

Avaliação de um assistente virtual inteligente (chatbot) no procedimento de ressuscitação cardiopulmonar

Evaluation of an intelligent virtual assistant (chatbot) in cardiopulmonary resuscitation

Evaluación de un asistente virtual inteligente (chatbot) en reanimación cardiopulmonar

B.el. Lucas Andrada Carrazzoni Goes¹, B.el. João Vitor de Amorim Silva¹, Dr. Amadeu Sá de Campos Filho¹

¹ Universidade Federal de Pernambuco, Centro de Ciências Médicas, Recife (PE), Brasil.

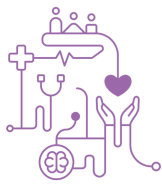
Autor correspondente: Dr. Amadeu Sá de Campos Filho
E-mail: amadeu.campos@ufpe.br

Resumo

Objetivo: Avaliar um chatbot para auxiliar na aprendizagem no Procedimento de Ressuscitação Cardiopulmonar (RCP). Métodos: Estudo transversal, prospectivo, analítico e observacional. A validação ocorreu mediante avaliação por especialistas, utilizando-se o Índice de Validação de Conteúdo (IVC) e o método de Delphi. A validação pelo usuário final ocorreu através das avaliações de usabilidade, que utilizou o instrumento Single Ease Question, o protocolo System Usability Scale e experiência do usuário, na qual foi utilizado o User Experience Questionnaire. Resultados: Os especialistas validaram o conteúdo do chatbot por notas atribuídas a alguns quesitos, com a média do IVC de 0,942. A eficiência relativa geral calculada foi 90,80% e a satisfação alcançou resultados excelentes, com pontuação de 80,875 no SUS. Conclusões: Houve elevado grau de concordância entre os profissionais, bom incremento no conhecimento dos usuários sobre RCP e poucas dificuldades no uso do chatbot, apesar de fatores limitantes.

Descritores: Ressuscitação Cardiopulmonar; Inteligência Artificial; Educação em saúde.

Abstract



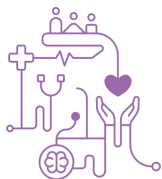
Objective: Evaluate a chatbot to assist in learning the Cardiopulmonary Resuscitation Procedure (CPR). **Methods:** Cross-sectional, prospective, analytical and observational study. Validation occurred through evaluation by experts, using the Content Validation Index (CVI) and the Delphi method. Validation by the end user occurred through usability evaluations, which used the Single Ease Question instrument, the System Usability Scale protocol and user experience, in which the User Experience Questionnaire was used. **Results:** The experts validated the chatbot's content using scores assigned to some questions, with an average CVI of 0.942. The calculated overall relative efficiency was 90.80% and satisfaction achieved excellent results, with a SUS score of 80.875. **Conclusions:** There was a high degree of agreement among professionals, a good increase in users' knowledge about CPR and few difficulties in using the chatbot, despite limiting factors.

Keywords: Cardiopulmonary Resuscitation; Artificial Intelligence; Health Education.

Resumen

Objetivo: Evaluar un chatbot para ayudar en el aprendizaje del Procedimiento de Reanimación Cardiopulmonar (RCP). **Métodos:** Estudio transversal, prospectivo, analítico y observacional. La validación se produjo mediante evaluación por parte de expertos, utilizando el Índice de Validación de Contenido (IVC) y el método Delphi. La validación por parte del usuario final ocurrió a través de evaluaciones de usabilidad, con el instrumento Single Ease Question, el protocolo System Usability Scale (SUS) y la experiencia de usuario, con el User Experience Questionnaire. **Resultados:** Los expertos validaron el contenido del chatbot mediante puntuaciones asignadas a algunas preguntas, con un IVC promedio de 0,942. La eficiencia relativa global fue del 90,80% y la satisfacción alcanzó excelentes resultados, con una puntuación SUS de 80,875. **Conclusiones:** Hubo un alto grado de acuerdo entre los profesionales, un buen aumento del conocimiento de los usuarios sobre RCP y pocas dificultades con el chatbot, apesar de los factores limitantes.

Descriptorios: Reanimación cardiopulmonar; Inteligencia artificial; Educación para la salud



Introdução

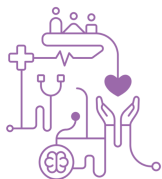
A evolução das tecnologias e o maior acesso à internet entre a população e a sua adoção para as diversas áreas é apontada como uma importante forma de disseminação da informação. Na área da saúde, diversas pesquisas elucidam a aplicabilidade das tecnologias na aquisição de informações por usuários de diferentes faixas etárias, apontando benefícios como a promoção de um ambiente interativo e a possibilidade de promoção à saúde ⁽¹⁾.

Entre as novas tecnologias, há os sistemas de autoatendimento baseados em conversação, métodos e técnicas de inteligência artificial (chatbots), que oferecem informações em tempo real e permitem uma interação com o conteúdo para a obtenção de apoio no processo de aprendizado, propiciando a aquisição de conhecimentos e o crescimento com a experiência do aprendizado ⁽²⁾.

Na área de saúde, seu uso ainda é incipiente, sendo mais relacionado ao cadastro, agendamento de exames, consultas e triagem. Um exemplo de chatbot já testado e utilizado atualmente é o Coronabot, tendo aplicação na triagem de pacientes com suspeitas de infecção por coronavírus ⁽³⁾ e outro exemplo é um chatbot para o auxílio na perda e manutenção de peso ⁽⁴⁾. Contudo, no âmbito da prática de manobras de saúde, especificamente nas ações para realizar a ressuscitação cardiopulmonar, a utilização de chatbot ainda é restrita.

Quando se trata do atendimento pré-hospitalar em casos de urgência e emergência, os primeiros minutos após um acidente são determinantes para o prognóstico da vítima, período chamado de “Hora de Ouro” pelo Dr. R. Adams Cowley⁽⁵⁾. Assim, é importante que a população leiga possua conhecimentos básicos sobre o assunto para poder iniciar o procedimento de forma precoce, antes da chegada do serviço de emergência especializado, diminuindo possíveis sequelas e aumentando a sobrevivência do acidentado.

Diante desse cenário, esse artigo tem como objetivo avaliar o uso de um chatbot desenvolvido para o ensino do procedimento de ressuscitação cardiopulmonar com usuário sem conhecimentos prévios sobre o assunto e que, no final, se sinta mais



confiante perante uma situação de emergência e que tome condutas baseadas no conhecimento adquirido.

Método

Desenho do estudo

Estudo transversal, prospectivo, analítico, observacional com abordagem quali-quantitativa.

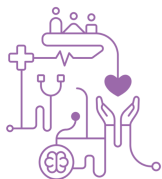
Amostra

A validação do conteúdo foi realizada com 9 profissionais de saúde com experiência em procedimentos de RCP e a validação de usabilidade e experiência do usuário foi realizada com 20 usuários finais que iriam utilizar o chatbot e que não tinham conhecimento sobre procedimentos de RCP

Validação do Chatbot com especialistas e com o usuário final

A validação do conteúdo utilizado pelo Chatbot ocorreu, como definido por Pasquali⁽⁶⁾, mediante a avaliação por especialistas da parte teórica. No âmbito da técnica utilizada para se obter o consenso entre os especialistas, utilizou-se o método Delphi. Nessa técnica, para o conteúdo ser validado precisa obter-se o consenso entre os especialistas, com um índice de validade de conteúdo (IVC) superior a 0,8, sobre o conteúdo empregado no projeto. Após essa etapa foi realizada a avaliação de usabilidade com o usuário final através do teste de Usabilidade e Experiência do Usuário.

Segundo a ISO 9241⁽⁸⁾ a usabilidade é uma medida na qual um produto pode ser usado por usuários específicos para alcançar objetivos específicos com eficácia, eficiência e satisfação em um contexto específico de uso. Dentro daqueles propósitos, temos a eficácia, que compreende a acurácia com a qual os usuários do sistema conseguiram atingir determinadas metas, estabelecendo-se pontuações de 0 (não conseguiu atingir a meta) à 1 (conseguiu atingir a meta). Nesse sentido, determinamos 3 tarefas que deveriam ser alcançadas pelos usuários do chatbot: 1) Saber o que fazer se a vítima não



estiver respirando; 2) Saber quantas compressões por ciclo e quantos ciclos são necessários durante a reanimação cardiopulmonar e 3) Saber como a vítima deve ser posicionada caso ela volta a respirar normalmente. Além disso, temos também a eficiência como um outro propósito incluído nas medidas de usabilidade de um sistema. A eficiência de acordo com a ISO⁽⁷⁾ é calculada a partir do tempo que se levou para a realização de cada uma das tarefas.

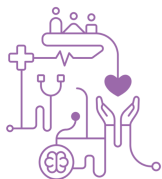
Além da eficácia e eficiência, temos a satisfação como um outro objetivo que deve ser incluído nos propósitos da usabilidade. Assim, a satisfação pode ser medida a partir de questionários de satisfação padronizados, os quais devem ser respondidos pelos usuários do programa em questão logo após o uso do chatbot, sendo que, aqui, teremos 2 aspectos de satisfação a serem analisados: a nível de cada tarefa (dificuldade) e a nível geral do teste (SUS).

Em relação à satisfação a nível de cada tarefa, utilizou-se o questionário *Single Ease Question* (SEQ), o qual leva em conta o quão fácil ou difícil foi concluir aquela tarefa para o usuário, podendo atribuir uma pontuação que vai de 1 (muito difícil) a 7 (muito fácil).

Instrumentos

Em relação a validação do chatbot com o público-alvo foram utilizados os seguintes instrumentos: questionário de usabilidade e de experiência do usuário. O questionário de usabilidade foi baseado nas métricas de usabilidade da ISO⁽⁷⁾ que é composta pela eficácia, eficiência e satisfação. O questionário usado para medir a satisfação de cada tarefa individualmente foi o instrumento *Single Ease Question*, o qual leva em conta o quão fácil ou difícil foi concluir aquela tarefa para o usuário, podendo ele atribuir uma pontuação que vai de 1 (muito difícil) a 7 (muito fácil) e o questionário utilizado para medir a satisfação geral foi protocolo *System Usability Scale* (SUS) composto por 10 questões onde cada usuário responderia numa escala de 1 a 5, sendo o número mais baixo equivalente a “discordo totalmente” e o mais alto equivalente a “concordo totalmente”

Para a avaliação de experiência do usuário foi utilizado o questionário resumido do protocolo *User Experience Questionnaire* (UEQ) que permite uma avaliação rápida feita



pelos usuários finais cobrindo uma impressão preferencialmente abrangente da experiência do usuário e avalia as qualidades pragmáticas e hedônicas de um produto ⁽⁸⁾. Cada item do UEQ consiste em um par de termos com significados opostos e os usuários classificam cada item em uma escala *Likert* de 7 pontos. As respostas são escaladas de -3 (concordo totalmente com o termo negativo) a +3 (concordo totalmente com o termo positivo).

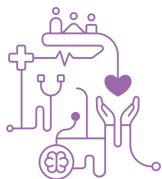
Coleta de Dados

Devido a restrição de contato presencial ocasionado pela pandemia do coronavírus, a coleta de dados da avaliação foi realizada remotamente por meio do instrumento citado acima enviado por e-mail para os especialistas e através de encontros remotos na plataforma do google meet para a validação do chatbot com os usuários finais.

Análise de dados

A análise dos dados da avaliação do conteúdo foi baseada no método de Índice de Validade de Conteúdo (IVC), o qual, segundo os autores, mede a proporção ou porcentagem de juízes que estão em concordância sobre determinados aspectos do instrumento e de seus itens ⁽⁹⁾. Ele permite, inicialmente, analisar cada item individualmente e depois o instrumento como um todo. O IVC utiliza uma escala do tipo *Likert* com pontuação de um a quatro de forma que as pontuações três e quatro foram consideradas representativas, sendo: 1 – Discordo totalmente; 2 – Discordo parcialmente; 3 – Neutro; 4 – Concordo parcialmente; 5 – Concordo totalmente.

Nessa validação, o escore do IVC é calculado por meio das somas de concordância dos itens que foram marcados por “4” ou “5” pelos especialistas divididos pelo número total de respostas. Os itens que receberam pontuação “1” ou “2” devem ser revisados ou eliminados. Com a participação de cinco ou menos sujeitos, todos devem concordar para ser representativo. No caso de seis ou mais, recomenda-se uma taxa não inferior a 0,78 ⁽¹⁰⁾.



Para verificar a validade de novos instrumentos de uma forma geral, alguns autores sugerem uma concordância mínima de 0,80 ⁽¹¹⁾. No entanto, neste caso, adotamos a concordância maior ou igual a 0,80. Em razão do grau de concordância adotado, não houve a necessidade de uma segunda rodada de avaliação.

Para a análise de Usabilidade e de Experiência do Usuário cada instrumento tinha seu cálculo pré-definido e validado pelos pesquisadores que criaram os instrumentos e também disponibilizavam uma calculadora com fórmula de acha o resultado final da análise.

Aspectos éticos

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde da Universidade Federal de Pernambuco. Aprovado sob Certificado de Apresentação para Apreciação Ética (CAAE) no 27115619.5.0000.5208 (Parecer no 3.824.448), foi dado prosseguimento com a realização da coleta dos dados. Todos os participantes assinaram, de forma livre e esclarecida, o termo de consentimento, aceitando participar do estudo.

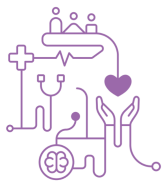
Resultados e Discussão

Perfil do público-alvo

O público-alvo para validação do chatbot foi dividido em 2 grupos. O Grupo 1 composto por estudantes de medicina do primeiro período que não tiveram contato com o tema da pesquisa e o grupo 2 composto por pessoas leigas que nunca fizeram o processo de ressuscitação cardiopulmonar.

Desenvolvimento do conteúdo

O conteúdo utilizado no assistente Virtual Inteligente foi elaborado e adaptado para ser usado em dispositivos móveis baseado na sequência do Suporte Básico de Vida do adulto realizada por leigos da Diretriz de Ressuscitação Cardiopulmonar e Cuidados Cardiovasculares de Emergência da Sociedade Brasileira de Cardiologia e da American



Heart Association ⁽¹²⁾. Entre os aspectos fundamentais do Suporte Básico de Vida (SBV) no adulto, foi escolhida a sequência de passos para RCP realizada pelo leigo para ser adaptada e compor o conteúdo utilizado no assistente virtual inteligente.

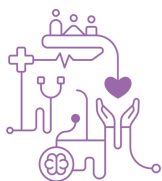
De acordo com o perfil educacional e as características de aprendizagem do público, o conteúdo foi adaptado em relação à linguagem de comunicação e as mídias utilizadas nos objetos de aprendizagem para que houvesse uma maior aceitação, engajamento e aprendizagem do público. Para atingir todos os tipos de público com os vários níveis de escolaridade, foram utilizadas mais de um tipo de mídia, como por exemplo vídeos, imagens e áudios que transcrevem o texto em som.

Validação do conteúdo

Para a validação do conteúdo por especialistas, foi composto um grupo avaliado contendo 9 profissionais, sendo 5 médicos(as) e 4 enfermeiros(as). Desses, 4 tem doutorado e 5 tem residência, com idades variando entre 26 e 59 anos e todos possuem mais de 5 anos de experiência em procedimentos de ressuscitação cardiopulmonar e em pesquisa na área de atuação.

Quadro 01 - Resultado da validação do conteúdo com as questões seguido do seu respectivo IVC

Domínio / Itens	IVC	Interpretação
1. Linguagem		
1.1- A linguagem está adequada para orientar o público leigo diante de uma emergência.	0,78	Bom
1.2- Não existem trechos que apresentam ambiguidade em sua instrução.	0,75	Bom
2. Conteúdo		
2.1- As recomendações estão corretas cientificamente.	0,88	Bom
2.2- As recomendações são claras e concisas.	1,00	Excelente
2.3- O conteúdo é apresentado de forma organizada e expressa uma única ideia.	1,00	Excelente
2.4- Os referenciais teóricos estão atualizados e adequados.	1,00	Excelente
3. Método		



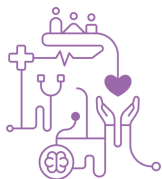
3.1- O chatbot de orientação ao leigo (assistente virtual) pode auxiliar nessa finalidade (Suporte básico de vida).	1,00	Excelente
4. Aplicabilidade		
4.1- O fluxograma da árvore de decisão para o RCP possui aplicabilidade prática.	1,00	Excelente
5. Impressão geral		
5.1- Recomendaria essa ferramenta para o público leigo.	1,00	Excelente
5.2- A ferramenta tem potencial didático para estudantes da área de saúde.	1,00	Excelente
Média	0,942	

Na avaliação, os especialistas analisaram os itens presentes na terceira parte do formulário de avaliação no Quadro 1, sendo que sete dos dez itens do conteúdo foram considerados “excelentes” com a pontuação do IVC 1,00 e apenas três itens tiveram índice de validade de conteúdo 0,75, 0,78 e 0,88, sendo interpretados como conteúdos “bons”. Quando analisada a média da validação de conteúdo verificou-se um excelente valor (0,942), indicando elevada concordância.

De acordo com o Quadro 1, o IVC médio foi verificado considerando média entre todos os índices de validação de conteúdo individuais, obtendo dessa forma um valor de 0,92 como IVC geral, portanto, superior ao valor objetivado de 0,8 e, conseqüentemente, com índice de concordância adequada por parte dos especialistas.

Analisamos as sugestões dos especialistas para o aperfeiçoamento do chatbot e adotando-as em sua totalidade o uso de palavras mais simples para melhor adequar o texto ao público leigo de diversos graus de instrução; implementação de áudio e vídeo para permitir tanto a transmissão de informação ao usuário sem que este tenha que parar o atendimento para ler o texto, quanto melhor instruir o leigo a respeito de como realizar as compressões; mudança nas imagens, para possibilitar um melhor entendimento por parte do usuário.

Além disso os especialistas recomendaram a melhora concernente a algumas palavras nos textos presentes no chatbot em questão, o que inclui tanto a adaptação com termos mais corretos como a substituição de palavras que possam ser de difícil compreensão



para leigos, além de correção de palavras com ortografia errada (as quais podem prejudicar o entendimento do público). Um exemplo da necessidade dessas correções fornecida por um dos profissionais foi a substituição da palavra “Ventilação” por “Respiração boca a boca” em uma das caixas de texto presente no chatbot, já que a primeira palavra poderia não ser compreendida por um leigo, havendo o risco de impossibilidade da realização da RCP. Além disso, foi sugerida também a necessidade de melhor descrição de alguns passos RCP presentes no chatbot, assim como a importância da inclusão de imagens que ilustrem certos procedimentos a fim de que sua execução seja facilitada.

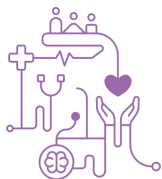
Desenvolvimento do chatbot

O chatbot foi desenvolvido pelos autores na plataforma manychats (<https://manychat.com/>) através do modelo orientado por regras. Isso significa que é preciso criar um script de perguntas e respostas e organizá-las na árvore de decisão de maneira coerente. O chatbot possuía vários tipos de recursos multimídias como áudio, vídeo, imagens e texto para atingir qualquer tipo de público inclusive os que não sabiam ler pois todo texto era transformado em áudio. O usuário também tinha a opção de clicar em botões para selecionar a opção de resposta desejada

Validação do chatbot juntamente com o público-alvo

Avaliação de usabilidade

Em relação a avaliação da usabilidade do chatbot em questão, a maior parte deles disse que não teve grandes dificuldades e que achou o sistema bastante intuitivo. Contudo, 2 usuários relataram falta de entendimento devido ao pouco uso de tecnologia, não sendo algo tão cotidiano para eles. Ademais, também foram solicitadas sugestões que possam trazer melhorias na usabilidade do chatbot. Assim, foram sugeridas melhorias na nitidez da disposição dos textos contendo as instruções, já que ajudariam na realização das tarefas, maior clareza nas perguntas, maior detalhamento no conteúdo



referente à respiração e a importância de considerar o nervosismo da pessoa que está prestando o socorro.

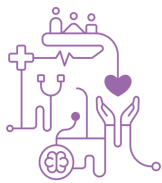
Eficácia e eficiência

O cálculo do resultado da eficácia para cada tarefa foi a média de todos os resultados em cada tarefa propostas e o resultado final das tarefas 1, 2 e 3, respectivamente, foram 1, 1 e 0,8. Assim, foi observado que terceira tarefa teve uma eficácia menor, o que provavelmente foi causado por uma variedade de respostas na árvore de decisão, já que uma das respostas possíveis levaria ao conhecimento que o faz saber o necessário para cumprir a terceira tarefa (saber como a vítima deve ser posicionada caso ela volte a respirar), enquanto uma das outras possíveis respostas não o instrui a esse respeito, impossibilitando o cumprimento da última atividade.

De acordo com ISO-9241 a Eficiência Relativa Geral calculada foi de 90,80%, o que significa que o tempo que realizar a tarefa no chatbot foi acima da média. Nesse sentido, essa elevada porcentagem demonstra que o chatbot é bastante eficiente no que se propõe, uma vez que, além de permitir que grande parte dos usuários pudessem concluir efetivamente as tarefas num menor tempo possível, fez com que eles alcançassem esse resultado em um espaço de tempo consideravelmente curto, para a maior parte dos usuários.

Satisfação

Dessa forma, em relação à primeira tarefa, obteve-se uma média de 5,65. Enquanto isso, na segunda tarefa, a média foi de 5,15 e, na terceira tarefa, a média foi de 5,60. Assim, pode-se perceber que todas as três tarefas tiveram uma média muito semelhante, mas as tarefas 1 e 3 foram mais fáceis de serem executadas, dada a maior média obtida a partir das notas fornecidas pelos usuários e, conseqüentemente, a maior proximidade da nota 7. Enquanto isso, a tarefa 2 mostrou-se mais difícil de ser concluída, já que obteve uma média um pouco menor que as duas outras tarefas, fator que pode ser explicado pela maior habilidade de atenção e memória que essa atividade exige, que



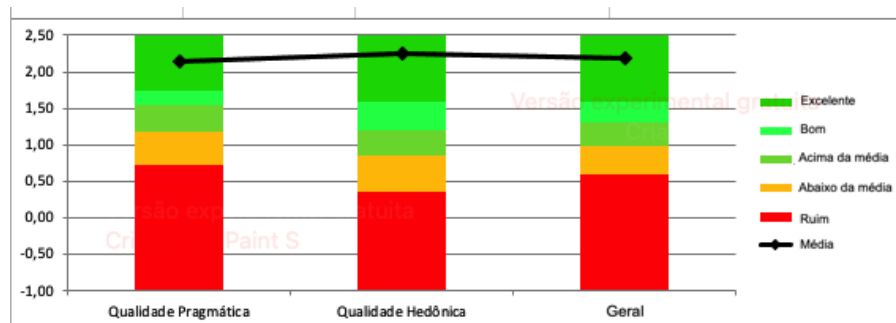
seria lembrar a quantidade de compressões por ciclo e quantos ciclos são necessários na reanimação cardiopulmonar (RCP). Além disso, a partir da análise do perfil demográfico e tecnológico dos participantes que tiveram maior dificuldade na execução da tarefa 2, não se encontrou características em comum entre essas pessoas que pudessem explicar o porquê dessa maior complexidade para a conclusão e aprendizado do que é exigido nessa atividade em específico.

No que tange à satisfação a nível geral do teste, utilizou-se o SUS, o qual pôde ser calculado por meio das notas dadas pelos usuários em um Google Forms. Assim, a partir das notas dadas pelos usuários, as quais variavam de 1 (Discordo completamente) a 2 (Concordo completamente), foi possível calcular o SUS, que teve como resultado médio 80,875. Dessa maneira, pode ser considerado que o resultado foi bastante satisfatório, uma vez que a média de pontuação no System Usability Scale é de 68 pontos e que pontuações acima de 80,3 são classificadas como excelentes, maior faixa de pontuação possível.

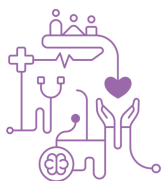
Experiência do usuário

Ao analisar as respostas do questionário sobre a experiência do usuário, pode-se notar na Figura 01 que a média geral de todos os itens das qualidades pragmáticas (2,138) e hedônica (2,250) foram acima da média geral do Benchmark sendo considerada excelente.

Figura 01 - Média geral do Benchmark da análise da experiência do usuário



Dessa forma, pode-se concluir que a qualidade hedônica do assistente virtual inteligente na ressuscitação cardiopulmonar foi maior que a sua qualidade pragmática, o



que reflete o seu caráter mais original, interessante e excitante de ser usado em relação a sua eficiência, facilidade de uso, capacidade de conduzir o usuário. Essa situação pode ser melhor compreendida se levarmos em consideração que o chatbot foi construído numa plataforma já existente, e que não foi construída com esse intuito, fazendo com que a adaptação da mesma não torne a experiência do usuário tão fluida e eficiente quanto poderia ser. Corroborar-se esse contexto por meio de algumas respostas do questionário de usabilidade que citam, como possíveis melhorias ao chatbot, situações que são inerentes aos mensageiros utilizados e que possuem limitadas ou inexistentes possibilidades de personalização por parte do produtor do chatbot, como a melhor disposição do texto e das imagens em relação a tela.

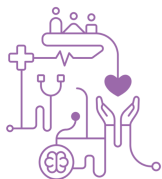
Tabela 01 - Resultado da análise de cada variável das qualidades pragmática e hedônica

Item	Média	Variança	Desv. Padrão	No.	Negativo	Positivo	Escala
1	2,2	2,1	1,5	20	Obstrutivo	Conduzidor	Qualidade Pragmática
2	2,2	2,5	1,6	20	Complicado	Fácil	Qualidade Pragmática
3	2,2	2,0	1,4	20	Ineficiente	Eficiente	Qualidade Pragmática
4	2,1	2,4	1,6	20	Confuso	Evidente	Qualidade Pragmática
5	2,5	1,9	1,4	20	Aborrecido	Excitante	Qualidade Hedônica
6	2,5	1,8	1,4	20	Desinteressante	Interessante	Qualidade Hedônica
7	2,4	1,8	1,4	20	Convencional	Original	Qualidade Hedônica
8	1,7	2,3	1,5	20	Comum	Vanguardista	Qualidade Hedônica

Ao analisar cada variável individual que compõem as qualidades pragmática e hedônica, pode-se notar, na Tabela 1, que as variáveis da qualidade pragmática mantiveram uma média satisfatória constante, com uma variação de 0.1 pontos, onde a variação da pontuação da média ficou entre 2.1 e 2.2. Em contrapartida, a média das variáveis da qualidade hedônica sofreu uma maior variação na pontuação de 0.7, onde a variação da pontuação da média ficou entre 1.7 e 2.5.

Ademais, outra questão digna de discussão é a média obtida pela classificação entre *leading edge* e *usual* (“vanguardista” e “comum” quando traduzidos para o português, respectivamente), que foi o menor valor quando comparado com os outros quesitos. Esse fato pode ser explicado pela crescente adoção de uso, por parte das empresas, das tecnologias de chatbot para o atendimento ao usuário.

Conclusão



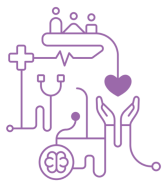
O Chatbot proposto foi desenvolvido e validado para ensinar o manejo do paciente com parada cardiorrespiratória para as pessoas que ainda não foram introduzidas para esse assunto, com o objetivo de introduzi-las a esse assunto e prepará-las para quando forem expostas a essa situação.

O processo de validação do conteúdo demonstrou, de maneira geral, um elevado grau de concordância entre os especialistas (IVC médio de 0,942), evidenciando a adequação do conteúdo utilizado para a construção do chatbot. Ademais, a validação do aprendizado obtido com a utilização da tecnologia desenvolvida revelou um incremento médio de 33% no conhecimento dos participantes, o que aconteceu em 75% deles. A avaliação da usabilidade obteve resultados que demonstraram, na média, a ausência de dificuldades para o uso da tecnologia em grande parte, devido à maior presença desse tipo de interação no cotidiano, exceto em usuários menos familiares com as novas tecnologias. No âmbito da eficiência, foi obtido uma Eficiência Relativa Geral calculada foi de 90,80%, o que demonstra que o chatbot é bastante eficaz para a função que foi projetada, visto que fez com que grande parte dos usuários concluíssem efetivamente as tarefas. Outrossim, o nível de satisfação dos usuários com a utilização do chatbot, obteve um valor do *System Usability Scale* de 80,875, sendo assim classificado como excelente. A avaliação da interface obteve um ótimo desempenho e aceitação.

Pode ser indicado como fatores limitantes do presente projeto a adaptação de plataformas já existentes para um uso específico, o que diminui as possibilidades de customização podendo levar a prejuízos a experiência do usuário.

Referências

1. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of Synchronous Text-Based Dialogue Systems in Mental Health Interventions: Systematic Review. *J Med Internet Res* [Internet]. 1º de agosto de 2017 [citado 29 de janeiro de 2024];19(8). Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28784594/>
2. Rocha JSY. Uso de tecnologias da informação e comunicação na educação em saúde. *Medicina (Ribeirão Preto)*. 2015;48(3):214–23.
3. De Campos Filho AS, Vasconcelos Cursino JR, Do Nascimento JWA, De Souza RR, Da Silva Lima Roque G, De Souza Cavalcanti AR. Content and Usability Validation of an Intelligent Virtual Conversation Assistant Used for Virtual Triage During the COVID-19 Pandemic in Brazil.



- Comput Inform Nurs [Internet]. 2 de novembro de 2022 [citado 29 de janeiro de 2024];40(11):779–85. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/35234699/>
4. Holmes S, Moorhead A, Bond R, Zheng H, Coates V, Mctear M. Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces? Em: Proceedings of the 31st European Conference on Cognitive Ergonomics [Internet]. New York, NY, USA: Association for Computing Machinery; 2019. p. 207–14. (ECCE '19). Disponível em: <https://doi.org/10.1145/3335082.3335094>
 5. Clark DE. R A Cowley, the “Golden Hour,” the “Momentary Pause,” and the “Third Space” - PubMed. Am Surg [Internet]. 1º de dezembro de 2017 [citado 29 de janeiro de 2024];83(12):1401–6. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29336762/>
 6. Pasquali L. Psicometria: Teoria dos testes na psicologia e na educação. São Paulo; 2013.
 7. ISO 9241-11:2018(en), Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts [Internet]. [citado 28 de agosto de 2022]. Disponível em: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>
 8. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) [Internet]. 2008 [citado 29 de janeiro de 2024];5298 LNCS:63–76. Disponível em: https://link.springer.com/chapter/10.1007/978-3-540-89350-9_6
 9. Coluci MZO, Alexandre NMC, Milani D. Construção de instrumentos de medida na área da saúde. Cien Saude Colet [Internet]. março de 2015 [citado 29 de janeiro de 2024];20(3):925–36. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232015000300925&lng=pt&tlng=pt
 10. Lynn MR. Determination and quantification of content validity. 1986 [citado 29 de janeiro de 2024];35(6):382–5. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/3640358/>
 11. Grant JS, Davis LL. Selection and use of content experts for instrument development. Res Nurs Health [Internet]. junho de 1997 [citado 29 de janeiro de 2024];20(3):269–74. Disponível em: [https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1098-240X\(199706\)20:3<269::AID-NUR9>3.0.CO;2-G](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G)
 12. Merchant RM, Topjian AA, Panchal AR, Cheng A, Aziz K, Berg KM, et al. Part 1: Executive Summary: 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation [Internet]. 20 de outubro de 2020 [citado 29 de janeiro de 2024];142(2):S337–57. Disponível em: <https://www.ahajournals.org/doi/abs/10.1161/CIR.0000000000000918>