiro aspecto analisado é se o jogo possui múltiplos personagens na história, se são apresentados e representam alguma importância na história do jogo e sua conclusão. O quarto aspecto avaliado é se o jogo apresenta qualquer tipo de escolha de personalidade ao personagem principal ou se existe alguma decisão que, quando tomada durante o decorrer da história, pode alterar o final ou resultar em mais de um final, dependendo dessas escolhas. O quinto e último aspecto é se o personagem principal do jogo pode ser customizado de qualquer forma para representar o mundo real. Cada ponto identificado acima retira ou não a heurística ou seção específica que é interligada a característica analisada pelo questionário.

Finalmente, a terceira e última seção do questionário pretende identificar perspectivas pessoais dos avaliadores em relação ao jogo alvo a ser avaliado. Assim, o avaliador pode identificar o seu nível de concordância sobre cada um dos aspectos do jogo em relação ao tipo do jogo selecionado. Juntamente com os dados da primeira seção, essa seção é utilizada para coletar dados sobre diferentes aspectos de jogos, a partir da visão do avaliador, utilizando o assistente de customização para a avaliação do jogo.

O questionário é um formulário online que é integrado no site do assistente PEACH. Através desse questionário online, o assistente PEACH é capaz de produzir e entregar aos

avaliadores um conjunto heurístico específico para o jogo e propósito desejado de acordo com as respostas dadas pelos avaliadores. PEACH é um assistente de customização desenvolvido como uma web page baseada na plataforma Google Sites e pode ser acessada através do link: <https://sites.google.com/icompu.ufam.edu.br/peach>.

7.3 Estudo de viabilidade da PEACH e primeiras impressões

Para que possamos avaliar se o assistente desenvolvido pode ser útil e pode ajudar realmente os avaliadores durante o processo de avaliação da Playability dos jogos digitais, um estudo experimental foi realizado com o Assistente de Customização PEACH. Esse estudo experimental foi desenvolvido para avaliar o próprio website e o processo de customização desenvolvido para a PEACH e o conjunto de heurísticas CustomCheck4Play através da avaliação de alguns tipos de jogos. O principal objetivo desse estudo é capturar as perspectivas dos usuários em relação à utilização do assistente de customização. Para isso, aspectos qualitativos do processo de customização pela percepção dos avaliadores foram coletados. Os resultados serão utilizados para identificar melhorias e objetivos que não foram identificados ou cobertos até o momento pelas técnicas desenvolvidas.

21 graduandos em Ciência da Computação da Universidade Federal do Amazonas (UFAM) participaram desse estudo experimental realizado dentro da universidade durante um dos períodos de aula sobre Introdução à Interação Humano-Computador no mês de Junho de 2019. Todos os participantes deste estudo não participaram de nenhum estudo anterior relacionado a esta pesquisa ou outra pesquisa que tivesse relação com o assunto em questão. Para caracterizar melhor o grupo de estudo, todos os participantes responderam a um questionário de caracterização. O questionário conteve duas questões, primeiro sobre os níveis de conhecimento em jogos em termos gerais. Essa questão foi subdividida em três possíveis respostas: *Heavy Player* (Pessoas que sabem muito sobre muitos tipos diferentes de jogos e dedicam um tempo especial de suas vidas para estudar e jogar esses jogos); *Casual Player* (Pessoas que gostam de algum tipo de jogo específico, mas não dedicam um tempo de suas vidas a ele); Não gosta de jogos.

Para a segunda questão, era desejado identificar se os participantes tinham algum conhecimento prévio / experiência com desenvolvimento de jogos. Essa pergunta teve quatro diferentes respostas possíveis: *Sim, eu já trabalhei em projetos de desenvolvimento de jogos na indústria*; *Sim, já trabalhei em projetos de desenvolvimento de jogos na academia*; *Sim, mas*

apenas por mim mesmo como curiosidade ou passatempo; Não, eu não tenho nenhum conhecimento / experiência anterior com desenvolvimento de jogos. O questionário de caracterização pode ser encontrado no APÊNDICE E.

Todos os participantes possuíam conhecimento prévio em Interação Humano-Computador, Engenharia de Software e avaliações de usabilidade e utilização de aplicativos e softwares digitais. Os avaliadores possuíam níveis diferentes no conhecimento de jogos digitais como um todo, assim como no conhecimento de inspeção de softwares digitais.

De acordo com o objetivo de entender qualitativamente a qualidade geral do Assistente de Customização PEACH e seus pontos de melhoria, todos os avaliadores utilizaram o assistente PEACH para customizar dois jogos de diferentes tipos e gêneros. O primeiro jogo selecionado é o jogo de tipo casual e desenvolvido localmente: *Leap of Cat*, que apresenta um tema e história simples e sem maiores modificações durante toda a sua extensão, além disso, o jogo já é bem conhecido por conta de estudos anteriores realizados nesta pesquisa com o mesmo, desta forma, os defeitos do jogo já estão mapeados, facilitando a identificação de falso-positivos. O segundo jogo selecionado foi o jogo de tipo de Ação/Emulação que seria posteriormente avaliado em um estudo comparativo da CustomCheck4Play: *Amazing Rope Police / Amazing Ninja Frog Rope*. Os dois jogos foram selecionados por serem livremente acessíveis para a plataforma Android através da Google Play Store em qualquer dispositivo compatível assim como na plataforma iOS através da Apple App Store em todos os dispositivos compatíveis. Além disso, os dois jogos, que foram avaliados anteriormente ao estudo, apresentam diversos problemas relacionados a Playability, mesmo assim, com avaliações extremamente positivas em suas respectivas lojas de aplicativos.

Antes do estudo experimental ser conduzido, foi necessário introduzir os conceitos fundamentais sobre a Playability para todos os participantes do estudo. Para isso, um treinamento de aproximadamente 30 minutos foi conduzido com todos os participantes. E todos os conceitos principais foram introduzidos. O treinamento consistiu na apresentação dos conceitos principais da Playability para que todos os participantes pudessem ter o mesmo nível de conhecimento sobre Playability.

Todos os participantes foram redirecionados para o Website do Assistente PEACH de forma que, no início do estudo, cada avaliador pudesse desenvolver seu próprio conjunto de heurísticas customizado para cada um dos jogos selecionados. Cada avaliador customizou o conjunto heurístico utilizando o Website do Assistente PEACH. E, após customizar o conjunto, os avaliadores utilizaram estes conjuntos para realizar uma curta avaliação da Playability nos jogos previamente selecionados, em busca de alguns defeitos, de forma que os avaliadores

pudessem ao final darem sua percepção sobre o uso do assistente PEACH. Para o processo de avaliação da Playability dos jogos, foi solicitado aos avaliadores que identificassem no máximo 2 a 3 defeitos em cada jogo com tempo máximo de 10 minutos para cada jogo avaliado. Realizada a customização e o curto processo de avaliação dos jogos, cada avaliador respondeu a um questionário de feedback que seria capaz de avaliar a experiência dos avaliadores com o assistente PEACH, assim como os pensamentos deles sobre o processo de avaliação em que usou o conjunto heurístico customizado (APÊNDICE F). Não foram realizadas análises quanto à correteza ou consistência nas customizações feitas no conjunto pelos avaliadores, uma vez que o objetivo desse estudo era focado na validação do uso do processo de customização do assistente PEACH apenas.

Para a análise dos resultados qualitativos apresentados a seguir, cada uma das entrevistas e questionários de opinião foram analisados individualmente onde o autor marcou e identificou sentenças relevantes em cada uma das mesmas. Após a demarcação das sentenças significantes, as mesmas foram classificadas em termos de categorias que deveriam ser observadas ou em pontos de melhoria que deveriam ser implementados em uma futura melhoria. A partir dessa categorização das sentenças dos usuários, foi feita uma análise geral dos pontos observados e transcritos em cada uma das categorias desenvolvidas. Para prevenir a omissão de informações, as sentenças não foram sintetizadas e sempre que possível, pelo menos duas sentenças foram destacadas, uma positiva e outra negativa. Na ausência de uma sentença negativa, pontos de melhoria foram levantados a partir das frases positivas. Os resultados identificados nas respostas dos questionários são apresentados na próxima seção.

7.4 Resultados

Interface de uso e a usabilidade da PEACH: muitos avaliadores fizeram o processo de avaliação e customização do conjunto heurístico através dos seus smartphones. Por isso, alguns deles, mais especificamente os Avaliadores 01, 08 e 21, reclamaram que as heurísticas disponibilizadas pelo website da PEACH ficaram muito ruins de serem lidas, uma vez que o aspecto de resolução da tela era muito pequeno. De acordo com o que pode ser observado no comentário feito pelo Avaliador 01:

“A interface de usuário e usabilidade do site podem ser melhoradas, ..., o website, pelo menos na sua versão móvel, possui uma usabilidade ruim por conta de necessitar de muito esforço para ler as heurísticas.” – Avaliador 01.

O Avaliador 09 também mencionou que se encontrou perdido em seu primeiro uso, pois não havia entendido que poderia rolar a página para baixo para poder ver o resto do questionário. O Avaliador 03 também indicou que o assistente é muito extenso e poderia se utilizar de alguns cortes de caminho dentro do website e um conjunto menor de ações e processo de avaliação.

O Avaliador 06 descreveu que a página de introdução do assistente PEACH poderia ser mais amigável para apresentar o objetivo do assistente e do processo de avaliação. Na visão deles, os benefícios de utilizar a PEACH não estavam ligados diretamente ao assistente, mas sim ao conjunto e suas heurísticas que estavam muito bem estruturadas.

Descrição dos gêneros e entendimento dos avaliadores sobre os gêneros e tipos de jogos: alguns avaliadores reportaram que, como os jogos possuem diferentes e variadas descrições para apenas um único gênero de jogo, poderia ser incluída uma seção no website que desse uma pequena introdução aos gêneros e tipos de jogos utilizados no questionário do assistente PEACH, assim como suas definições para um melhor direcionamento. Isso pode ser observado no comentário feito pelo Avaliador 02:

“..., o assistente poderia definir melhor algumas informações sobre os tipos e gêneros de jogos assim como uma descrição de como identificar estes tipos e gêneros em qualquer jogo” – Avaliador 02.

O Avaliador 04 sugeriu que as heurísticas poderiam ser apresentadas de forma mais amigável, usando gráficos, imagens, exemplos e outros métodos que chamassem mais a atenção dos usuários. Na opinião do avaliador, tabelas podem ser muito restritivas e não são amigáveis aos usuários.

Simplicidade e entendimento das heurísticas: todos os avaliadores que participaram desse estudo indicaram que as heurísticas eram simples o suficiente para serem compreendidas sem que houvesse a necessidade de ajudas extras. Além disso, identificaram que as heurísticas poderiam ser capazes de apoiar os avaliadores que não possuíssem uma compreensão completa da metodologia de avaliação da Playability dos jogos digitais. Isso pode ser observado pelo comentário feito pelo Avaliador 07:

“O assistente aponta para heurísticas específicas de maneira simples e muito resumida também de maneira a avaliar apenas o importante sem ser extensiva. Além disso, o conjunto e o assistente dão uma boa visão sobre o design dos jogos e seus problemas, até mesmo para os desenvolvedores ...” – Avaliador 07.

Porém, o Avaliador 05 indicou que teve algumas dificuldades com as heurísticas customizadas para o jogo de tipo Casual, uma vez que algumas delas não eram aplicáveis ao

contexto do jogo sendo avaliado. O Avaliador 06 também sugeriu que, de maneira similar ao resto do assistente de customização PEACH, as heurísticas poderiam ter uma apresentação que introduziria cada uma das heurísticas aos usuários e seu uso.

A maioria dos avaliadores concordou e indicou que não existiam heurísticas faltantes nem que sentiram a falta de alguma categoria de heurísticas para os jogos avaliados. Entretanto, os Avaliadores 02, 08, 12, 16 e 18 identificaram pontos de melhoria que poderiam ser executados e modificados no assistente, assim como no conjunto de heurísticas. O Avaliador 02 sentiu falta de heurísticas que pudessem avaliar aspectos mais visuais dos jogos. O Avaliador 08 pediu heurísticas relacionadas com o contexto da Experiência do Jogador (do inglês, “*Player Experience*”). O Avaliador 12 pediu heurísticas relacionadas à linguagem do jogo em relação ao mundo real e r relacionadas para serem mais amigáveis para com o jogador. O avaliador 16 pediu heurísticas que pudessem avaliar jogos multigêneros, heurísticas unificadas para avaliar os jogos multigêneros. Por fim, o Avaliador 18 pediu por heurísticas que pudessem relacionar o mundo real ao mundo digital do jogo em questão e indicar como os dois poderiam ser relacionados e avaliados a partir dessa relação entre os dois mundos. Porém, esses avaliadores não especificaram as situações nas quais essas novas heurísticas poderiam ser utilizadas ou aplicadas.

Entendimento dos itens do questionário e sua sequência: maioria dos avaliadores indicou que o questionário fácil de ser aplicado e possui a sequência correta e lógica de perguntas em sua extensão. Porém, o Avaliador 03 sugeriu que os itens do questionário poderiam incluir a resposta intermediária (“Eu não sei responder a este quesito”). Além disso, o avaliador sugeriu que os itens do questionário poderiam ser sequenciados de uma forma diferente, uma vez que no início do questionário os avaliadores ainda não haviam utilizado o jogo e poderiam responder erroneamente a certos quesitos. Esse mesmo problema também foi reportado pelo Avaliador 20. Porém, deve ser definido e lembrado aqui que o processo de customização dos jogos para a sua avaliação é pretendida e direcionada para ser utilizada por um avaliador que possua conhecimento prévio do jogo a ser avaliado.

Alguns avaliadores também indicaram que os itens de avaliação da PEACH eram muito gerais ao invés de específicos para o objetivo proposto e não indicavam como cada uma das questões influenciaria na customização das heurísticas. Alguns avaliadores também indicaram que a PEACH poderia ser melhor e mais utilizada se houvesse uma maneira de selecionar algumas heurísticas e reportar os defeitos encontrados no jogo diretamente na ferramenta. Ainda mais, a PEACH poderia possuir uma página de reportes e configurações, em

que os avaliadores poderiam avaliar as heurísticas customizadas e reportar ao sistema sua percepção sobre a qualidade e precisão da customização.

De forma bem similar, os avaliadores comentaram que seria interessante uma caixa de sugestões na qual poderiam relatar os problemas com o processo de customização e sugerir como o conjunto e o assistente de customização em si poderiam ser melhorados. Mais ainda, os avaliadores solicitaram que houvesse uma maneira de ver heurísticas retiradas do conjunto após o processo de customização, de modo a melhor indicar e entender o processo de customização do assistente. Finalmente, o Avaliador 20 indicou que seria interessante se, após o processo de customização, o assistente indicasse defeitos comuns da Playability que poderiam ser identificados no tipo e gênero de jogo selecionado.

7.5 Conclusão

Esse capítulo apresentou o conceito do assistente de customização PEACH e a sua construção, assim como o processo de customização definido dentro do assistente e sua utilização. Por fim, apresentou o estudo de validação de uso do processo de customização do assistente PEACH, indicando se era capaz de corretamente guiar seus usuários durante o processo de customização do conjunto. O estudo qualitativo desenvolvido para o assistente PEACH demonstrou que o assistente foi útil para os avaliadores se guiarem no processo de customização do conjunto de heurísticas, além de apoiar os avaliadores no entendimento de todo o processo de avaliação.

Porém, alguns pontos indicados pelos avaliadores ainda necessitam ser aprimorados e melhor desenvolvidos: primeiramente o assistente PEACH poderia possuir um meio de reportar os defeitos da Playability identificados nos jogos diretamente pelo assistente, facilitando e apoiando no processo de identificação de defeitos. Os avaliadores também sugeriram que o assistente PEACH poderia sugerir defeitos da Playability comumente identificados no tipo e gênero de jogo sendo avaliado. Outros avaliadores também identificaram a necessidade de melhoria do design do assistente, incluindo seções mais amigáveis e mais atrativas aos usuários.

Dessa forma, os pontos indicados pelos avaliadores na análise qualitativa, ficam registrados como possíveis trabalhos futuros para o assistente PEACH em uma nova versão de melhoria do assistente. Nesse momento, o assistente PEACH pode apoiar suficientemente os avaliadores no processo de customização das heurísticas e definição das heurísticas a serem utilizadas de acordo com o jogo sendo avaliado.

CAPÍTULO 08 – ESTUDO COMPARATIVO CUSTOMCHECK4PLAY

Este capítulo apresenta um segundo estudo comparativo realizado com a técnica CustomCheck4Play e seus resultados. Esse estudo foi desenvolvido de forma a avaliar a utilização da técnica em comparação à outra técnica da literatura após a utilização do assistente PEACH para a customização das heurísticas. Além disso, um jogo de tipo e gênero diferente dos estudos anteriores foi utilizado de forma a observar a capacidade da técnica em avaliar diferentes tipos de jogos, bem como o objeto de comparação modificado de modo a validar o conjunto de forma mais abrangente.

8.1 Estudo Experimental

O estudo de viabilidade desenvolvido para a CustomCheck4Play, apresentado no Capítulo 6, levou em consideração a avaliação da viabilidade de se utilizar a CustomCheck4Play como uma metodologia para avaliar a Playability de jogos digitais. Para isso, um estudo comparativo foi executado entre a CustomCheck4Play e o conjunto de heurísticas proposto por Barcelos et al. (2011). Nesse estudo, ficou demonstrado que a CustomCheck4Play era mais eficiente e eficaz para a avaliação da Playability de jogos digitais no contexto proposto a partir da avaliação de um jogo específico de tipo casual. Porém, de acordo com o objetivo geral da técnica CustomCheck4Play, foi buscado o desenvolvimento de uma técnica que pudesse se adaptar a diferentes tipos e gêneros de jogos digitais. Neste contexto, um único experimento com apenas um tipo de jogo não é o suficiente para indicar que a técnica proposta realmente cumpre objetivo proposto.

Além disso, para melhor e de forma mais abrangente, validar o uso e a qualidade dos resultados produzidos pela técnica CustomCheck4Play, o estudo foi realizado com um conjunto da literatura diferente do conjunto avaliado anteriormente. Assim, mostrando a capacidade de indicar de forma mais abrangente a qualidade dos resultados produzidos pela técnica de avaliação e como a técnica pode produzir resultados úteis para os seus usuários ou não.

Por fim, de acordo com as alterações realizadas no processo de customização da técnica através do assistente PEACH, é necessário identificar se o conjunto ainda desempenha uma boa performance na análise dos defeitos da Playability de jogos digitais. Mesmo que os resultados entre os dois estudos não possam ser comparados, o estudo ainda deve demonstrar

se a CustomCheck4Play é uma opção viável para a avaliação da Playability de jogos digitais após ser customizada através do assistente PEACH.

Assim como no estudo anterior (apresentado no CAPÍTULO 06), foi pretendido avaliar quantitativamente e qualitativamente o desempenho da CustomCheck4Play em relação à sua eficiência e eficácia em comparação a um conjunto da literatura. O conjunto da literatura selecionado para esse estudo foi o conjunto desenvolvido por Pinelle *et al.* (2018a & 2018b) chamado de *Game Heuristics* (ou, abreviadamente, *GH*). Além de observar os quesitos quantitativos da comparação entre a CustomCheck4Play e o conjunto heurístico GH, era desejado observar aspectos qualitativos da avaliação e utilização da CustomCheck4Play para a avaliação da Playability de jogos digitais. Esses aspectos qualitativos que serão analisados poderão ajudar no futuro com as modificações necessárias para que a CustomCheck4Play possa se adaptar a todas as situações necessárias para avaliar todos os tipos de jogos pré-definidos. Esses aspectos qualitativos da avaliação foram coletados a partir de um questionário pós-uso ao qual todos os participantes responderam ao final do processo de avaliação.

O estudo experimental foi conduzido em duas rodadas em duas universidades diferentes e de estados diferentes no Brasil, sendo possível cobrir um número maior, e significativo, de participantes e uma maior diversidade de experiências e conhecimentos entre todos os avaliadores.

Em cada rodada do estudo, os dois conjuntos foram utilizados. Metade dos participantes utilizaria a CustomCheck4Play e a outra metade, o conjunto heurístico *GH*. Os grupos foram balanceados e divididos partindo do princípio da seleção randômica. Desse modo, era pretendido garantir que ambos os grupos possuíssem participantes de mesmos níveis de conhecimento e, quando não foi possível alcançar a igualdade, o grupo que avaliaria o conjunto heurístico *GH* foi favorecido.

Para realizar esse balanceamento, todos os participantes responderam a um questionário de caracterização antes de o estudo ser realizado. O questionário continha duas questões, sendo a primeira sobre os níveis de conhecimento em jogos em termos gerais. Essa questão foi subdividida em três possíveis respostas: *Heavy Player* (pessoa que sabe muito sobre muitos tipos diferentes de jogos e dedicam um tempo especial de suas vidas para estudar, jogar e completar os desafios desses jogos); *Casual Player* (pessoas que gostam de alguns tipos específicos de jogos, e que jogam esporadicamente em busca de diversão e descontração); Não gostam de jogos.

Para a segunda questão, era necessário identificar se os participantes tinham algum conhecimento prévio/experiência com o desenvolvimento de jogos. Essa pergunta teve quatro

diferentes repostas possíveis: *Sim, eu já trabalhei em projetos de desenvolvimento de jogos na indústria; Sim, já trabalhei em projetos de desenvolvimento de jogos na academia; Sim, mas apenas para mim mesmo como curiosidade ou passatempo; Não, eu não tenho nenhum conhecimento/experiência anterior com desenvolvimento de jogos.* O questionário de caracterização utilizado pode ser encontrado no APÊNDICE E.

8.2 Processo Experimental

Todo processo de avaliação da Playability começa pela definição e seleção do jogo o qual será avaliado em busca de defeitos da Playability. Para o propósito desse estudo, foi selecionado um tipo de jogo de Ação/*Third-Person Shooter (TPS)* produzido por uma empresa de desenvolvimento de jogos indie. Há duas versões diferentes do mesmo jogo para as plataformas iOS e Android com os nomes de: *Amazing Rope Police*²⁹, para usuários da plataforma iOS, e *Amazing Rope Frog Ninja Hero*³⁰ para os usuários da plataforma Android.

O jogo apresenta uma temática simples e história com poucas mudanças durante sua extensão ou com mudanças nas suas mecânicas ou na game play, sem existir uma clara indicação de mudança de história de acordo com as ações realizadas no mundo do jogo. Além disso, possui poucos objetivos primários e quase nenhum objetivo secundário visível no mapa, sendo de rápida execução e de curta duração. Esse jogo foi escolhido por estar gratuitamente disponível para ambas as plataformas (Android e iOS) através da *Google Play Store*, e através da *Apple App Store*, para qualquer dispositivo compatível com ele. Além disso, após uma curta avaliação, foram identificados diversos defeitos de Playability no jogo. Tornou-se simples o trabalho dos avaliadores de identificar problemas no jogo por toda a extensão de sua utilização.

Para esse estudo, o conjunto heurístico a ser utilizado na avaliação foi previamente customizado pela perspectiva da CustomCheck4Play através do assistente PEACH. Para esse cenário, o assistente PEACH retornou um conjunto de heurísticas que não contemplava a categoria de '*Progressão de história*' e todas as suas heurísticas relacionadas foram cortadas sem nenhuma outra alteração em outras categorias e suas heurísticas. Esse comportamento do assistente ocorreu como previsto, uma vez que esse tipo de jogo não deve apresentar histórias muito elaboradas ou com grandes evoluções, assim como não apresenta uma fidelidade a uma

²⁹ Disponível em: <https://apps.apple.com/br/app/amazing-rope-police/id1450015565>

³⁰ Disponível em: <https://play.google.com/store/apps/details?id=com.assassingames.ninjadfrogrope>

história linear, deixando os jogadores mais livres para realizar a exploração do mundo e realizar os objetivos na ordem que preferirem.

Esse experimento foi projetado de forma unidirecional, o design cruzado não foi selecionado por causa da propagação do conhecimento entre as seções de avaliação dos conjuntos. Avaliadores de duas instituições diferentes participaram desse estudo de modo que houvesse um número maior de participantes e uma maior diversificação de conhecimento entre eles.

8.2.1 Primeira rodada

A primeira rodada do estudo experimental foi conduzida na Universidade Federal do Amazonas, UFAM (Manaus, Amazonas), com alunos de graduação em Ciência da Computação e alunos de graduação em Engenharia de Software. Todos os participantes tinham conhecimento prévio em Interação Humano-Computador e Engenharia de Software. 23 estudantes participaram desta primeira rodada do estudo empírico durante o mês de Julho de 2019. Todos os participantes deste estudo não participaram de nenhum estudo anterior relacionado a esta pesquisa ou outra pesquisa que tivesse relação com o assunto em questão. Todos os participantes foram aleatoriamente designados para ambos os grupos (princípio do desenho aleatório) após responder ao questionário de caracterização, respeitando seu nível de conhecimento conforme informado por eles. No final do balanceamento, o grupo que utilizaria a CustomCheck4Play possuía 12 participantes, e o grupo que utilizaria o conjunto heurístico *GH*, 11 participantes. Além disso, todos os 23 participantes completaram com sucesso o estudo empírico pretendido.

8.2.2 Segunda Rodada

A segunda rodada do estudo empírico foi conduzida na Pontifícia Universidade Católica do Rio Grande do Sul, PUCRS (Porto Alegre, Rio Grande do Sul), com alunos de graduação em Engenharia de Software. Todos os participantes tinham conhecimento prévio em Engenharia de Software e estavam cursando a matéria de Interação Humano-Computador. 44 alunos participaram da segunda rodada do estudo empírico durante o mês de Junho de 2019. Todos os participantes deste estudo não participaram de nenhum estudo anterior relacionado a esta pesquisa ou outra pesquisa que tivesse relação com o assunto em questão. Todos os

participantes foram aleatoriamente designados para ambos os grupos (princípio do desenho aleatório) após responderem ao questionário de caracterização, respeitando seu nível de conhecimento conforme informado por eles. No final do balanceamento, o primeiro grupo tinha 21 participantes que utilizariam a CustomCheck4Play e o segundo grupo tinha 23 participantes que utilizariam o conjunto heurístico *GH*. Além disso, todos os 44 participantes completaram com sucesso o estudo empírico pretendido.

8.2.3 Seleção de variáveis

Esse estudo utilizou os mesmos conceitos de variáveis independentes e dependentes, considerando a eficiência e eficácia. O conceito das variáveis dependentes de eficiência e eficácia foram os mesmos utilizados nos estudos anteriores. Da mesma forma, as definições de discrepâncias, problemas, falso-positivos e o processo de definição dos defeitos foram os mesmos dos utilizados nos estudos anteriores.

8.2.4 Hipóteses

O experimento foi realizado com o intuito de testar as seguintes hipóteses (nulas e alternativas respectivamente) de acordo com os fatores de tratamentos selecionados para ele:

H01: Não existe diferença em termos de eficiência entre a técnica CustomCheck4Play e o conjunto heurístico *GH*.

HA1: Existe uma diferença em termos de eficiência ao usar a técnica CustomCheck4Play e o conjunto heurístico *GH*.

H02: Não existe diferença em termos de eficácia entre a técnica CustomCheck4Play e o conjunto heurístico *GH*.

HA2: Existe uma diferença em termos de efetividade ao usar a técnica CustomCheck4Play e o conjunto heurístico *GH*.

8.2.5 Execução do Estudo

Antes que qualquer experimento pudesse ser executado, os participantes tinham que ser introduzidos aos conceitos fundamentais da Playability e a seu processo de avaliação. Foi realizada uma sessão de treinamento sobre os conceitos básicos da Playability e seu processo

de avaliação com todos os participantes em ambas as rodadas. Após o treinamento, foi considerado que todos os participantes tinham níveis semelhantes de conhecimento em Playability e avaliações utilizando heurísticas.

Após o treinamento da Playability, os participantes foram instruídos a jogar e usar a solução designada ao mesmo para identificar problemas no jogo. Além disso, os participantes foram instruídos a escrever todos os problemas encontrados em uma tabela fornecida para eles, bem como a heurística violada para cada problema (APÊNDICE B).

Em ambas as rodadas, todo o processo de inspeção foi feito *in loco*, com a presença de dois organizadores de inspeção. No início do processo de inspeção, a solução designada a cada participante, a tabela de especificação de problemas e algumas instruções breves foram entregues a cada um deles. Eles foram instruídos a retornar a tabela de especificação dos problemas preenchida ao final do processo de inspeção. Em ambas as rodadas, o tempo de inspeção foi limitado a 2 horas. Além disso, cada participante deveria preencher e entregar individualmente sua tabela de especificação de problemas para os organizadores da inspeção. Os organizadores da inspeção podiam apenas interferir se os participantes não tivessem entendido completamente o sentido de algumas palavras nas instruções ou se dúvidas quanto ao processo de inspeção surgissem. Se a dúvida se referisse à solução designada ou ao próprio jogo, os organizadores não interfeririam. Por exemplo, se os participantes não compreendessem o processo de avaliação do jogo, os organizadores poderiam interferir para explicar como conduzir o processo de avaliação. Mas se os inspetores tivessem alguma dúvida sobre regras do jogo avaliado ou qualquer coisa relacionada ao jogo, os organizadores poderiam interferir.

Cada participante realizou a avaliação do jogo individualmente e anotou os tempos inicial e final na tabela de especificação dos problemas. Todas as tabelas com discrepâncias identificadas foram entregues aos pesquisadores no final da avaliação. Depois de receber todas as tabelas de discrepâncias, os autores criaram uma lista de discrepâncias de ambas as rodadas, sem duplicatas. A lista produzida sem discrepâncias duplicadas foi analisada por dois especialistas que classificaram cada discrepância como um problema ou um falso-positivo. Em três situações (avaliadores 12, 20 e 43) o número de defeitos encontrados foi superior ao número de discrepâncias porque, nos três casos, os avaliadores identificaram dois defeitos distintos em uma mesma descrição. Sendo assim, tais defeitos foram classificados em dois defeitos distintos, elevando em 1 o número de defeitos encontrados em relação às discrepâncias. A Tabela 8.1 mostra a lista final contendo o número de problemas únicos encontrados por cada inspetor.

Tabela 8.1. Valores quantitativos absolutos por avaliador.

AV	ND	FP	D	T (h)	E (D/T)	TD
AV - 01	16	1	15	0,83	18,0	
AV - 02	12	0	12	1,00	12,0	
AV - 03	19	3	16	1,12	14,3	
AV - 04	9	1	8	1,25	6,4	
AV - 05	10	2	8	1,17	6,9	
AV - 06	7	0	7	0,82	8,6	
AV - 07	7	0	7	1,15	6,1	
AV - 08	11	0	11	0,87	12,7	
AV - 09	8	0	8	0,92	8,7	
AV - 10	11	0	11	0,92	12,0	
AV - 11	20	3	17	0,50	34,0	
AV - 12	19	0	20	0,82	24,5	
AV - 13	8	0	8	0,43	18,5	
AV - 14	3	1	2	0,17	12,0	
AV - 15	9	1	8	0,55	14,5	
AV - 16	11	2	9	0,48	18,6	
AV - 17	15	2	13	0,58	22,3	CustomCheck4Play = 284
AV - 18	7	2	5	0,17	30,0	
AV - 19	13	1	12	0,67	18,0	
AV - 20	9	0	10	0,50	20,0	
AV - 21	6	1	5	0,50	10,0	
AV - 22	9	1	8	0,42	19,2	
AV - 23	4	1	3	0,33	9,0	
AV - 24	5	1	4	0,58	6,9	
AV - 25	5	1	4	0,50	8,0	
AV - 26	7	1	6	0,50	12,0	
AV - 27	4	1	3	0,30	10,0	
AV - 28	9	0	9	0,62	14,6	
AV - 29	12	0	12	0,62	19,5	
AV - 30	9	2	7	0,57	12,4	
AV - 31	6	1	5	0,58	8,6	
AV - 32	8	0	8	0,45	17,8	
AV - 33	3	0	3	0,33	9,0	
AV - 34	24	9	15	0,92	16,4	Conjunto GH = 216
AV - 35	7	0	7	0,77	9,1	
AV - 36	13	1	12	0,85	14,1	
AV - 37	13	0	13	1,08	12,0	
AV - 38	11	4	7	0,78	8,9	
AV - 39	7	2	5	0,87	5,8	

AV	ND	FP	D	T (h)	E (D/T)	TD
AV - 40	12	0	12	1,03	11,6	
AV - 41	20	4	16	0,48	33,1	
AV - 42	28	6	22	1,20	18,3	
AV - 43	6	0	7	0,92	7,6	
AV - 44	9	2	7	0,72	9,8	
AV - 45	1	0	1	0,50	2,0	
AV - 46	2	0	2	0,53	3,8	
AV - 47	18	6	12	0,48	24,8	
AV - 48	5	3	2	0,50	4,0	
AV - 49	5	0	5	0,52	9,7	
AV - 50	7	1	6	0,50	12,0	
AV - 51	5	2	3	0,67	4,5	
AV - 52	6	2	4	0,52	7,7	
AV - 53	4	0	4	0,35	11,4	
AV - 54	5	2	3	0,50	6,0	
AV - 55	9	3	6	0,47	12,9	
AV - 56	4	1	3	0,67	4,5	
AV - 57	3	1	2	0,47	4,3	
AV - 58	6	2	4	0,58	6,9	
AV - 59	5	1	4	0,40	10,0	
AV - 60	5	0	5	0,58	8,6	
AV - 61	4	0	4	0,42	9,6	
AV - 62	1	0	1	0,50	2,0	
AV - 63	5	1	4	0,25	16,0	
AV - 64	5	2	3	0,25	12,0	
AV - 65	6	3	3	0,25	12,0	
AV - 66	6	0	6	0,47	12,9	
AV - 67	6	0	6	0,48	12,4	

Legenda: AV – Avaliador; ND – Número de Discrepâncias; FP – Falso-Positivos; D – Defeitos; T(h) – Tempo total (em horas); E (D/T) – Eficiência individual (Defeitos / Tempo); TD – Total de Defeitos (por técnica contando duplicatas)

8.3 Resultados

Para facilitar o relato dos resultados encontrados, o primeiro grupo, que usou a técnica CustomCheck4Play, será referenciado como Grupo 1 e, da mesma forma, o segundo grupo, que usou o conjunto *Game Heuristics* (GH), será referenciado como Grupo 2. Em cada grupo, as duas rodadas do estudo são consideradas e unificadas pela mesma abreviação. Os resultados são apresentados a seguir, separados em resultados quantitativos e qualitativos, agregando as duas rodadas.

8.3.1 Resultados quantitativos

De maneira geral, 131 defeitos únicos foram encontrados no jogo (defeitos sem duplicatas e identificados no jogo, possíveis melhorias a serem feitas nele), considerando os dois métodos de avaliação, a CustomCheck4Play e o GH. Na primeira rodada, os participantes que utilizaram a CustomCheck4Play encontraram 62 defeitos, e os participantes que utilizaram o GH encontraram 67. Na segunda rodada, participantes que utilizaram a CustomCheck4Play encontraram 52 defeitos, e os participantes utilizando o GH encontraram 55.

A Tabela 8.2 apresenta as médias calculadas para a eficiência e a eficácia de cada técnica, unificando ambas as rodadas. Uma avaliação estatística e testes estatísticos foram realizados para avaliar a diferença e a significância entre os conjuntos que foram avaliados, com respeito a sua eficiência e eficácia.

Tabela 8.2. Médias calculadas de acordo com os resultados quantitativos.

Conjuntos	TD	M	EM (%)	MT	AEF (P/H)
CustomCheck4Play	84	8.12	6.57%	0.64	14.39
GH	95	4.18	4.85%	0.60	10.49

Legend: **TD** – Total de Defeitos; **M** – Média de Problemas por Participante; **EM** – Eficácia Média por participante; **MT** – Média de Tempo por Participante (em Horas); **AEF** – Eficiência Média por participante (Problemas/Hora)

De acordo com a Tabela 8.2, participantes que utilizaram o GH encontraram mais problemas que os participantes que utilizaram a CustomCheck4Play (no número total de problemas), porém, avaliadores individuais utilizando a CustomCheck4Play encontraram mais defeitos em média (média de problemas encontrados por participante), de forma mais rápida (eficiência média), assim como cada participante foi mais efetivo em encontrar os defeitos (eficácia média).

Para os procedimentos dos testes estatísticos, como o tamanho das amostras de cada grupo é menor que 50 participantes, foi realizado um teste de normalidade de *Shapiro-Wilk* para escolher o teste estatístico adequado para ser possível comparar as amostras corretamente. Como mostra a Tabela 8.3, nenhuma das duas variáveis (eficiência ou eficácia) foram consideradas normalmente distribuídas, desta forma, um teste estatístico de Mann-Withney com um nível de confiança $\alpha=0.05$ foi executado nas duas variáveis. Todos os testes estatísticos, plotagem dos gráficos e análises de dados foram desenvolvidos e rodados

utilizando a ferramenta SPSS³¹ para Análises Estatísticas, uma ferramenta amplamente utilizada tanto no âmbito acadêmico como no âmbito industrial. A seguir, os resultados individuais para cada variável, eficiência e eficácia, são mostrados e comparados para identificar sua performance na avaliação da Playability de jogos digitais.

Tabela 8.3. Valores do teste de normalidade para a eficiência e eficácia.

	Conjunto	p-value
Eficiência	Teste de Shapiro-Wilk	
	CustomCheck4Play	0.007
	GH	0.001
Eficácia	Teste de Shapiro-Wilk	
	CustomCheck4Play	0.125
	GH	< 0.001

8.3.1.1 Eficiência

Primeiramente, para entender a distribuição de dados entre os dois conjuntos, foi utilizada uma análise de *boxplot* (diagrama de caixa). O *boxplot* para eficiência é mostrado na Figura 8.1. É possível observar que a CustomCheck4Play teve um desempenho melhor em termos de eficiência em sua média e nos seus extremos, mas com uma maior dispersão e menos *outliers*. Os resultados do teste estatístico de Mann-Whitney confirmam essa observação, mostrando uma diferença estatisticamente significativa entre a CustomCheck4Play e o GH ($p\text{-value} = 0,010$), confirmando que a CustomCheck4Play é mais eficiente que o GH quando usado para avaliar problemas de Playability no jogo especificamente selecionado. De acordo com esses resultados, é possível rejeitar a hipótese nula (H_0) e aceitar a hipótese alternativa (H_A) formuladas nesse estudo experimental. Além disso, de forma a compreender melhor as diferenças entre os dois grupos, uma análise de “*effect size*” foi desenvolvida de acordo com o que foi definido no teste por Cohen’s (Cohen J., 1988) para as duas variáveis. Para a variável de eficiência, o valor do “*effect size*” foi de $d=0.81$, e, de acordo com a escala desenvolvida por Cohen’s (Cohen J., 1988), esse resultado demonstra que existe uma **grande** diferença entre os dois grupos em relação a sua eficiência. Para propósitos práticos, isso significa que existe uma **grande vantagem** em utilizar a CustomCheck4Play ao invés do GH, pois a CustomCheck4Play é capaz de encontrar mais problemas de uma maneira mais rápida durante o processo de avaliação.

³¹ <https://www.ibm.com/br-pt/products/spss-statistics>

8.3.1.2 Eficácia

O *boxplot* para eficácia é mostrado na Figura 8.1. É possível observar claramente que a CustomCheck4Play teve um desempenho melhor em termos de Eficácia, com uma maior dispersão e sem apresentar nenhum *outlier*. Os resultados do teste estatístico Mann-Whitney confirmaram essa observação, mostrando uma diferença estatística significativa entre a CustomCheck4Play e o GH ($p\text{-value} = 0,008$). Os resultados mostram que a CustomCheck4Play é mais eficaz que o GH quando utilizados para acessar os problemas da Playability no jogo especificamente selecionado. Sendo assim, de acordo com esse resultado, é possível rejeitar a hipótese nula (H_0) e aceitar a hipótese alternativa (H_A) formuladas nesse estudo experimental. Para os valores da eficácia, o valor do “*effect size*” foi de $d=0.36$, e, considerando a escala desenvolvida por Cohen’s (Cohen J., 1988), esse resultado representa uma **pequena** diferença entre os grupos. Na prática, existe um **pequeno ganho** no número de problemas identificados quando utilizando a CustomCheck4Play ao invés do GH.

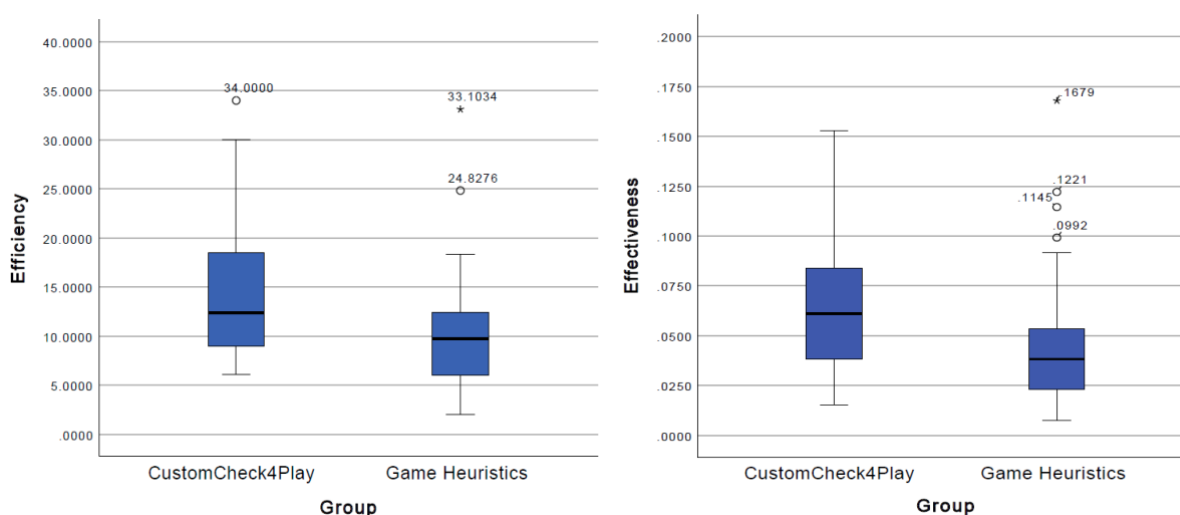


Figura 8.1. Boxplot para os valores da Eficiência e Eficácia deste estudo.

8.3.2 Resultados qualitativos

Em cada rodada, ao final do estudo, os participantes preencheram um questionário para analisar sua percepção sobre o uso do conjunto CustomCheck4Play. Em ambas as rodadas, o questionário foi fornecido no final do estudo, individualmente, para cada participante preenchê-lo no final da sessão. O questionário é disponibilizado no APÊNDICE G. Além dos questionários, foram realizadas entrevistas com 8 participantes do estudo (4 participantes de cada rodada) os quais utilizaram a CustomCheck4Play. Apenas o áudio das entrevistas foi

gravado e então as respostas foram transcritas para que pudessem ser analisadas por um pesquisador através de um processo de codificação aberta. O questionário também foi avaliado através do mesmo processo de codificação aberta aplicado às entrevistas gravadas.

Foram utilizados dois métodos diferentes para obter dados qualitativos por dois motivos. O primeiro motivo é que o questionário aplicado focou em perguntas fechadas, utilizando escalas e diretamente perguntando se a experiência com o método de avaliação foi boa ou não, se o método, de modo geral, foi fácil de ser utilizado, e assim por diante. De outra forma, as entrevistas foram desenvolvidas para melhor entender aspectos diferentes e específicos do método de avaliação e o entendimento dos avaliadores. Por exemplo, inicialmente, foi questionado se o avaliador percebera a presença das frases de apoio nas heurísticas, e se sim, se a utilizou, como e onde foram utilizadas, finalmente chegando à questão da sua experiência utilizando os componentes do método de avaliação. Dessa forma, foi possível um melhor e mais profundo entendimento da percepção de uso dos avaliadores, assim como uma percepção mais ampla de todo o método a partir da sua visão.

A codificação aberta (“*open coding*”) foi feita por dois pesquisadores utilizando o software ATLAS.ti³², um software para análise qualitativa amplamente difundido e utilizado para realizar esse tipo de avaliação. Um dos pesquisadores analisou os dados coletados, codificando cada evento descrito pelos participantes com um código genérico que representava o acontecimento em questão. O segundo pesquisador analisou os códigos e as sentenças relacionadas aos códigos desenvolvidos, ajudando a aprimorar o entendimento das sentenças, o código criado e as suas relações. Por fim, no final do processo de análise, os códigos foram analisados mais uma vez. Quando os códigos não refletiam completamente a sentença dos avaliadores, os códigos eram modificados para palavras e frases que melhor refletissem o acontecimento e o entendimento deles antes que uma checagem final das transcrições fosse feita, identificando qualquer situação a qual não estivesse ainda codificada. Depois da criação dos códigos e revisão dos mesmos, um processo de classificação dos códigos foi feito onde cada classe de códigos resumia uma situação, ou um tipo de problema, ou tipo de melhoria, que poderia ser feita no processo de avaliação ou no conjunto de heurísticas em si. Após a classificação dos códigos, foram analisadas quais as situações descrevem aquele evento da classe de códigos, sejam essas situações positivas ou negativas.

A seguir, os resultados são apresentados considerando os grupos de códigos criados durante a análise qualitativa feita nas transcrições de entrevistas e questionários. A numeração

32 <https://atlasti.com>

dos avaliadores foi feita de forma sequencial, a partir de uma lista de participantes ordenada de forma alfabética e então numeradas na sequência em que estavam. Quando a resposta avaliada, ou a sentença utilizada como exemplo, veio de uma resposta colocada nos questionários, a letra ‘Q’ entre parênteses foi adicionada após a numeração do participante. De forma análoga, quando a resposta foi dada durante as entrevistas, incluímos a letra ‘E’ entre parênteses ao lado de sua numeração de identificação. O *Codebook*, contendo todos os códigos e grupos criados juntamente com as associações feitas entre códigos e as sentenças utilizadas para cada código, está apresentado no APÊNDICE H.

8.3.2.1 Entendimento das heurísticas e Especificidade das heurísticas

A maioria dos participantes entendeu de forma geral as heurísticas do conjunto apresentado e os mesmos não acharam que houvesse a necessidade de mudanças ou um grande suporte para melhorar seu entendimento, como reportado e exemplificado por esta frase do Avaliador 28:

“Eu não tive nenhuma dificuldade durante o processo de avaliação, eu acho que todas as heurísticas estavam muito claras e bem descritas...” – Avaliador 28 (Q).

Um participante (novato – conhecimento baixo em jogos e em desenvolvimento de jogos) achou que as heurísticas relacionadas à avaliação de controles e a mecânica de controles eram ambíguas e poderiam guiar os participantes para o entendimento incorreto delas. Porém, ele não foi capaz de especificamente apontar em quais heurísticas teve dificuldades e o porquê, de acordo com a frase a seguir extraída de sua entrevista:

“...alguns defeitos que eu encontrei, eu classifiquei os mesmo em duas ou três heurísticas porque as mesmas eram muito similares uma a outra, por exemplo, um defeito que encontrei sobre os controles não serem fáceis de entender e dominar, eu classifiquei três heurísticas diferentes para este defeito. Também, talvez existisse alguma heurística que eu não entendi completamente, mas não me lembro qual seria a mesma, esta heurística eu precisei ler e reler e prestar mais atenção a ela para conseguir classificá-la.” – Avaliador 01 (E).

8.3.2.2 Viabilidade das Sentenças de Ajuda

Após analisar os códigos desenvolvidos para a seção de sentenças de ajuda e relacioná-las diretamente entre si através de um mapa mental, fomos capazes de identificar duas situações específicas em que as sentenças de ajuda obtiveram algum efeito. A primeira situação foi

quando os participantes sentiram a necessidade de melhor entender uma heurística específica quando ela não estava perfeitamente descrita. A segunda situação foi quando as sentenças de ajuda ativamente apoiaram os participantes em dúvida quanto à decisão de utilizar a heurística correta para avaliar um defeito identificado. As duas situações podem ser exemplificadas pelas frases extraídas de comentários a seguir:

“Quando eu cheguei em uma heurística para ser avaliada que possuía a sentença de ajuda ... ela me ajudou a entender que a heurística na verdade estava avaliando uma coisa diferente da que eu inicialmente havia compreendido, por isso eu acabei mudando minha resposta na revisão” – Avaliador 13 (E).

“...quando eu encontrava mais de uma heurística que poderia ser utilizada para descrever um defeito específico que eu havia identificado, eu lia a sentença de ajuda, se ela existisse nas heurísticas, para determinar qual heurística melhor se encaixaria naquele defeito.” – Avaliador 05 (Q).

Por fim, ainda houve o caso de quando a sentença de suporte não foi utilizada de forma alguma durante a avaliação do jogo, pois os participantes claramente compreenderam as heurísticas apenas com as palavras e frases utilizadas para as próprias heurísticas, como ilustrado pela seguinte frase extraída:

“Eu percebi a existência da seção de sentenças de suporte para algumas heurísticas, porém, as mesmas foram indiferentes para mim, eu fui capaz de entendê-las apenas pela própria heurística, sem olhar para sua frase de suporte.” – Avaliador 08 (E).

8.3.2.3 Tamanho do Conjunto de Heurísticas e Heurísticas Faltantes

Os participantes indicaram que sentiram falta de heurísticas ligadas à avaliação gráfica do jogo. Isso está diretamente ligado ao fato de o conjunto heurístico não possuir heurísticas que avaliam bugs ou ‘glitches’ (por exemplo, erros em um software computacional ou sistema que o fazem impróprios para uso de alguma forma).

“...eu senti falta de heurísticas que poderiam avaliar a seção visual gráfica do jogo e seus componentes.” – Avaliador 08 (E).

Além disso, participantes utilizaram heurísticas de forma incorreta, mesmo quando sabiam não ser o uso apropriado para ela. Isso acontecia quando uma heurística apropriada para o defeito não existia para ser utilizada, ou uma heurística a qual poderia ser corretamente utilizada para, de modo similar, identificar aproximadamente o defeito identificado.

“... eu não fui capaz de encontrar uma heurística para tanto, então eu usei uma heurística que não era totalmente correta para identificar o defeito, eu simplesmente encontrei uma forma de encaixá-la no defeito que eu havia identificado.” – Avaliador 08 (E).

Os participantes também identificaram que existiam heurísticas faltantes relacionadas com a consistência de informações, como, por exemplo, diferentes línguas utilizadas para descrever um pedaço de informação a qual não era clara o suficiente.

“... existiam seções do jogo que estavam em Inglês e outras partes que estavam em Chinês ao mesmo tempo, ..., eu cadastrei isso como um problema, porém não existia heurística que pudesse ser relacionada neste defeito.” – Avaliador 13 (E).

8.3.2.4 Categorização do Conjunto Heurístico

Em relação à categorização do conjunto heurístico, existem três situações e opiniões diferentes no assunto. A primeira é quando a categorização não ajuda de forma alguma os participantes da seção de avaliação.

“Eu honestamente não utilizei categorização do conjunto heurístico, eu notei sua existência, mas eu fui diretamente utilizando as heurísticas e avaliando o jogo em si só.” – Avaliador 28 (Q).

A segunda situação ocorre quando a categorização do conjunto ajudou os participantes a melhor compreender uma heurística do conjunto heurístico.

“Eu acho que dentro do conjunto heurístico, a categorização me ajudou a entender mais ou menos o que aquelas heurísticas estavam querendo dizer de acordo com a categoria onde elas estavam classificadas assim como me ajudou a tirar minhas dúvidas ao comparar a heurística com as outras heurísticas daquela categoria, para entender melhor o objetivo central da mesma.” – Avaliador 18 (E).

A terceira situação é quando a categorização ajudou os participantes a melhor entender qual seção do jogo especificamente estariam avaliando naquele momento com tais heurísticas (introdução, tutorial, assim por diante).

“... a classificação me deu uma ideia de qual parte do jogo eu estaria avaliando naquele momento, qual categoria de heurísticas e o que aquela heurística estaria realmente querendo dizer para aquela parte ...” – Avaliador 08 (E).

Alguns participantes também constataram que necessitavam utilizar a categorização das heurísticas inicialmente para não se perderem, mas que, com o passar do tempo e do

processo de avaliação, a categorização era menos necessária, sendo mais fácil de identificar e localizar as heurísticas desejadas.

8.3.3 Ameaças à validade

Todo estudo experimental apresenta algumas ameaças à sua validade que precisam ser identificadas para que possam ser tratadas durante todo o processo experimental. As ameaças associadas a esse trabalho são separadas em quatro categorias e apresentadas abaixo: **validade interna, validade externa, validade de conclusão e validade de construto** (Wohlin *et al.*, 2012).

Validade Interna: Com relação à medição do tempo, pode haver um erro de medição feito pelos inspetores durante a avaliação. Essa ameaça pode ter um impacto significativo em nossos resultados da primeira rodada, já que não tivemos nenhum controle sobre como os inspetores registraram o tempo ou se o tempo registrado estava correto. Na tentativa de controlar essa ameaça, a tabela fornecida tinha lugares distintos em que e os inspetores precisavam registrar seus tempos, e essa etapa específica foi mencionada durante a fase de treinamento. Com relação ao treinamento dado, pode haver efeitos colaterais durante a inspeção. Isso pode acontecer se o treinamento dado aos inspetores do Conjunto 1 fosse melhor ou pior do que o treinamento dado para os inspetores do Conjunto 2, ou entre a primeira e a segunda seção. Para evitar essa ameaça, o mesmo treinamento foi dado para os inspetores do Conjunto 1 e 2, bem como para as rodadas um e dois. Além disso, o treinamento dado usava apenas exemplos simples que não eram relacionados nem ao Conjunto 1 ou ao Conjunto 2 além de não estarem relacionados ao jogo avaliado. Por fim, em relação aos níveis de conhecimento, esse estudo foi realizado por meio do conhecimento autodeclarado. Os próprios inspetores preencheram o questionário sem qualquer confirmação independentemente de sua avaliação.

Validade Externa: Para fins de generalização, os jogos utilizados para a avaliação não representam todos os tipos de jogos existentes. Além disso, os participantes desse estudo não eram da indústria do jogo, nem eram especialistas em Jogabilidade. Portanto, os resultados são limitados até que seja realizada uma investigação mais aprofundada. No entanto, essa não é uma grande ameaça, já que o público-alvo do nosso conjunto são avaliadores não especialistas e jogadores em geral. Portanto, a amostra usada representa nossa população.

Validade de Conclusão: Essa ameaça seria válida se, por exemplo, o tamanho da amostra utilizado não fosse grande o suficiente para conter representatividade estatística. Porém, do ponto de vista estatístico, uma amostra acima de 45 participantes pode ser

considerada relevante. Dessa forma, como o estudo foi executado com uma população de 67 avaliadores, é possível considerar que a população é suficiente para que não seja uma ameaça impeditiva ou que seu peso possa ser considerado relevante nesse contexto.

Validade de construção: Para este tipo de ameaça, consideramos a definição de eficácia e eficiência usada para esse estudo. Tais indicadores são amplamente utilizados na literatura para esse tipo de estudo e adotamos a mesma definição que Korhonen, Paavilainen e Saarenpää (2009) e Valentim, Conte e Maldonado (2015).

8.4 Discussão

De acordo com os resultados apresentados nesse capítulo, e avaliando-os em relação aos resultados apresentados no CAPÍTULO 06, em que o primeiro estudo da CustomCheck4Play foi executado, é possível considerar que a CustomCheck4Play, como uma metodologia de avaliação da Playability customizada, apresenta resultados suficientemente satisfatórios para que ela seja considerada uma opção viável. O estudo apresentado nesse capítulo também suporta o entendimento de que a CustomCheck4Play é capaz de avaliar corretamente e de forma eficiente e eficaz, pelo menos, mais que um único tipo e gênero de jogo. Esse também foi um dos objetivos principais da metodologia desenvolvida a qual se pretendia ser customizável, de acordo com o tipo e gênero de jogo sendo avaliado. Porém, ainda existem alguns tipos e gêneros de jogos digitais que devem ser investigados, de forma a melhor compreender se a customização executada pela CustomCheck4Play, em conjunto com a PEACH, é o suficiente para avaliar todos os tipos propostos dentro da metodologia da CustomCheck4Play e da ferramenta PEACH.

Por fim, o estudo apresentado nesse capítulo também inicialmente avaliou o suporte da customização realizado pela ferramenta PEACH, uma vez que, o conjunto de heurísticas proveniente da customização da ferramenta PEACH foi o conjunto utilizado durante o estudo. Neste sentido, a ferramenta PEACH foi capaz de entregar um conjunto heurístico que corretamente pode ser encaixada e utilizada para avaliar corretamente os tipos e gêneros de jogos propostos até o momento. De acordo com a avaliação qualitativa realizada nesse estudo, alguns pontos de melhoria ainda podem ser realizados no conjunto heurístico, customização e apresentação do conjunto.

CAPÍTULO 09 – CONCLUSÕES E TRABALHOS FUTUROS

Este capítulo apresenta as conclusões dos achados evidenciados nesta dissertação e as discussões geradas a partir dos achados evidenciados. Destacam-se também as oportunidades de trabalhos futuros que podem ser desenvolvidos a partir dos resultados desta dissertação.

9.1 Conclusões

Este trabalho apresentou o processo de desenvolvimento de duas técnicas para a avaliação da Playability em jogos digitais com foco em avaliadores não especialistas e conjuntos heurísticos customizáveis, focados nos tipos e gêneros de jogos avaliados. Além disso, foi apresentada uma ferramenta de customização dos conjuntos heurísticos a qual realiza o suporte da técnica de avaliação da Playability gerando um conjunto de heurísticas que pode, corretamente, e de forma mais específica, avaliar os aspectos necessários da Playability dependendo do tipo e gênero de jogo sendo avaliado.

Inicialmente, foi realizado um mapeamento sistemático o qual permitiu identificar as principais técnicas e metodologias de avaliação da Playability para Jogos Digitais com as quais pode-se verificar, juntamente com a utilização de um “*Hands-On*”, os defeitos e lacunas existentes na avaliação da Playability de Jogos Digitais. Foi discutido no CAPÍTULO 01 e CAPÍTULO 02 as necessidades de desenvolvimento de técnicas de avaliação da Playability e as necessidades da indústria em relação às técnicas a serem propostas. Os resultados provenientes dessa fase, permitiram a identificação da lacuna, da necessidade de uma técnica de avaliação que levasse em consideração os avaliadores não especialistas, para diminuição de custos e tempo para sua realização. Além disso, foi através dessa fase que foi possível identificar a lacuna de necessidade de desenvolvimento de um conjunto que levasse em consideração também os diferentes aspectos de cada tipo e gênero de jogos digitais.

Portanto, como a primeira versão da técnica de avaliação da Playability em jogos digitais, foi proposto nesse trabalho a NExPlay, uma técnica que visa diminuir custos e tempo de avaliação através do uso de avaliadores não especialistas. A NExPlay é uma técnica subdividida em três categorias para melhor classificação das heurísticas que a compõem (veja CAPÍTULO 03). Por fim, após a proposta e desenvolvimento da NExPlay, um estudo experimental comparativo foi desenvolvido para testar a eficácia e eficiência da técnica em

relação a outro conjunto da literatura. Nessa avaliação, a NExPlay não provou ter melhor desempenho, porém, mostrou ser capaz de avaliar jogos digitais de gênero e escopo aberto.

Após a proposta da NExPlay e a identificação da necessidade de foco em escopos de avaliação menores a partir do tipo e gênero de jogo sendo avaliado, foi desenvolvida a CustomCheck4Play. A CustomCheck4Play é uma técnica de avaliação voltada para a customização das categorias de heurísticas de modo a se adaptar aos diferentes tipos e gêneros de jogos existentes. Essa customização é feita através de um questionário o qual simultaneamente modifica as heurísticas apresentadas para melhor atender às necessidades especificadas. A técnica também teve como objetivo corrigir os defeitos encontrados no estudo realizado com a NExPlay, de forma a melhorara sua interação. Após sua definição, um estudo experimental foi realizado com ela, em busca de avaliar sua eficiência e eficácia em relação a outro conjunto da literatura. Os resultados demonstraram que a CustomCheck4Play obteve melhores resultados de eficiência e eficácia no jogo selecionado, além de ter sido capaz de tratar as limitações identificadas com a avaliação do conjunto NExPlay.

Após a avaliação inicial da CustomCheck4Play, seu processo de customização foi utilizado para que fosse possível desenvolver uma ferramenta que apoiaria e facilitaria a customização dos conjuntos heurísticos através de um website que guiaria os usuários no processo de customização. Dentro dessa ferramenta, os usuários poderiam mais facilmente compreender o processo de avaliação, customização, identificação dos defeitos, reporte deles e avaliação dos defeitos em relação às heurísticas entregues a eles. Desta forma, a ferramenta PEACH foi desenvolvida e apresentada no CAPÍTULO 07.

Por fim, um segundo estudo comparativo experimental foi realizado. Nele, unimos a customização da ferramenta PEACH com o processo de avaliação proposto pela CustomCheck4Play. Foram utilizados tipos e gêneros diferentes dos estudos realizados anteriormente, assim como um conjunto comparativo diferente. Assim, houve um entendimento mais abrangente e com uma amostra maior de avaliadores no processo de avaliação, tanto para a metodologia da CustomCheck4Play quanto para a ferramenta PEACH.

9.2 Contribuições

As principais contribuições desta pesquisa são destacadas a seguir:

- **Uma visão geral sobre conjuntos heurísticos existentes na literatura e seus defeitos e lacunas:** Através dos resultados obtidos com o mapeamento sistemático desenvolvido,

fomos capazes de avaliar os defeitos e lacunas até então existentes na literatura. Além disso, o mapeamento fornece uma breve descrição de outros trabalhos relacionados com a avaliação da Playability e suas pesquisas e objetivos centrais. De acordo com os objetivos propostos por esta pesquisa, a base de conhecimento é formada por essa contribuição;

- **Desenvolvimento de uma técnica inicial para a avaliação da Playability de Jogos Digitais através de avaliadores não especialistas:** Foi desenvolvida uma técnica inicial para a avaliação da Playability em jogos digitais de forma a demonstrar o potencial de uso de avaliadores não especialistas para a condução dessas avaliações. Mesmo a técnica não obtendo valores significativamente relevantes, ela se mostrou viável para a avaliação de jogos. Os resultados dessa etapa foram publicados no artigo: **“NExPlay – Playability Assessment for Non-Expert Evaluators”** aceito e apresentado na *International Conference on Enterprise and Information Systems (ICEIS)* no ano de 2018 – Prémio de melhor artigo da trilha de Interação Humano-Computador;
- **Desenvolvimento de uma técnica de avaliação da Playability em jogos digitais configurável de acordo com tipo e gênero de jogo:** Foi proposto neste trabalho uma segunda técnica para a avaliação da Playability de Jogos Digitais de modo configurável em que seus resultados demonstraram que ela foi estatisticamente melhor em termos de eficiência e eficácia que outro conjunto heurístico da literatura. Os resultados desse estudo foram publicados no artigo: **“Straight to the Point - Evaluating What Matters for You: A Comparative Study on Playability Heuristic Sets”** aceito e apresentado na *International Conference on Enterprise and Informations Systems (ICEIS)* no ano de 2020 – Prémio de melhor artigo de estudante da trilha de Interação Humano-Computador;
- **Desenvolvimento de uma ferramenta inicial para o apoio no processo de customização do conjunto heurístico utilizado pela técnica CustomCheck4Play:** Foi proposto nesse trabalho uma ferramenta de apoio à customização do conjunto heurístico de modo a apoiar o uso da técnica CustomCheck4Play na avaliação de jogos digitais de tipos e gêneros diferentes. A ferramenta proposta, juntamente com o seu estudo de viabilidade e o estudo comparativo de sua utilização a em conjunto com a CustomCheck4Play será publicada posteriormente em um Journal paper já produzido, mas ainda não encaminhado a uma revista.

9.3 Trabalhos Futuros

A partir da análise dos resultados e conclusões desenvolvidas para o trabalho apresentado aqui, são formulados alguns trabalhos futuros os quais podem ainda ser realizados para aprimorar o uso e resultados entregues pelas técnicas e metodologias desenvolvidas. A seguir, alguns trabalhos futuros são apresentados, divididos de acordo com a técnica/metodologia ou área na qual o trabalho pode ser feito:

- **CustomCheck4Play:** ainda existem melhorias que poderiam ser realizadas na técnica e conjunto de heurísticas da CustomCheck4Play:
 - Primeiramente, a inclusão de um método e heurísticas com as quais o usuário poderia identificar e relacionar bugs gráficos e glitches visuais que podem acontecer durante o uso do jogo e sua avaliação;
 - Uma segunda melhoria seria na seção das sentenças de ajuda, em que a sentença não é apresentada a todo momento para o usuário, mas em uma *tooltip* que pode ser clicada para abrir ou não a sentença quando necessário, diminuindo a quantidade de informação apresentada ao usuário;
 - Além disso, haveria inclusão da sentença de ajuda, em todas as heurísticas, mesmo que ela já seja completamente compreendida pelo usuário, assim melhorando o entendimento dos usuários quanto ao local onde a mesma pode ser aplicada ou como ela pode ser entendida de forma mais fácil;
- **PEACH:** para a ferramenta PEACH, existem diversas melhorias que ainda podem ser executadas:
 - A inclusão, na própria ferramenta, de um método com o qual os usuários possam registrar defeitos identificados, de acordo com o jogo avaliado, e, a partir da customização do conjunto, a ferramenta já direcionaria o usuário a uma tela que lhe permitisse identificar os defeitos de acordo com as heurísticas apresentadas para o jogo em questão;
 - A inclusão de um processo de sugestão de novas heurísticas e mudanças que poderiam ser realizadas em heurísticas existentes no conjunto heurístico dentro da própria ferramenta que identificaria qual o conjunto customizado foi utilizado pelo usuário, a seção do jogo na qual ele necessitaria desta nova heurística e por quê;

- De acordo com a primeira sugestão, também é possível incluir uma seção onde os usuários pudessem incluir um anexo (uma foto ou vídeo) do defeito identificado no jogo, de modo a apoiar os desenvolvedores e gerentes do processo de avaliação do jogo a desenvolver as melhorias e identificar os locais com problemas;
- Por fim, pode ser incluída uma maneira para que os desenvolvedores e gerentes de avaliação possam classificar e identificar os defeitos mais importantes e relevantes para os desenvolvedores e para o jogo em si.
- **Avaliações e oportunidades de pesquisa:** Existem oportunidades que vão além da melhoria e implementação de características nas técnicas desenvolvidas, sendo elas:
 - Seria interessante que um estudo de grande escala com a CustomCheck4Play em comparação a uma segunda técnica seja efetuado. Esse estudo seria conduzido de modo a focar na avaliação do conhecimento dos participantes desse estudo em relação à eficiência e à eficácia desses avaliadores, no fim, sendo possível identificar se avaliadores com um menor nível de conhecimento são capazes de avaliar a Playability dos jogos de forma semelhante ou igual a avaliadores experientes e que já aplicaram anteriormente esse tipo de avaliação;
 - Uma maior investigação dos fatores de impacto de cada um dos gêneros e tipos de jogos poderia ser realizada através de uma série de estudos com cada um deles, de modo a identificar as possíveis necessidades de alterações que a CustomCheck4Play ainda necessita. Além do mais, ainda é necessário atacar uma das limitações deste trabalho relacionada ao tamanho dos jogos avaliados, uma vez que os jogos avaliados nos estudos têm contextos pequenos e curtos. Desta forma, não sabemos se o tamanho do jogo sendo avaliado pela técnica pode influenciar no seu desempenho, sendo possível investigar este ponto juntamente com os gêneros e tipos de jogos deste ponto.

REFERÊNCIAS BIBLIOGRÁFICAS

- A. Kitchenham, B. (2007) *Kitchenham, B.: Guidelines for performing Systematic Literature Reviews in software engineering. EBSE Technical Report EBSE-2007-01.*
- Aker, Ç., Rizvanoglu, K., İnal, Y. (2020) *Revisiting Heuristics for Evaluating Player Experience in Different Gaming Platforms: A Multi-Modal Approach.* In: *Game User Experience And Player-Centered Design.* Springer, Cham , p. 123-161.
- Aleem, S., Capretz, L. F. and Ahmed, F. (2016) “Game development software engineering process life cycle: a systematic review,” *Journal of Software Engineering Research and Development*, 4(1), p. 6. doi: 10.1186/s40411-016-0032-7.
- Arruda, E. P. (2014) *Fundamentos para o Desenvolvimento de Jogos Digitais: Série Tekne.* Bookman Editora.
- Arsenault, D. (2009) "Video game genre, evolution and innovation". *Eludamos. Journal for Computer Game Culture* 3.2 (2009): 149-176.
- Barbosa, S. and Silva, B. (2010) *Interação humano-computador.* Elsevier Brasil.
- Barcelos, T. S. *et al.* (2011) “Análise Comparativa De Heurísticas Para Avaliação De Jogos Digitais,” in *Proceedings of the 10th Brazilian Symposium on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction.* Porto Alegre, Brazil, Brazil: Brazilian Computer Society (IHC+CLIH '11), pp. 187–196. Available at: <http://dl.acm.org/citation.cfm?id=2254436.2254469>.
- Basili, V. R. and Rombach, H. D. (1988) “The TAME project: towards improvement-oriented software environments,” *IEEE Transactions on Software Engineering*, 14(6), pp. 758–773. doi: 10.1109/32.6156.
- Bevan, N. (2001) “International Standards for HCI and Usability,” *International Journal of Human-Computer Studies.* Duluth, MN, USA: Academic Press, Inc., 55(4), pp. 533–552. doi: 10.1006/ijhc.2001.0483.
- Bjork, S. and Holopainen, J. (2004) *Patterns in Game Design (Game Development Series).* 1st ed. Charles River Media. Available at: <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/1584503548>.
- Caillois, R. and Barash, M. (2001) *Man, Play, and Games.* University of Illinois Press (Sociology/Sport). Available at: <https://books.google.com.br/books?id=bDjOPsjzfC4C>.
- Chang, A., Montalvo, L., Paz, F. (2019) *Usability and Playability Heuristics for Augmented Reality Video Games in Smartphones.* In: *International Conference on Human-Computer*

Interaction. Springer, Cham, p. 19-29.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences*. 2nd ed – Chapter 3. Hillsdale, NJ: Lawrence Earlbaum Associates.

Crawford, C. (1984) *The Art of Computer Game Design: Reflections of a Master Game Designer*. McGraw-Hill Osborne Media. Available at: <https://books.google.com.br/books?id=H0k1NgAACAAJ>.

Desurvire, H., Caplan, M. and Toth, J. A. (2004) “Using Heuristics to Evaluate the Playability of Games,” in *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI EA '04), pp. 1509–1512. doi: 10.1145/985921.986102.

Desurvire, H. and Wiberg, C. (2008) “Evaluating user experience and other lies in evaluating games,” in *CHI'08 : Workshop on Evaluating User Experiences in Games*.

Desurvire, H. and Wiberg, C. (2009) “Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration,” in Ozok, A. A. and Zaphiris, P. (eds.) *Online Communities and Social Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 557–566.

Duarte, C. and Veloso, A. (2016) “Análise heurística de jogos digitais multijogador em rede: O caso do ‘Curve Fever 2,’” in *Atas da Videojogos 2016 - 9ª Conferência de Ciências e Artes dos Videojogos*. Lisboa - Portugal, pp. 71–77.

Entertainment Software Engineering (ESA). 2009. Essential Facts about the Computer and Video Game Industry of 2009. ESA. Available at: <https://library.princeton.edu/sites/default/files/2009.pdf> [Accessed: 25 July 2019].

Entertainment Software Engineering (ESA). 2019. Essential Facts about the Computer and Video Game Industry of 2019. ESA. Available at: <https://www.theesa.com/wp-content/uploads/2019/05/2019-Essential-Facts-About-the-Computer-and-Video-Game-Industry.pdf> [Accessed: 25 July 2019].

Falbo, R. A. (1998) *Integração de Conhecimento em um Ambiente de Desenvolvimento de Software*. Tese de defesa, pp. 124 - 131. COPPE/UFRJ - Rio de Janeiro, RJ.

Federoff, M. A. (2002). *Heuristics and usability guidelines for the creation and evaluation of fun in video games* (Doctoral dissertation, Indiana University).

Kapp, K. M. (2012) *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education*. 1^o Edition. Wiley (Pfeiffer essential resources for training and HR professionals). Available at: <https://books.google.com.br/books?id=M2Rb9ZtFxccC>.

Korhonen, H. and Koivisto, E. M. I. (2006) “Playability Heuristics for Mobile Games,” in *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices*

- and Services*. New York, NY, USA: ACM (MobileHCI '06), pp. 9–16. doi: 10.1145/1152215.1152218.
- Korhonen, H. and Koivisto, E. M. I. (2007) “Playability Heuristics for Mobile Multi-player Games,” in *Proceedings of the 2Nd International Conference on Digital Interactive Media in Entertainment and Arts*. New York, NY, USA: ACM (DIMEA '07), pp. 28–35. doi: 10.1145/1306813.1306828.
- Korhonen, H., Paavilainen, J. and Saarenpää, H. (2009) “Expert Review Method in Game Evaluations: Comparison of Two Playability Heuristic Sets,” in *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*. New York, NY, USA: ACM (MindTrek '09), pp. 74–81. doi: 10.1145/1621841.1621856.
- Manzoni, F. S., Ferreira, B. M. and Conte, T. U. (2018) “NExPlay - Playability Assessment for Non-experts Evaluators,” in *Proceedings of the 20th International Conference on Enterprise Information Systems - Volume 2: ICEIS*. SciTePress, pp. 451–462. doi: 10.5220/0006695604510462.
- Manzoni, F. S., Conte, T. U., Silveira, M. S., Barbosa, S. D. J. (2020) " Straight to the Point - Evaluating What Matters for You: A Comparative Study on Playability Heuristic Sets" in *Proceedings of the 22th International Conference on Enterprise and Information Systems - Volume 2: ICEIS*, SciTePress, pp. 499-510. doi: 10.5220/0009381304990510.
- Melo, W., Shull, F., & Travassos, G. H. (2001). "Software review guidelines". COPPE/UFRJ Systems Engineering and Computer Science Program Technical Report ES-556/01.
- Nacke, L. *et al.* (2009) “Playability and player experience research,” in *Proceedings of DiGRA 2009: Breaking new ground: Innovation in games, play, practice and theory*.
- Nielsen, J. (1994a) “Enhancing the Explanatory Power of Usability Heuristics,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '94), pp. 152–158. doi: 10.1145/191666.191729.
- Nielsen, J. (1994b) “Usability Inspection Methods,” in *Conference Companion on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '94), pp. 413–414. doi: 10.1145/259963.260531.
- Nielsen, J. and Molich, R. (1990) “Heuristic Evaluation of User Interfaces,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '90), pp. 249–256. doi: 10.1145/97243.97281.
- ISO - International Organization for Standardization. (2011 - Revisada em 2017) "Norm ISO/IEC 25010:2011: Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models." ISO Organization for Standardization.

- Orlando, G. (2007) *Console Portraits: A 40-Year Pictorial History of Gaming*, Wired News.
- Pinelle, D. *et al.* (2009) “Usability Heuristics for Networked Multiplayer Games,” in *Proceedings of the ACM 2009 International Conference on Supporting Group Work*. New York, NY, USA: ACM (GROUP '09), pp. 169–178. doi: 10.1145/1531674.1531700.
- Pinelle, D., Wong, N. and Stach, T. (2008a) “Heuristic Evaluation for Games: Usability Principles for Video Game Design,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '08), pp. 1453–1462. doi: 10.1145/1357054.1357282.
- Pinelle, D., Wong, N. and Stach, T. (2008b) “Using Genres to Customize Usability Evaluations of Video Games,” in *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*. New York, NY, USA: ACM (Future Play '08), pp. 129–136. doi: 10.1145/1496984.1497006.
- Politowski, C. *et al.* (2016) “Software Engineering Processes in Game Development: a Survey about Brazilian Developers’ Experiences,” in *SBGames 2016*.
- Prensky, M. (2007) *Digital Game-Based Learning*. 2^o Edition. Paragon House. Available at: <https://books.google.com.br/books?id=e-CjQAAACAAJ>.
- Pressman, R. S. (2016) *Software engineering: a practitioner’s approach*. 7th Editio. New York, NY, USA: McGraw-Hill, Inc.
- Rajanen, M. and Tapani, J. (2018) “A Survey of Game Usability Practices in North American Game Companies,” in *The 27th International Conference on Information Systems Development (ISD2018)*.
- Rogers, S. (2014) *Level Up! The guide to great video game design*. John Wiley & Sons.
- Sánchez, J. L. G. *et al.* (2012) “Playability: analysing user experience in video games,” *Behaviour & Information Technology*. Taylor & Francis, 31(10), pp. 1033–1054. doi: 10.1080/0144929X.2012.710648.
- de Souza Filho, J. C., Monteiro, I. T., Jucá, P. M. (2019) *Game for PLAYability Heuristic Evaluation (G4H-PLAY): Adapting G4H Gamification to Allow the Use of Playability Heuristics*. In: *International Conference on Human-Computer Interaction*. Springer, Cham, p. 68-84.
- Valentim, N., Conte, T. and Maldonado, J. (2015) “Evaluating an Inspection Technique for Use Case Specifications: Quantitative and Qualitative Analysis,” *ICEIS 2015 - 17th International Conference on Enterprise Information Systems, Proceedings*. doi: 10.5220/0005374100130024.
- Wohlin, C. *et al.* (2012) *Experimentation in software engineering*. Springer Science & Business Media.

Wohlin, C. (2014) *Guidelines for snowballing in systematic literature studies and a replication in software engineering*. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. London, UK. ACM. p. 1-10. DOI: <http://dx.doi.org/10.1145/2601248.2601268>.

APÊNDICE A – QUESTIONÁRIO DE CARACTERIZAÇÃO

O formulário de caracterização dos participantes foi criado através da ferramenta online Google Forms e é apresentada a seguir apenas suas perguntas utilizadas e repassadas aos avaliadores.

- 1- E-mail válido para contato com o avaliador.
- 2- Com que frequência você joga jogos digitais? (Por exemplo: videogames, celular, computador, web e outros)
 - a. Não joga nenhum tipo de jogos digitais.
 - b. Joga casualmente quando tenho tempo.
 - c. Joga pelo menos uma vez por semana.
 - d. Joga pelo menos 3 vezes por semana.
 - e. Joga pelo menos 5 vezes por semana.
 - f. Joga todos os dias na semana.
- 3- Qual gênero de jogos digitais você geralmente joga?
- 4- Qual sua experiência com Inspeções de Software?
 - a. Nenhuma experiência – Não possui conhecimento em IHC/IES ou que possui algumas noções sobre inspeção de softwares adquiridas através de leituras/palestras, mas sem experiência prática.
 - b. Baixa – Participou de pelo menos um projeto como Inspetor de Software em sala de aula.
 - c. Média – Participou de 1 a 4 projetos como Inspetor de Software na indústria.
 - d. Alta – Participou de 5 projetos como Inspetor de Software na indústria ou mais.

APÊNDICE B – TABELA DE IDENTIFICAÇÃO E CLASSIFICAÇÃO DOS DEFEITOS

Avaliador: _____

Técnica utilizada: _____

Hora de Início Jogo: _____ Hora final jogo: _____

Hora de Início de avaliação: _____ Hora final de avaliação: _____

Numeração	Problema encontrado (Descrição e localização no Jogo)	Nº heurística

APÊNDICE C – QUESTIONÁRIO DE AVALIAÇÃO E OPINIÃO PÓS-USO DOS CONJUNTOS HEURÍSTICOS

O questionário de opinião foi criado através da ferramenta online Google Forms e é apresentada a seguir apenas suas perguntas utilizadas e repassadas aos avaliadores.

- 1- Endereço de e-mail para contato:
 - 2- Nome completo:
 - 3- Solução utilizada:
 - a. NExPlay
 - b. Barcelos et al.
 - 4- Aponte como você utilizou a técnica de inspeção (identificação de problemas utilizando as heurísticas):
 - a. Primeiramente utilizei o jogo anotando os possíveis problemas identificados por mim e depois disto associei estes problemas encontrados com as heurísticas específicas a cada problema.
 - b. Utilizei o jogo de forma que tentei encontrar as heurísticas as características dentro do jogo. Ou seja, utilizei as heurísticas como guia.
 - c. Outras formas:
-
- 5- Aponte resumidamente os pontos positivos e negativos que você encontrou em relação a técnica utilizada:
 - 6- Sobre o conjunto de heurísticas utilizado, você encontrou dificuldades no entendimento das heurísticas? Você acha que faltou alguma heurística neste conjunto? Explique:
 - 7- Você aplicaria as heurísticas utilizadas novamente? Em qual(is) contexto(s) aplicaria? Justifique:

APÊNDICE D – QUESTIONÁRIO DE MODULARIZAÇÃO DA TÉCNICA CUSTOMCHECK4PLAY

CustomCheck4Play - Questionário de Modularização da Técnica

Informações iniciais

Este é o questionário para modularização da técnica antes de utilizá-la para a avaliação da jogabilidade de jogos digitais. Nas próximas seções iremos fazer algumas perguntas sobre o jogo que você deseja avaliar e quais as informações que você deseja conhecer depois de avaliá-lo.

Qual é o Tipo/Gênero PRINCIPAL do jogo a ser avaliado? *

Mesmo que o jogo seja classificado com mais de um tipo/gênero, marque apenas aquele que é considerado o principal, ou seja, a primeira classificação existente para o mesmo.

Por favor, escolha ...

O jogo está em fase de desenvolvimento? *

Se o jogo ainda não foi lançado como um jogo totalmente aberto ao público (em fases alpha, beta, open beta) a resposta deve ser marcada como 'sim'.

sim

não

Características do Gameplay

Das opções abaixo, quais delas se fazem presentes no jogo a ser avaliado? *

Abaixo você irá encontrar uma sequência de características comuns aos jogos. Para cada uma das opções, marque aquelas que seu jogo apresenta durante sua extensão.

História progressiva

Customização de personagens

Ajudas de Contexto

Heurísticas a serem utilizadas

Heurísticas da categoria INTRODUÇÃO

Introdução		
Nº	Heurística	Ajuda
H1	O jogo apresenta uma história inicial que justifique as ações do jogador apresentando um objetivo principal.	
H2	O jogo deve apresentar um tutorial para familiarização do jogador com as mecânicas e gameplay.	
H3	As terminologias utilizadas para todos os objetos/funções são entendidas pelo jogador.	

Categoria Apresentação de Personagem

Apresentação de personagens		
Nº	Heurística	Ajuda
H4	É possível customizar o personagem principal de modo a criarmos uma identidade única a partir do mesmo.	
H5	O personagem principal consegue ter uma personalidade e características que prendem nossa atenção acreditando que o mesmo é real.	
H6	Os personagens do jogo criam uma relação pessoal com o jogador caracterizando-os como indivíduos reais.	

Categoria Introdução ao Gameplay

Introdução ao Gameplay		
Nº	Heurística	Ajuda
H7	Todas as representações visuais do jogo devem ser de fácil compreensão pelo jogador.	
H8	O jogador consegue obter facilmente informações sobre o estado atual do jogo, pontuações, vida e tudo mais a sua volta de maneira simples.	
H9	Os controles devem ser simples para que o aprendizado seja rápido.	

Categoria Desenvolvimento da Gameplay

Desenvolvimento da Gameplay		
Nº	Heurística	Ajuda
H10	O jogo apresenta feedback ao jogador de forma apropriada.	Ou seja, o jogo faz o uso correto de músicas/efeitos sonoros/vibrações de controle/avisos visuais quando necessário ao jogador assim como possui um padrão sendo seguido para estas ações.
H11	Atividades desenvolvidas durante o jogo são variadas de modo a diminuir fadiga.	Não existem tarefas repetitivas.
H12	As habilidades necessárias para se alcançar um objetivo atual ou futuro são conhecidas e ensinadas.	
H13	O jogo é balanceado de modo a aplicar pressão no jogador sem frustrá-lo.	Ou seja, desafios são experiências positivas ao jogador e o mantém interessado.

Categoria Progresso da História

Progresso da História		
Nº	Heurística	Ajuda
H14	É possível conquistar um objetivo principal de diversas maneiras.	O jogo apresenta diversos objetivos secundários e opcionais os quais complementam um (ou mais) objetivo(s) principal(ais).
H15	É possível entender a cronologia da história de forma quase natural e as ações do jogador interferem na mesma.	
H16	O jogo apresenta uma história imersiva que prenda a atenção do jogador.	

Categoria Evolução da Gameplay

Evolução da Gameplay		
Nº	Heurística	Ajuda
H17	Controles devem ser claros e fisicamente confortáveis, além de suas ações de resposta serem imediatas.	
H18	O jogador não deve perder qualquer objeto ganho através de grande esforço.	
H19	O jogador percebe alterações no mundo do jogo através de suas ações e progresso	
H20	Os inimigos representam desafios e surpresas inesperadas ao jogador independente do nível selecionado.	
H21	O jogador consegue aprender como não errar em atividades semelhantes.	O jogo não penaliza de forma imediata nos primeiros erros dando oportunidade para o jogador aprender com estes erros.
H22	O jogador consegue se divertir com o jogo.	
H23	Os gráficos do jogo são satisfatórios/atrativos.	Mesmo que feito de forma simples os gráficos devem ser apresentados de forma atrativa para os jogadores.
H24	O Design do jogo é atrativo o suficiente para manter a atenção do jogador.	O design do jogo não causa fadiga no jogador.

Categoria Pausas de Jogo

Pausas de Jogo		
Nº	Heurística	Ajuda
H25	O jogador deve ser capaz de interromper facilmente a partida a qualquer momento.	Ou seja, interrupções são esperadas e planejadas.
H26	É possível salvar o estado atual do jogo para retomar o mesmo posteriormente do mesmo ponto salvo.	
H27	Todas as vezes que o jogador iniciar/recomeçar/continuar o jogo, o mesmo possui informações suficientes para começar a jogar.	Não existe a necessidade de consultar manuais, documentações ou tutoriais constantemente para continuar jogando.

Categoria Ajudas de Contexto e Recuperação de Erros

Ajudas de Contexto e Recuperação de erros		
Nº	Heurística	Ajuda
H28	Jogadores devem receber ajudas de contexto durante as fases para que o jogador não se encontre preso na mesma.	
H29	As ajudas de contexto não devem facilitar demais as fases de modo que não hajam desafios ao jogador.	
H30	O jogo deve ser capaz de evitar erros não intencionais dos jogadores.	Por exemplo, não permitir que o jogador perca seu slot de salvamento do jogo quando acidentalmente o mesmo apertar um botão errado.

Categoria Dificuldades e Níveis Progressivos

Dificuldade e Níveis Progressivos		
Nº	Heurística	Ajuda
H31	O desafio do jogo poder ser ajustado de acordo com as habilidades dos jogadores.	
H32	A dificuldade do jogo é gradual e acompanha o desenvolvimento do jogador.	O jogo segue a dificuldade escolhida pelo jogador, porém, esta se ajusta ao longo das fases, quanto mais fases o mesmo consegue passar, mais difícil e desafiador o jogo se torna.

Categoria Configurabilidade e Menus

Configurabilidade e Menus		
Nº	Heurística	Ajuda
H33	Layouts e menus devem ser intuitivos e organizados de forma que o jogador possa manter o foco no jogo.	
H34	O jogo deve permitir customização das configurações e controles.	
H35	Os controles devem ser expansíveis para jogadores mais habilidosos.	

APÊNDICE E – QUESTIONÁRIO DE CARACTERIZAÇÃO DO ESTUDO REALIZADO PARA A CUSTOMCHECK4PLAY

O formulário de caracterização dos participantes foi criado através da ferramenta online Google Forms e é apresentada a seguir apenas suas perguntas utilizadas e repassadas aos avaliadores.

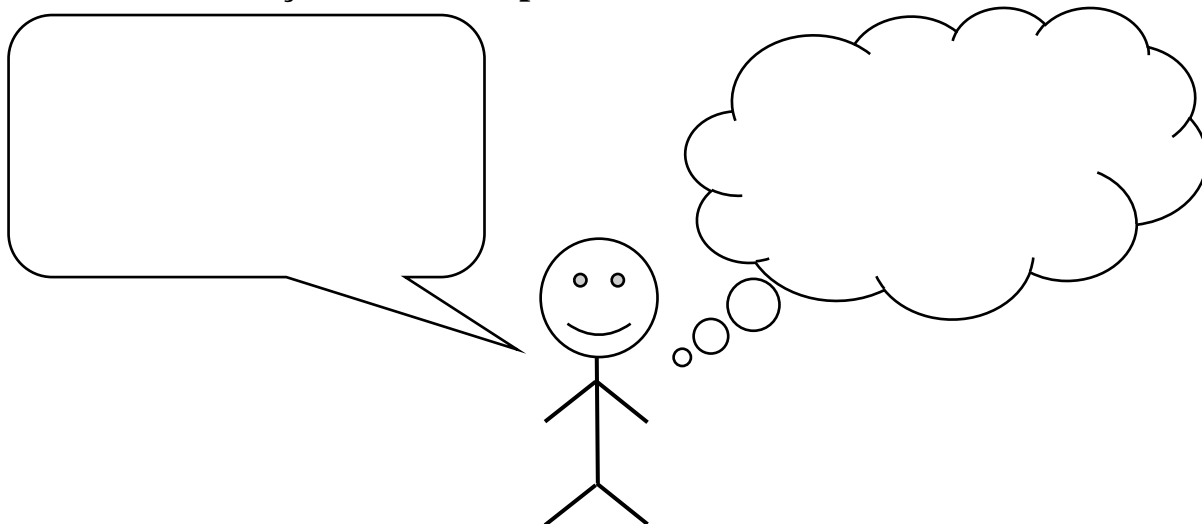
- 1- E-mail válido para contato com o avaliador.
- 2- Nome completo.
- 3- Em relação a sua experiência com jogos, você se considera um jogador:
 - a. Heavy Player (Jogadores que conhecem muito sobre diversos tipos de jogos e dedicam um tempo significativo para um jogo, principalmente quando existem novas atualizações)
 - b. Casual Player (Jogadores que gostam de alguns jogos específicos, porém os utilizam como meios de distração e para ocasionalidades, sem se importar com a total completude do jogo ou novas atualizações e missões complementares)
 - c. Não gosto de jogos
- 4- Você tem ou já teve alguma experiência com o desenvolvimento de jogos, seja para qualquer tipo de dispositivo ou ambiente?
 - d. Sim, já trabalhei em projetos de desenvolvimento de jogos na indústria de modo geral.
 - e. Sim, trabalhei no desenvolvimento de jogos no âmbito acadêmico (seja durante cursos acadêmicos ou complementares).
 - f. Já trabalhei no desenvolvimento de jogos como Hobby e por curiosidade de aprendizagem fora de qualquer âmbito acadêmico ou industrial.
 - g. Não, nunca trabalhei com o desenvolvimento de jogos.

APÊNDICE F – QUESTIONÁRIO DE FEEDBACK DO ESTUDO REALIZADO PARA A PEACH

Endereço de e-mail:

Nome:

Modelo de avaliação da User Experience – 3E



1. Sobre o conjunto de heurísticas utilizado, você encontrou dificuldades no entendimento das heurísticas? Se sim, em quais?

2. Você acha que faltou alguma heurística neste conjunto que poderia ter sido utilizada para avaliar algum dos jogos? Explique.

3. Aponte resumidamente os pontos positivos e negativos que você encontrou em relação a técnica e ferramenta utilizada.

4. Se você possuir algum tipo de sugestão para a ferramenta, o seu processo, para as heurísticas ou algum destes itens, por favor, deixe as mesmas a seguir e explique as mesmas.

APÊNDICE G – QUESTIONÁRIO DE FEEDBACK DO ESTUDO REALIZADO PARA A CUSTOMCHECK4PLAY

O formulário de caracterização dos participantes foi criado através da ferramenta online Google Forms e é apresentada a seguir apenas suas perguntas utilizadas e repassadas aos avaliadores.

- 1- E-mail válido para contato com o avaliador:
- 2- Nome completo:
- 3- Solução utilizada:
 - a. CustomCheck4Play
 - b. Barcelos *et al.*
- 4- Sobre o conjunto de heurísticas utilizado, você encontrou dificuldades no entendimento das heurísticas? Se sim, em quais?
- 5- Você acha que faltou alguma heurística neste conjunto que poderia ter sido utilizada durante sua avaliação? Explique:
- 6- Aponte resumidamente os pontos positivos e negativos que você encontrou em relação a técnica utilizada:
- 7- Se existisse a oportunidade, você aplicaria as heurísticas utilizadas novamente? Se sim, em que contexto? Para qualquer jogo? Apenas para grandes projetos?

APÊNDICE H – CODEBOOK DESENVOLVIDO PARA A ANÁLISE QUALITATIVA DO ESTUDO COMPARATIVO DA CUSTOMCHECK4PLAY

◇ Uso da ajuda.

8 Codes:

○ [CAT] Uso da ajuda.

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Linked Codes:

- ← is associated with – ○ A seção de ajuda melhora o entendimento de algumas heurísticas.
- ← is associated with – ○ Seção de ajuda apoiou na decisão da heurística correta para o defeito.
- ← is associated with – ○ Seção de ajuda não foi utilizada.

Groups:

- ◇ Uso da ajuda.

○ A seção de ajuda melhora o entendimento de algumas heurísticas.

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

- 📄 9 AV_13
- 📄 11 AV_28

Quotations:

- 🗣️ 9:22 talvez essa ajuda até atrapalharia pelo fato de mudar nosso entendimento.
- 🗣️ 9:23 me ajudou a perceber que a heurística era outra coisa do que eu estava pensando e mudei minha respos...
- 🗣️ 11:8 A seção de ajuda das heurísticas ajudou ou facilitou de alguma forma no entendimento das heurísticas...

Linked Codes:

- is associated with → ○ [CAT] Uso da ajuda.
- ← is part of – ○ A seção de ajuda pode atrapalhar nas heurísticas que estão claras.
- ← is associated with – ○ Certas heurísticas não precisam de ajuda.

Groups:

- ◇ Uso da ajuda.

○ A seção de ajuda é percebida no conjunto.

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

- 📄 1 AV_01
- 📄 3 AV_05
- 📄 4 AV_08
- 📄 9 AV_13
- 📄 10 AV_17

Quotations:

- 🗣️ 1:9 Eu percebi a existência dessa seção
- 🗣️ 3:8 percebi que existia uma seção de ajuda para algumas heurísticas
- 🗣️ 4:9 Sim, percebi que existia uma seção de ajuda para algumas heurísticas
- 🗣️ 9:16 Quando chegava em uma heurística que possuía isso, eu lembro de um ponto que isso aqui me ajudou a p...
- 🗣️ 10:8 Porque as heurísticas em si já eram o suficiente para mim. A heurística em si já é bem autoexplicati...

Groups:

◊ Uso da ajuda.

○ **A seção de ajuda pode atrapalhar nas heurísticas que estão claras.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 9 AV_13

Quotations:

🗨️ 9:21 porque as que estão claras, talvez essa ajuda até atrapalharia pelo fato de mudar nosso entendimento...

Linked Codes:

– is part of → ○ A seção de ajuda melhora o entendimento de algumas heurísticas.

Groups:

◊ Uso da ajuda.

○ **Heurísticas estavam claras e bem descritas.**

Created: 29/08/19 by Felipe S. Manzoni, **Modified:** 29/08/19 by Felipe S. Manzoni

Used In Documents:

📄 11 AV_28

Quotations:

🗨️ 11:7 acho que tudo ficou bem claro e que todas as heurísticas estavam bem descritas
 🗨️ 11:10 Não seria necessário estar em todas as heurísticas porque várias delas estão bem explicadas e bem fá...

Groups:

◊ Uso da ajuda.

○ **Seção de ajuda apoiou na decisão da heurística correta para o defeito.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 9 AV_13

Quotations:

🗨️ 3:11 quando eu encontrava mais de uma heurística que poderia se encaixar no problema, eu lia a ajuda, se...
 🗨️ 9:18 Quando chegava em uma heurística que possuía isso, eu lembro de um ponto que isso aqui me ajudou a p...

Linked Codes:

– is associated with → ○ [CAT] Uso da ajuda.

Groups:

◊ Uso da ajuda.

○ **Seção de ajuda apoiou quando em dúvida entre heurísticas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05

Quotations:

3:24 A seção de ajuda das heurísticas ajudou ou facilitou de alguma forma no entendimento das heurísticas...

Groups:

◊ Uso da ajuda.

○ **Seção de ajuda não foi utilizada.**

Created: 26/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

1 AV_01 2 AV_03 4 AV_08 10 AV_17

Quotations:

1:8 Eu percebi a existência dessa seção, mas não utilizei a mesma em nenhum momento. Ela foi indiferente...

2:5 04. A seção de ajuda das heurísticas ajudou ou facilitou de alguma forma no entendimento das heurís...

4:8 Sim, percebi que existia uma seção de ajuda para algumas heurísticas, mas as mesmas foram indiferent...

10:7 A seção de ajuda das heurísticas ajudou ou facilitou de alguma forma no entendimento das heurísticas...

Linked Codes:

– is associated with → ○ [CAT] Uso da ajuda.

← is cause of – ○ eu conseguia entender apenas lendo a pró

Groups:

◊ Uso da ajuda.

No code group

65 Codes:

○ **A classificação das heurísticas não ajudou os avaliadores.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

○ **A classificação ajudou a entender melhor as heurísticas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

11 AV_28

Quotations:

11:1 E esta classificação das heurísticas ajudou durante a avaliação ou facilitou na identificação do que...

○ **A classificação atrapalhou por estar muito no início.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

2 AV_03

Quotations:

2:15 Eu achei que ela me atrapalhou, até porque ela ficou muito no início.


○ **A classificação das heurísticas atrapalhou o avaliador.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 2 AV_03

Quotations:

 2:14 E esta classificação das heurísticas ajudou durante a avaliação ou facilitou na identificação do que...


○ **A classificação das heurísticas não focou nos aspectos do final do jogo.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 2 AV_03

Quotations:

 2:8 Ela deu ênfase no início do jogo e toda a classificação inicial. Ficou nesses detalhes assim, mais,...


○ **A classificação do jogo ajuda a compreender qual seção do jogo estamos avaliando.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 4 AV_08

Quotations:

 4:18 ela me dava uma situada no que que eu estou exatamente avaliando com cada uma das heurísticas, que p...


○ **A divisão das heurísticas não foi interligada com a primeira página.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 9 AV_13

Quotations:

 9:28 Eu notei que existia uma divisão, mas eu não entendi que era por conta dessa primeira página e que e...


○ **A ordem dos defeitos na tabela foi confundida com a sequência das categorias no conjunto.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 9 AV_13

Quotations:

 9:25 Porque, na folha que foi dado para mim, tinham três campos que deveriam ser preenchidos: tinha a ord...


○ **A primeira página foi compreendida como a classificação das heurísticas.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 2 AV_03  3 AV_05  4 AV_08  10 AV_17

Quotations:

 2:6 Ficou entendido que a primeira página do conjunto eram as características, ou a classificação, das h...

- ⌚ 3:13 Ficou entendido que a primeira página do conjunto eram as características, ou a classificação, das h...
- ⌚ 4:10 Ficou entendido que a primeira página do conjunto eram as características, ou a classificação, das h...
- ⌚ 10:10 Ficou entendido que a primeira página do conjunto eram as características, ou a classificação, das h...

○ **A repetição de heurísticas apoiou na escolha da heurística mais apropriada ao defeito.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 9 AV_13

Quotations:

- ⌚ 3:3 até de certa forma repetitivas, porque isso ajudou, isso ajudava a clarear mais a observação, a part...
- ⌚ 3:5 mas ao mesmo tempo me fazia me prender a qual a heurística que melhor se encaixa nesse problema aqui...
- ⌚ 9:9 acho que o tamanho foi uma coisa positiva, porque era bem completo e quando eu precisava de alguma c...

○ **As heurísticas do conjunto foram compreendidas.**

Created: 26/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01 📄 2 AV_03 📄 3 AV_05 📄 4 AV_08 📄 9 AV_13 📄 10 AV_17 📄 11 AV_28

Quotations:

- ⌚ 1:5 02. Houve dificuldade de entendimento de alguma heurística do conjunto? Não sei se é porque já te...
- ⌚ 2:4 02. Houve dificuldade de entendimento de alguma heurística do conjunto? Não.
- ⌚ 3:6 Houve dificuldade de entendimento de alguma heurística do conjunto? Eu acho que de modo geral não,...
- ⌚ 4:7 02. Houve dificuldade de entendimento de alguma heurística do conjunto? Não, eu compreendi todas...
- ⌚ 9:12 Houve dificuldade de entendimento de alguma heurística do conjunto? Sim, eu entendi o que as heurí...
- ⌚ 10:6 Houve dificuldade de entendimento de alguma heurística do conjunto? Sim, compreendi todas as heurí...
- ⌚ 10:9 Porque as heurísticas em si já eram o suficiente para mim. A heurística em si já é bem autoexplicati...
- ⌚ 11:6 Houve dificuldade de entendimento de alguma heurística do conjunto? Não tive dificuldades, acho qu...

○ **Característica de bug identificado.**

Created: 26/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 4 AV_08 📄 9 AV_13 📄 11 AV_28

Quotations:

- ⌚ 4:4 travar durante o jogo
- ⌚ 4:5 coisas que estavam fora do meu campo de visão e eu conseguia atingir
- ⌚ 9:3 o jogo, para mim, ele instalou duas vezes no meu celular, tinham dois ícones,
- ⌚ 9:4 isso é um bug que existe, ou, por exemplo, tem partes do jogo que estão em linguagens diferentes, is...
- ⌚ 9:5 Eu também notei que um dos meus colegas quando ele instalou o jogo apareceu uma

notificação em chinê...

🕒 11:4 eu estava no carro e ficava presa dentro do carro

Linked Codes:

← is a – ○ Falta heurística sobre objetos fora do campo de visão.

← is a – ○ Falta heurística sobre travamento do jogo.

○ **Categorias ajudaram a compreender o processo de avaliação, quando utilizar cada uma delas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 4 AV_08

Quotations:

🕒 4:19 E esta classificação das heurísticas ajudou durante a avaliação ou facilitou na identificação do que...

○ **Categorias na primeira página atrapalham o avaliador.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01

Quotations:

🕒 1:19 Na verdade só, tipo, aquela primeira página lá, eu não usei e ela chegava tipo a me incomodar as vez...

○ **Certas heurísticas não precisam de ajuda.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 11 AV_28

Quotations:

🕒 3:10 Não acho que faltou a seção de ajuda em nenhuma das heurísticas e como era bastante coisa então tipo...

🕒 11:10 Não seria necessário estar em todas as heurísticas porque várias delas estão bem explicadas e bem fá...

Linked Codes:

– is associated with → ○ A seção de ajuda melhora o entendimento de algumas heurísticas.

○ **Classificação criou uma expectativa de ordem de atividades no jogo.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01

Quotations:

🕒 1:23 Sim, ela tipo, antes de eu começar a jogar eu lia aquilo e eu meio que: “hummm, deve ter algo que mei...


○ **Classificação das heurísticas ajudou na avaliação.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 4 AV_08

Quotations:

 4:15 então acho que a ordem não altera tanto, mas eles estarem em grupos ajuda bastante.


○ **Classificação das heurísticas ajudou na definição da heurística correta ao defeito.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 3 AV_05

Quotations:

 3:15 eu utilizei mais no início, assim, para tentar definir qual seria o mais apropriado para um defeito...


○ **Classificação das heurísticas foi pouco utilizada.**


Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 3 AV_05  10 AV_17

Quotations:

 3:14 E esta classificação das heurísticas ajudou durante a avaliação ou facilitou na identificação do que...

 10:11 E esta classificação das heurísticas ajudou durante a avaliação ou facilitou na identificação do que...


○ **Classificação dentro do conjunto é percebida.**


Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni


Used In Documents:

 1 AV_01  3 AV_05  4 AV_08

Quotations:

 1:12 Eu percebi que as heurísticas estavam subdivididas ou classificadas

 3:15 eu utilizei mais no início, assim, para tentar definir qual seria o mais apropriado para um defeito...

 4:15 então acho que a ordem não altera tanto, mas eles estarem em grupos ajuda bastante.


○ **Classificação dentro do conjunto não foi entendida.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 9 AV_13

Quotations:

 9:27 por eu não ter entendido a página inicial, eu nem me toquei que ele estava dividido, eu não notei es...

○ **Com o tempo, a classificação não é mais necessária, vai direto para a heurística.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 3 AV_05

Quotations:

- ⌚ 3:25 eu utilizei mais no início, assim, para tentar definir qual seria o mais apropriado para um defeito...

- **Confusão das categorias com as próprias heurísticas.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

- 📄 1 AV_01
- 📄 11 AV_28

Quotations:

- ⌚ 1:11 confundindo a primeira página com aquela introdução de tipo: "Todo jogo deve ter uma história", é is...
- ⌚ 11:13 acho que se ela estivesse em outro lugar eu acabaria confundindo elas com heurísticas por serem bem...

- **Defeito atribuído a mais de uma heurística.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

- 📄 1 AV_01
- 📄 9 AV_13

Quotations:

- ⌚ 1:17 inclusive tem muito defeito que eu coloquei uma ou mais heurísticas, teve um que eu botei 3, que era...
- ⌚ 9:10 Mas no fim, certos erros que eu identifiquei, acabaram ficando classificados em diferentes heurístic...

- **Defeito de usabilidade da técnica.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

- 📄 1 AV_01

Quotations:

- ⌚ 1:20 Na verdade só, tipo, aquela primeira página lá, eu não usei e ela chegava tipo a me incomodar as vez...

- **Defeitos identificados não seguem uma sequência lógica.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

- 📄 2 AV_03
- 📄 4 AV_08

Quotations:

- ⌚ 2:11 Preferi fazer aleatoriamente, até porque tu procurar e achar, evidente, um defeito com aquela classi...
- ⌚ 4:14 Mesmo seguindo eu acho que não teria nem me ajudado nem me atrapalhado porque teve momentos em que e...

- **Dificuldade de identificar missões do jogo.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

- 📄 2 AV_03

Quotations:

⌚ 2:13 ai no final já do tempo é que eu entendi que ele tinha que ir com um NPC, pegar a instrução, pra que...

○ **Dificuldade de identificar o objetivo do jogo.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 9 AV_13

Quotations:

⌚ 3:21 Houve dificuldade de entendimento do jogo em sí? Ou seja, foi difícil de avaliar o jogo dado? O jo...

⌚ 9:34 Mas uma coisa que eu acho que pecou bastante, que eu até sublinhei no meu review, foi que eu não sab...

○ **Dificuldade em como utilizar as heurísticas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 9 AV_13

Quotations:

⌚ 9:13 só tive problema em entender como utilizar elas, tipo, essa parte inicial aqui e etc. Como botar tud...

○ **É necessário incluir ajuda para heurísticas que não estão claras.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 9 AV_13

Quotations:

⌚ 9:20 Na minha percepção, seria interessante naquelas que não estão claras

○ **Entende as heurísticas como um conjunto.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 11 AV_28

Quotations:

⌚ 11:2 além de mais ou menos ver o que essas heurísticas tinham em comum entre elas dentro dessas seções.

○ **eu conseguia entender apenas lendo a pró**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 4 AV_08 📄 10 AV_17

Quotations:

⌚ 4:17 eu conseguia entender apenas lendo a própria heurística.

⌚ 10:9 Porque as heurísticas em si já eram o suficiente para mim. A heurística em si já é bem autoexplicati...

Linked Codes:

– is cause of → ○ Seção de ajuda não foi utilizada.



○ **Falta heurística sobre a parte gráfica.**

Created: 26/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 4 AV_08

Quotations:

-  4:1 parte de gráficos não estava muito bem coberta.
-  4:16 Apenas, novamente, a falta de heurísticas da parte gráfica para adicionar.

Linked Codes:

- ← is part of – ○ Falta heurística sobre objetos fora do campo de visão.
- ← is part of – ○ Falta heurística sobre travamento do jogo.
- is cause of → ○ Uso de heurística inadequada na falta da correta.


○ **Falta heurística sobre consistência de linguagem utilizada.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 9 AV_13

Quotations:

-  9:6 tem partes do jogo que estão em inglês e outras em português, no momento eu não saberia apontar quai...


○ **Falta heurística sobre objetos fora do campo de visão.**

Created: 26/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 4 AV_08

Quotations:

-  4:6 ou algumas interações existem coisas que estavam fora do meu campo de visão e eu conseguia atingir,...

Linked Codes:

- is a → ○ Característica de bug identificado.
- is part of → ○ Falta heurística sobre a parte gráfica.


○ **Falta heurística sobre travamento do jogo.**

Created: 26/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 4 AV_08

Quotations:

-  4:3 A parte de gráficos estava deficiente, tipo, travar durante o jogo e essas coisas,

Linked Codes:

- is a → ○ Característica de bug identificado.
- is part of → ○ Falta heurística sobre a parte gráfica.

○ **Falta uma classificação de severidade para os defeitos.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 2 AV_03

Quotations:

- 2:9 Acho que apenas faltou uma classificação de nota (severidade) para os defeitos.

- **Faltaram heurísticas para avaliação de bugs.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

- 2 AV_03
- 9 AV_13
- 11 AV_28

Quotations:

- 2:2 Não, mas eu reavaliei isso depois por causa que eu procurei tipo, uma heurística que era um 'bug' e...
- 9:2 mas neste quesito eu acho que o jogo tem mais do que tipo o que está dentro dele, por exemplo, teve...
- 11:3 mas eu vi a maioria delas, mas eu acho que talvez, bugs que aconteceram no jogo não apareceram por a...

- **Foi necessário reler algumas das heurísticas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

- 3 AV_05

Quotations:

- 3:7 talvez exista uma, mas eu não lembro exatamente qual era a heurística, essa eu precisei reler e pres...

- **Heurísticas do conjunto cobriram todos os aspectos necessários do jogo.**

Created: 29/08/19 by Felipe S. Manzoni, **Modified:** 29/08/19 by Felipe S. Manzoni

Used In Documents:

- 11 AV_28

Quotations:

- 11:5 O número, e tipo, de heurísticas no conjunto foi o suficiente para realizar a avaliação do jogo? Ou...

- **Heurísticas do conjunto eram repetitivas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

- 3 AV_05
- 9 AV_13

Quotations:

- 3:2 Ou seja, faltou alguma heurística ou característica a ser avaliada pelo conjunto no jogo? Assim, e...
- 3:20 algumas categorias de heurísticas, em algumas elas se tornam bastante redundantes.
- 9:11 mais de uma heurística também e isso eu fiz porque as heurísticas eram mais específicas e tinham vár...

- **Heurísticas necessárias estavam presentes na hora da avaliação.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

- 1 AV_01
- 2 AV_03
- 9 AV_13
- 10 AV_17

Quotations:

- 🗨️ 1:1 Sim, tudo que eu precisei eu encontrei lá, não faltou nada na minha opinião.
- 🗨️ 2:3 Não faltaram heurísticas dentro dos erros que eu achei/procurei.
- 🗨️ 9:1 O número, e tipo, de heurísticas no conjunto foi o suficiente para realizar a avaliação do jogo? Ou...
- 🗨️ 9:7 Mas as coisas que eu mesmo avaliei e etc, com isso aqui, ajudou bastante e foi o suficiente.
- 🗨️ 10:1 O número, e tipo, de heurísticas no conjunto foi o suficiente para realizar a avaliação do jogo? Ou...

○ **Heurísticas sobre controles estão ambíguas.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01

Quotations:

- 🗨️ 1:18 se era fácil ou se eram entendíveis os controles, aí coube uma 3 nesse defeito aí.

○ **Heurísticas utilizadas como um checklist.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 4 AV_08

Quotations:

- 🗨️ 3:17 Para o fim eu já estava mais fazendo um checklist e comparando com o que eu tava vendo do que explor...
- 🗨️ 4:13 Eu joguei primeiro, depois eu fui lendo heurística por heurística vendo se eu tinha sentido/visto/ob...

○ **Houve dificuldade no entendimento do jogo.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 2 AV_03 📄 9 AV_13

Quotations:

- 🗨️ 2:12 Houve dificuldade de entendimento do jogo em sí? Ou seja, foi difícil de avaliar o jogo dado? A pr...
- 🗨️ 9:35 Houve dificuldade de entendimento do jogo em sí? Ou seja, foi difícil de avaliar o jogo dado? Não...

○ **Jogador explorou as mecânicas para compreender as mesmas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 9 AV_13

Quotations:

- 🗨️ 3:23 tanto que eu fiquei explorando ele por um bom tempo, sem buscar objetivo nenhum, só explorando mesmo...
- 🗨️ 9:33 só joguei meio que testando tipo assim, oque acontece se eu fizer isso e isso e isso, dessa forma eu...

○ **Leitura prévia das heurísticas ajudou no processo de avaliação.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 9 AV_13

Quotations:

🕒 9:31 Do jeito que eu fiz, eu já sabia que quando existia um defeito ele era dessa heurística aqui e que e...

○ **Mapa do jogo não foi percebido.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05

Quotations:

🕒 3:22 Tanto que eu nem percebi que por exemplo, o mapa você conseguia abrir ele para ver tudo que ele podi...

○ **Não é necessário incluir novas heurísticas.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01 📄 3 AV_05 📄 9 AV_13 📄 10 AV_17

Quotations:

- 🕒 1:2 Achei que o conjunto estava ok e eu não incluiria nenhuma heurística.
- 🕒 3:1 Assim, eu acho que elas abrangiam bastante coisas e até, até de certa forma repetitivas, porque isso...
- 🕒 9:7 Mas as coisas que eu mesmo avaliei e etc, com isso aqui, ajudou bastante e foi o suficiente.
- 🕒 9:9 acho que o tamanho foi uma coisa positiva, porque era bem completo e quando eu precisava de alguma c...
- 🕒 10:5 Também não acho que faltou heurísticas, estava bem completo.

○ **Não foi entendido como localizar as missões.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01

Quotations:

🕒 1:22 Eu não conseguia entender onde era a localização dessas missões no mapa.

○ **O conjunto é extenso.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 9 AV_13 📄 10 AV_17

Quotations:

- 🕒 3:12 e como era bastante coisa então tipo, algumas delas eu realmente passei por cima,
- 🕒 3:19 Comentários extras: Eu acredito que seria interessante “enxugar” algumas categorias de heurísticas...
- 🕒 9:8 E eu acho que o tamanho do conjunto, mesmo sendo um pouco extenso, mais me ajudou por ser bem comple...
- 🕒 10:2 achei um pouco volumoso pela quantidade de heurísticas
- 🕒 10:3 O tamanho me atrapalhou um pouco, porque fica sendo um pouco mais extenso para fazer a avaliação e v...


- **O jogo foi erroneamente compreendido.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 1 AV_01

Quotations:

 1:21 Então, eu acho que cheguei a jogar, mais ou menos, 1 hora, e eu entendi que deveríamos fazer algumas...


- **Ordem cronológica das heurísticas e categorias foi percebida.**


Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

 1 AV_01  10 AV_17

Quotations:

 1:14 antes de eu começar a jogar eu lia aquilo e eu meio que: “hummm, deve ter algo que meio nessa ordem a...

 10:13 Mas essa classificação dentro do conjunto elas me ajudaram por estarem meio em ordem cronológica do...


- **Primeira página compreendida como a descrição das seções básicas do jogo.**

Created: 29/08/19 by Felipe S. Manzoni, **Modified:** 29/08/19 by Felipe S. Manzoni

Used In Documents:

 11 AV_28

Quotations:

 11:11 Ficou entendido que a primeira página do conjunto eram as características, ou a classificação, das h...


- **Primeira página do conjunto não é compreendida como a descrição das categorias do conjunto.**


Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni


Used In Documents:


 1 AV_01  9 AV_13  11 AV_28

Quotations:

 1:10 Não, é porque eu estou confundindo a primeira página com aquela introdução de tipo: “Todo jogo deve...

 9:24 Ficou entendido que a primeira página do conjunto eram as características, ou a classificação, das h...

 9:32 A única coisa seria especificar melhor essa primeira página aqui porque teria ajudado mais se eu tiv...

 11:12 Ela não me atrapalhou por estrar aqui mas acho que se ela estivesse em outro lugar eu acabaria confu...

- **Processo de avaliação através da memória das heurísticas.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

 1 AV_01  3 AV_05  9 AV_13

Quotations:

- 🕒 1:15 Primeiro eu li e eu ia jogando, e quando algo me incomodava no jogo, eu olhava lá tipo, o que mais e...
- 🕒 3:18 fui ver as heurísticas, tentando minimamente compreendê-las e então tentar a associar com alguma coi...
- 🕒 9:29 eu fui lendo uma heurística por uma, antes de eu começar a avaliar o jogo, somente para me situar, e...
- 🕒 9:30 Seguir a sequência de heurísticas do conjunto ajudou a avaliar o jogo, ou você preferiu identificar...

○ **Processo de avaliação condiz com conhecimento prévio.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01

Quotations:

- 🕒 1:4 não é nada diferente do que eu já havia feito antes.

○ **Processo de avaliação feito identificando defeitos antes das heurísticas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05

Quotations:

- 🕒 3:16 Na verdade eu fiz um processo bem misto, eu comecei tentando explorar o jogo primeiro e depois eu im...

○ **Processo de avaliação foi fácil e simples de aplicar.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 10 AV_17

Quotations:

- 🕒 10:12 eu parti direto para as heurísticas porque ficou bem autoexplicativo o que eu deveria fazer e como e...

○ **Processo de verificação por cada heurística é trabalhoso.**

Created: 27/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 1 AV_01

Quotations:

- 🕒 1:16 Eu tentei, no início eu comecei pegar uma heurística e procurar, só que foi dando muito trabalho

○ **Repetição de heurísticas no conjunto tomavam muito tempo dos avaliadores.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 3 AV_05 📄 10 AV_17

Quotations:

- 🕒 3:4 mas ao mesmo tempo me fazia me prender a qual a heurística que melhor se encaixa nesse problema aqui...

🗨️ 10:4 porque fica sendo um pouco mais extenso para fazer a avaliação e você acaba tendo que dedicar mais t...

○ **Seção de ajuda não é necessária em todas as heurísticas.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 9 AV_13 📄 11 AV_28

Quotations:

🗨️ 9:19 E não que em todas ajudaria mas apenas se fosse necessário, porque tipo, eu não sei se todas precisa...

🗨️ 11:9 Não seria necessário estar em todas as heurísticas porque várias delas estão bem explicadas e bem fá...

○ **Seguir a sequência das heurísticas na avaliação cria uma padronização.**

Created: 29/08/19 by Felipe S. Manzoni, **Modified:** 29/08/19 by Felipe S. Manzoni

Used In Documents:

📄 10 AV_17 📄 11 AV_28

Quotations:

🗨️ 10:14 Seguir a sequência de heurísticas do conjunto ajudou a avaliar o jogo, ou você preferiu identificar...

🗨️ 11:14 Seguir a sequência de heurísticas do conjunto ajudou a avaliar o jogo, ou você preferiu identificar...

○ **Seguir a sequência das heurísticas não ajudou.**

Created: 28/08/19 by Felipe S. Manzoni, **Modified:** 28/08/19 by Felipe S. Manzoni

Used In Documents:

📄 2 AV_03

Quotations:

🗨️ 2:10 Não, a princípio eu tentei fazer a sequência, mas depois eu comecei...como eu comecei a achar erros...

○ **Uso de heurística inadequada na falta da correta.**

Created: 26/08/19 by Felipe S. Manzoni, **Modified:** 27/08/19 by Felipe S. Manzoni

Used In Documents:

📄 4 AV_08

Quotations:

🗨️ 4:2 eu não consegui achar, eu encaixei em uma heurística mas não achei que aquela era a certa, só dei um...

Linked Codes:

← is cause of – ○ Falta heurística sobre a parte gráfica.