

Avaliação Heurística Interdisciplinar e Testes de Usabilidade de um Jogo Sério em Educação Médica

Mônica F. Silva¹, André P. Freire¹,
Miriam M. de C. Graciano², Juliana G. Greggi¹, Marluce R. Pereira¹

¹Instituto de Ciências Exatas e Tecnológicas – Universidade Federal de Lavras (UFLA) – Caixa Postal 3037 – CEP 37200-900 – Lavras – MG – Brazil

²Faculdade de Ciências da Saúde – Universidade Federal de Lavras (UFLA) – Caixa Postal 3037 – CEP 37200-900 – Lavras – MG – Brazil

monica.faria15@gmail.com

{ apfreire, miriam.graciano2, juliana, marluce}@ufla.br

Abstract. *Serious Games are games used to help students and teachers in learning. In the medical field, the InsuOnline game was implemented to train clinicians to use insulin in the treatment of diabetes mellitus, in the context of Primary Health Care. This project aims to investigate empirical evidence of problems in InsuOnline through heuristic evaluation by an interdisciplinary team with experts in usability and medical education, and usability tests having medical students as game users. The results provide better subsidies for developers to use recommendations based on empirical data, allowing the improvement of techniques and methods for the design and effective assessment of Serious Games.*

Resumo. *Jogos sérios são usados para ajudar alunos e professores no aprendizado. Na área de educação médica, o jogo InsuOnline foi implementado para treinar médicos clínicos a usar a insulina no tratamento do diabetes mellitus, no contexto da Atenção Primária à Saúde. Este projeto teve como objetivo investigar evidências empíricas de problemas no InsuOnline por meio de avaliação heurística com time interdisciplinar de especialistas em usabilidade e educação médica, e testes de usabilidade tendo estudantes de medicina como usuários do jogo. Os resultados auxiliam a fornecer melhores subsídios para que desenvolvedores possam usar recomendações baseadas em dados empíricos, permitindo a melhoria de técnicas e métodos para o design e avaliação efetiva de jogos sérios.*

1. Introdução

Os jogos digitais denominados “jogos sérios” são voltados a conteúdos e finalidades específicos. O jogador pode ser treinado para realizar tarefas, utilizar seus conhecimentos para resolver problemas e conhecer novas problemáticas [Machado et al. 2011]. Os jogos ajudam no processo de ensino/aprendizagem, motivando estudantes a aprender conteúdos específicos [Fialho 2008, Boyle et al. 2016, Bai et al. 2020].

A motivação que um jogador apresenta com relação ao jogo sério pode ser comprometida por uma interface mal elaborada e uma baixa usabilidade, prejudicando a experiência do jogador e, conseqüentemente, a diversão e o aprendizado [Federoff 2002, Schmidt and De Marchi 2017a].

Muitos estudos de jogos digitais já foram realizados para investigar uma série de atributos para melhorar a indústria de jogos e conhecer os impactos do uso de jogos sobre os usuários [Caroux et al. 2015, Federoff 2002, Boyle et al. 2016, Schmidt and De Marchi 2017a]. Alguns aspectos importantes de serem analisados são: design, recursos financeiros, qual o contexto em que o serviço/software será utilizado, funções que são disponibilizadas, experiência do usuário ao utilizar o jogo e seu grau de usabilidade.

A educação médica sobre diabetes mellitus e uso de insulina, muitas vezes é realizada através de aulas expositivas, palestras ou grupos de discussão, gerando pouco resultado para o aprendizado dos estudantes de medicina no contexto da Atenção Primária à Saúde. Por isso, são necessários novos métodos educacionais e um deles é a gamificação. Os jogos sérios têm o propósito de criar um ambiente de resolução de problema que pode ser integrado com a educação contínua e contribuem para a melhoria do processo de aprendizado. Desta forma, ensinar utilizando um jogo sério gera motivação, entusiasmo e permite aprender resolvendo problemas de forma ativa e interativa, melhorando o processo de aprendizado [Diehl et al. 2017, Asadzandi et al. 2020].

Alguns trabalhos relacionados a jogos sérios para a área médica já vêm sendo realizados [Siregar and Julen 2016, Schmidt and De Marchi 2017b, Dekker and Williams 2017, Ee et al. 2018, Georgieva-Tsaneva and Serbezova 2020, Lemos et al. 2020]. Especificamente, o uso de gamificação para diabetes já ocorreu em outros trabalhos [Asadzandi et al. 2020], com destaque para o InsuOnline.

O jogo sério InsuOnline é um jogo para treinamento de médicos ou estudantes de Medicina para o manejo de insulina. Ele já vem ganhando atenção nesta área de pesquisa sobre testes de usabilidade. Trabalhos com o objetivo de relatar o resultado da avaliação do InsuOnline, analisando sua eficácia educacional, vêm sendo desenvolvidos desde sua criação [Diehl et al. 2013, Diehl et al. 2015a].

Este artigo apresenta os resultados da avaliação de usabilidade da versão 1.8.4 do InsuOnline, realizada antes da pandemia, com uma combinação de métodos: avaliação heurística interdisciplinar com especialistas em usabilidade e em educação médica e teste de usabilidade com um grupo de 10 estudantes de medicina.

A avaliação heurística foi realizada utilizando as 10 heurísticas de Pinelli para avaliação de jogos [Pinelle et al. 2008] e os graus de severidade de Nielsen [Nielsen 1993], atribuídos aos problemas identificados pelos especialistas.

Os resultados fornecem uma caracterização dos tipos de problemas encontrados por diferentes métodos, fornecendo a desenvolvedores e pesquisadores um melhor entendimento sobre a natureza dos problemas encontrados pelos usuários e como eles são afetados pelos problemas, a partir do uso de diferentes métodos.

O restante deste artigo está organizado da seguinte forma: a Seção 2 apresenta técnicas de se fazer teste de usabilidade, a Seção 3 apresenta a metodologia utilizada

neste trabalho, a Seção 4 mostra os resultados obtidos das avaliações, a Seção 6 apresenta uma breve análise sobre o estudo feito. Finalmente, a Seção 6 apresenta a conclusão e propostas de trabalhos futuros.

2. Avaliação de Usabilidade

De acordo com a International Organization for Standardization (ISO), a usabilidade pode ser encarada como uma medida de como um produto pode ser usado por usuários específicos para alcançar objetivos específicos como eficácia, eficiência e satisfação, num contexto de utilização específico [ISO9241-11 1998, Nielsen 2012].

Neste trabalho, as técnicas utilizadas para avaliar usabilidade foram coleta da opinião de especialistas por meio da avaliação heurística e observação de usuários por meio do teste de usabilidade.

A **avaliação heurística** é um método de inspeção de usabilidade criada por Nielsen e Molich [Nielsen 1993], tendo como característica ser um método relativamente barato e rápido para encontrar problemas em interfaces de usuários em aplicações desktop. Neste tipo de avaliação, os avaliadores têm uma liberdade significativa em como eles conduzem a avaliação. Avaliadores exploram a interface enquanto procuram por instâncias que estão relacionadas a um conjunto de princípios de usabilidade chamado heurística. Porém, as heurísticas de Nielsen e Molich não abordam várias questões de usabilidade importantes no contexto de jogos, como o uso de ângulos de câmera apropriados ao exibir o mundo do jogo ou fornecer mapeamentos de controle intuitivo. Por isso, [Pinelle et al. 2008] criou um conjunto de heurísticas voltadas para avaliação de jogos.

Teste de usabilidade é uma técnica de pesquisa utilizada para avaliar um produto ou serviço. Consiste em observar os usuários que estão testando um sistema individualmente e permitir que eles resolvam problemas encontrados por si mesmos. Os testes são realizados com usuários representativos do público-alvo. Cada participante tenta realizar as tarefas típicas enquanto o analista observa, ouve e faz anotações [Nielsen 2012].

3. Metodologia

Segundo Nielsen, um avaliador pode realizar uma avaliação heurística sozinho, porém, um avaliador sozinho não irá perceber a maior parte dos problemas em uma interface. Em média, um único avaliador encontra cerca de 35% dos problemas de usabilidade em interfaces. Como cada avaliador tende a encontrar problemas diferentes uns dos outros, Nielsen recomenda um grupo de 3 a 5 avaliadores, assim, 75% a 85% dos problemas de interface considerados graves podem ser encontrados pelo grupo de avaliadores [Nielsen 1993]. As avaliações ocorreram com 3 grupos de avaliadores: o primeiro grupo, composto por 4 avaliadores e 2 aplicadores, tendo os aplicadores responsabilidade de fazer a interação com o sistema enquanto os avaliadores apontavam possíveis problemas e, o outro aplicador, anotar os problemas apontados em uma lista, para posteriormente, os avaliadores analisarem o grau de severidade individualmente. Neste grupo, contou-se com a presença de um avaliador da área médica entendedor do assunto abordado no jogo. O outro grupo de avaliadores também era composto por 4 avaliadores e 2 aplicadores, sendo suas tarefas as mesmas do grupo anterior, porém, neste grupo não havia nenhum avaliador entendedor do assunto abordado no jogo. E, o último grupo, composto pelos usuários.

Este projeto foi aprovado sob o código CAAE 54830016.6.0000.5148, pelo Comitê de Ética em Pesquisa em Seres Humanos (COEP).

As avaliações da versão 1.8.4 do jogo sério InsuOnline foram realizadas em laboratório de pesquisa de uma instituição de ensino superior, antes da pandemia.

3.1. Avaliação Heurística Interdisciplinar

Foi realizada avaliação heurística para avaliar a usabilidade e interatividade com o jogo. Os especialistas foram professores de Ciência da Computação com experiência nesse tipo de avaliação e um docente com formação na área médica e experiência com ensino em Medicina. Para o especialista da área médica, foi feito um treinamento prévio, antes da data marcada para esta avaliação. O projeto teve, como objetivo, a avaliação de aspectos relacionados com a usabilidade de interfaces a fim de detectar problemas de projeto e fazer recomendações para eliminar tais problemas.

No presente trabalho, os especialistas fizeram a avaliação em 2 horas, apontando falhas no jogo de acordo com as heurísticas de Pinelli [Pinelle et al. 2008], descritas a seguir:

1. Fornecer respostas consistentes para as ações dos usuários.
2. Permitir que os usuários personalizem configurações de vídeo, áudio, nível de dificuldade e velocidade.
3. Fornecer comportamento previsível e razoável para unidades controladas pelo computador.
4. Fornecer visão sem obstruções que sejam apropriadas para as ações atuais dos usuários.
5. Permitir que os usuários pulem ações repetidas e que não fornecem jogabilidade.
6. Fornecer mapeamentos de entrada personalizáveis e intuitivos.
7. Fornecer controles que sejam fáceis de gerenciar, e que tenham um nível apropriado de sensibilidade e responsividade.
8. Fornecer informações sobre o status do jogo para os usuários.
9. Fornecer instruções, treinamento e ajuda.
10. Fornecer representações visuais que sejam fáceis de interpretar e minimizem a necessidade de microgerenciamento.

Dois colaboradores foram responsáveis por conduzir o teste e anotar as falhas encontradas pelos 4 avaliadores voluntários. Cada avaliador atribuiu uma nota de severidade para cada falha apontada. Os níveis de severidade dados pelos especialistas foram propostos por [Nielsen 1993]:

- 0** - Não é considerado, totalmente, um problema de usabilidade;
- 1** - Problema apenas estético: não necessita ser consertado a menos que tenha tempo extra disponível no projeto;
- 2** - Problema menor de usabilidade: o conserto deste problema deverá ser de baixa prioridade;
- 3** - Problema maior de usabilidade: é importante consertá-lo, para isso deverá ser dada alta prioridade;
- 4** - Catástrofe de usabilidade: é obrigatório consertá-lo antes do produto ser divulgado.

O método escolhido foi a Avaliação Heurística Colaborativa por ter a facilidade de reunir os especialistas em um mesmo local. Assim, estes efetuam a avaliação no mesmo local e ao mesmo tempo e, uma lista de problemas é feita em conjunto, mas cada avaliador atribui níveis de severidade individualmente [Petrie 2010].

3.2. Teste de usabilidade

Nesta etapa, a amostra que avaliou o jogo foi composta por 10 alunos do curso de Medicina, cursando o primeiro, o segundo ou o terceiro período do curso. Alguns desses estudantes já viram, em algum momento do curso, a matéria que fala sobre diabetes porém, o assunto sobre insulina, não foi abordado no início do curso.

A avaliação com usuários é importante para saber se o jogo treina adequadamente os usuários nas tarefas propostas e no ambiente em que será utilizado. A avaliação de interface é necessária para analisar a qualidade de uso do software [Prates 2003].

Os usuários tiveram, como tarefa, jogar em 20 minutos a primeira fase do jogo sério. Durante a realização dos testes, foi aplicada a metodologia *Think Aloud* que é utilizada para identificar problemas de usabilidade com base nas ações e dificuldades dos usuários utilizando o sistema em questão [Renzi 2010]. Segundo Yoshida [Yoshida 2008], o método *Think Aloud* requer que os participantes falem ao pesquisador o que estão pensando e fazendo, ao executar uma tarefa. Os participantes são geralmente instruídos a manter o pensamento em voz alta, agindo como se eles estivessem sozinhos em um ambiente, falando para si mesmos. O método *Think Aloud* é gravado e depois transcrito para análise de conteúdo. Uma das principais vantagens do método é os avaliadores poderem relatar seus pensamentos ao mesmo tempo em que estão envolvidos no desempenho de uma tarefa, evitando assim, efeitos de latência que permitam lembrar de informações enquanto pensam em voz alta.

Nesta metodologia, contou-se com a ajuda de celulares para gravar a voz de cada participante. Também foi utilizado o software livre ShadowPlay [Nvidia] que permite fazer a gravação do jogo e, assim, registrar toda a ação do usuário (desde clicadas de mouse até ações do teclado) para analisar como cada participante se saiu no jogo, como por exemplo: o que deveriam fazer, para onde deveriam ir, o que fizeram, etc.

Os usuários poderiam indicar, a qualquer momento, quando encontrassem algum problema. Os problemas também foram registrados em uma ficha pelo pesquisador, que trabalhou somente como um "facilitador" da sessão e procurou interferir o mínimo possível com a realização das tarefas por cada participante.

4. Resultados

Esta seção apresenta os resultados da avaliação heurística com especialistas da área de computação e médica, somente com especialistas da área de computação e teste de usabilidade com estudantes de Medicina.

4.1. Avaliação heurística com Especialistas das Áreas de Computação e Médica

Durante a avaliação foram coletados 39 problemas, sendo que a Tabela 1 apresenta os 10 problemas com maior média de severidade coletados na avaliação heurística, as heurísticas aplicáveis, as severidades atribuídas pelos avaliadores e a média aritmética (\bar{m}) das severidades.

Local do problema	Descrição do problema	Heurísticas	Avaliadores				
			A1	A2	A3	A4	\bar{m}
Jogar	Informações incorretas nas informações passadas pelas instruções do jogo (momento da leitura de prontuário) - O prontuário deve ser lido antes do início da consulta.	H8	4	4	3	4	4
Jogar	Volume do áudio é maior do que a "voz" do personagem.	H2	4	2	3	3	3
Jogar-Diálogo	Não é visível que o usuário deve escolher o que falar para a paciente.	H8	3	3	3	3	3
Jogar-Diálogo	Poucas opções de escolha para o diálogo. Por ser um jogo educativo, poderia haver mais opções a serem escolhidas, e cada opção mudaria o humor da paciente.	H1	4	4	2	4	3
Jogar-Diálogo	Não é visível para o usuário qual personagem está falando no diálogo.	H8	4	3	3	3	3
Jogar-Diálogo	Opções de diálogo são ruins. São muito ruins ou não são bem pensados em relação à uma educação integrada. Assim, foca apenas em ensinar sobre insulina e não em um atendimento como um todo.	H8	4	3	4	4	4
Jogar-Diálogo	Botão do canto superior esquerdo da tela não é informativo sobre suas ações.	H10	3	4	3	3	3
Jogar-Diálogo	Botão "Voltar" dentro do botão no canto superior esquerdo não volta para o início do jogo. Apenas sai do menu.	H10	3	4	3	4	3
Jogar-Diálogo	Não há opção de reiniciar fase.	H2	3	2	4	3	3
Jogar-Exames	Dica dada pelo mentor dá a resposta final para o usuário em relação à consulta, não realizando, assim, nenhum fim educativo.	H3	4	2	2	4	3

Tabela 1. Resultado Resumido da Avaliação com os Especialistas com Participação de um Avaliador da Área Médica

Dentre os problemas encontrados, a heurística mais violada foi a heurística H10, "Fornecer representações visuais que são fáceis de interpretar e que minimizem a necessidade de microgestão", com 13 ocorrências. Esta heurística diz que representações

visuais devem ser concebidas de modo que elas sejam fáceis de interpretar, minimizando confusões e oclusões. Assim, os usuários podem diferenciar elementos importantes de elementos irrelevantes. Neste jogo, essa heurística não é fortemente seguida, por exemplo, em um dos casos em que o botão esquerdo superior da tela não é sugestivo às ações que o responde.

Em segundo lugar, tem-se a heurística H8, “Fornecer aos usuários informações sobre o estado de jogo”, com 11 ocorrências. Nela, devem ser fornecidas informações suficientes que permitam aos usuários tomar decisões apropriadas durante o jogo. Durante a avaliação, os avaliadores se sentiram “perdidos” para tomar algumas decisões, devido à falta de informações que não foram dadas e/ou não estavam claras aos avaliadores. Por exemplo, essa violação ocorre quando é aberto o diálogo entre o paciente e o médico e os avaliadores não sabiam quem estava falando e o que deveriam fazer naquele momento.

As heurísticas H4, H5 e H6, não foram relacionadas aos problemas encontrados pelos especialistas.

Analisando as severidades dadas pelos avaliadores, a maioria dos problemas encontrados apresentou “severidade 2” - problema de baixa prioridade; em segundo ficou a “severidade 3” - problema de alta prioridade.

4.2. Avaliação heurística com Especialistas somente da Área de Computação

Durante a avaliação heurística sem a participação de um avaliador da área médica foram coletados 32 problemas. A Tabela 2 apresenta os 10 problemas com maior média aritmética das severidades coletados na avaliação heurística.

Após as análises dos 32 problemas coletados, conclui-se que a heurística mais violada foi a heurística H10, “Fornecer representações visuais que são fáceis de interpretar e que minimizem a necessidade de microgestão”, com 15 ocorrências. Um exemplo em que essa heurística não está sendo seguida, diz respeito à localização do botão voltar na tela de anotações não estar visível ao usuário, deixando-o confuso. A segunda heurística mais violada foi a H9, “Fornecer instruções, treinamento e ajuda”, com 4 ocorrências. Um exemplo em que essa heurística não está sendo seguida é facilmente identificada na abordagem de se iniciar o diálogo entre paciente e médico. Não há instruções suficientes para realizar esta tarefa. As heurísticas H5 e H6, não foram relacionadas aos problemas encontrados pelos especialistas.

Analisando as severidades atribuídas pelos avaliadores, a maioria dos problemas encontrados apresentou “severidade 3” - problema de alta prioridade; em segundo lugar, ficou a “severidades 4” - catastrófica.

4.3. Resultados do teste de usabilidade

Durante o teste com usuários foram coletados 57 problemas, sendo que 4 deles foram relatados por 9 dos 10 usuários que realizaram o teste de usabilidade do jogo. Os problemas foram: poucas opções de diálogo; o exame físico deveria ser mais detalhado; na tela de desafio, o botão “confirmar” fica apagado com a cor cinza; quando se inicia o diálogo, fica confuso como proceder após o primeiro diálogo. Com 8 ocorrências tem-se 4 problemas: “Fazer a dosagem de insulina” não explica claramente o que deve ser feito; Após finalizar a fase, não fica claro que pode ver seu progresso; Círculo ao lado no nome do primeiro caso não é botão de opção; Onde está o mentor 1, não é tão chamativo o ícone piscando.

Tabela 2. Resultado Resumido da Avaliação com os Especialistas sem a Participação de um Avaliador da Área Médica

Local do problema	Descrição do problema	Heurísticas	Avaliadores				
			A1	A2	A3	A4	\bar{m}
Modal de 'Casos'	A seta abaixo da descrição não possui ação	H1	4	1	3	2	3
Modal de 'Casos'	A barra de rolagem não possui utilidade até o momento	H1	3	1	3	2	2
Modal de 'Casos'	Não há indicação de dificuldade. Não há, por exemplo, um limite máximo apresentado para o usuário	H2	2	3	4	4	3
Jogo	A movimentação do personagem através da barra de rolagem é ruim	H7	3	2	3	2	3
Jogo	O personagem já inicia a fase no meio do cenário	H4	3	0	0	2	1
Anotações	A localização do botão "Voltar" na aba de 'Anotações' não está boa.	H10	3	3	3	3	3
Exames	Dois botões para a mesma ação 'Continuar' e 'X'.	H10	2	3	4	4	3
Desafio	Momento em que o 'Desafio' é apresentado poderia ser melhorado, sendo apresentado em momentos mais propícios.	H10	2	2	3	2	2
Desafio	Cores dos botões 'Confirmar' e 'Continuar' está muito ruim	H10	3	3	4	2	3
Jogo	Não há um bom retorno para o usuário em relação ao status do sistema, como, por exemplo, quando uma fase se encerra.	H8	3	3	3	3	3

A Tabela 3 apresenta mais 10 problemas coletados no teste de usabilidade e a localização do problema. O conjunto completo dos problemas encontrados durante a avaliação heurística e os testes com usuários podem ser acessados neste [documento](#).

5. Discussão

Após todas as análises feitas, foi possível concluir que, na presença do avaliador da área médica, foram encontrados problemas que, sem sua presença, não foram percebidos, tais como o prontuário não sendo exibido antes da consulta, falta de informações sobre o paciente, falta de informação passada pelo médico, dentre outras. Do mesmo modo, os especialistas da área da computação identificaram problemas que, antes, passaram despercebidos - como na tela do desafio, em que a nuvem branca por trás da questão não deixa os usuários totalmente concentrados no problema sugerido. Outro problema encontrado, e que antes não havia se manifestado, foi na tela do exame em que o botão "fecha" e o botão "continuar" desenvolvem a mesma funcionalidade.

Tabela 3. Síntese dos Tipos de Problemas Encontrados no Teste com os Usuários

Localização do problema	Problema	Número de Ocorrências
Selecionar Caso	Dúvida se pode escolher outros casos	4
Diálogo	Diálogos são “grosseiros”	3
Jogo	As vozes dos personagens são bem estranhas	5
Diálogo	Tipo de consulta Anamnese, falta mais diálogo sobre a rotina do paciente	1
Diálogo	“Por que ele tá no lugar do Dr. Braga, como ele foi parar ali?” Não conta no jogo e o paciente deveria saber.	2
Jogo	Falta de organização no ambiente de saúde. A enfermeira e o médico não estavam corretos em seus comportamentos.	1
Como Jogar	Não fica explícito no jogo que pode se movimentar os pacientes tanto com mouse como com teclado	2
Jogo	Muitos ícones são abertos no decorrer da fase, dificultando, assim, encontrar informações	5
Como Jogar	Instruções de “como jogar” não estão claras	1
Jogo	As informações passadas durante o decorrer do jogo são repetitivas	1

O InsuOnline já foi comparado a métodos tradicionais de ensino (palestras, mesa redonda) para analisar sua eficácia educacional. Como resultado, o jogo é tão eficaz quanto uma atividade educacional tradicional para a educação médica em terapia com insulina e pode ser um bom método de ensino/aprendizagem [Diehl et al. 2015b]. Foram também realizados testes de experiência dos usuários médicos [Diehl et al. 2017] para analisar aceitação e efetividade educacional.

Entende-se que o jogo é capaz de atender ao objetivo que propõe, sem causar muitos problemas aos usuários que se dispõem a utilizá-lo para aprimorar seus conhecimentos sobre o assunto. Porém, é importante que todos os problemas ressaltados pelos avaliadores sejam corrigidos, para que as necessidades dos usuários sejam melhor atendidas e que o objetivo proposto seja alcançado mais rapidamente. A análise dos resultados permitiu obter conclusões sobre os efeitos de diferentes características do jogo, visando a construção de jogos mais adequados para estudantes de Medicina e realizar a correlação com sua motivação de aprendizado.

Os resultados deste estudo também apresentaram aspectos importantes para o planejamento de atividades de avaliação de jogos sérios na área médica e outras áreas similares. Houve problemas específicos que foram identificados somente por usuários e outros somente na avaliação heurística, como é esperado pela complementariedade dos métodos em avaliações de usabilidade. Os dados mostram a importância de efetuar testes com usuários finais, e do envolvimento de estudantes para planejar adequadamente o design

do jogo.

Entretanto, o presente estudo teve importantes contribuições ao mostrar o impacto do envolvimento de profissional de educação da área fim do jogo educacional em uma avaliação heurística. Houve esforço maior para treinamento com o uso do método e sobre aspectos de usabilidade e sobre como conduzir uma avaliação heurística. Porém, os resultados foram muito promissores, e mostraram que, com uma equipe interdisciplinar e apoio de profissionais da área de Interação Humano-Computador, foi possível envolver uma pessoa sem experiência prévia em avaliação de usabilidade de jogos educacionais e trazer melhorias significativas nos resultados. Diversos problemas relacionados ao uso e aspectos de aprendizado foram potencializados, e diversos problemas de usabilidade identificados por especialistas em Interação Humano-Computador tiveram a descrição pedagógica com significativo detalhamento com os pontos de vista de docente do curso de Medicina.

6. Conclusão e Trabalhos Futuros

No presente trabalho foi possível observar a importância de se fazer testes de usabilidade em jogos educacionais, já que esses games, conhecidos como jogos sérios, estão ganhando cada vez mais espaço entre alunos e professores no que diz respeito ao ensino. Este projeto teve como base o software sobre manejo de insulina, conhecido como In-suOnline, versão 1.8.4. Nele, puderam ser verificados requisitos básicos de usabilidade, como interação do jogo com o usuário, utilidade de se usar um game como metodologia de ensino/aprendizagem, efetividade do software sobre os seus objetivos iniciais, dentre outros aspectos.

Os testes com estudantes de Medicina foram muito importantes para a coleta de dados. Sem os testes com *think aloud*, muitas informações teriam sido perdidas. Embora este método seja trabalhoso, pelo fato de que os pesquisadores devem prestar muita atenção nas ações dos usuários e, ainda mais, nas análises dos dados coletados, este é um método que proporciona resultados consideráveis em estudos como este.

Um aspecto crucial observado neste projeto são os diferentes problemas encontrados em suas avaliações. A avaliação heurística interdisciplinar com a presença do avaliador da área médica fez com que o restante dos avaliadores prestasse mais atenção a problemas relacionados à didática do jogo. Com isso, houve mais relatos de problemas relacionados ao comportamento do médico no jogo, às falas dos personagens, ao assunto abordado e como ele foi passado. Já os especialistas sem experiência em educação médica tiveram outros focos. Os problemas encontrados por estes foram mais técnicos, como por exemplo, botões diferentes com as mesmas funcionalidades, ícones que não são sugestivos para indicar suas funções, plano de fundo que pode distrair os usuários de suas tarefas, dentre outras. No teste com os usuários foram encontrados problemas relacionados tanto com a parte técnica do software quanto à parte didática, como por exemplo, siglas da área médica que não foram entendidas por eles, não saber “voltar” a uma tela anterior, não encontrar informações para desenvolverem a tarefa, dentre outras.

Como trabalhos futuros, pode-se aplicar o mesmo teste abordado nesse trabalho a outros jogos sérios, e com outras combinações de equipes para avaliação heurística: com somente especialistas da área de Computação, equipe mista de especialistas da área de Computação e da área de conhecimento do jogo, e com usuários que utilizarão o software.

Referências

- Asadzandi, S., Sedghi, S., Bigdeli, S., and Sanjari, M. a. (2020). A systematized review on diabetes gamification. *Medical Journal of the Islamic Republic Of Iran*, 34(1). 2
- Bai, S., Hew, K. F., and Huang, B. (2020). Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review*, 30:100322. 1
- Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., and Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, 94:178–192. 1, 2
- Caroux, L., Isbister, K., Le Bigot, L., and Vibert, N. (2015). Player–video game interaction: A systematic review of current concepts. *Computers in Human Behavior*, page 366–381. 2
- Dekker, M. R. and Williams, A. D. (2017). The use of user-centered participatory design in serious games for anxiety and depression. *Games for Health Journal*, pages 327–333. 2
- Diehl, L. A., de Souza, R. M., Gordan, P. A., Esteves, R. Z., and Coelho, I. C. (2015a). User assessment of "insuonline," a game to fight clinical inertia in diabetes: A pilot study. *Games Health J.*, pages 335–343. 2
- Diehl, L. A., Gordan, P. A., Esteves, R. Z., and Coelho, I. C. (2015b). Effectiveness of a serious game for medical education on insulin therapy: a pilot study. pages 470–473. 9
- Diehl, L. A., Souza, R. M., Alves, J. B., Gordan, P. A., Esteves, R. Z., Jorge, M. L., and Coelho, I. C. (2013). Insuonline, a serious game to teach insulin therapy to primary care physicians: Design of the game and a randomized controlled trial for educational validation. *JMIR research protocols*. 2
- Diehl, L. A., Souza, R. M., Gordan, P. A., Esteves, R. Z., and Coelho, I. C. M. (2017). Insuonline, an electronic game for medical education on insulin therapy: A randomized controlled trial with primary care physicians. *J Med Internet Res*, 19(3):e72. 2, 9
- Ee, R. W. X., Yap, K. Z., and Yap, K. Y.-L. (2018). Herbopolis – a mobile serious game to educate players on herbal medicines. *Complementary Therapies in Medicine*, 39:68–79. 2
- Federoff, M. (2002). Heuristics and usability guidelines for the creation and evaluation of fun in video games. Master's thesis, Department of Telecommunications, Indiana University, Bloomington, Indiana, USA. 2
- Fialho, N. N. (2008). Os jogos pedagógicos como ferramentas de ensino. In *Congresso Nacional de Educação*, pages 12298–12306. 1
- Georgieva-Tsaneva, G. and Serbezova, I. (2020). Virtual reality and serious games using in distance learning in medicine in bulgaria. *International Journal of Emerging Technologies in Learning (iJET)*, 15(19):223–230. 2

- ISO9241-11 (1998). Iso 9241-11: Ergonomic requirements for office work with visual display terminals (vdts) — part 11: Guidance on usability. Acesso em: julho de 2021. 3
- Lemos, W., Junior, I., and Filho, A. C. (2020). Uma proposta de um serious game no auxílio do aprendizado da anatomia humana. pages 655,664. Anais do XXVIII Simpósio Brasileiro de Informática na Educação (SBIE 2017), Proceedings of the XX-VIII Brazilian Symposium on Computers in Education (SBIE 2017). 2
- Machado, L. S., Moraes, R. M., Nunes, F. L. S., and Costa, R. M. E. S. (2011). Serious games baseados em realidade virtual para educação médica. *Revista Brasileira de Educação Médica*, pages 254–262. 1
- Nielsen (2012). Usability 101: Introduction to usability. retrieved 20 june 2012. Disponível online em <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>, último acesso em julho 2021. 3
- Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann, San Francisco, CA. 2, 3, 4
- Nvidia. Shadowplay - recording and capture your greatest gaming moments. Disponível em <http://www.geforce.com/geforce-experience/shadowplay>, acesso em julho de 2021. 5
- Petrie, H. L.; Buykx, L. (2010). Collaborative heuristic evaluation: improving the effectiveness of heuristic evaluation. Proceedings of UPA 2010 International Conference. Omnipress. 5
- Pinelle, D., Wong, N., and Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1453–1462. 2, 3, 4
- Prates, R. O. Barbosa, S. D. J. (2003). Avaliação de interfaces de usuário - conceitos e métodos. In *Jornada de Atualização em Informática, SBC*, chapter 6. 5
- Renzi, A. B.; Freitas, S. (2010). Aplicação de think-aloud protocol em teste de usabilidade na procura de livros em livrarias online: recomendações de leitores. In *10º USIHC—Congresso Internacional de ergonomia e usabilidade*. 5
- Schmidt, J. D. E. and De Marchi, A. C. B. (2017a). Usability evaluation methods for mobile serious games applied to health: a systematic review. *Universal Access in the Information Society*, 16(4):921–928. 2
- Schmidt, J. D. E. and De Marchi, A. C. B. (2017b). Usability evaluation methods for mobile serious games applied to health: a systematic review. *Universal Access in the Information Society*, pages 921–928. 2
- Siregar, P. and Julen, N. (2016). Virtual patients and serious games in medicine. *Diagnostic Pathology*, 1(8). 2
- Yoshida, M. (2008). Think-aloud protocols and type of reading task: The issue of reactivity in l2 reading research. pages 199–209. In Selected Proceedings of the 2007 Second Language Research Forum, Somerville, MA, USA. 5