

UNIVERSIDADE DE PASSO FUNDO
INSTITUTO DE CIÊNCIAS EXATAS E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
COMPUTAÇÃO APLICADA

UMA METODOLOGIA PARA
AVALIAÇÃO DA USABILIDADE DE
INTERFACES 3D PARA IDOSOS

Manoela Rogofski Brum Postal

Passo Fundo

2018

UNIVERSIDADE DE PASSO FUNDO
INSTITUTO DE CIÊNCIAS EXATAS E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

**UMA METODOLOGIA PARA AVALIAÇÃO DA USABILIDADE DE
INTERFACES 3D PARA IDOSOS**

Manoela Rogofski Brum Postal

Dissertação apresentada como requisito parcial
à obtenção do grau de Mestre em Computação
Aplicada na Universidade de Passo Fundo.

Orientador: Rafael Rieder

Passo Fundo
2018

Dados Internacionais de Catalogação na Publicação (CIP)

P857m	Postal, Manoela Rogofski Brum Uma metodologia para avaliação da usabilidade de interfaces 3D para idosos / Manoela Rogofski Brum Postal. – 2018. 99 f. : il. color. ; 30 cm. Orientador: Rafael Rieder. Dissertação (Mestrado em Computação Aplicada) – Universidade de Passo Fundo, 2018. 1. Idosos. 2. Interface de programas aplicativos (Software). 3. Projeto de sistema centrado no usuário. 4. Interação homem-máquina I. Rieder, Rafael, orientador. II. Título. CDU: 004.4
-------	---

Catalogação: Bibliotecário Luís Diego Dias de S. da Silva

Número do Registro no CRB 10/2241

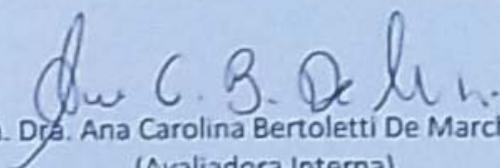
**ATA DE DEFESA DO
TRABALHO DE CONCLUSÃO DE CURSO DO ACADÊMICO**

MANOELA ROGOFSKI BRUM POSTAL

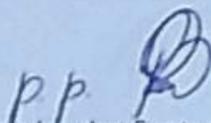
Aos dez dias do mês de janeiro do ano de dois mil e dezenove, às 9 horas, realizou-se, no Instituto de Ciências Exatas e Geociências, prédio BS, da Universidade de Passo Fundo, a sessão pública de defesa do Trabalho de Conclusão de Curso "Uma Metodologia para Avaliação da Usabilidade de Interfaces 3D para Idosos", de autoria de Manoela Rogofski Brum Postal, acadêmica do Curso de Mestrado em Computação Aplicada do Programa de Pós-Graduação em Computação Aplicada - PPGCA/UPF. Segundo as informações prestadas pelo Conselho de Pós-Graduação e constantes nos arquivos da Secretaria do PPGCA, a aluna preencheu os requisitos necessários para submeter seu trabalho à avaliação. A banca examinadora foi composta pelos doutores Rafael Rieder, Ana Carolina Bertoletti De Marchi e Fátima de Lourdes dos Santos Nunes Marques. Concluídos os trabalhos de apresentação e arguição, a banca examinadora considerou a candidata APROVADA. Foi concedido o prazo de até quarenta e cinco (45) dias, conforme Regimento do PPGCA, para a acadêmica apresentar ao Conselho de Pós-Graduação o trabalho em sua redação definitiva, a fim de que sejam feitos os encaminhamentos necessários à emissão do Diploma de Mestre em Computação Aplicada. Para constar, foi lavrada a presente ata, que vai assinada pelos membros da banca examinadora e pela Coordenação do PPGCA.



Prof. Dr. Rafael Rieder - UPF
Presidente da Banca Examinadora
(Orientador)



Prof. Dra. Ana Carolina Bertoletti De Marchi - UPF
(Avaliadora Interna)



Prof. Dra. Fátima de Lourdes dos Santos Nunes Marques - USP
(Avaliadora Externa)



Prof. Dr. Rafael Rieder
Coordenador do PPGCA

Verso do Termo de Apresentação (página em branco)

Dedico este trabalho, fruto carinhoso da dedicação, à minha família, ao meu marido e ao meu orientador, pois sem auxílio e apoio dos quais, ele não existiria.

AGRADECIMENTOS

Agradeço a Universidade de Passo Fundo pela oportunidade de conduzir esta pesquisa e me especializar na área que escolhi. Agradeço aos colegas de mestrado e do grupo de pesquisa de Sistemas Interativos e de Informação pelo auxílio nas diversas áreas exploradas no decorrer do desenvolvimento deste trabalho, bem como a todos os professores do Instituto de Ciências Exatas e Geociências que auxiliaram na pesquisa de acordo com suas áreas de especialização.

Agradeço ao CREATI – Centro de Referência e Atenção ao Idoso de Passo Fundo, RS, ao Solar Santa Tereza e aos demais participantes voluntários do experimento para validação da pesquisa conduzida. Tenho consciência de que cada colaboração individual constitui peça fundamental na completude deste trabalho.

Um agradecimento especial ao meu orientador que dispôs da atenção e auxílio necessários sempre que precisei. Aos meus pais, agradeço pelo constante incentivo e suporte às minhas conquistas acadêmicas, e proverem a mim tudo o que estava ao seu alcance. E agradeço ao meu marido, por ajudar e apoiar minha aspiração de tornar-me Mestre.

UMA METODOLOGIA PARA AVALIAÇÃO DA USABILIDADE DE INTERFACES 3D PARA IDOSOS

RESUMO

Profissionais de saúde podem utilizar interfaces 3D como ferramentas de apoio para a reabilitação de idosos, oferecendo a estes recursos divertidos e benéficos para a prática de atividades físicas e cognitivas. Nesse contexto, é preciso estabelecer mecanismos de avaliação destas interfaces, para alcançar o equilíbrio entre funcionalidade, facilidade de uso e bem-estar do indivíduo. O objetivo deste trabalho é propor um método para avaliação da usabilidade de interfaces 3D para idosos. Para tanto, desenvolveu-se uma revisão sistemática de literatura com objetivo de identificar quais técnicas e instrumentos são utilizados em intervenções com idosos para avaliar sistemas de realidade virtual. A partir dos resultados da revisão e de experiência prévia, desenvolveu-se um método e um questionário, denominado UQE 3D, envolvendo conceitos inerentes a interfaces 3D. O método envolve o uso de um questionário sociodemográfico, MEEM e GDS-15 para caracterização e seleção da amostra, e para avaliação da interface utiliza-se o SUS e o UQE 3D. Para validação, realizou-se um experimento envolvendo um público heterogêneo de 30 idosos (60+ anos), sendo sete destes idosos institucionalizados. Os resultados apontaram para um panorama positivo, obtendo como média geral amostral para o UQE 3D um escore de 82,60 (num intervalo de 1 a 100), sendo que para o grupo de idosos ativos o escore foi 88,52 e, 63,17 para os idosos institucionalizados. A comparação entre os escores do UQE 3D e do SUS não apresentaram diferença estatística significativa, sugerindo que o UQE 3D é equivalente ao SUS ao analisar a usabilidade de uma interface 3D.

Palavras-chave: Avaliação, Idosos, Interfaces 3D, Usabilidade.

A USABILITY EVALUATION METHODOLOGY OF 3D USER INTERFACES FOR THE ELDERLY

ABSTRACT

Health professionals could use three-dimensional user interfaces as a tool to support elderly's rehabilitation, allowing them entertaining and benefits for the practice of physical and cognitive activities. In this context, it is necessary to establish mechanisms to evaluate these interfaces, to get a balance between functionality, easiness of use and welfare of the user. Our approach aims to propose a usability evaluation method of three-dimensional user interfaces for the elderly. Therefore, we developed a systematic review to identify the techniques and tools used in interventions with the elderly to evaluate virtual reality systems. From the review's results and our previous experience, a method and a questionnaire denominated UQE 3D was developed, wrapping concepts inherent to three-dimensional user interfaces. Our method uses a demographic questionnaire, MEEM and GDS-15 for description and sample selection, and SUS, UQE 3D and interview for the interface's evaluation. For validation, we executed an experiment considering a heterogeneous group with 30 subjects (60+ years), where seven subjects were institutionalized elderly. Results pointed a positive scenario, getting an overall sample mean for the UQE 3D a score of 82.60 (in a range of 1 to 100), where the score for the active elderly was 88.52 and for the institutionalized elderly it was 63.17. The difference between the UQE 3D and the SUS scores did not represent statistically significant, suggesting that the UQE 3D is equivalent to the SUS regarding the evaluation of three-dimensional user interfaces.

Keywords: Elderly, Evaluation, Three-dimensional user interface, Usability.

LISTA DE FIGURAS

Figura 1 - Diagrama do processo de Revisão Sistemática	27
Figura 2 - Idoso utilizando o ambiente Virtual Studio (à esquerda) e ambiente onde eram realizadas as gravações (à direita) [16].	28
Figura 3 – Imagem da tela da aplicação [17].	30
Figura 4 – Imagem do cenário virtual gerado (à esquerda) e exemplo de tarefa de interação (à direita) [18].	30
Figura 5 – Idosos realizando as tarefas de interação [13].	36
Figura 6- Ambientes disponíveis no Gymcentral [14].	37
Figura 7 – Headset utilizado para registrar os sinais cerebrais e as expressões faciais dos usuários [38].	41
Figura 8 – CS, Jogo virtual utilizado para realizar os testes [38].	42
Figura 9 – Telas do aplicativo Medication Assistant [39].	43
Figura 10 – Telas do aplicativo Medication Assistant resultante do processo de avaliação [40].	44
Figura 11 – Imagens exemplo da interface do sistema [41].	45
Figura 12 – Idosos participando da sessão de demonstração do sistema [41].	46
Figura 13 – Relação entre os princípios e características definidas pelos autores [3] [4] [48] [57] com as perguntas do UQE 3D.	60
Figura 14 – Relação das perguntas do SUS com os princípios e características definidas pelos autores [3] [4] [48] [57].	60
Figura 15 - Visão da interface do jogo Motion Rehab AVE 3D em primeira pessoa.	64
Figura 16 – Exemplos de movimentos para interagir com a interface [60].	65
Figura 17 – Idosos interagindo com o sistema no experimento de avaliação.	65
Figura 18 – Dispositivos utilizados para visualização e interação no ambiente virtual.	66
Figura 19 – Esquema explicativo da lógica para cálculo do SUS, replicada no UQE 3D.	69

LISTA DE TABELAS

Tabela 1 – Pontos fracos elencados pelos autores [7] [17] e suas respectivas soluções implementadas no método.	22
Tabela 2 – Primeira versão do questionário desenvolvido.	23
Tabela 3 – Principais itens elencados pelos profissionais para serem alterados no questionário.	24
Tabela 4 – Tabela comparativa dos instrumentos utilizados para avaliação em cada estudo. .	50
Tabela 5 – Resumo dos aspectos e abordagens de avaliação utilizados pelos estudos selecionados.	51
Tabela 6 – Estrutura do método proposto	53
Tabela 7 – Perguntas do UQE 3D e suas respectivas relações com os princípios e características definidas pelos autores [3] [4] [48] [57]	59
Tabela 8 – Perguntas da Entrevista Semiestruturada e suas relativas questões do UQE 3D ...	62
Tabela 9 – Valores de referência para consistência interna do questionário. [64]	67
Tabela 10 – Simulação do melhor caso, utilizada para definir o score do questionário UQE 3D.	70
Tabela 11 - Dados da análise estatística para definição do Intervalo de Confiança.	71
Tabela 12 – Estatísticas de grupo para o Teste T de Student.	72
Tabela 13 - Resumo das respostas afirmativas para as questões da entrevista.	73
Tabela 14 – Distribuição e frequência das respostas assinaladas no UQE 3D para o Grupo A.	73
Tabela 15 - Distribuição e frequência das respostas assinaladas no UQE 3D para o Grupo B.	74

LISTA DE ABREVIATURAS

2D – duas dimensões, bidimensional

3D – três dimensões, tridimensional

3DUI – Three-dimensional user interface

CEC-D10 – Escala de Depressão dos Estudos Epidemiológicos

CSQ-8 – Questionário de Satisfação do Cliente-8

DK – Development Kit

EEG – Eletroencefalograma

FFA – FitForAll

GDS-15 – Escada de Depressão Geriátrica

HMDs – Head-Mounted Displays

ICF-US I – International Classification of Functioning based Usability Scale

MEEM – Mini Exame do Estado Mental

OMS – Organização Mundial da Saúde

RS – Revisão Sistemática de Literatura

RV – Realidade Virtual

SUS – System Usability Scale

TCLE – Termo de consentimento Livre e Esclarecido

TV – Televisão

UQE 3D – 3DUI Usability Questionnaire for Elderly

WHO – World Health Organization

SUMÁRIO

1.	INTRODUÇÃO.....	17
2.	FUNDAMENTAÇÃO METODOLÓGICA.....	21
3.	REVISÃO SISTEMÁTICA DE LITERATURA	25
3.1.	MÉTODO	25
3.2.	ESTUDOS SELECIONADOS	27
3.2.1.	Orso <i>et al.</i>	27
3.2.2.	Hagiya, Horiuchi e Yazaki	28
3.2.3.	Tost <i>et al.</i>	30
3.2.4.	Himmelsbach <i>et al.</i>	31
3.2.5.	Sacramento, Bacellar e Ferreira	31
3.2.6.	Le, Thompson e Demiris	32
3.2.7.	Al-khomsan <i>et al.</i>	34
3.2.8.	Konstantinidis <i>et al.</i>	35
3.2.9.	Hsieh e Huang	35
3.2.10.	Báez <i>et al.</i>	37
3.2.11.	Schuermans <i>et al.</i>	39
3.2.12.	Chiu <i>et al.</i>	39
3.2.13.	Meza-Cubo <i>et al.</i>	41
3.2.14.	Ferreira <i>et al.</i>	42
3.2.15.	Teixeira <i>et al.</i>	43
3.2.16.	Epelde <i>et al.</i>	44
3.2.17.	Discussão dos Estudos Seleccionados	46
4.	MATERIAIS E MÉTODOS.....	53
4.1.	PRÉ-TESTE.....	54
4.2.	TESTE.....	56
4.3.	PÓS-TESTE.....	57
5.	RESULTADOS E DISCUSSÕES.....	63
5.1.	USO DO MÉTODO	63
5.1.1.	Participantes.....	63
5.1.2.	Tarefa	64
5.1.3.	Dispositivos.....	66
5.2.	AVALIAÇÃO DO MÉTODO - VALIDAÇÃO E DISCUSSÃO.....	66
5.2.1.	Alfa de Crombach, KMO e Bartlet.....	67

5.2.2.	Intervalo de Confiança.....	68
5.2.3.	Teste T de Student e Teste F de Fisher	71
5.2.4.	Distribuição e Frequência	72
6.	CONCLUSÃO.....	79
	REFERÊNCIAS.....	83
	APÊNDICE A - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	89
	APÊNDICE B - QUESTIONÁRIO SOCIODEMOGRÁFICO E DE CARACTERIZAÇÃO DA AMOSTRA	91
	APÊNDICE C - QUESTIONÁRIO UQE 3D	93
	ANEXO A - MINI EXAME DO ESTADO MENTAL – MEEM	95
	ANEXO B - ESCALA DE DEPRESSÃO GERIÁTRICA (GDS-15) – VERSÃO REDUZIDA	97
	ANEXO C - QUESTIONÁRIO SUS	99

1. INTRODUÇÃO

A usabilidade se caracteriza como a variedade e o grau com que as funcionalidades de um sistema podem ser utilizadas eficientemente, de forma intuitiva pelo usuário [1]. Os sistemas interativos somente podem ser considerados úteis, práticos e eficazes se possuírem harmonia entre funcionalidade e usabilidade. Portanto, é importante avaliar a usabilidade de um sistema para que este possa melhor atender seus objetivos.

De acordo com Cockton [2], avaliar a usabilidade é fundamental para que se estabeleça uma relação entre a qualidade de um sistema interativo e a qualidade de interação. O autor menciona que, quando a avaliação de usabilidade aponta a utilidade de um sistema ou dispositivo, os métodos e as métricas contribuem para determinar a extensão de usabilidade, mensurando a robustez, os objetivos e a confiabilidade. Para isso é necessário a utilização de um método ou protocolo que englobe e possa avaliar de forma confiável todas essas questões.

Segundo Nielsen [3], as características principais da usabilidade que devem ser avaliadas são a facilidade e a eficiência em se realizar tarefas básicas, a facilidade em reutilizar recursos, o reestabelecimento de serviços aos erros cometidos, e a satisfação vivenciada pelo participante através do uso do sistema.

Interfaces 3D, como sistemas de Realidade Virtual (RV), vêm se popularizando na área de games, e necessitam de avaliação de usabilidade do seu processo interativo. Vários conceitos e objetivos em avaliação de usabilidade de interface tradicionais e tridimensionais são similares, mas as abordagens de aplicação utilizadas no contexto de interfaces 3D geralmente são diferentes. Avaliações para interfaces tradicionais podem utilizar variadas abordagens, desde estudos informais com usuários até experimentos formais, baseados em tarefas e avaliações heurísticas. Com relação às interfaces 3D, em sua maior parte, é difícil avaliar a sua usabilidade sem o auxílio de usuários reais interagindo com tarefas representativas [4].

Serious games para idosos com interfaces 3D são exemplos de aplicações que vêm sendo utilizados em intervenções clínicas de reabilitação [5]. Um *serious game* segue os mesmos princípios de entretenimento de jogos digitais interativos, porém seu objetivo é transmitir também um conteúdo educacional ou de treinamento ao usuário. No caso de idosos, eles podem ajudar no estímulo à prática de atividades benéficas ao corpo humano e aumentar o interesse do paciente por tratamentos e processos de reabilitação, que muitas vezes são lentos e dolorosos [6].

Esforços têm sido aplicados para projetar e desenvolver jogos destinados ao público idoso, pois estas ferramentas têm se mostrado eficientes para melhorar o tempo de reação, a percepção visual, as habilidades cognitivas, a autoconfiança e, conseqüentemente, a qualidade de vida e o bem-estar dessa categoria de usuário [7]. Fua et al. [8] relatou melhorias na memória e atenção dos idosos com o uso de um *serious game* desenvolvido com este objetivo. Além disso, o autor menciona que jogos de computador são capazes de melhorar o tempo de reação e habilidades motoras, como coordenação e destreza, e que essas melhoras podem ter impacto direto no cotidiano dos idosos.

Aplicações para idosos necessitam, como qualquer outro sistema, de um método de avaliação que ofereça a possibilidade de avaliar a qualidade de interação, como meio de desmistificar a falta de acesso, prática ou medo por parte desse público [9]. Além disso, é importante que a interface seja avaliada quanto à sua usabilidade para assegurar que a solução seja adequada ao perfil do idoso. Problemas de usabilidade estão relacionados a aspectos de uma interface que podem causar problemas na aprendizagem de uso, no próprio uso eficiente desta interface ou ainda no grau de satisfação do usuário.

Uma pesquisa nesse contexto se justifica, primeiramente, pelo aumento no número de idosos que procuram jogos digitais e ambientes virtuais 3D para entretenimento, reabilitação ou treinamento – independentemente de serem *serious games* ou serem jogos destinados para sua faixa etária. De acordo com Hall e Marston [10], devido ao aumento da expectativa de vida, jogos para o público idoso têm se apresentado como valiosas ferramentas de promoção da saúde e educação para a saúde. Além disso, promovem a interação social com indivíduos da mesma faixa etária, e permitem uma conexão mais próxima com o público jovem [11].

Quando se trabalha com um público específico como idosos, é importante utilizar ferramentas que se enquadrem na classificação do contexto específico [12], e pensadas para contribuir em qualidade a esta faixa etária. Neste caso, desenvolver um método para o contexto específico da população idosa pode garantir que a mesma seja capaz de identificar as necessidades de acessibilidade e percepção deste público, sendo capaz de certificar se uma interface 3D, por exemplo, esteja ou não adequada ao perfil do idoso. Um método apropriado a este perfil pode ajudar no desenvolvimento de sistemas de RV melhores para um público que, gradativamente, se aproxima da tecnologia através do entretenimento com seus descendentes, de *serious games* para reabilitação [13] e de aplicações interativas que estimulam a prática de exercícios físicos [14] [15]. Diante desse cenário, a avaliação da usabilidade de interfaces 3D é fundamental para garantir uma boa experiência ao público

idoso, minimizando efeitos inerentes às limitações provenientes do avanço de idade. Nesse sentido, este trabalho realizou uma revisão sistemática de literatura, que mostrou que não existe uma tendência no que diz respeito à avaliação da usabilidade de interfaces realizadas em experimentos com usuários idosos. Não foram encontrados métodos, ferramentas ou técnicas que considerem a classificação de contexto específico salientada por Günter [12] e LaViola *et al.* [4]. Os trabalhos elencados também não mencionaram adaptações de instrumentos que utilizam o contexto específico do público alvo. Quanto às interfaces 3D, também não foram identificadas técnicas de avaliação projetadas especificamente para o público idoso, bem como algumas abordagens não definem a dimensionalidade do sistema aplicado em seus experimentos.

Em vista disso, este trabalho propõe um método para avaliação da usabilidade de interfaces 3D específicas para idosos. O método utiliza técnicas de avaliação reconhecidas pela bibliografia e encontradas como mais utilizadas em uma Revisão Sistemática que buscou instrumentos comumente utilizados em intervenções com idosos. O objetivo do método é avaliar interfaces 3D desenvolvidas para o público idoso, com o auxílio dos próprios idosos como sujeitos nos experimentos de avaliação. Para tanto, criou-se também novo instrumento de avaliação para integrar o método, que atenta para a acessibilidade por parte do público específico, capaz de identificar as opiniões e as sensações dos sujeitos a respeito da aplicação considerando sua linguagem.

Dessa forma, esse documento está assim organizado: uma fundamentação metodológica para demonstrar o caminho percorrido para o desenvolvimento do método está descrita no Capítulo 2; uma revisão sistemática de literatura é descrita pelo Capítulo 3, mostrando os trabalhos relacionados selecionados; o método desenvolvido e o questionário de contexto específico criado encontram-se no Capítulo 4; os resultados e discussões do experimento de validação do método proposta são apresentados no Capítulo 5; e, por fim, o Capítulo 6 apresenta as conclusões e os trabalhos futuros.

2. FUNDAMENTAÇÃO METODOLÓGICA

Diante da necessidade de se avaliar a usabilidade de interfaces 3D destinadas ao público idoso, buscou-se identificar, primeiramente, qual método tem sido utilizado na literatura para avaliar interfaces em intervenções com o público idoso. De acordo com LaViola et al. [4], a maneira mais fácil de avaliar a usabilidade de uma interface 3D é com o auxílio de usuários reais interagindo com tarefas representativas. Para tanto, desenvolveu-se uma Revisão Sistemática de Literatura que será descrita em detalhes no Capítulo 3.

É importante salientar que a revisão foi pensada de uma forma ampla, não envolvendo apenas experimentos com interfaces 3D, mas considerando qualquer tipo de dispositivo computacional. Tomou-se como objetivo identificar quais métodos, ferramentas ou técnicas são utilizadas para intervenções com idosos, envolvendo avaliação de dispositivos computacionais para, a partir daí, elaborar uma forma de avaliar a usabilidade de interfaces 3D específica para idosos.

Além da revisão, foi utilizado também como fundamentação teórica os conceitos de Realidade Virtual elencados por Burdea e Coiffet [16], os conceitos de avaliação de interfaces 3D catalogados por LaViola et al. [4], a fundamentação para elaboração de questionários citados por Günter [12] e as variáveis de usabilidade de Nielsen [3] [17].

O método proposto por este trabalho (Capítulo 4) foi desenvolvido pensando na avaliação de interfaces 3D desenvolvidas para idosos, para serem aplicadas em experimentos com usuário idosos, propostas e testadas em experimentos anteriores por Simor [7] e Brum e Rieder [18]. Ambos os autores descreveram seus métodos divididos em três etapas: Pré-teste, Teste e Pós-teste.

A etapa de Pré-Teste contempla aplicação de instrumentos para seleção e classificação dos participantes, mais especificamente os questionários MEEM, GDS-15 e um questionário sociodemográfico e de caracterização da amostra. Na primeira etapa também é explicado ao participante como sucederá o experimento e entregue o Termo de Consentimento Livre e Esclarecido para ser assinado.

Na etapa de Teste, os participantes interagem com a/as tarefa/as de interação destinadas ao experimento, e avaliam a interface 3D posteriormente.

Para avaliação na etapa de Pós-Teste, são utilizados os questionários SUS e um desenvolvido pelos autores. Os principais apontamentos dos autores em relação aos seus métodos e aos experimentos desenvolvidos são referentes ao tempo dos procedimentos

(tempo de aplicação do método, tempo de espera para realizar o teste, tempo utilizado para a interação com o jogo), pois tempos muito longos causam o desinteresse dos participantes; referente ao número de questionários envolvidos; à falta de um momento para treinamento, em que os participantes poderão perder o medo de utilizar a tecnologia e sanar dúvidas básicas; e à dificuldade dos idosos em entenderem termos referentes à tecnologia, por tratar-se de uma área que, na maior parte dos casos, não faz parte do seu cotidiano.

Atentando para os pontos elencados pelos autores [7] [18] e pela revisão sistemática, desenvolveu-se o método em proposição, tentando resolver os pontos fracos elencados na Tabela 1.

Tabela 1 – Pontos fracos elencados pelos autores [7] [18] e suas respectivas soluções implementadas no método.

PONTOS FRACOS A SEREM RESOLVIDOS	SOLUÇÃO IMPLEMENTADA NO MÉTODO
Tempo de espera para participar do teste	Marcar dia e horário para cada sessão de teste com cada participante
Tempo de aplicação do método	Dividir o método em dois dias para aplicação e organizá-la de forma a minimizar o tempo de aplicação de cada etapa
Tempo para interação com a interface	Minimizar o tempo de interação para que seja suficiente, mas não exaustivo
Número de questionários envolvidos	Por ora, manteve-se o número de questionários para avaliar em detalhes a interface
Falta de um momento de treinamento	Criado um momento para interação com a interface para familiarização com a tecnologia
Dificuldade dos idosos em relação aos termos específicos	Maior detalhamento em cada pergunta e resposta e 5evisão dos questionários de avaliação por especialistas, inserindo a técnica de entrevista na etapa Pós-teste

Em relação ao questionário desenvolvido para integrar o método, com perguntas referentes à usabilidade de interfaces 3D utilizando dispositivos de interação não convencionais, pode-se observar na Tabela 2 a primeira versão do questionário. As perguntas foram detalhadas para facilitar o entendimento do público alvo, assim como as respostas. Este questionário foi encaminhado para dois profissionais da área da saúde que trabalham com idosos e dois profissionais da área da computação. O objetivo foi elencar pontos a serem aperfeiçoados no questionário para que ele pudesse alcançar a melhor compreensão do público alvo, e contemplar todas as questões de avaliação.

Já a Tabela 3 resume os principais apontamentos da revisão executada pelos profissionais de acordo com suas áreas de atuação.

Tabela 2 – Primeira versão do questionário desenvolvido.

Nº	Afirmações	Marque suas respostas abaixo!				
		1	2	3	4	5
1	Senti-me confortável durante a interação no jogo, foi agradável	Fiquei muito desconfortável, algo realmente me incomodou	Fiquei desconfortável, algo me incomodou um pouquinho	Nem confortável, Nem desconfortável	Fiquei confortável	Fiquei muito confortável
2	Senti-me dentro do jogo, quase esquecendo que estava no mundo real	Não me senti dentro do jogo	Quase me senti dentro do jogo	Senti-me dentro do jogo, mas sabia que estava no mundo real	Quase esqueci do mundo real, mas não esqueci	Por alguns momentos realmente esqueci
3	Senti-me bem, a interação com o game e com os equipamentos não provocou situações de mal-estar, como enjoo, dor de cabeça, tontura ou náusea	Tive bastante mal-estar	Tive um leve desconforto	Não tive reações negativas nem positivas	Me senti bem	Extremamente bem
4	Senti-me orientado geograficamente dentro do ambiente 3D	Não! Não sabia mais o que era esquerda e direita!	Não, foi relativamente e difícil me localizar lá dentro do jogo	Mais ou menos	Sim, foi relativamente fácil	Sim, parecia que eu vivia lá e sabia em que lado estaria qualquer coisa!
5	Consegui realizar naturalmente as tarefas no game	Não! Não sabia o que fazer com meus braços!	Não, foi um pouco artificial o que eu fazia	Mais ou menos	Sim, foi relativamente natural	Sim, foi como caminhar ou comer! Totalmente automático!
6	Gostei da interação por gestos	Nem um pouco	Não muito	Tanto faz	É interessante	Gostei muito
7	Achei a interação por gestos cansativa	Muito cansativo	Cansativo	Relativo	Levemente cansativo	Nem um pouco cansativo
8	O capacete de visualização interferia na forma como encarei o jogo	Interferiu de forma muito negativa	Interferiu de forma um pouco negativa	Não interferiu	Interferiu de forma um pouco positiva	Interferiu de forma muito positiva
A imagem do jogo permitiu ou me ajudou a...						
9	Utilizar e interagir facilmente com o jogo	Extremamente difícil	Relativamente difícil	Mais ou menos	Relativamente fácil	Extremamente fácil
10	Saber facilmente o que fazer para executar as tarefas	Não	Dificultou um pouco	A imagem do jogo não interferiu	Facilitou	Absolutamente sim
11	Visualizar, interpretar e entender facilmente os elementos interativos do jogo (aspectos visuais)	Dificultou muito, estava confuso	Dificultou um pouco	A imagem do jogo não interferiu	Facilitou	Ajudou muito!
12	Ouvir e assimilar facilmente os elementos sonoros do jogo	Dificultou muito, estava confuso	Dificultou um pouco	A imagem do jogo não interferiu	Facilitou	Totalmente
Outras questões						
13	Foi divertido jogar o jogo	Não, foi chato!	Não foi legal	Nem bom, nem ruim	Foi legal	Muito divertido!
14	A temática do jogo está associada à minha idade	Totalmente fora de conceito!	Não muito	Tanto faz	Um pouco	Muito associada! O jogo me soou familiar!

Tabela 3 – Principais itens elencados pelos profissionais para serem alterados no questionário.

Questão Referente	PROFISSIONAIS DA ÁREA DA COMPUTAÇÃO
Todas	A forma descritiva de resposta pode confundir os usuários
Todas	Cuidar o sentido das perguntas (negativa x positiva)
Algumas	Padronizar o uso de expressões como jogo e <i>game</i>
1	Avaliar apenas um aspecto (agradável ou confortável)
2	Trocar a expressão “mundo real”
6	Especificar o tipo de interação
8	Trocar o termo “encarei”
9-12	Reforçar que as questões se relacionam à imagem
11	Trocar a expressão “elementos interativos do jogo”
12	Trocar expressão “elementos sonoros”
12	Focar a questão em “ouvir” ou “assimilar”
14	Trocar termo “soou”
Questão Referente	PROFISSIONAIS DA ÁREA DA SAÚDE
Todas	Alterar a forma descritiva de respostas para um padrão de escala
	Acrescentar a questão: você considerou fácil se localizar no jogo? (as respostas mostrarão o grau de facilidade ou dificuldade)
	Adicionar questões referentes a tempos de descanso e da tarefa
4	Trocar expressão “orientado geograficamente”
6	Realocar a questão para ficar mais próxima da questão sobre diversão
8	Dar um sentido (positivo ou negativo) para a questão
9	Avaliar apenas um aspecto
9, 10 e 11	Resumir em duas questões
11	Avalia muitos aspectos no mesmo item
12	Não está relacionada à imagem

Buscando atender aos apontamentos dos profissionais que revisaram o questionário, foram realizadas as alterações propostas e a versão final pode ser vista no Apêndice C - Questionário UQE 3D.

3. REVISÃO SISTEMÁTICA DE LITERATURA

Para desenvolver um método a ser aplicado com idosos, foi necessário identificar primeiramente se já existia algum desenvolvimento neste sentido. E, caso não existisse, seria interessante verificar ferramentas e técnicas que estão sendo utilizadas para intervenções com este público em avaliações envolvendo interfaces 3D. Para tanto, desenvolveu-se uma Revisão Sistemática de Literatura descrita nas próximas subseções.

3.1. MÉTODO

Como estratégia metodológica para verificar as contribuições para esta área de pesquisa, procurou-se fundamentar esse processo a partir das concepções sobre Revisão Sistemática de Literatura (RS). De acordo com Dyba [19], uma RS é uma revisão planejada para responder a uma pergunta específica. Este processo utiliza métodos explícitos e sistemáticos para identificar, selecionar e avaliar criticamente os estudos, e para coletar e analisar os dados dos estudos incluídos na revisão.

Desta forma, inicialmente elaborou-se uma questão de pesquisa para apoiar o processo de planejamento da RS: Quais são os instrumentos de avaliação de interfaces que estão sendo utilizados em intervenções com idosos?

Para responder a esta questão, foi estabelecida uma estratégia de busca utilizando a seguinte expressão: (evaluation OR validation) AND (instrument OR tool OR technique OR method) AND (Virtual Reality OR software OR application) AND (Elderly OR Older OR Senior).

Neste processo, foram elencadas quatro bases digitais que tradicionalmente concentram periódicos e trabalhos científicos na área de Ciência da Computação: ScienceDirect, ACM DL Digital Library, IEEE Xplore Digital Library e Springer Link. As buscas foram filtradas por data (intervalo 2012-2017) e, sempre que possível, utilizando metadados (por exemplo, argumentos “tak” e “near” nas ferramentas de busca das bases Science Direct e Springer Link). O ano 2012 foi definido como período de publicação inicial levando em consideração que as tecnologias computacionais não eram amplamente difundidas ao público idoso antes desse período. O termo “usabilidade” não foi incluído na estratégia de busca objetivando não afunilar a pesquisa apenas para este viés, uma vez que a ideia é

encontrar métodos e técnicas utilizados para avaliar interfaces com idosos, e não necessariamente apenas avaliar a usabilidade delas.

De acordo com a estratégia de busca, definiram-se os critérios de inclusão e exclusão dos trabalhos. Para inclusão foram definidos os seguintes critérios:

- Artigos que apresentaram alguma interface de software;
- Artigos que apresentaram alguma avaliação da interface;
- Artigos que realizaram a avaliação com idosos.

Para exclusão foram elaborados os seguintes critérios:

- Artigos que não apresentaram algum processo de avaliação da interface;
- Artigos que não utilizaram idosos para o processo de avaliação;
- Artigos que já foram encontrados e analisados em fontes diferentes.

Efetuada a busca nas fontes citadas, obteve-se como retorno 252 resultados, sendo 106 trabalhos oriundos da base de pesquisa IEEE Xplore Digital Library, 88 da ACM DL Digital Library, 7 da Springer Link e 51 trabalhos da Science Direct. Esta primeira fase de busca dos estudos foi realizada entre os dias 13 de março de 2017 a 10 de abril de 2017.

A partir de então, foi realizada uma primeira triagem com o objetivo de selecionar os artigos que trouxessem, de fato, um experimento em que houvesse a avaliação de alguma interface com o público idoso. Para tanto, foram analisados os títulos das obras em função dos critérios de pesquisa, culminando num montante de 45 artigos supostamente relacionados.

A segunda triagem se deu com a leitura dos resumos dos artigos para selecionar abordagens que realizam alguma intervenção com idosos para avaliação de alguma interface. Com isso, foram excluídos 17 artigos, os quais não apresentavam tal característica, permanecendo somente 28 trabalhos para uma análise aprofundada.

Na terceira triagem houve a leitura completa dos materiais, momento no qual se observou que ~57,14% dos artigos selecionados na segunda triagem apresentavam avaliações de interfaces com idosos. Portanto, a presente pesquisa baseia-se nas informações oriundas de 16 artigos, selecionados no processo de RS. O processo descrito pode ser visualizado na Figura 1.

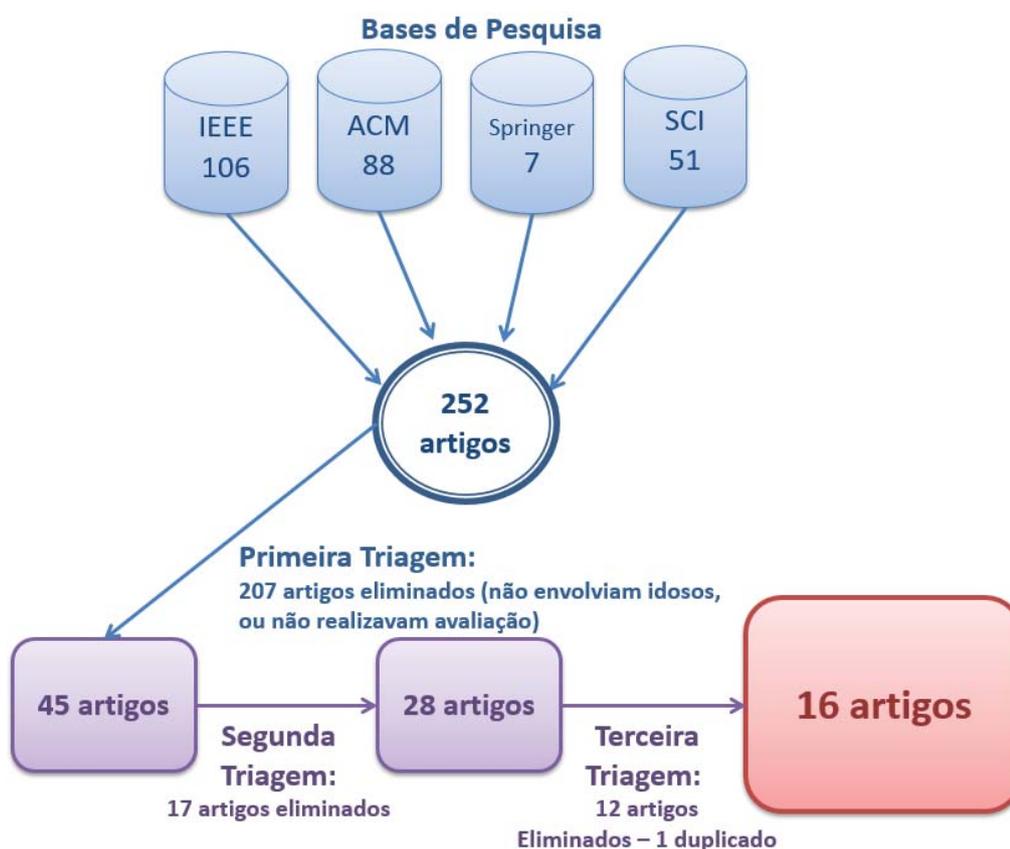


Figura 1 - Diagrama do processo de Revisão Sistemática

3.2. ESTUDOS SELECIONADOS

As próximas seções apresentam os 16 estudos selecionados após o processo de revisão sistemática – os quais são trabalhos relacionados com a abordagem proposta.

3.2.1. Orso *et al.*

No trabalho de Orso *et al.* [20], os autores projetaram e criaram uma solução interativa para uma Smart TV destinada ao público idoso. Para isso desenvolveram dois

softwares, Virtual Studio e o SeniorChannel. O primeiro é um programa de edição de vídeo que permite aos idosos desenvolver conteúdos para serem disponibilizados no SeniorChannel, para serem acessados por outros idosos.

Um grupo composto por seis idosos, com idades entre 60 e 82 anos, realizaram um conjunto de tarefas que envolvia a criação de um videoclipe com o Virtual Studio e uso da ferramenta SeniorChannel, como pode ser visto na Figura 2. Na sequência, um questionário com escala Likert de seis níveis era preenchido.

Os resultados mostraram que os idosos foram capazes de operar com sucesso e autonomia um programa de software dedicado para o gerenciamento da gravação e edição de um programa de TV após um treinamento relativamente curto. Além disso, expressaram opiniões positivas quanto à compreensão das diferentes características da interface. Além disso, os idosos estavam positivamente envolvidos com o conteúdo e dispostos a participar ativamente na produção do conteúdo de TV. Os usuários também elogiaram o papel que um canal de TV digital interativa pode ter na vida da comunidade local.



Figura 2 - Idoso utilizando o ambiente Virtual Studio (à esquerda) e ambiente onde eram realizadas as gravações (à direita) [20].

3.2.2. Hagiya, Horiuchi e Yazaki

No trabalho de Hagiya, Horiuchi e Yazaki [21], os autores criaram um tutor de digitação para smartphone dirigido ao público idoso que é capaz de detectar dificuldades de digitação e prover instruções ao usuário.

Foram realizadas duas avaliações para cada uma das duas versões do aplicativo. Para a primeira avaliação da primeira versão foram selecionados seis participantes (três homens e três mulheres) com idades entre 65 e 72 anos e para a segunda avaliação da primeira versão foram selecionados o mesmo número de participantes com idades entre 66 e 74 anos. O procedimento de avaliação foi o mesmo para ambos os grupos: foi mostrado aos participantes um tutorial de como utilizar o smartphone e o teclado, para que escrevessem 10 frases utilizando um aplicativo de e-mail. O período de interação com o smartphone foi limitado a 60 minutos. Após, eles respondiam a um questionário utilizando escala Likert de cinco níveis e participavam de uma entrevista.

Os resultados das avaliações da primeira versão do aplicativo mostraram que em relação ao questionário houve uma resposta positiva ao item "A instrução é útil para melhorar minha habilidade". O método de instrução também foi avaliado positivamente, com a instrução por voz sendo a mais favorecida. As entrevistas mostraram que para operações que podem ser entendidas apenas a partir da posição e do rótulo de um objeto, instruções detalhadas podem ser desnecessárias, entretanto, para operações mais complexas, repetir instruções ocasionalmente é um comportamento útil de um software para esse público.

Para a avaliação da segunda versão do software houve a participação de 20 usuários (onze homens e nove mulheres) com idades entre 65 e 74 anos. Metade do grupo recebeu instruções a respeito do Tutor de Digitação (grupo A) e a outra metade não recebeu. Os usuários ficaram em teste por duas semanas. Uma vez por semana conversavam com um operador e ao final das duas semanas passaram por processos de entrevistas. Todos os participantes enviaram mais de seis mensagens por dia e foram avaliados de acordo com o nível de habilidade estimado, número de dificuldades de entrada, velocidade de digitação e precisão de digitação.

O nível de habilidade estimado para o grupo A no último dia foi melhor do que no primeiro, o número de erros para o grupo A foi menor do que para o grupo B desde o princípio, mas ambos os grupos apresentaram melhora significativa no número de erros ao final das duas semanas; os testes foram realizados utilizando o teste U de Mann-Whitney ($\alpha \leq 0.05$) e mostraram diferença significativa do terceiro ao décimo dia de uso do sistema. A velocidade de digitação do grupo A foi maior do que do grupo B, mas ambos tiveram um aumento na velocidade no decorrer das duas semanas de teste; o teste U de Mann-Whitney ($\alpha \leq 0.05$) mostra uma diferença significativa para todos os dias após o segundo. Para a precisão de digitação não houve diferenças significativas entre os dois grupos. A Figura 3 mostra a interface final da aplicação.

3.2.3. Tost *et al.*

Tost *et al.* [22] descrevem em seu trabalho o SmartAgeing, um jogo sério 3D voltado para a detecção precoce de deficiências cognitivas leves em pessoas com idades entre 50 e 80 anos, para avaliação de deficiências cognitivas em pessoas já diagnosticadas com demência neurodegenerativa. O jogo possui cinco tarefas diárias num ambiente familiar abordando competências cognitivas como memória, funções executivas, atenção dividida, memória de longo prazo e orientação espacial e atenção. O jogo registra todas as ações dos usuários e, a partir desses dados, compõe um conjunto de índices para cada tarefa do jogo separadamente. Ao final da execução, esses índices são analisados para dar uma pontuação geral das habilidades cognitivas do paciente. Um exemplo de cenário virtual e tarefas executadas no jogo podem ser vistos na Figura 4.

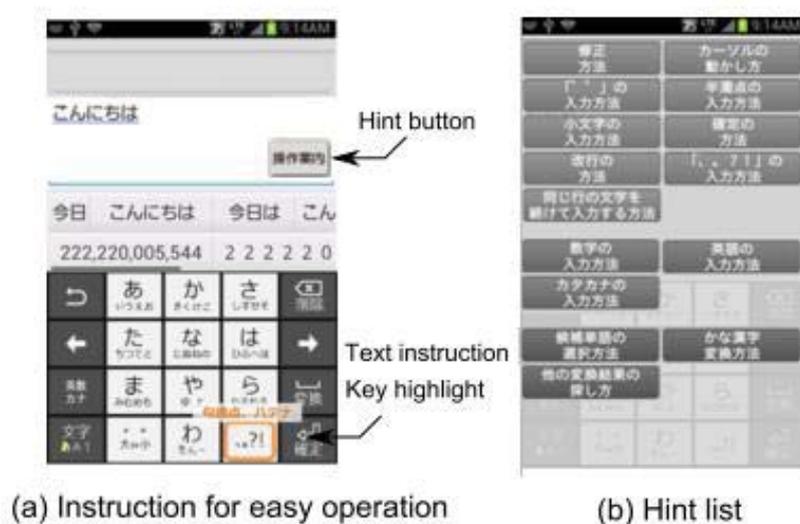


Figura 3 – Imagem da tela da aplicação [21].



Figura 4 – Imagem do cenário virtual gerado (à esquerda) e exemplo de tarefa de interação (à direita) [22].

O software está em fase de avaliação preliminar por grupo de 1000 usuários diagnosticados com alguma imparidade cognitiva leve ou doenças neurodegenerativas. A amostra é estratificada de acordo com o gênero, em grupos de idades (50-60, 61-70, 71-80), e nível de escolaridade. É utilizado o Mini Exame do Estado Mental (MEEM) como critério de inclusão e também a Subescala do Domínio da Memória Escala de Avaliação da Demência Clínica [23]. Os critérios de inclusão são história negativa para doenças neurológicas e psiquiátricas, escore MEEM > 24 e desempenho dentro dos intervalos normativos para idade e escolaridade nos testes de lápis e caneta. O estudo não apontou resultados.

3.2.4. Himmelsbach *et al.*

Himmelsbach *et al.* [24] investigam quais fatores influenciam as escolhas dos idosos com relação aos tipos de interação preferidos. O objetivo considerar essas preferências no desenvolvimento de novos sistemas móveis multimodais.

Para tanto, houve a participação de 19 idosos com idades entre 66 e 84 anos, que levaram um telefone para casa e interagiram com ele durante três semanas. Foi utilizada a técnica de reconstrução do dia [25] como uma forma de avaliação, através da qual os participantes relatam sistematicamente suas atividades e experiências do dia anterior com procedimentos destinados a reduzir os vieses de recordação. As técnicas de incidentes críticos [26] e entrevistas telefônicas semanais também foram utilizadas. No final do estudo foi realizada uma entrevista centrada em problemas. Eles puderam escolher entre realizar as interações através de toque ou fala.

Como resultado foi identificado que deficiências leves ou graves influenciam nas escolhas dos usuários em relação a qual modalidade de interação utilizar, por exemplo, uma deficiência leve no controle motor faz com que a interação preferida seja através da fala. Além disso, quando a interação ocorre em espaço público existem considerações a respeito de privacidade e das condições ambientais que incentivam a escolha dos idosos por um ou outro tipo de interação. No entanto, de modo geral, a interação por fala foi a mais aceita.

3.2.5. Sacramento, Bacellar e Ferreira

O objetivo do trabalho de Sacramento, Bacellar e Ferreira [27] foi avaliar a interação dos idosos com uma rede social popular (Facebook) baseado no fator da comunicabilidade.

Adotou-se uma abordagem qualitativo-exploratória, baseada em estudo de caso com múltiplas unidades de análise e que consiste na observação e na aplicação do MAC-g [28] na interação com o Facebook por pessoas idosas e não-idosas, em abordagem comparativa. A avaliação foi restrita a conversas de chat, ao compartilhamento de informação na linha de tempo de outra pessoa e a ação de curtir a página de uma pessoa pública. Participaram adultos com idades entre 30 e 50 anos e idosos entre 70 e 90 anos.

O método foi dividido em quatro etapas:

- Definição do sistema de redes sociais e das características a serem analisadas;
- Definição da versão do Facebook analisada;
- Definição do método de avaliação;
- Avaliação da capacidade de comunicação.

Os resultados da avaliação de comunicabilidade revelaram diferenças consideráveis entre idosos e não idosos no que se refere à quantidade de avarias de comunicação e falhas completas nas incidências encontradas e na duração da conclusão das tarefas.

Uma das principais diferenças entre esses dois perfis (adultos e idosos) foi relacionada à linguagem utilizada. Nas interações feitas pelos idosos haviam muitas identificações “o que é isso?” com relação aos objetos do software. Eles tiveram dificuldades para entender os símbolos (frequentemente representados por ícones) disponíveis na interface. “O que é isso?” pode indicar desconhecimento por parte do usuário quanto ao significado referido pelo designer e gerar conflitos entre o sistema de significado conhecido pelos usuários e o especificado pelo sistema.

Dificuldades de compreensão da linguagem do Facebook também podem ter influenciado sobre as baixas taxas de conclusão das tarefas desconhecidas obtidas por idosos em comparação com participantes não-idosos. Esses resultados confirmam que os idosos apresentaram dificuldades em entender meta-mensagens usadas pelos designers de redes sociais.

3.2.6. Le, Thompson e Demiris

No trabalho de Le, Thompson e Demiris [29] é realizada uma pesquisa bibliográfica para pesquisar técnicas utilizadas com idosos, com comparação de três

diferentes métodos de avaliação, buscando avaliar o desempenho do usuário, sua experiência de interação e a comunicação através da visualização.

As três abordagens que eles utilizam são benchmark, insights e questionários. Benchmarks focam na avaliação da eficácia através de métodos empíricos baseados em tarefas. Os participantes recebem tarefas definidas para serem concluídas durante o processo interativo, com registro de métricas de precisão e tempo de conclusão. Estas medidas quantitativas permitem comparações entre designs de visualização. A abordagem de insights utiliza o protocolo Think Aloud, através do qual os participantes são convidados a explorar abertamente a visualização da informação de uma maneira pensada em voz alta. Com orientação mínima dos moderadores, os participantes verbalizam insights gerados através da interação com a visualização. Os pesquisadores posteriormente categorizaram essas percepções, para quantificar e comparar os resultados. A terceira abordagem utilizada, questionários, envolve pedir aos participantes que classifiquem sua experiência interagindo com a visualização. Esta experiência aborda a percepção de usabilidade e aprendizado das visualizações de informações. O questionário é preenchido após o processo interativo. Utilizou-se a Escala de Usabilidade do Sistema (SUS), um questionário de avaliação de usabilidade padrão que consiste em 10 perguntas agregadas em uma escala de 0 a 100.

Para os testes foram recrutados 21 idosos com 60 anos ou mais (15 mulheres e seis homens) que, após receberem orientação sobre o estudo, avaliaram e exploraram duas visualizações diferentes com cinco tarefas abrangendo toda a amplitude de interações disponíveis dentro de cada visualização. Os participantes passaram pelas três abordagens de avaliação, sendo as duas primeiras enquanto exploravam a visualização e a última, após o término da exploração. Os participantes receberam uma bonificação de \$15 para participarem do estudo.

Quanto à avaliação de benchmark, ao comparar os tempos de conclusão da tarefa entre as visualizações, não foi encontrado significância estatística para nenhuma das tarefas (95% CI: [-6.3, 39.5]). O tempo de conclusão da tarefa foi altamente variável, com os participantes diferindo em motivação e persistência na realização de tarefas. Portanto, não foram encontradas evidências para concluir que o tipo de visualização afetou a proporção de participantes que são capazes de executar corretamente as tarefas.

Quanto aos insights, foi possível perceber que os participantes geraram em média mais percepções com a segunda visualização (IC 95%: [14,9, 25,1]) em relação à primeira visualização (IC 95%: [10,8, 16,7]). Ao categoriza-los por meio da taxonomia funcional de tarefas de baixo nível utilizada no trabalho, pode-se perceber que a maioria dos insights

representava uma caracterização da distribuição entre as visualizações. Esses insights frequentemente apresentavam tendências nos dados (aumentando, decrescendo ou permanecendo variáveis).

Na avaliação com o questionário, os participantes classificaram a Visualização A no SUS com pontuação média 66,1 (IC 95%: [54,6-77,7]), enquanto a Visualização - B foi avaliada com pontuação média 73,3 (IC 95%: [64,0-82,5]). Considerando que uma pontuação de 68 na escala SUS é considerada média no percentil 50, os autores concluíram que os participantes perceberam a Visualização B como um pouco acima da média na usabilidade, enquanto a Visualização A estava próxima à media. Um teste t pareado comparando os escores SUS não mostrou diferenças estatisticamente significativas no escore SUS entre visualizações ($p = 0,12$), apesar da visualização B ser superior a A por um escore médio de 5,9 (IC 95%: [-1.6, 13.5]).

3.2.7. Al-khomsan *et al.*

O trabalho de Al-khomsan *et al.* [30] tem o objetivo de avaliar a usabilidade do aplicativo Twitter usando a plataforma Android.

A avaliação de usabilidade foi conduzida do ponto de vista do usuário, focando em como ele interage com o aplicativo para executar as tarefas representativas. Dez participantes idosos participaram do experimento, no entanto não foram definidos o gênero e a idade dos mesmos. Os participantes testaram em um cenário focado na funcionalidade básica do Twitter, enquanto a ferramenta Morae (<http://www.techsmith.com/morae.html>) registrou o comportamento do usuário através do software Mobizen (<https://www.mobizen.com/>). Ao mesmo tempo, o observador observou o participante e somente após a conclusão de todas as tarefas, os participantes preencheram o questionário SUS.

Foram identificadas dificuldades dos usuários em relação à usabilidade do aplicativo através do smartphone e eles não conseguiram concluir todas as tarefas. A avaliação determinou que características como cores muito brilhantes, ícones em lugares inapropriados e símbolos muito pequenos dificultavam a interação dos idosos com a interface. Além disso notou-se a necessidade de um maior número de feedbacks em relação às atitudes dos usuários, informando ao usuário o motivo do seu erro, e também a disponibilização de uma descrição sobre os ícones.

3.2.8. Konstantinidis *et al.*

Na pesquisa de Konstantinidis *et al.* [14], os autores apresentam a concepção, implementação e avaliação da plataforma FitForAll (FFA) que disponibiliza *serious games* e exercícios físicos para idosos.

Os critérios de inclusão para ambos os grupos de teste (intervenção e controle) foram: idade ≥ 60 , habilidades linguísticas fluentes, visão e audição normais ou corrigidas para o normal, exame e permissão formal de um cardiologista e termo de consentimento assinado, obtidos antes do início do ensaio.

Participaram do experimento 116 idosos, em média de 2-3 vezes por semana, durante 7-8 semanas. Cada usuário realizou exercícios aeróbicos durante 20 minutos, exercícios de resistência durante 10 minutos, exercícios de flexibilidade durante 10 minutos e um conjunto de equilíbrio exercícios direcionados (Figura 5). Foram utilizadas medidas subjetivas e objetivas para avaliar a plataforma FFA. A avaliação foi realizada principalmente em duas frentes, o nível de usuário (usabilidade, aderência) e o nível de eficácia (impacto em termos de aptidão, impacto sobre a qualidade de vida). A experiência do usuário foi avaliada através do questionário Inventário de Usabilidade de Software (SUMI) [31]. Além disso, os participantes também preenchem o SUS, e apontavam melhorias na qualidade de vida por meio do questionário WHOQoL-BREF [32].

Os valores resultantes do questionário SUMI para a Eficiência, Utilidade, Controle, Aprendizagem e pontuação global são todos acima de 60% (64,84; 65,73; 61,69; 60,09; e 68,33, respectivamente). A pontuação em relação ao questionário SUS ficou também acima dos 50% (77,7 normalizado). Os usuários opinaram positivamente sobre a ferramenta, 85,4% dos usuários perceberam que a plataforma FFA lhes permite controlar melhor a sua saúde. Além disso, 87,1% dos usuários relataram que com um máximo de 5 dias de familiarização com a plataforma foram capazes de usá-la sem ajuda, enquanto outros 92,2% relataram que consideravam isso como uma plataforma de nível comercial.

3.2.9. Hsieh e Huang

Hsieh e Huang [33] realizaram um pesquisa com foco em avaliar as experiências operacionais em pessoas idosas de modo a explorar o grau de carga física e mental gerado por determinados softwares; avaliar os graus de fadiga e carga de trabalho mental quando pessoas idosas operam diferentes dispositivos de hardware; e analisar a relação entre o desempenho de

determinados softwares , a carga de trabalho mental dos idosos e o desconforto subjetivo quando os idosos operam esses sistemas.

Houve a participação de 20 idosos com idades superiores a 65 anos, sendo 10 homens e 10 mulheres sem experiência, utilizando os seguintes softwares: Angry-birds, Jogo de quebra-cabeças e Facebook. Os sujeitos tiveram a liberdade de interagir com os programas livremente por cinco minutos e então preencheram um questionário com Escala Likert de cinco níveis.



Figura 5 – Idosos realizando as tarefas de interação [14].

Foi verificado que devido às limitações de visão e de sensibilidade manual dos idosos seria interessante que um computador projetado para idosos tenha o cursor ampliado. No entanto, ressaltam que os resultados experimentais não são suficientemente abrangentes.

3.2.10. Báez *et al.*

Báez *et al.* [15] descrevem em seu trabalho o desenho e validação de um ambiente de aptidão virtual que visa manter os idosos fisicamente e socialmente ativos. O ambiente de fitness virtual (Gymcentral) foi concebido para permitir e motivar os adultos mais velhos a seguir exercícios personalizados a partir de casa, com um grupo (heterogêneo) de amigos remotos e sob a supervisão remota de um treinador. O sistema conta com cenários específicos de recepção, vestiário e sala de aula (Figura 6).



Figure 3. Main features of the Trainee Application

Figura 6- Ambientes disponíveis no Gymcentral [15].

A avaliação do sistema contou com a participação de 36 idosos com idades entre 65 e 87 anos considerados com um nível de fraqueza não-frágil, transitualmente frágil ou leve. O nível de fragilidade foi medido utilizando o Groningen Frailty Indicator [34], um questionário validado que procura as limitações auto referidas em idosos. Todos os participantes obtiveram uma aprovação formal por escrito por seu médico para participar do estudo. Ambos os médicos e os participantes receberam um esboço escrito e explicação do estudo antes de participar. O procedimento durou um total de 10 semanas, incluindo uma semana no início para implantação técnica, testes de aplicação e a coleta de questionários iniciais, e uma semana no final para a administração dos questionários finais.

Utilizando um procedimento de distribuição aleatória correspondente com base no nível de idade e fragilidade, os participantes foram atribuídos a um grupo experimental (social) ou de controle. Os participantes do grupo social receberam a versão completa da App

Trainee, incluindo o programa personalizado, juntamente com os recursos sociais e de persuasão (a condição com o conjunto mais complexo de recursos). O grupo de controle, ao contrário, tinha uma versão básica da aplicação, que incluía o programa de exercícios personalizado, mas sem características persuasivas, sociais e de auto monitoramento.

A usabilidade foi avaliada pelo instrumento SUS em dois momentos: no início do estudo (após o tutorial na aplicação, quando os participantes usaram o aplicativo pela primeira vez), e no final do estudo. Estas medidas foram obtidas para ambos os grupos, a fim de comparar a usabilidade dos diferentes níveis de complexidade da interface. Para avaliar a aceitação da tecnologia e a utilidade dos recursos foram elaborados questionário com cinco níveis de respostas. Os questionários não foram apresentados no trabalho.

Uma análise mista de variância entre sujeitos foi conduzida para comparar a pontuação do SUS entre os participantes no grupo experimental e no grupo controle. O resultado sugere que, embora no início houvesse uma notável diferença na usabilidade das duas versões do Gymcentral (com a versão básica tendo uma melhor desempenho), a usabilidade da aplicação obteve uma melhora significativamente maior ao longo do tempo no grupo social do que no grupo de controle ($F(1, 34) = 8.286, p = .007$). Embora inicialmente a aplicação completa tenha sido relatada como mais difícil de usar, no final do estudo sua usabilidade percebida aumentou para atingir um nível comparável ao da versão básica.

Também foi conduzida uma análise mista de variância entre sujeitos para comparar cada uma das dimensões de aceitação de tecnologia no início e no final do estudo. Pode-se perceber que a ansiedade em relação à aplicação diminuiu significativamente ao longo do tempo para o social, mas não para o grupo de controle ($F(1, 34) = 5.543, p = .024$). Além disso, testes de comparação múltipla com correção de Bonferroni mostraram que a atratividade da aplicação ($F(1,34) = 7.448, p = .01$), a satisfação do usuário ($F(1,33) = 6.561, p = .015$) e a utilidade percebida da aplicação ($F(1,34) = 4.291, p = .046$) aumentaram significativamente para o grupo social, mas não para o grupo controle. Embora os resultados da análise de variância sugiram que, considerados em conjunto, ambos os grupos relataram ter gostado da aplicação significativamente mais no final do estudo, as comparações pós sugerem que essa diferença foi significativa para o grupo social, mas não para o grupo de controle. Ainda, a usabilidade da aplicação foi menor para o grupo Social no início do estudo, refletindo as dificuldades iniciais dos participantes para lidar com uma interface de usuário mais complexa. No entanto, no fim do programa de treinamento, a usabilidade aumentou significativamente, aproximando-se do topo da escala.

3.2.11. Schuurmans *et al.*

Schuurmans *et al.* [35] descrevem o contexto teórico e a metodologia de um estudo que examina a viabilidade de um sistema de auto-gestão de comprimidos para idosos com transtornos de humor recorrentes. O eCare Homeplatform objetiva melhorar a consciência e o conhecimento dos pacientes sobre os transtornos de humor recorrentes e seu tratamento, promover a autogestão através do uso de uma ferramenta simples de monitoramento e facilitar o contato on-line com seu clínico por videoconferência.

A aplicação ainda não foi executada. Os autores apenas descrevem o procedimento previsto.

Para a avaliação do projeto serão selecionados pacientes com idades superiores a 60 anos e com diagnóstico de depressão ou transtorno bipolar e que estejam em tratamento por um período mínimo de 6 meses. A aceitabilidade do sistema será medida de duas maneiras. Primeiramente, utilizando a análise de arquivo de log para determinar os níveis de uso para as funções do sistema (monitoramento de humor, biblioteca e videoconferência). Depois, por intermédio do Questionário de Credibilidade / Expectativa (CEQ) [36] que busca medir expectativas prévias em relação ao sistema e credibilidade das opções do sistema.

Na metade e no final do procedimento de testes, com duração total de quatro meses, os participantes do estudo avaliarão a usabilidade da interface do sistema ECH usando a Escala SUS e preencherão o Questionário de Satisfação do Cliente-8 (CSQ-8) [37]. Uma entrevista qualitativa semiestruturada também será administrada aleatoriamente para dez pacientes e cinco clínicos envolvidos no estudo de viabilidade. O objetivo será avaliar questões como impressão geral, experiências pessoais com o sistema, valor agregado ao tratamento, contato com o clínico, estratégias de autogestão e quaisquer falhas no sistema.

3.2.12. Chiu *et al.*

O trabalho de Chiu *et al.* [38] descreve a avaliação de um treinamento de um dispositivo móvel touchscreen, em que trinta e nove idosos participaram (25 mulheres e 14 homens), com 50 anos ou mais, durante oito semanas. Eles foram recrutados de um centro comunitário em uma baixa área de uso de Internet no sul de Taiwan.

Os participantes passaram por uma sessão inicial na qual se explicava como se procederia o estudo e então assinavam um termo de consentimento. Quatro participantes recusaram participar do estudo e dois participantes foram excluídos devido ao analfabetismo.

Um participante morreu antes do término estudo e dois participantes participaram das sessões de treinamento menos de cinco vezes. Dez participantes adicionais abandonaram antes da última sessão e não completaram a pesquisa final.

Assim, um total de 20 participantes (51% dos 39 participantes iniciais) completaram o curso de treinamento e o pré- e pós-questionário que foi utilizado para análise estatística. Além disso, 16 participantes concordaram em ser entrevistados e seis entrevistas com grupos focais foram realizadas no dia seguinte ao final do curso de treinamento. Após as entrevistas iniciais, os participantes foram convidados a participar de um curso de treinamento baseado em aplicativos para as oito semanas seguintes, em que cada sessão teria a duração de 90 minutos. Além disso, um mês após o fim do curso de formação, os participantes foram convidados a participar de entrevistas de grupo focadas para partilharem as suas experiências em relação à tecnologia ou com a Internet e as aplicações que utilizavam.

Além do questionário para identificar os dados sociodemográficos do usuário, o Attitudes Toward Computers Questionnaire [39] foi usado para avaliar as atitudes dos participantes em relação ao uso da Internet. Com este questionário, foi possível medir principalmente a ansiedade e a confiança, incluindo perguntas como "É difícil para mim entender o uso do computador / Internet" e "Eu acredito que com mais prática eu possa navegar na Internet mais facilmente".

Quando à atividade de navegar na internet, foi abordada a perspectiva de como isso poderia influenciar nas suas vidas. Os idosos determinaram que positivamente poderia ampliar sua rede social, aumentar sua busca por informações úteis, enriquecer suas atividades de lazer, melhorar sua saúde física e aumentar sua felicidade. Porém, a perspectiva negativa foi mensurada quanto às questões de que o uso da Internet poderia deteriorar a visão, aumentar a fadiga e piorar a capacidade de comunicação face a face.

A avaliação também considerou o viés psicológico do participante, por meio da Escala de Depressão dos Estudos Epidemiológicos (CES-D10) [40] e da Escala de Inventário de Felicidade da China [41], ambos com três níveis de resposta.

Os autores observaram que os participantes que tiveram uma visão positiva da internet continuaram utilizando-a regularmente após o término do treinamento. Além disso, o interesse poderia ser ainda maior se as aplicações tinham alguma relação com sua vida diária. Os participantes ainda declararam que o interesse em aprender seria maior se tivessem alguém a quem pedir auxílio e tirar dúvidas enquanto estão utilizando o aparelho.

3.2.13. Meza-Cubo *et al.*

No trabalho de Meza-Cubo *et al.* [42], os autores construíram, treinaram e validaram um padrão de rede neural para identificar emoções agradáveis e desagradáveis. A rede neural foi treinada utilizando sinais de eletroencefalograma (EEG) registrados durante a apresentação de estímulos visuais que induziam emoções conhecidas. Então, validou-se o uso de uma rede neural treinada para avaliar as emoções em idosos enquanto eles utilizam uma aplicação. As emoções dos usuários foram registradas utilizando o dispositivo EMOTIV EPOC + Headset 2, um headset sem fio com 14 sensores capazes de detectar sinais cerebrais e expressões faciais dos usuários (Figura 7).



Figura 7 – Headset utilizado para registrar os sinais cerebrais e as expressões faciais dos usuários [42].

A aplicação utilizada pelos idosos foi o CS, um sistema que usa uma metáfora de jogo de tabuleiro (Figura 8). Para validar o uso da rede neural houve a participação de 23 idosos, sendo 10 homens e 13 mulheres, com idades entre 60 e 83 anos. Foram registrados auto relatos dos participantes através de um questionário aplicado imediatamente ao final do processo interativo, e análises qualitativas adicionais por observações das avaliações para contrastar com os resultados obtidos. Para a avaliação da análise qualitativa, as sessões de atividade do CS foram gravadas em vídeo e analisadas para identificar expressões verbais e não verbais de adultos mais velhos durante a realização da atividade.

Alguns participantes relataram que iniciaram os testes ansiosos, mas todos eles relataram que houve um engajamento com o jogo. Os resultados da avaliação qualitativa mostraram que a maioria dos idosos demonstrou reações (expressões e atitudes) agradáveis

durante a atividade. Apenas um participante mostrou reações agradáveis e desagradáveis e alguns outros permaneceram neutros. A rede neural foi capaz de determinar os participantes que demonstraram reações agradáveis e classificou todos os outros como tendo demonstrado reações desagradáveis. Portanto, os autores concluem que é possível avaliar a de idosos enquanto eles usam um aplicativo CS usando a rede neural treinada para reconhecer emoções agradáveis e desagradáveis de sinais EEG com uma precisão de reconhecimento do método variando de 60,87% a 82,61%, o que, segundo os autores, está de acordo com resultados de estudos anteriores relacionados.

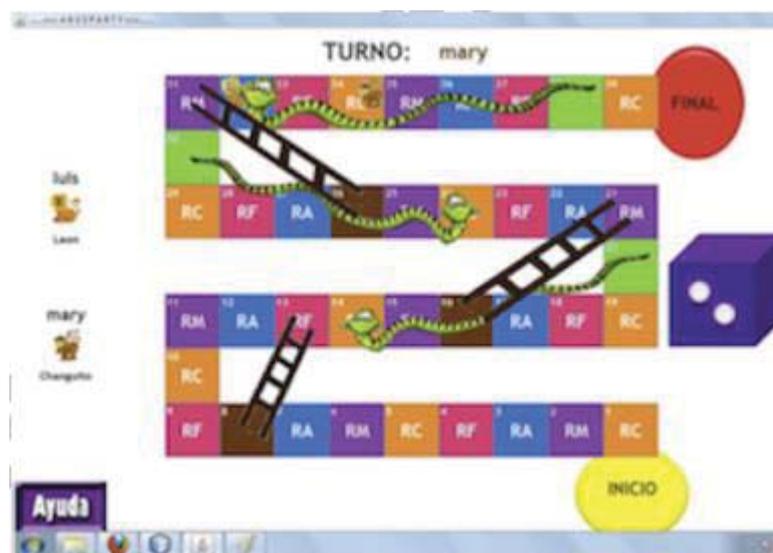


Figura 8 – CS, Jogo virtual utilizado para realizar os testes [42].

3.2.14. Ferreira *et al.*

No trabalho de Ferreira *et al.* [43], os autores desenvolveram a aplicação intitulada “Assistente de Medicação” seguindo um método iterativo centrado no usuário idoso. A avaliação é feita a cada iteração, considerando as funcionalidades do sistema.

A avaliação do primeiro protótipo foi conduzida com um grupo heterogêneo de especialistas, três mulheres e um homem com idades entre 25 e 60 anos. O Protocolo Think Aloud foi utilizado, com gravações dos áudios. O objetivo desta primeira avaliação foi reunir opiniões e sugestões a respeito das questões de design da interface.

A avaliação do segundo protótipo foi realizada com um grupo de usuários composto por três mulheres e um homem com idades entre 57 e 76 anos. Este teste foi dividido em duas etapas: etapa de interação e etapa de avaliação da usabilidade. Na primeira etapa, o usuário deveria seguir um conjunto de tarefas de interação com o aplicativo por meio

de um smartphone e enquanto isso, o observador deveria fazer anotações acerca das dúvidas, dificuldades, comentários, sugestões e problemas que o usuário manifestasse por meio do protocolo Think Aloud. Na segunda etapa, aplicava-se um questionário concernente à experiência de interação do usuário com a interface. O questionário buscava identificar questões como layout, tamanho e cor das letras, demais características do aplicativo e interação com o mesmo.

Através da avaliação foi possível identificar que cor e tamanho da fonte, interações com o aplicativo e complexidade do mesmo se salientaram como pontos negativos da aplicação, os usuários apresentaram dificuldades em relação a estas características. Por outro lado, o auxílio prestado pelo aplicativo nas situações de esquecimento de tomar a medicação ou com relação à data de validade dos medicamentos e características como menu de ajuda, foram considerados pontos fortes do aplicativo. Os participantes elogiaram a facilidade de uso e a utilidade do aplicativo. A Figura 10 mostra as principais telas do aplicativo (Tela Inicial, Menu Principal, Menu de Ajuda, Alarmes de Medicação e Informações sobre Medicamento, respectivamente).



Figura 9 – Telas do aplicativo Medication Assistant [43].

3.2.15. Teixeira *et al.*

Teixeira *et al.* [44] traz a continuação do trabalho de Ferreira *et al.* [43], portanto, trabalha sobre o mesmo projeto e um novo protótipo é produzido para ser avaliado.

No decorrer do desenvolvimento, foram elaborados três protótipos e uma avaliação de cada um deles. A avaliação com o primeiro protótipo ocorreu com um grupo heterogêneo de engenheiros e profissionais da saúde. A avaliação com o segundo protótipo utilizava o protocolo Think Aloud e um questionário sobre a experiência de interação do

usuário com idosos, três mulheres e um homem, com idades entre 57 e 76 anos. A primeira e a segunda avaliações são as mesmas descritas no trabalho de Ferreira *et al.*

Para a avaliação com o terceiro protótipo, o usuário foi guiado por um conjunto de tarefas a serem executadas ao se interagir com o aplicativo. Então ele utilizava o protocolo de avaliação ICF-US I (International Classification of Functioning based Usability Scale), com um nível de detalhamento maior para cada item do questionário. Participaram desta fase da avaliação 10 idosos (nove mulheres e um homem) com idades entre 57 e 82 anos.

Os resultados mostraram que os usuários notaram melhoras no menu de ajuda e na interação por voz. No entanto, destacaram que os comandos por voz e a interação por toque precisam ser melhorados. Outras sugestões resumem-se em adicionar funcionalidades como planejar o dia, gerir e inserir medicação

O menu de ajuda e a utilidade do sistema foram elencados como pontos positivos da aplicação. Comandos de voz e deslizar e a funcionalidade de status da aplicação ainda são considerados pontos fracos. A Figura 10 mostra a versão final da interface.

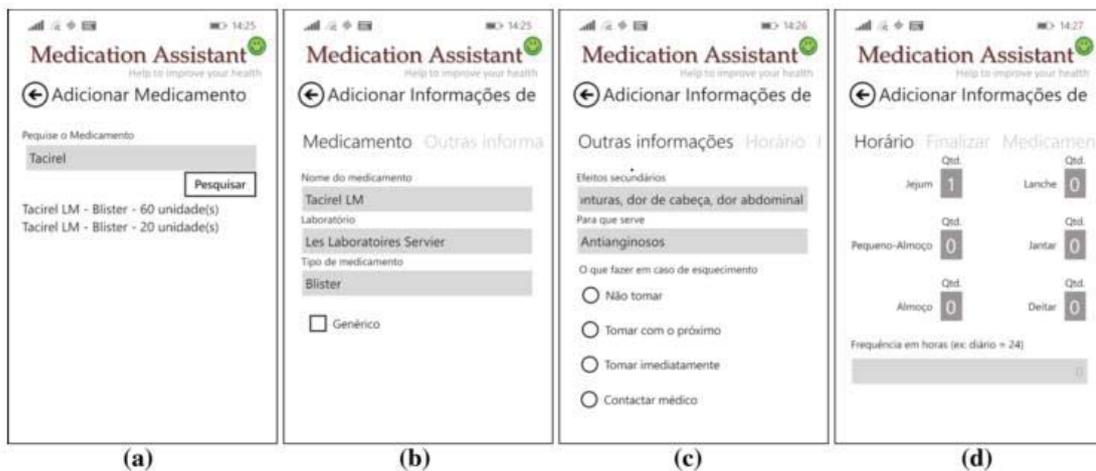


Figura 10 – Telas do aplicativo Medication Assistant resultante do processo de avaliação [44].

3.2.16. Epelde *et al.*

Neste artigo, Epelde *et al.* [45] descrevem uma nova abordagem capaz de integrar todos os tipos de sistemas interativos, local ou remotamente, de maneira que possibilita a personalização da interface de acordo com as necessidades de cada grupo de usuários. É possível escolher o dispositivo de interação, a tecnologia de interação e a apresentação da interface do usuário. A Figura 11 mostra um exemplo de interface.



Figura 11 – Imagens exemplo da interface do sistema [45].

O sistema foi avaliado ao final de cada iteração de desenvolvimento. A avaliação preliminar foi realizada com 40 usuários, suas idades e sexo não foram especificadas. Os usuários interagiram com os protótipos e então passaram por entrevistas individuais e em grupos. Para a avaliação final, os testes foram realizados por 20 usuários em suas casas por um período de três a quatro semanas e por 47 usuários em associações de idosos, com idades entre 52 a 89 anos. Foram apresentadas as funcionalidades do sistema (Figura 12) e os usuários foram então convidados a responderem alguns questionários que envolviam questões sociodemográficas, sociais, diárias, de satisfação com sua vida e para a avaliação das funcionalidades da plataforma. Os usuários foram também solicitados a discutirem em voz alta entre si sobre as funcionalidades do sistema e foram feitas anotações a respeito.

Os resultados discutidos no artigo se referem somente aos usuários que participaram dos testes em associações de idosos. A maioria dos idosos manifestaram que a aplicação poderia ser útil para manter contato com seus familiares e melhorar sua vida social, bem como melhorar a sua qualidade de vida. Todos eles manifestaram que a interface estava legível e a maioria deles considerou-a agradável ou muito agradável e ressaltaram que não provocava cansaço nos olhos. Eles ficaram impressionados com a tecnologia de controle por

voz, mas deram conselhos a respeito do fornecimento da tecnologia em outras línguas e da precisão da interação por voz com usuários idosos.



Figura 12 – Idosos participando da sessão de demonstração do sistema [45].

3.2.17. Discussão dos Estudos Selecionados

A Tabela 4 mostra os resultados desta Revisão Sistemática, no que tange às ferramentas utilizadas em cada abordagem para avaliação com idosos. A tabela mapeia os questionários de usabilidade utilizados (reconhecidos pela literatura) e questionários diversos (também validados na literatura ou criados pelos próprios autores). O resumo também identifica os trabalhos que utilizaram entrevistas no seu processo de avaliação do sistema, e outros métodos que os autores possam ter utilizado, que não se encaixam nas categorias “questionários” ou “entrevistas”. A expressão “n.e.” (não especificado) identifica os casos em que os autores declaram ter usado algum questionário, mas não o identificam, nem mesmo sua fonte. A Tabela 5 resume as idades dos participantes e outras características relacionadas à avaliação, como a tridimensionalidade da interface e dispositivos utilizados, e mais detalhes a respeito dos métodos de avaliação utilizados (usando a classificação de Laviola *et al.* [4]) e sobre os aspectos de interação que foram avaliados em cada um dos trabalhos.

Para análise dos resultados em relação aos idosos será utilizada a definição do *World Health Organization* no que diz respeito à classificação de idosos como pessoas com 60 anos ou mais [46]. Nesse contexto, é importante ressaltar que WHO reconhece os idosos não apenas pela idade, mas por novas posturas, perda de posturas anteriores ou incapacidade de realizar contribuições ativas para a sociedade [47].

Segundo Laviola *et al.* [4], os métodos de avaliação de interfaces podem ser categorizados em: caminhada cognitiva, avaliação heurística, avaliação formativa, avaliação somativa, questionários e entrevistas e demonstrações. A partir dos resultados da Revisão Sistemática, pode-se observar que a maioria dos autores utilizou avaliações formativas (75%) e pelo menos dois métodos diferentes para avaliação da interface (~87%), tendência recomendada por Laviola *et al.* [4]. Apenas quatro trabalhos (25%) utilizam avaliação somativa, e quando utilizam o fazem para a comparação de grupos de usuários.

Em alguns casos, entrevistas (~37%) são usadas durante o processo de avaliação. Além disso, os estudos selecionados frequentemente utilizam questionários (~93%), mas não mencionam possíveis dificuldades que idosos ao utilizarem esse tipo de técnica. A respeito disso, sugere-se um estudo mais aprofundado, ou maneiras de flexibilizar o uso da técnica, pois isso reflete nas condições de acessibilidade do público, que pode variar de acordo com a região que se aplica o método e o grau de escolaridade do idoso [47]. Nota-se que as abordagens utilizam questionários criados pelos próprios autores ou questionários genéricos do ponto de vista da faixa etária a que são aplicados, e não mencionam qualquer tipo de adaptação de linguagem para facilitar a comunicação com o público alvo.

De acordo com Guillemain, Bombardier e Beaton [48] e Günter [12], é importante aplicar um processo de adaptação transcultural para questionários destinados a um alvo específico, permitindo uma compreensão adequada da linguagem e evitando ambiguidades e subavaliações. No caso de instrumentos destinados a pessoas idosas, é necessário que haja uma abordagem ainda mais sensível para essa questão, uma vez que se trata de uma população que não compreende alguns domínios de tecnologia [49]. Os recursos tecnológicos são relativamente novos ou desconhecidos para os idosos, pouco usados, dificultando a compreensão de certos termos e abstração de determinadas situações e expressões. Uma alternativa para essa dificuldade foi apontada no trabalho de Brum e Rieder [18] e realizada em Teixeira *et al.* [44]. Consiste em usar questionários mais detalhados, usando uma linguagem que oferece uma definição mais clara sobre o que ela pretende fazer.

Considerando que exista carência de métodos específicas para avaliação de interfaces 3D de contexto específico ao público idoso, mostra-se relevante destinar esforços para a estruturação de um instrumento de avaliação de interface destinado a pessoas com 60 ou mais anos. Isso se justifica também pelo emprego recente de tecnologia para melhorar a qualidade de vida e auxiliar em processos de reabilitação para essa faixa etária – além das facilidades em se adquirir equipamentos eletrônicos e aplicativos. Analisando os resultados do trabalho, pode-se observar que apenas dois deles utilizam interfaces 3D (~12%). Isso pode ser

devido justamente a carência de esforços direcionados para a avaliação destas tecnologias de forma específica para o público idoso, uma vez que a avaliação deste tipo de interface geralmente tende a apresentar maiores dificuldades aos usuários [50].

Medidas subjetivas ou medidas de desempenho de usuário são utilizadas por diferentes autores [21] para verificar as dificuldades dos idosos, uma vez que medidas de desempenho baixas podem significar que o idoso não está conseguindo realizar uma tarefa ou aprender a melhor forma de realizá-la. De acordo com Nielsen [3], facilidade e eficiência em realizar tarefas básicas é uma característica da usabilidade, portanto, é fundamental que uma interface permita a realização de uma tarefa e sua aprendizagem, independente das possíveis limitações dos usuários, nesse caso, advindas da idade avançada.

A percepção visual do usuário é importante para que este perceba o que deve fazer em determinado cenário e também para estabelecer o sentido de imersão e presença, uma das principais variáveis de usabilidade [3]. O trabalho de Sacramento, Bacellar e Ferreira [27] trouxe um exemplo claro de tal importância, uma vez que os idosos apresentaram dificuldades em entender objetos na interface que os adultos jovens assimilaram facilmente.

Tal peculiaridade também foi percebida no trabalho de Al-Khomsan et al [30], uma vez que os usuários idosos não conseguiram concluir todas as tarefas e mostraram dificuldades com a usabilidade de um aplicativo comumente utilizado com facilidade por outras faixas etárias. Nesta pesquisa, identificou-se que as principais características da interface que dificultaram a interação dos idosos foram cores, locais dos ícones e tamanhos dos símbolos e ícones. Além disso, notou-se a necessidade de um maior número de feedbacks em relação às atitudes dos usuários, informando ao usuário o motivo do seu erro, e também a disponibilização de uma descrição sobre os ícones, no caso de usuários idosos.

Em contrapartida, quando a interface é desenvolvida com atenção às particularidades exigidas por esse público, o nível de satisfação é melhorado. Além disso, a dificuldade em interagir com a interface, o tempo de assimilação e entendimento do funcionamento do sistema são minimizados [14] [15].

As pesquisas de Ferreira et al. [43], Teixeira et al. [44] e Al-Khomsan et al. [30] sugerem que cores muito brilhantes ou muito impactantes influenciam negativamente na experiência do usuário com uma interface, sendo mais adequado a este perfil, cores mais discretas. Além disso, nos mesmos trabalhos, e também no trabalho de Hagiya, Horiuchi e Yazaki [21], pode-se observar que funcionalidades de ajuda e feedbacks ao usuário são pontos positivos em interfaces destinadas ao público idoso.

Como observado no trabalho de Himmelsbach [24] e Hsieh e Huang [33], deficiências leves ou graves influenciam na escolha dos idosos sobre qual método de interação utilizar, de forma que uma deficiência leve no controle motor, pode fazer com que o usuário prefira a interação através da fala. Portanto, é necessário atentar para o modo de interação da interface que se está utilizando e para as deficiências que os usuários possuem, procurando oferecer alternativas para a conclusão da tarefa, e proporcionando uma melhor experiência ao usuário.

Embora esta pesquisa tenha encontrado muitas características de usabilidade das interfaces que se mostraram importantes a serem consideradas, uma vez que influenciaram na experiência do usuário com a interface, apenas ~37% dos trabalhos avaliaram a usabilidade dos sistemas que utilizaram. Além disso, nenhum trabalho mencionou realizar a avaliação da interface levando em consideração as diferenças existentes entre as interfaces 2D e 3D, pois muitos deles ao menos diferenciavam sua interface de acordo com a sua dimensionalidade. De acordo com LaViola *et al.* [4], vários conceitos e objetivos na avaliação de usabilidade de interfaces tradicionais e tridimensionais são similares, porém, as abordagens de aplicação utilizadas no contexto de interfaces 3D geralmente precisam ser diferentes, pois muitas das características avaliadas em interfaces tradicionais se relacionam com características não estudadas neste tipo de interface, como imersão e presença, que são inerentes à interfaces tridimensionais. É necessário, portanto, algum método capaz de verificar essa relação, identificando se há ou não falhas de usabilidade na interface, ou seja, se as características comuns da interface interferem na percepção de imersão e presença do usuário influenciando de forma negativa ou positiva na sua experiência.

Em suma, esta pesquisa identificou uma carência de métodos de avaliação de interfaces para a faixa etária específica abordada, que atinja as necessidades específicas deste público, uma vez que não há padronização na escolha dos métodos utilizados. A análise mais aprofundada dos resultados de cada autor permitiu identificar também que, características de usabilidade precisam ser analisadas e de forma distinta para este público, pois foi constatado pelos autores que são fatores que influenciam a experiência do usuário idoso de forma diferente se comparado a outras faixas etárias, apontando para a necessidade de existir algum método que avalie a usabilidade da interface, no contexto específico do público idoso, que leve em consideração, também, as características específicas da dimensionalidade da interface. Dentro desse contexto, a pesquisa também apontou como tendência a avaliação da acessibilidade das interfaces, pois ~43% dos trabalhos avaliaram essa questão, o que se mostra bastante obvio levando em consideração as necessidades específicas da faixa etária.

Cabe ressaltar que nossas impressões e indicativos podem contribuir para definir etapas metodológicas capazes de avaliar não apenas a usabilidade, mas também outras questões relacionadas à idade avançada (como acessibilidade) e inovações tecnológicas disponíveis para idosos (como interfaces 3D). Esta pesquisa pode ajudar na disseminação e estabelecimento de aplicações adequadas para este público que, de acordo com a OMS [46], tende a aumentar rapidamente.

Tabela 4 – Tabela comparativa dos instrumentos utilizados para avaliação em cada estudo.

REFERÊNCIA	QUESTIONÁRIOS DE USABILIDADE	OUTROS QUESTIONÁRIOS	ENTREVISTA	OUTRO MÉTODO
Hagiya, Horiuchi e Yazaki	Não	Sim (n.e.)	Sim	Não
Tost <i>et al.</i>	Não	MEEM e Escala de Avaliação de Demência Clínica	Não	Não
Himmelsbach <i>et al.</i>	Não	Não	Sim	Técnica de Reconstrução do Dia, Técnica de Incidentes Críticos
Sacramento, Bacellar e Ferreira	Não	MAC-G	Sim	Não
Al-khomsan <i>et al.</i>	SUS	Não	Não	Ferramenta Morae e Software Mobizen
Konstantinidis <i>et al.</i>	SUS	SUMI	Não	WHOQoL-BREF
Hsieh e Huang	Não	Sim (n.e.)	Não	Não
Báez <i>et al.</i>	SUS	Groningen Frailty Indicator e outros questionários criados por eles (n.e.)	Não	Não
Schuurmans <i>et al.</i>	SUS	CEQ E CSQ-8	Sim	Não
Chiu <i>et al.</i>	Não	Attitudes Toward Computers Questionnaire e CES-D 10	Sim	Não
Meza-Cubo <i>et al.</i>	Não	Sim (n.e.)	Não	Observação das expressões dos participantes (vídeo)
Ferreira <i>et al.</i>	Não	Sim (n.e.)	Não	Think Aload
Epelde <i>et al.</i>	Não	Sim (n.e.)	Sim	Não
Teixeira <i>et al.</i>	ICF-US	Não	Não	Não
Le, Thompson e Demiris	SUS	Não	Não	Benchmark e Think Aload
Orso <i>et al.</i>	Nao	Sim (n.e.)	Nao	Nao

Tabela 5 – Resumo dos aspectos e abordagens de avaliação utilizados pelos estudos selecionados.

REFERÊNCIA	IDADE	INTERFACE	DISPOSITIVO	ASPECTO DE INTERAÇÃO AVALIADO	MÉTODO DE AVALIAÇÃO
Hagiya, Horiuchi e Yazaki	65-74	2D	Smartphone	Utilidade percebida e acessibilidade	Formativa, Questionário, Entrevista e Somativa
Tost <i>et al.</i>	50-80	3D	Interface touch	Eficiência	Formativa
Himmelsbach <i>et al.</i>	66-84	2D	Smartphone	Acessibilidade	Formativa, Entrevista
Sacramento, Bacellar e Ferreira	30-50 70-90	2D	n.e.	Acessibilidade e comunicabilidade	Somativa e Formativa
Al-khomsan <i>et al.</i>	60+	Ne	n.e.	Usabilidade	Formativa e Questionários
Konstantinidis <i>et al.</i>	60+	2D	Smartphone	Experiência do usuário, engajamento, usabilidade, e eficiência	Formativa e Questionários
Hsieh e Huang	65+	2D	Computador	Acessibilidade	Questionários
Báez <i>et al.</i>	65-87	3D	Tablet e sensores	Usabilidade, aceitação, acessibilidade e uso de terminologia	Questionários e Somativa
Schuurmans <i>et al.</i>	60+	2D	Tablet	Aceitação, satisfação, usabilidade	Caminhada Cognitiva, Questionários e Entrevista
Chiu <i>et al.</i>	50+	2D	Tablet ou Smartphone	Aceitação, acessibilidade e as atitudes	Formativa, questionários e entrevistas
Meza-Cubo <i>et al.</i>	60-83	2D	Headset, sensores e computador	Emoção	Formativa e questionário
Ferreira <i>et al.</i>	57-76	2D	Smartphone	Usabilidade e utilidade percebida	Formativa, Heurística e questionários
Epelde <i>et al.</i>	52-89	2D	SmartTV	Usabilidade e utilidade percebida	Demonstração, Questionário e entrevistas
Teixeira <i>et al.</i>	57-82	2D	Smartphone	Usabilidade e utilidade percebida	Formativa, Heurística e questionários
Le, Thompson e Demiris	60+	Ne	n.e.	Desempenho, experiência de interação e comunicação do usuário	Formativa, Somativa e Questionários
Orso <i>et al.</i>	60-82	2D	Computador e SmartTV	Engajamento/envolvimento e acessibilidade	Formativa, Caminhada Cognitiva, Questionário

4. MATERIAIS E MÉTODOS

Para o desenvolvimento da abordagem proposta por este trabalho, utilizaram-se os conceitos de avaliação de interfaces 3D catalogados por LaViola *et al.* [4]. De acordo com os autores, a avaliação de interfaces 3D é ainda uma área pouco explorada devido à recente inserção de tecnologias de RV no mercado, o que torna difícil prever a sua usabilidade sem o auxílio de usuários reais interagindo com tarefas representativas.

Vários conceitos e objetivos em avaliação de usabilidade de interface tradicionais e tridimensionais são similares, porém as abordagens de aplicação utilizadas no contexto de interfaces 3D geralmente são diferentes. Uma das consideráveis diferenças em sistemas de RV é a presença de dispositivos de interação não convencionais durante o processo interativo, como equipamentos multissensoriais que permitem ao usuário interagir na aplicação com movimentos naturais do corpo. Dispositivos de visualização imersiva, como *head-mounted displays* (HMDs), por exemplo, precisam ser avaliados por técnicas que não interfiram na experiência do usuário e no nível de imersão da interação, o que pode prejudicar a sensação de presença deste durante o processo interativo no ambiente virtual.

Diante da amplitude de variáveis a serem consideradas em um experimento para a avaliação da usabilidade de interfaces 3D, e com base na revisão apresentada pelo Capítulo 2, mostrou-se necessário desenvolver um método de avaliação de usabilidade que toma por base um protocolo de procedimentos (Tabela 6), englobando assim questões de especificidade de público e do processo interativo.

Tabela 6 – Estrutura do método proposta

DIA	ETAPA	PROCEDIMENTOS	DURAÇÃO MÉDIA
1	Pré-Teste	Questionários	15 minutos
		Treinamento	5 minutos
2	Teste	Interação	1 minuto
		Descanso (opcional)	2 minutos
	Pós-Teste	Questionários	5 minutos
		Entrevista	10 minutos

Devido a especificidade do público alvo (idosos) e consequente necessidade em se aplicar uma ferramenta específica para avaliação de interfaces 3D [4], somado ao fato da RS

não ter detectado instrumento para avaliação nesse contexto específico, desenvolveu-se também um questionário para avaliação da usabilidade para integrar a etapa pós-teste do método. Buscou-se desenvolver uma ferramenta que pudesse identificar as variáveis a serem avaliadas e as características de usabilidade definidas como importantes pelos autores da literatura [1] [3] [4] [16], considerando situações do processo interativo que interfaces 3D propiciam. Tal necessidade se justifica por tratar-se de um grupo que apresenta determinadas necessidades específicas ou, em alguns casos, não apresenta grande familiaridade com tecnologias de RV. Portanto, a criação de uma ferramenta de contexto específico para este público, que abranja questões específicas de entendimento e de abstração, pode garantir, por exemplo, que o método consiga ser claramente entendida pelos sujeitos, minimizando erros e vieses na pesquisa por desentendimentos semânticos.

Conforme protocolo apresentado na Tabela 6, o método proposto se divide em três etapas: “Pré-teste”, “Teste” e “Pós-teste”, detalhado nas próximas seções. Recomenda-se que a aplicação dos procedimentos se divida em duas sessões, com duração máxima de 20 minutos cada. O objetivo é não tornar a sessão de avaliação cansativa, tampouco tediosa para os sujeitos. Por este motivo, organizou-se o método de acordo com os tempos médios descritos na Tabela 6, porém os tempos podem ser reajustados de acordo com as necessidades do contexto em que o método for aplicado, diminuindo os tempos de uma etapa para aumentar os tempos de outra, quando for necessário. Por ser direcionada a um público que pode estar tendo contato pela primeira vez com tecnologias de RV, esses detalhes podem colaborar para a obtenção de resultados mais precisos. Assim, a etapa de “Pré-teste” é realizada na primeira sessão, e as etapas de “Teste” e “Pós-teste” são realizadas na segunda sessão.

4.1. PRÉ-TESTE

Neste momento ocorre o primeiro contato do observador do experimento com o sujeito. O condutor deve explicar o experimento, como ele será conduzido e seu objetivo, pedindo-lhe também que leia e assine o Termo de Consentimento Livre e Esclarecido (TCLE)(Apêndice A). Ademais, esta etapa objetiva selecionar e caracterizar os participantes, verificando seu conhecimento prévio da tecnologia e limitações físicas e cognitivas que possam vir a influenciar no teste de usabilidade. Para tanto, utilizam-se os seguintes instrumentos:

- Escala de Depressão Geriátrica – GDS-15 [51] [52] (Anexo B);
- Mini Exame do Estado Mental – MEEM [53] [54](Anexo A);

- Questionário Sociodemográfico e de Caracterização da Amostra [18] (Apêndice B).

Os scores aplicados para os questionários MEEM e GDS-15 foram definidos de acordo com a literatura [51] [52] [53] [54].

O formulário GDS-15 é uma avaliação superficial do participante para analisar se possui algum grau de depressão. De acordo com Sheikh e Yesavage [52], um sujeito depressivo tende a apresentar dados não confiáveis, pois há possibilidade de que seu estado psicológico interfira no resultado no teste. Portanto, este teste é utilizado para selecionar participantes que apresentam graus severos de depressão, não indicados para continuar o experimento. As perguntas do questionário verificam se o participante está satisfeito consigo mesmo e com a sua vida, respondendo apenas “sim” ou “não”. O ponto de corte adotado foi cinco, definindo não caso/caso, ou seja, conta-se um ponto a cada resposta negativa para o sentido da pergunta (não necessariamente um ponto para cada “não” respondido) para sujeitos saudáveis. Para casos de depressão leve, o ponto de corte adotado foi 10 [55].

O MEEM é um teste rápido para avaliar a função cognitiva do usuário (em torno de sete minutos). De fácil aplicação, não requer material específico e é medido através de pontuações. Assim como o GDS-15, serve para seleção de participantes e que, por motivos cognitivos, não são indicados para seguir no experimento. Perguntas deste questionário envolvem questões de orientação espacial, temporal, memória imediata e de evocação, cálculo, linguagem-nomeação, repetição, compreensão, escrita e cópia de desenho. O ponto de corte definido foi 25 para idosos com escolaridade, 18 para idosos com baixa e média escolaridade e 13 para analfabetos [53] [56].

O questionário sociodemográfico e de caracterização da amostra tem o objetivo de coletar informações gerais do público alvo. As perguntas do questionário envolvem questões sobre escolaridade, dificuldades físicas e perguntas que denotam o grau de familiaridade que o participante tem com as tecnologias que estão sendo testadas. As perguntas podem ser alteradas de acordo com o tipo de jogo ou aplicação a que for aplicado, no entanto recomenda-se que envolva, no mínimo, as seguintes questões:

- Escolaridade: para fins de pontos de corte;
- Grau de conhecimento prévio da tecnologia utilizada;

- Limitações físicas: para exclusão de indivíduos que não sejam considerados aptos a realizar as tarefas que o jogo propõe, ou inclusão de indivíduos, quando os critérios de inclusão sejam que o indivíduo apresente determinada dificuldade.

O TCLE é um termo que explica ao participante que ele tem a liberdade de não fazer o teste e desistir antes do término. Ele também esclarece que as informações coletadas serão exclusivamente para estudo e que sua integridade estará preservada.

Após o preenchimento dos questionários, os participantes dispõem de um tempo médio de cinco minutos para treinar com os equipamentos que serão utilizados e o sistema a ser avaliado na etapa “Teste”. Eles recebem explicações sobre o funcionamento do sistema de RV e podem interagir livremente com a aplicação. Este procedimento se mostra necessário para que a interação da etapa “Teste” não obtenha o ônus da novidade, e de modo que o participante seja mais crítico quanto ao funcionamento do sistema [18]. Além disso, auxilia a evitar quebras de presença na imersão dos participantes durante o experimento da etapa “Teste”, porque permite que no treinamento as dúvidas referentes aos equipamentos e ao processo interativo sejam sanadas. Este procedimento foi inserido no método devido aos apontamentos de Simor [7] e Brum e Rieder [18]; na literatura pesquisada não foi encontrado qualquer referência ao tempo ideal ou quaisquer outros detalhes a serem levados em consideração ao executar esta etapa.

4.2. TESTE

Nesta etapa os participantes apenas interagem com o sistema em avaliação, realizando tarefas de interação específicas, orientados a utilizar o protocolo Think Aloud [3]. Desta forma o observador poderá identificar dificuldades e impressões dos participantes, tomando nota durante o processo interativo.

Recomenda-se que o tempo do procedimento não ultrapasse dois minutos por tarefa de interação, especialmente nos casos de tarefas que trabalham com repetição de movimentos [7] [18]. Tratando-se de usuários idosos, um tempo muito prolongado pode cansar o participante, correndo o risco de desestimular sua sinceridade e a continuidade da participação na pesquisa, ou ainda causar descontentamento com a tarefa, interferindo nas suas impressões sobre a usabilidade do sistema.

4.3. PÓS-TESTE

Nesta etapa, o participante transmite suas impressões e opiniões a respeito da interface e da interação por meio de dois questionários e uma entrevista semiestruturada. Os questionários propostos são:

- System Usability Scale - SUS [57] (Anexo C);
- 3DUI Usability Evaluation Questionnaire for Elderly – UQE 3D (Apêndice C).

O Questionário SUS foi incluído no método proposto por ser o instrumento mais utilizado nos resultados da RS – embora não seja preparado para avaliar a usabilidade de interfaces 3D, tampouco projetado para um público específico como os idosos. No próximo capítulo, discutir-se-á a validade de mantê-lo ou não no método proposta, em virtude do nível de especificidade de nossa abordagem.

O UQE 3D é o questionário desenvolvido especificamente para a proposta desse método, e tem por objetivo avaliar interfaces 3D considerando o contexto específico de público idoso. Esta ferramenta foi pensada para ser aplicada diretamente em experimentos que envolvam sujeitos com 60 anos ou mais [12], e também revisada por quatro profissionais: dois da área de Saúde, que têm contato frequente com idosos, e dois da área da Computação. A revisão visou a adequação de termos ao contexto específico do público alvo e verificação dos sentidos das afirmações.

Uma parte do UQE 3D busca identificar os seguintes aspectos em relação à interface: **conforto** e **bem-estar**, **imersão**, **presença** e **percepção** (elementos visuais e sonoros intuitivos). Tais aspectos investigados baseiam-se em métricas de avaliação de interfaces 3D consideradas importantes por LaViola *et al.* [4] e nos princípios básicos de RV definidos por Burdea e Coiffet [16].

Para LaViola *et al.* [4], interfaces 3D devem ser intuitivas, oferecendo recursos intuitivos e de bom feedback; além e não serem intrusivas, propiciando a sensação de conforto e bem-estar no uso (minimizando o *cybersickness*). Os autores ainda salientam que presença e imersão são conceitos essenciais para avaliação de interfaces 3D.

Burdea e Coiffet [16] definem três conceitos primordiais da RV como sendo: Imersão, Interação e Imaginação. Huang [58] analisa e explica esses conceitos afirmando que a Imersão se divide entre imersão mental e sensorial (a primeira congruente ao que definiu LaViola *et al.* [4], sendo uma consequência da segunda, que se dá pela interação com a interface por meio de estímulos sensoriais). Já a Interação promove a sensação de imersão,

uma vez que profere a reciprocidade do sistema em contato com o usuário. E Imaginação tem relação com o envolvimento, pois estimula a abstração e a capacidade da mente humana de perceber e imaginar de forma criativa, de acordo com os estímulos recebidos.

Outra parte do UQE 3D considera métricas de usabilidade de interface definidas por Nielsen [3] [17]: **facilidade de utilizar a aplicação, facilidade de realizar a tarefa, clareza do procedimento a ser realizado, e relação entre cenário e tarefa**. Ainda, optou-se por acrescentar as variáveis **tempo de execução da tarefa e intervalos de descanso**, para identificar se a duração das tarefas (ou descanso) são consideradas satisfatórias e/ou suficientes pelos participantes. O questionário utiliza a Escala Likert com cinco níveis de resposta.

Cabe salientar que as afirmações do UQE 3D são adaptáveis para o contexto do experimento. Por exemplo, para esse experimento, a questão 7, que se refere à movimentos corporais realizados durante o processo interativo, foi ajustada para “movimentos com os braços” – uma vez que a tarefa realizada pelo grupo de participantes envolvia apenas este tipo de interação. Outros exemplos de adaptação, como “movimento com as pernas” e “movimentos com as mãos”, podem ser utilizados conforme o objetivo da avaliação.

Isso também ocorre com relação ao tipo de interface que se utiliza na avaliação; no caso em que a interface 3D não seja especificamente um jogo. Termos como “sistema”, “ambiente virtual” ou “simulação” podem ser utilizados, desde que o usuário esteja familiarizado com a expressão. De acordo com Günter [12] é necessário adaptar a linguagem de um questionário para que este atenda melhor as demandas de entendimento do público alvo.

Portanto, o UQE 3D foi desenvolvido visando as características de usabilidade e de RV já mencionadas, de forma que possam ser adaptadas para cada contexto em que for aplicado (jogo, simulação, etc.). Como o experimento desenvolvido para validação utilizará um *serious game*, o questionário está adaptado para o contexto “jogo”.

A Tabela 7 mostra as afirmações e o Apêndice C apresenta o questionário. A Figura 13 representa as relações entre as afirmações do questionário e as definições dos autores. Além disso, utilizou-se como referência o questionário desenvolvido por Simor [7]. A principal diferença entre o questionário SUS e o UQE 3D é a especificidade do UQE 3D para sistemas de RV, uma vez que o questionário avalia variáveis de 3DUI, que são fracamente avaliadas pelo SUS, ou mesmo, não são avaliadas, como é o caso das variáveis Imersão e Presença (Figura 14). Por este motivo, discutir-se-á a necessidade de manter o questionário SUS neste método.

Tabela 7 – Perguntas do UQE 3D e suas respectivas relações com os princípios e características definidas pelos autores [3] [4] [16] [17]

Nº	QUESTÃO	VARIÁVEIS E CARACTERÍSTICAS INVESTIGADAS
1	Senti-me confortável com os equipamentos durante o jogo.	Conforto [4] [16]
2	Senti-me parte integrante do jogo enquanto jogava.	Imersão e Presença [4] [16]
3	O jogo provocou situações de mal-estar como enjoo, dor de cabeça,	Bem-estar [4] [16]
4	Eu sabia onde estava no espaço geográfico dentro do jogo.	Imersão, Presença e Percepção Visual [4] [16]
5	Foi fácil me localizar no espaço geográfico dentro do jogo.	Percepção Visual [4] [16] e Facilidade em utilizar a aplicação [3] [17]
6	Foi fácil realizar as tarefas do jogo.	Facilidade em utilizar a aplicação e Facilidade em realizar a tarefa [3] [17]
7	Cansou-me fazer movimentos durante o jogo.	Bem-estar [4] [16] e Tempo de execução
8	Usar o capacete de visualização foi bom para o jogo.	Percepção visual, Imersão e Elementos visuais [4] [16]
9	As imagens me ajudaram a entender o jogo.	Relação entre cenário e tarefa [3] [17], Elementos visuais e Percepção visual [4] [16]
10	As imagens me ajudaram a entender como jogar.	Clareza do procedimento a ser realizado [3] [17], Relação entre cenário e tarefa [3] [17], Percepção visual e Elementos visuais [4] [16]
11	Os sons emitidos pelo jogo me ajudaram a entender como jogar.	Clareza do procedimento a ser realizado [3] [17] e Elementos sonoros [4] [16]
12	Achei o jogo divertido.	Imersão e Bem-estar [4] [16]
13	Gostei de interagir por gestos no jogo.	Imersão e Bem-estar [4] [16]
14	Achei o jogo adequado para minha idade.	Bem-estar e Presença (indiretamente) [4] [16]
15	Achei suficiente o tempo que tive para jogar.	Tempo de execução
16	Achei suficiente o tempo que tive para descansar após o jogo.	Tempo de descanso

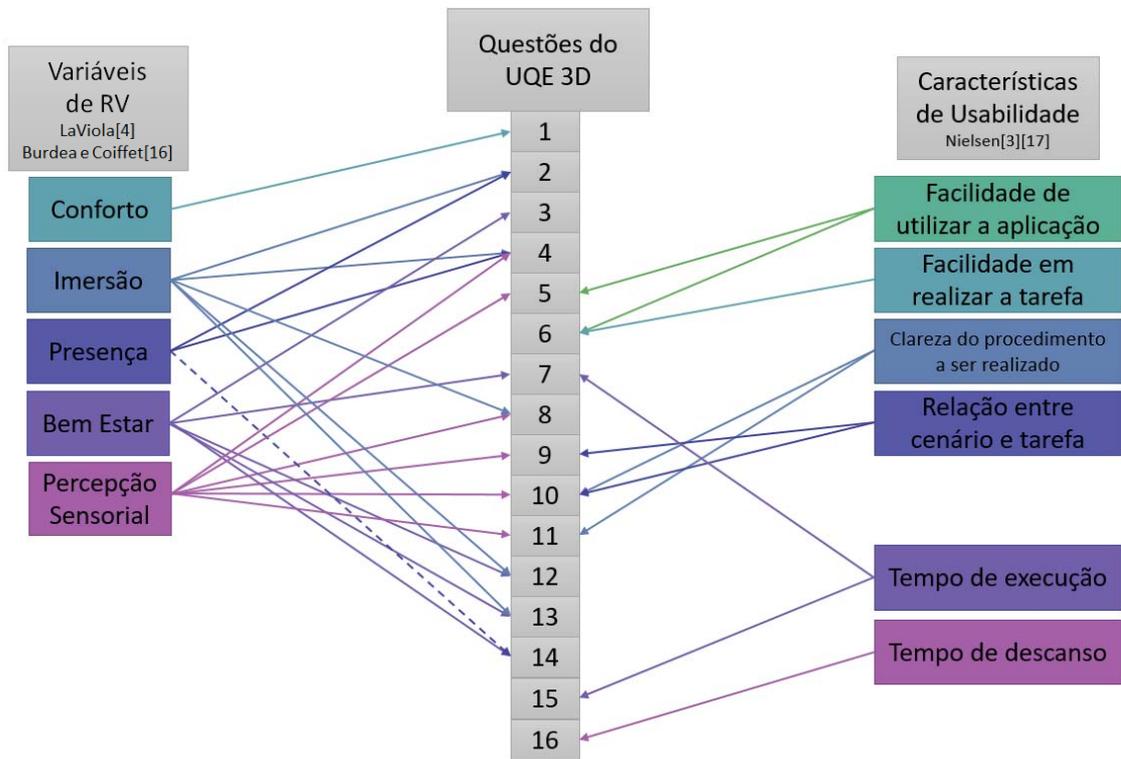


Figura 13 – Relação entre os princípios e características definidas pelos autores [3] [4] [16] [17] com as perguntas do UQE 3D.

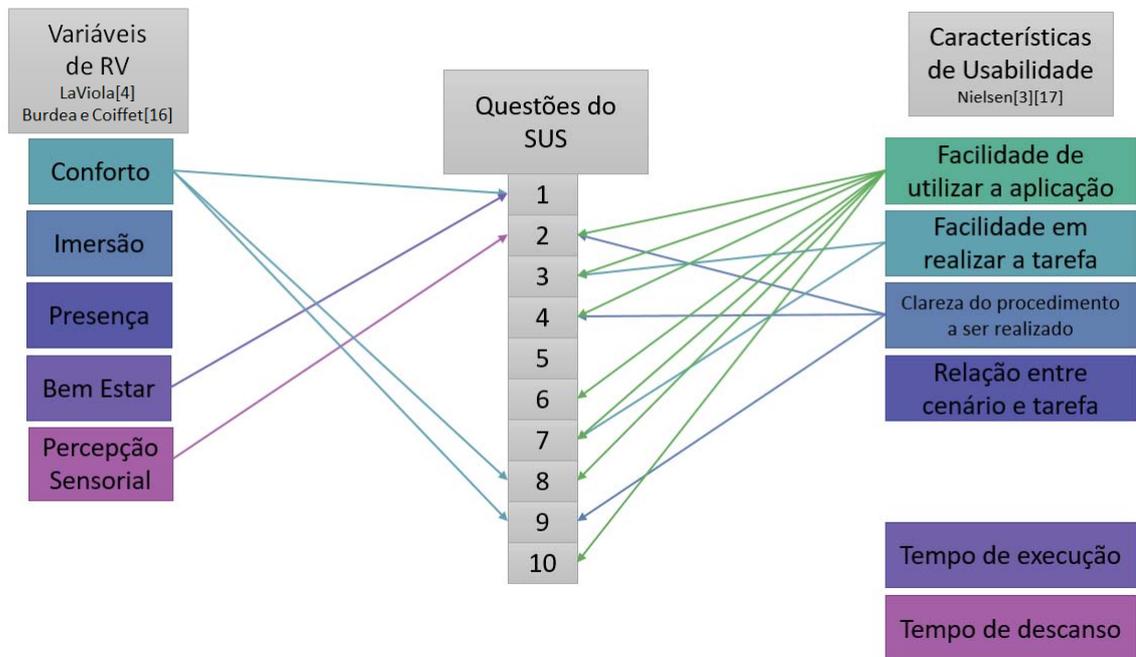


Figura 14 – Relação das perguntas do SUS com os princípios e características definidas pelos autores [3] [4] [16] [17].

Durante o preenchimento, o participante pode comentar sobre o teste abertamente, permitindo uma coleta de informações complementares por parte do observador.

Uma vez que o UQE 3D abrange boa parte das informações que se objetiva captar dos participantes, adotou-se também ao método a aplicação de uma entrevista semiestruturada para validar as respostas do questionário. Para tanto, formulou-se perguntas semelhantes para obter respostas sugestivas às mesmas variáveis e características investigadas com o questionário.

A Tabela 8 mostra as perguntas da entrevista, bem como a relação entre o UQE 3D e a entrevista. Algumas perguntas da entrevista não se relacionam textualmente e de forma direta com o UQE 3D, porém, com base nas respostas dos participantes, pode-se entender sua experiência e associar com as afirmações do questionário. Por exemplo, se a resposta para a pergunta 4 da entrevista for afirmativa, significa que o usuário se sentiu imerso no jogo, pois conseguiu sentir-se presente no ambiente e se localizar no cenário (afirmações 4 e 5 do UQE 3D); também significa que a interface está intuitiva (afirmações 9, 10 e 11 do UQE 3D), pois ele saberia facilmente como utilizar o jogo, sem precisar recorrer a outras pessoas. Se houver incoerência quanto às respostas do questionário é possível realizar uma análise subjetiva do comportamento do usuário enquanto interagia e pode-se levar em consideração as pontuações do MEEM e GDS-15 e as respostas destes usuários ao questionário sociodemográfico e de caracterização da amostra para identificar quais fatores podem ter influenciado a discrepância.

As entrevistas são de suma importância para o experimento devido ao contexto específico ao qual ele se aplica. Em geral, pode-se captar informações adicionais valiosas do público idoso por intermédio de entrevistas, sobretudo tratando-se de aplicações com dispositivos de RV para pessoas mais longevas. A entrevista ajuda o pesquisador a identificar desentendimentos de semântica em relação aos termos referentes à tecnologia e a explicá-los, pois, por vezes, eles não fazem parte do cotidiano dos participantes. É uma forma de assegurar o melhor entendimento da sensação que o usuário vivenciou ao longo do processo interativo.

Tabela 8 – Perguntas da Entrevista Semiestruturada e suas relativas questões do UQE 3D

Nº	QUESTÃO DA ENTREVISTA	QUESTÕES RELACIONADAS DO UQE 3D
1	Você gostou de utilizar seus braços para interagir com o jogo?	7 e 13
2	Você gostou de usar o capacete de visualização? Por quê?	2 e 8
3	Se você tivesse um jogo como esse, teria vontade de jogá-lo?	12
4	Você conseguiria utilizar esse jogo sozinho, sem ninguém para explicar como o jogo funciona?	4, 5, 9, 10 e 11
5	Você conseguiu entender o que fazer dentro do jogo? O que dificultava?	2, 9, 10 e 11
6	Teve algum momento em que você ficou em dúvida sobre o que fazer dentro do jogo?	9, 10 e 11
7	O que você achou cansativo no jogo?	7
8	O jogo é adequado à sua faixa etária? Que pontos podem ser melhorados?	14
9	O jogo estimulou a prática de atividades benéficas ao corpo humano? Física ou mental?	–
10	O jogo contribuiu para aumentar o seu interesse em ferramentas computacionais para educação, bem-estar e saúde?	–
11	O jogo permitiu aumentar a sua conexão com outras pessoas que usam tecnologia, como jovens ou outros idosos?	–

5. RESULTADOS E DISCUSSÕES

Este método vem sendo formulado desde os trabalhos de Simor [7] e Brum e Rieder [18]. Portanto, foi elaborado usando, além de todo referencial teórico mencionado, os resultados dos experimentos preliminares conduzidos por esses autores.

Buscando uma validação concludente, aplicou-se o método proposto nesse trabalho em um experimento envolvendo idosos e tecnologias de RV. Os resultados do experimento são detalhados nas próximas subseções.

5.1. USO DO MÉTODO

Para melhor desenvolvimento didático, esta subseção se refere à aplicação do método a um experimento com o objetivo de testá-lo e avalia-lo. Os resultados dessa avaliação e a discussão dos resultados encontram-se na próxima subseção.

5.1.1. Participantes

De acordo com Benyon [59], um número reduzido de participantes para o experimento pode ser justificado, se a avaliação se destina a um público alvo relativamente homogêneo. De acordo com Nielsen [60], 20 participantes são suficientes para se obter números estatisticamente significativos, enquanto que ao se tratar de uma pesquisa qualitativa de usabilidade, cinco participantes já são suficientes.

Dada a especificidade do contexto da pesquisa, e consequente dificuldade em se conseguir um considerado número de sujeitos, foram recrutados voluntariamente 30 participantes com 60 anos ou mais ($71,4 \text{ anos} \pm 10,29$). Buscou-se compor grupos de idosos de diferentes contextos para verificar a amplitude de aplicabilidade do método. De acordo com Bertolucci et al. [53] existe diferença entre idosos oriundos de diferentes contextos, e essa diferença se relaciona não com a idade do sujeito, mas, em geral, de acordo com a sua escolaridade [47]. Assim, a avaliação foi dividida em dois grupos:

- Grupo A, Idosos Ativos: indivíduos que conduzem atividades diariamente, participavam de programas sociais específicos para sua idade, e sem qualquer nível de demência, com escores MEEM > 18 e GDS-15 < 5 ;

- Grupo B, Idosos Não Ativos: indivíduos institucionalizados, e que apresentam algum tipo de declínio cognitivo leve ou sintomatologia depressiva leve, com escores MEEM > 18 e GDS-15 < 10.

Os critérios de inclusão comuns a ambos os grupos consideraram que os indivíduos deveriam ser alfabetizados e sem comprometimento motor.

O Grupo A foi composto por 23 idosos (67,47 anos \pm 6,45), quatro homens e 19 mulheres, e o Grupo B foi composto por sete participantes (84,28 anos \pm 10,32), dois homens e cinco mulheres.

5.1.2. Tarefa

Para validar a abordagem proposta, realizou-se uma avaliação com um dos níveis de jogo do Motion Rehab AVE 3D (BR 51 2016 001373-7) [5] [13], um *serious game* para auxiliar profissionais de saúde em atividades de reabilitação motora e cognitiva de idosos, com suporte a RV (visualização imersiva e interação espacial). A Figura 15 mostra um dos cenários do jogo, com a visão do jogador em primeira pessoa. O UQE 3D está, portanto, com terminologias adaptadas para o contexto da aplicação – jogo – e de acordo com a tarefa de interação (Figura 15) e com os dispositivos utilizados (Figura 18).



Figura 15 - Visão da interface do jogo Motion Rehab AVE 3D em primeira pessoa.

A tarefa selecionada para o experimento consiste em utilizar os braços para encostar com mãos virtuais em bolas de vôlei de praia (objetos de contexto) que vêm em direção ao personagem. Em alguns momentos, lápis (objetos distratores) são apresentados no lugar das bolas, exigindo que o jogador não encoste com as mãos virtuais (pode baixar os braços ou desviar-se corporalmente, Figura 16). A Figura 17 ilustra o movimento corporal dos idosos durante o processo interativo no jogo.

Para esta avaliação, o participante pode interagir várias vezes na mesma fase do game (o nível escolhido do jogo dura 30 segundos), não superando o tempo de dois minutos de interação, estabelecidos pelo protocolo do método.

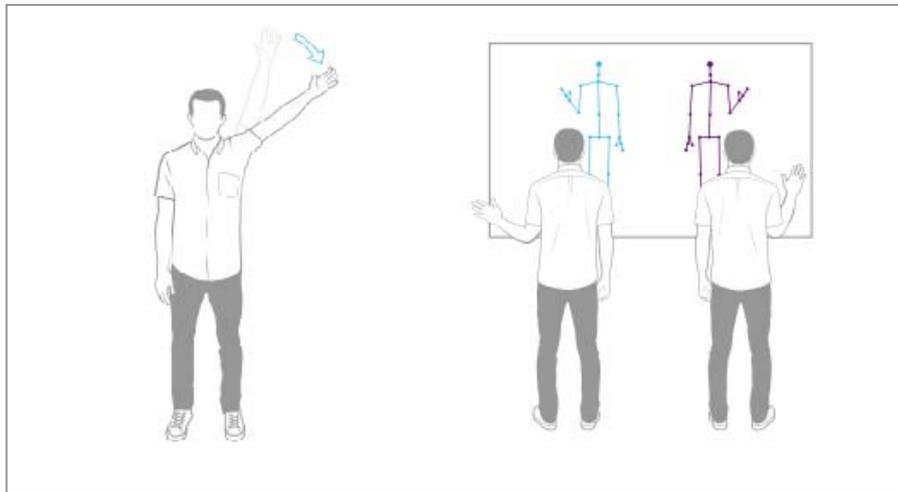


Figura 16 – Exemplos de movimentos para interagir com a interface [61].



Figura 17 – Idosos interagindo com o sistema no experimento de avaliação.

5.1.3. Dispositivos

O dispositivo de visualização utilizado é o HMD modelo Oculus Rift DK 1 (Figura 18, esquerda), um capacete de RV com rastreamento de cabeça que permite a visualização do espaço 3D em 360°. Para capturar os movimentos gestuais e corporais do usuário, utilizou-se o sensor de movimentos Microsoft Kinect One (Figura 18, direita).



Figura 18 – Dispositivos utilizados para visualização e interação no ambiente virtual.

5.2. AVALIAÇÃO DO MÉTODO - VALIDAÇÃO E DISCUSSÃO

Para validar o método proposto, primeiramente buscou-se validar o UQE 3D. Para tanto, foi necessário verificar se o instrumento proposto expressava de forma clara o que buscava identificar, se possuía coerência interna e se atingia o seu público alvo.

De acordo com Hoss e Caten [62], a avaliação interna de um instrumento de medição atribui-lhe um conceito de qualidade. Esta qualidade é mensurada em duas características essenciais: confiabilidade e validade [63]. A validade pode ser descrita como o grau em que os elementos do instrumento representam o conceito a ser avaliado e são relevantes para tal. Confiabilidade se refere à precisão da mensuração independente do que se é medido. A consistência interna, como medida de confiabilidade de um questionário, representa quão diferente os itens medem o mesmo conceito [64].

Após a avaliação da consistência interna, foi calculado um escore para o UQE 3D baseado no cálculo de escore do questionário SUS, e, em seguida, um Intervalo de Confiança para aplicar ao UQE 3D como medida de avaliação de escores dos indivíduos do experimento.

Assim, foi possível executar os testes T de Student e F de Fisher para comparar os grupos, utilizando os escores oriundos da etapa anterior.

Finalmente, utilizou-se a técnica de Distribuição e Frequência para avaliar as respostas dos indivíduos e relacioná-las com as respostas da entrevista. Esta comparação foi realizada para reafirmar as respostas obtidas pelo questionário, e também como mecanismo de identificação da adequação do instrumento ao público alvo.

5.2.1. Alfa de Cronbach, KMO e Bartlett

Para avaliar a consistência interna do questionário desenvolvido, utilizou-se a técnica de Análise Fatorial aplicada de forma exploratória e, como confirmação, utilizou-se o Alfa de Cronbach [62]. Foram definidas duas hipóteses:

- H0: Os dados não são adequados para a análise fatorial;
- H1: Os dados são adequados para a análise fatorial.

Com a aplicação das técnicas citadas, o UQE 3D obteve o valor de 0,669 para o teste de KMO e 0 (zero) para o teste de Bartlett. Ambos os testes verificam os graus de ajuste dos dados à análise fatorial. Da mesma forma, o Alfa de Cronbach avalia a consistência interna do questionário verificando se os itens têm coerência entre si. O UQE 3D obteve pontuação de 0,868 para o alfa de Cronbach. Ambos os testes foram realizados com população de 30 indivíduos.

Os valores aceitáveis para o teste de KMO devem ser maiores que 0,5 para indicarem adequação ao método, enquanto para o teste de Bartlett devem ser menores que 0,1 para negar a hipótese nula [62]. Os valores de referenciais para o Alfa de Cronbach podem ser visualizados na Tabela 9 [65].

Tabela 9 – Valores de referência para consistência interna do questionário. [65]

Alfa de Cronbach	Consistência interna
Maior do que 0,80	Quase perfeito
De 0,80 a 0,61	Substancial
De 0,60 a 0,41	Moderado
De 0,40 a 0,21	Razoável
Menor do que 0,21	Pequeno

De acordo com os resultados, pode-se perceber que o UQE 3D apresenta boa consistência interna, com dados adequados à análise. Além disso, o teste não mostrou necessidade de se excluir nenhuma das questões do questionário.

5.2.2. Intervalo de Confiança

Após a validação interna do UQE 3D, foi definido um intervalo de confiança para o questionário, baseado nos resultados do experimento. Para tanto, foi utilizado a mesma lógica que John Brooke utilizou para definir a média do questionário SUS [66], que é considerada uma forma intuitiva de cálculo [67].

De acordo com essa lógica, primeiramente, é necessário encontrar um valor do fator multiplicativo para o questionário. Para isso, simula-se uma avaliação com o questionário considerando sempre a melhor resposta possível para cada questão, em uma escala Likert de 1 a 5. Desta forma, para cada questão do questionário faz-se a subtração do valor 1 da resposta do sujeito, para os casos em que as afirmações são positivas ($x - 1$, sendo que x é 5 para a melhor resposta esperada); e para os casos em que a afirmação é negativa, subtrai-se a resposta do sujeito de 5 ($5 - x$, sendo que x é 1 para a melhor resposta esperada). De acordo com essa lógica, todas as questões podem ter como resultado 4 para a melhor resposta esperada. A partir daí, soma-se todos esses valores e obtém-se o escore máximo possível. O fator multiplicativo (que para o questionário SUS é 2,5) é o resultado de 100 dividido pelo escore máximo ($100 / \text{escore máximo}$). A Figura 19 esquematiza o cálculo realizado e a Tabela 10 mostra a simulação.

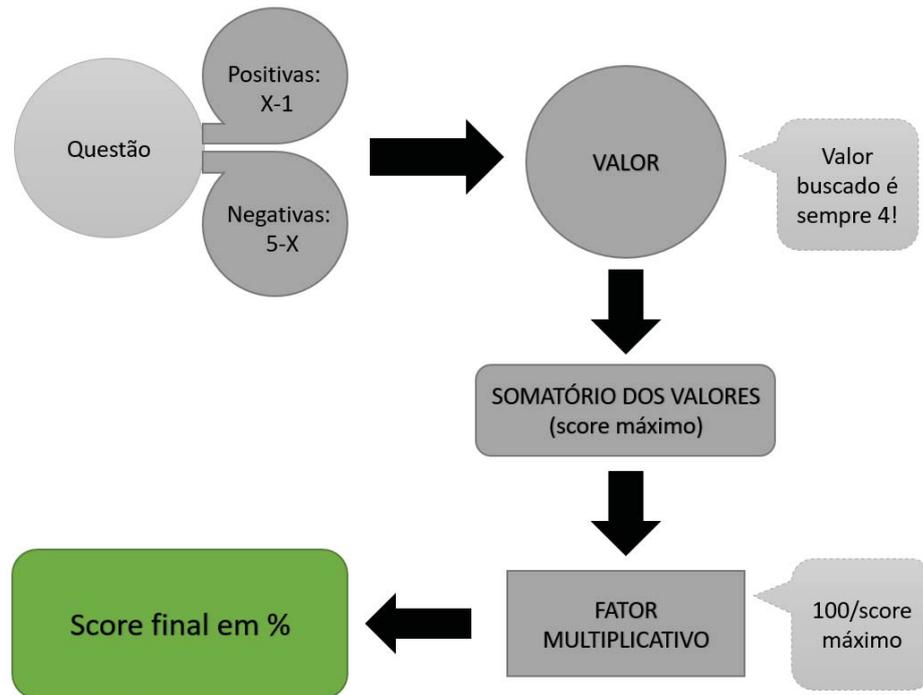


Figura 19 – Esquema explicativo da lógica para cálculo do SUS, replicada no UQE 3D.

Dessa forma, ao realizar o cálculo de um sujeito, basta multiplicar o seu escore (somatório dos valores de cada questão, resultantes do primeiro cálculo: $x - 1$ para as positivas e $5 - x$ para as negativas) pelo fator multiplicativo do questionário para obter o seu escore final, entre 1 e 100.

Após identificado o fator multiplicativo para o UQE 3D, buscou-se um intervalo de confiança para aplicar ao questionário uma medida de qualidade. Para definir uma média para o questionário (semelhante como existe para o SUS), seria necessária uma amostra maior. Portanto, optou-se por desenvolver um Intervalo de Confiança.

Para o cálculo do intervalo foi utilizado $\alpha = 0,05$ e os dados obtidos do experimento, para os Grupos: A – Idosos Ativos, B – Idosos Institucionalizados e Geral (A + B). A Tabela 11 mostra os dados resultantes da análise estatística para cada um dos grupos e seus respectivos Intervalos de Confiança.

Pode-se perceber na Tabela 11 que existe considerável diferença entre os Escores Médios (média amostral) do Grupo A (Idosos Ativos) e do Grupo B (Idosos Institucionalizados). Desta forma, um Intervalo de Confiança definido de forma genérica, considerando os dois grupos (Grupo Geral) pode não ser considerado adequado a indivíduos com características semelhantes aos pertencentes ao Grupo B deste experimento. Portanto,

para fins de aplicação do UQE 3D, recomenda-se utilizar os Intervalos de Confiança de forma específica para cada grupo de usuários (Idosos Ativos e Idosos Não-Ativos), quando as características do grupo de usuários se adequem às características dos grupos deste experimento. Caso exista um grupo de usuários com características não condizentes com nenhum dos grupos aplicados, pode-se desenvolver um novo intervalo de confiança utilizando os mesmos parâmetros de confiabilidade abordados aqui.

Tabela 10 – Simulação do melhor caso, utilizada para definir o score do questionário UQE 3D.

Questão	Sentido da Questão	Cálculo	Resposta Esperada	Score
1	(+)	$x - l$	5	4
2	(+)	$x - l$	5	4
3	(-)	$5 - x$	1	4
4	(+)	$x - l$	5	4
5	(+)	$x - l$	5	4
6	(+)	$x - l$	5	4
7	(-)	$5 - x$	1	4
8	(+)	$x - l$	5	4
9	(+)	$x - l$	5	4
10	(+)	$x - l$	5	4
11	(+)	$x - l$	5	4
12	(+)	$x - l$	5	4
13	(+)	$x - l$	5	4
14	(+)	$x - l$	5	4
15	(+)	$x - l$	5	4
16	(+)	$x - l$	5	4
Score máximo				64 (100%)
Fator multiplicativo UQE 3D				100/64 = 1,5625

Tabela 11 - Dados da análise estatística para definição do Intervalo de Confiança.

Dados Estatísticos	Grupo A	Grupo B	Geral
Alfa (α)	0,05	0,05	0,05
Desvio Padrão Amostral	9,810082845	9,059677311	14,45357399
Amostra	23	7	30
Média Amostral	88,52	63,17	82,60
Intervalo de Confiança	[84,28 ; 92,76]%	[54,79 ; 71,55]%	[77,81 ; 88,00]%

5.2.3. Teste T de Student e Teste F de Fisher

Ainda para a validação do UQE 3D, buscando-se identificar se ele pode ser considerado apropriado, do ponto de vista estatístico, utilizou-se o Teste T de Student [68] para comparar se existe diferença significativa entre o SUS e o UQE 3D para os Grupos A, B e Geral. Para tanto, utilizou-se os escores médios de cada questionário (SUS e UQE 3D), que é factível, uma vez que os escores são calculados de maneira semelhante.

Como pode-se observar na Tabela 12, as comparações entre os grupos A, B e Geral mostraram que não existe diferença significativa entre os resultados da aplicação do SUS e do UQE 3D para qualquer um dos casos. Este resultado reforça o pressuposto de que o UQE 3D é equivalente ao SUS ao trabalhar as questões de usabilidade, com a diferença de que engloba conceitos específicos referentes a interfaces 3D e dispositivos não-convencionais de interação.

Devido à diferença ligeiramente maior nos valores de média e desvio padrão apresentada entre o SUS e o UQE 3D na análise sobre o grupo Geral, em relação aos outros grupos, realizou-se ainda – apenas sobre este grupo – o Teste F de Fisher [69]. Este teste compara duas variâncias populacionais para analisar a diferença entre elas, levando em consideração os valores de média e de desvio padrão. De acordo com o Teste F de Fisher, o desvio padrão amostral apresentado pelo subgrupo SUS é considerado estatisticamente igual ao desvio padrão amostral apresentado pelo subgrupo UQE 3D, ou seja, a diferença apresentada entre eles não é considerada estatisticamente significativa, levando em consideração o tamanho da amostra.

Tabela 12 – Estatísticas de grupo para o Teste T de Student.

Grupo A			
Questionário	Amostra	Média	Desvio Padrão
SUS	23	88,0435	8,08184
UQE 3D	23	88,5200	9,80923
<i>p = 0,8580 > 0,05</i>		Não existe diferença significativa	
Grupo B			
Questionário	Amostra	Média	Desvio Padrão
SUS	7	71,4286	10,78745
UQE 3D	7	63,1700	9,05828
<i>p = 0,1470 > 0,05</i>		Não existe diferença significativa	
Grupo Geral			
Questionário	Amostra	Média	Desvio Padrão
SUS	30	84,1667	11,16748
UQE 3D	30	82,6050	14,45316
<i>p = 0,6410 > 0,05</i>		Não existe diferença significativa	

O valor de p calculado para o grupo dos idosos ativos (Grupo A) foi o maior em comparação com os outros grupos. Isso pode se justificar pela maior facilidade com que os idosos ativos entendem o processo interativo, e a facilidade com que executam o processo de abstração das afirmações dos questionários em relação à experiência proporcionada pelo ambiente virtual.

5.2.4. Distribuição e Frequência

As análises estatísticas realizadas mostraram que existe diferença entre os dados oriundos dos experimentos realizados com o Grupo A (Idosos Ativos) e com o Grupo B (Idosos Institucionalizados). Portanto, optou-se por fazer a análise da distribuição e frequência das respostas do UQE 3D separadamente, de acordo com cada um dos grupos, permitindo estabelecer uma relação entre eles, além da relação com as respostas da entrevista, que estão resumidas na Tabela 13.

Na Tabela 14 e na Tabela 15 é possível verificar a distribuição e frequência das respostas ao UQE 3D, considerando o Grupo A e o Grupo B, respectivamente. A partir dessa

distribuição, é possível observar pelas respostas que todos os integrantes do Grupo A sentiram-se confortáveis em utilizar os equipamentos, e não se cansaram com a interação; já no grupo B, apenas 42,9% responderam terem se sentido confortáveis, e 42,9% responderam “mais ou menos”. Além disso, a maior parte dos indivíduos do Grupo B (85%, seis indivíduos) não se cansaram com a interação, porém uma pessoa respondeu de forma medial à questão. Essa divergência pode se justificar pela diferença do estilo de vida e acesso que os idosos do Grupo B (idosos institucionalizados) têm às tecnologias computacionais em geral.

Tabela 13 - Resumo das respostas afirmativas para as questões da entrevista.

Questões	1	2	3	4	5	6	7	8	9	10	11
Grupo A	91,2%	91,2%	86,9%	56,6%	100%	8,6%	0%	82,6%	91,2%	82,6%	82,6%
Grupo B	85%	71%	57%	0%	100%	42%	0%	42%	100%	0%	57%

Tabela 14 – Distribuição e frequência das respostas assinaladas no UQE 3D para o Grupo A.

Opções de resposta	1	2	3	4	5	Sem resposta
Questões						
1				43,5%	52,2%	4,3%
2			17,4%	26,1%	56,5%	
3	95,7%		43,3%			
4			13%	26,1%	60,9%	
5			13%	17,4%	69,6%	
6			8,7%	4,3%	87%	
7	91,3%	4,3%	4,3%			
8			13%	34,8%	52,2%	
9			8,7%	17,4%	69,6%	4,3%
10			8,7%	17,4%	73,9%	
11	4,3%	8,7%	4,3%	26,1%	56,5%	
12			13%	47,8%	39,1%	
13			4,3%	34,8%	60,9%	
14		4,3%	8,7%	34,8%	52,2%	
15		4,3%	4,3%	26,1%	60,9%	4,3%
16			4,3%	17,4%	78,3%	

Tabela 15 - Distribuição e frequência das respostas assinaladas no UQE 3D para o Grupo B.

Opções de resposta	1	2	3	4	5	Sem resposta
Questões						
1		14,3%	42,9%	42,9%		
2	28,6%	14,3%	42,9%	14,3%		
3	57,1%	28,6%	14,3%			
4	14,3%		28,6%	57,1%		
5		14,3%	28,6%	57,1%		
6		14,3%	28,6%	57,1%		
7	71,4%	14,3%	14,3%			
8				85,7%	14,3%	
9			42,9%	57,1%		
10		14,3%	42,9%	42,9%		
11	14,3%	14,3%	14,3%	57,1%		
12				85,7%	14,3%	
13				100%		
14	14,3%	42,9%	28,6%	14,3%		
15		42,9%	28,6%	28,6%		
16		28,6%	28,6%	42,9%		

Quanto à imersão, 82% dos participantes do Grupo A se sentiram parte integrante do jogo, e 87% também conseguiram se localizar facilmente no espaço geográfico dentro do jogo. No Grupo B, apenas uma pessoa se sentiu parte integrante do jogo, ao passo que 57% da amostra soube se localizar no espaço geográfico dentro do jogo (conforme respostas ao UQE 3D). Essa diferença pode se justificar por uma dificuldade de abstração dos integrantes do Grupo B, uma vez que apresentaram escores do MEEM e GDS-15 mais baixos e, por isso, foram mantidos em um grupo separado para a avaliação, ou por terem menos contato com qualquer tipo de tecnologia.

Nas entrevistas, 71% do Grupo B, mencionaram terem gostado de utilizar o capacete de visualização por se sentirem mais presentes no ambiente virtual. E para a questão 8 do UQE 3D, de maneira unânime responderam que “usar o capacete de visualização foi bom para o jogo”. Portanto, é possível também que estes indivíduos não tenham conseguido abstrair exatamente o sentido das perguntas 2, 4 e 5 do questionário, o que

também pode se justificar pelos escores do MEEM e GDS-15 mais baixos, gerando a diferença quanto a resposta à questão 2 do UQE 3D para esse grupo.

Ainda quanto à imersão, na entrevista, 91,2% dos participantes do Grupo A manifestaram que gostaram de interagir com o jogo por gestos, e também de utilizar o capacete de visualização, mencionando que se sentiram dentro do jogo e muito próximos dos objetos com os quais deveriam interagir. O UQE 3D reafirma essas respostas para o Grupo A considerando, em média, 87% de respostas positivas às questões referentes (8 e 12).

No Grupo A, 86,9% dos sujeitos declararam durante a entrevista que gostariam de jogar o jogo com frequência, (56,5% sem pedir ajuda a alguém para conseguir jogar o jogo); no Grupo B, esse número foi de 57%. Quanto à facilidade da interação e intuitividade do jogo, todos os integrantes de ambos os grupos mencionaram que conseguiram facilmente entender o que deveriam realizar dentro do jogo (tarefas de interação). Da mesma forma, no questionário, 87% do grupo A e 85,7% do Grupo B, em média, responderam afirmativo para as questões que exploraram a diversão na experiência e a facilidade de utilizar a aplicação (questões 5, 6, 9, 10, 11 e 12).

No Grupo A, 82,6% e no Grupo B, 57% da amostra consideraram que o jogo é adequado a sua faixa etária e que contribuiu para aumentar o seu interesse em relação à tecnologia. Além disso, relataram aumentar a sua conexão com outras pessoas que usam tecnologias, sejam elas idosos ou jovens. Dentre estes, cinco pessoas declararam que poderiam conversar sobre este assunto com os jovens, aumentando a integração; três pessoas manifestaram que poderiam indicar a experiência para outros idosos; e duas pessoas comentaram que por meio do contato com os dispositivos já se sentiam mais próximas da tecnologia e, por isso, dos jovens.

Além disso, 91,2% do Grupo A e todos os integrantes enquadrados no Grupo B acreditaram que o jogo estimulou a prática de atividades benéficas ao corpo fisicamente (pelos movimentos e pela precisão de movimentos que exigia), e mentalmente (pelo estímulo à concentração e ao raciocínio rápido).

Nenhum dos participantes encontrou algo que pudesse definir como inadequado ou cansativo no jogo e apenas três pessoas do Grupo B (42%) declararam na entrevista terem sentido incerteza sobre a tarefa em algum momento enquanto jogavam. As questões 9, 10 e 11 do UQE 3D exploraram o quanto as imagens e os sons ajudaram a entender o jogo e a tarefa de interação, e obtiveram uma média de respostas afirmativas de 87% para o Grupo A e 52,3% para o Grupo B. Perante isso, é possível entender que as imagens e os sons do jogo não foram suficientes para impedir que os indivíduos do Grupo B vivenciassem o sentimento de

incerteza ou insegurança quanto à interação, gerando as dúvidas que os levaram a responder negativamente na entrevista e no questionário. Na entrevista, alguns idosos do Grupo B manifestaram não conseguir ouvir com clareza os sons emitidos pelo jogo, o que justifica a pontuação negativa para a questão 11. Quanto às questões 9 e 10, novamente pode-se assumir que essa diferença seja consequência da falta de contato desta amostra com a tecnologia, ou pelos escores do MEEM e GDS-15 mais baixos.

O questionário sociodemográfico aplicado na etapa Pré-Teste apontou que nenhum dos idosos que participaram do experimento conheciam as tecnologias de RV utilizadas, e apenas dois idosos manifestaram jogar jogos de computador com frequência. Ainda assim, o método identificou que os idosos conseguiram usufruir do sistema de forma fácil e intuitiva, considerando os aspectos de usabilidade de interfaces 3D.

O método proposto foi capaz de identificar, portanto, que a interface avaliada possui boa usabilidade para os idosos, uma vez que permitiu a realização das tarefas de interação de forma fácil e intuitiva, proporcionando imersão à experiência virtual, e promovendo diversão e bem-estar aos participantes.

Além disso, identificou que, para idosos institucionalizados (que possuem algum tipo de declínio cognitivo leve ou sintomatologia depressiva leve) existe uma diferença na percepção do ambiente virtual, motivo pelo qual, sugere-se a aplicação do Intervalo de Confiança diferenciado para fins de avaliação de uma interface, quando os grupos se equivalerem aos grupos utilizados neste experimento. Experimentos que utilizem grupos de participantes com características diferentes são uma ótima oportunidade de incremento do método.

Ainda, os resultados oriundos da aplicação do método indicam que o uso de entrevistas é importante como abordagem ao público idoso, por conseguir identificar melhor os sentimentos do sujeito com relação à interface, do que o uso apenas de questionários.

Subjetivamente, portanto, é possível afirmar que o método foi validado em uma etapa inicial por este experimento, uma vez que ele foi capaz de identificar o que se propõe a identificar. É importante salientar que, se o questionário desenvolvido for utilizado separadamente, ele pode atingir o mesmo resultado que o método atingiria, mas pode não ser capaz de prover validade aos resultados, pois, não será possível relacionar os seus resultados com os resultados das outras etapas do método, como os valores obtidos para o MEEM e GDS-15, as respostas do questionário sociodemográfico e as respostas dos participantes às perguntas da entrevista. Como apurado pelo experimento conduzido, sem as informações obtidas pelas demais etapas do método, não seria possível alcançar um conhecimento

completo da experiência de usabilidade vivenciada pelos idosos do Grupo B do experimento desenvolvido.

No entanto, é necessário que se realizem novos experimentos para atestar o que este experimento mostrou, bem como avaliações com especialistas da área de usabilidade de interfaces 3D para efetivamente validar o método e/ou identificar possíveis vieses a serem corrigidos. Além disso, pode ser interessante conduzir experimentos que avaliem a amplitude de efetividade da aplicação do UQE 3D, em relação ao método como um todo, para confirmar a necessidade da aplicação do método por completo.

O SUS como questionário de usabilidade não demonstrou resultados além dos identificados pelo UQE 3D, principalmente por não ser destinado a avaliar interfaces 3D ou para um público específico, pois não avalia variáveis 3DUI, como presença e imersão. Em concordância, a similaridade estatística das médias entre esses dois questionários sugere que o uso do UQE 3D é suficiente para identificar os quesitos de usabilidade buscados em uma avaliação de interfaces 3D para idosos. Portanto, o SUS não faz parte do método proposto.

Por intermédio da análise estatística, e por ser capaz de identificar os nuances da avaliação com idosos, é possível considerar que o método produziu resultados satisfatórios e pode ser útil para a avaliação da usabilidade de interfaces 3D para idosos. Além disso, o fato de o UQE 3D ter apresentado um escore muito semelhante estatisticamente ao escore do questionário SUS, pode significar que o UQE 3D é equivalente ao SUS ao analisar a usabilidade de uma interface e, portanto, pode ser aplicado em intervenções com idosos para a avaliação de interfaces 3D.

Além disso, o questionário desenvolvido foi avaliado em experimento envolvendo o público idoso, uma vez que o foco de sua concepção foi para este público. Porém, observou-se que, em relação a sua sintaxe e semântica, embora seja de fácil compreensão para o público idoso, por ter sido desenvolvido com este intuito, também pode ser aplicado a outros públicos por não estar intrinsecamente relacionado com o público em questão. Portanto, o UQE 3D pode ser bem utilizada ao ser direcionada para outras faixas etárias que não estejam familiarizadas com a tecnologia. Vale ressaltar novamente que este questionário foi desenvolvido com foco no público idoso, logo, ao mudar a especificidade do público, é necessário atentar para a adaptação da linguagem utilizada no questionário [12].

Uma limitação perscrutada em relação ao método se refere ao seu tempo médio de aplicação. Como mencionado, tempos muito longos de aplicação do método podem gerar desinteresse nos participantes. A solução encontrada foi a divisão do protocolo de aplicação do método em dois dias, porém seria interessante futuros esforços para identificar novas

formas de minimização do tempo. Novos experimentos podem ser conduzidos para investigar a exclusão de etapas, ou mesmo a execução delas em ambiente doméstico, tendo os participantes a incumbência de preencher os questionários e trazê-los ao avaliador no dia do experimento, por exemplo.

Com relação à validação do método, uma limitação encontrada é o tamanho da amostra. A amostra foi suficiente para a definição dos intervalos de confiança, para verificar que o método é capaz de identificar o que se propõe e para certificar que é capaz de atingir diferentes categorias de idosos e captar as suas percepções em relação à interface. No entanto, embora a especificidade do público alvo torne difícil o desenvolvimento de experimentos com um público maior, seria interessante direcionar esforços para tanto, o que poderia confirmar a validação do método e identificar novos pontos a serem melhorados.

6. CONCLUSÃO

O presente trabalho apresentou a proposta de um método para avaliação da usabilidade de interfaces 3D para idosos, oferecendo como principal contribuição a criação de um questionário de avaliação de usabilidade específico para o público idoso, o UQE 3D.

O método consiste de uma sequência de passos divididos em três etapas: Pré-Teste, Teste e Pós-Teste. O Pré-Teste engloba a aplicação dos questionários MEEM, GDS-15 e um questionário para caracterização da amostra, e de uma sessão de treinamento com os dispositivos a serem utilizados na etapa teste. Sugere-se que a etapa Teste e Pós-Teste ocorram num segundo dia, quando o participante testará a aplicação da qual se deseja avaliar a usabilidade (Etapa Teste) e posteriormente, responderá os questionários Pós-Teste e participará de uma entrevista semiestruturada para a avaliação da usabilidade da interface.

A proposição deste método envolveu o desenvolvimento de uma Revisão Sistemática de Literatura com o objetivo de identificar quais instrumentos são comumente utilizados em intervenções com idosos, envolvendo algum tipo de tecnologia virtual. A RS apontou 16 trabalhos relacionados, e demonstrou um padrão de uso de questionários para avaliações com o público idoso.

Para desenvolver o método foram considerados, portanto, os resultados da RS, os resultados de dois experimentos de avaliação conduzidos previamente, os conceitos e técnicas para avaliação de interfaces 3D e o processo de criação de questionários de autores conhecidos. O UQE 3D foi também revisado por profissionais da área da Computação e da Saúde antes de ser aplicado diretamente aos idosos.

O método foi aplicado em um experimento envolvendo 30 sujeitos idosos de ambos os sexos e os resultados foram analisados para fins de avaliação do próprio método. Foram realizados testes estatísticos para verificar a consistência interna do UQE 3D, para elaboração de um cálculo de score e um intervalo de confiança para aplicar aos resultados do questionário e para verificar se o questionário pode ser comparado ao SUS em termos estatísticos. Além disso, as respostas do UQE 3D foram comparadas com as respostas das entrevistas e entre os grupos de usuários ativos e institucionalizados para fins de validação dos resultados e do próprio método.

A principal contribuição deste trabalho é o método proposto, considerando o uso do questionário UQE 3D criado para avaliar interfaces 3D desenvolvidas para o público idoso. Tal método pode ser considerado validado por conseguir identificar os fatores que se

propôs. O método, com todos os seus instrumentos, é passível de ser aplicado desde o desenvolvimento das interfaces, no *design* centrado ao usuário, para identificar se a interface está atendendo aos requisitos elencados, como também ao final do processo de desenvolvimento para a validação da interface.

Entre outras contribuições relevantes, foi possível identificar que ao se trabalhar com um público específico, como idosos, entrevistas são importantes para determinar nuances que os questionários possam não captar ou mesmo para reafirmar ou identificar com mais precisão o que o questionário percebeu.

O método proposto é, portanto, uma sequência mínima de passos, identificadas até o momento como importantes para se avaliar a usabilidade de interfaces específicas para o público idoso (em intervenções com este público) capazes de identificar os nuances da percepção dos sujeitos em relação à interface. O UQE 3D se mostrou eficaz e eficiente ao identificar a usabilidade de interfaces 3D para este público.

Este método apresenta também uma contribuição especial para a área de RV, que é uma ferramenta capaz de avaliar e assegurar a qualidade de interfaces 3D para o público idoso. Como soluções neste âmbito ainda não são amplamente difundidos entre a população, o uso do UQE 3D, por exemplo, pode ajudar na melhoria da interface, e pode impulsionar o desenvolvimento destas tecnologias para idosos. Ademais, o UQE 3D se mostrou apto a ser aplicado em interações com dispositivos de RV, conseguindo identificar as principais percepções dos usuários em relação à experiência de interação.

Ainda, o questionário pode ser aplicado a outro público que não idosos, apesar do foco de sua concepção ter sido o público idoso. Assim, a proposta pode ser direcionada para outras faixas etárias que não estejam familiarizadas com a tecnologia, uma vez que o UQE 3D foi criado para ser de fácil assimilação a quem não conhece ou não convive diariamente com a tecnologia de RV.

Como trabalhos futuros, sugere-se a avaliação com uma amostra heterogênea, consideravelmente maior, para reafirmar os resultados deste trabalho e difundir o uso deste método. Recomenda-se também a avaliação do mesmo jogo em seus outros níveis, além de experimentos com diferentes aplicações com suporte a RV (outros games, simuladores, aplicações de indústria, entre outros). Além disso, é interessante realizar um estudo avaliativo especificamente com grupos de idosos institucionalizados, verificando o impacto da RV para o bem-estar dessas pessoas. Ademais, ele pode também ser aplicado a outras faixas etárias que não estejam familiarizadas com a tecnologia para averiguar a sua adaptação a outros contextos. Ainda, pode ser interessante conduzir experimentos que avaliem a amplitude de

efetividade da aplicação isolada do UQE 3D, em relação ao método como um todo, para confirmar a necessidade da aplicação do método por completo. Ainda, como trabalhos futuros, pode-se investigar maneiras de aplicação da avaliação da interface utilizando abordagens automatizadas, como por exemplo, o rastreamento óptico do usuário e do ambiente onde ele está para identificar relações de imersão e presença. Este tipo de abordagem pode assegurar também, que não haja qualquer tipo de interferência de nível intrapessoal na relação de avaliar a interação com o sistema. Com relação ao UQE 3D, pode ser interessante aproximar as questões 3 e 7, pois ambas possuem sentido negativo, e tal reorganização pode facilitar o processo de abstração e avaliação por parte do usuário.

REFERÊNCIAS

- [1] KARRAY, F. et al. Human-computer interaction: Overview on state of the art. *Citeseer*, Waterloo, 2008.
- [2] COCKTON, G. Usability Evaluation. *Encyclopedia of Human-Computer Interaction.*, n. The Interaction Design Foundation, 2012.
- [3] NIELSEN, J. *Usability 101: Introduction to Usability*, 2012. Disponível em: <<http://www.nngroup.com/articles/usability-101-introduction-to-usability>>. Acesso em: 30 maio 2017.
- [4] LAVIOLA, JR., J. J. et al. *3D User Interfaces: theory and practice*. Boston: Addison-Wesley Professional, 2017. 478 p. ISBN 0-201-75867-9.
- [5] FIORIN, M. R. et al. Motion Rehab: um jogo sério para idosos com sequelas de Acidente Vascular Encefálico. *XIV Workshop de Informática Médica*, 2014.
- [6] BROEREN, J. et al. Virtual rehabilitation after stroke. *Studies in health technology and informatics*, v. 136, p. 77, 2008.
- [7] SIMOR, F. W. *Método de avaliação da usabilidade para 'serious games' com interação gestual para idosos. Dissertação (Mestrado em Computação Aplicada)*. 2016. 78p.
- [8] FUA, K. C. et al. Designing serious games for elders. *Proceedings of the 8th International Conference of the Foundations of Digital Games*, Chania, Crete, Greece, p. 291-297, 2013.
- [9] CARVALHO, R. N. S. D.; ISHITANI, L. Fatores motivacionais para desenvolvimento de mobile serious games com foco no público da terceira idade: uma revisão de literatura. *ETD-Educação Temática Digital*, v. 15, n. MISC, p. 16--32, 2013. ISSN 1.
- [10] HALL, A. K.; MARSTON, H. R. Games for health in the home: Gaming and older adults in the digital age of healthcare. *Handbook of Smart Homes, Health Care and Well-Being*, p. 579-588, 2017.
- [11] OSMANOVIC, S.; PECCHIONI, L. Beyond entertainment: motivations and outcomes of video game playing by older adults and their younger family members. *Games and Culture*, v. 11, n. 1-2, p. 130-149, 2016.
- [12] GÜNTER, H. Como Elaborar um Questionário (Série: Planejamento de Pesquisa nas Ciências Sociais). *UnB, Laboratório de Psicologia Ambiental*, Brasília, v. 15, 2015. ISSN 1.
- [13] TROMBETTA, M. et al. Motion Rehab AVE 3D: A VR-based exergame for post-stroke rehabilitation. *COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE*, v. 151, p. 15-20, 2017.
- [14] KONSTANTINIDIS, E. I. et al. Design, implementation, and wide pilot deployment of FitForAll: An Easy to use exergaming platform improving physical fitness and life quality of senior citizens. *IEEE Journal of Biomedical and Health Informatics*, v. 20, n. 1, p. 189-200, 2016.
- [15] BAEZ, M. et al. Online Group-exercises for Older Adults of Different Physical Abilities. *Submitted to mobileCHI*. [S.l.]: [s.n.]. 2016.
- [16] BURDEA, G. C.; COIFFET, P. *Virtual reality technology*. [S.l.]: John Wiley & Sons, 2003.
- [17] NIELSEN, J. Iterative User-Interface Design. *Computer*, v. 26, n. 1993.11, p. 32-41,

- Novembro 1993. ISSN 10.1109/2.241424.
- [18] BRUM, M. R.; RIEDER, R. *Um Protocolo de Avaliação de Usabilidade de Interfaces 3D para Idosos*. 2016. 24p.
- [19] DYBA, T.; DINGSOYR, T.; HANSEN, G. K. Applying systematic reviews to diverse study types: An experience report. *First International Symposium on Empirical Software Engineering and Measurement*, Madrid, p. 225-234, Setembro 2007.
- [20] ORSO, V. et al. Involving Older Adults in Designing Interactive Technology: The Case of SeniorChannel. *Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter*, 2015. 102-109.
- [21] HAGIYA, T.; HORIUCHI, T.; YAZAKI, T. Typing Tutor: Individualized Tutoring in Text Entry for Older Adults Based on Input Stumble Detection. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016. 733-744.
- [22] TOST, D. et al. SmartAgeing: a 3D serious game for early detection of mild cognitive impairments. *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, 2014. 294-297.
- [23] WASHINGTON UNIVERSITY SCHOOL OF MEDICINE. MAPI Research Institute (2008) Clinical Dementia Rating Questionnaire. *Alzheimer's Disease Research Center (Knight ADRC)*, Missouri.
- [24] HIMMELSBACH, J. et al. Enabling accessibility through multimodality? In: MUM '15. *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. [S.l.]: [s.n.]. 2015. p. 195-199.
- [25] KAHNEMAN, D. et al. A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 2004. 1776-1780.
- [26] FLANAGAN, J. C. The critical incident technique. *Psychological bulletin*, v. 51, p. 327, 1954. ISSN 4.
- [27] SACRAMENTO, C.; BACELLAR, S.; FERREIRA, L. Elderly 's Comprehension on Facebook: a Communicability Evaluation. *Proceedings of the 7th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*. [S.l.]: [s.n.]. 2016. p. 315-322.
- [28] DANTA, A. et al. Interaction Evaluation using the Mac-g in Sloodle. *Procedia Computer Science*, v. 67, p. 160-170, 2015.
- [29] LE, T.; THOMPSON, H. J.; DEMIRIS, G. A Comparison of Health Visualization Evaluation Techniques With Older Adults. *IEEE Comput Graph Appl*. [S.l.]: [s.n.]. 2015. p. 1-26.
- [30] AL-KHOMSAN, M. et al. Usability evaluation of Twitter on Android platform for elderly Arab users using morae evaluation tool. In: ICITST 2015. *2015 10th International Conference for Internet Technology and Secured Transactions*. [S.l.]: [s.n.]. 2016. p. 446-447.
- [31] KIRAKOWSKI, J.; CORBETT, M. SUMI: The software usability measurement inventory. *British journal of educational technology*, v. 24, p. 210-212, 1993. ISSN 3.
- [32] SAXENA, S. et al. The WHO quality of life assessment instrument (WHOQOL-Bref): the importance of its items for cross-cultural research. *Quality of life Research*, v. 10, p. 711-721, 2001. ISSN 8.
- [33] HSIEH, J. C.; HUANG, C. C. Performance evaluation on simplified computer for the elderly people. In: IMIS 2013. *Proceedings - 7th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. [S.l.]: [s.n.]. 2013. p. 701-705.

- [34] STEVERINK, N. et al. Measuring frailty: Developing and testing the GFI (Groningen Frailty Indicator). *Gerontologist*, v. 41, p. 236-237, 2001.
- [35] SCHUURMANS, J. et al. Tablet-based support for older adults with severe mood disorders treated in an ambulatory geriatric psychiatry setting: Protocol of a feasibility study of the eCare@Home platform. *Internet Interventions*, v. 6, p. 22-28, 2016.
- [36] DEVILLY, G. J.; BORKOVEK, T. D. Psychometric properties of the credibility/expectancy questionnaire. *Journal of behavior therapy and experimental psychiatry*, v. 31, p. 73-86, 2000. ISSN 02.
- [37] ATTKISSON, C. C.; ZWICK, R. The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and program planning*, v. 5, p. 233-237, 1982. ISSN 3.
- [38] CHIU, C.-J. et al. The attitudes, impact, and learning needs of older adults using apps on touchscreen mobile devices: Results from a pilot study. *Computers in Human Behavior*, v. 63, p. 189-197, 2016.
- [39] JAY, G. M.; WILLIS, S. L. Influence of direct computer experience on older adults' attitudes toward computers. *Journal of Gerontology*, v. 47, p. 250-257, 1992. ISSN 4.
- [40] KOHOUT, F. J. et al. Two shorter forms of the CES-D depression symptoms index. *Journal of aging and health*, v. 5, p. 179-193, 1993. ISSN 2.
- [41] LU, L.; LIN, Y. Y. Family roles and happiness in adulthood. *Personality and individual differences*, v. 25, p. 195-207, 1998. ISSN 2.
- [42] MEZA-KUBO, V. et al. Assessing the user experience of older adults using a neural network trained to recognize emotions from brain signals. *Journal of Biomedical Informatics*, v. 62, p. 202-209, 2016.
- [43] FERREIRA, F. et al. Elderly centered design for interaction - The case of the S4S Medication Assistant. *Procedia Computer Science*, v. 27, p. 398-408, 2013.
- [44] TEIXEIRA, A. et al. Design and development of Medication Assistant: older adults centred design to go beyond simple medication reminders. *Universal Access in the Information Society*, p. 1-16, 2016.
- [45] EPELDE, G. et al. Providing universally accessible interactive services through TV sets: Implementation and validation with elderly users. *Multimedia Tools and Applications*, v. 67, n. 2, p. 497-528, 2013.
- [46] ORGANIZAÇÃO MUNDIAL DA SAÚDE. Global Age-friendly Cities: A Guide. *World Health Organization*, 2008. Disponível em: <<http://www.who.int/ageing/GuiaAFCPortuguese.pdf>>. Acesso em: julho 2018.
- [47] WHO. Indicators for the Minimum Data Set Project on Ageing: A Critical Review in sub-Saharan Africa. *World Health Organization*, 2001. Disponível em: <http://www.who.int/healthinfo/survey/ageing_mds_report_en_daressalaam.pdf>. Acesso em: Agosto 2018.
- [48] GUILLEMIN, F.; BOMBARDIER, C.; BEATON, D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of clinical epidemiology*, v. 46, n. 12, p. 1417-1432., 1993.
- [49] CHAI, J. et al. Application of frontal EEG asymmetry to user experience research. *Engineering Psychology and Cognitive Ergonomics*, v. 8532, p. 234-243, 2014.
- [50] BOWMAN, D. A.; GABBARD, J. L.; HIX, D. A Survey of Usability Evaluation in Virtual Environments: classification and comparison of methods. *Presence: Teleoperators and Virtual Environments*, v. 11, n. 4, p. 404-424, 2002.

- [51] YESAVAGE, J. A. et al. development and validation of a geriatric depression screening scale: a preliminary report. *Journal of psychiatric research*, v. 17, n. Elsevier, p. 37--49, 1982. ISSN 1.
- [52] SHEIK, J. I.; YESAVAGE, J. A. Geriatric depression scale (GDS): recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 1986.
- [53] BERTOLUCCI, P. H. F. et al. O Mini-exame do Estado Mental em uma população geral: impacto da escolaridade. *Arquivos de Neuropsiquiatria*, São Paulo, v. 52, p. 01--07, Março 1994. ISSN 1.
- [54] BRUCKI, S. M. et al. Sugestões para o Uso do mini Exame do Entado Mental no Brasil. *Arquivos de Neuro-psiquiatria*, 2013.
- [55] ALMEIDA, O. P.; ALMEIDA, S. A. Confiabilidade da versão brasileira da Escala de Depressão em Geriatria (GDS) versão reduzida. *Arq Neuropsiquiatr*, v. 57, p. 421--6, 1999. ISSN 2B.
- [56] LOURENÇO, R. A.; VERAS, R. P. Mini-Mental State Examination: psychometric characteristics in elderly outpatients. *Revista de Saúde Pública*, v. 40, p. 712--719, 2006. ISSN 4.
- [57] US DEPARTMENT OF HEALTH AND HUMAN SERVICES. Usability.gov. *Usability.gov*. Disponível em: <<https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>>.
- [58] HUANG, H.-M.; RAUCH, U.; LIAW, S.-S. Investigating learners' attitudes toward virtual reality learning environments. *Computers & Education*, 2010. 1171--1182.
- [59] BENYON, D. *Interação Humano-Computador*. São Paulo: Pearson, 2011.
- [60] NIELSEN, J. How Many Test Users in a Usability Study? *Nielsen Normal Group*, 2012. Disponível em: <<http://www.nngroup.com/articles/how-many-test-users/>>. Acesso em: fevereiro 2018.
- [61] MICROSOFT. Kinect for Windows. *Microsoft Windows Dev Center*. Disponível em: <<https://developer.microsoft.com/en-us/windows/kinect>>. Acesso em: agosto 2018.
- [62] HOSS, M.; CATEN, C. S. T. Processo de Validação Interna de um Questionário em uma Survey Research. *Produto & Produção*, v. 11, n. 2, p. 104-119, junho 2010.
- [63] HUBLEY, A. M.; ZUMBO, B. D. A dialectic on validity: where we have been and. *Journal of General Psychology*, v. 23, n. 3, p. 207-215, 1996.
- [64] NUNNALLY, J. C. *Psychometric theory*. 2. ed. New York: McGraw-Hill, 1978.
- [65] LANDIS, R. J.; KOCK, G. G. The measurement of observer agreement for categorical data. *Biometrics*, p. 159-174, 1977.
- [66] BOUCINHA, M.; TAROUÇO, L. M.. Avaliação de Ambiente Virtual de Aprendizagem com o uso do SUS-System Usability Scale. *RENOTE*, v. 11, n. 3, 2013.
- [67] UX RESEARCH. Measuring and Interpreting System Usability Scale (SUS). *UIUX Trend*, 2017. Disponível em: <<https://uiuxtrend.com/measuring-system-usability-scale-sus/>>. Acesso em: 21 out. 2018.
- [68] SNEDECOR, G. W.; COCHRAN, W. G. *Statistical Methods*. 8. ed. [S.l.]: [s.n.], 1989.
- [69] LOMAX, R. G.; HAHS-VAUGHN, D. L. *Statistical Concepts: A Second Course*. [S.l.]: [s.n.], 2013. 10 p.
- [70] *GTK+ 3 Reference Manual*. Disponível em: <<https://developer.gnome.org/gtk3/>>. Acesso em: 31 Dez. 2013.
- [71] CASTRO, A. A. *Método.org*, maio 2011. Disponível em:

<<http://método.org/wpcontent/uploads/2010/08/meta1.PDF>>. Acesso em: 05 março 2017.

- [72] ISHINO, K.; HAGIWARA, M. A feeling estimation system using a simple electroencephalograph. *IEEE International Conference on Systems, Man and Cybernetics*, v. 3, p. 4204-4209, 2003.
- [73] PETRANTONAKIS, P. C.; HADJILEONTIADIS, L. J. Emotion recognition from EEG using higher order crossings. *IEEE Transactions on Information Technology in Biomedicine*, v. 14, n. 2, p. 186-197, 2010.
- [74] ZHANG, Q.; LEE, M. Analysis of positive and negative emotions in natural scene using brain activity and GIST. *Neurocomputing*, v. 4, n. 72, p. 1302-1306, 2009.
- [75] LIN, Y. P. et al. EEG-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, v. 57, n. 7, p. 1798-1806, 2010.
- [76] CHIU, C. J.; ET AL. The attitudes, impact, and learning needs of older adults using apps on touchscreen mobile devices: Results from a pilot study. *Computers in Human Behavior*, v. 63, p. 189-197, 2016.

APÊNDICE A - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado(a) a participar da pesquisa “UMA METODOLOGIA PARA AVALIAÇÃO DA USABILIDADE DE INTERFACES 3D PARA IDOSOS, de responsabilidade do professor DR. RAFAEL RIEDER e da mestrandia MANOELA ROGOFSKI BRUM POSTAL, pesquisadores do Grupo de Sistemas Interativos e de Informação da Universidade de Passo Fundo (UPF).

Esta pesquisa se justifica devido à tendência do mercado em ofertar *games* e ambientes virtuais 3D destinados ao público idoso, tanto para reabilitação, quanto para entretenimento ou treinamento. Nesse contexto, é importante que soluções de software como estas, que oferecem interfaces 3D, sejam avaliadas quanto à sua usabilidade por meio de um método que assegure a facilidade de uso e a adequação ao perfil do público alvo.

O objetivo desta pesquisa é validar um método de avaliação de usabilidade de interfaces 3D para idosos com o uso de dispositivos de Realidade Virtual, como Oculus Rift e Kinect. Para tanto, o(a) participante é convidado(a) a testar as funcionalidades de um *game* utilizando um dispositivo de visualização e um sensor de movimentos.

Como benefícios, esta pesquisa possibilitará que interfaces 3D desenvolvidas para idosos atendem as necessidades de uso e acesso dessa faixa etária, facilitando o uso e a interação com o computador. Como consequência, tecnologias como jogos de computador, por exemplo, poderão oferecer entretenimento com mais conforto, bem-estar e satisfação, contribuindo para uma melhor qualidade de vida do idoso.

Sua participação na pesquisa será de 2 (dois) encontros, com duração aproximada de 20 (vinte) minutos, durante o período diurno, nas dependências do Campus I da UPF, Prédio B5, em data e local a serem definidos. Esta observação será registrada em papel e, eventualmente, por um software de captura de telas do computador e/ou por vídeo.

Cabe lembrar que o objetivo deste estudo **não é** avaliar o participante, **mas sim** avaliar o software e os equipamentos que o(a) participante utilizará durante os testes. O uso que se faz dos registros efetuados durante o teste é estritamente limitado a atividades de pesquisa e desenvolvimento, e sua contribuição permitirá verificar a qualidade do sistema em questão.

Não são esperados desconfortos ou riscos durante a sessão. Porém, caso sejam identificados sinais de desconforto, o pesquisador fará o encaminhamento para os profissionais especializados da área de Psicologia ou de Enfermagem da UPF.

Você terá a garantia de receber esclarecimentos sobre qualquer dúvida relacionada à pesquisa e poderá ter acesso aos seus dados em qualquer etapa do estudo. Sua participação não é obrigatória e você pode desistir a qualquer momento, retirando seu consentimento.

Não há previsão de ressarcimento de gastos ou pagamento pela participação na pesquisa. Em caso de necessidade de auxílio para transporte, este será providenciado pelos responsáveis em questão.

Os dados relacionados à sua identificação não serão divulgados, e os registros gravados serão posteriormente destruídos. Os resultados da pesquisa serão divulgados em formato de trabalho de conclusão de curso e também em eventos científicos, mas você tem a partir desse momento a garantia do sigilo e da confidencialidade dos dados.

Caso você tenha dúvidas sobre a pesquisa ou se sinta constrangido(a), incomodado(a) ou prejudicado(a) em sua dignidade ou autonomia, você pode entrar em contato com os pesquisadores RAFAEL RIEDER ou MANOELA ROGOFSKI BRUM POSTAL, pelos telefones (54) 9204-7070 ou 9175-0530, ou consultar o Comitê de Ética em Pesquisa da UPF, pelo telefone (54) 3316-8157, no horário das 08h às 12h e das 13h30 às 17h30, de segunda a sexta-feira.

Dessa forma, se você concorda em participar da pesquisa como consta nas explicações e orientações acima, coloque seu nome e a data nos locais abaixo indicados. Desde já, agradecemos sua colaboração e solicitamos a sua assinatura de autorização neste termo, que será também assinado pela responsável em duas vias, sendo que uma ficará com você e outra com os pesquisadores.

Passo Fundo, ____/ ____/ ____.

Assinatura do(a) Participante

Nome completo do(a) participante (por extenso)

Assinatura do Pesquisador
Professor Dr. Rafael Rieder

Assinatura da Pesquisadora
*Mestranda Manoela Rogofski Brum
Postal*

APÊNDICE B - QUESTIONÁRIO SOCIODEMOGRÁFICO E DE CARACTERIZAÇÃO DA AMOSTRA

Data: ____/____/____

ID do participante: _____

Idade: _____ anos

Gênero: () Masculino () Feminino

Escolaridade (marque com um X):

	Completo	Incompleto
Ensino Fundamental (Antigo 1º Grau, Primário + Ginásio)		
Ensino Médio (Antigo 2º grau, Colegial Científico/ Clássico)		
Ensino Superior (Antigo 3º grau, Faculdade)		
Pós-Graduação (Especialização ou Mestrado ou Doutorado)		

Já jogou algum jogo de computador (*game*)? () Sim () Não

Se sim, com qual frequência? _____

Se sim, utilizou algum dos equipamentos a seguir relacionados? (Marque aqueles que já interagiu em games)

- () Kinect
- () Oculus Rift
- () Smart TV 3D
- () Relógio inteligente

Faz algum tipo de tratamento médico? () Sim () Não

Se sim, qual tratamento? _____

Você faz uso de alguma medicação? () Sim () Não

Se sim, qual medicação? _____

Possui alguma dificuldade de movimento (motora)? () Sim () Não

Se sim, qual dificuldade? _____

APÊNDICE C - QUESTIONÁRIO UQE 3D

Data: ____/____/____

ID do participante: _____

Nº	Afirmações	Assinale as respostas abaixo que melhor descrevem como você se sentiu em relação a interação no jogo				
		1 NADA	2 POUCO	3 MAIS OU MENOS	4 MUITO	5 EXTREMA- MENTE
1	Senti-me confortável com os equipamentos durante o jogo.	1	2	3	4	5
2	Senti-me parte integrante do jogo ao longo do processo interativo.	1	2	3	4	5
3	O jogo provocou situações de mal-estar como enjoo, dor de cabeça, tontura ou náusea.	1	2	3	4	5
4	Consegui saber onde eu estava durante o jogo.	1	2	3	4	5
5	Foi fácil me localizar durante o jogo.	1	2	3	4	5
6	Foi fácil realizar as tarefas do jogo.	1	2	3	4	5
7	Cansou-me fazer movimentos durante o jogo.	1	2	3	4	5
8	Usar o capacete de visualização me ajudou no jogo.	1	2	3	4	5
9	As imagens me ajudaram a entender como jogar.	1	2	3	4	5

10	As imagens me ajudaram a desempenhar as tarefas do jogo.	1	2	3	4	5
11	Os sons emitidos me ajudaram a entender como jogar.	1	2	3	4	5
Nº	Afirmações	Assinale as respostas abaixo que melhor descrevem como você se sentiu em relação a interação no jogo				
		1 NADA	2 POUCO	3 MAIS OU MENOS	4 MUITO	5 EXTREMA- MENTE
12	Achei o jogo divertido.	1	2	3	4	5
13	Gostei de interagir por gestos no jogo.	1	2	3	4	5
14	Achei o jogo adequado para minha idade.	1	2	3	4	5
15	Achei suficiente o tempo que tive para jogar.	1	2	3	4	5
16	Achei suficiente o tempo que tive para descansar após o jogo.	1	2	3	4	5

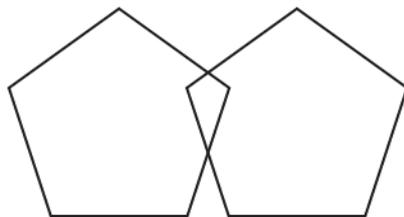
ANEXO A - Mini Exame do Estado Mental – MEEM

Folstein, Folstein e McHugh (1975)

Data: ____/____/____

ID do participante: _____

Questões	Pontos
* Qual é o (ano) (estação) (data) (dia/mês) e (mês).	5
* Onde estamos (país) (estado) (cidade) (rua ou local) (andar).	5
* Nomear três objetos (pente, rua, azul). Pedir para prestar atenção, pois terá que repetir mais tarde. Pergunte pelas três palavras após tê-las nomeado. Repetir até que evoque corretamente e anotar número de vezes (5): ____	3
* Subtrair: 100-7 e 7 dos seguintes números: 93 – 86 – 79 – 72 – 65	5
* Perguntar pelas 3 palavras anteriores (pente-rua-azul)	3
* Identificar lápis e relógio de pulso	2
* Repetir: “Nem aqui, nem ali, nem lá”.	1
* Seguir o comando de três estágios: “Pegue o papel com a mão direita, dobre ao meio e ponha no chão”.	3
* Ler “em voz baixa” e executar: FECHE OS OLHOS	1
* Escrever uma frase (um pensamento, ideia completa).	1
* Copiar o desenho:	1



ANEXO B - Escala de Depressão Geriátrica (GDS-15) – Versão Reduzida

Yesavage *et al.* (1983)

Data: ____/____/____

ID do participante: _____

Responda **Sim** ou **Não** consoante se tem sentido de há uma semana para cá:

Está satisfeito(a) com sua vida?	S	N
Interrompeu muitas de suas atividades?	S	N
Acha sua vida vazia?	S	N
Aborrece-se com frequência?	S	N
Sente-se bem com a vida na maior parte do tempo?	S	N
Teme que algo de ruim lhe aconteça?	S	N
Sente-se alegre a maior parte do tempo?	S	N
Sente-se desamparado com frequência?	S	N
Prefere ficar em casa a sair e fazer coisas novas?	S	N
Acha que tem mais problemas de memória que outras pessoas?	S	N
Acha que é maravilhoso estar vivo(a)?	S	N
Sente-se inútil?	S	N
Sente-se cheio(a) de energia?	S	N
Sente-se sem esperança?	S	N
Acha que os outros tem mais sorte que você?	S	N

ANEXO C - Questionário SUS

Data: ____/____/____

ID do participante: _____

QUESTIONÁRIO DE AVALIAÇÃO DA USABILIDADE DO SISTEMA						
Nº	Afirmação:	Marque sua resposta aqui!				
1	Eu acho que gostaria de usar esse jogo com frequência.	1	2	3	4	5
2	Eu acho o jogo muito complexo, e não deveria ser.	1	2	3	4	5
3	Eu achei o jogo fácil de usar.	1	2	3	4	5
4	Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o jogo.	1	2	3	4	5
5	Eu acho que as várias funções do jogo estão bem integradas.	1	2	3	4	5
6	Eu acho que o jogo apresenta problemas.	1	2	3	4	5
7	Eu imagino que as pessoas aprenderão como usar esse jogo rapidamente.	1	2	3	4	5
8	Eu achei o sistema atrapalhado de usar.	1	2	3	4	5
9	Eu me senti confiante ao usar o sistema.	1	2	3	4	5
10	Eu precisei aprender coisas novas antes de conseguir usar o sistema.	1	2	3	4	5