International Journal of Innovative Research and Scientific Studies, 9(6) 2026, pages: 10-20

IJIRSS
International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Usability of KAI: A generative AI chatbot to support mental health among university students in Latin America

Saray Galeano Ospino[1*], Luz Marcela Restrepo Tamayo[2], Rodrigo Arias[3], Santiago Mazuera Clavijo[1]

[1]*Corporación Universitaria Adventista, Antioquia. Medellín, Colombia.*
[2]*Universidad de Medellín, Antioquia. Medellín, Colombia.*
[3]*Universidad Adventista del Plata, Libertador San Martin. Entre Ríos, Argentina.*

Corresponding author: Saray Galeano Ospino (*Email: sagaleano@unac.edu.co*)

Abstract

Mental health among university students in Latin America has been compromised by stress, depression, and anxiety, which have been exacerbated by the pandemic. Despite the vulnerability of this population and the potential of digital tools, the actual adoption of chatbots for psychological support remains low due to doubts about their effectiveness and ease of use. The purpose of this study is to analyze the usability of KAI, a virtual assistant powered by generative artificial intelligence designed to assess symptoms using the DASS-21 scale and offer personalized recommendations. The methodology employed a mixed-methods descriptive and exploratory approach with 112 students from three Latin American universities, using the System Usability Scale (SUS) and qualitative interviews with psychology students. The main results indicate that the chatbot has outstanding usability and high acceptability, with a positive perception of its ease of use and security, although significant differences in evaluation were identified among the participating institutions. The main conclusion highlights that KAI is an intuitive and technologically viable tool that functions as a scalable emotional support resource, provided it maintains strict professional referral protocols in cases of high severity.

Keywords: Chatbots, Generative artificial intelligence, Mental health, college students, System usability.

DOI: 10.53894/ijirss.v9i6.11705
Funding: This study received no specific financial support.
History: Received: 12 March 2026 / **Revised:** 11 May 2026 / **Accepted:** 15 May 2026 / **Published:** 25 May 2026

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

College students face multiple factors that can trigger stress and compromise their mental health and overall well-being. Depression, anxiety, and loneliness are common conditions among this group; in fact, global estimates suggest that

more than 30% of the student population in higher education exhibits clinically significant symptoms [1]. Unlike elementary and secondary school students, college students are more vulnerable because this stage of education involves intensified academic pressure and financial burdens, in addition to the complex transition to adulthood [2].

Concerns about mental health among college students have driven the adoption of tools such as artificial intelligence (AI), used particularly to strengthen psychological support and counseling services. The rapid expansion of AI tools in our daily lives coincides with an ongoing mental health crisis on college campuses, particularly regarding social anxiety and participation-related distress [3]. One of the most widely used applications of AI is chatbots, which can communicate with and guide people through various tasks [4]. In this way, generative chatbots, such as ChatGPT and Google Gemini, have evolved from niche applications to widespread adoption, permeating basic academic functions such as advising and peer collaboration [3].

Generative AI chatbots can act as virtual therapists, offering counseling and serving as a support line for people with mental health issues in a more affordable way than traditional psychotherapy [5]. Although they do not hold medical degrees, these chatbots have made their way into conversations about healthcare and have been applied in fields such as surgery, oncology, and orthopedics, proving to be versatile and widely useful [6]. This interactive and personalized format allows users to ask sensitive questions and receive immediate, personalized responses in a confidential and non-judgmental space. These features distinguish chatbots from other sources of digital information and make them particularly relevant for addressing intimate health concerns [6].

Despite the potential of these technologies, the actual use of chatbots for psychological support remains low, with only 5% of students reporting that they have sought help through these channels [4]. Regarding their use, the literature reports some limitations, such as students perceived lack of need and persistent doubts about the effectiveness of AI compared to traditional therapy [4]. On the other hand, although AI can facilitate the self-disclosure of intimate emotions [7] ethical concerns remain regarding data security, the rigor of evidence-based content, and the systems' responsiveness to emergency situations [2]. Thus, perceptions of usefulness and ease of use remain key factors limiting the widespread adoption of these tools in academic settings [5].

Given the above, the purpose of this research is to analyze the usability of KAI, a generative AI-powered chatbot designed to assist college students exhibiting symptoms of stress, depression, and anxiety, which incorporates the DASS-21 scale as a tool to assess these symptoms. The study seeks to determine whether this generative AI-based virtual assistant can serve as a scalable and effective mental well-being resource that complements existing university support services [3].

To this end, this study adopted a mixed-methods approach with a descriptive and exploratory design to evaluate the user experience and the perceived usefulness of the chatbot. The evaluation phase focused on the participation of university students from three Latin American universities, who interacted with the designed tool. To measure KAI's usability, the System Usability Scale (SUS) [8] was administered. Additionally, qualitative data were collected through interviews with a special group of psychology students in their final semesters, focusing on identifying barriers to adoption and the accuracy of the chatbot's responses in contexts of emotional support for stress, depression, and anxiety.

Quantitative data were analyzed using descriptive statistics to determine frequencies and trends in the perception of the system's ease of use and security. To ensure the reliability of these results, Cronbach's alpha coefficient was calculated, confirming high internal consistency of the instrument. Additionally, an inferential analysis was conducted to validate whether the response patterns were statistically significant.

The results from the SUS scale show that the chatbot has high acceptability and outstanding usability among the evaluators. The descriptive analysis of the responses reveals a significant concentration in the highest categories of the scale for positive items, particularly highlighting the perception of ease of use, speed of learning, and integration of system functions. Conversely, the negatively worded items confirm that users did not perceive the chatbot as complex, inconsistent, or dependent on constant technical support. These data, supported by a Cronbach's alpha coefficient that validates the instrument's reliability, suggest that the tool is intuitive and creates a safe interactive environment, satisfactorily meeting the technical requirements to serve as a resource for emotional support in the face of stress and anxiety in the university setting.

Considering the elements described above, this article is structured as follows: Section 2 describes the materials and methods used in this research. Section 3 presents the results obtained, which are analyzed in Section 4. Finally, Section 5 presents the conclusions and outlines future lines of research.

2. Materials and Methods

The KAI chatbot is a digital tool developed with the aim of identifying symptoms of depression, stress, and anxiety in college students and providing them with personalized relaxation tools. Its development aligns with the growing need to offer psycho-emotional support in university settings, where the prevalence of these symptoms intensified following the pandemic.

KAI is designed as a modular system that integrates psychometric assessment with generative AI capabilities to recommend personalized emotional support tools. Once the student logs in, they take the DASS-21 test, which measures their levels of depression, anxiety, and stress [9]. The results of this test are processed by the system's back end, which uses the gpt-oss AI model (an open-source model from OpenAI). This AI model does not operate generically but uses the specific diagnosis (the mild and moderate severity levels for each factor) to generate personalized relaxation recommendations that have been predefined by a mental health specialist for each severity level (Table 1). This mechanism ensures that the suggested tools are relevant to the user's current emotional state.

Table 1.
Relaxation tools defined for the symptoms identified by the DASS-21.

Factor	Severity	Tool	Description
Stress	Mild	Deep Breathing	Inhale slowly for 4 seconds, hold for 4, exhale for 4. Repeat this cycle 5 times.
	Moderate	Biblical Meditation	Choose a biblical text from the Psalms. Read it mindfully, taking in each word. Choose a particular word or phrase that you find meaningful. Close your eyes, repeat it slowly while taking deep breaths. Remain in a state of relaxation for a few moments, noting any impressions that arise from the chosen word.
Anxiety	Mild	5-4-3-2-1 Technique	Name 5 things you see, 4 you can touch, 3 you hear, 2 you smell, and 1 you taste.
	Moderate	Progressive muscle relaxation	Systematically tense and relax different muscle groups.
	Moderate	Behavioral activation	Plan enjoyable activities, even if you don't feel motivated at first.
	Mild	Social connection	Take a moment to connect with friends and family, even if only briefly.

The KAI chatbot incorporates differentiated intervention protocols based on the severity level of the DASS-21 diagnosis. When the results for any of the factors (Depression, Anxiety, or Stress) fall within the normal range, the system sends a positive message to reinforce the student's well-being. Conversely, when a severe or extremely severe level is identified, the chatbot suspends self-care recommendations and instructs the student to immediately contact the University Support Center to seek professional help.

The study sample consisted of a total of 112 students from three higher education institutions: UAP, UNACH, and UPEU. The distribution was divided into two main academic areas: the Psychology program, which represented the majority at 59.82% (n=67), and the Systems Engineering or Computer Science program, at 40.18% (n=45). It is worth noting that the largest concentration of participants came from UAP, accounting for 68.75% of the total sample (Table 2).

Table 2.
Student sample.

Program	UAP	UNACH	UPEU	Total
Systems Engineering or Computer Science	10 (8.93%)	15 (13.39%)	20 (17.86%)	45 (40.18%)
Psychology	67 (59.82%)	0 (0.00%)	0 (0.00%)	67 (59.82%)
Total	77 (68.75%)	15 (13.39%)	20 (17.86%)	112 (100.0%)

A group of ten eighth-semester psychology students from the UAP validated the chatbot's functionality in terms of usability and the recommendations generated based on the DASS-21 results for each severity level of anxiety, stress, and depression. The validation of KAI took place during the month of November 2025 (15 days). The main purpose of this test was to capture as much data as possible on users' actual use of the chatbot at different times, to evaluate its performance and usefulness under real-world conditions [10].

To evaluate the impact of the KAI chatbot, both groups used the SUS [8]. The objective of the SUS scale was to determine users' perceptions regarding ease of use, efficiency, and satisfaction during their interaction with the developed tool. In terms of functionality, the system integrates modules that enable the collection, processing, and visualization of data in real time, optimizing decision-making and reducing the operational burden. Methodologically, this involves conducting controlled tests with a representative group of users who perform specific, predefined tasks; subsequently, their responses to the SUS scale are collected, and the scores obtained are analyzed to establish the system's level of acceptance and identify potential improvements.

This instrument uses a 10-item Likert scale that provides an overall measure of a system's usability. The odd-numbered items (1, 3, 5, 7, and 9) are phrased positively, so they contribute positively to the score using the formula (response - 1), while the even-numbered items (2, 4, 6, 8, and 10) are phrased negatively, contributing using the formula (5 - response). Next, the individual contributions of each item are summed, with a possible range of 0 to 40 points. Finally, the SUS score is obtained by multiplying this sum by 2.5, resulting in a scale of 0 to 100, where values above 68 indicate acceptable or good usability, and values above 80 are considered excellent. Demographic questions such as age, gender, program, and semester were also included, along with other questions to better contextualize the students' responses. Table 3 lists the questions used in the survey.

Table 3.
Chatbot usability survey

Section	No.	Item / Question	Response Type
Demographics	1	Age	Numeric
	2	Gender	Female or Male
	3	Program of Study/Major	Open
	4	Semester/Academic Year	Numeric / Multiple Choice
Context	5	How often do you use digital apps or tools to support your mental health?	Never, Rarely, Sometimes, Almost always, Always
	6	What is your overall comfort level when interacting with AI chatbots?	Likert scale (1=Very uncomfortable to 5=Very comfortable)
	7	Have you previously received formal psychological or psychiatric support?	Yes/No
	8	How many times did you interact with the KAI chatbot during the trial period?	Numeric
Usability (SUS)	9	I think I would like to use this system frequently.	Likert scale (1=Strongly disagree to 5=Strongly agree)
	10	I found the system unnecessarily complex.	
	11	I thought the system was easy to use.	
	12	I think I would need help from a technical expert to use this system.	
	13	I found that the system's various features were well integrated.	
	14	I think there were too many inconsistencies in this system.	
	15	I imagine most people would learn to use this system very quickly.	
	16	I found the system very complicated to use.	
	17	I felt very confident using the system.	
	18	I had to learn a lot before I could start using this system.	

Additionally, to validate the chatbot's performance and recommendations among the special group of psychology students, an interview was structured in which these participants answered the following questions:

- Given that the AI uses the DASS-21 assessment to generate the recommendation, is the result presented to the user perceived as truly personalized and consistent with the assessment?
- If you were giving an initial recommendation, is there any relaxation tool that the chatbot does not include?
- Is there any relaxation tool that should be replaced or modified in its description?
- At what severity level (Normal, Mild, Moderate, Severe, or Extremely Severe) would you recommend that the chatbot stop providing tools and instead refer the student to campus wellness services?
- Are any of the tool descriptions confusing or in need of modification for a non-expert user?
- Is there any essential technique for the initial management of anxiety, stress, or depression that is missing from the list?
- Do you know of any other technological application that does the same or something similar? If so, please specify it.
- Additional Comments

All questions were hosted on an online form to facilitate the consolidation and analysis of responses. Initially, a descriptive statistical analysis was conducted for each group for the post-test evaluation. This stage includes measures of central tendency, measures of dispersion, and graphs related to user behavior, the number of interactions with the platform, and the performance of the AI implemented to recommend therapies. This descriptive analysis allows us to observe usage trends and interpret the behavior of the algorithms under real-world conditions, so that functional aspects can be refined before final implementation [11]. Since SUS uses a Likert scale, the nonparametric Kruskal–Wallis test was applied to compare usability scores across the three universities. Data analysis was performed using the R programming language with the reshape2, effsize, dplyr, tidyr, psych, and ggplot2 packages.

This study strictly adhered to ethical principles for research. The research protocol was reviewed and approved by the Ethics Committee of the participating educational institutions. The confidentiality and anonymity of the information collected via digital forms were guaranteed, ensuring that sensitive data related to mental health was handled securely and used solely for academic purposes.

3. Results

The following sections present the results of the descriptive and inferential statistical analysis of the data.

3.1. Descriptive Statistical Analysis

Figure 1 shows the distribution of participants by gender at each of the universities included in the study. Unlike UNACH and UPEU, female participation predominates at UAP, while at UNACH and UPEU most participants are male.

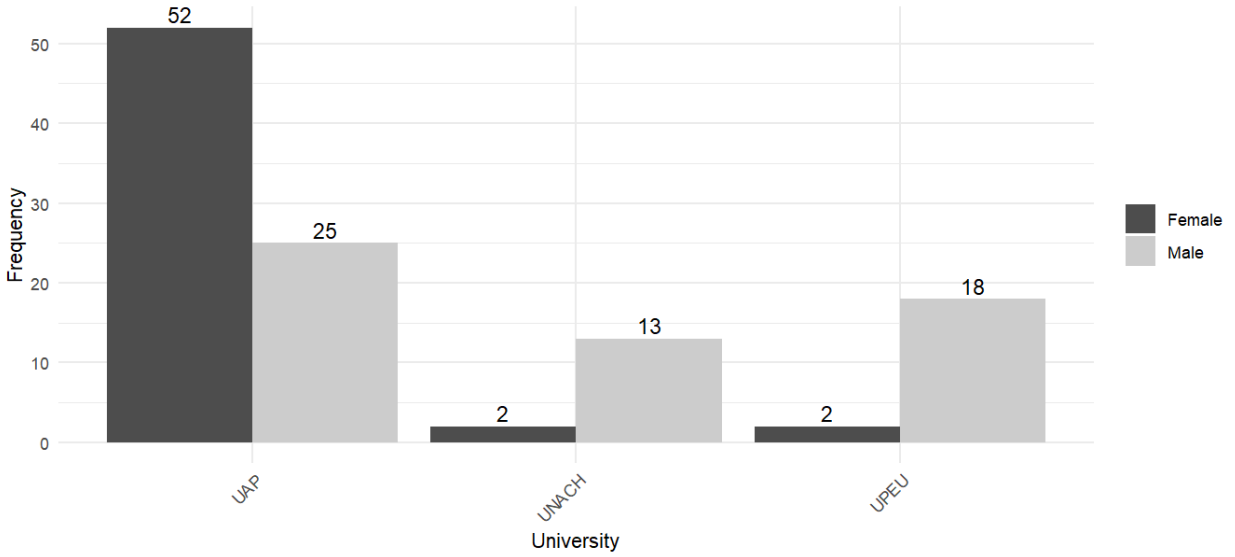

Figure 1.
Number of responses by gender per university.

Of the 77 participants from UAP, 87% (n=67) are students in the Psychology program, while the remaining 13% (n=10) are Systems Engineering students. Meanwhile, the participants from UNACH (n=15) and UPEU (n=20) are all Systems Engineering or Computer Science students. Figure 2 shows the percentage composition of the sample by academic semester at each institution.

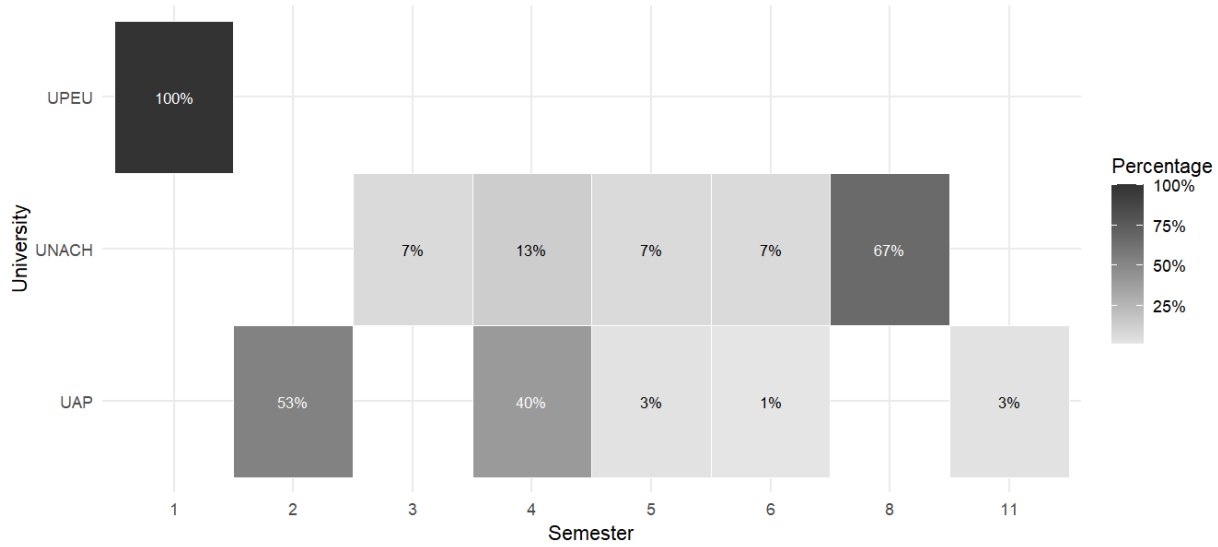

Figure 2.
Percentage of participation by academic semester by university.

Since these are higher education institutions, the participants' ages range from 17 to 31 years. Given that 100% of the UPEU participants are first-semester students, there is less variability in this group compared to the other universities, whose participants are at different levels of study. The age distribution by university is reported in Figure 3.

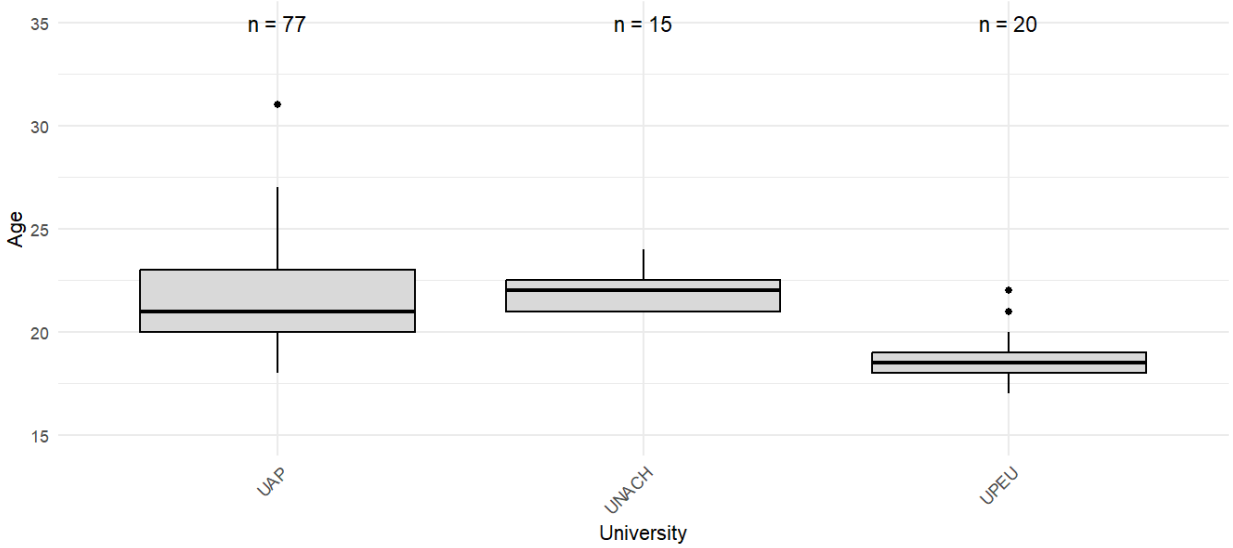

Figure 3.
Age by university.

The survey instrument was used to assess the level of comfort with interacting with the chatbot. Figure 4 presents the percentage distribution of the comfort level reported by participants when interacting with the chatbot. Fifty percent of the participants (n=56) reported feeling "Comfortable," while only 6.25% (n=7) reported feeling "Uncomfortable" or "Very uncomfortable."

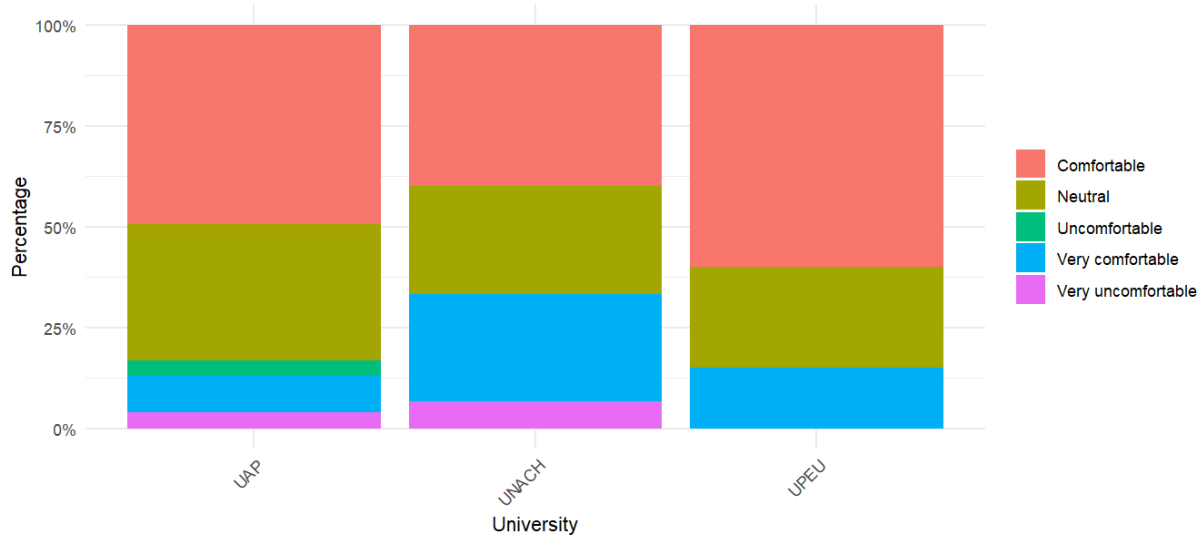

Figure 4.
Perception of comfort when using the chatbot.

Of the 112 participants, 58.9% (n = 66) reported having received psychological and/or psychiatric support, a relevant aspect given that all participants evaluated the usability of the mental health-oriented chatbot. Regarding the use of the chatbot, most participants interacted briefly with the tool during the evaluation process. 58.0% (n = 65) reported a single interaction, and 19.6% (n = 22) indicated they had interacted twice. The remaining 22.3% recorded more than two interactions, indicating that a smaller group explored the system more thoroughly before providing their usability assessment. Figure 5 shows the distribution of SUS scores (0–100) by university using box plots. For each institution, the median, interquartile range, and extreme values are presented, along with the mean and standard deviation reported at the top of each graph.

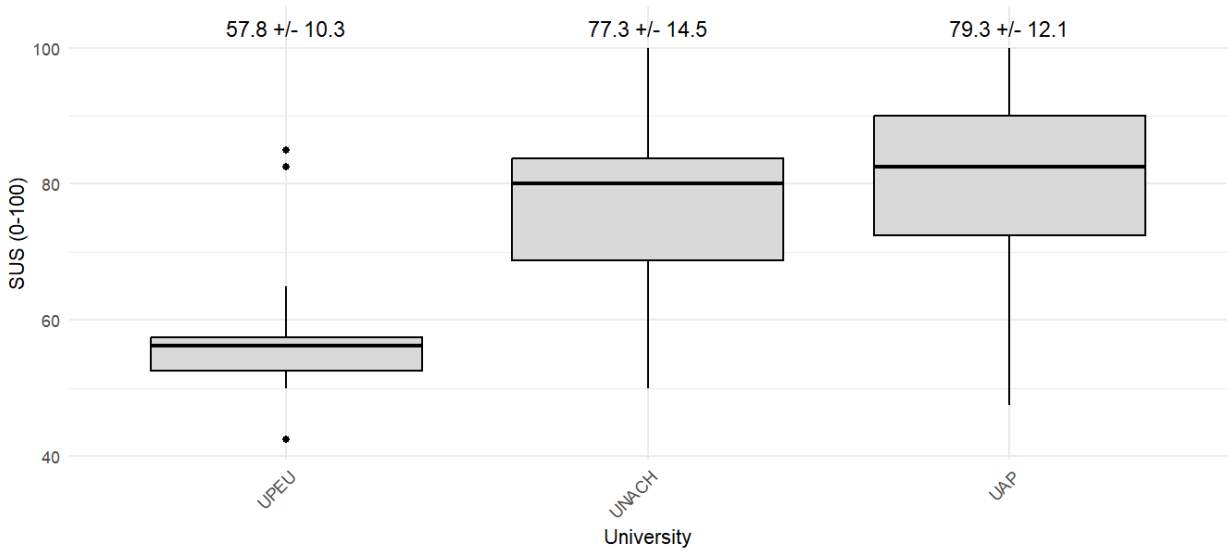

Figure 5.
SUS Scores by University.

The internal consistency of the instrument was evaluated using Cronbach's alpha coefficient, yielding a value of $\alpha = 0.783$, indicating adequate reliability of the scale in this sample. This result suggests that the instrument's items exhibit a satisfactory level of internal consistency for measuring the chatbot's usability.

Figure 6 presents the frequency distribution of responses for each of the ten items comprising the SUS scale. There is a general trend toward favorable ratings of the system in the positive items (1, 3, 5, 7, and 9), where responses at levels 4 and 5 predominate. In particular, items related to ease of use, quick learning, and confidence in use show a significant concentration in the high categories of the scale. Conversely, the negatively worded items (2, 4, 6, 8, and 10) show a higher concentration in the low categories (1 and 2), which is consistent with a reduced perception of complexity, inconsistency, or the need for technical support.

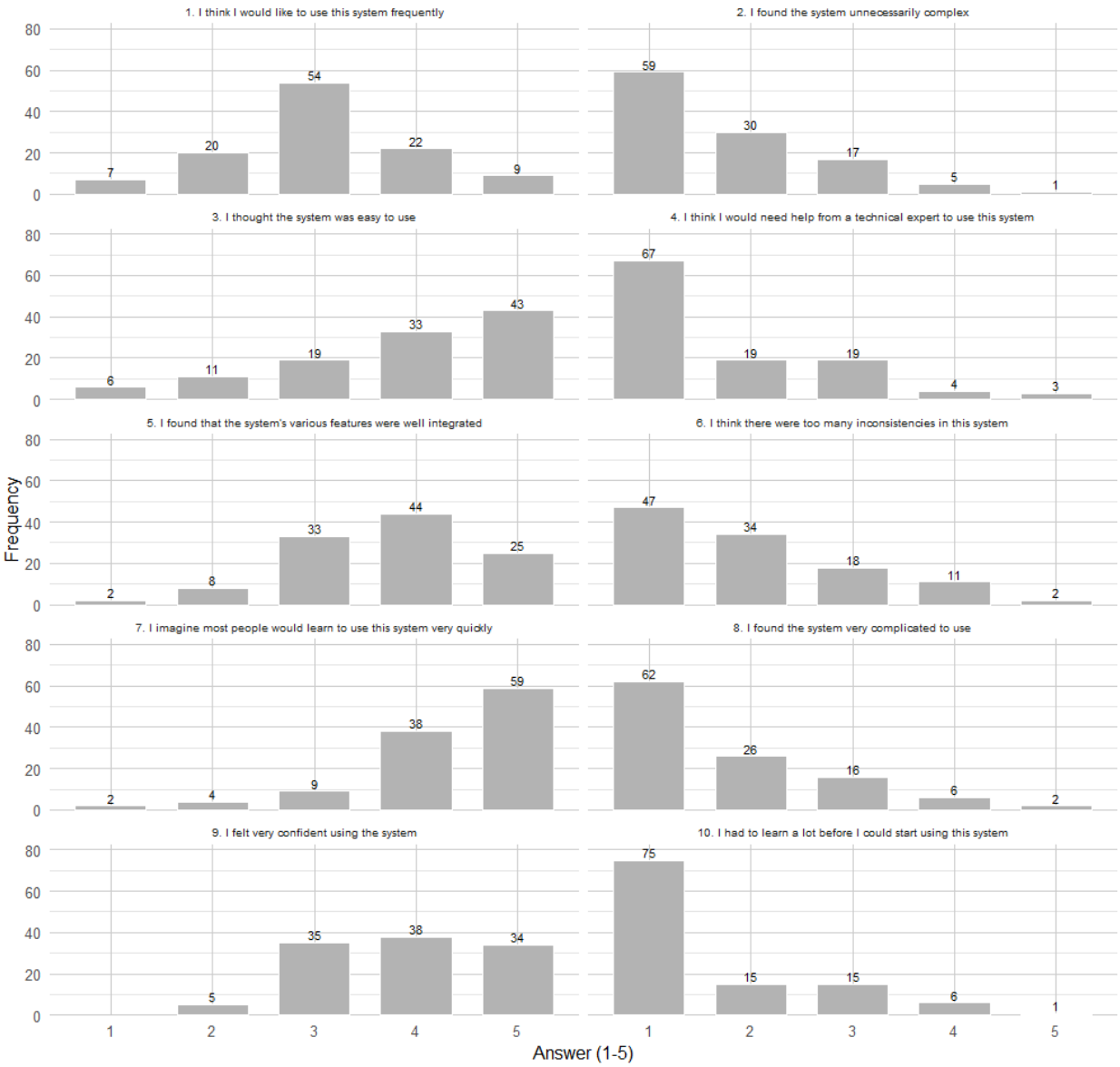

Figure 6.
Responses to each of the SUS questions.

The analysis of the responses to the open-ended questions posed to the UAP psychology focus group was conducted thematically and inductively. This type of analysis seeks similarities and differences within the dataset to identify interpretive patterns in a flexible and direct manner [12]. This form of analysis is suitable for capturing heterogeneous impressions of the same reality and opening the discussion to diverse interpretive frameworks. The flexibility of thematic analysis facilitates the identification of exceptional cases or singular data points that contrast with the interpretive patterns and thus help to understand, weigh, and make sense of them [13].

The coding process yielded four emerging themes related to the chatbot's usability. To ensure quality, verbatim quotes were used to support the interpretation of the findings and ensure the representativeness of the voices of the trainee experts.

The first theme concerns the consistency and personalization of the diagnosis. This theme addresses the question: Is the result presented to the user perceived as truly personalized and consistent with the DASS-21 diagnosis? Regarding this theme, participants expressed mixed perceptions about the AI's ability to reflect their emotional state: *"Actually, yes, it is consistent with each person's responses." "It feels pretty accurate." "I feel like it's very easy to manipulate the data if you know what's being asked; it doesn't help provide real insight." "Sometimes the recommendations are very general or generic."*

The second theme concerns the relevance and adequacy of the relaxation tools recommended by the chatbot. This topic addresses the questions: Are there any tools that KAI doesn't include? and Is there any essential tool missing for managing anxiety, stress, or depression? The students identified technical gaps and proposed additions to the chatbot's repertoire: *"There should be a few more for depression since I felt that almost all of them were for anxiety." "I would recommend the 4-step breathing technique." "I would include techniques such as walking, resting, words of encouragement, or talking to someone close to you." "I would recommend a more tactful approach when presenting the results, as they can sound somewhat alarming."*

The third topic addresses the thresholds for safety and professional referral. This topic addresses the question: At what level of severity would you recommend that the chatbot stop providing tools and instead refer the user to campus wellness services? The students' responses reflect a consensus on the limits of automated intervention: *"Starting at the moderate level, it should stop providing tools so that the professional can take the initiative." "At severe and extremely severe levels, mandatory referral is necessary." "A person with depression or anxiety sometimes does not seek help on their own; the bot should facilitate that direct contact." "I believe that at the moderate level, referral to university wellness services should already occur."*

The final topic concerns the evaluation of the interface and the user experience within the system. This topic addresses questions regarding: relaxation tools that need to be replaced, and additional comments. Students' comments focused on ease of use and barriers to access: *"The fact that it asks you to create a profile or Google account right from the start discourages you from using it." "It's innovative, and I think it could be very useful for many people." "The app 'throws' you to the end of the message; there should be an easier way to go back to the beginning." "I suggest that the questions be more specific to the current moment or the current week."*

3.2. Inferential Statistical Analysis

The nonparametric Kruskal–Wallis test was performed to compare usability scores among the three universities. The results indicate statistically significant differences between the groups,$H(2) = 30.1$,$p < .001$. The effect size was large ($\eta_H^2 = 0.258$), indicating that approximately 25.8% of the variability in usability scores is associated with the university of origin.

Post hoc analysis using the Dunn test with Holm correction showed that UPEU had significantly lower scores than UAP ($p < .001$) and UNACH ($p < .001$). No significant differences were found between UAP and UNACH ($p = .537$). Thus, the results indicate that the perceived usability of the system is significantly lower at UPEU compared to the other two universities, which show similar levels of perceived usability.

4. Discussion

The main objective of this research was to evaluate the use and utility of KAI as a generative AI-based chatbot for mental health support among university students in Latin America. Through the implementation, it was determined that the tool is not only technologically viable but that students perceive it as an accessible resource for emotional self-management. The integration of structured protocols allowed for a positive response in user-machine interaction, thus fulfilling the purpose of identifying how these technologies can act as a first level of care or triage within the university well-being ecosystem.

A notable finding of this research was the discrepancy between the general use of AI and its specific application in mental health. Although participants reported a high frequency of using tools such as ChatGPT for academic purposes, its use for emotional support was significantly lower. This phenomenon is closely related to the findings of Rackoff, et al. [4] who reported that, although nearly half of college students use chatbots, only a small percentage (5–8%) use them to seek psychological help. This suggests that, in both the Latin American and U.S. contexts, skepticism persists regarding the therapeutic capacity of AI compared to traditional psychotherapy, posing a challenge for large-scale implementation.

The research also revealed that the main benefits perceived by students are anonymity and immediate availability, which reduce the barriers of social stigma. In this regard, the data aligns with the concept of "pocket doctors" proposed by Mina, et al. [6] who argue that in societies where mental health or certain health issues are taboo, AI offers a safe and judgment-free space. This perception is fundamental, since, as Reichenpfader, et al. [2] indicate, a chatbot's ability to provide evidence-based responses and 24/7 access is a critical requirement for students to trust these digital tools as a viable complement to university counseling services.

Additionally, a positive impact was observed on the perception of social support and a reduction in feelings of loneliness among frequent users. This finding aligns with the study by Kim, et al. [14] on the therapeutic potential of social chatbots to alleviate social anxiety and loneliness through empathetic communication. Furthermore, the effectiveness of these interventions in this study appears to be mediated by the user's context; as Wang, et al. [1] note, factors such as financial stress can moderate the effectiveness of AI-based therapy, underscoring the need for college chatbots to be sensitive to the socioeconomic realities of the region.

Despite the benefits observed in managing mild and moderate symptoms, it is important to emphasize that generative AI chatbots are not a substitute for professional human intervention in scenarios of acute crisis or severe suicide risk. This finding is consistent with the warnings of Reichenpfader, et al. [2] who identify the management of emergency responses as a non-negotiable technical and ethical requirement for any digital mental health tool aimed at students. The literature agrees that, while AI can act as a 'pocket doctor' for wellness consultations [6] it lacks the clinical judgment necessary to conduct a lethal risk assessment. In this regard, our results align with the position of Reyes-Portillo, et al. [15] who suggest that the real utility of these tools lies in their function as triage systems or complementary support, where the programming of robust safety protocols must ensure immediate referral to human emergency services upon the detection of suicidal ideation.

Regarding perceived usability, the results indicate significant variations depending on the students' academic program. It was observed that students in technical and engineering fields reported greater ease of use and more intuitive navigation compared to their peers in the social sciences and arts. This trend is consistent with the findings of Reichenpfader, et al. [2] who note that prior familiarity with technological interfaces and AI literacy are critical prerequisites that shape the user

experience. In this regard, usability depends not only on interface design but also on the digital competence inherent to the student's academic profile, which facilitates a smoother adoption in fields accustomed to using algorithmic tools.

On the other hand, a high perception of usefulness was observed among Psychology and Systems students, who highlighted the system's 24/7 availability and the immediacy of its responses. When comparing these findings with the study by Mina, et al. [6] it is confirmed that students in disciplines with a heavy academic workload prioritize the chatbot's ability to offer anonymous support and structured responses during times of crisis. This suggests that, although technical usability is higher among students in technology-related fields, pragmatic usability—understood as the effectiveness in meeting a specific need—is common across different programs. Such effectiveness depends on the tool integrating evidence-based protocols and empathetic communication, aspects also highlighted by Kim, et al. [14] in their analysis of the potential of social chatbots.

Although the results of this study are promising, some limitations are also evident. One of them lies in the exclusive use of self-report instruments to measure complex variables such as perceived utility and emotional well-being. These measures are susceptible to social desirability bias, where students may have reported a more positive perception to meet the study's implicit expectations. Likewise, the concept of "usability" may have been influenced by the participants' prior digital literacy, making it difficult to isolate whether acceptance was due to the chatbot's intrinsic characteristics or the user's technological competence.

Furthermore, due to the scope of this study, which was more oriented toward the descriptive analysis of the collected data, it was not possible to establish a causal relationship between the use of the chatbot and changes in students' mood. The lack of a randomized control group prevents us from ruling out the influence of extraneous variables, such as changes in academic workload, external social support, or seasonal events on the university calendar, which could have affected the results independently of the technological intervention. This is an aspect considered for future work.

The representativeness of the findings is limited by the size and nature of the sample, which focused on students in specific programs (Psychology and Systems) within a particular geographical context in Latin America. Consequently, the results cannot be generalized to the global university population, to students in other disciplinary areas with different stress levels, or to individuals outside the higher education system.

Additionally, the cross-sectional nature of the study limits the observation of long-term effects, preventing us from determining whether the perceived utility and reduction in symptoms are sustainable over time or whether they result from a novelty effect due to the use of generative AI. Furthermore, the sample size may not have been sufficient to identify subtle statistical differences between subgroups, suggesting that the conclusions should be interpreted as preliminary trends requiring validation in larger-scale longitudinal studies.

5. Conclusions and Future Work

This research demonstrates that the implementation of the KAI chatbot, based on generative AI, constitutes a viable and highly acceptable technological tool for mental health support in the university setting in Latin America. The results obtained through the SUS scale, with usability rated as outstanding and reliable internal consistency (alpha = 0.783), confirm that the system is intuitive, secure, and easy for students to learn. By integrating the DASS-21 psychometric assessment with personalized recommendations, the chatbot effectively functions as a first-contact or triage resource, enabling the timely identification of symptoms of stress, anxiety, and depression.

Likewise, the study highlights the importance of automated safety protocols, which suspend the chatbot's intervention and mandatorily refer students to professional services when critical severity levels are detected. However, the significant variability in usability perception across different institutions suggests that contextual factors, such as the level of training and the digital culture of each academic program, influence the adoption of these tools. In conclusion, although generative AI does not replace traditional psychotherapy, it represents a scalable, low-cost complement to strengthen student well-being strategies.

As future lines of research, we identify the need to expand the repertoire of emotional intervention techniques, incorporating a greater number of specific tools for managing depression and refining their descriptions to improve diagnostic clarity. It is essential to explore mechanisms that reduce initial access barriers, such as the requirement for Google accounts, which may discourage anonymous and confidential use of the platform. Additionally, it is suggested to implement features that allow for smoother bidirectional navigation, making it easier for users to return to the system's home screen after receiving recommendations.

Finally, future research should focus on longitudinal studies to assess the chatbot's actual impact on reducing clinical symptoms in the long term. It is also recommended to integrate more advanced natural language processing (NLP) capabilities that allow the system to proactively detect crisis risks during free-form interaction, and not just through structured tests. Expanding this model to other geographic areas and academic disciplines will allow for the validation of its robustness and the adaptation of content to the specific needs of diverse university populations.

References

[1] Y. Wang, X. Li, Q. Zhang, D. Yeung, and Y. Wu, "Effect of a cognitive behavioral therapy–based AI chatbot on depression and loneliness in chinese university students: Randomized controlled trial with financial stress moderation," *JMIR mHealth and uHealth*, vol. 13, p. e63806, 2025, https://doi.org/10.2196/63806.

[2] D. Reichenpfader, L. Affolter, E. Gabarron, and K. Denecke, *Opportunities, requirements, and risks of a student mental health chatbot: A qualitative user-centered, multi-method approach. In Healthcare of the Future 2025: Redefining Healthcare Delivery in the Digital Era (pp. 67-72)*. 1 Oliver's Yard, 55 City Road, London, EC1Y 1SP: SAGE Publications Pvt. Ltd, 2025.

[3] D. Akiba, "ChatGPT told me to say it: AI chatbots and class participation apprehension in university students," *Education Sciences*, vol. 15, no. 7, p. 897, 2025, https://doi.org/10.3390/educsci15070897.

[4] G. N. Rackoff, Z. Z. Zhang, and M. G. Newman, "Chatbot-delivered mental health support: Attitudes and utilization in a sample of US college students," *Digital Health*, vol. 11, p. 20552076241313401, 2025, https://doi.org/10.1177/20552076241313401.

[5] A. P. Citoler and E. Lee, *End users' perception on an ai chatbot in a snus cessation mobile application. In MEDINFO 2025—Healthcare Smart× Medicine Deep*. Amsterdam, The Netherlands: IOS Press, 2025.

[6] A. Mina *et al.*, "AI chatbots as 'pocket doctors': Intimate health support for young women in Lebanon," *BMC Public Health*, vol. 25, no. 1, p. 3989, 2025, https://doi.org/10.1186/s12889-025-25386-1.

[7] M. Kian *et al.*, "Engagement and disclosures in LLM-Powered cognitive behavioral therapy exercises: A factorial design comparing the influence of a robot vs. chatbot over time," in *2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) (pp. 1173-1180). IEEE*, 2025.

[8] J. Brooke, *The system usability scale: A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), Usability Evaluation in Industry*. London, UK: Taylor & Francis, 1996.

[9] P. F. Lovibond and S. H. Lovibond, "The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories," *Behaviour research and therapy*, vol. 33, no. 3, pp. 335-343, 1995, https://doi.org/10.1016/0005-7967(94)00075-U.

[10] N. Singh, D. Patel, A. Raj, and M. S. K. Shubham, "CI/CD pipeline for web applications," *Int. J. Res. Appl. Sci. Eng. Technol*, vol. 11, no. 5, pp. 5218-5226, 2023, https://doi.org/10.22214/ijraset.2023.52867.

[11] V. Vangala, "Blue-green and canary deployments in DevOps: A comparative study," *International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence*, vol. 15, no. 1, pp. 1047-1063, 2024, https://doi.org/10.5281/zenodo.15483641.

[12] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, no. 2, pp. 77-101, 2006, https://doi.org/10.1191/1478088706qp063oa.

[13] V. Pérez-Samaniego, E. López-Cañada, and J. Monforte, "Physical activity and disability: A qualitative study with women in an adapted gym," *Movimento (ESEFID/UFRGS)*, vol. 23, no. 3, 2017, https://doi.org/10.22456/1982-8918.68973.

[14] M. Kim *et al.*, "Therapeutic potential of social chatbots in alleviating loneliness and social anxiety: quasi-experimental mixed methods study," *Journal of medical Internet research*, vol. 27, p. e65589, 2025, https://doi.org/10.2196/65589.

[15] J. A. Reyes-Portillo *et al.*, "Generative AI–powered mental wellness chatbot for college student mental wellness: Open trial," *JMIR Formative Research*, vol. 9, no. 1, p. e71923, 2025, https://doi.org/10.2196/71923.