Assistente Virtual na Educação em Saúde dos Homens
Virtual Assistant in Men’s Health Education
Asistente Virtual en Educación para la Salud del Hombres

Amadeu Sá de Campos Filho¹, Jose Ricardo Vasconcelos Cursino², Thomé Décio Pinheiro Barros Júnior³, Évelyn Cristina Morais Pessôa Lima⁴.

1. Center for Medical Sciences. Federal University of Pernambuco, Brazil. Doctor in Computer Science at the Informatics Center of the Federal University of Pernambuco;
2. São Miguel University Center, Brazil. Bachelor in Nursing;
3. Center for Medical Sciences. Federal University of Pernambuco, Brazil. Doctor in Medicine (Urology) from the Federal University of São Paulo;
4. Nossa Senhora das Graças College of the University of Pernambuco, Brazil. Bachelor in Nursing.

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

Resumo

Objetivo: Este artigo tem como objetivo oferecer um método eficaz de educação em saúde por meio do IChatSUS, um Agente Inteligente de Conversação, que possa esclarecer dúvidas sobre questões de saúde do homem. Método: A pesquisa foi observacional descritiva com abordagem quantitativa e qualitativa por meio de um estudo transversal. Foi realizado um estudo de caso com 20 participantes para avaliar a Usabilidade do IChatSUS e a experiência do usuário por meio do System Usability Score e User Experience Questionnaire. Resultado: Quanto ao perfil tecnológico, 90% da amostra tinha acesso à internet e 95% relataram não conhecer chatbots. A usabilidade do IChatSUS na escala SUS revelou a necessidade de melhorias. Ao avaliar a Experiência do Usuário, os resultados foram otimistas. Conclusão: A utilização do IChatSUS como método de educação em saúde demonstrou excelente desempenho, mas necessidade de revisões. É preciso expandir sua utilização em diversos ambientes e públicos para uma efetividade maior do IChatSUS.

Descritores: Educação em Saúde; Tecnologia Educacional; Saúde do Homem.

Abstract

Objective: This article aims to offer an effective method of health education through IChatSUS, an Intelligent Conversation Agent that can ask about men's health issues. Method: It was a descriptive observational with a qualitative and qualitative approach through a cross-sectional study. It was conducted with 20 participants to assess the
Usability of IChatSUS and the user experience through the System Usability Score and User Experience Questionnaire. **Results:** Regarding the sample, 90% of the sample had access to the internet, 95% of the users had a technological profile and did not know about chatbots. The usability of IChatSUS revealed a need for improvement. When evaluating the User Experience, the results were optimistic. **Conclusion:** The use of health education IChatSUS had an excellent performance but reveals the need for maintenance. It is necessary to expand the public and environment of the ChatSUS's application for greater use of it. **Keywords:** Health Education; Educational Technology; Men's Health.

**Resumen**

**Objetivo:** Este artículo tiene como objetivo ofrecer un método efectivo de educación para la salud a través de IChatSUS, un Agente de Conversación Inteligente que puede preguntar sobre problemas de salud de los hombres. **Método:** A fue observacional descriptivo con un enfoque cualitativo y cualitativo a través de un estudio transversal. Se llevó a cabo con 20 participantes para evaluar la usabilidad de IChatSUS y la experiencia del usuario a través de estudios de System Usability Score y User Experience Questionnaire. **Resultados:** En cuanto a la muestra, 90% tenía acceso a internet, el 95% de los usuarios tenía perfil tecnológico y desconocía sobre chatbots. La usabilidad de IChatSUS reveló una necesidad de mejora. Al evaluar la Experiencia de Usuario, los resultados fueron optimistas. **Conclusión:** El uso de IChatSUS para la educación en salud es excelente como rendimiento, pero requiere mantenimiento. Es necesario ampliar el público y el ambiente para mejor uso de IChatSUS. **Descriptores:** Educación en Salud; Tecnología Educacional; Salud del Hombre.

**Introduction**

Male health care has attracted the focus of health agencies in recent decades. Since the publication of the National Policy for Integral Attention to Men's Health (NPIAMH) on August 27, 2009, the need for preventive measures that objective greater care for the male public has been notorious\(^1\), and according to Brito and Santos\(^2\), this need is caused by the high morbidity/mortality rates common to this public, problems that could be avoided if, effectively, there was an early diagnosis and treatment.
In Brazil, there are several problems and difficulties demonstrated by men seeking health care or performing diagnosis and basic health care due to some factors such as: i. verbalization of feeling since they have difficulties in verbalizing their own needs, because talking about their health problems can mean a possible demonstration of weakness, of feminization towards others and the idea of feminization associated with health care; ii. the opening hours of the health units, since men generally work when the basic health units are open; the severity of the health problems, which limits care actions; iii. Shame: they are ashamed of exposing themselves to another man or woman, probably this shame is associated with the historical issue and the lack of habit of exposing themselves to a health professional; iv. Cultural factors where machismo is associated with the idea that “man does not get sick”; v. The difficulty of socialization in health environments of these users, which comes from a heavy cultural load of invulnerability and the failure to implement programs in primary care that promote the inclusion of these.

Therefore, these factors are considered to be responsible for these behaviors that end up for hindering a more global follow-up, including prevention actions for the male.\(^{(3,4)}\)

One of the ways to combat these difficulties is the use of new communication and information technologies (CITs) such as Artificial Intelligence together with machine learning through intelligent conversation agents (Chatbots). Because they are wide-use and easy to interact with, such technologies facilitate the dissemination and exchange of information for the construction of knowledge in various areas, which includes health education\(^{(5,6)}\).

One way to use chatbots to share information about men's health and clarify doubts from the population is the development of intelligent conversation alums also called chatbots. They prove to be a technology of optimal use by simulating a dialogue with the user, so that it is like a human interaction through messages\(^{(7)}\). Chatbots are resources that patients can access for information, suggestions, and 24/7 care to improve their health. Chatbots can provide a better health experience.

This technology is a tool for the dissemination of effective knowledge has already been used in several areas of health, such as: medical clinic, psychiatry and women's health\(^{(8,9,10)}\). On the other hand, in Brazil, a large part of the population does not have
adequate technological knowledge to use the applications either, because of the lack of digital skills or the complexity that the application is developed.\textsuperscript{(11)} Thus, it is necessary to survey the technological profile of the target audience and also the level of usability of the application to understand the gaps in the lack of engagement in the use of educational technologies.

In this scenario, this article aims to evaluate the effectiveness, user experience and acceptance of the use of a chatbot in men’s health education within the Brazilian Unified Health System (IChatSUS) encouraging the search for health care and improving their engagement in the search for information to answer their doubts about men’s health.

**Methods**

The research is descriptive observational with a quantitative and qualitative approach through a cross-sectional study. In the data collection, the following methods and techniques were used: interview, providing instructions, observation and screen capture of the mobile phone.

The study was conducted at the Gouveia de Barros polyclinic and at the Testing and Counseling Center (CTA) in the city of Recife, Brazil. The study population consisted of users who sought care at the service where the data were collected. The sample consisted of 20 men who agreed to participate in the research, provided that they met the selection criteria explained below. They were randomly invited to participate in the experiment as they arrived at the polyclinic. After they accepted the invitation, they read and signed the Free and Informed Consent Form.

**Criteria of Inclusion and Exclusion**

The inclusion criteria were the patient being male who sought care in the unit, had a smartphone and was willing to perform the research. On the other hand, patients under 18 years of age and unable to verbally expose their opinion about the object of study were excluded.

**Instruments of Data Collection**
For data collection, two forms were used: the first to obtain sociodemographic data containing 18 questions related to the collection of the sample profile, such as age, family income, level of education, internet access, among others. The second form was a combination of two evaluation methods, referring to Usability and User Experience.

For a data evaluation, two validated questionnaires were used, System Usability Score (SUS) and User Experience Questionnaire (UEQ). The SUS\(^{(12)}\) consists of a questionnaire with 10 items and 5 answer options, ranging from 'Totally agree' to 'Totally disagree'. With the data results in hands, it is possible to calculate the performance of the system, the SUS determines 68 points as an average score to have a good usability, any value below that requires a greater commitment to the usability of your system.\(^{(13)}\) The UEQ, as well as the SUS, uses a Likert questionnaire with 7 items, containing at the ends two semantically opposed words, such as Unpleasant - pleasant, in which the user will choose an option that refers to the closest to his experience.\(^{(14)}\)

**Procedures for data collection**

Data were collected from face-to-face interaction, in a place that promoted the confidentiality and participation of the researched subject, through the completion of the tool. The tasks defined for users to perform were: 'Seek self-care information', 'Look for information on ways to get involved in pregnancy', 'Seek paternity leave information', 'Seek information about your rights as a man in health', 'Seek information on Prostate Cancer', 'Seek information about HIV/AIDS', 'Seek information about Syphilis', 'Seek information about HPV', 'Look for information about Herpes', 'Seek information about Hepatitis'.

To avoid identifying the participants, a code with the acronym H (man) and a number from 0 to 20 were adopted to name the user. During the entire performance of the tasks, the screen was captured to obtain the time of use of the application. After use, the SUS and UEQ questionnaires were applied to obtain the information regarding the use of IChatSUS.

**Data analysis and interpretation**
The research used methods of analysis of the differentiated data for each tool. For the sociodemographic questionnaire, the evaluation method consisted of descriptive statistical analysis mediated by tabulation of data in Microsoft Excel and later presented in graphs and tables. On the other hand, the data referring to the usability questionnaires and user experience of the IChatSUS, because they are validated forms, the analysis followed the procedures recommended by the authors of each questionnaire.

In addition to the two evaluations described above, we decided to insert one more question to qualitatively evaluate the preference of users over the method of interaction with IChatSUS due to their low level of education.

**Ethical aspects**

Before performing data collection, the project was submitted to the research ethics committee of the Center for Medical Sciences of the Federal University of Pernambuco. Approved under CAAE number 23347119.8.0000.5208, data collection continued.

**Results and Discussion**

**Conversation flowchart**

With the analysis of the results of the literature review, it was possible to establish the subjects that caused the most doubts to men and thus proceed to the second stage of the research. This step consisted in the creation of the decision tree flowchart for the establishment of the lines of dialogue between the user and the chatbot. These lines of dialogue aimed to perform the interaction the user can be led to specific outcomes based on their responses and facilitate the recognition of errors in the execution of IChatSUS. The decision tree flowchart was developed in a mental map building application called Whimsical® (https://whimsical.com/), because it is free and easily accessible, also allowing the elaboration together.

The decision tree was divided into three groups of subjects: "Self-care, Citizenship and Diseases". Each group was divided into themes with specific questions for each doubt
defined as can be seen in “Figure 1”. The contents of the decision tree were validated by 04 experts in men's health.

**Figure 1** - Simplified IChatSUS decision tree, Recife-PE, 2021.

---

**Development of IChatSUS**

IChatSUS was developed by Google's Dialogflow® platform (https://dialogflow.cloud.google.com/), which is a natural language processing platform where it facilitates the design and integration of a conversational user interface with mobile apps and web applications. After choosing the development platform, the conversation platform was defined on which the chatbot would be inserted. Telegram® (https://web.telegram.org/) was the choice because it had not cost of using Chatbot; it can be used with no Application Programming Interface (API), because it has an own tool to implement chatbots and, finally, it doesn't need another app to make the chatbot functioning. After writing the chatbot source code and inserting it into the platform, tests were performed by the research team to correct errors and failures in the chatbot. After the error tracking and their correction were performed, the implementation and data collection stage were continued.

**Case study**

The first data collected in the interview were sociodemographic and technological data, with the objective of tracing the profile and digital skills of the public who made use of and performed the IChatSUS evaluation. When analyzing the profile of the research sample, the following data were obtained.
Sociodemographic profile

When measuring the age of the sample, it was observed that 7 (35%) men are between 20 and 30 years old, the most expressive age groups. Furthermore 5 (25%) are men from 31 to 40, 6 (35%) between 41 and 50 years old and 2 (10%) over 50.

When analyzing the data regarding schooling, according to the Brazilian Education Guidelines Law, of the research participants, it was observed that 03 (15%) men had incomplete elementary school, 01 (5%) had completed elementary school, 01 (5%) incomplete secondary education, 10 men (50% of the sample) had completed secondary education, 03 (15%) had incomplete higher education and 02 (10%) had completed higher education. In this scenario, we can see that most of the sample was literate.

When comparing the data regarding schooling with those obtained by the 2019 IBGE census, it can be observed that the sample obtained has a higher level of education, since the highest percentage in IBGE are those with incomplete elementary school, with 32.2%.(15)

In relation to family income, two results were obtained: those who had 0 to 1 minimum wage with 45% of the sample and those surveyed who had income between 2 and 3 minimum wages, which in turn represents 55% of the total surveyed.

Technological Profile

When questioning users about internet access, it was observed that only 10% of the sample did not have it. When comparing the access data with the 2019 IBGE census, it is possible to notice that the sample surveyed has a higher percentile than the data obtained by the agency in the northeast region, such data reveal that 79.1% of the houses in the urban area have access to the network.(16) In addition, the frequency and time of average internet use among the 90% of users who had internet was also evaluated.

When analyzing the frequency of internet use during the week, it was observed that, of the 18 users who had internet access, 15 (83.3%) reported daily use, either through WIFI, via network cable or mobile data. Regarding the daily time of internet use, of the 18 users, 3 (16.7%) spend more than 8 hours online, 2 (11.1%) go from 7 to 8
hours, 2 (11.1%) spent from 5 to 6 hours per day, 5 (27.8%) go from 3 to 4 hours, 2 (11.1%) go from 1 to 2 hours and 4 (22.2%) of these users use the internet for less than 1 hour per day.

**IChatSUS usability assessment**

The time of use of IChatSUS was on average 10 minutes and 12 seconds of interaction with IChatSUS, the maximum interaction time was 14 minutes and 17 seconds, the minimum time was 6 minutes and 49 seconds and standard deviation of 2 minutes and 11 seconds. “Graphic 1” demonstrates the individual usage time of IChatSUS.

To observe the performance of IChatSUS, we saw the need to evaluate usability and experience of use, so that it was possible to analyze the desired aspects through methodologies validated and previously evaluated in other chatbots.

The IChatSUS, according to the analysis of the results of the SUS questionnaire, obtained 51 points, which revealed that a review is necessary in terms of the usability of IChatSUS. This result is part of this sociodemographic and technological's findings about the profile of users, because it showed that they did not have digital skills in the use of mobile devices. On the other hand, such a need for improvements in usability was also observed in the report of the users, such as the H9 user "to facilitate interaction, it would be better to add colors and images, in addition to having a more popular language.", as well as the H11 user who suggested "Could add more photos". These are means of solving the problem with usability.
User experience assessment

The UEQ questionnaire, although extensive, contains a more complete analysis of the results of the user experience. Evaluating aspects such as attractiveness, transparency, efficiency, control, stimulation and innovation, UEQ creators have developed a spreadsheet that automatically calculates the collected results.

To be able to evaluate, the UEQ establishes the need to use two opposing words to conduct the assessment of the item addressed. Below you can see “Table 1”, which contains data regarding the answers for each pair of words.

Table 1 - Individual results of each pair of words, as well as definition of which domain each pair corresponds to. Source: UEQ data analysis spreadsheet, Recife-PE, 2021.

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. Dev.</th>
<th>Left</th>
<th>Right</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.6</td>
<td>1.0</td>
<td>1.0</td>
<td>Annoying</td>
<td>Enjoyable</td>
<td>Attractiveness</td>
</tr>
<tr>
<td>2</td>
<td>2.8</td>
<td>0.5</td>
<td>0.7</td>
<td>Not understandable</td>
<td>Understandable</td>
<td>Perspicuity</td>
</tr>
<tr>
<td>3</td>
<td>2.2</td>
<td>3.0</td>
<td>1.7</td>
<td>Creative</td>
<td>Dull</td>
<td>Novelty</td>
</tr>
<tr>
<td>4</td>
<td>2.5</td>
<td>2.1</td>
<td>1.4</td>
<td>Easy to learn</td>
<td>Difficult to learn</td>
<td>Perspicuity</td>
</tr>
<tr>
<td>5</td>
<td>2.8</td>
<td>0.6</td>
<td>0.8</td>
<td>Valuable</td>
<td>Inferior</td>
<td>Stimulation</td>
</tr>
<tr>
<td>6</td>
<td>1.7</td>
<td>4.1</td>
<td>2.0</td>
<td>Boring</td>
<td>Exciting</td>
<td>Stimulation</td>
</tr>
<tr>
<td>7</td>
<td>2.3</td>
<td>1.4</td>
<td>1.2</td>
<td>Not interesting</td>
<td>Interesting</td>
<td>Stimulation</td>
</tr>
<tr>
<td>8</td>
<td>2.1</td>
<td>3.2</td>
<td>1.8</td>
<td>Unpredictable</td>
<td>Predictable</td>
<td>Dependability</td>
</tr>
<tr>
<td>9</td>
<td>2.4</td>
<td>1.9</td>
<td>1.4</td>
<td>Fast</td>
<td>Slow</td>
<td>Efficiency</td>
</tr>
<tr>
<td>10</td>
<td>1.4</td>
<td>5.8</td>
<td>2.4</td>
<td>Inventive</td>
<td>Conventional</td>
<td>Novelty</td>
</tr>
<tr>
<td>11</td>
<td>2.1</td>
<td>3.6</td>
<td>1.9</td>
<td>Obstructive</td>
<td>Supportive</td>
<td>Dependability</td>
</tr>
<tr>
<td>12</td>
<td>2.8</td>
<td>0.6</td>
<td>0.8</td>
<td>Good</td>
<td>Bad</td>
<td>Attractiveness</td>
</tr>
<tr>
<td>13</td>
<td>2.6</td>
<td>1.8</td>
<td>1.4</td>
<td>Complicated</td>
<td>Easy</td>
<td>Perspicuity</td>
</tr>
<tr>
<td>14</td>
<td>2.3</td>
<td>1.7</td>
<td>1.3</td>
<td>Unlikely</td>
<td>Pleasing</td>
<td>Attractiveness</td>
</tr>
<tr>
<td>15</td>
<td>1.1</td>
<td>6.2</td>
<td>2.5</td>
<td>Usual</td>
<td>Leading edge</td>
<td>Novelty</td>
</tr>
<tr>
<td>16</td>
<td>2.2</td>
<td>2.5</td>
<td>1.6</td>
<td>Unpleasant</td>
<td>Pleasant</td>
<td>Attractiveness</td>
</tr>
<tr>
<td>17</td>
<td>2.9</td>
<td>0.5</td>
<td>0.7</td>
<td>Secure</td>
<td>Not secure</td>
<td>Dependability</td>
</tr>
<tr>
<td>18</td>
<td>2.5</td>
<td>1.1</td>
<td>1.1</td>
<td>Motivating</td>
<td>Demotivating</td>
<td>Stimulation</td>
</tr>
<tr>
<td>19</td>
<td>2.2</td>
<td>3.4</td>
<td>1.9</td>
<td>Meets expectations</td>
<td>Does not meet expectations</td>
<td>Dependability</td>
</tr>
<tr>
<td>20</td>
<td>2.6</td>
<td>0.8</td>
<td>0.9</td>
<td>Inefficient</td>
<td>Efficient</td>
<td>Efficiency</td>
</tr>
<tr>
<td>21</td>
<td>2.0</td>
<td>3.7</td>
<td>1.9</td>
<td>Clear</td>
<td>Confusing</td>
<td>Perspicuity</td>
</tr>
<tr>
<td>22</td>
<td>2.4</td>
<td>2.3</td>
<td>1.5</td>
<td>Impractical</td>
<td>Practical</td>
<td>Efficiency</td>
</tr>
<tr>
<td>23</td>
<td>2.5</td>
<td>1.9</td>
<td>1.4</td>
<td>Organized</td>
<td>Cluttered</td>
<td>Efficiency</td>
</tr>
<tr>
<td>24</td>
<td>2.4</td>
<td>1.6</td>
<td>1.3</td>
<td>Attractive</td>
<td>Unattractive</td>
<td>Attractiveness</td>
</tr>
<tr>
<td>25</td>
<td>2.4</td>
<td>2.2</td>
<td>1.5</td>
<td>Friendly</td>
<td>Unfriendly</td>
<td>Attractiveness</td>
</tr>
<tr>
<td>26</td>
<td>1.6</td>
<td>5.7</td>
<td>2.4</td>
<td>Conservative</td>
<td>Innovative</td>
<td>Novelty</td>
</tr>
</tbody>
</table>

By analyzing “Table 1”, you can see the domains that each pair of words discusses. A key point to be emphasized is the need to change words, due to the diversity of instruction of the sample surveyed. This point can be observed in the pair of words 'Common/Vanguardist', referring to innovation; these terms can generate doubts and thus provide a less accurate result in the domain addressed.
Another evaluation conducted by UEQ is a benchmark with other technologies of the same branch that exists in the questionnaire database, which resulted in a table expressing the performance achieved by IChatSUS. This result is expressed in “Table 2”.

Table 2 - IChatSUS Benchmark according to UEQ evaluation sheet, Recife-PE, 2021.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>Comparassion to benchmark</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>attractiveness</td>
<td>2.41</td>
<td>Exceelle</td>
<td>In the range of the 10% best results</td>
</tr>
<tr>
<td>transparency</td>
<td>2.46</td>
<td>Exceelle</td>
<td>In the range of the 10% best results</td>
</tr>
<tr>
<td>efficiency</td>
<td>2.48</td>
<td>Exceelle</td>
<td>In the range of the 10% best results</td>
</tr>
<tr>
<td>control</td>
<td>2.29</td>
<td>Exceelle</td>
<td>In the range of the 10% best results</td>
</tr>
<tr>
<td>stimulation</td>
<td>2.29</td>
<td>Exceelle</td>
<td>In the range of the 10% best results</td>
</tr>
<tr>
<td>innovation</td>
<td>1.56</td>
<td>Exceelle</td>
<td>In the range of the 10% best results</td>
</tr>
</tbody>
</table>

As observed, in all respects the IChatSUS was in the range of the best 10%, being worth mentioning the innovation item that was also among the 75% worst (which may have been caused by the terms used, as previously mentioned). These results demonstrate the effectiveness of Chatbot in the transmission of knowledge in the health area, especially when analyzing the use profile of the sample used.

Another key point is the observance of the percentage of users who have never used a chatbot before (“graphic” 1), when commencing with the results of “Table 2” it is possible to notice that IChatSUS performed excellently.

Evaluation of interface interaction

Another question evaluated is the method of interaction with IChatSUS. The available methods were the writing of the question, selecting a button with the desired question and entering a code referring to the desired question. As a result, 65% of men reported that clicking on the desired option is the best method of interaction with the chatbot, although 25% preferred entering a code referring to the desired question.

By analyzing the data, you can associate the preference by "clicking on the desired option" with the ease of interaction at the time of using the chatbot. Even though most existing chatbots use the question writing method as a means of interaction, it is notorious that users prefer the convenience of just clicking to get their answers.

Conclusions
With the results obtained by the research, it can be noted that the use of chatbot in education was again proven as an effective means in the discussion of knowledge. As highlighted by King(7), ease of access and easy interaction are advantages found with the use of this medium.

From the perspective of the men involved in the research, the use of this means of knowledge propagation is, according to H3, "Interesting, leaves man more integrated over his rights." This report emphasizes that the user understood the real meaning of the creation of IChatSUS, which demonstrates that the objective was achieved.

However, it is necessary to pay close care of the result obtained about usability, with the score of 51, according to the SUS it is necessary to reevaluate the way the information is passed and the interface, as suggested in the H9 report "to facilitate the interaction it would be better to add colors and images, besides having a more popular language." Such changes can be made and applied to IChatSUS, requiring a retesting to confirm the solution of the problem.

The study had some difficulties and limitations, among them: the availability of the researcher's schedule with the opening hours of the health unit where the research was carried out and the public's engagement in accepting to participate in the research. All usability difficulties and limitations found will be resolved in the next version of IChatSUS as reported by users.

Acknowledgments

We want to acknowledge the expertise support from the Telehealth Center (NUTES) of Medical Sciences Center of Federal University of Pernambuco.

References


