



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELOS DE DECISÃO E SAÚDE

BIANCA OLIVEIRA ISMAEL DA COSTA

**SIMULAÇÃO COMPUTACIONAL PARA TREINAMENTO PERCEPTIVO-
VISUAL DA CLASSIFICAÇÃO DOS RESÍDUOS FARÍNGEOS NA
VIDEOENDOSCOPIA DA DEGLUTIÇÃO**

JOÃO PESSOA

2024

BIANCA OLIVEIRA ISMAEL DA COSTA

**SIMULAÇÃO COMPUTACIONAL PARA TREINAMENTO PERCEPTIVO-
VISUAL DA CLASSIFICAÇÃO DOS RESÍDUOS FARÍNGEOS NA
VIDEOENDOSCOPIA DA DEGLUTIÇÃO**

Tese de doutorado a ser apresentada ao Programa de Pós-Graduação em Modelos de Decisão e Saúde – Nível Doutorado, do Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba.

Linha de Pesquisa: Modelos em Saúde.

Orientadores:

Prof. Dr. Leandro de Araújo Pernambuco

Prof.^a Dra. Liliane dos Santos Machado

JOÃO PESSOA

2024

Catálogo na publicação
Seção de Catalogação e Classificação

C838s Costa, Bianca Oliveira Ismael da.

Simulação computacional para treinamento perceptivo-visual da classificação dos resíduos faríngeos na videoendoscopia da deglutição / Bianca Oliveira Ismael da Costa. - João Pessoa, 2024.
116 f. : il.

Orientação: Leandro de Araújo Pernambuco, Liliane dos Santos Machado.
Tese (Doutorado) - UFPB/CCEN.

1. Deglutição. 2. Simulação computacional - Saúde.
3. Treinamento perceptivo-visual. 4. Educação em saúde.
5. Resíduos faríngeos - Videoendoscopia. I. Pernambuco, Leandro de Araújo. II. Machado, Liliane dos Santos.
III. Título.

UFPB/BC

CDU 612.312(043)

Aos vinte e três dias do mês de agosto do ano de dois mil e vinte e quatro, às 13h20min, no Auditório da Escola Técnica de Saúde (CPT/ETS/UFPB), instalou-se a banca examinadora de tese de Doutorado da aluna BIANCA OLIVEIRA ISMAEL DA COSTA. A banca examinadora foi composta pelos professores Dra. ROSANE SAMPAIO SANTOS, UTP, e REGIVAN HUGO NUNES SANTIAGO, UFRN, examinadores externos à instituição, Dr. LEONARDO WANDERLEY LOPES, UFPB, e Dr. RONEI MARCOS DE MORAES, UFPB, examinadores internos, Dra. LILIANE DOS SANTOS MACHADO, UFPB, como orientadora e Dr. LEANDRO DE ARAÚJO PERNAMBUCO, UFPE, orientador e presidente da banca examinadora. Dando início aos trabalhos, o presidente da banca cumprimentou os presentes, comunicou aos mesmos a finalidade da reunião e passou a palavra à candidata para que se fizesse, oralmente, a exposição do trabalho de tese intitulado “SIMULAÇÃO COMPUTACIONAL PARA TREINAMENTO PERCEPTIVO-VISUAL DA CLASSIFICAÇÃO DOS RESÍDUOS FARÍNGEOS NA VIDEOENDOSCOPIA DA DEGLUTIÇÃO”. Concluída a apresentação a candidata foi arguida pela Banca Examinadora que emitiu o parecer “**APROVADA**”. Sendo assim, após a candidata proceder às devidas correções exigidas pela Banca Examinadora no prazo máximo de **30 dias** e depositar as cópias finais da tese, deverá a Coordenação expedir uma declaração de conclusão do Programa e a Universidade Federal da Paraíba deverá expedir o respectivo diploma de Doutora em Modelos de Decisão e Saúde, na forma da lei. E, para constar, o Prof. Hemílio Fernandes Campos Coêlho, Sr. Coordenador do Programa de Pós-Graduação em Modelos de Decisão e Saúde, lavrou a presente Ata, que vai assinada por ele e pelos demais membros da Banca Examinadora. João Pessoa, 23 de agosto de 2024.

Prof. Leandro de Araújo Pernambuco
Orientador e Presidente



Documento assinado digitalmente
LEANDRO DE ARAUJO PERNAMBUCO
Data: 04/09/2024 09:28:18-0300
Verifique em <https://validar.iti.gov.br>

Prof. Liliane dos Santos Machado
Orientadora

Prof. Leonardo Wanderley Lopes
Examinador Interno

Prof. Ronei Marcos de Moraes
Examinador Interno



Documento assinado digitalmente
ROSANE SAMPAIO SANTOS
Data: 03/09/2024 13:48:53-0300
Verifique em <https://validar.iti.gov.br>

Profa. Rosane Sampaio Santos
Examinadora Externa à Instituição

Prof. Regivan Hugo Nunes Santiago
Examinador Externo à Instituição



Documento assinado digitalmente
REGIVAN HUGO NUNES SANTIAGO
Data: 04/09/2024 07:59:55-0300
Verifique em <https://validar.iti.gov.br>

Emitido em 23/08/2024

ATA Nº 01/2024 - PPGMDS (11.01.14.54)
(Nº do Documento: 1)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 17/09/2024 16:12)
RONEI MARCOS DE MORAES
PROFESSOR DO MAGISTERIO SUPERIOR
337967

(Assinado digitalmente em 04/09/2024 10:03)
LILIANE DOS SANTOS MACHADO
PROFESSOR DO MAGISTERIO SUPERIOR
2379027

(Assinado digitalmente em 04/09/2024 10:10)
LEONARDO WANDERLEY LOPES
PRO-REITOR(A)
2634755

Para verificar a autenticidade deste documento entre em <https://sipac.ufpb.br/documentos/> informando seu número: **1**,
ano: **2024**, documento (espécie): **ATA**, data de emissão: **04/09/2024** e o código de verificação: **3fac10b7ff**

DEDICATÓRIA

A **Deus**, meu alicerce em todos os momentos da vida. Por revigorar minhas forças durante as provas e conduzir meus passos sob a Sua vontade.

À minha mãe, **Damiana**, pelas orações e incentivos durante mais uma fase importante da minha vida. Obrigada por sempre interceder por mim. E ao meu pai, **Severino**, pelo apoio e estímulo depositados. Por acreditar neste sonho deste o surgimento dele. Gratidão por me concederem o acesso à educação, o que me permitiu chegar até aqui. Dedico a vocês essa conquista que é nossa.

A **Ricardo Luna**. Obrigada por ser e estar presente durante este desafio. Por apoiar e acreditar no meu trabalho e nas minhas escolhas como ninguém.

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Leandro Pernambuco, por acreditar na minha capacidade de desenvolver este trabalho. Obrigada por toda dedicação e parceria nestes oito anos, por tantas oportunidades de aprendizado. És inspiração como professor, pesquisador e ser humano.

A minha orientadora, Profa. Dra. Liliane Machado, por todos os ensinamentos, pelas inúmeras sugestões e discussões durante o desenvolvimento do trabalho, e pela paciência. Gratidão pelas contínuas injeções de ânimo que me fizeram persistir.

Aos meus familiares e amigos pela paciência e compreensão nos diversos momentos em que precisei me ausentar do convívio social para me dedicar ao desenvolvimento desta tese. Obrigada por todo o apoio e carinho.

Aos colegas do Laboratório de Tecnologias para o Ensino Virtual e Estatística (LabTEVE/UFPB). Um agradecimento especial a Raul, pela amizade construída durante esses quatro anos e meio. A Desiré, pela parceria e colaboração no trabalho. A Zildomar, Dandara, Lucas e Elvis. Enfim, a todos os membros do laboratório.

Às Dras. Roberta Gonçalves da Silva e Dra. Suely Mayumi Motonaga Onofri, coordenadoras do LADis – Laboratório de Disfagia da UNESP/Marília-SP, pela parceria e cessão do banco de dados em videoendoscopia da deglutição utilizado neste estudo.

À Dra. Thaís Coelho Alves, colaboradora deste estudo, que com a sua expertise sobre videoendoscopia da deglutição forneceu dados para a pesquisa. Obrigada pelo tempo disponibilizado para contribuir com este trabalho.

Aos colegas do PPGMDS, principalmente a Nadjenny, Alexandra, Mary, Karol e Maxsuel, pelo companheirismo especialmente nos anos iniciais do doutorado, quando compartilhamos muitos momentos de estudo e conhecimento.

Aos docentes e técnicos-administrativos do PPGMDS, pela disponibilidade de sempre e contribuição para a minha formação acadêmica.

Aos membros da banca examinadora, pela disponibilidade e contribuições valiosas. As considerações foram e serão imprescindíveis na construção deste trabalho.

Aos profissionais de saúde que prontamente se dispuseram a participar e compor a amostra do estudo.

À Capes – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo apoio financeiro.

RESUMO

O acúmulo de resíduos faríngeos é um parâmetro que está relacionado à eficiência da deglutição e pode ser visualizado por meio do exame de videoendoscopia da deglutição (VED). O treinamento para classificar este parâmetro perceptivo-visual aumenta a confiabilidade da classificação e é recomendado na literatura. O uso de simuladores computacionais pode contribuir no treinamento profissional, possibilitando a avaliação automática do desempenho do participante por meio de situações que replicam experiências reais do contexto clínico, o que pode resultar em uma melhor qualidade da classificação dos resíduos faríngeos. O argumento desta tese é de que é possível proporcionar e avaliar o treinamento para a classificação dos resíduos faríngeos na VED por meio da simulação computacional. Sendo assim, o presente trabalho teve como objetivo propor e desenvolver um simulador computacional de treinamento perceptivo-visual capaz de avaliar a habilidade do usuário em realizar a classificação dos resíduos faríngeos na VED. Para tanto, foram utilizadas informações revisadas na literatura buscando compreender como o treinamento tradicional é realizado e identificar aspectos importantes para a elaboração de um treinamento computacional. Foram identificados os requisitos necessários para tal treinamento, utilizados para propor e implementar um modelo de decisão para avaliar a habilidade perceptivo-visual do usuário. Além disso, a extração de variáveis e valores para o treinamento foi feita a partir do tratamento de vídeos e da classificação de um banco de dados de resíduos faríngeos em VED por especialistas. O resultado desta pesquisa resultou no desenvolvimento do simulador DeglutSIM para treinar a classificação dos resíduos faríngeos na VED, utilizando recortes de vídeos rotulados por especialistas e um índice de desempenho baseado na confiabilidade intra-avaliador e precisão. O DeglutSIM foi implementado com um sistema baseado em regras (SBR) com base na lógica clássica como modelo de decisão inteligente para a avaliação do desempenho do usuário, fundamentado em classificações realizadas por especialistas. O simulador conta com duas etapas para a avaliação: nivelamento, onde casos mais fáceis de serem classificados são apresentados; e prática, em que todos os tipos de casos são apresentados. As medidas de confiabilidade intra-avaliador e precisão são utilizadas para mensurar o desempenho do usuário nestas etapas. A avaliação da usabilidade do simulador foi feita através da aplicação do questionário *System Usability Scale* (SUS). Como resultado da avaliação de usabilidade, um *score* SUS médio de 92,7 de um *score* máximo de 100, classificado como “melhor imaginável”, foi obtido. A confiabilidade entre os avaliadores, medida pelo cálculo do coeficiente Kappa ponderado, mostrou-se entre moderada

e perfeita na maioria das comparações, o que demonstra a consistência na avaliação da usabilidade do sistema entre os usuários. Deste modo, o DeglutSIM demonstrou ser uma ferramenta com potencial para auxiliar o treinamento perceptivo-visual na classificação dos resíduos faríngeos na VED, apresentando alto grau de usabilidade e concordância entre os avaliadores, o que evidencia a robustez da avaliação realizada.

Palavras-chave: Deglutição; Simulação por Computador; Treinamento; Educação em Saúde; Avaliação de treinamento.

ABSTRACT

The accumulation of pharyngeal residues relates to swallowing efficiency and can be visualized using fiberoptic endoscopic evaluation of swallowing (FEES). Training to classify this visual-perceptual parameter increases the reliability of the classification and is recommended in the literature. Using computer simulators can contribute to professional training by making it possible to automatically evaluate the participant's performance through situations that replicate real experiences in the clinical context, potentially improving the quality of pharyngeal residue classification. This thesis argues that it is possible to provide and measure training for the classification of pharyngeal residues in FEES through computer simulation. Therefore, this study aimed to propose and develop a perceptual-visual training computer simulator capable of evaluating the user's ability in classifying pharyngeal residues in FEES. To achieve this, information reviewed in the literature was used to understand how traditional training is conducted and to identify important aspects for the development of computer-based training. The necessary requirements for such training were identified and used to propose and to implement a decision model to assess the user's perceptual-visual ability. Additionally, the extraction of variables and values for training was based on the processing of videos and the classification of a database of pharyngeal residues in FEES by experts. The outcome of this research resulted in the development of the DeglutSIM simulator to train the classification of pharyngeal residues in FEES, using video clippings labeled by experts and a performance index based on intra-rater reliability and accuracy. DeglutSIM was implemented with a rule-based system (RBS) based on classical logic as an intelligent decision model for evaluating user performance, based on classification made by experts. The simulator has two evaluation stages for evaluation: leveling, where cases that are easier to classify are presented; and practice, in which all types of cases are presented. The simulator's usability was assessed using the System Usability Scale (SUS) questionnaire. As a result of the usability evaluation, an average SUS score of 92.7 was obtained out of a maximum score of 100, classified as the "best possible". Inter-rater reliability, as measured by the weighted Kappa test, was moderate to perfect in most of the comparisons, which demonstrates the consistency of the system's usability evaluation among users. In this way, DeglutSIM proved to be a tool with the potential to help perceptual-visual training in the classification of pharyngeal residues in FEES, with a high degree of usability and agreement between the evaluators, which shows the validity of the evaluation.

Keywords: Deglutition; Computer Simulation; Training; Health Education; Training assessment.

LISTA DE FIGURAS

Figura 1: Funcionamento de um sistema baseado em regras.	33
Figura 2: Fluxograma do processo de desenvolvimento do banco de dados PhaReD-FEES .	38
Figura 3: Processo de escolha do modelo de decisão inteligente para o simulador.....	46
Figura 4: Visão geral do índice considerado na avaliação do desempenho do usuário no modelo de decisão.	47
Figura 5: Protótipo de tela do simulador.	50
Figura 6: Fluxograma demonstrativo do funcionamento do sistema, onde RFV significa resíduos faríngeos em valéculas; RFS, resíduos faríngeos em seios piriformes; e RFVS, resíduos faríngeos em valéculas e seios piriformes.	64
Figura 7: Tela inicial.....	65
Figura 8: Seleção de fase.....	66
Figura 9: Seleção de etapa.....	66
Figura 10: Instruções na etapa Ensino.	67
Figura 11: Tela principal da etapa Ensino.	67
Figura 12: Tela referente ao nível de gravidade I da YPRSR.	68
Figura 13: Instruções da etapa Nivelamento na fase RFV.	69
Figura 14: Tela principal da etapa Nivelamento com âncora aberta na fase RFV.....	69
Figura 15: Tela principal da etapa Nivelamento na fase RFVS.	70
Figura 16: Tela principal do Nivelamento RFVS, com âncora RFS aberta.	70
Figura 17: Feedback negativo na etapa Nivelamento.	71
Figura 18: Feedback positivo da etapa Nivelamento.	71
Figura 19: Feedback final desempenho Insatisfatório.....	72
Figura 20: Feedback final desempenho Bom.	73
Figura 21: Feedback final desempenho Excelente.	73
Figura 22: Exemplos de casos simulados durante testes do modelo de decisão inteligente. ..	75
Figura 23: Profissional de saúde realizando simulação no DeglutSIM.	77
Figura 24: Boxplot com a usabilidade dos participantes por item SUS.....	82

LISTA DE TABELAS

Tabela 1: Comparações entre os acertos e a correspondência com o coeficiente Kappa.	47
Tabela 2: Comparações entre os acertos e a correspondência com a Acurácia.	48
Tabela 3: Descrição da pontuação final dentro do simulador.	49
Tabela 4: Distribuição dos participantes quanto ao gênero, anos de atuação em Disfagia, experiência em VED e formação profissional (n=10).	78
Tabela 5: Scores SUS e classificação correspondente de acordo com cada participante da pesquisa e média geral dos scores.	79
Tabela 6: Cálculo do Kappa ponderado para cada par de avaliadores dos itens SUS.	80

LISTA DE QUADROS

Quadro 1: Descrição dos estudos que realizam treinamento da classificação dos resíduos faríngeos na VED.....	24
Quadro 2: Critérios sugeridos por Fehring (1987) com adaptação para a área de Disfagia.....	40
Quadro 3: Quantidade de trechos de vídeos separadas pela classificação dos resíduos faríngeos nas regiões de valéculas e seios piriformes.....	41
Quadro 4: Classificação para os valores do coeficiente Kappa (LANDIS; KOCH, 1977) e acurácia (HOSMER; LEMESHOW, 2000).....	44
Quadro 5: Requisitos do especialista.....	57
Quadro 6: Requisitos computacionais para a etapa Ensino.....	59
Quadro 7: Requisitos computacionais para a etapa Nivelamento.....	60
Quadro 8: Requisitos computacionais para a etapa Prática.....	61

LISTA DE ABREVIATURAS E SIGLAS

AVC	Acidente Vascular Cerebral
CCM	Centro de Ciências Médicas
CEP	Comitê de Ética em Pesquisa
CNE	Conselho Nacional de Educação
IHC	Interação Humano-Computador
LabTEVE	Laboratório de Tecnologias para o Ensino Virtual e Estatística
RCE	Requisitos computacionais do Ensino
RCN	Requisitos computacionais do Nivelamento
RCP	Requisitos computacionais da Prática
RFS	resíduos faríngeos em seios piriformes
RFV	resíduos faríngeos em valéculas
RFVS	Resíduos faríngeos em valéculas e seios piriformes
SBR	Sistema baseado em regras
SPSS	<i>Statistical Package for the Social Sciences</i>
ST	Simulador de treinamento
SUS	<i>System Usability Scale</i>
UFPB	Universidade Federal da Paraíba
VED	Videoendoscopia da deglutição
VFD	Videofluoroscopia da deglutição
YPRSRS	<i>Yale Pharyngeal Residue Severity Rating Scale</i>

SUMÁRIO

1. INTRODUÇÃO	15
1.1. PROBLEMA DE PESQUISA.....	15
1.2. ARGUMENTO DE TESE.....	18
1.3. MOTIVAÇÃO E RELEVÂNCIA	18
1.4. OBJETIVOS	20
1.5. CONTRIBUIÇÕES	21
2. FUNDAMENTAÇÃO TEÓRICA	22
2.1. AVALIAÇÃO E TREINAMENTO DA CLASSIFICAÇÃO DOS RESÍDUOS FARÍNGEOS NA VED.....	22
2.1.1. Estrutura do treinamento	22
2.1.2. Avaliação do treinamento	26
2.1.3. Treinamento mediado por tecnologias digitais.....	27
2.2. SIMULAÇÃO COMPUTACIONAL PARA TREINAMENTO EM SAÚDE..	29
2.3. MODELOS DE DECISÃO INTELIGENTES PARA A AVALIAÇÃO DE HABILIDADES EM SAÚDE	32
2.4. CONSIDERAÇÕES	35
3. MÉTODO	37
3.1. DESCRIÇÃO DO PROCESSO METODOLÓGICO	37
3.2. REVISÃO DA LITERATURA	37
3.3. DEFINIÇÃO DE PARÂMETROS PARA A AVALIAÇÃO DE UM TREINAMENTO PARA A CLASSIFICAÇÃO DOS RESÍDUOS FARÍNGEOS NA VED.....	38
3.3.1. Extração dos requisitos, variáveis e valores para o sistema de avaliação	38
3.3.2. Estrutura do treinamento e métricas para a avaliação do usuário.....	42
3.4. PROPOSIÇÃO DE UM MODELO DE DECISÃO INTELIGENTE	44
3.4.1 Confiabilidade intra-avaliador	47
3.4.2 Precisão.....	48
3.4.3 Pontuação final	48
3.5. DESENVOLVIMENTO DO ST	49
3.5.1. Definição do conteúdo.....	49
3.5.2. Construção das telas	50
3.5.3. Implementação do simulador e incorporação ao modelo de decisão	51
3.6. VALIDAÇÃO DA USABILIDADE DO ST.....	51
3.6.1. Cenário, população e amostra.....	51
3.6.2. Coleta de dados.....	53
3.6.3. Análise dos dados	54
3.6.4. Aspectos Éticos da Pesquisa	55
4. RESULTADOS E DISCUSSÃO	56
4.1. REVISÃO DE ESCOPO DA LITERATURA	56
4.2. DESENVOLVIMENTO DO SIMULADOR DE TREINAMENTO DEGLUTSIM.....	56

4.2.1. Análise dos Requisitos.....	56
4.2.2. Fluxo do sistema.....	62
4.2.3. DeglutSIM	65
4.3. MODELO DE DECISÃO INTELIGENTE	74
4.4. AVALIAÇÃO DO ST	76
4.4.1. Caracterização da amostra	77
4.4.2. Usabilidade	79
4.4.3. Aspectos positivos, negativos e sugestões	84
 5. CONSIDERAÇÕES FINAIS.....	 89
 6. REFERÊNCIAS	 91
 7. APÊNDICES	 102
APÊNDICE 1 – Dados demográficos do usuário.....	102
APÊNDICE 2 – Avaliação da usabilidade	103
APÊNDICE 3 – Termo de consentimento livre e esclarecido	105
APÊNDICE 4 – Categorização das respostas qualitativas dos profissionais sobre a experiência no simulador DeglutSIM	108
 8. ANEXOS.....	 112
ANEXO 1 – Registro do programa de computador DeglutSIM junto ao Instituto Nacional de Propriedade Industrial (INPI).....	112
ANEXO 2 – Parecer de aprovação do comitê de ética e pesquisa	113

1. INTRODUÇÃO

1.1. PROBLEMA DE PESQUISA

A avaliação da deglutição é composta por duas vertentes, a avaliação clínica, que envolve entrevista inicial, avaliação estrutural e avaliação funcional da deglutição; e a avaliação instrumental, que considera principalmente os exames de imagem videofluoroscopia e videoendoscopia da deglutição (VFD e VED), além de outros exames complementares para o diagnóstico em disfagia orofaríngea. Com base nas impressões obtidas durante a avaliação clínica da deglutição, o profissional pode decidir sobre a necessidade de realizar uma avaliação de caráter perceptivo-visual, como a avaliação instrumental (FURKIM, 2005; SANTOS, 2015).

Uma avaliação instrumental da disfagia refere-se ao uso de equipamentos especializados de imagem para a investigação de mecanismos envolvidos na deglutição. A videoendoscopia da deglutição é um exame instrumental que tornou-se bem estabelecido com o passar dos anos, e que registra a deglutição dos pacientes através de vídeos em cores utilizando um laringoscópio flexível, passado por via transnasal na faringe. O exame permite a visualização em tempo real de estruturas da região laringofaríngea e da dinâmica que ocorre durante a fase faríngea da função de deglutição (HISS, 2003; LANGMORE, 2017).

Por propiciar informações importantes sobre o processo de deglutição, o diagnóstico e o planejamento terapêutico da disfagia orofaríngea costumam considerar a classificação de parâmetros da deglutição visualizados por meio da VED (BRADY, 2007). Estes parâmetros atribuem classificações a variáveis perceptivo-visuais – aspectos que podem ser interpretados através da visão. Os aspectos englobam o tempo (duração percebida em um evento), o espaço (localização percebida em relação à anatomia, tamanho e proporção de um indicador clínico), o volume (quantidade de bolo alimentar ou secreções) e as respostas do paciente, como tosse ou engasgo (SWAN et al., 2019).

A diretriz estabelecida por Langmore (2001) para realizar a VED a descreve em três partes, sendo que na parte 1 observa-se a anatomia, secreções, o movimento de algumas estruturas e sensibilidade; na parte 2, é realizada a avaliação direta da deglutição, à medida que o paciente ingere diversas consistências do bolo alimentar; e na parte 3, são testadas mudanças dietéticas e posturais à medida que são percebidas alterações na deglutição. Um parâmetro que pode ser observado na parte 2 e costuma receber atenção nos estudos sobre VED é o acúmulo de resíduos faríngeos.

O acúmulo de resíduos faríngeos acontece quando a amplitude do movimento de deglutição não ocorre em sua totalidade, impedindo a eliminação completa do bolo alimentar pela hipofaringe em direção ao esôfago. O resíduo indica uma redução da depuração do bolo alimentar e pode ser uma consequência da diminuição da constrição faríngea, da retração da base de língua e/ou da excursão laríngea. Tais condições podem estar associadas a acometimentos neurológicos, câncer e/ou tratamentos oncológicos, ou pelo processo de envelhecimento. A formação de resíduos faríngeos em região de valéculas ou de seios piriformes após a deglutição está relacionada à ineficiência da deglutição. Desta forma, a visualização do resíduo é significativa para o diagnóstico em disfagia orofaríngea e indicativa para o tratamento da deglutição (LANGMORE, 2017).

Escala baseada em percepção visual já foram propostas para classificar os resíduos faríngeos na VED (CURTIS et al., 2022; KANEOKA et al., 2013; NEUBAUER; RADEMAKER; LEDER, 2015; SABRY; COYLE; ABOU-ELSAAD, 2021). Apesar do crescente número de publicações nesse sentido, ressalta-se a variabilidade na interpretação e a falta de consenso sobre critérios padronizados para a classificação (PISEGNA; LANGMORE, 2016; ROMMEL; HAMDY, 2016; SWAN et al., 2019).

Analisar a gravidade dos resíduos faríngeos é considerada uma tarefa complexa (LANGMORE, 2017). A classificação de imagens da deglutição baseia-se essencialmente no julgamento visual, sendo, portanto, de natureza subjetiva. Além disso, essa análise é influenciada por fatores como a experiência do avaliador, a consistência do bolo alimentar e a gravidade da disfagia (FROWEN; COTTON; PERRY, 2008; HIND et al., 2009; PILZ et al., 2016; PISEGNA, 2022; SCOTT et al., 1998).

Uma tentativa de melhorar a confiabilidade da classificação dos parâmetros perceptivo-visuais da deglutição é a proposta de treinamento. A literatura sugere que treinar a habilidade perceptivo-visual da classificação dos resíduos faríngeos na VED apresenta resultados favoráveis para a confiabilidade intra e interavaliadores (CURTIS et al., 2022; GERSCHKE et al., 2019; KANEOKA et al., 2013; NEUBAUER; RADEMAKER; LEDER, 2015). Um estudo constatou que um treinamento de 3 horas, onde os participantes classificaram resíduos faríngeos até obterem resultados próximos aos especialistas, mostrou excelente confiabilidade interavaliadores (KANEOKA et al., 2013), enquanto outro estudo ofereceu treinamento por meio de uma palestra instrucional em vídeo de 30 minutos e demonstrou concordância substancial entre os participantes (WARNECKE et al., 2009).

Os requisitos para o treinamento da classificação dos resíduos faríngeos variam de acordo com o público-alvo ao qual está direcionado. Para participantes experientes, são considerados a quantidade de anos de experiência e formação específica. Quando se trata de um treinamento voltado à profissionais ou estudantes inexperientes, cursos com conteúdo sobre anatomia e fisiologia da deglutição, fundamentos da VED e instrução sobre a escala que será utilizada são requeridos (BRADY, 2007; CURTIS et al., 2022; KANEOKA et al., 2013; SABRY; COYLE; ABOU-ELSAAD, 2021).

Assim como ocorre no ensino em várias práticas na área de saúde, no âmbito da educação em VED, a observação em tempo real é um elemento recorrente do processo de formação (BRADY et al., 2018). Contudo, um desafio inerente ao método de observação é que a realização do exame em um ambiente onde o estudante se encontra fisicamente presente nem sempre oferece a oportunidade de testemunhar manifestações clínicas mais atípicas. Além disso, as demandas imediatas de assistência ao paciente têm prioridade sobre o ensino, portanto a observação ao vivo pode não proporcionar uma experiência de aprendizagem suficientemente envolvente (LENCHUS, 2010).

Considerando a influência do treinamento perceptivo-visual para a classificação dos resíduos faríngeos, ressalta-se a importância da determinação de um modelo de treinamento para capacitar estudantes e profissionais de saúde que seja capaz de mensurar o desempenho da habilidade perceptivo-visual. Nos treinamentos atuais, existem inconsistências quanto à estrutura do treinamento fornecido, forma de apresentação dos exames para a classificação (vídeo/foto), quantidade de exames apresentados para a avaliação, bem como a falta de critérios padronizados para considerar a habilidade para realizar esse tipo de análise. Os avanços tecnológicos e o treinamento computacional vêm se tornando parte integrante do desenvolvimento de habilidades nas profissões da área da saúde (LIU et al., 2016), o que pode permitir oportunidades de treinamento estendidas para classificar parâmetros da deglutição na VED, em particular, os resíduos faríngeos.

A simulação computacional é o processo de utilização de modelos matemáticos e algoritmos computacionais para imitar ou reproduzir o comportamento de sistemas ou processos do mundo real. Na área de saúde, simuladores de treinamento computacionais oferecem experiências controladas no ambiente profissional, sendo uma ferramenta de ensino-aprendizagem eficiente que possibilita um espaço seguro para erros e repetições de procedimentos sem riscos ao paciente (ALANAZI; NICHOLSON; THOMAS, 2017; MACHADO; COSTA; MORAES, 2018). Além disso, os simuladores computacionais

viabilizam a avaliação de desempenho em tempo real, baseada na interação usuário-máquina (ANDRADE et al., 2022b; MORAES; FERREIRA; MACHADO, 2021).

Desenvolver um simulador de treinamento para treinar a habilidade perceptivo-visual de classificar os resíduos faríngeos na VED pode trazer contribuições na formação de estudantes ou no aprimoramento de habilidades de profissionais de saúde na área de Disfagia. O treinamento computacional pode ajudar a melhorar a consistência das avaliações, melhorar o uso do tempo dos profissionais, diminuir despesas com o treinamento e permitir uma avaliação mais rápida e precisa do desempenho dos avaliadores, se comparado ao treinamento tradicional.

Diante do exposto, considera-se a necessidade da implementação de um treinamento perceptivo-visual para a classificação dos resíduos faríngeos na VED por meio da simulação computacional, utilizando estratégias de treinamento que permitam atingir precisão e confiabilidade aceitáveis para os avaliadores.

1.2. ARGUMENTO DE TESE

Defende-se a tese de que é possível proporcionar e avaliar o treinamento perceptivo-visual para a classificação do parâmetro resíduos faríngeos na VED através da simulação computacional.

1.3. MOTIVAÇÃO E RELEVÂNCIA

O estudo da deglutição por meio da VED é considerado padrão de referência por alguns autores (BRADY; DONZELLI, 2013; NOORDALLY et al., 2011). No exame, imagens reais da deglutição são fornecidas, possibilitando a identificação do bolo alimentar ou líquidos residuais e a sua quantidade na região laringofaríngea após a deglutição (PRIKLADNICKI; SANTANA; CARDOSO, 2022). A visualização do acúmulo do bolo alimentar neste dado momento é denominada como o parâmetro perceptivo-visual “resíduos faríngeos”, medida que recebe atenção durante os exames de VED devido à sua importância para a eficiência da deglutição (LANGMORE, 2017).

Apesar da vasta utilização e aplicabilidade no contexto clínico, existem preocupações quanto à qualidade das propriedades psicométricas e a reprodutibilidade da classificação dos parâmetros perceptivo-visuais da VED, visto a variabilidade na sua interpretação (PILZ et al., 2016; PISEGNA; LANGMORE, 2016; SWAN et al., 2019). Diversas escalas para mensurar

parâmetros da deglutição vem sendo propostas, mas a criação de mais categorias e escores provavelmente não garantirá a melhora na reprodutibilidade (PISEGNA et al., 2018).

Além disso, fatores como o nível de conhecimento em disfagia orofaríngea e o desenvolvimento de habilidades adequadas para a interpretação da VED são importantes para esta avaliação, de forma que profissionais sem treinamento podem acabar realizando recomendações inadequadas que podem acarretar eventos adversos relacionados ao manejo da disfagia (SCHINDLER et al., 2022), como uma indicação dietética imprópria para o estado clínico do paciente, o que pode colocar em risco sua saúde pulmonar. De forma geral, a deficiência no manejo da disfagia pode ter consequências negativas significativas para a saúde e a qualidade de vida do paciente, como o aumento do tempo de internação hospitalar (ALTMAN; YU; SCHAEFER, 2010).

O treinamento como estratégia de aprendizagem pretende melhorar o desempenho de indivíduos por meio de situações que facilitem a aquisição e retenção de conhecimentos e habilidades (BORGES-ANDRADE; ABBAD; MOURÃO, 2009). Treinar profissionais de saúde para classificar parâmetros da deglutição na VED através da adoção de um protocolo padronizado é recomendado pela literatura, e pode aumentar a confiabilidade da avaliação (CURTIS et al., 2022; GERSCHKE et al., 2019; KANEOKA et al., 2013; LANGMORE, 2017; NEUBAUER; RADEMAKER; LEDER, 2015; SCHINDLER et al., 2022).

Considerando as novas tecnologias, o desenvolvimento de simuladores de treinamento computacionais é aplicável para a classificação do parâmetro resíduos faríngeos na VED. A proposição de um simulador de treinamento (ST) com este propósito deve envolver o conhecimento sobre como é realizado comumente o treinamento. Criar um modelo de decisão que atue como modelo de avaliação no contexto do treinamento trará uma contribuição importante para facilitar o aprendizado, definir métricas relacionadas ao treinamento, e contribuirá para o desenvolvimento e aprimoramento de pesquisas futuras, inclusive envolvendo outros parâmetros considerados na avaliação da deglutição na VED.

A estruturação de um ST com este propósito se torna desafiadora ao considerar um procedimento de análise subjetiva, sem métricas bem estabelecidas e onde não há consenso entre os pesquisadores da área, como é o caso dos parâmetros da deglutição analisados pela VED. Dessa forma, acredita-se que o desenvolvimento de um ST para classificar os resíduos faríngeos na VED pode auxiliar na capacitação de estudantes e profissionais de saúde na realização deste procedimento avaliativo, permitindo o conhecimento da influência das variáveis envolvidas, como a estrutura de treinamento (quantidade de sessões, número de horas,

número de exames apresentados, método de ensino), as características dos exames apresentados durante o treinamento e o método utilizado para avaliar a habilidade do participante.

Considerando o exposto, nota-se a importância de propor e utilizar métodos apropriados para avaliar o aprendizado na classificação dos resíduos faríngeos na VED. A avaliação do desempenho neste tipo de análise se mostra um tópico interessante e necessário no cenário mundial, onde a tecnologia poderá auxiliar permitindo uma capacitação mais acessível financeiramente, que proporciona maior flexibilidade de tempo para a sua utilização e oferece *feedback* para o participante.

1.4. OBJETIVOS

O objetivo principal deste trabalho consiste em propor e desenvolver um simulador computacional para treinamento perceptivo-visual capaz de avaliar a habilidade do usuário em realizar a classificação dos resíduos faríngeos na videoendoscopia da deglutição (VED).

Para alcançar este objetivo, será apresentada uma proposta de simulador computacional que visa proporcionar treinamento e avaliação da habilidade nesta avaliação em saúde. Com esta proposta, visa-se representar computacionalmente o momento de classificação diagnóstica da gravidade dos resíduos faríngeos na deglutição dos pacientes durante a realização do exame de VED. Esta abordagem busca preparar estudantes e profissionais de saúde envolvidos no diagnóstico em disfagia orofaríngea para realizarem uma conduta mais assertiva. Nesta perspectiva, os objetivos específicos delineados durante o projeto de tese foram:

- I.** Identificar as características do treinamento perceptivo-visual para a classificação dos parâmetros da VED;
- II.** Definir parâmetros para a avaliação de um treinamento para a classificação dos resíduos faríngeos na VED;
- III.** Propor um modelo de decisão inteligente para a avaliação do desempenho do usuário em um simulador de treinamento (ST), com foco na classificação dos resíduos faríngeos na VED;
- IV.** Implementar um simulador contendo um sistema de avaliação para o treinamento da classificação dos resíduos faríngeos na VED, utilizando os parâmetros previamente definidos.
- V.** Validar a usabilidade do ST.

1.5. CONTRIBUIÇÕES

Este trabalho apresenta algumas contribuições, que podem ser vistas das óticas científica, educacional, tecnológica e social. A primeira contribuição refere-se à proposição de métricas para o treinamento da classificação dos resíduos faríngeos na VED. As métricas existentes na literatura atualmente não estão bem estabelecidas.

A segunda contribuição, considerando a ótica tecnológica, refere-se à apresentação de um modelo de decisão que possa ser usado em um ST desenvolvido com o propósito apresentado, e que atua no processo de avaliação do desempenho do usuário. Neste sentido, ainda não foi encontrado nenhum trabalho envolvendo ST's voltados à classificação dos parâmetros da deglutição na VED. Espera-se que o modelo desenvolvido possa ser utilizado em outros ST's.

Por fim, considerando as óticas educacional e social, o ST pode atuar na capacitação profissional de estudantes e profissionais de saúde que assistem pacientes com disfagia. O ST pode ser utilizado como recurso pedagógico para desenvolver a habilidade perceptivo-visual dos usuários/profissionais, de modo a tornarem-se mais seguros e assertivos para o diagnóstico funcional da deglutição.

2. FUNDAMENTAÇÃO TEÓRICA

Neste tópico apresenta-se a fundamentação teórica que norteia a construção desta tese. Para isto, são discutidos em subtópicos os seguintes temas: 1) Avaliação e treinamento da classificação dos resíduos faríngeos na VED; 2) Simulação computacional para treinamento de habilidades em saúde; e 3) Modelos de decisão para a avaliação de habilidades em saúde.

2.1. AVALIAÇÃO E TREINAMENTO DA CLASSIFICAÇÃO DOS RESÍDUOS FARÍNGEOS NA VED

2.1.1. Estrutura do treinamento

Treinar a realização da VED tem sido uma recomendação vigente no cenário mundial, de forma que em alguns países instituições desenvolveram currículos de treinamento estruturados e certificados com esta finalidade (DZIEWAS et al., 2017). A Sociedade Alemã de Neurologia e a Sociedade Alemã de Acidente Vascular Encefálico propõem um currículo de treinamento envolvendo habilidades práticas e habilidades clínicas de base cognitiva (pensamento crítico, tomada de decisão, resolução de problemas, conhecimento básico sobre disfagia), consideradas cruciais para compreender e interpretar a deglutição a partir de uma visão endoscópica. A proposta envolve a identificação e classificação de alguns parâmetros da deglutição, como os resíduos faríngeos, a penetração/aspiração e o escape prematuro do alimento (DZIEWAS et al., 2016).

De igual forma, a Associação Americana de Fonoaudiologia (*ASHA*, do inglês) desenvolveu um tutorial para recomendar diretrizes de melhores práticas para fonoaudiólogos que realizam e interpretam/classificam os parâmetros da VED em adultos (LANGMORE et al., 2022). Programa semelhante foi proposto no Japão (YOSHIDA et al., 2020).

Estas recomendações, embora pretendam orientar a prática profissional da VED, buscando estabelecer melhores padrões de atendimento em seus respectivos países, possuem foco maior nas habilidades práticas ou psicomotoras do procedimento, como a inserção do endoscópio pela via nasal. Os programas consideram que os parâmetros da deglutição devem ser identificados e classificados. Porém, o nível de desempenho necessário nas classificações para certificar a habilidade dos participantes não costuma ser claramente informado.

O interesse reduzido no aprimoramento da classificação dos parâmetros relacionados aos eventos da deglutição em alguns destes programas pode ser um reflexo da cultura da realização do procedimento em cada nação. A VED foi desenvolvida por uma fonoaudióloga, e

nos Estados Unidos e em alguns países da Europa, a parte prática do exame está dentro do escopo deste profissional (LANGMORE; SCHATZ; OLSEN, 1988). Na Alemanha, o neurologista comumente realiza o procedimento, apesar de qualquer fonoaudiólogo treinado poder realizar (DZIEWAS et al., 2016). Na maior parte da América do Sul, incluindo o Brasil, o médico otorrinolaringologista assumiu a liderança, com o fonoaudiólogo desempenhando um papel de apoio diagnóstico. Nesta conjuntura, o fonoaudiólogo pode participar da análise e classificação dos parâmetros da deglutição, mas não realiza a parte prática do exame.

Apesar de a condução da VED requerer habilidade nos domínios cognitivo e psicomotor, os estudos que envolvem treinamento em VED acabam por priorizar habilidades psicomotoras devido à natureza técnica do procedimento, que exige coordenação mão-olho e a manipulação correta do endoscópio. Dessa forma, embora o conhecimento teórico e habilidades de interpretação dos achados do exame sejam importantes, a habilidade na realização técnica da VED é essencial para a segurança física do paciente durante o exame e previne efeitos adversos para a saúde (ROBINSON, 2021).

O médico é o profissional qualificado para diagnosticar e tratar doenças que podem estar relacionadas à disfagia orofaríngea e o fonoaudiólogo é o profissional que tem maior conhecimento sobre os distúrbios da deglutição e que oferece tratamentos para esta condição, além de liderar pesquisas em deglutição e disfagia (LANGMORE, 2017). As lacunas que existem na formação e qualificação de profissionais de saúde que possam avaliar e tratar pacientes com disfagia indicam a necessidade de ampliar a instrução e o treinamento em VED considerando todas as habilidades que este procedimento exige, para que mais pacientes possam receber a assistência que necessitam, independente da formação de base do profissional (BAIJENS et al., 2016; DZIEWAS et al., 2016; SCHINDLER et al., 2022).

Em outro cenário, estudos que propõem e procuram validar escalas de classificação dos parâmetros da deglutição tendem a incluir como fase adicional do trabalho o treinamento perceptivo-visual para melhorar a habilidade perceptivo-visual de classificar os parâmetros utilizando estas escalas. O Quadro 1 apresenta estudos encontrados envolvendo treinamento para classificação do parâmetro resíduos faríngeos e os descreve quanto aos objetivos, estrutura do treinamento, características dos exames do treinamento e resultados obtidos a partir do treinamento.

Quadro 1: Descrição dos estudos que realizam treinamento da classificação dos resíduos faríngeos na VED (continua).

Autor(es), ano, país(es)	Objetivo(s)/ parâmetro(s) considerado(s)	Estrutura do treinamento	Características dos exames do treinamento	Resultados relacionados ao treinamento
KANEOKA et al., 2013, Estados Unidos	Examinar a confiabilidade e validade da <i>Boston Residue and Clearance Scale</i> (BRACS). Considerou o parâmetro resíduos faríngeos.	Sessão de 3 horas para classificação independente dos exames pelos participantes e comparação à especialista de referência. Se havia discrepância na pontuação, o treinamento continuava. Quando a pontuação dos participantes e do especialista estava dentro de uma faixa de 3 pontos na pontuação da BRACS nas três consistências consideradas, os participantes foram considerados proficientes.	Não relata características e quantidade dos vídeos usados no treinamento. No experimento após o treinamento, foram usados 63 exames de VED de 51 pacientes, com diferentes patologias de base e níveis de gravidade distintos de resíduos faríngeos.	Os participantes foram considerados proficientes após três horas de treinamento, quando a pontuação de todos eles e do especialista estava dentro de uma faixa de 3 pontos na pontuação total da BRACS nas três consistências consideradas.
NEUBAUER; RADEMAKER; LEDER, 2015, Estados Unidos	Desenvolver, padronizar e obter evidências de validade da <i>Yale Pharyngeal Residue Severity Rating Scale</i> . Considerou o parâmetro resíduos faríngeos.	Treinamento em etapa única que incluiu: definições escritas, representações visuais, explicações verbais e perguntas/respostas esclarecedoras sobre o sistema de classificação. Não referiu tempo de duração do treinamento.	Não relata as características das imagens usadas no treinamento. No experimento após treinamento, foram usadas 25 imagens, sendo uma sem resíduo e três imagens de cada um dos quatro níveis da escala para valéculas e seios piriformes.	Proficiência determinada por análise de confiabilidade usando o coeficiente kappa. Os participantes apresentaram maior confiabilidade intra-avaliador para classificação de resíduos em valécula. A confiabilidade interavaliador e a comparação com os especialistas foram favoráveis ao treinamento na classificação de resíduos faríngeos em valéculas e seios piriformes.
GERSCHKE et al., 2019, Alemanha	Validar a versão alemã da <i>Yale Pharyngeal Residue Severity Rating Scale</i> e investigar o impacto da experiência e treinamento do avaliador. Considerou o parâmetro resíduos faríngeos.	Tutorial em vídeo de oito minutos, incluindo uma introdução audiovisual ao conceito das escalas, representações e descrições dos níveis de gravidade e exercícios de classificação com imagens. Nenhuma imagem que seria usada no experimento principal foi usada no treinamento.	Não relata as características das imagens usadas no treinamento. No experimento principal, foram usadas 30 imagens, sendo 15 de valéculas e 15 de seios piriformes.	Proficiência determinada pela análise de confiabilidade usando o coeficiente kappa. Os participantes treinados apresentaram maior confiabilidade interavaliador para classificação de resíduos em valécula ($k = 0,964$).

Quadro 1: Descrição dos estudos que realizam treinamento da classificação dos resíduos faríngeos na VED (conclusão).

Autor(es), ano, país(es)	Objetivo(s)/ parâmetro(s) considerado(s)	Estrutura do treinamento	Características dos exames do treinamento	Resultados relacionados ao treinamento
CURTIS et al., 2022, Estados Unidos e Nova Zelândia	Descrever o desenvolvimento da <i>Visual Analysis of Swallowing Efficiency and Safety</i> (VASES) - e explorar a viabilidade de treinar novatos para interpretar a VED usando a VASES. Considerou os parâmetros resíduos faríngeos e penetração/ aspiração	Uma fase de treinamento dividida em cinco partes: Apresentação de regras e definições operacionais; 2) Prática de classificação de cinco vídeos; 3) Vídeo de 60 minutos; 4) Prática de classificação de cinco vídeos; 5) Sessão em grupo ao vivo de 60 minutos com discussão com um dos avaliadores de referência. Número médio de 6 horas para concluir (mínimo de 4h, máximo de 20h).	35 vídeos de exames de VED, sendo 10 deles repetidos para permitir análise de confiabilidade intravaliador.	Proficiência determinada pela precisão nas classificações, considerando a avaliação dos especialistas. A análise da confiabilidade (Kappa de Cohen - κ_w), foi investigada como medida secundária. Houve melhora na precisão da classificação das medidas de resultado. A confiabilidade interavaliadores aumentou para quatro das sete medidas de resultado, atingindo o ponto de corte estabelecido ($\geq 50\%$ das sete medidas demonstrarem aumento na precisão).
ROCCA et al., 2024, Itália	Avaliar a eficácia do treinamento para a avaliação de resíduos faríngeos em quadros e vídeos de VED por meio da <i>Yale Pharyngeal Residue Severity Rating Scale</i> . Considerou o parâmetro resíduos faríngeos.	Treinamento em 3 etapas, num total de 160 minutos: (1) Videoaula gravada de 75 minutos envolvendo módulos teóricos sobre VED, resíduos faríngeos, aplicação da escala e estudos de caso; (2) Prática de classificação em 6 pares de quadros e vídeos de VED; (3) Reuniões em grupo com um especialista para esclarecimento de dúvidas.	30 pares de vídeos e frames (15 para valéculas e 15 para seios piriformes) e 6 pares de vídeos para análise de validade e confiabilidade, e 6 pares de vídeos e frames para o treinamento.	Proficiência determinada pela validade de construto, que considera a concordância entre a avaliação dos participantes e a dos especialistas; e a confiabilidade interavaliadores. A validade de construto melhorou significativamente após treinamento no grupo treinamento para vídeos dos seios piriformes. No grupo de estudantes, melhorou significativamente para as valéculas e vídeos de seios piriformes. Não foram encontradas diferenças significativas na confiabilidade interavaliadores em nenhum grupo.

A partir da análise dos estudos apresentados no Quadro 1, observa-se que existe heterogeneidade na estrutura dos treinamentos (quantidade de sessões, número de horas, número de exames apresentados, método de ensino), nas características dos exames apresentados para serem classificados no treinamento e no método utilizado para avaliar a habilidade do participante.

As características dos exames apresentados tanto no treinamento quanto para a avaliação do treinamento diferem nos estudos apresentados, e muitas vezes não são descritas. Os critérios de escolha da variabilidade dos casos para a avaliação do treinamento nos estudos variam entre diagnóstico do paciente, níveis de gravidade na escala de classificação adotada ou a região anatômica.

Em síntese, não existe uma estrutura de referência nos treinamentos propostos para classificar os parâmetros da deglutição na VED.

2.1.2. Avaliação do treinamento

No que concerne às medidas de resultado utilizadas para a avaliação do treinamento para a classificação dos resíduos faríngeos na VED, a medida que costuma ser considerada na maioria dos estudos descritos no Quadro 1 é a confiabilidade, seja intra ou interavaliadores. A precisão também é considerada em alguns estudos (CURTIS et al., 2022; KANEOKA et al., 2013).

Como os estudos apresentados têm como foco principal a validação das escalas propostas, a avaliação do desempenho do participante normalmente ocorre após o treinamento. Os resultados são publicados com o intuito de evidenciar que ao oferecer treinamento, os participantes com o perfil selecionado (iniciante, avançado) são capazes de utilizar a escala de forma confiável em sua prática clínica.

A análise de confiabilidade interavaliadores, descrita na maioria dos estudos apresentados como medida de resultado do treinamento, traz dados de incerteza. Com esta medida, torna-se difícil distinguir se houve aumento na precisão das classificações com o treinamento ou se os resultados de avaliadores inexperientes se influenciaram mutuamente para uma classificação propensa a erros, por exemplo. Já a medida de confiabilidade intra-avaliador mostra-se uma medida importante, visto que representa a consistência na classificação do avaliador em momentos distintos. Esta medida foi considerada em dois estudos envolvendo treinamento na classificação dos resíduos faríngeos na VED (CURTIS et al., 2022; NEUBAUER; RADEMAKER; LEDER, 2015).

Sobre a medida de acurácia ou precisão, Kaneoka e colaboradores (2013) estabeleceram um critério para considerar a proficiência do participante que o considera proficiente após a pontuação da escala BRACS (*Boston Residue and Clearance Scale*) estar dentro de uma faixa de até 3 pontos de diferença em relação à de um especialista. Outro estudo estabeleceu como critério de proficiência um incremento de pelo menos 50% na precisão das sete medidas consideradas no estudo (resíduos faríngeos em valéculas, seios piriformes, epiglote, vestíbulo laríngeo, pregas vocais e subglote; e o escore da *Penetration-Aspiration Scale*) para tornar o treinamento viável, baseando o acerto na classificação do especialista, resultado que foi atingido pós-treinamento (CURTIS et al., 2022).

Warnecke e colaboradores (2009) desenvolveram um treinamento utilizando um protocolo para avaliação da disfagia por meio da VED em pacientes com AVC agudo, considerando, entre os achados clínicos, os resíduos faríngeos. Embora tenham levado em conta a precisão das classificações, não foi estabelecido um ponto de corte para determinar a proficiência dos participantes através desta medida durante o treinamento. Ainda assim, durante o treinamento houve uma taxa de acerto de 88,8%, enquanto na avaliação do treinamento propriamente dita, houve 96% de acerto. Os autores julgaram a taxa de acerto adequada após a análise dos resultados.

A taxa de acerto (precisão) nos estudos mencionados é calculada a partir da comparação das classificações dos participantes com as classificações de um ou mais especialistas de referência, conforme a literatura recomendada para treinamentos em saúde (ANDRADE et al., 2022a).

Considerando o exposto, é de suma importância utilizar medidas de precisão e confiabilidade intra-avaliador no treinamento para classificar resíduos faríngeos na VED. Ao fornecer um índice de desempenho composto por essas medidas no relatório de desempenho do participante ao final do treinamento, garante-se respostas precisas e consistentes ao longo do tempo do treinamento. Este fator é especialmente relevante em contextos de saúde, onde a precisão dos dados pode impactar diretamente o diagnóstico e os resultados do tratamento dos pacientes.

2.1.3. Treinamento mediado por tecnologias digitais

O treinamento na área da saúde por meio de tecnologias digitais representa uma inovação na forma que os profissionais são instruídos. Essas tecnologias permitem a aquisição de conhecimentos e habilidades de maneira flexível, independente da sua localização

geográfica, o que facilita a educação continuada e melhora a educação inicial dos profissionais de saúde, contribuindo para melhorar a qualidade do atendimento à população (FREHYWOT et al., 2013).

Dentre as tecnologias digitais voltadas para o ensino e aprendizagem, o termo *e-learning* ou aprendizagem eletrônica refere-se a todo o conjunto de conhecimentos que podem ser acessados por meio da internet ou em uma intranet (BINTI; ZAIDI; HOQUE, 2019). Este formato permite o acesso a cursos, materiais didáticos, vídeos e a participação em atividades interativas *online*, contando com ferramentas que facilitam o processo de educação à distância.

No âmbito da educação em VED, um estudo avaliou a eficácia de um treinamento *online*, interdisciplinar e interativo que foi projetado para melhorar a capacidade de interpretar parâmetros da deglutição de forma precisa. Estudantes de residência e de pós-graduação em Fonoaudiologia tiveram ganhos de conhecimento semelhantes em treinamentos *online* e presenciais, com desempenho similar ao de especialistas. Contudo, os autores destacaram algumas vantagens do curso *online* em relação ao curso presencial, como a possibilidade de escolha do horário mais conveniente, a opção de controle da mídia e do ritmo do curso e a padronização do conteúdo do curso, que muitas vezes não é possível no curso presencial, dado o formato de sala de aula (BRADY et al., 2018).

Considerando ainda estudos que empregaram tecnologias digitais no ensino em VED, Gerschke e colaboradores (2019) conduziram um treinamento para a classificação da gravidade dos resíduos faríngeos através de uma plataforma *online*, o que demonstra a possibilidade de utilizar recursos computacionais para este propósito. Dziejwas e colaboradores (2017) também incorporaram com sucesso um método *e-learning* de aprendizagem com duração de doze horas sobre fisiopatologia, diagnóstico e tratamento da disfagia orofaríngea para complementar o treinamento prático no programa que propuseram para capacitar profissionais na realização da VED.

Um exemplo adicional de treinamento no estudo da deglutição, é o *Modified Barium Swallow Impairment Profile* (MBSImP), que envolve a videofluoroscopia da deglutição (VFD). O treinamento apresenta um protocolo padronizado para análise e avaliação de habilidade neste exame. Originalmente concebido como um pacote combinado de atividades *online* e ensino didático presencial, o MBSImP evoluiu para um formato de aprendizagem autodirigida totalmente *online*, atualmente disponível por meio de uma aplicação dedicada (MARTIN-HARRIS et al., 2008).

Nesse contexto, um outro programa de *e-learning* para estudo da deglutição utilizando a VFD foi desenvolvido para apoiar o treinamento de profissionais nesta avaliação. O programa ajudou a superar algumas barreiras como o acesso limitado à educação formal, demandas clínicas e a complexidade do exame, que reduzem as oportunidades de treinamento no local de trabalho. Resultados obtidos no estudo indicam que fonoaudiólogos veem o *e-learning* como uma ferramenta conveniente para uso, proporcionando treinamento *online* gratuito e eficaz, flexibilidade e acesso a recursos especializados que, de outra forma, seriam de difícil obtenção (BURNS et al., 2021).

Estudo realizado posteriormente corroborou ainda mais estes achados demonstrando que o treinamento *online* independente se mostrou tão eficaz quanto o treinamento apoiado por pares e o treinamento facilitado por especialistas para o desenvolvimento de habilidades de análise do exame para avaliadores iniciantes, apoiando a utilização deste método para preparar os alunos para o treinamento prático na VFD (EDWARDS et al., 2024).

Portanto, a integração de tecnologias digitais no treinamento de profissionais de saúde, especialmente no campo da deglutição, utilizando dados de avaliação instrumental da disfagia como a VED, oferece vantagens como flexibilidade, acessibilidade e padronização do conteúdo. A aplicação destas tecnologias não apenas complementa ou introduz o treinamento prático, mas assegura uma formação mais abrangente e contribui significativamente na melhora da qualidade do atendimento aos pacientes.

2.2. SIMULAÇÃO COMPUTACIONAL PARA TREINAMENTO EM SAÚDE

A maneira de ensinar no campo da saúde passou por uma transformação considerável globalmente. A educação médica tem experimentado alterações, adições e novas abordagens no processo de ensino-aprendizagem. O modelo tradicional de ensino, que se apoia em técnicas passivas, como aulas teóricas e provas escritas, tem demonstrado ser insuficiente para a retenção de conhecimento e na sua aplicação prática. É reconhecida, então, a necessidade de os estudantes vivenciarem situações reais, desenvolver ou melhorar habilidades e tomar decisões de forma ativa. Isso inclui atividades como a simulação, o aprendizado baseado em resolução de problemas, discussões de casos e trabalho em equipe (BRANDÃO; COLLARES; MARIN, 2014).

A utilização de tecnologias digitais é uma estratégia interessante que pode complementar a formação prática de estudantes e profissionais que desejam se aperfeiçoar. Os simuladores computacionais para treinamento em saúde podem oferecer uma nova forma de

aprendizagem, onde se utiliza a interatividade para tornar a simulação mais realista. Estas ferramentas utilizam um ambiente virtual gerado por computador para recriar uma situação ou procedimento do mundo real (SANTOS; MACHADO, 2009).

Aplicações neste sentido são desenvolvidas com o objetivo de treinar habilidades cognitivas, psicomotoras e afetivas. A insegurança dos profissionais sem experiência ao lidar com pacientes pela primeira vez, os custos envolvidos e os riscos ao paciente envolvendo o treinamento tradicional podem ser minimizados por meio do uso destes recursos (ANDRADE et al., 2022a).

Com o uso de simuladores computacionais, é possível exercitar habilidades cognitivas e alcançar proficiência em determinado procedimento antes de realizá-lo em pacientes reais, repetindo-o por quantas vezes forem necessárias, até que estejam devidamente capacitados (MORAES; MACHADO, 2007). Além disso, estas ferramentas podem ser utilizadas para simular uma maior variabilidade de casos clínicos, propiciando a oportunidade de desenvolver habilidades que seriam mais difíceis de serem treinadas em situações reais. Ainda é possível extrair dados para a avaliação do desempenho durante o treinamento de forma automatizada e fornecer *feedback* imediato ao participante, o que é denominado como “avaliação *online*” (MORAES; MACHADO, 2014; SOUZA et al., 2022).

A simulação computacional aumenta a confiança e habilidade ao oferecer aprendizado interativo e prático, o que por consequência promove melhora das competências clínicas. A identificação dos pontos fracos e pontos de melhoria do participante permite correções por meio do *feedback*, promovendo segurança, refinamento de habilidades e a identificação de lacunas de conhecimento e treinamento (ALMEIDA et al., 2021).

Considerando que muitas diretrizes educacionais ao redor do mundo sugerem a inclusão de tecnologias digitais no ensino da graduação para o ensino em cursos da área de saúde, um estudo validou o protótipo de um simulador para o teste de emissões otoacústicas, mais conhecido como Teste da Orelhinha, como uma ferramenta facilitadora do ensino da realização da avaliação auditiva de bebês para alunos do curso de Fonoaudiologia. Uma impressão da cabeça de um boneco com o tamanho de um recém-nascido foi associada a um aplicativo contendo a representação visual da tela do aparelho utilizado no exame, funcionando interligados a um fone de ouvido (MASSIH et al., 2023).

Para investigar a viabilidade de simular a deglutição do bolo alimentar, Farazi e colaboradores (2015) construíram um modelo 3D baseado em dados clínicos. O intuito do trabalho foi propiciar a visualização da cinemática da deglutição para estudantes. De acordo

com os autores, o modelo tem potencial para se tornar uma plataforma interativa 3D para o treinamento do *Modified Barium Swallow Impairment* (MBSImp), método de avaliação da deglutição a partir de imagens videofluoroscópicas.

Com o intuito de fornecer modelos tridimensionais realistas para simular o fluxo peristáltico em pacientes com disfagia, Gastelum et al. (2016) construíram um modelo 3D do trato gastrointestinal superior a partir de imagens coloridas obtidas de um banco de dados do *National Institutes of Health*. Este estudo discutiu o uso do modelo para simular a deglutição em função das propriedades do bolo alimentar, podendo servir como base para estudos posteriores voltados ao treinamento de estudantes para realizarem a VED ou outros procedimentos de reabilitação. Nenhum estudo envolvendo simuladores computacionais para classificar parâmetros da deglutição visualizados por meio da VED foi encontrado.

Considerando toda a interatividade presente nestes simuladores apresentados, a interação humano-computador (IHC) é a área da computação que se dedica ao estudo do *design*, avalia e desenvolve interfaces para proporcionar interação dos sistemas computacionais com os usuários finais de forma eficiente e intuitiva. A IHC também pode ser definida como o que acontece quando um ser humano e um sistema computacional unem-se para cumprir tarefas no intuito de realizar um objetivo. Esta interação pode ser representada por um processo de manipulação, comunicação, troca, influência, entre outros (DIX, 2009).

Por sua vez, a usabilidade é um conceito dentro da IHC que está relacionada à facilidade que um usuário tem ao interagir com determinado sistema sem demandar esforço excessivo (KRUG, 2013). No campo da saúde, existe uma preocupação maior com a usabilidade em tecnologias destinadas ao treinamento, uma vez que estão frequentemente ligadas a procedimentos relevantes para a vida humana. Os avanços nesta área exigem uma interação homem-máquina mais fidedigna, para que assim sejam garantidos segurança do paciente, eficiência no procedimento e satisfação por parte dos profissionais/usuários (SÁ et al., 2022).

Estudantes representam o principal público-alvo nos estudos envolvendo simuladores computacionais em saúde, mas é essencial que profissionais mantenham o aprimoramento contínuo de suas habilidades ao longo da vida. O acesso às ferramentas de educação permanente torna-se, portanto, fundamental para uma atuação eficaz. Simulações que não atingem os objetivos propostos se tornam irrelevantes para a prática clínica/educacional, o que torna o teste de *software*, incluindo a avaliação da usabilidade do simulador, uma etapa indispensável durante o desenvolvimento destas ferramentas (ANDRADE et al., 2022a; MACHADO; COSTA; MORAES, 2018).

2.3. MODELOS DE DECISÃO INTELIGENTES PARA A AVALIAÇÃO DE HABILIDADES EM SAÚDE

A tomada de decisão é um processo que envolve a escolha de uma ação ou estratégia que seja mais adequada para solucionar um problema identificado, e deve ser realizado de forma inteligente (FEENEY; WILLIAMSON; BISHOP, 2002). Na área da saúde, muitas vezes esse processo considera a subjetividade e experiência dos profissionais envolvidos, o que deve ser reduzido o quanto possível a partir da produção de evidências e desenvolvimento de métodos apropriados para minimizar os riscos de uma conduta (TANAKA; TAMAKI, 2012).

Considerando o desenvolvimento de simuladores computacionais, um modelo de decisão inteligente é capaz de atuar na avaliação dos processos de tomada de decisão do usuário (MORAIS, 2011). Na avaliação de habilidades em saúde, é possível retornar informações sobre o desempenho do usuário com base em dados coletados previamente. Os dados podem ser extraídos do conhecimento de especialistas da área, da literatura ou de forma combinada. Este último é recomendado, pois apesar da literatura científica ser um conhecimento tradicional e confiável para a formação profissional, a experiência prática pode agregar observações que contribuem para o realismo da simulação (ANDRADE et al., 2022a).

A avaliação de habilidades em um simulador computacional pode ocorrer *offline* ou *online*. A avaliação *offline* registra as ações do usuário no simulador e compara suas respostas com as de um ou mais especialistas, retornando um relatório desta comparação de forma não automática. Por outro lado, a avaliação *online* monitora as ações de acordo com os dados recolhidos dos especialistas, e o *feedback* ao usuário ocorre em tempo real ou em até um segundo após a finalização do treinamento (MACHADO; MORAES, 2010).

A avaliação de habilidades *online* geralmente ocorre utilizando como recursos mecanismos de alerta em tempo real, indicadores de progresso e/ou relatórios sobre o desempenho do usuário na simulação. Esses indicadores são formas de medir se o usuário alcançou o desempenho desejado. Além disso, eles também funcionam como agentes motivadores, à medida que incorporam elementos presentes em jogos eletrônicos (ANDRADE et al., 2022b).

Para avaliar habilidades dentro de um simulador computacional, um dos métodos que podem ser empregados é o sistema baseado em regras (SBR). Em simuladores, as regras são representações formais que descrevem uma relação de causa e efeito, frequentemente utilizando uma estrutura lógica que conecta condições a ações. O SBR é uma forma simples de inteligência artificial projetada para simular a capacidade de decisão de um especialista humano. Neste tipo

de sistema, o conhecimento obtido é armazenado em uma base de dados utilizando uma representação lógica, as regras. O gerenciamento das regras é realizado por um motor de inferência que emprega fatos obtidos desde o início ou após a dedução. A base de conhecimento e o mecanismo de regras são capazes de resolver problemas complexos, fornecer recomendações ou tomar decisões de maneira semelhante ao especialista humano (CASTILLO; GUTIÉRREZ; HADI, 1997). A Figura 1 ilustra o funcionamento um SBR.

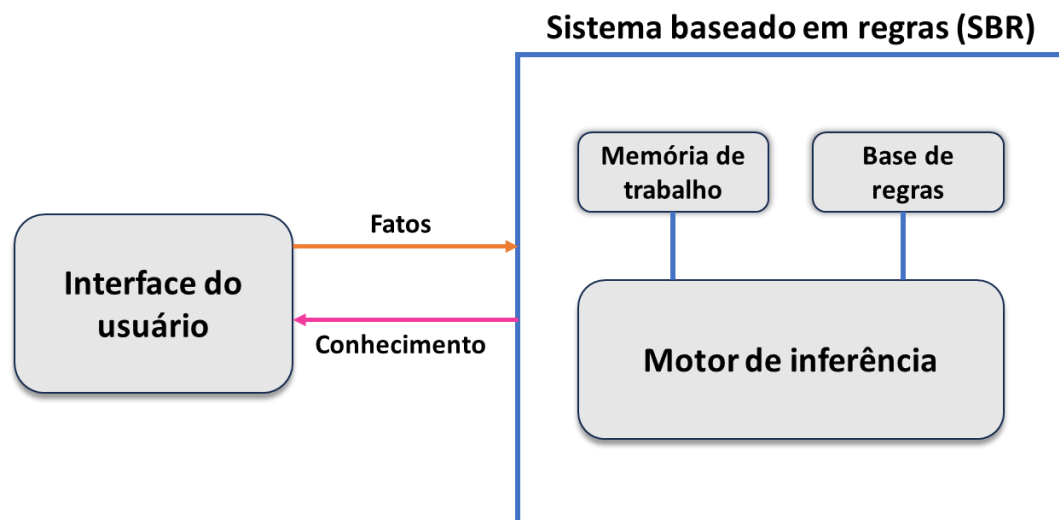


Figura 1: Funcionamento de um sistema baseado em regras.

Fonte: Adaptado de Durkin, 1990.

A construção das regras dentro de um SBR é dada no formato *se-então* (condição-ação), que são utilizadas no intuito de levar a conclusões. A base de regras é responsável por armazenar o conhecimento, ou seja, o conjunto de regras *se-então* definidas por um ou mais especialistas. Já a memória de trabalho armazena o conhecimento concreto, representado pelo conhecimento que pode ser considerado fato antes do processo de inferência. De posse da base de regras e do conhecimento armazenado na memória de trabalho, o motor de inferência é capaz de inferir conclusões. Aplicações envolvendo SBR no campo da saúde normalmente são desenvolvidas nas áreas de diagnóstico e avaliação, que utilizam regras no estilo condição-ação (COSTA, 2017; COSTA; MACHADO; MORAES, 2014).

A lógica clássica é amplamente utilizada em SBRs devido à sua simplicidade e formalismo bem estabelecidos, que facilitam a construção, manutenção e interpretação das regras. Além disso, sua natureza binária (verdadeiro-falso) garante previsibilidade e precisão no processamento das condições e ações. Apesar disso, outras formas de lógica podem ser

utilizadas, como por exemplo a lógica *fuzzy*, que pode ser empregada para lidar com incertezas ou informações imprecisas, permitindo que o sistema tome decisões em casos em que os dados não sejam imprecisos ou não estejam claramente definidos. A escolha do tipo de lógica a ser utilizada dependerá das características do problema a ser resolvido e da necessidade de flexibilidade ou precisão do sistema.

Os sistemas baseados em regras são uma abordagem simples e de fácil aplicação para a resolução de problemas onde a lógica e as regras são bem compreendidas e estabelecidas. Além disso, destacam-se pela sua flexibilidade, considerando que as regras podem ser acrescidas, removidas e modificadas conforme necessidade, sem influenciar o mecanismo de inferência.

A qualidade de *software* é um aspecto importante a ser considerado nos simuladores computacionais, e deve priorizar três características: confiabilidade – funcionar e fornecer o necessário, no nível exigido; segurança – funcionar de forma segura, eliminando possíveis falhas; e eficiência – funcionar da melhor maneira possível, com requisitos mínimos de recursos, como por exemplo, uso da memória (LIGÊZA, 2006).

Oliveira (2020) desenvolveu um simulador de realidade virtual para treinar profissionais de saúde no procedimento de reanimação neonatal. O simulador REANIME foi construído com um método de avaliação do usuário que utilizou, dentre os vários modelos de decisão, um sistema baseado em regras *fuzzy*. Os resultados preliminares foram satisfatórios na utilização das regras para a capacitação dos profissionais de saúde, o que tornou o treinamento através do REANIME confiável para o objetivo almejado.

No campo da Fonoaudiologia, o simulador de treinamento “Ouvindo Vozes”, para capacitar estudantes e profissionais na avaliação perceptivo-auditiva da voz, foi proposto em 2023. Nesse simulador, um SBR baseado na lógica clássica foi utilizado para avaliar o usuário durante o treinamento considerando tarefas de diversos níveis de dificuldade, tendo como base a comparação das respostas do usuário com as respostas dos especialistas (PAIVA; MACHADO; LOPES, 2023). O modelo aplicado ao simulador “Ouvindo vozes” cumpriu o objetivo de verificar e direcionar o progresso do usuário dentro da ferramenta (PAIVA, 2022).

Avaliar habilidades em simuladores computacionais significa avaliar o desempenho do usuário para identificar se eles estão preparados para realizar o procedimento em questão em situações da vida real. A habilidade perceptivo-visual para a classificação de variáveis envolvendo um exame de imagem, como no caso da VED, considera habilidades não técnicas. Segundo Huang et al. (2020), habilidades não técnicas em simuladores possuem relação com aspectos como a investigação diagnóstica e a tomada de decisão clínica.

A classificação dos parâmetros da deglutição, como os resíduos faríngeos, requer habilidades de percepção visual e cognitiva e tem como característica a subjetividade, pois depende da inspeção e interpretação da imagem pelo avaliador (como em outros exames de imagem na área da saúde) (KRUPINSKI, 2010). Estas habilidades podem ser aprimoradas com a experiência e variam de acordo com a velocidade do pensamento e a capacidade de reconhecer padrões e codificar, reter e recuperar informações (ABERNETHY et al., 1994).

Reconhecendo esta questão, um estudo sobre treinamento em exames de VFD considerou a influência da habilidade perceptivo-visual no treinamento (EDWARDS et al., 2023). Embora a percepção visual não fosse o foco principal, os autores destacaram a necessidade de fonoaudiólogos identificarem visualmente padrões e alterações relacionadas à deglutição no exame.

Em conclusão, a aplicação de modelos de decisão inteligentes, como os Sistemas Baseados em Regras (SBR), apresenta-se como uma abordagem eficaz para a avaliação de habilidades em simuladores computacionais na área da saúde, podendo ser uma alternativa para aprimorar a habilidade perceptivo-visual de parâmetros da deglutição, como os resíduos faríngeos visualizados na VED. Este modelo inteligente permite a incorporação de conhecimento com base na literatura científica e no conhecimento de especialistas, minimizando a subjetividade no processo de avaliação. A utilização de um SBR não só facilita tomadas de decisões precisas, mas também oferece flexibilidade para adaptação e melhoria contínua do sistema.

2.4. CONSIDERAÇÕES

Nesta seção do trabalho foram apresentados os principais estudos e conceitos relacionados ao treinamento para a classificação dos resíduos faríngeos na VED, com foco no processo de avaliação do treinamento. Os resultados encontrados demonstram a variabilidade existente na estrutura dos treinamentos encontrados na literatura, assim como na utilização de medidas de resultado para verificar a eficácia dos treinamentos. Apesar disso, a integração das tecnologias digitais no treinamento de profissionais da saúde no campo da deglutição, considerando avaliações instrumentais como a VED é uma realidade. Estudos neste sentido mostram os benefícios como flexibilidade, acessibilidade e padronização do conteúdo, além de destacarem a importância destas ferramentas para complementar ou introduzir o treinamento prático e melhorar a qualidade do atendimento aos pacientes com disfagia.

O ensino na Saúde está mudando globalmente, substituindo métodos tradicionais por práticas interativas e tecnologias digitais, como simuladores computacionais. Esses simuladores permitem que estudantes e profissionais pratiquem em ambientes virtuais seguros, desenvolvendo habilidades cognitivas. Além disso, proporcionam *feedback* imediato, reduzem riscos e custos do treinamento tradicional, e melhoram a confiança e habilidade dos participantes. Nenhum estudo envolvendo o desenvolvimento de simuladores computacionais para o treinamento profissional na classificação de parâmetros da VED foi encontrado.

Em relação à avaliação de habilidades, a avaliação *online* com base em dados de especialistas pode complementar ou introduzir o treinamento prático na classificação dos resíduos faríngeos na VED. Esta avaliação, principalmente quando realizada em associação com um modelo de decisão inteligente baseado em regras, fornece *feedback* em tempo real sobre a aquisição de habilidades do usuário, que pode ser aprimorada ao longo do tempo de uso da ferramenta. Apesar de ainda não existirem propostas de treinamento específicos em simuladores computacionais para desenvolver a habilidade perceptivo-visual de classificação de parâmetros da deglutição na VED, tecnologias digitais para este objetivo já foram desenvolvidas. Estes recursos demonstram a importância do treinamento para aprimorar esta habilidade, que pode ser potencializada com a experiência e pela capacidade de identificar padrões, além de codificar, armazenar e recuperar informações.

Quanto à avaliação de habilidades, os sistemas baseados em regras (SBR) são flexíveis e permitem a adição ou modificação de regras conforme necessário, e por este motivo são bastante aplicados em casos de treinamento de diagnósticos e avaliações em saúde. A qualidade do *software* em termos de usabilidade também é crucial para garantir que estas ferramentas atinjam os objetivos esperados.

Nenhum método específico para treinar estudantes e profissionais de saúde para classificarem os parâmetros da deglutição foi encontrado em um simulador computacional. A importância desta classificação no diagnóstico em disfagia orofaríngea, os riscos de recomendações inadequadas associadas a este diagnóstico e a falta de trabalhos nesta direção motivam a criação de um sistema de avaliação para esta tarefa.

3. MÉTODO

3.1. DESCRIÇÃO DO PROCESSO METODOLÓGICO

Para alcançar os objetivos propostos para este trabalho, algumas etapas foram executadas. Didaticamente, estas etapas foram divididas em: 1) Revisão da literatura; 2) Definição de parâmetros para a avaliação de um treinamento para a classificação dos resíduos faríngeos na VED; 3) Proposição de um modelo inteligente; 4) Desenvolvimento do ST DeglutSIM 4) Validação da usabilidade do ST. Estas etapas foram consideradas para chegar ao objetivo do trabalho, que é de propor e desenvolver um simulador computacional de treinamento perceptivo-visual que seja capaz de avaliar a habilidade do usuário em realizar a classificação do parâmetro resíduos faríngeos na VED, e serão descritas com mais detalhes nos tópicos seguintes.

3.2. REVISÃO DA LITERATURA

Para responder ao primeiro objetivo específico deste trabalho, foram utilizadas informações revisadas na literatura. Para identificar as características do treinamento, foi realizada uma revisão de escopo que buscou compreender como este é realizado, considerando a estrutura do treinamento fornecido, quantidade de exames utilizados, medidas de resultados consideradas, entre outros aspectos importantes para a elaboração de um treinamento que possa tornar o avaliador competente para classificar os parâmetros da VED.

Uma busca abrangente foi realizada no intuito de determinar o escopo da literatura que contém o treinamento da habilidade perceptivo-visual para analisar os parâmetros funcionais da VED, fornecendo uma visão geral do tema e indicando lacunas existentes.

As bases de dados eletrônicas pesquisadas foram a MEDLINE (PubMed), Cochrane Library, Embase (Elsevier), Web of Science (Clarivate), Scopus (Elsevier) e CINAHL Full Text (EBSCO). Estudos não publicados e literatura cinzenta também foram considerados no Google Scholar, ProQuest e MedNar. Todas as buscas foram feitas no modo avançado, sem restrição de idioma ou ano. As palavras-chave utilizadas na busca combinou grupos de palavras relacionados a “disfagia”, “VED” e “treinamento”, adaptadas para cada base de dados eletrônica. A busca foi concluída em 16 de dezembro de 2021.

3.3. DEFINIÇÃO DE PARÂMETROS PARA A AVALIAÇÃO DE UM TREINAMENTO PARA A CLASSIFICAÇÃO DOS RESÍDUOS FARÍNGEOS NA VED

3.3.1. Extração dos requisitos, variáveis e valores para o sistema de avaliação

Nesta etapa do estudo, buscou-se obter requisitos, variáveis e valores para o sistema de avaliação de um simulador de treinamento para a classificação da gravidade dos resíduos faríngeos na avaliação de deglutição por VED. A extração dos requisitos foi feita a partir de consulta à literatura, especificada no subtópico anterior, somada a uma entrevista a uma fonoaudióloga especialista com experiência em realização da VED. Para a extração das variáveis e valores, foi vista a necessidade do desenvolvimento de um banco de dados específico para a classificação dos resíduos faríngeos na VED, que foi denominado PhaReD-FEES (do inglês, *Pharyngeal Residue Database in Fiberoptic Endoscopic Evaluation of Swallowing*), estruturado em quatro etapas (Figura 2):

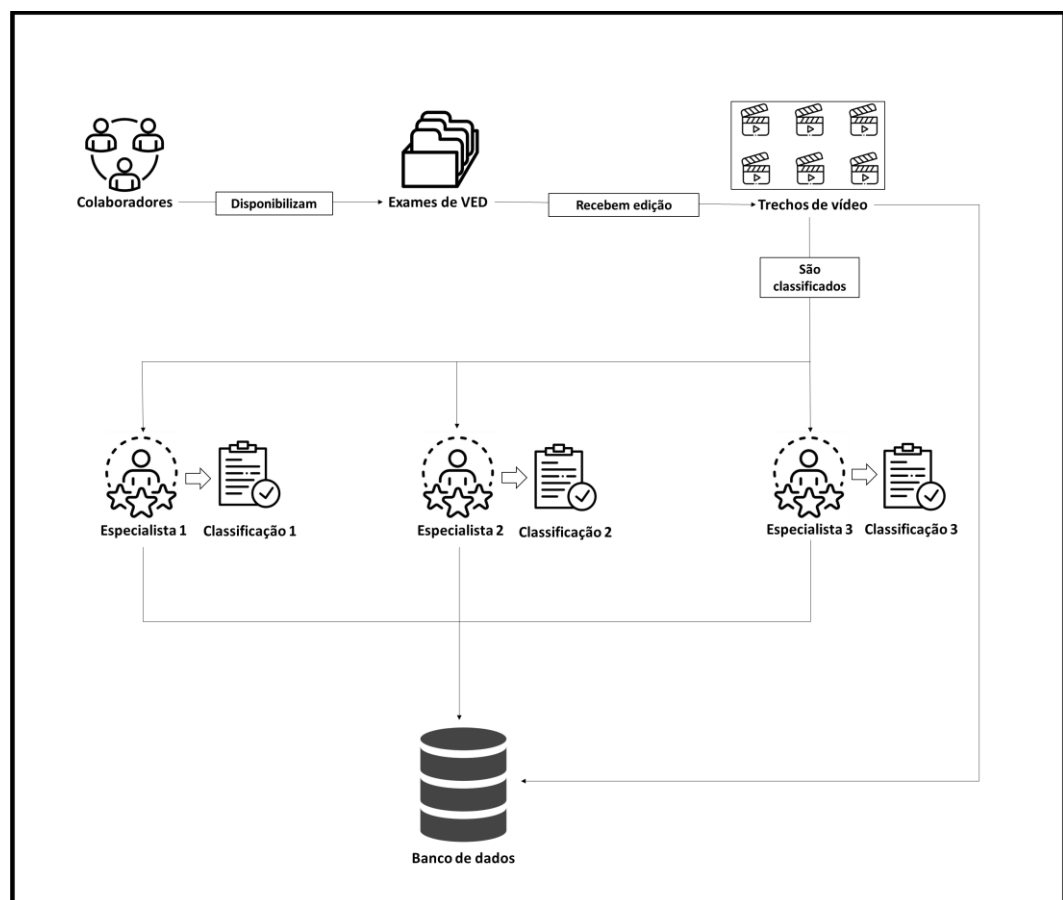


Figura 2: Fluxograma do processo de desenvolvimento do banco de dados PhaReD-FEES

1. Coleta e Edição de vídeos

- Foram selecionados aleatoriamente 52 exames de VED provenientes de uma base particular com exames de VED de indivíduos jovens e idosos, com ou sem doença neurológica, pertencente ao Laboratório de Disfagia da Universidade Estadual Paulista – Campus de Marília/SP (LADis/UNESP). Estes vídeos abrangiam uma variedade de resoluções entre 720x406 e 320x180 pixels.
- Esses vídeos passaram por uma avaliação independente por três fonoaudiólogos especialistas, que editaram os vídeos para manter apenas os momentos imediatamente antes, durante e após a deglutição.
- Os trechos de vídeo recortados continham até 11 segundos, com base na visualização do *whiteout* (bloqueio temporário à passagem da luz no aparelho de VED que impede a visualização direta dos eventos da deglutição), utilizado como referência para identificar o momento da formação dos resíduos, que ocorre imediatamente após este evento. O objetivo foi isolar os momentos do exame em que os resíduos faríngeos após a deglutição ficaram visíveis nas regiões de valéculas e seios piriformes.
- Vários trechos de vídeos poderiam ser extraídos de um exame realizado em um único paciente, visto que durante um exame são ofertadas diferentes consistências alimentares e pode ser visualizada mais de uma deglutição em cada oferta.
- Os trechos de vídeo foram extraídos sem alterações em sua resolução ou contraste, utilizando um *software* online de edição (<https://online-video-cutter.com/>). Apenas trechos que apresentavam boa qualidade de visualização das estruturas anatômicas de interesse foram mantidos.

2. Seleção de especialistas

- Foram selecionados três especialistas que cumpriram os critérios de Fehring (FEHRING, 1987) adaptados para a área de Disfagia. Esses critérios contemplam tanto a titulação acadêmica quanto a experiência clínica.
- O critério definido para participação neste estudo foi: Possuir no mínimo 10 pontos, de um total de 12 pontos (Quadro 2).

Quadro 2. Critérios sugeridos por Fehring (1987) com adaptação para a área de Disfagia.

Critério de Fehring (adaptado)	Especialista 1	Especialista 2	Especialista 3
Ter titulação de mestre (4 pontos)	x	x	x
Ser mestre, com dissertação envolvendo diagnóstico em Disfagia (4 pontos)	x	x	x
Ter experiência (pelo menos 1 ano) com diagnósticos em Disfagia na prática clínica, ensino ou pesquisa (2 pontos)	x	x	x
Ter ou estar cursando especialização em Disfagia (2 pontos)	x	x	
Total	12	12	10

Fonte: Fehring (1987) adaptado.

- Os especialistas passaram por treinamento para padronizar a classificação utilizando a *Yale Pharyngeal Residue Severity Rating Scale* (YPRSRS) versão traduzida e adaptada para o português brasileiro, uma escala validada para classificar a gravidade dos resíduos nas regiões de valéculas e seios piriformes, em cinco níveis: nenhum, vestígio, leve, moderado e grave (NEUBAUER; RADEMAKER; LEDER, 2015; VENITE et al., 2024). O treinamento buscou minimizar variações interpessoais na classificação.

3. Classificação dos resíduos faríngeos

- A YPRSRS foi utilizada para a classificação dos resíduos faríngeos nas regiões de valéculas e seios piriformes, separadamente, pelos três especialistas.
- Cada trecho de vídeo recebeu uma classificação de 1 a 5, referente aos cinco níveis da escala. As classificações que não foram possíveis de serem feitas, foram classificados como “0 – Não é possível avaliar”.

- Divergências significativas entre os especialistas foram discutidas em reuniões de consenso. Para divergências menores (± 1 ponto na escala), ambas as classificações foram mantidas.
- Ao final, 92 trechos de vídeos foram incluídos no banco de vídeos, recebendo pelo menos uma classificação para valéculas e para seios piriformes (Quadro 3).

Quadro 3. Quantidade de trechos de vídeos separadas pela classificação dos resíduos faríngeos nas regiões de valéculas e seios piriformes.

	Quantidade de trechos de vídeos por classificação					
	Nenhum	Vestígio	Leve	Moderado	Grave	0
Valéculas	13	20	20	14	20	22
Seios Piriformes	23	27	17	14	16	18

4. Estrutura e disponibilização do banco de dados

- O banco de dados final é composto por 92 trechos de vídeos recortados e classificados, acompanhados de documentação com descrição das classificações.
- O banco de dados será disponibilizado publicamente em repositório de acesso aberto, permitindo acesso a pesquisadores, clínicos e outros interessados na área de Disfagia. Uma publicação neste sentido será encaminhada para um periódico indexado.

De 118 trechos inicialmente selecionados, 26 foram excluídos devido a um consenso insuficiente entre os especialistas ou problemas técnicos que inviabilizaram a análise. O banco final apresenta uma distribuição equilibrada entre as categorias de gravidade, o que o torna uma ferramenta valiosa para estudos futuros, incluindo o treinamento de algoritmos de inteligência artificial.

O PhaReD-FEES foi o banco de dados desenvolvido para este estudo. Através deste banco, foi possível definir parâmetros (variáveis e valores) para a avaliação de um treinamento para a classificação de resíduos faríngeos na VED em uma simulação computacional. Ademais, o produto que será gerado com esse procedimento metodológico preenche uma lacuna

significativa no acesso a dados padronizados de vídeos de VED. Além de fomentar colaborações interdisciplinares e permitir análises replicáveis, o banco de dados é expansível, permitindo a inclusão de novos dados. Essa flexibilidade contribui diretamente para avanços científicos e melhorias na assistência ao paciente com disfagia.

3.3.2. Estrutura do treinamento e métricas para a avaliação do usuário

A partir do conhecimento disponível na literatura sobre treinamento na classificação dos parâmetros da VED e do conhecimento de especialistas nesta temática, foi definida a estruturação do ST e estabelecidas as métricas para a avaliação do usuário.

Constituído o banco de dados da pesquisa, foram selecionados recortes de vídeos de VED para o parâmetro da deglutição resíduos faríngeos (em valéculas e em seios piriformes) que serviram como base para o treinamento: Recortes de vídeos com a classificação da gravidade do resíduo unânime entre os especialistas e que eram mais representativos de cada nível (etapa 01 – ensino); recortes de vídeos de exames com classificações unânimes entre os especialistas e que fossem dos níveis extremos – nível I (“nenhum”) e do nível V (“grave”) –, que possuem menor dificuldade para serem identificados visualmente (etapa 02 – nivelamento); e recortes de vídeos de exames de todos os níveis, tanto para valécula quanto para seios piriformes, para serem apresentados junto com os exames da etapa 02, repetidos para a análise de confiabilidade intra-avaliador dentro do sistema (etapa 03 – prática).

Duas medidas foram utilizadas para mensurar a habilidade perceptivo-visual do usuário no ST, são elas: a confiabilidade intra-avaliador (entre o usuário e ele mesmo, em momentos distintos) e a acurácia (precisão das classificações do usuário em comparação às dos especialistas). Essas medidas compuseram o índice de desempenho utilizado para verificar a aquisição da habilidade do usuário. Considerando que o objetivo do ST é desenvolver habilidades perceptivo-visuais para classificar parâmetros da deglutição, essas medidas foram definidas para classificar o desempenho do usuário quanto à precisão e à consistência nas classificações dos exames. A melhora no desempenho das habilidades possibilitará uma avaliação instrumental da deglutição mais padronizada e precisa, o que refletirá em melhores condições de diagnóstico em disfagia orofaríngea e melhor definição das estratégias de reabilitação da função de deglutição.

A confiabilidade intra-avaliador refere-se à consistência das medidas avaliadas nas mesmas condições em dois momentos diferentes, pelo mesmo avaliador. Para analisar esta medida, será adotado o coeficiente Kappa. O coeficiente Kappa é utilizado comumente para

medir o desempenho, considerando o número de casos corretamente avaliados. Este coeficiente assume valores entre 0 e 1, sendo que quanto mais próximo de 1, mais precisa será a avaliação. A escolha pelo Kappa deve-se ao fato deste coeficiente possuir mais rigor em relação às outras medidas de concordância. Destaca-se, ainda, que a literatura considera o Kappa como método mais utilizado quando se trata de variáveis nominais ou ordinais (FONSECA; SILVA; SILVA, 2007). A fórmula geral do coeficiente kappa é:

$$\kappa = \frac{P_0 - P_e}{1 - P_e}$$

Onde:

- P_0 é a proporção de concordância observada entre as duas avaliações, em momentos distintos;
- P_e é a proporção de concordância esperada ao acaso.

A acurácia ou precisão é uma medida de desempenho utilizada comumente para verificar a eficácia de modelos de classificação, representando a proporção de previsões corretas realizadas pelo modelo em relação ao total de observações. No presente estudo, esta medida está relacionada à concordância entre a avaliação do usuário e a avaliação dos especialistas. Sendo assim, a acurácia será definida como a relação entre o número de exames corretamente classificados e o número total de exames avaliados. A fórmula para calcular a acurácia é dada por:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN}$$

Onde:

- VP = verdadeiro positivo;
- VN = verdadeiro negativo;
- FP = falso positivo;
- FN = falso negativo.

As medidas de confiabilidade intra-avaliador e precisão foram adotadas para compor o índice de desempenho do usuário no ST. A classificação da confiabilidade intra-avaliador

através dos valores do coeficiente Kappa seguirá a proposta de Landis e Koch (1977), e para os valores de acurácia, será utilizada a proposta de Hosmer e Lemershow (2000) (Quadro 4).

Quadro 4: Classificação para os valores do coeficiente Kappa (LANDIS; KOCH, 1977) e acurácia (HOSMER; LEMERSHOW, 2000).

Medida	Valores	Classificação
Coeficiente Kappa	0,81 – 1,00	Perfeita ou quase perfeita
	0,61 – 0,80	Boa
	0,41 – 0,60	Moderada*
	0,21 – 0,40	Regular
	0,00 – 0,20	Discreta
	< 0,00	Pobre
Acurácia	> 90%	Excelente
	80 – 90%	Boa
	70 – 80%	Aceitável*
	60 – 70%	Pobre

Legenda: *Classificação aceitável para a medida.

Após consultar a literatura sobre o tema e um especialista para levantamento dos requisitos, foi constatada a necessidade de determinação de um índice de desempenho para a avaliação do usuário na classificação dos resíduos faríngeos na VED envolvendo as medidas de: Confiabilidade intra-avaliador, com base no coeficiente Kappa; e Acurácia ou precisão nas classificações.

3.4. PROPOSIÇÃO DE UM MODELO DE DECISÃO INTELIGENTE

Durante a fase inicial deste estudo, foram realizados levantamentos teóricos na literatura e consultas com uma especialista com experiência na realização da VED, para a definição dos requisitos (do especialista e computacionais) do simulador. Foram identificados dois aspectos a serem priorizados no sistema: a avaliação da habilidade perceptivo-visual e o fornecimento de *feedback* imediato ao usuário, ambos com potencial impacto no aprendizado. Detalhes sobre o levantamento dos requisitos podem ser encontrados na seção Resultados e Discussão.

Além disso, foram analisadas as classificações (variáveis de avaliação) dos níveis de gravidade dos resíduos faríngeos de acordo com a *Yale Pharyngeal Residue Severity Rating Scale* (YPRSRS): I – Nenhum, II – Vestígio, III – Leve, IV – Moderado e V – Grave; bem como os valores atribuídos pelos especialistas a estas variáveis. Estas análises permitiram a

construção de regras para um sistema de avaliação e a definição de um modelo inteligente capaz de avaliar a precisão e a confiabilidade intra-avaliador dos participantes.

As variáveis de avaliação consideradas na escala YPRSRS são do tipo categóricas e consideram um conjunto de conceitos descritivos e visuais. Pela diferença visual sutil entre as classificações em muitos casos se torna difícil distinguir visualmente os níveis da escala, e os juízes costumam discordar na faixa de 1 nível para mais ou para menos. Este fenômeno costuma ocorrer frequentemente na prática clínica e foi previsto pelos pesquisadores durante o delineamento do estudo. Ao receber as classificações realizadas pelos três especialistas, o fenômeno foi constatado.

Visando solucionar este problema, a base de conhecimentos deste trabalho utilizou uma regra composta por uma disjunção que considera uma ou duas condições com apenas uma conclusão:

$$\begin{aligned} & \text{SE } (<\text{condição1}>) \text{ ENTÃO } <\text{conclusão1}> \\ & \text{SE } (<\text{condição1}> \text{ OU } <\text{condição2}>) \text{ ENTÃO } <\text{conclusão1}>. \end{aligned}$$

A disjunção é representada pelo operador “OU”, e indica que para a conclusão ser verdadeira, basta que uma das condições seja verdadeira. Deste modo, a conclusão será ativada se condição1 para verdadeira ou se condição2 para verdadeira, ou ainda se ambas forem verdadeiras. A disjunção é descrita pela seguinte fórmula lógica:

$$P \vee P_q \rightarrow R$$

Onde:

P e **P_q** são as condições;

V é o operador lógico “OU” (disjunção);

R é a conclusão que ocorre se pelo menos uma das condições for verdadeira.

Em termos práticos, a aplicação da lógica para este trabalho foi a seguinte:

$$\begin{aligned} & \text{SE } (<\text{Nível II}>) \text{ ENTÃO } <\text{correto}> \text{ SENÃO } <\text{incorreto}> \\ & \text{SE } (<\text{Nível II}> \text{ OU } <\text{Nível III}>) \text{ ENTÃO } <\text{correto}> \text{ SENÃO } <\text{incorreto}>. \end{aligned}$$

Isso quer dizer que, caso um especialista tenha atribuído a classificação “II - leve” para um exame e os outros dois especialistas atribuam a classificação “III - moderado”, ou vice-versa, ambos os valores (I e II) foram considerados corretos. Nos casos em que os três

especialistas atribuíram “II-leve” para um exame, apenas este valor foi considerado correto dentro do sistema de avaliação. Nos casos em que cada especialista atribuiu um nível diferente, o recorte de vídeo foi desconsiderado para a avaliação dentro do simulador.

A análise das variáveis de avaliação classificadas pelos especialistas nas etapas anteriores revelou a ausência de características de subjetividade e imprecisão, enquadrando-as no domínio da lógica clássica. Foi desenvolvido um Sistema Baseado em Regras (SBR), com o objetivo de ser aplicado ao treinamento da classificação dos resíduos faríngeos na VED. Este tipo de sistema é uma forma de inteligência artificial que utiliza regras para representar o conhecimento de forma codificada dentro do sistema. O modelo proposto atuará no que diz respeito à avaliação do desempenho do usuário, contabilizando os acertos ao comparar as respostas com as dos especialistas (precisão) e as respostas do próprio usuário em fases distintas (confiabilidade intra-avaliador). Foi definido que a combinação destes dois aspectos geraria um índice com valores pré-determinados, que é retornado para o usuário em forma de estrelas (*feedback*). A Figura 3 ilustra o processo de escolha do modelo de decisão inteligente.

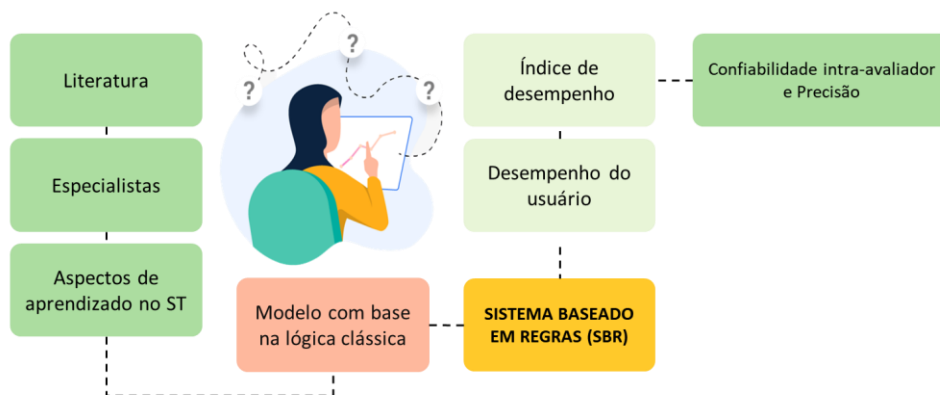


Figura 3: Processo de escolha do modelo de decisão inteligente para o simulador

O modelo de decisão inteligente é composto por três componentes: 1) Confiabilidade intra-avaliador; 2) Precisão das classificações e 3) Pontuação final. A Figura 4 apresenta uma visão geral do *feedback* avaliativo retornado para o usuário ao final da simulação.

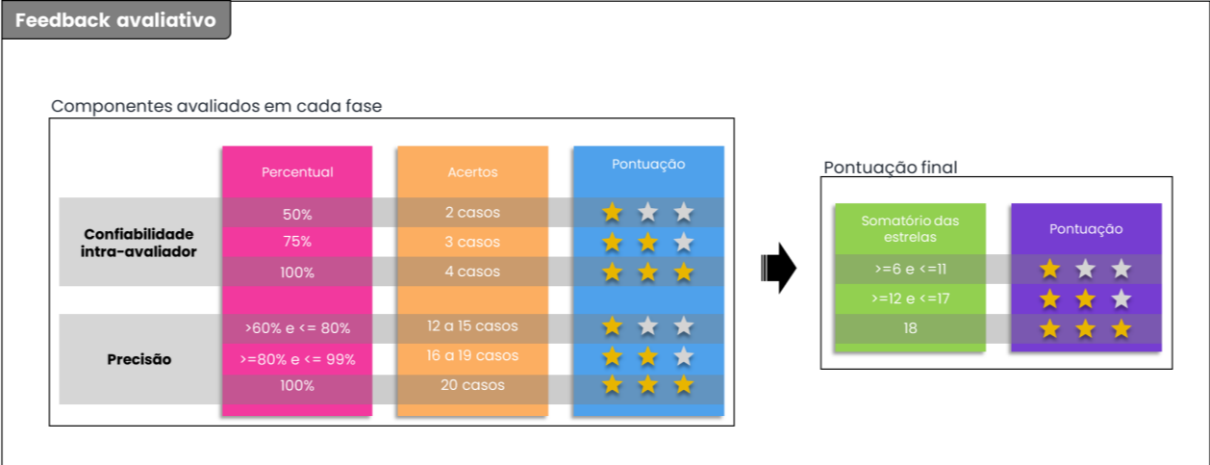


Figura 4: Visão geral do índice considerado na avaliação do desempenho do usuário no modelo de decisão.

Fonte: Elaborado pelo autor, 2024.

3.4.1 Confiabilidade intra-avaliador

O componente confiabilidade intra-avaliador está representado dentro do sistema como a quantidade de acertos dos quatro itens que foram apresentados ao usuário na etapa Nivelamento e que se repetem na etapa de avaliação propriamente dita, a etapa Prática. O usuário tem a possibilidade de acertar entre 0 e 4 casos neste componente. Caso acerte a classificação de nenhum ou um dos quatro casos, não ganhará nenhuma estrela. Caso acerte dois (50%), ganhará uma estrela; três (75%), duas estrelas; e quatro (100%), receberá o máximo, correspondente a três estrelas.

A quantidade de estrelas para a confiabilidade intra-avaliador foi determinada pelos pesquisadores baseando-se nos valores de classificação do coeficiente Kappa estabelecidos por Landis e Koch (1977). A Tabela 1 descreve as comparações e equivalências:

Tabela 1: Comparações entre os acertos e a correspondência com o coeficiente Kappa.

Confiabilidade intra-avaliador			
Acertos	Percentual	Classificação Kappa	Estrelas
2	50%	Moderada (0,41 – 0,60)	*
3	75%	Boa (0,61 – 0,80)	**
4	100%	Perfeita ou Quase perfeita (0,81 – 1,00)	***

3.4.2 Precisão

O componente precisão verifica a acurácia do participante na etapa Prática à medida que compara a sua resposta com a(as) resposta(s) dos especialistas. O usuário tem a possibilidade de acertar entre 0 e 20 casos neste componente, que é independente e não possui relação com nenhuma etapa anterior do sistema. Caso acerte menos do que 12 casos, o usuário não pontuará nenhuma estrela neste componente.

A quantidade de estrelas para a precisão foi determinada pelos pesquisadores baseando-se nos valores determinados para a medida de acurácia propostos por Hosmer e Lemeshow (HOSMER; LEMESHOW, 2000). A Tabela 2 descreve estas comparações:

Tabela 2: Comparações entre os acertos e a correspondência com a Acurácia.

Precisão das classificações			
Acertos	Percentual	Classificação Acurácia	Estrelas
12 a 15	>60% e <80%	Pobre (60 – 70%) a Aceitável (70 – 80%)	*
16 a 19	>=80% e <=99%	Boa (80 – 90%) a Excelente (90%)	**
20 casos	100%	Excelente (>90%)	***

3.4.3 Pontuação final

A pontuação final é um índice calculado pela soma das estrelas obtidas nas fases RFV, RFS e RFVS. As estrelas que podem ser obtidas na pontuação final (*feedback*) e o intervalo correspondente estão detalhados na Figura 22.

De maneira geral, a pontuação final reflete as medidas de confiabilidade intra-avaliador e precisão que juntas compõem o índice de desempenho utilizado para avaliar o usuário. Com base no somatório das estrelas obtidas em ambas as medidas, em cada uma das três fases, é possível determinar o desempenho geral do usuário na simulação. A Tabela 3 apresenta uma descrição da pontuação final no simulador.

Tabela 3: Descrição da pontuação final dentro do simulador.

Pontuação final			
Soma de estrelas	Nº de estrelas	Classificação do usuário	<i>Feedback</i>
<6	0	-	“Seu desempenho pode ser melhorado. Recomendamos que faça o treinamento novamente.”
>=6 e <=11	1	Ruim	“Seu desempenho pode ser melhorado. Recomendamos que faça o treinamento novamente.”
>=12 e <=17	2	Bom	“Seu desempenho foi satisfatório, mas ainda há espaço para aprimoramento. Caso julgue necessário, pode repetir o treinamento.”
18	3	Excelente	“Parabéns! Seu desempenho no treinamento foi excelente.”

3.5. DESENVOLVIMENTO DO ST

3.5.1. Definição do conteúdo

Para a elaboração de um ST com o propósito de desenvolver habilidades específicas, alguns elementos devem ser definidos durante o seu processo de desenvolvimento, como: objetivo, assunto principal, público-alvo, cenário e graduação dos desafios. Estes elementos levam a refletir sobre a incorporação do conteúdo da aplicação, mantendo o foco no interesse do usuário e ressaltando a importância do tema no processo educacional (MACHADO; COSTA; MORAES, 2018). Para este estudo, foram estabelecidos:

- 1) Objetivo: Desenvolver a habilidade perceptivo-visual para profissionais e estudantes da área de saúde classificarem os resíduos faríngeos na VED;
- 2) Assunto principal: Treinamento perceptivo-visual na classificação dos resíduos faríngeos na VED, com acompanhamento do desempenho em tempo real;

- 3) Público-Alvo: Profissionais e estudantes de pós-graduação da área de saúde que desejam atingir competência na classificação dos resíduos faríngeos na VED;
- 4) Cenário: Interfaces práticas, interativas, de fácil usabilidade;
- 5) Graduação de desafios: Possui uma fase de nivelamento, que apresenta casos fáceis de serem avaliados, pré-requisito para a fase seguinte, que abarca todos os tipos de casos (fáceis e difíceis de serem avaliados).

3.5.2. Construção das telas

Os protótipos de tela da aplicação foram construídos para integrar a interface do simulador. De acordo com o objetivo da aplicação, a interface deveria permitir que os usuários aprendessem através da observação de imagens, vídeos representativos e definições percentuais e operacionais de cada nível da escala YPRSRS. O esboço de uma tela construída inicialmente para compor a etapa Ensino está ilustrada na Figura 5. A tela apresenta as imagens de todos os níveis funcionando como botões, onde clicando no botão disponível o usuário seria direcionado para outra tela contendo os itens a serem observados.



Figura 5: Protótipo de tela do simulador.

Fonte: Elaborado pelo autor, 2023.

3.5.3. Implementação do simulador e incorporação ao modelo de decisão

O simulador foi desenvolvido utilizando o motor de jogo Unity 3D e a linguagem de programação C#. O Unity é uma plataforma voltada para a criação de jogos e aplicações interativas em 2D e 3D. A escolha deste motor de jogo se deu pela gratuidade da ferramenta e pela existência de uma comunidade de desenvolvedores ampla e ativa. Dada a natureza dos requisitos do sistema, optou-se pela utilização do modelo 2D para a simulação.

Posteriormente, as especificações obtidas nas fases anteriores foram traduzidas em regras conforme o modelo de avaliação inteligente SBR. Essas regras foram incorporadas ao código do simulador para permitir a avaliação da capacidade perceptivo-visual do usuário na classificação dos resíduos faríngeos na VED. O simulador está disponível para execução em plataformas *desktop*, requerendo um computador com acesso à internet e sistema operacional *Windows 10* ou versões mais recentes. O programa foi desenvolvido para atender às necessidades específicas deste trabalho, sendo registrado junto ao Instituto Nacional da Propriedade Industrial (INPI), sob o número BR512024004449-3 (Anexo 1).

3.6. VALIDAÇÃO DA USABILIDADE DO ST

Esta foi uma pesquisa exploratória e de abordagem quantitativa que buscou mensurar o grau de satisfação de profissionais de saúde quanto à usabilidade do simulador computacional para treinamento perceptivo-visual da classificação dos resíduos faríngeos na videoendoscopia da deglutição (VED).

No contexto computacional, usabilidade refere-se à facilidade com que os usuários podem interagir com um sistema, levando em conta aspectos como facilidade de aprendizagem, facilidade de memorização, satisfação do usuário e eficiência do sistema (BROOKE, 1996). A avaliação da usabilidade em um simulador de treinamento computacional é crucial para garantir que os usuários possam aprender de forma eficaz. Uma pesquisa sobre a usabilidade do sistema através de um questionário foi realizada por meio de um experimento controlado em que os participantes interagiram com o simulador computacional de forma individual.

3.6.1. Cenário, população e amostra

Foram selecionados por conveniência indivíduos com o perfil do público-alvo do estudo, que eram profissionais de saúde envolvidos no tratamento de pacientes com disfagia orofaríngea: fonoaudiólogos com pós-graduação (em curso ou finalizada) ou que atuam clinicamente na área de Disfagia (subárea da Fonoaudiologia), e médicos com residência

médica em Otorrinolaringologia em curso ou finalizada. Quatorze pessoas com o perfil descrito acima foram contatadas para participar do experimento dentro da sua disponibilidade de tempo por duas semanas, sendo possível a participação de onze pessoas ao final.

Para a seleção dos indivíduos, foi realizado contato individual prévio para explicação dos objetivos, procedimentos da pesquisa e sobre a necessidade de assinatura de um Termo de Consentimento Livre e Esclarecido para a sua anuência de participação na pesquisa. O contato foi realizado via e-mail, ligação telefônica ou por meio de aplicativo de mensagens eletrônicas instantâneas (*Whatsapp*). Foram agendados encontros presenciais com os participantes que tinham disponibilidade nas dependências do Laboratório de Tecnologias para o Ensino Virtual e Estatística, da Universidade Federal da Paraíba (LabTEVE/UFPB).

O procedimento foi estruturado para avaliar a experiência do usuário, nas seguintes etapas, descritas a seguir:

1) Início do experimento

Os participantes foram recebidos no laboratório, onde foi feito um breve esclarecimento sobre os objetivos do estudo e como funcionaria a pesquisa. A pesquisadora forneceu informações gerais sobre o simulador e as tarefas que seriam realizadas. Em seguida, os participantes assinaram um Termo de consentimento livre e esclarecido (TCLE), garantindo que estavam cientes dos objetivos, procedimentos e riscos da pesquisa.

2) Interação com o simulador

Cada participante foi convidado a interagir com o simulador de forma individual. Durante o experimento, o simulador foi apresentado em um computador no ambiente controlado do laboratório. O usuário foi orientado a utilizar o simulador para aprender sobre a classificação dos resíduos faríngeos na VED utilizando a *Yale Pharyngeal Residue Severity Rating Scale* (YPRSRS), para posteriormente realizar a classificação individual dos exames apresentados. Os participantes ficaram livres com relação ao tempo de interação com o simulador. Porém, após 30 minutos, a pesquisadora apresentou duas opções para o participante: prosseguir até o final para visualizar o *feedback* que a ferramenta apresentaria sobre o seu desempenho; ou encerrar a interação e visualizar as possibilidades de *feedback* através de imagens de captura de tela do simulador, apresentados pela pesquisadora.

Durante o teste, a pesquisadora observou as reações dos participantes e registrou pontos importantes sobre suas interações, como dificuldades, comportamentos e sugestões espontâneas.

3) Levantamento

Após a interação com o simulador, os participantes foram convidados a responder a dois questionários, o que ajudou a mensurar a percepção sobre a usabilidade do sistema. A pesquisadora também coletou dados qualitativos adicionais: comentários, observações ou sugestões dos participantes sobre pontos positivos e negativos da experiência de uso do simulador.

3.6.2. Coleta de dados

Após participar da simulação, os usuários responderam questões relacionadas aos dados demográficos e formação profissional (Apêndice 1) e ao questionário SUS (do inglês, *System Usability Scale*) em sua versão traduzida para o Português brasileiro (Apêndice 2). O SUS é uma ferramenta desenvolvida em 1986 (BROOKE, 1996) e possui dez questões com objetivo de mensurar a usabilidade de diversos produtos e serviços, como websites, sistemas e aplicativos móveis.

As dez questões contidas no SUS são graduadas em escalas do tipo *Likert*, com valores de um a cinco, classificadas como: “discordo fortemente”, “discordo”, “não concordo nem discordo”, “concordo” e “concordo fortemente”. As questões 11, 12 e 13 do questionário foram incluídas pelos pesquisadores e eram dissertativas e não obrigatórias. Após o preenchimento, foi feito o cálculo da pontuação total, que gerou um número único.

Para calcular o *score*, primeiro foi somado o *score* de cada item da escala, de 1 a 5. Para os itens 1, 3, 5, 7 e 9, o *score* individual é composto pela nota recebida menos 1. Para os itens 2, 4, 6, 8 e 10, a contribuição é 5 menos a nota recebida. Multiplicou-se a soma de todos os *scores* por 2,5 para obter o valor total do SUS. O *score* SUS pode resultar em um número que varia de 0 a 100.

Conforme estudos envolvendo a escala SUS ao longo dos anos, a média adotada é de 68 pontos. Um resultado abaixo de 50 pontos é considerado inaceitável; abaixo de 20 pontos, pobre; abaixo de 12, pior imaginável. Em relação às pontuações acima da média estabelecida nos estudos, a partir de 71 é bom; a partir de 85 é excelente e a partir de 90, melhor imaginável (BANGOR; KORTUM; MILLER, 2009; BROOKE, 2013).

3.6.3. Análise dos dados

Os dados quantitativos foram analisados por meio do *software* estatístico *Statistical Package for the Social Sciences* (SPSS), versão 21.0 para Windows® e apresentados de forma descritiva (frequência simples e relativa, média, desvio padrão). Para os itens considerados da escala SUS, foram utilizadas média, mediana, desvio padrão e variância para os valores de cada item e *score* total da escala. O cálculo do coeficiente Kappa Ponderado foi utilizado para verificar a confiabilidade interavaliadores considerando os 10 itens da escala. A ponderação quadrática foi escolhida para penalizar discordâncias mais graves de forma acentuada, garantindo maior precisão na classificação da escala SUS.

O coeficiente Kappa ponderado é uma medida de concordância entre avaliados que considera o grau de discordância entre eles. A fórmula considera uma matriz de pesos, que ajusta o cálculo de acordo com a gravidade das discordâncias entre as categorias. O kappa ponderado (k_c) é calculado pela seguinte fórmula:

$$k_c = \frac{P_0 - P_e}{1 - P_e}$$

Onde:

- P_0 é a proporção de concordância ponderada observada, calculada pela soma dos produtos das frequências observadas (n_{ij}) e os pesos (w_{ij}) atribuídos às categorias;
- P_e é a proporção de concordância ponderada esperada ao acaso, baseada nas distribuições marginais dos avaliadores.

Os dados qualitativos foram tratados e analisados através de análise temática. As informações advindas dos “pontos positivos”, “pontos negativos” e “sugestões”, assim como as anotações realizadas pela pesquisadora durante a simulação foram transcritas, o conteúdo explorado e por fim, sintetizado e agrupado de acordo com a compreensão dos núcleos de sentido por meio de classes que possuem características comuns.

3.6.4. Aspectos Éticos da Pesquisa

Esta pesquisa foi aprovada pelo Comitê de Ética em Pesquisa do Centro de Ciências Médicas da Universidade Federal da Paraíba (CEP/CCM/UFPB), sob o parecer 6.879.883 (Anexo 2). O projeto atende as normas e diretrizes que regulamentam a pesquisa envolvendo seres humanos dispostas na Resolução 466/2012, do Conselho Nacional de Educação (CNE). De acordo com esta resolução, é garantido o anonimato do participante por meio da apresentação do Termo de Consentimento Livre e Esclarecido (TCLE) (Apêndice 3).

A pesquisa trouxe como benefício a construção e validação de uma simulação computacional que poderá auxiliar estudantes e profissionais de saúde no treinamento da classificação dos parâmetros da deglutição avaliados durante a VED. A aquisição desta habilidade é importante para o diagnóstico em disfagia orofaríngea e o seu aprimoramento trará a contribuição de desenvolver condutas profissionais mais assertivas nesta área de conhecimento.

Em relação aos riscos, existia a possibilidade de os participantes da pesquisa sentirem-se intimidados ou desconfortáveis no fornecimento dos dados necessários. Neste caso, eles poderiam desistir de participar a qualquer momento, sem nenhum ônus, e receberiam o apoio necessário da pesquisadora responsável.

4. RESULTADOS E DISCUSSÃO

Visando exibir o alcance dos objetivos propostos para este trabalho, esta seção foi dividida em partes, relativas à: 1) Revisão da literatura; 2) Desenvolvimento do simulador; 3) Modelo de decisão e 4) Validação da usabilidade do simulador.

4.1. REVISÃO DE ESCOPO DA LITERATURA

Para o desenvolvimento deste trabalho, a primeira etapa teve foco na identificação de trabalhos existentes envolvendo treinamento para a análise de parâmetros da VED. Considerando este objetivo, foi realizada uma revisão de escopo, que consiste em uma forma sistemática de revisão que tem como objetivo mapear e identificar as evidências em determinado tema. As revisões de escopo buscam identificar conceitos, teorias, fontes e lacunas de conhecimento (TRICCO et al., 2018), assim como implicações para a tomada de decisão (TRICCO et al., 2016).

Foram buscados por meio de estratégias específicas trabalhos que realizassem treinamento para analisar parâmetros que são considerados na avaliação endoscópica da deglutição, com o objetivo de identificar especificações básicas sobre o exame e explorar o conteúdo dos treinamentos. Para isso, um protocolo a priori foi publicado (<https://doi.org/10.1590/1982-0216/202224111021>) (COSTA et al., 2022) para relatar a metodologia utilizada, incluindo os objetivos, critérios de inclusão e métodos para analisar e apresentar os dados numa revisão de escopo. O manuscrito final da revisão, contendo os resultados encontrados, está publicado no periódico *Dysphagia* (<https://doi.org/10.1007/s00455-023-10614-w>) (COSTA et al., 2024).

Os achados da revisão de escopo também serviram como base para a construção da revisão narrativa apresentada na Fundamentação Teórica deste trabalho, fornecendo subsídios para discutir os principais conceitos relacionados ao treinamento para a classificação dos resíduos faríngeos na VED.

4.2. DESENVOLVIMENTO DO SIMULADOR DE TREINAMENTO DEGLUTSIM

4.2.1. Análise dos Requisitos

Após consulta prévia à literatura, onde foi verificada a variabilidade existente na estrutura dos treinamentos, foi realizada uma consulta a uma fonoaudióloga especialista com experiência na classificação dos parâmetros da VED, no intuito de levantar requisitos para

desenvolvimento do ST. Para tal, a questão norteadora foi: “O que deve conter em um simulador computacional para treinamento da classificação dos parâmetros da deglutição pela VED?”. Os requisitos levantados estão resumidos no Quadro 5.

Quadro 5: Requisitos do especialista

Nº	Requisitos
RE01	Disponibilizar os exames em vídeo, e não somente fotos como em alguns treinamentos
RE02	Disponibilizar amostras do parâmetro escolhido em todos os níveis de gravidade considerados
RE03	Possibilidade de <i>play/pause</i> no vídeo
RE04	Não mostrar volume e consistência dos alimentos ofertados para os pacientes durante os exames – não impacta no aprendizado
RE05	Não levar em conta o tempo para avaliar, mas a quantidade de exames, visto que o aprendizado ocorre por repetição
RE06	Possibilitar a consulta de imagens âncoras (imagens de referência) da escala escolhida para ser utilizada durante o treinamento
RE07	Avaliar o efeito do treinamento considerando a precisão e a confiabilidade intra-avaliador
RE08	Requerer pontuação para passar de fase
RE09	<i>Feedback</i> de forma nominal (ex.: ruim, bom, excelente) ou percentual (ex.: acertou x%)

Considerando os itens sinalizados pela especialista, os requisitos computacionais para desenvolvimento do sistema foram propostos. A estruturação do ST compreende três fases e três etapas.

As duas primeiras fases possuem o mesmo nível de dificuldade, pois o usuário deverá classificar a gravidade dos resíduos faríngeos nas regiões de valéculas e seios piriformes de forma isolada. Na terceira fase, a classificação será feita pelo usuário conforme as condições reais de realização do exame, exigindo a classificação da gravidade dos resíduos faríngeos para ambas as regiões faríngeas.

A primeira etapa (Ensino) visa instruir o usuário sobre os níveis de gravidade considerados para a classificação. Cada nível será ilustrado por um vídeo, rotulado pelos especialistas, acompanhado pelas referências da escala escolhida (imagem âncora, descrição operacional, percentual). Na segunda etapa (Nivelamento), o usuário enfrentará um desafio que requer acertar a classificação de todos os quatro casos apresentados (considerados mais fáceis) para avançar para a próxima etapa. A terceira etapa (Prática) corresponde à avaliação propriamente dita. A precisão do usuário em classificar 20 casos será avaliada, incluindo os

quatro casos previamente acertados na etapa Nivelamento, para permitir a análise da confiabilidade intra-avaliador.

Durante as etapas de Nivelamento e Prática, o usuário poderá consultar as imagens âncoras e pausar/reproduzir os vídeos conforme necessário. O *feedback* será apresentado de forma nominal (frases) e por meio de estrelas (0-3).

Para as fases 01 e 02, correspondentes ao treinamento das classificações de resíduos faríngeos em valéculas e seios piriformes, respectivamente, três etapas serão consideradas. Já para a fase 03, que considerará a classificação de valéculas e seios piriformes em um mesmo exame, apenas as fases de Nivelamento e Prática serão consideradas. A seguir está a descrição dos requisitos computacionais e suas correspondências, dentro de cada etapa do sistema, com os requisitos do especialista (Quadros 6, 7 e 8).

Os Quadros 6, 7 e 8 descrevem uma série de requisitos para a progressão do usuário dentro do simulador de treinamento. Cada **requisito** é identificado por um código único (como RCE01, RCE02...) que define uma ação específica que o usuário deve realizar. A **resposta** ocorre quando uma ação é concluída, por exemplo, ao assistir o vídeo de um caso clínico, o usuário confirma que concluiu a tarefa. O **pré-requisito** indica as condições que devem ser atendidas antes de realizar a ação, garantindo que o usuário siga uma sequência lógica, como assistir o vídeo de gravidade “Nenhum” antes de passar para o vídeo de gravidade “Vestígio”. A **descrição** fornece informações sobre o que o usuário deve fazer e os resultados esperados, como o avanço para o próximo nível. A **prioridade** sinaliza a importância do requisito. O **tipo** indica o tipo da variável, de acordo com as características dos valores que ela pode assumir. A coluna **parâmetro** estabelece os possíveis valores ou respostas que o sistema aceita. Finalmente, o **requisito do especialista** vincula o requisito computacional aos requisitos do especialista que precisam ser cumpridos. Os elementos descritos criam um fluxo organizado e lógico de atividades de aprendizado e avaliação dentro do simulador de treinamento.

Quadro 6: Requisitos computacionais para a etapa Ensino

Requisito	Ação	Resposta	Pré-requisito	Descrição	Prioridade	Tipo	Requisito do espec.
RCE01	Assistir ao vídeo de gravidade “Nenhum”	Assistir	-	O usuário indica que assistiu ao vídeo do caso nível “Nenhum”	Obrigatório	Discreta	RE01, RE02, RE03, RE04
		Não Assistir					
RCE02	Assistir ao vídeo de gravidade “Vestígio”	Assistir	Assistir ao vídeo de gravidade “Nenhum”	O usuário indica que assistiu ao vídeo do caso nível “Vestígio”	Obrigatório	Discreta	
		Não Assistir					
RCE03	Assistir ao vídeo de gravidade “Leve”	Assistir	Assistir ao vídeo de gravidade “Vestígio”	O usuário indica que assistiu ao vídeo do caso nível “Leve”	Obrigatório	Discreta	
		Não Assistir					
RCE04	Assistir ao vídeo de gravidade “Moderado”	Assistir	Assistir ao vídeo de gravidade “Leve”	O usuário indica que assistiu ao vídeo do caso nível “Moderado”	Obrigatório	Discreta	
		Não Assistir					
RCE05	Assistir ao vídeo de gravidade “Grave”	Assistir	Assistir ao vídeo do caso nível “Moderado”	O usuário indica que assistiu ao vídeo do caso nível “Grave”	Obrigatório	Discreta	
		Não Assistir					
RCE06	Passar para o Nivelamento	Passar	Assistir ao vídeo do caso nível “Grave”	O usuário indica que concluiu o nível Ensino e segue para a etapa Nivelamento	Obrigatório	Discreta	
		Não Passar					

Quadro 7: Requisitos computacionais para a etapa Nivelamento

Requisito	Ação	Resposta	Pré-requisito	Descrição	Prioridade	Tipo	Parâmetro	Requisito do espec.
RCN01	Consultar imagens âncoras no Nivelamento	Consultar	Concluir a etapa Ensino	O usuário consulta as imagens âncoras para auxiliar na avaliação do caso a ser classificado	Opcional	Discreta	Verdadeiro ou falso	RE06
		Não consultar						
RCN02	Assistir ao vídeo do caso para classificar a gravidade	Assistir	Concluir a etapa Ensino	O usuário assiste o vídeo do caso que irá realizar a classificação	Obrigatório	Discreta	Verdadeiro ou falso	RE01, RE03, RE04
RCN03	Classificar o caso apresentado	Classificar	Concluir a etapa Ensino	O usuário irá classificar o caso apresentado de acordo com a gravidade	Obrigatório	Ordinal	Nenhum, Vestígio, Leve, Moderado ou Grave	RE01, RE03, RE04
		Não classificar						
RCN04	Avançar para a etapa Prática	Avançar	1) Concluir a etapa Ensino 2) Classificar corretamente os casos apresentados no Nivelamento (RCN03)	Soma das classificações realizadas corretamente pelo usuário	Obrigatório	Ordinal	0-4	RE08, RE09

4.2.2. Fluxo do sistema

Um fluxograma sobre o funcionamento do simulador está apresentado na Figura 6. O treinamento ocorre em diferentes etapas e fases, conforme descrito a seguir.

O treinamento inicia com uma etapa instrucional – **Ensino** (etapa 01), onde o usuário visualiza exemplos de exames de VED de todos os níveis das escalas adotadas. A YPRSRS possui cinco níveis: “I - nenhum”, “II - vestígio”, “III - leve”, “IV - moderado” e “V - grave”, considerados para duas regiões faríngeas específicas, valéculas e seios piriformes. Todos os níveis da escala, acompanhados de vídeo representativo, imagem âncora, definição operacional e percentual são apresentados ao usuário nesta etapa.

A etapa 01 ocorre nas fases 01 e 02, relativas a **resíduos faríngeos em valéculas (RFV)** e **resíduos faríngeos em seios piriformes (RFS)**. A fase 03 considera **resíduos faríngeos em valéculas e seios piriformes (RFVS)** e não possui etapa de Ensino. Diferente da etapa 01 (Ensino), as etapas 02 (Nivelamento) e 03 (Prática) ocorrem nas três fases (RFV, RFS e RFVS) com objetivos diferentes.

Na etapa 02 – **Nivelamento**, são expostos para a avaliação exames que foram considerados simples de serem classificados visualmente, alocados nos níveis I ou V de forma unânime pelos especialistas (quatro em cada tentativa, sendo dois de cada nível da escala). Caso faça a classificação errada, o usuário recebe o *feedback* negativo, solicitando que classifique novamente os exames, até que classifique todos corretamente.

Por último, na etapa 03 – **Prática**, o usuário deve classificar exames de todos os níveis da escala. Julgará no mínimo quatro exames de cada um dos cinco níveis da YPRSRS, totalizando 20 exames em cada fase. Em cada uma das três fases, os exames apresentados e acertados na etapa **Nivelamento** da fase correspondente (RFV, RFS ou RFVS) aparecerão para serem classificados novamente de forma aleatória, embaralhados entre os 20 exames apresentados, para que seja possível a análise de confiabilidade intra-avaliador. Não é necessário que o usuário acerte todas as classificações dos exames apresentados nesta etapa para concluir a fase em que se encontra. Contudo, o módulo da fase RFVS estará disponível somente após a conclusão das fases RFV e RFS.

A classificação (nível) dos exames que serão apresentados dentro do sistema durante a etapa **Prática** se dá de forma aleatória. Alguns exames possuem mais de uma classificação, nos casos em que houve divergência na classificação dos especialistas, com variação na faixa de 1 nível (I-II, II-III, III-IV...). Mesmo considerando esta diferença, o sistema garante o mínimo de 4 exames para cada nível.

As fases RFV e RFS possuem as etapas 01 (**Ensino**), 02 (**Nivelamento**) e 03 (**Prática**). Já a fase RFVS é composta pelas etapas 02 e 03. Nela, o usuário deve classificar resíduo faríngeo em valécula e em seios piriformes em cada exame, tal qual ocorre em condições reais. Por esta razão, a fase RFVS somente é desbloqueada após a conclusão das fases RFV e RFS.

A etapa 01 é destinada à instrução sobre a escala utilizada. Nela, são mostradas definições sobre cada nível da escala: imagem âncora (representação visual), descrição (representação escrita) e um vídeo demonstrativo do nível, previamente selecionado. A fase RFVS somente é desbloqueada quando o usuário completa as fases RFV e RFS, que já contemplam uma etapa de Ensino. O usuário pode optar por iniciar o processo dentro do simulador tanto pela fase RFV quanto pela RFS.

A etapa 02 tem como objetivo avaliar o desempenho do usuário para a classificação de exames considerados mais simples de serem avaliados, após a instrução recebida. Neste momento, o usuário realiza a classificação do(s) parâmetro(s) e recebe *feedback* sobre seu desempenho, devendo acertar a classificação de todos os exames. O acerto significa a concordância entre a classificação do usuário e a dos especialistas, e o erro significa a discordância entre as avaliações. O erro implica repetir classificações em novos exames, randomizados, até que a taxa de acerto seja de 100%.

Já a etapa 03 avalia o desempenho do usuário após o treinamento. Nela, o usuário julgará 20 exames distribuídos por todos os níveis da escala, incluindo os exames classificados corretamente na etapa 02, para análise de confiabilidade intra-avaliador. Ao final, o usuário receberá *feedback* de acordo com o seu relatório de desempenho, que será baseado em medidas de acurácia e confiabilidade intra-avaliador.

Nas fases do sistema as etapas precisam ser todas cumpridas para avançar para outra fase. Apesar disso, o usuário pode retornar à fase Ensino a qualquer momento.

Fonte: Elaborado pelo autor, 2023.

4.2.3. DeglutSIM

Ao executar o DeglutSIM – Simulador de treinamento perceptivo-visual nos parâmetros da deglutição, o usuário terá acesso ao Menu do simulador (Figura 7).



Figura 7: Tela inicial

Fonte: Elaborado pelo autor, 2023.

Ao clicar em “iniciar”, o usuário será direcionado para um menu de opções, onde escolherá entre as fases disponíveis para que possa ser iniciada a simulação (Figura 8). As opções disponíveis neste momento serão as fases RFV e RFS. Ao optar por uma das duas fases disponíveis, o usuário será direcionado para uma tela de seleção da etapa, onde terá a opção de selecionar a etapa Ensino inicialmente (Figura 9).



Figura 8: Seleção de fase.

Fonte: Elaborado pelo autor, 2023.

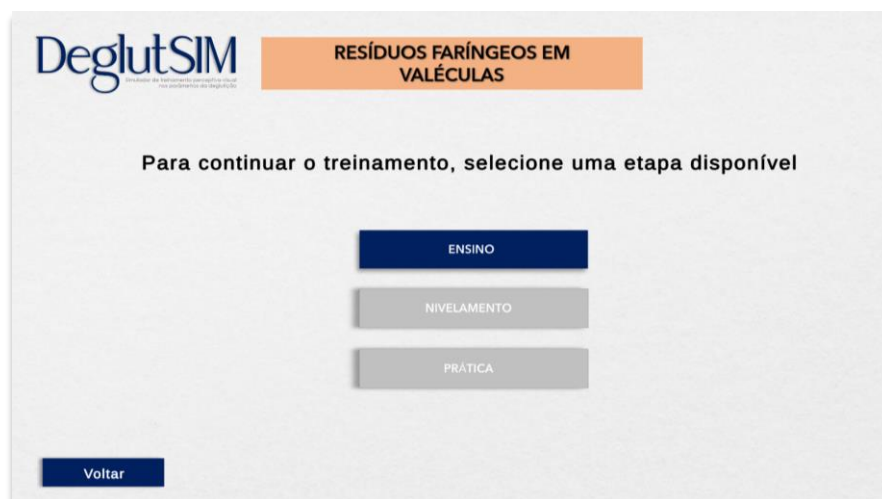


Figura 9: Seleção de etapa.

Fonte: Elaborado pelo autor, 2023.

Etapa Ensino

Ao acessar a etapa Ensino, o usuário é instruído quanto aos objetivos da referida etapa, sobre as funcionalidades disponíveis nesta etapa e os requisitos para a sua finalização. Após sinalizar que leu a mensagem, a tela principal da etapa é exibida (Figura 10 e Figura 11).



Figura 10: Instruções na etapa Ensino.

Fonte: Elaborado pelo autor, 2023.



Figura 11: Tela principal da etapa Ensino.

Fonte: Elaborado pelo autor, 2023.

As imagens âncoras da escala YPRSRS funcionam como botões na tela principal da etapa Ensino. À medida que o usuário clica nestes botões, uma tela contendo todos os itens disponibilizados pela escala (imagem âncora, descrição operacional e percentual) para cada nível é exibido ao usuário, assim como um recorte de vídeo referente ao nível apresentado. A Figura 12 demonstra esta visualização na fase RFV.



Figura 12: Tela referente ao nível de gravidade I da YPRSR.

Fonte: Elaborado pelo autor, 2023.

À medida que visualiza um nível de gravidade e clica em “continuar”, o botão correspondente ao nível YPRSRs subsequente é habilitado para clique, até que todas as opções sejam visualizadas (I a V). Desta forma, o usuário estará habilitado a acessar a fase Nivelamento.

Etapa Nivelamento

Tal como na etapa Ensino, ao acessar a etapa Nivelamento o usuário recebe instruções sobre o objetivo da etapa, sobre como realizar a classificação dos resíduos faríngeos e sobre as funcionalidades disponíveis (Figura 13). A tela de classificação nesta etapa dispõe das opções de níveis de gravidade, consulta às instruções, consulta às âncoras com possibilidade de manter abertas ou ocultar e sinalização do progresso do usuário na etapa (Figura 14).



Figura 13: Instruções da etapa Nivelamento na fase RFV.

Fonte: Elaborado pelo autor, 2023.



Figura 14: Tela principal da etapa Nivelamento com âncora aberta na fase RFV.

Fonte: Elaborado pelo autor, 2023.

As fases RFV e RFS possuem design semelhante nesta etapa. A interface é idêntica, com a devida indicação sobre a etapa em que o usuário se encontra no canto superior direito da tela e mudança de cores para situar o usuário sobre a mudança (RFV laranja, RFS roxo). Já a fase RFVS possui um design diferente, com opções de botões de classificação referente à valéculas e seios piriformes dos dois lados, e opções de âncoras também de valéculas e de seios piriformes, que só podem ser abertas de forma assíncrona (Figuras 15 e 16).



Figura 15: Tela principal da etapa Nivelamento na fase RFVS.

Fonte: Elaborado pelo autor, 2023.



Figura 16: Tela principal do Nivelamento RFVS, com âncora RFS aberta.

Fonte: Elaborado pelo autor, 2023.

Ao realizar a classificação de quatro casos, o usuário recebe o *feedback* sobre o seu desempenho nesta etapa. Caso tenha acertado a classificação dos quatro exames, recebe *feedback* positivo (Figura 17). Nos casos em que o usuário errou a classificação de pelo menos um caso, recebe *feedback* negativo (Figura 18).

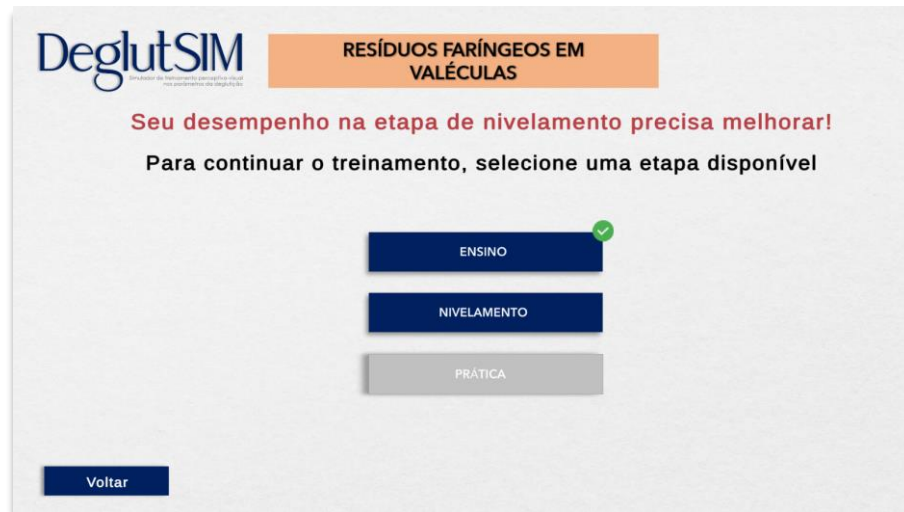


Figura 17: *Feedback* negativo na etapa Nivelamento.

Fonte: Elaborado pelo autor, 2023.

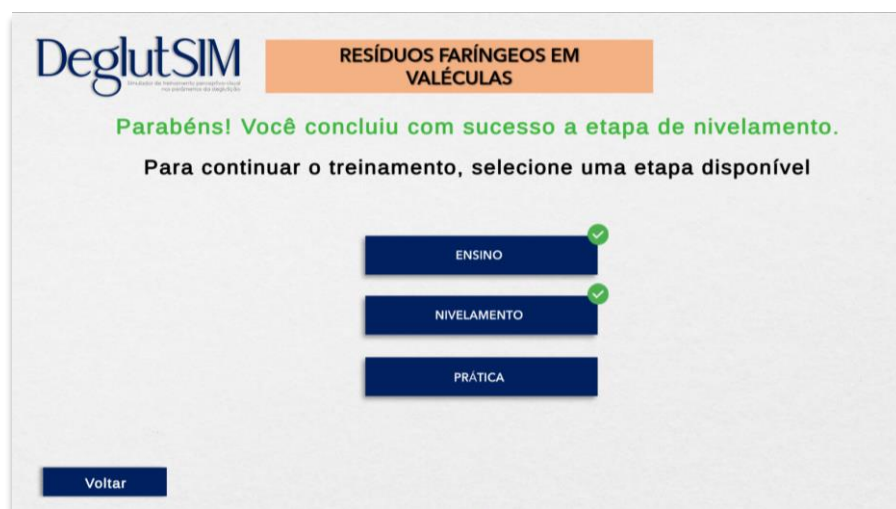


Figura 18: *Feedback* positivo da etapa Nivelamento.

Fonte: Elaborado pelo autor, 2023.

Etapa Prática

Ao chegar na etapa Prática, o usuário visualiza as instruções, de forma semelhante à etapa anterior, e classifica 20 exames para a avaliação propriamente dita dentro do ST. As fases RFV, RFS e RFVS possuem design semelhante à etapa Nivelamento para cada fase correspondente, com a devida indicação sobre a etapa em que o usuário se encontra no canto superior direito da tela.

Relatório de Desempenho

Após finalizar a etapa Prática, um relatório de desempenho é apresentado ao usuário. As Figuras 19, 20 e 21 ilustram um exemplo de relatório de desempenho para um *feedback* Insatisfatório, Bom e Excelente, respectivamente.



Figura 19: *Feedback* final desempenho Insatisfatório.

Fonte: Elaborado pelo autor, 2023.



Figura 20: Feedback final desempenho Bom.

Fonte: Elaborado pelo autor, 2023.



Figura 21: Feedback final desempenho Excelente.

Fonte: Elaborado pelo autor, 2023.

4.3. MODELO DE DECISÃO INTELIGENTE

A variabilidade e os dados inconsistentes encontrados na literatura sobre o treinamento para a classificação dos resíduos faríngeos na VED utilizando escalas validadas justificou a necessidade do desenvolvimento de um modelo de avaliação para o treinamento desta habilidade perceptivo-visual específica. O desenvolvimento do modelo de decisão utilizando um SBR com estrutura baseada na lógica clássica, que considera até duas respostas possíveis para cada classificação, buscou mitigar a subjetividade frequentemente associada à classificação dos resíduos faríngeos na VED. Esse avanço possibilitou a implementação de um simulador capaz de fornecer *feedback* imediato aos usuários.

Foram realizados testes preliminares com o simulador, com o objetivo de verificar erros de navegação, inconsistência de dados, mecanismos de interação e a resposta do modelo de decisão inteligente integrado ao simulador. Para o teste, todas as combinações possíveis de acertos para cada possibilidade de *feedback* foram listadas, o que totalizou 4.096 casos. Pelo menos dois casos para *feedback* de 0, 1, 2 e 3 estrelas, selecionados aleatoriamente, foram verificados. A Figura 22 mostra o exemplo de um de caso para cada tipo de *feedback* possível na avaliação.

O modelo de decisão proposto foi desenvolvido para avaliar o desempenho dos usuários no simulador de treinamento perceptivo-visual para classificação dos resíduos faríngeos na videoendoscopia da deglutição (VED). Baseado na lógica clássica, o modelo mitigou a subjetividade frequentemente associada à classificação, ao adotar critérios objetivos e mais abrangentes do que os convencionais para mensurar confiabilidade intra-avaliador e precisão. Esses critérios, fundamentados em índices amplamente aceitos como a confiabilidade intra-avaliador através do coeficiente kappa e a medida de acurácia, proporcionaram uma estrutura robusta para avaliação realizada pelo sistema. Além disso, o modelo permitiu a geração de relatórios detalhados de desempenho, com *feedback* imediato ao usuário, visando contribuir para a formação de estudantes e profissionais de saúde na prática de avaliação dos resíduos faríngeos.

	VALÉCULAS		SEIOS PIRIFORMES		VALÉCULAS E SEIOS PIRIFORMES		PONTUAÇÃO FINAL	FEEDBACK
	PRECISÃO	CONFIABILIDADE INTRA-AVALIADOR	PRECISÃO	CONFIABILIDADE INTRA-AVALIADOR	PRECISÃO	CONFIABILIDADE INTRA-AVALIADOR		
 Usuário 1	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	"Seu desempenho pode ser melhorado. Recomendamos que faça o treinamento novamente."
 Usuário 2	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	"Seu desempenho pode ser melhorado. Recomendamos que faça o treinamento novamente."
 Usuário 3	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	"Seu desempenho foi satisfatório, mas ainda há espaço para aprimoramento. Caso julgue necessário, pode repetir o treinamento."
 Usuário 4	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★	"Parabéns! Seu desempenho no treinamento foi excelente."

Figura 22: Exemplos de casos simulados durante testes do modelo de decisão inteligente.

Fonte: Elaborado pelo autor, 2024.

Durante os testes preliminares do simulador, foi assegurado que o sistema fosse rigoroso na obtenção das estrelas no relatório de desempenho. Para confiabilidade intra-avaliador, era necessário acertar pelo menos 2 de 4 casos (50%) para conquistar uma estrela; para precisão, acertar 12 de 20 casos (60%); e para a pontuação final, no mínimo 6 de 18 estrelas (33,3%). Essa exigência refletiu a adoção de medidas estatísticas robustas neste estudo, o coeficiente kappa e a medida de acurácia, que foram utilizadas como parâmetros no modelo de avaliação. A precisão diagnóstica em casos de pacientes com disfagia desempenha um papel essencial na prevenção de complicações graves, como desidratação, desnutrição e danos pulmonares decorrentes da aspiração laringotraqueal, que podem levar a quadros de pneumonia aspirativa e até mesmo à morte do paciente. Visto isso, a rigorosidade dos critérios foi projetada para alinhar os resultados do treinamento às demandas da prática clínica, promovendo maior assertividade e segurança no diagnóstico em disfagia orofaríngea.

A integração desse modelo ao simulador demonstrou seu potencial para impactar positivamente a prática clínica e o ensino na área de Disfagia. Ao estruturar o treinamento de forma progressiva e incluir etapas específicas para a avaliação da habilidade perceptivo-visual, o simulador promoveu uma abordagem que reflete a prática clínica real, mas em um ambiente controlado, que permite erros por parte do treinado sem acarretar prejuízos para a saúde do paciente.

Embora o presente trabalho não apresente uma análise quantitativa do desempenho dos usuários, o modelo garantiu que as métricas implementadas no simulador fossem coerentes com os objetivos propostos, permitindo avaliações precisas, rápidas e consistentes. Futuros estudos podem explorar a aplicação do sistema em contextos clínicos e educacionais, avaliando empiricamente o impacto do modelo de decisão na melhoria da acurácia diagnóstica e na redução da variabilidade entre avaliadores.

4.4. AVALIAÇÃO DO ST

Finalizados os passos de implementação e os testes preliminares do simulador, o passo seguinte dedicou-se à avaliação da usabilidade da ferramenta. O objetivo foi verificar a satisfação do usuário quanto à usabilidade de um simulador computacional para treinamento perceptivo-visual na videoendoscopia da deglutição (VED).

O público-alvo do estudo foram profissionais de saúde envolvidos no tratamento de pacientes com disfagia, recrutados por conveniência. A Figura 23 ilustra uma avaliação realizada nas dependências do LabTEVE.



Figura 23: Profissional de saúde realizando simulação no DeglutSIM.

Fonte: Do autor, 2024.

4.4.1. Caracterização da amostra

Um estudo piloto foi inicialmente realizado com um participante, com o objetivo de testar a aplicação dos instrumentos de coleta de dados. O teste piloto ajudou a corrigir problemas iniciais que pudessem comprometer a qualidade dos dados, permitindo testar a logística do estudo. No caso deste estudo, a realização do teste foi importante para adequar o tempo necessário para a coleta de dados, pois o primeiro participante (U01) relatou um tempo exaustivo para a conclusão da simulação, o que não tinha sido claramente informado. As devidas adaptações foram realizadas e aplicadas para os participantes seguintes.

Os dados coletados do primeiro participante (U01) foram excluídos do estudo, resultando em uma amostra final composta por 10 participantes (U02-U11) com média de idade de $33,1 \pm 7,23$ anos, todos graduados em Fonoaudiologia. Nenhum dos participantes recebeu treinamento para classificar resíduos faríngeos na VED durante a sua formação profissional, seja na graduação ou na pós-graduação. Dentre os participantes, quatro tinham experiência na realização da VED. A média de tempo de experiência em VED nos participantes que possuíam experiência foi de 5,7 (mínimo 4, máximo 7) anos. A distribuição dos participantes quanto ao

gênero, anos de atuação em Disfagia, experiência em VED e formação profissional está descrita na Tabela 4.

Embora os resultados obtidos forneçam dados preliminares relevantes, é importante destacar que o tamanho reduzido da amostra limita a generalização dos achados. Estudos futuros devem incluir uma avaliação com definição da amostra estatística, de forma a garantir maior robustez e representatividade dos resultados.

Tabela 4: Distribuição dos participantes quanto ao gênero, anos de atuação em Disfagia, experiência em VED e formação profissional (n=10).

	<i>n (%)</i>
Gênero	
Feminino	8 (80,0)
Masculino	2 (20,0)
Anos de atuação em Disfagia	
<5 anos	3 (30,0)
≥5 anos	5 (50,0)
≥10 anos	2 (20,0)
Experiência em VED	
Sim	4 (40,0)
Não	6 (60,0)
Pós-Graduação <i>Stricto Sensu</i>	
Doutorado	1 (10,0)
Mestrado	3 (30,0)
Não possui	6 (60,0)
Pós-Graduação <i>Lato Sensu</i>	
Especialização em Disfagia	6 (60,0%)
Residência Multiprofissional em Saúde Hospitalar <i>em curso</i>	2 (20,0%)
Residência Multiprofissional em Saúde Hospitalar <i>concluída</i>	1 (10,0%)
Especialização em Voz	1 (10,0%)

4.4.2. Usabilidade

A usabilidade é um conjunto fundamental de propriedades e atributos que, ao focar na interação entre seres humanos e sistemas computacionais, assegura a qualidade do projeto e a satisfação dos usuários. Esta análise, centrada no usuário, possibilita ajustar e atualizar o sistema para minimizar erros e atender às expectativas e necessidades (MARTINS et al., 2012; MORATO et al., 2020). A usabilidade avaliada por meio da escala SUS (BROOKE, 1996) permite mensurar aspectos de facilidade de aprendizagem, facilidade de memorização e satisfação do usuário, além de estar relacionada com a eficiência do sistema. Com as respostas obtidas da aplicação da escala, foram calculadas a média individual de cada participante, assim como a média das notas individuais dos participantes. A Tabela 5 apresenta o *score* de usabilidade do simulador DeglutSIM.

Tabela 5: Scores SUS e classificação correspondente de acordo com cada participante da pesquisa e média geral dos scores.

Participante	Score SUS	Classificação
U02	95	Melhor imaginável
U03	92,5	Melhor imaginável
U04	100	Melhor imaginável
U05	100	Melhor imaginável
U06	97,5	Melhor imaginável
U07	85	Excelente
U08	100	Melhor imaginável
U09	75	Bom
U10	85	Excelente
U11	97,5	Melhor imaginável
Total	92,7	Melhor imaginável

A média apontada na literatura para o *score* SUS é de 68 (BANGOR; KORTUM; MILLER, 2009). A Tabela 5 demonstra que o DeglutSIM obteve um *score* de usabilidade de 92,7, dentro da classificação “melhor imaginável”, que se encontra na faixa de aceitabilidade e na classificação mais alta da escala (BROOKE, 2013). O maior índice individual apresentado foi o dos usuários 04, 05 e 08, que atingiram 100 pontos, e o menor, pertence ao usuário 09, com 75 pontos e classificação “bom”.

No intuito de medir o grau de confiabilidade entre os avaliadores, foi realizado o teste Kappa ponderado para verificar o grau de discordância, atribuindo pesos diferentes a diferentes níveis de discordância entre eles. A Tabela 6 mostra o cálculo do kappa para cada par de avaliadores.

Tabela 6: Cálculo do Kappa ponderado para cada par de avaliadores dos itens SUS.

Classificações	Kappa ponderado	Erro padrão Assintótico	Z	Sig.	Limite inferior	Limite Superior
U02 - U03	,684	,118	2,639	,008	,453	,916
U02 - U04	,750	,142	2,734	,006	,471	1,029
U02 - U05	,750	,142	2,734	,006	,471	1,029
U02 - U06	,798	,143	2,973	,003	,517	1,079
U02 - U07	,526	,208	2,062	,039	,118	,934
U02 - U08	,750	,142	2,734	,006	,471	1,029
U02 - U09	,551	,119	2,722	,006	,317	,784
U02 - U10	,775	,087	3,383	<,001	,605	,946
U02 - U11	,697	,146	2,597	,009	,412	,982
U03 - U04	,850	,072	3,058	,002	,709	,991
U03 - U05	,850	,072	3,058	,002	,709	,991
U03 - U06	,798	,080	2,948	,003	,641	,954
U03 - U07	,531	,164	2,096	,036	,209	,853
U03 - U08	,850	,072	3,058	,002	,709	,991
U03 - U09	,659	,110	3,162	,002	,443	,875
U03 - U10	,667	,094	2,860	,004	,483	,850
U03 - U11	,899	,064	3,321	<,001	,774	1,024
U04 - U05	1,000*	,000	3,162	,002	1,000	1,000
U04 - U06	,950	,047	3,128	,002	,858	1,042
U04 - U07	,700	,161	2,556	,011	,384	1,016
U04 - U08	1,000	,000	3,162	,002	1,000	1,000
U04 - U09	,550	,084	2,850	,004	,385	,715
U04 - U10	,650	,095	2,900	,004	,465	,835
U04 - U11	,950	,047	3,128	,002	,858	1,042
U05 - U06	,950	,047	3,128	,002	,858	1,042
U05 - U07	,700	,161	2,556	,011	,384	1,016
U05 - U08	1,000*	,000	3,162	,002	1,000	1,000
U05 - U09	,550	,084	2,850	,004	,385	,715
U05 - U10	,650	,095	2,900	,004	,465	,835
U05 - U11	,950	,047	3,128	,002	,858	1,042
U06 - U07	,745	,166	2,742	,006	,420	1,070
U06 - U08	,950	,047	3,128	,002	,858	1,042
U06 - U09	,583	,073	2,953	,003	,440	,727
U06 - U10	,694	,100	3,135	,002	,498	,890
U06 - U11	,899	,064	3,035	,002	,774	1,024
U07 - U08	,700	,161	2,556	,011	,384	1,016
U07 - U09	,489	,125	2,431	,015	,245	,733
U07 - U10	,396*	,171	1,810	,070	,061	,730
U07 - U11	,643	,164	2,366	,018	,321	,965
U08 - U09	,550	,084	2,850	,004	,385	,715
U08 - U10	,650	,095	2,900	,004	,465	,835
U08 - U11	,950	,047	3,128	,002	,858	1,042
U09 - U10	,605	,099	2,812	,005	,411	,800
U09 - U11	,583	,095	2,953	,003	,398	,769
U10 - U11	,592	,096	2,674	,007	,403	,780

O cálculo do kappa ponderado mostrou que na comparação entre os avaliadores em pares, há confiabilidade entre moderada e perfeita na maioria dos casos. Na comparação entre os usuários 04-05, 04-08, 05-08 houve uma concordância perfeita ($\kappa = 1,00$; $p = 0,002$). Entre as comparações, apenas entre os usuários 07-10 houve uma confiabilidade fraca ($\kappa=0,39$; IC 95%; $p=0,70$; IC 95% [0,061 – 0,730]).

Uma observação válida é que os usuários 07 e 09 possuem um intervalo de tempo de atuação significativo. Enquanto o usuário 07 possui 0,33 anos de atuação em Disfagia, o usuário 10 possui cinco anos de atuação. Desta forma, a experiência clínica pode interferir nos aspectos de avaliação da ferramenta, o que pode ter influenciado nesta diferença entre as avaliações.

Outra observação feita pela pesquisadora é que o usuário 07 é um profissional de saúde em aperfeiçoamento profissional (fonoaudiólogo residente) e compareceu ao local da realização do teste acompanhado de sua supervisora, que participou da simulação antes dele. A observação do desempenho da supervisora no simulador pode ter influenciado as respostas deste participante no teste, visto que pelo momento de formação em que se encontra, este profissional tende a colocar-se no lugar de aprendiz com maior frequência e acredita precisar de maior apoio do que seus supervisores para desenvolver uma habilidade clínica.

A partir de uma análise centralizada nos aspectos específicos que o SUS avalia (facilidade de aprendizagem, facilidade de memorização, satisfação do usuário e eficiência do sistema), é possível realizar uma análise focada nos itens relacionados a cada aspecto. Um boxplot apresentado na Figura 23 mostra a análise item a item da escala SUS. Esta representação gráfica fornece uma visualização sobre a distribuição de um conjunto de dados através dos quartis (Q1, Q2 e Q3), com destaque para a mediana (Q2), assim como variabilidade dos dados e valores discrepantes. A análise a partir desta representação pode auxiliar na identificação de padrões de desempenho.

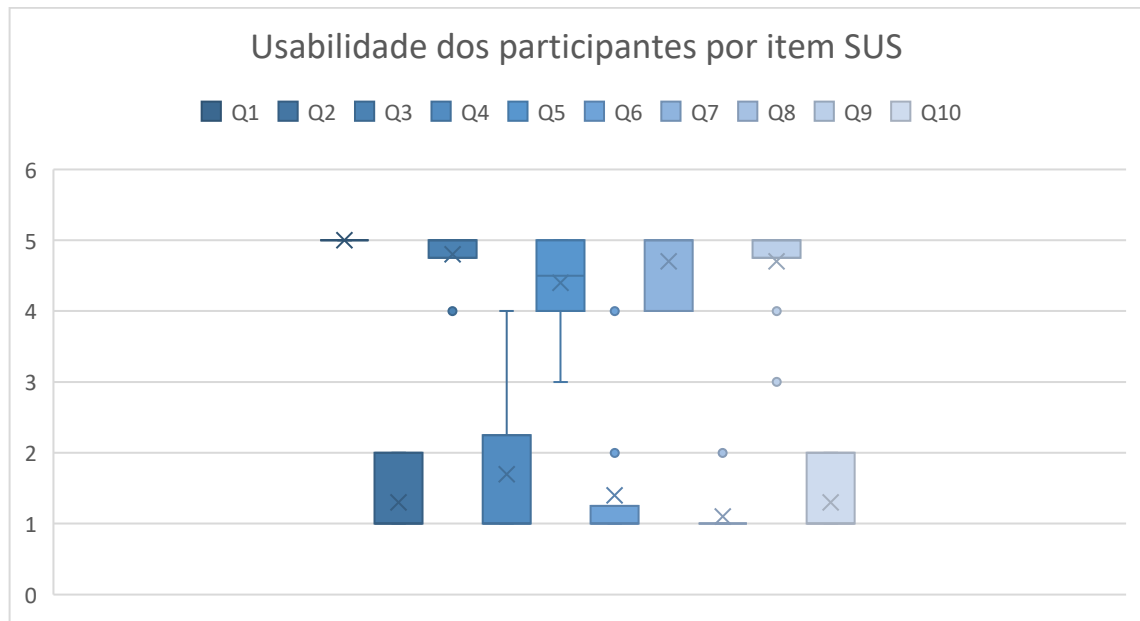


Figura 24: Boxplot com a usabilidade dos participantes por item SUS.

Fonte: Elaborado pelo autor, 2024.

Com a análise individual dos itens da escala, observa-se que os itens ímpares apresentam valores de mediana mais próximos de 5, e os itens pares, valores mais próximos de 1. De forma geral, a maioria dos itens ímpares apresentam valores entre 4 e 5, e os itens pares, valores entre 1 e 2. Na escala SUS, para os itens ímpares, os valores próximos a 5 indicam uma avaliação positiva da usabilidade; e para os itens pares, os valores próximos a 1 é que indicam uma avaliação positiva, dada a natureza da frase de afirmação.

Dois itens apresentaram um valor discrepante na avaliação (*outlier*), os itens Q6 (“Acho que o sistema apresenta muitas inconsistências”) e Q9 (“Eu me senti muito confiante ao usar o sistema”). Estes itens referem-se, respectivamente, à eficiência do sistema e à satisfação do usuário quanto ao desempenho do sistema no quesito usabilidade.

Os valores discrepantes apresentados nas avaliações dos itens Q6 e Q9 pertencem ao usuário 07. Para o item Q6, o usuário “concorda” com a afirmação de que o sistema apresenta inconsistências. No item Q9, o usuário coloca “não concorda e nem discorda” da afirmação de que se sentiu confiante ao usar o sistema. A insegurança de profissionais com pouca ou nenhuma experiência é relatada na literatura e pode justificar este último achado (SALAS et al., 2005).

De maneira geral, ao considerar as respostas para todos os itens da escala SUS, verifica-se que os profissionais julgaram que quanto à facilidade de aprendizagem, facilidade de memorização, satisfação do usuário e desempenho do sistema, o sistema apresenta uma

usabilidade “melhor alcançável”. Logo, o simulador DeglutSIM se mostrou adequado para o treinamento da classificação dos resíduos faríngeos na VED, porém, apresenta espaço para melhorias.

Para fins de comparação, o treinamento para classificar parâmetros da VED mais próximo ao que foi desenvolvido neste estudo foi realizado anteriormente utilizando a plataforma *Adobe Presenter*. Esta plataforma permite a criação de conteúdo interativo *online*, com recursos como sobreposições de áudio em slides e vídeos. Neste treinamento, os participantes receberam *feedback* imediato, tanto por escrito quanto auditivo, sobre a precisão de suas respostas, considerando parâmetros relativos ao fluxo do bolo alimentar na deglutição (resíduos faríngeos e penetração/aspiração) de uma escala não validada. Apesar do estudo anterior apresentar resultados positivos ao exibir resultados estatisticamente significantes ao comparar os treinamentos presencial e *online*, o *feedback* apresentado ao participante não era oportuno. Além disso, não foi considerada a medida de confiabilidade intra-avaliador, importante para garantir a consistência nas classificações, e não foi utilizado um modelo de decisão para a avaliação do treinamento (BRADY et al., 2018).

O DeglutSIM é um simulador computacional desenvolvido com o propósito de treinar profissionais que desejam se aprimorar e estudantes na classificação dos resíduos faríngeos visualizados na VED. Esta ferramenta possui como recurso o *feedback* em momento oportuno, ao final do treinamento, por meio de um relatório de desempenho. O usuário conta com a possibilidade de repetir o processo de treinamento quantas vezes forem necessárias, e todas as vezes que o treinamento for iniciado, novos casos serão apresentados de forma aleatória e randomizada dentro do sistema. No relatório de desempenho, o *feedback* é exibido ao usuário em formato de estrelas, o que traz características de jogos eletrônicos e pode motivar o usuário a obter pontuação máxima, aprimorando cada vez mais a habilidade perceptivo-visual de classificar os resíduos faríngeos.

O treinamento no DeglutSIM considera as medidas de precisão e confiabilidade intra-avaliador, importantes para garantir a consistência e a precisão das classificações dos resíduos faríngeos pelos profissionais durante esta avaliação diagnóstica. Isto garante maior segurança para os pacientes com disfagia, à medida que minimiza as chances de um diagnóstico errôneo neste sentido e recomendações inadequadas relacionadas a este diagnóstico funcional da função de deglutição.

Obter o cálculo das medidas de resultado ao final do treinamento facilitadas por um modelo de decisão inteligente permite o *feedback* automático e quanto à aquisição de

habilidades do usuário. Porém, como a aquisição de conhecimento pode ocorrer em curto, médio ou longo prazo, é recomendável investigar como ocorre esta construção do conhecimento ao longo do tempo (ANDRADE et al., 2022b). Estudos futuros devem ser desenvolvidos nesta direção.

4.4.3. Aspectos positivos, negativos e sugestões

Sobre as questões dissertativas contidas no questionário, emergiram duas categorias, a saber “O que foi bom” e “O que pode melhorar”, com duas subcategorias do sistema cada uma, relacionado ao “Método de Ensino/Avaliação” e à “Parte técnica do sistema”.

Na categoria “O que foi bom”, foram unidas as respostas individuais dos participantes para a pergunta “Quais foram os pontos positivos na sua experiência com o simulador?” e as observações individuais da pesquisadora no momento da simulação, de acordo com comentários feitos pelos participantes. Já para a categoria “O que pode melhorar”, foram unidas as respostas individuais dos participantes para as perguntas “Quais foram os pontos negativos na sua experiência com o simulador?” e “Você tem sugestões para melhorar o simulador?”, unidas com as observações individuais da pesquisadora (Apêndice 4).

Na categoria “O que foi bom”, subcategoria “Método de Ensino/Avaliação”, os profissionais expõem com propriedade as percepções relacionadas a aspectos de ensino/aprendizagem em saúde e os benefícios do uso do simulador para a prática clínica. Algumas destas percepções podem ser constatadas nos relatos seguintes:

Aprimoramento das habilidades de avaliação e monitoramento; mantém o aluno e o profissional atualizado; melhora a qualidade do trabalho profissional. (U03)

Treinar a classificação na VED, ampliando os conhecimentos para a prática. (U06)

Treinar o olhar clínico na interpretação do exame, correlacionando com os achados da avaliação clínica; Proporcionar maior segurança para a conduta diagnóstica e terapêutica. (U09)

O simulador permite com que o participante desenvolva um olhar apurado para imagens apresentadas, a medida em que são respondidos os questionamentos. (U11)

As percepções expressas nas frases supramencionadas corroboram com achados de um estudo realizado com simulação envolvendo a identificação da deglutição em imagens de VFD. Os resultados demonstraram que métodos de aprendizagem *online* da análise de parâmetros da deglutição demonstram resultados positivos e destacam como vantagem a possibilidade de

avancar no material conforme o ritmo individual do aluno (EDWARDS et al., 2024). Programa desenvolvido no mesmo sentido verificou que os participantes obtêm ganhos relativos ao conhecimento, habilidades e confiança ao final do treinamento através da utilização destes recursos (BURNS et al., 2021).

Outro estudo realizado com fonoaudiólogos gestores de serviços de saúde examinou as percepções destes profissionais sobre um programa *e-learning* de treinamento para a análise de imagens videofluoroscópicas e o impacto do treinamento para a prestação de serviços relacionados à VFD no local de trabalho. Foram listados pelos gestores os seguintes benefícios do programa: ampla aceitação devido ao design e conteúdo; aprimoramento das oportunidades de treinamento; melhora da habilidade clínica e fortalecimento da força de trabalho e eficiência para o serviço fonoaudiológico. Em suma, estes profissionais perceberam o treinamento no programa *e-learning* como um recurso valioso para suplementar o treinamento na VFD (TAUBERT et al., 2021).

Estes resultados voltados à sistemas para o estudo da deglutição apresentam benefícios claros para a aprendizagem, tal qual outros estudos com sistemas computacionais envolvendo o ensino em outros campos da saúde, como Enfermagem (ANTONELI et al., 2018), Farmacologia (BARROS, 2015) e Cirurgia (MOURA; MORAES; MACHADO, 2016; PAIVA et al., 2018).

Ainda referente à categoria “O que foi bom”, sobre a subcategoria “Parte técnica do sistema”, os profissionais manifestaram satisfação com o uso do simulador e os aspectos envolvidos no sistema. As percepções estão descritas a seguir:

Software intuitivo, fácil e didático. (U02)

Identificação das estruturas; Sistema intuitivo. (U05)

Os vídeos e as imagens reais, tornam o sistema mais intuitivo. (U07)

Fácil utilização; muito embasado cientificamente; ótimo para a prática profissional; bem intuitivo; visualmente agradável. (U08)

Nestes relatos, observa-se a recorrência do adjetivo “intuitivo” entre os quatro participantes que comentaram sobre as características do sistema. Essa observação, juntamente com os termos “fácil”, “didático”, “ótimo” e “agradável”, corrobora que a aplicação atingiu o propósito para o qual foi desenvolvida considerando o elemento “cenário”, conforme preconizado no processo de design do ST (definição de conteúdo), que buscou desenvolver

“Interfaces práticas, interativas, de fácil usabilidade” no início do estudo. Este aspecto ganha ainda mais relevância no contexto de um ST na área de saúde, onde há impacto direto na vida dos pacientes pela atuação dos profissionais (MACHADO; COSTA; MORAES, 2018).

Na categoria “O que pode melhorar”, a subcategoria relacionada ao “Método de Ensino/Avaliação” recebeu relatos sobre uma possível falta de instrução ao participante durante o uso do sistema:

Na avaliação de valéculas, não fica claro se devemos considerar o espaço total ou apenas um lado (Direito ou esquerdo) para a classificação. (U02)

Dificuldade para identificar resíduos nos vídeos com qualidade ruim. Também sentiu falta de especificar se deve considerar o resíduo no espaço total ou apenas um lado (Direito ou esquerdo). (U09)

É difícil saber se devemos considerar o espaço total ou unilateral. Não acha que devem ser incluídos os exames que estão com a qualidade da imagem ruim, pois considera difícil para quem está aprendendo. (U10)

Nas quatro avaliações recebidas sobre o que poderia melhorar, os participantes apontaram a falta de instrução sobre considerar a presença de resíduos faríngeos, seja em valéculas ou em seios piriformes, antes de serem colocados em situação de avaliação dentro do sistema. A escala escolhida para mensurar a gravidade dos resíduos no treinamento proposto neste estudo não dispõe informações sobre como proceder em exames em que o paciente apresenta resíduo apenas de um lado das estruturas laríngeas, como nos casos em que existe paralisia de prega vocal unilateral. Apesar disso, no estudo de validação da escala os autores apontam que a YPRSRS, que possui fácil administração e interpretação, é generalizável para a aplicação para todos os tipos de pacientes (NEUBAUER; RADEMAKER; LEDER, 2015). Durante o treinamento, os participantes receberam a orientação para considerar todo o espaço valcular/de seios piriformes, tal qual os especialistas procederam para avaliar os exames. Espera-se que os autores da escala utilizada venham esclarecer estes apontamentos em estudos futuros com foco no aprimoramento do instrumento. Vale destacar que mesmo com a limitação apresentada e outras, colocadas por outros autores em um estudo que propõe outra escala neste sentido (SABRY; COYLE; ABOU-ELSAAD, 2021), a YPRSRS demonstrou ter maior eficácia em comparação com três escalas de classificação dos resíduos faríngeos na VED, o que norteou a escolha para este trabalho (MESSINA et al., 2024).

Sobre o outro apontamento realizado por dois usuários, em relação à qualidade ruim de alguns vídeos apresentados para classificação dentro do simulador, este foi um quesito

verificado e discutido entre os especialistas ao avaliar os exames de referência. Em um estudo anterior, que validou a escala YPRSRS em sua versão para o idioma alemão, os autores apontaram como limitação do estudo a seleção de imagens por consenso. Segundo eles, criar uma seleção “melhor das melhores” pode influenciar positivamente na probabilidade de concordância entre os avaliadores, mas não fornece uma amostra sobre a prática clínica conforme a vida real, onde o resíduo pode se distribuir de forma desigual nas cavidades faríngeas e nem sempre uma visão suficientemente clara do exame pode ser fornecida (GERSCHKE et al., 2019).

Ainda referente à categoria “O que pode melhorar”, considerando a subcategoria “Parte técnica do sistema”, na percepção dos profissionais, puderam ser constatados alguns problemas de ordem técnica com o uso do ST. Os apontamentos mais frequentes e que merecem destaque são descritos a seguir:

(1) Quanto aos botões play/pause:

O botão play/pause poderia ser único. (U2)

Usuário apresentou dificuldade com os botões play/pause, ficou apertando ‘play’ na intenção de pausar. (U03)

Melhorar o sistema de “pausar” (*Usuário apresentou dificuldade com os botões play/pause, ficou apertando ‘play’ na intenção de pausar*). (U04)

Usuário apresentou dificuldade com os botões play/pause, ficou apertando ‘play’ na intenção de pausar. (U08)

(2) Quanto a disposição das âncoras na fase RFVS:

Seria interessante poder deixar as duas âncoras abertas ao mesmo tempo na fase em que avalia valéculas e seios piriformes juntos (*esse usuário praticamente não abriu as âncoras nesta fase, perceptível que ficou cansado de ficar “de um lado para outro” quando estava avaliando*). (U02)

Na fase que julga valécula e seios piriformes, tentar otimizar as âncoras para aparecerem simultaneamente (*Este usuário relatou que se confundiu e perdeu a atenção ao avaliar valéculas+seios piriformes, devido à tarefa de ter que fechar uma âncora para visualizar a outra a cada caso avaliado*). (U09)

Sugestão: Sinalizar necessidade de maior atenção quando é necessário avaliar valéculas e seios piriformes ao mesmo tempo. Colocar nas instruções que “Do lado laranja será valéculas e do lado roxo, seios piriformes”, com a âncora e botões de classificação de cada um do lado correspondente, ajudaria nesta última fase. Ou então, um tutorial em vídeo com cerca de 10 segundos mostrando como realizar a avaliação. (U10)

(3) Quanto ao progresso na avaliação:

O progresso não é medido em números, poderia ter uma barra de progresso numérica para que se possa acompanhar (ex.: 1/20, 2/20...). (U02)

Relatou sentir falta do avanço do progresso na Prática (tinha avaliado 3 exames e ainda não tinha avançado – só avança a cada 4 vídeos nesta etapa). (U08)

Teve a impressão de que a classificação dos vídeos não estava sendo salva na Prática, pois a barra de progresso não estava avançando (só avança após classificar 4 exames). (U10)

(4) Quanto à velocidade de reprodução do vídeo:

Interessante se pudesse alterar a velocidade do vídeo. (U08)

Incluir possibilidade de modificar a velocidade do vídeo – *slow motion*. (U10)

Verifica-se nestas observações aspectos que irão nortear as medidas para a execução de melhorias de ordem técnica no uso do sistema. A tecnologia envolvida no desenvolvimento de simuladores de treinamento computacionais está em constante aprimoramento no sentido de oferecer inovação no aprendizado em saúde. Destaca-se também o desenvolvimento de práticas processuais mais sofisticadas e simuladores avançados que representem com cada vez mais precisão cenários que favoreçam a aquisição de habilidades em saúde (SAKAKUSHEV et al., 2017).

Sobre o apontamento de nº4, a alteração da velocidade do vídeo para uma análise quadro-a-quadro, utilizando os recursos de câmera lenta (*slow motion*, do inglês) conforme necessidade é um recurso utilizado por alguns pesquisadores para a realização da classificação dos resíduos faríngeos na VED (CURTIS et al., 2022; PISEGNA et al., 2018, 2020). Devido a duração de uma deglutição completa ser curta, em torno de um segundo, algumas informações podem ser perdidas. Deste modo, a função *slow motion* é considerada importante para a análise de parâmetros da deglutição que podem ser vistos na VED, como os resíduos faríngeos. Este é um aprimoramento que deve ser considerado nas próximas fases do sistema.

Simuladores de treinamento computacionais necessitam de melhoria contínua para garantir que estas ferramentas estejam sempre atualizadas e adaptadas às necessidades específicas do usuário (MORATO et al., 2020). Desta forma, o levantamento desta e das demais observações guiará os próximos passos para aprimoramento do sistema.

5. CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo o desenvolvimento de um simulador de treinamento perceptivo-visual para a classificação dos resíduos faríngeos na VED que permitisse avaliar a habilidade perceptivo-visual do usuário na realização desta classificação diagnóstica. O simulador DeglutSIM, resultado principal deste trabalho, foi projetado para ensinar ao usuário a classificação dos resíduos faríngeos em regiões de valéculas e seios piriformes em diferentes níveis de gravidade, utilizando uma escala validada psicometricamente e amplamente utilizada (*Yale Pharyngeal Residue Severity Rating Scale*) (NEUBAUER; RADEMAKER; LEDER, 2015). Para a avaliação do treinamento no ambiente virtual, foram dispostos recortes de vídeos rotulados por três especialistas, que serviram como base para considerar o acerto das respostas do usuário. Ainda foi proposto um índice de desempenho a partir de medidas de confiabilidade intra-avaliador e precisão, para garantir respostas consistentes e confiáveis no treinamento da classificação diagnóstica dos resíduos faríngeos.

A pesquisa atingiu os cinco objetivos propostos. O primeiro, de identificar as características do treinamento, foi realizado a partir de uma revisão de escopo e uma revisão narrativa da literatura, que estabeleceram as bases para a estruturação do treinamento proposto. O segundo objetivo, de definir parâmetros para a avaliação do treinamento, foi atendido através do levantamento de requisitos com uma profissional fonoaudióloga e dos achados da literatura. Além disso, foram obtidos valores para as variáveis do estudo com três especialistas, resultando deste processo o desenvolvimento de um banco de dados de vídeos de VED com foco na classificação dos resíduos faríngeos (PhaReD-FEES): mais um produto desta tese e que será disponibilizado em um repositório de acesso aberto para a comunidade científica. O terceiro objetivo, propor um modelo de decisão inteligente para a avaliação do desempenho do usuário, foi alcançado com a implementação de um SBR que respondeu de forma satisfatória aos testes realizados. O quarto objetivo, de implementar o simulador DeglutSIM, foi cumprido ao desenvolver o simulador, integrar os parâmetros de avaliação definidos com o modelo de decisão e oferecer um treinamento eficaz e interativo para os usuários. Por fim, o quinto objetivo, de validar a usabilidade do simulador, foi atendido com os resultados positivos da análise de usabilidade dos usuários, evidenciados por uma pontuação SUS média de 92,7, classificado como “melhor imaginável”.

Considerando as tecnologias existentes para o treinamento e as vantagens do uso da simulação computacional no âmbito da saúde, a pesquisa apresenta o DeglutSIM como uma ferramenta inovadora para preparar estudantes e profissionais de saúde envolvidos no

diagnóstico em disfagia orofaríngea para realizarem uma conduta mais assertiva ao classificarem corretamente a gravidade dos resíduos faríngeos na VED. Além dos resultados positivos, algumas áreas foram sugeridas para melhorias, como a unificação dos botões *play/pause*, a possibilidade de visualizar simultaneamente as imagens âncoras da escala de resíduos faríngeos, a inclusão de uma barra de progresso numérica e a opção de alterar a velocidade de reprodução dos vídeos.

A observação de que a experiência clínica pode influenciar na habilidade perceptivo-visual de classificar os parâmetros da deglutição, observada em um estudo anterior com treinamento de imagens de VFD, e que parece ter refletido na avaliação do sistema proposto, destacam a importância de adaptar o treinamento para diferentes níveis de experiência clínica do usuário, como profissionais e estudantes de graduação/pós-graduação. Além disso, a presença de valores discrepantes na avaliação de itens específicos relacionados à satisfação do usuário e eficiência do simulador sugere a necessidade de refinamentos no conteúdo e nas instruções fornecidas durante o treinamento. O acréscimo de um módulo sobre anatomia endoscópica, por exemplo, poderia beneficiar profissionais e estudantes com pouca experiência clínica.

O argumento e hipótese de pesquisa desta tese, de que é possível proporcionar e avaliar o treinamento para a classificação dos resíduos faríngeos na VED por meio da simulação computacional foi confirmada. Espera-se que este trabalho abra novas oportunidades para pesquisas futuras, tanto no treinamento para a classificação de parâmetros da deglutição na VED quanto em estudos voltados à classificação destes parâmetros utilizando outros exames instrumentais da deglutição.

O DeglutSIM demonstrou ser uma ferramenta valiosa para o treinamento de profissionais da saúde, contribuindo para a melhoria das habilidades de diagnóstico na prática clínica. A implementação das sugestões de melhoria identificadas pode potencializar ainda mais a aceitação do sistema, garantindo um treinamento de alta qualidade, mais disponível e acessível financeiramente e, consequentemente, maior segurança para os pacientes com disfagia. Ressalta-se, ainda, o pioneirismo deste estudo ao propor um simulador de treinamento computacional desenvolvido especificamente para aprimorar a habilidade perceptivo-visual de classificação dos resíduos faríngeos na VED através de uma escala validada, com avaliação automática e *feedback* em tempo real por meio de um modelo de decisão inteligente.

6. REFERÊNCIAS

ABERNETHY, B. et al. Visual-Perceptual and Cognitive Differences Between Expert, Intermediate, and Novice Snooker Players. *Applied cognitive psychology*, v. 8, p. 185–206, 1994.

ALANAZI, A. A.; NICHOLSON, N.; THOMAS, S. The Use of Simulation Training to Improve Knowledge, Skills, and Confidence Among Healthcare Students: A Systematic Review. *The Internet Journal of Allied Health Sciences and Practice*. The Internet Journal of Allied Health Sciences and Practice, v. 15, n. 3, p. 2, 2017.

ALMEIDA, G. F. et al. Uso da simulação computacional para ensino e aprendizagem nos cursos de saúde. *Research, Society and Development*, v. 10, n. 15, p. e565101523522, 2021.

ALTMAN, K. W.; YU, G-P; SCHAEFER, S. D. Consequence of Dysphagia in the Hospitalized Patient Impact on Prognosis and Hospital Resources. *Archives of Otolaryngology- Head & Neck Surgery*, v. 136, n. 8, p. 784–789, 2010.

ANDRADE, J. R. et al. Can We Trust Virtual Simulators for Health Education? A Study on Evaluation and Indicators of Accuracy and Reliability. *Comunicações em Informática*, v. 6, n. 1, p. 5–8, 2022a.

ANDRADE, J. R. B. et al. Virtual simulations for health education: how are user skills assessed? *Revista Brasileira de Educação Médica*, v. 46, n. 4, 2022b.

ANTONELI, G. et al. O uso de simuladores no auxílio do ensino-aprendizagem na enfermagem. *Educação & Linguagem*, v. 21, n. 2, p. 25–42, 2018.

BAIJENS, L. W. J. et al. European society for swallowing disorders - European union geriatric medicine society white paper: Oropharyngeal dysphagia as a geriatric syndrome. *Clinical Interventions in Aging*, v. 11, p. 1403–1428, 7 out. 2016.

BANGOR, A.; KORTUM, P.; MILLER, J. Determining what individual SUS scores mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, v. 4, n. 3, p. 114-123, 2009.

BARROS, W. B. Z. G. O uso de simulações computacionais em Farmacologia. *Revista EDaPECI*, v. 15, n. 1, p. 74-87, 2015.

ZAIDI, N. N. M.; HOQUE, M. Application of e-Learning for Teaching Hadith in Higher Education Institutional Education in Malaysia: A Literature Review. *Journal of Quran Sunnah Education & Special Needs*, v. 3, n. 2, p.28-34, 2019.

BORGES-ANDRADE, J. E.; ABBAD, G. da S.; MOURÃO, L. *Treinamento, desenvolvimento e educação em organizações e trabalho*. Porto Alegre: Artmed, 2009.

BRADY, S. Use of Dysphagia Severity Scales During the Fiberoptic Endoscopic Exam of Swallowing: Treatment Decisions and Planning. *Perspectives on Swallowing and Swallowing Disorders (Dysphagia)*, v. 16, n. 2, p. 10-13, 2007.

BRADY, S.; DONZELLI, J. The modified barium swallow and the functional endoscopic evaluation of swallowing. *Otolaryngologic Clinics of North America*, v. 46, n. 6, p. 1009-1022, 2013.

BRADY, S. L. et al. Face-to-face versus online training for the interpretation of findings in the fiberoptic endoscopic exam of the swallow procedure. *Advances in Medical Education and Practice*, v. 9, p. 433-441, 2018.

BRANDÃO, C. F. S.; COLLARES, C. F.; MARIN, H. F. A simulação realística como ferramenta educacional para estudantes de medicina. *Sci Med*, v. 24, n. 2, p. 187-192, 2014.

BROOKE, J. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, v. 189, n. 194, p. 4-7, 1996.

BROOKE, J. SUS: A Retrospective. *Journal of usability studies*, v. 8, n. 2, 2013.

BURNS, C. L. et al. Speech–language therapists’ perceptions of an eLearning program to support training in videofluoroscopic swallow studies. *International Journal of Language & Communication Disorders*, v. 56, n. 2, p. 257–270, 2021.

CASTILLO, E.; GUTIÉRREZ, J. M.; HADI, A. S. *Expert systems and probabilistic network models*. New York: Springer, 1997.

COSTA, B. O. I. et al. Training to analyze functional parameters with fiberoptic endoscopic evaluation of swallowing: a scoping review. *Dysphagia*, v. 39, n. 2, p. 198-207, 2024.

COSTA, B. O. I. et al. Training for fiberoptic endoscopic evaluation of swallowing parameter analysis: a scoping review protocol. *Revista CEFAC*, v. 24, n. 1, p. e11021, 2022.

COSTA, T. K. de L. *Uma arquitetura para portais de serious games e ambientes virtuais com serviços de integração e acompanhamento de atividades*. 2017. Tese (Doutorado em Informática) – Universidade Federal da Paraíba, João Pessoa, 2017.

COSTA, T. K. L.; MACHADO, L.S.; MORAES, R.M. Inteligência artificial e sua aplicação em serious games para saúde. *RECIIS - Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, v. 8, n. 4, p. 525–539, 2014.

CURTIS, J. A. et al. Visual Analysis of Swallowing Efficiency and Safety (VASES): A Standardized Approach to Rating Pharyngeal Residue, Penetration, and Aspiration During FEES. *Dysphagia*, v. 37, n. 2, p. 417–435, 2022.

DIX, A. Human-computer interaction. In: LIU, L.; ÖZSU, M. T. (ed.). *Encyclopedia of database systems*. Boston, MA: Springer US, 2009. p. 1327–1331.

DZIEWAS, R. et al. Flexible endoscopic evaluation of swallowing (FEES) for neurogenic dysphagia: training curriculum of the German Society of Neurology and the German stroke society. *BMC Medical Education*, v. 16, n. 1, p. 1–9, 2016.

DZIEWAS, R. et al. European Society for Swallowing Disorders FEES accreditation program for neurogenic and geriatric oropharyngeal dysphagia. *Dysphagia*, v. 32, n. 6, p. 725–733, 2017.

EDWARDS, A. et al. Factors that influence development of speech pathology skills required for videofluoroscopic swallowing studies. *International Journal of Language & Communication Disorders*, v. 58, n. 5, p. 1645–1656, 2023.

EDWARDS, A. et al. Developing novice analysts' videofluoroscopic swallowing study skills in speech-language pathology: A randomised control trial comparing blended and online training approaches. *International Journal of Speech-Language Pathology*, v. 26, n. 2, p. 212–224, 2024.

FARAZI, M. R. et al. A 3D dynamic biomechanical swallowing model for training and diagnosis of dysphagia. In: IEEE International symposium on biomedical imaging, 12, 2015. *Proceedings of the IEEE 12th International Symposium on Biomedical Imaging*. [S.l.]: IEEE, 2015. p. 1385–1388.

FEENEY, M. E. F.; WILLIAMSON, I. P.; BISHOP, I. D. The role of institutional mechanisms in spatial data infrastructure development that supports decision-making. *Cartography*, v. 31, n. 2, p. 21–38, 2002.

FEHRING, R. Methods to validate nursing diagnoses. *Nursing Faculty Research and Publications*, v. 27, 1987.

FONSECA, R. J. R. M.; SILVA, P. J. S. P. S.; SILVA, R. R. Acordo inter-juízes: O caso do coeficiente kappa. *Laboratório de Psicologia*, v. 5, n. 1, p. 81–90, 2007.

FREHYWOT, S. et al. E-learning in medical education in resource constrained low-and middle-income countries. *Human Resources for Health*, v. 11, p. 1-15, 2013.

FROWEN, J. J.; COTTON, S. M.; PERRY, A. R. The stability, reliability, and validity of videofluoroscopy measures for patients with head and neck cancer. *Dysphagia*, v. 23, n. 4, p. 348–363, 2008.

FURKIM, A. M. Disfagia orofaríngea neurogênica. In: QUEIROZ, M. I. (org.). *Fundamentos em fonoaudiologia: aspectos clínicos da motricidade orofacial*. Rio de Janeiro: Guanabara Koogan, [s.d.], p. 121–132.

GASTELUM, A. et al. Building a three-dimensional model of the upper gastrointestinal tract for computer simulations of swallowing. *Medical & Biological Engineering and Computing*, v. 54, p. 525–534, 2016.

GERSCHE, M. et al. Validation of the German Version of the Yale Pharyngeal Residue Severity Rating Scale. *Dysphagia*, v. 34, p. 308–314, 2019.

HIND, J. A. et al. Comparison of trained clinician ratings with expert ratings of aspiration on videofluoroscopic images from a randomized clinical trial. *Dysphagia*, v. 24, p. 211–217, 2009.

HISS, S. G.; POSTMA, G. N. Fiberoptic Endoscopic Evaluation of Swallowing. *The Laryngoscope*, v. 113, n. 8, p. 1386–1393, 2003.

HOSMER, D.; LEMESHOW, S. *Applied logistic regression*. New York: Wiley, 2000.

HUANG, V. W.; JONES, C. B.; GOMEZ, E. D. State of the art of virtual reality simulation in anesthesia. *International Anesthesiology Clinics*, v. 58, n. 4, p. 31–35, 2020.

KANEOKA, A. S. et al. The Boston residue and clearance scale: Preliminary reliability and validity testing. *Folia Phoniatrica et Logopaedica*, v. 65, n. 6, p. 312–317, 2014.

KRUG, S. *Don't make me think: a common sense approach to web usability*. 2. ed. Berkeley: New Riders, 2013. v. 1.

KRUPINSKI, E. A. Current perspectives in medical image perception. *Attention, Perception & Psychophysics*, v. 72, n. 5, p. 1205–1217, 2010.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, v. 33, n. 1, p. 159–174, 1977.

LANGMORE, S. E. *Endoscopic evaluation and treatment of swallowing disorders*. New York: Thieme, 2001.

LANGMORE, S. E. History of Fiberoptic Endoscopic Evaluation of Swallowing for Evaluation and Management of Pharyngeal Dysphagia: Changes over the Years. *Dysphagia*, v. 32, n. 1, p. 27–38, 2017.

LANGMORE, S. E. et al. Tutorial on clinical practice for use of the fiberoptic endoscopic evaluation of swallowing procedure with adult populations: Part 1. *American Journal of Speech-Language Pathology*, v. 31, n. 1, p. 163–187, 2022.

LANGMORE, S. E.; SCHATZ, K.; OLSEN, N. Fiberoptic endoscopic examination of swallowing safety: a new procedure. *Dysphagia*, v. 2, n. 4, p. 216–219, 1988.

LENCHUS, J. D. End of the “see one, do one, teach one” era: the next generation of invasive bedside procedural instruction. *Journal of Osteopathic Medicine*, v. 110, n. 6, p. 340–346, 2010.

LIGÊZA, A. Principles of verification of rule-based systems. In: *Logical foundations for rule-based systems*. Berlin/Heidelberg: Springer-Verlag, 2006. v. 11, p. 191–198.

LIU, Q. et al. The effectiveness of blended learning in health professions: systematic review and meta-analysis. *Journal of Medical Internet research*, v. 18, n. 1, p. e2, 2016.

MACHADO, L. S.; MORAES, R. M. Intelligent decisionmaking in training based on virtual reality. *Computational Intelligence in Complex Decision Systems*, p. 85–123, 2010.

MACHADO, L. S.; COSTA, T. K. L.; MORAES, R. M. Multidisciplinaridade e o desenvolvimento de serious games e simuladores para educação em saúde. *Revista Observatório*, v. 4, n. 4, p. 149–172, 2018.

MARTIN-HARRIS, B. et al. MBS measurement tool for swallow impairment-MBSImp: establishing a standard. *Dysphagia*, v. 23, n. 4, p. 392–405, 2008.

MARTINS, A. I. et al. Assessment of ambient assisted living services in a living lab approach: a methodology based on ICF. In: 2nd INTERNATIONAL LIVING USABILITY LAB WORKSHOP ON AAL LATEST SOLUTIONS, TRENDS AND APPLICATIONS (AAL), 2012.

MASSIH, C. G. P. A. et al. Validação de simulador de emissões otoacústicas para o ensino em audiologia. *Arquivos de Ciências da Saúde da UNIPAR*, v. 27, n. 8, p. 4289–4306, 2023.

MESSINA, F. et al. Pharyngeal residue scoring in fiberoptic endoscopic evaluation of swallowing: reliability comparison and applicability among different scales. *Dysphagia*, p. 1-9, 2024.

MORAES, R. M.; FERREIRA, J. A.; MACHADO, L. S. A new bayesian network based on gaussian naive bayes with fuzzy parameters for training assessment in virtual simulators. *International Journal of Fuzzy Systems*, v. 23, n. 3, p. 849–861, 2021.

MORAES, R. M.; MACHADO, L. S. Multiple assessment for multiple users in virtual reality training environments. In: IBEROAMERICAN CONGRESS ON PATTERN RECOGNITION, 2007. *Proceedings [...]*. Berlin, Heidelberg: Springer, 2007.

MORAES, R. M.; MACHADO, L. S. Psychomotor skills assessment in medical training based on virtual reality using a weighted possibilistic approach. *Knowledge-Based Systems*, v. 70, p. 97–102, 2014.

MORAIS, A. M. de. *Planejamento e desenvolvimento de um serious game voltado ao ensino de saúde bucal em bebês*. 2011. Dissertação (Mestrado em Modelos de Decisão e Saúde) – Universidade Federal da Paraíba, João Pessoa, 2011.

MORATO, Y. C. et al. Análise do sistema de informação em imunizações do brasil sob a ótica das heurísticas de usabilidade. *Revista Cubana de Información en Ciencias de la Salud*, v. 31, n. 2, p. e1515, 2020.

MOURA, I. F. M. S.; MORAES, R. M.; MACHADO, L. DOS S. Assessment for surgical incision trajectories with SVM. *Revista Informática Teórica Aplicada*, v. 23, n. 1, p. 212–230, 2016.

NEUBAUER, P. D.; RADEMAKER, A. W.; LEDER, S. B. The Yale pharyngeal residue severity rating scale: an anatomically defined and image-based tool. *Dysphagia*, v. 30, n. 5, p. 521–528, 2015.

NOORDALLY, S. O. et al. A study to determine the correlation between clinical, fiberoptic endoscopic evaluation of swallowing and videofluoroscopic evaluations of swallowing after prolonged intubation. *Nutrition in Clinical Practice*, v. 26, n. 4, p. 457–462, 2011.

OLIVEIRA, A. C. C. DE. REANIME: Simulador de realidade virtual para avaliação de treinamento em reanimação neonatal. Dissertação—João Pessoa: Universidade Federal da Paraíba, 2020.

PAIVA, M. A. A. DE. Proposta de um simulador de treinamento para a avaliação perceptivo-auditiva da voz. Dissertação—João Pessoa: Universidade Federal da Paraíba, 2022.

PAIVA, M. A. A.; MACHADO, L. DOS S.; LOPES, L. W. Proposal of requirements for the development of a training simulator for the auditory-perceptual judgment of voice. *CODAS*, v. 35, n. 6, p. e20220209, 2023.

PAIVA, P. V. F. et al. SimCEC: A collaborative VR-based simulator for surgical teamwork education. *Computers in Entertainment*, v. 16, n. 2, p. 1–26, 2018.

PILZ, W. et al. Observers' Agreement on Measurements in Fiberoptic Endoscopic Evaluation of Swallowing. *Dysphagia*, v. 31, n. 2, p. 180–187, 2016.

PISEGNA, J. M. et al. Rethinking residue: determining the perceptual continuum of residue on fees to enable better measurement. *Dysphagia*, v. 33, n. 1, p. 100–108, 2018.

PISEGNA, J. M. et al. Residue ratings on FEES: trends for clinical application of residue measurement. *Dysphagia*, v. 35, n. 5, p. 834–842, 2020.

PISEGNA, J. M. What's the evidence? A commentary on FEES research. *Perspectives of the ASHA Special Interest Groups*, v. 7, n. 6, p. 1941–1959, 2022.

PISEGNA, J. M.; LANGMORE, S. E. Parameters of Instrumental Swallowing Evaluations: Describing a Diagnostic Dilemma. *Dysphagia*, v. 31, n. 3, p. 462–472, 2016.

PRIKLADNICKI, A.; SANTANA, M. G.; CARDOSO, M. C. Protocols and assessment procedures in fiberoptic endoscopic evaluation of swallowing: an updated systematic review. *Brazilian Journal of Otorhinolaryngology*, v. 88, n. 3, p. 445–470, 2022.

ROBINSON, H. F. Enhancing the learning and supervision framework for training in flexible endoscopic evaluation of swallowing. *Current Opinion in Otolaryngology and Head and Neck Surgery*, v. 29, n. 3, p. 204–212, 2021.

ROMMEL, N.; HAMDY, S. Oropharyngeal dysphagia: Manifestations and diagnosis. *Nature Reviews Gastroenterology and Hepatology*, v. 13, n. 1, p. 49–59, 2016.

SÁ, D. R. T. et al. Usabilidade das tecnologias biomédicas na Unidade de Terapia Intensiva e sua influência na assistência de enfermagem. *Global Academic Nursing Journal*, v.2, n. Sup. 3, p. e185-e185, 2022.

SABRY, A.; COYLE, J. L.; ABOU-ELSAAD, T. Mansoura fiberoptic endoscopic evaluation of swallowing residue rating scale (MFRRS): an anatomically based tool - a preliminary study. *Folia Phoniatrica et Logopaedica*, v. 73, n. 6, p. 478–490, 2021.

SAKAKUSHEV, B. E. et al. Striving for better medical education the simulation approach. *Folia Medica*, v. 59, n. 2, p. 123–131, 2017.

SALAS, E. et al. Using simulation-based training to improve patient safety: what does it take? *The Joint Commission Journal on Quality and Patient Safety*, v. 31, n. 7, p. 363–371, 2005.

SANTOS, A. D.; MACHADO, L. S. Realidade virtual aplicada ao ensino de medicina: taxonomia, desafios e resultados. In: *WORKSHOP DE REALIDADE VIRTUAL E AUMENTADA*, 2009. *Anais [...]*. 2009.

SANTOS, L. Á. *Eficácia e importância da avaliação clínica da deglutição*. Mestrado em Ciências Médicas—Ribeirão Preto: Universidade de São Paulo, 2015.

SCHINDLER, A. et al. Phoniaticians and otorhinolaryngologists approaching oropharyngeal dysphagia: an update on FEES. *European Archives of Oto-Rhino-Laryngology*, v. 279, n. 6, p. 2727–2742, 2022.

SCOTT, A. et al. A Study of Interrater Reliability when using videofluoroscopy as an assessment of swallowing. *Dysphagia*, v. 13, n. 4, p. 223–227, 1998.

SOUZA, F. H. B. et al. Virtual patient: development of a clinical cases simulator as a tool for medical education. In: *5th INTERNATIONAL CONFERENCE ON RESEARCH IN APPLIED SCIENCE*, 2022. *Proceedings [...]*. 2022.

SWAN, K. et al. Psychometric properties of visuoperceptual measures of videofluoroscopic and fibre-endoscopic evaluations of swallowing: a systematic review. *Dysphagia*, v. 34, n. 1, p. 2–33, 2019.

TANAKA, O. Y.; TAMAKI, E. M. O papel da avaliação para a tomada de decisão na gestão de serviços de saúde. *Ciência & Saúde Coletiva*, v. 17, p. 821–828, 2012.

TAUBERT, S. T. et al. Speech-language pathology managers' perceptions of a videofluoroscopic swallow study eLearning programme to support training and service delivery. *International Journal of Speech-Language Pathology*, v. 23, n. 1, p. 103–112, 2021.

TRICCO, A. C. et al. A scoping review on the conduct and reporting of scoping reviews. *BMC Medical Research Methodology*, v. 16, n. 1, p. 1–10, 2016.

TRICCO, A. C. et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, v. 169, n. 7, p. 467–473, 2018.

VENITE, R. S. et al. Translation and cross-cultural adaptation of the Yale Pharyngeal Residue Severity Rating Scale into Brazilian Portuguese. *Brazilian Journal of Otorhinolaryngology*, v. 90, p. 101470, 2024.

WARNECKE, T. et al. Towards a basic endoscopic evaluation of swallowing in acute stroke - Identification of salient findings by the inexperienced examiner. *BMC Medical Education*, v. 9, n. 1, p. 1–7, 2009.

YOSHIDA, M. et al. Safety and the effectiveness of a new education program for nurses to assess swallowing function using fiberoptic endoscopic evaluation of swallowing (FEES). *Japan Journal of Nursing Science*, v. 17, n. 2, 2020.

7. APÊNDICES

APÊNDICE 1 – Dados demográficos do usuário

Dados demográficos
Instituição:
Idade:
Sexo: [] Masculino [] Feminino
Graduação: [] Fonoaudiologia [] Medicina [] Outra
Pós-graduação Lato Sensu: [] Residência Multiprofissional <i>em curso</i> [] Residência Médica em Otorrinolaringologia <i>em curso</i> [] Residência Multiprofissional <i>concluída</i> [] Residência Médica em Otorrinolaringologia <i>concluída</i> [] Especialização em Disfagia [] Outra especialização: _____ [] Não possuo
Pós-graduação Stricto Sensu: [] Mestrado [] Doutorado [] Pós-doutorado [] Não possuo
Anos de atuação profissional com pacientes com disfagia:
Recebeu algum tipo de treinamento para classificar os resíduos faríngeos na videoendoscopia da deglutição? [] Sim [] Não

APÊNDICE 2 – Avaliação da usabilidade

Questionário SUS (System Usability Scale)					
	1 Discordo totalmente	2	3	4	5 Concordo totalmente
1. Acho que gostaria de usar esse sistema com frequência.					
2. Achei o sistema desnecessariamente complexo.					
3. Achei o sistema fácil de usar					
4. Acho que precisaria do apoio de um técnico para poder utilizar este sistema.					
5. Acho que as várias funções do sistema estão bem integradas.					
6. Acho que o sistema apresenta muitas inconsistências.					
7. Imagino que a maioria das pessoas aprenderia a usar este sistema muito rapidamente.					
8. Achei o sistema muito complicado de usar.					
9. Eu me senti muito confiante ao usar o sistema.					
10. Eu precisava aprender muitas coisas antes de começar a usar esse sistema.					

11. Na sua opinião, quais foram os pontos positivos da sua experiência no simulador?

12. Na sua opinião, quais foram os pontos negativos da sua experiência no simulador?

13. Você tem sugestões para melhorar o simulador?

APÊNDICE 3 – Termo de consentimento livre e esclarecido

Você está sendo convidado(a) a participar da pesquisa intitulada “Avaliando usabilidade de um simulador para treinamento da classificação dos resíduos faríngeos na videoendoscopia da deglutição”, coordenada por Bianca Oliveira Ismael da Costa. O objetivo deste estudo é verificar a satisfação do usuário quanto à usabilidade de um simulador computacional para treinamento perceptivo-visual da classificação dos resíduos faríngeos na videoendoscopia da deglutição (VED).

Caso aceite participar, você terá que contribuir respondendo dois questionários, um contendo dados pessoais e sobre sua formação, e outro sobre a sua satisfação relacionada ao uso do simulador, o que deve dispendar cerca de 30 minutos.

RISCOS E BENEFÍCIOS

Com sua participação nesta pesquisa, você pode se sentir intimidado ou desconfortável em responder as questões contidas nos questionários. Caso isto venha a ocorrer, a coleta de dados poderá ser interrompida a qualquer momento, sem que isso lhe acarrete nenhum ônus ou prejuízo. Você também receberá apoio sobre a questão apresentada, sob a responsabilidade do (a) pesquisador (a) responsável. Esta pesquisa tem como benefícios a obtenção de dados para a validação de um simulador computacional direcionado ao treinamento da classificação dos resíduos faríngeos no exame de VED, o que contribuirá para que os profissionais de saúde envolvidos no diagnóstico de distúrbios da deglutição desenvolvam habilidade clínica para realizar uma conduta e tomada de decisão clínica mais assertivas.

SIGILO, ANONIMATO E PRIVACIDADE

O material e informações obtidas podem ser publicados em aulas, congressos, eventos científicos, palestras ou periódicos científicos, sem sua identificação. Os pesquisadores se responsabilizam pela guarda e confidencialidade dos dados, bem como a não exposição individualizada dos dados da pesquisa. Sua participação é voluntária e você terá a liberdade de se recusar a responder quaisquer questões que lhe ocasionem constrangimento de alguma natureza.

AUTONOMIA

Você também poderá desistir da pesquisa a qualquer momento, sem que a recusa ou a desistência lhe acarrete qualquer prejuízo. É assegurada a assistência durante toda a pesquisa, e garantido o livre acesso a todas as informações e esclarecimentos adicionais sobre o estudo e suas consequências. Se com a sua participação na pesquisa for detectado que você apresenta alguma condição que precise de tratamento, você receberá orientação da equipe de pesquisa, de forma a receber um atendimento especializado. Você também poderá entrar em contato com os pesquisadores, em qualquer etapa da pesquisa, por e-mail ou telefone, a partir dos contatos dos pesquisadores que constam no final do documento.

DEVOLUTIVA DOS RESULTADOS

Os resultados da pesquisa poderão ser solicitados a partir de janeiro de 2025, através dos e-mails disponibilizados ao final deste documento. Ressalta-se que os dados coletados nesta pesquisa – seja informações de prontuários, gravação de imagem, voz, audiovisual ou material biológico – somente poderão ser utilizados para as finalidades da presente pesquisa, sendo que para novos objetivos um novo TCLE deve ser aplicado.

RESSARCIMENTO E INDENIZAÇÃO

Lembramos que sua participação é voluntária, o que significa que você não poderá ser pago, de nenhuma maneira, por participar desta pesquisa. De igual forma, a participação na pesquisa não implica em gastos a você. No entanto, caso você tenha alguma despesa decorrente da sua participação, tais como transporte, alimentação, entre outros, você será ressarcido do valor gasto. Se ocorrer algum dano decorrente da sua participação na pesquisa, você será indenizado, conforme determina a lei.

Após ser esclarecido sobre as informações da pesquisa, no caso de aceitar fazer parte do estudo, assine o consentimento de participação em todas as páginas e no campo previsto para o seu nome, que é impresso em duas vias, sendo que uma via ficará em posse do pesquisador responsável e a outra via com você.

Pesquisador (a) responsável (orientador (a)): Liliane dos Santos Machado

E-mail para contato:

Telefone para contato:

Assinatura do (a) pesquisador (a) responsável: _____

Outros pesquisadores:

Nome: Bianca Oliveira Ismael da Costa

E-mail para contato:

Telefone para contato:

Assinatura do (a) aluno (a) pesquisador (a): _____

CONSENTIMENTO DE PARTICIPAÇÃO

Declaro que fui devidamente esclarecido sobre a pesquisa, riscos e benefícios e concordo em participar, voluntariamente da pesquisa intitulada “Avaliando usabilidade de um simulador para treinamento da classificação dos resíduos faríngeos na videoendoscopia da deglutição” conforme informações contidas neste TCLE.

Local e data: _____

Assinatura: _____

O Comitê de Ética em Pesquisa em Seres Humanos (CEP) é composto por um grupo de pessoas que estão trabalhando para garantir que seus direitos como participante sejam respeitados, sempre se pautando pelas Resoluções 466/12 e 510/16 do Conselho Nacional de Saúde (CNS). O CEP tem a obrigação de avaliar se a pesquisa foi planejada e se está sendo executada de forma ética. Caso você achar que a pesquisa não está sendo realizada da forma como você imaginou ou que está sendo prejudicado de alguma forma, você pode entrar em contato com o Contato do pesquisador responsável ou com o Comitê de Ética do Centro de Ciências Médicas

Endereço:- Centro de Ciências Médicas, 3º andar, Sala 14, Campus I - Cidade Universitária

- Bairro Castelo Branco CEP: 58059-900 - João Pessoa-PB Telefone: (083) 3216-7308

E-mail: comitedeetica@ccm.ufpb.br

APÊNDICE 4 – Categorização das respostas qualitativas dos profissionais sobre a experiência no simulador DeglutSIM

<i>O que foi bom (pontos positivos)</i>	
<i>Método de Ensino/Avaliação</i>	<i>Parte técnica do sistema</i>
Aprimoramento das habilidades de avaliação e monitoramento; mantém o aluno e o profissional atualizado; melhora a qualidade do trabalho profissional. (U03) A ferramenta auxilia os profissionais na classificação das estases e é autoexplicativa, o que facilita o uso. (U04) Treinar a classificação na VED, ampliando os conhecimentos para a prática. (U06) O tutorial expando os níveis de resíduo com cada local considerado no procedimento é um ponto positivo. (U07) Adorei, estou me sentindo positivamente desafiada! (<i>durante o uso do simulador</i>). (U08) Treinar o olhar clínico na interpretação do exame, correlacionando com os achados da avaliação clínica; Proporcionar maior segurança para a conduta diagnóstica e terapêutica. (U09) Possibilidade de ver vários casos clínicos e refletir sobre os achados. (U10) O simulador permite com que o participante desenvolva um olhar apurado para imagens apresentadas, a medida em que são respondidos os questionamentos. (U11)	<i>Software</i> intuitivo, fácil e didático. (U02) Identificação das estruturas; Sistema intuitivo. (U05) Os vídeos e as imagens reais, tornam o sistema mais intuitivo. (U07) Fácil utilização; Muito embasado cientificamente; Ótimo para a prática profissional; Bem intuitivo; Visualmente agradável. (U08)

<i>O que pode melhorar (pontos negativos/sugestões)</i>	
<i>Método de Ensino/Avaliação</i>	<i>Parte técnica do sistema</i>
- Na avaliação de valéculas, não fica claro se devemos considerar o espaço total ou apenas um lado (D/E) para a classificação. (U02) - Dificuldade para identificar resíduos nos vídeos com qualidade ruim de nitidez; <i>Sentiu falta de especificar se deve considerar o resíduo no espaço total ou apenas um lado (D/E).</i> (U09) - É difícil saber se devemos considerar o espaço total ou unilateral; - Não acho que devem ser incluídos exames que estão com a qualidade da imagem ruim, pois considero difícil para quem está aprendendo. (U10)	- Poderia ter a possibilidade de colocar os vídeos em tela cheia ou em tamanho maior, para aproveitar os espaços vazios na tela; - Seria interessante poder deixar as duas âncoras abertas ao mesmo tempo na fase em que avalia valéculas e seios piriformes juntos (<i>esse usuário praticamente não abriu as âncoras nesta fase, perceptível que ficou cansado de ficar “de um lado para outro” quando estava avaliando</i>); - O progresso não é medido em números, poderia ter uma barra de progresso numérica para que se possa acompanhar (ex.: 1/20, 2/20...); - Poderia ter aviso sobre os exames com a classificação errada no nivelamento; - O botão play/pause poderia ser único. (U2) - Os erros e acertos não são identificados, dificultando o <i>feedback</i> do treinamento. Poderiam ser indicados os erros e acertos individualmente (caso a caso) depois da última etapa (ex.: exame 1, classificação 5, marcou 4, errou; exame 2, classificação 3, marcou 3, acertou...). <i>Usuário apresentou dificuldade com os botões play/pause, ficou apertando ‘play’ na intenção de pausar.</i> (U03) - O sistema de pausar os vídeos para analisar melhor as estases. - Deixar exposto na tela a referência da classificação (as âncoras); - Melhorar o sistema de “pausar” (<i>Usuário apresentou dificuldade com os botões play/pause, ficou apertando ‘play’ na intenção de pausar</i>). (U04) - Difícil de passar na fase de nivelamento, gerando desmotivação. Sugestão de identificar a questão errada ao final da tentativa, caso a pessoa não passe na fase do nivelamento. (U06)

	• A “obrigatoriedade” de acertar todos os níveis de resíduo na fase de nivelamento, seja em valéculas, seja em seios piriformes ou em ambos concomitantemente. • A ausência de uma indicação específica sobre qual o vídeo que foi marcado errado no nivelamento. (U07) • Interessante se pudesse alterar a velocidade do vídeo; • Acrescentar um botão de “continuar” (<i>na parte que mostra as etapas ensino, nivelamento, prática</i>), pois a opção do botão de “voltar” dá impressão de querer desistir do treinamento. • <i>Usuário apresentou dificuldade com os botões play/pause, ficou apertando ‘play’ na intenção de pausar;</i> • <i>Relatou sentir falta do avanço do progresso na Prática (tinha avaliado 3 exames e ainda não tinha avançado – só avança a cada 4 vídeos nesta etapa).</i> (U08) • Dificuldade para avaliar de forma simultânea com as duas âncoras abertas; • O tempo de duração das etapas, em especial as etapas isoladas de valéculas e seios piriformes (<i>teve a impressão que nestas duas fases iniciais tinham mais exames para serem avaliados do que na última</i>); • Diminuir o tempo das etapas/quantidade de vídeos; • Ter a opção para dar o zoom nas imagens menos nítidas e mais “distantes”; • Na fase que julga valécula e seios piriformes, tentar otimizar as âncoras para aparecerem simultaneamente (<i>Este usuário relatou que se confundiu e perdeu a atenção ao avaliar valéculas+seios piriformes, devido à complexidade da tarefa de ter que fechar uma âncora para visualizar a outra a cada caso avaliado</i>). (U09) • Concluir todas as etapas de uma vez, sem ter que passar por um nivelamento; • Sinalizar necessidade de maior atenção quando é necessário avaliar valéculas e seios piriformes ao mesmo tempo. Colocar nas instruções que “Do lado laranja será valéculas e do lado roxo, seios piriformes”, com a âncora e botões de classificação de cada um do lado correspondente, ajudaria nesta última fase. Ou então, um
--	---

	tutorial em vídeo com cerca de 10 segundos mostrando como realizar a avaliação. • Incluir possibilidade de modificar a velocidade do vídeo – <i>slow motion</i>; • <i>Teve a impressão de que a classificação dos vídeos não estava sendo salva na Prática, pois a barra de progresso não estava avançando (só avança após classificar 4 exames).</i> (U10)
--	---

8. ANEXOS

ANEXO 1 – Registro do programa de computador DeglutSIM junto ao Instituto Nacional de Propriedade Industrial (INPI)



REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS

Certificado de Registro de Programa de Computador

Processo Nº: BR512024004449-3

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de 01/08/2024, em conformidade com o §2º, art. 2º da Lei 9.609, de 19 de Fevereiro de 1998.

Título: DeglutSIM

Data de criação: 01/08/2024

Titular(es): UNIVERSIDADE FEDERAL DA PARAIBA

Autor(es): LILIANE DOS SANTOS MACHADO; JOSÉ RAUL DE BRITO ANDRADE; LEANDRO DE ARAÚJO PERNAMBUCO; BIANCA OLIVEIRA ISMAEL DA COSTA; DESIRÉ DOMINIQUE DINIZ DE MAGALHÃES

Linguagem: C#; UNITY

Campo de aplicação: ED-04; SD-09

Tipo de programa: AP-01; DS-04; SM-02

Algoritmo hash: SHA-512

Resumo digital hash:
8a79654ceaa4b2e4aa5f2f6f1a9d21e099c5c15593b3e7b0c0784f641a633b01838c97630d9051f72b440a4527bfd0bafbe
8bd6e643a8cd8e3633f0ff1ab439b

Expedido em: 26/11/2024

Aprovado por:
Carlos Alexandre Fernandes Silva
Chefe da DIPTO

ANEXO 2 – Parecer de aprovação do comitê de ética e pesquisa

UFPB - CENTRO DE CIÊNCIAS
MÉDICAS DA UNIVERSIDADE
FEDERAL DA PARAÍBA / CCM



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: AVALIANDO USABILIDADE DE UM SIMULADOR DE TREINAMENTO PARA CLASSIFICAR RESÍDUOS FARÍNGEOS NA VIDEOENDOSCOPIA DA DEGLUTIÇÃO

Pesquisador: BIANCA OLIVEIRA ISMAEL DA COSTA

Área Temática:

Versão: 2

CAAE: 79036724.0.0000.8069

Instituição Proponente: UNIVERSIDADE FEDERAL DA PARAIBA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 6.879.883

Apresentação do Projeto:

Intervenção/Experimental.

QP1: Um simulador computacional para treinamento da habilidade perceptivo-visual de classificar dos resíduos faríngeos na VED consegue obter resultados positivos em relação ao grau de satisfação do usuário no quesito usabilidade do sistema? H1-0: Um simulador computacional para treinamento da habilidade perceptivo-visual de classificar dos resíduos faríngeos na VED é capaz de obter resultados positivos quanto à usabilidade do sistema na percepção do usuário. H1-1: Um simulador computacional para treinamento da habilidade perceptivo-visual de classificar dos resíduos faríngeos na VED não é capaz de trazer resultados positivos quanto à usabilidade do sistema na percepção do usuário. QP2: As características sociodemográficas e a experiência do usuário no uso de dispositivos tecnológicos são fatores que influenciam na satisfação do usuário quanto à usabilidade do sistema? H2-0: As características sociodemográficas e a experiência do usuário no uso de dispositivos tecnológicos são fatores que influenciam na satisfação do usuário quanto à usabilidade do sistema. H2-1: As características sociodemográficas e a experiência do usuário no uso de dispositivos tecnológicos são fatores que não influenciam na satisfação do usuário quanto à usabilidade do sistema.

4.1. Delineamento do estudo: Trata-se de uma pesquisa exploratória e de abordagem quantitativa. 4.2.

População do estudo: A população do estudo será constituída por

Endereço: Centro de Ciências Médicas, 3º andar, Sala 14 - Cidade Universitária Campus 1

Bairro: CASTELO BRANCO **CEP:** 58.051-900

UF: PB **Município:** JOAO PESSOA

Telefone: (83)3218-7308

E-mail: comitedeetica@ccm.ufpb.br

UFPB - CENTRO DE CIÊNCIAS
MÉDICAS DA UNIVERSIDADE
FEDERAL DA PARAÍBA / CCM



Continuação do Parecer: 6.879.883

Justificativa de Ausência	TCLE.pdf	21/03/2024 11:35:09	BIANCA OLIVEIRA ISMAEL DA COSTA	Aceito
Folha de Rosto	FolhadeRosto.pdf	21/03/2024 11:32:19	BIANCA OLIVEIRA ISMAEL DA COSTA	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

JOAO PESSOA, 11 de Junho de 2024

Assinado por:

MARCIA ADRIANA DIAS MEIRELLES MOREIRA
(Coordenador(a))

Endereço: Centro de Ciências Médicas, 3º andar, Sala 14 - Cidade Universitária Campus 1
Bairro: CASTELO BRANCO CEP: 58.051-900
UF: PB Município: JOAO PESSOA
Telefone: (83)3216-7308 E-mail: comitedeetica@ccm.ufpb.br