


# When Ai Joins The Counseling Room: Students' Trust, Disclosure, And Therapeutic Alliance In Ai-Assisted Counseling

Neni Noviza\*

Universitas Islam Negeri Raden Fatah Palembang, Indonesia

[neninoviza\\_uin@radenfatah.ac.id](mailto:neninoviza_uin@radenfatah.ac.id)\*

<p><b>Submitted:</b> 2026-03-30</p> <p><b>Revised:</b> 2026-04-10</p> <p><b>Published:</b> 2026-04-21</p> <p><b>Keywords:</b> AI-Assisted Counseling, Self-Disclosure, Students' Trust, Therapeutic Alliance</p> <p><b>Copyright holder:</b> © Author/s (2026)</p> <p><b>This article is under:</b></p>  <p><b>How to cite:</b> Noviza, N. (2026). When AI Joins the Counseling Room: Students' Trust, Disclosure, and Therapeutic Alliance in AI-Assisted Counseling. <i>Bulletin of Counseling and Psychotherapy</i>, 8(2). <a href="https://doi.org/10.51214/002026081884000">https://doi.org/10.51214/002026081884000</a></p> <p><b>Published by:</b> Kuras Institute</p> <p><b>E-ISSN:</b> 2656-1050</p>	<p><b>ABSTRACT:</b> The growing use of artificial intelligence in psychological support has positioned AI-assisted counseling as a promising approach to expand access, reduce stigma, and provide immediate support for students. However, counseling is fundamentally relational, requiring trust, self-disclosure, and therapeutic alliance to function effectively. This study investigated these three dimensions and their interrelationships in AI-assisted counseling among Indonesian university students. A cross-sectional field survey was conducted with 1,284 undergraduate students from 12 universities across Indonesia. Data were collected using demographic questions, the Trust in AI-Assisted Counseling Scale (8 items), the Self-Disclosure in AI-Assisted Counseling Scale (7 items), and the Therapeutic Alliance in AI-Assisted Counseling Scale (9 items), all rated on a 5-point Likert scale. The instruments demonstrated good reliability, with Cronbach's alpha values of 0.88, 0.86, and 0.90, and were reviewed by three experts and pilot-tested with 92 students. Data were analyzed using descriptive statistics, Pearson correlation, and multiple regression. Students reported moderately positive levels of trust (M = 3.68, SD = 0.64), self-disclosure (M = 3.39, SD = 0.71), and therapeutic alliance (M = 3.57, SD = 0.62). Trust was positively associated with self-disclosure and therapeutic alliance, and both variables jointly predicted therapeutic alliance.</p>
---	---

## INTRODUCTION

The development of artificial intelligence (AI) has brought major changes across various fields, including education, healthcare, and psychological support (Chen, 2024). In the context of mental health, AI-based systems such as chatbots, virtual assistants, and generative conversational platforms are increasingly viewed as tools capable of expanding access to services (H. Li, Zhang, et al., 2023), providing immediate initial support, and helping to overcome longstanding barriers such as stigma, cost, waiting time, and the limited availability of professionals. A growing body of research suggests that mental health chatbots have the potential to support psychoeducation, self-reflection, emotion monitoring, and early intervention, particularly for adolescents and young adults who are familiar with digital technology (Haque & Rubya, 2023; Koulouri et al., 2022; Mitsea & Drigas, 2023).

Nevertheless, the presence of AI in counseling cannot be understood merely as a technical innovation (Nazir, 2026; Ping, 2024). Counseling is fundamentally a relational process grounded in the quality of the helping relationship, psychological safety, trust, and the client's willingness to disclose personal concerns (Podolan & Gelo, 2024). Therefore, a central question in the

development of AI-assisted counseling is not only whether AI can provide fast and accurate responses, but also whether it can be accepted within the relational dimension that lies at the core of counseling practice. The digital mental health literature shows that the effectiveness of digital interventions is determined not solely by technological features, but also by the quality of the relationship formed between users and the system, including what has been referred to as the digital therapeutic alliance (Lotringen et al., 2021; Tong et al., 2022; Tremain et al., 2020).

In the traditions of counseling and psychotherapy, trust, disclosure, and therapeutic alliance are three closely interconnected elements (Krause et al., 2011). Trust serves as the foundation that allows clients to feel psychologically safe (Hunt et al., 2021; Seager, 2007). Self-disclosure enables problems, emotions, and personal experiences to be expressed more honestly (Wang et al., 2017). Therapeutic alliance, in turn, reflects the quality of collaboration, relational bonding, and shared understanding of the goals of help (Wampold & Flückiger, 2023). In AI-based services, these three elements become even more important because users interact with systems that do not possess full human presence. Several studies have shown that mental health chatbots can evoke perceptions of support and comfort, yet the quality of this experience is strongly shaped by users' perceptions of empathy, safety, responsiveness, and system reliability (Chaudhry & Debi, 2024; Jung et al., 2025; Shen et al., 2024; Wang, 2025).

This issue is particularly relevant among university students. As a group that intensively uses digital technology, students also face academic pressure, identity-related concerns, complex social relationships, and an increasing need for psychological support. Previous studies indicate that students perceive mental health chatbots as more accessible and less hindered by stigma than traditional services (Almaskari et al., 2025; Hoffman et al., 2024; Lawati et al., 2026; Miqdadi et al., 2025; Rackoff et al., 2025). However, they also express concerns regarding privacy, response accuracy, emotional sensitivity, and AI's capacity to genuinely understand human problems. This suggests that the acceptance of AI in mental health services is ambivalent: it offers accessibility, yet still raises important questions about the quality of the helping relationship that can be established.

To date, most research on AI in mental health has focused on acceptance, usability, the general effectiveness of chatbots, or broad ethical and clinical evaluations. In contrast, studies that specifically examine the relational dynamics of AI-assisted counseling, particularly the relationships among trust, disclosure, and therapeutic alliance among university students, remain limited. Yet university students constitute a highly relevant population because they stand at the intersection of psychological support needs, high digital familiarity, and changing help-seeking patterns in the AI era.

Against this backdrop, the present study offers novelty by positioning the relational dimension as the primary focus rather than merely examining technology acceptance. This study also simultaneously investigates the relationships among students' trust in AI-assisted counseling, self-disclosure, and perceptions of therapeutic alliance using a large field-based sample drawn from multiple universities across Indonesia. Accordingly, this study aims to: (1) describe students' levels of trust, self-disclosure, and therapeutic alliance in the context of AI-assisted counseling; (2) examine the relationships among these three variables; and (3) analyze the contribution of trust and self-disclosure in predicting therapeutic alliance. This study proposes three hypotheses: H1, trust is positively associated with self-disclosure; H2, trust is positively associated with therapeutic alliance; and H3, trust and self-disclosure jointly predict therapeutic alliance.

## **METHODS**

### **Research Design**

This study employed a cross-sectional field survey design to examine students' trust, self-disclosure, and therapeutic alliance in the context of AI-assisted counseling (Alqurashi, 2025; Spector, 2019). A cross-sectional approach was chosen because the study aimed to capture

students' perceptions, attitudes, and behavioral tendencies at a specific point in time regarding the use of artificial intelligence in counseling interactions.

The study was conducted as a multi-site field study involving students from public and private universities across various regions of Indonesia. This design was selected to enhance sample representativeness and to provide a broader picture of how Indonesian university students respond to the presence of AI in counseling services within higher education settings.

The study focused on three main variables: students' trust in AI-assisted counseling, students' self-disclosure in AI-assisted counseling, and students' perceptions of therapeutic alliance in AI-assisted counseling settings. In addition, demographic data and prior experience with AI-based applications were collected to provide contextual background for the findings. To help readers understand the overall flow of the study, Figure 1 presents a summary of the research design, participant characteristics, instruments, data collection procedures, data analysis techniques, ethical considerations, and inclusion criteria used in this study.

### Participants

The participants in this study were 1,284 undergraduate students recruited from 12 universities in Indonesia, consisting of public universities, Islamic public universities, and private universities. The universities were selected to represent variation in institutional type, geographic region, and student background. Participants were drawn from universities located in Sumatra, Java, Kalimantan, Sulawesi, and Bali–Nusa Tenggara, thereby providing a more nationally representative picture.

The inclusion criteria were as follows: (1) being an active undergraduate student at one of the participating universities; (2) being between 18 and 25 years old; (3) willingness to participate voluntarily; and (4) having at least basic exposure to digital applications, including prior use of or familiarity with AI-based platforms such as chatbots, virtual assistants, or generative AI. Respondents who submitted incomplete questionnaires or failed the response consistency screening were excluded from the final analysis.

Of the 1,352 responses initially received, 68 responses were excluded due to incomplete data, uniform response patterns, or failure to meet the inclusion criteria. Thus, the final valid sample analyzed in this study consisted of 1,284 responses. The final sample included 487 male students (37.9%) and 797 female students (62.1%). The average age of participants was 20.41 years ( $SD = 1.62$ ), ranging from 18 to 25 years. In terms of year of study, participants ranged from first-year to fourth-year students, with the largest proportions coming from second- and third-year students. Regarding AI experience, 1,109 students (86.4%) reported having used AI-based applications, while 175 students (13.6%) reported limited or no direct experience.

Table 1. Distribution of Participants by University

No.	University	Region	Number of Participants	Percentage
1	Universitas Sriwijaya	Sumatra	126	9.8
2	Universitas Islam negeri Raden Fatah Palembang	Sumatra	118	9.2
3	Universitas PGRI Palembang	Sumatra	94	7.3
4	Universitas Negeri Semarang	Java	168	13.1
5	Universitas Negeri Malang	Java	142	11.1
6	Universitas PGRI Kanjuruhan Malang	Java	136	10.6
7	Universitas Mulawarman	Kalimantan	96	7.5
8	Universitas Lambung Mangkurat	Kalimantan	82	6.4
9	Universitas Negeri Makassar	Sulawesi	104	8.1

10	Institut Agama Islam Negeri Kendari	Sulawesi	88	6.9
11	Universitas PGRI Mahadewa Indonesia	Bali–Nusa Tenggara	64	5.0
12	Universitas Pendidikan Ganesha	Bali–Nusa Tenggara	66	5.1
Total			1,284	100.0

Source: Authors' own data based on a field survey of 1,284 undergraduate students from 12 universities in Indonesia.

Table 2. Demographic Characteristics of Participants

Variable	Category	N	%
Gender	Male	487	37.9
	Female	797	62.1
Age	18–19 years	356	27.7
	20–21 years	601	46.8
	22–23 years	267	20.8
	24–25 years	60	4.7
Year of study	First year	251	19.5
	Second year	382	29.8
	Third year	401	31.2
	Fourth year or above	250	19.5
Experience using AI	Yes	1,109	86.4
	No/Limited	175	13.6

Source: Authors' own data based on questionnaire responses from the study participants.

## Instruments

Data were collected using a structured self-report questionnaire consisting of four sections: demographic information, the Trust in AI-Assisted Counseling Scale, the Self-Disclosure in AI-Assisted Counseling Scale, and the Therapeutic Alliance in AI-Assisted Counseling Scale (Boateng et al., 2018).

## Demographic Information

The first section gathered participants' background information, including age, gender, year of study, type of university, university region, and previous experience with AI-based applications.

## Trust in AI-Assisted Counseling Scale

Students' trust in AI-assisted counseling was measured using an 8-item scale developed in this study based on conceptual dimensions including perceived reliability, psychological safety, responsiveness, and confidence in AI-supported counseling interactions. Example items included: *"I believe that AI-assisted counseling can provide useful initial support for students' personal problems"* and *"I feel that AI-based counseling tools can be trusted to provide consistent responses."* Each item was rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores indicated greater trust in AI-assisted counseling. In this study, the scale demonstrated good internal consistency, with a Cronbach's alpha of 0.88.

## Self-Disclosure in AI-Assisted Counseling Scale

Students' willingness to disclose their thoughts, feelings, and personal problems in AI-assisted counseling was measured using a 7-item scale. The items emphasized aspects of comfort, openness, hesitation, and perceived emotional risk when sharing personal concerns through AI-based

counseling systems. Example items included: *“I feel comfortable sharing personal concerns through an AI-assisted counseling platform”* and *“I am willing to discuss emotional difficulties through AI-assisted counseling services.”* Responses were recorded using the same 5-point Likert scale. Negatively worded items were reverse-scored prior to analysis. Higher scores indicated greater self-disclosure. The scale demonstrated good reliability, with a Cronbach’s alpha of 0.86.

### Therapeutic Alliance in AI-Assisted Counseling Scale

Students’ perceptions of therapeutic alliance were measured using a 9-item scale adapted to fit the context of AI-assisted counseling. This scale assessed perceptions of collaboration, emotional support, shared engagement, and alignment of tasks and goals in counseling interactions involving AI. Example items included: *“AI-assisted counseling can help me feel understood during the early stages of help-seeking”* and *“I feel that AI-assisted counseling can support a meaningful helping relationship when used appropriately.”* All items were rated on a 5-point Likert scale. Higher scores reflected more positive perceptions of therapeutic alliance in AI-assisted counseling. In this study, the scale demonstrated strong internal consistency, with a Cronbach’s alpha of 0.90. Before being used in the main data collection, all instruments were reviewed by three experts in counseling, educational psychology, and digital mental health to ensure content relevance and clarity of wording. In addition, a pilot test involving 92 students from a university not included in the main sample was conducted to refine item wording, improve readability, and assess initial reliability.

Table 3. Structure of the Research Instruments

Variable	Number of Items	Indicators	Scale	Cronbach’s Alpha
Trust in AI-Assisted Counseling	8	Reliability, confidence, psychological safety	Likert 1–5	0.88
Self-Disclosure in AI-Assisted Counseling	7	Openness, comfort, willingness to share	Likert 1–5	0.86
Therapeutic Alliance in AI-Assisted Counseling	9	Support, collaboration, feeling understood	Likert 1–5	0.90
Total	24			

Source: Developed by the authors based on the study instrument design and pilot testing results.

### Procedure

Data were collected between August and October 2025. After obtaining institutional permission from the participating universities, the researchers coordinated with lecturers, program administrators, student organizations, and academic staff to distribute the questionnaires to eligible students. The questionnaires were administered both online and offline to increase participation coverage. The online version was distributed through institutional communication channels such as academic email, class groups, and student organization networks. Offline data collection was conducted with the assistance of research assistants who visited selected classes and student activity centers to invite students to participate.

Before completing the questionnaire, all participants were given an informed consent form explaining the purpose of the study, the voluntary nature of participation, the protection of confidentiality, and their right to withdraw at any time without penalty. Participants who agreed to participate then proceeded to complete the questionnaire, which took approximately 15–30 minutes. To ensure data quality, several screening procedures were applied. First, incomplete responses were automatically flagged. Second, uniform response patterns across all items were examined to identify careless responding. Third, excessively short completion times that indicated

insufficient attention were also reviewed. Only data meeting the quality standards were included in the final analysis.

### Data Analysis

The data were analyzed using IBM SPSS Statistics. The analysis was conducted in several stages. First, descriptive statistics were used to describe participants' demographic characteristics and the general tendency of the three main study variables. These analyses included means, standard deviations, frequencies, and percentages. Second, the psychometric quality of the instruments was assessed through item-total correlations, Cronbach's alpha coefficients, and inter-construct correlations. In addition, statistical assumptions such as normality, linearity, and multicollinearity were examined before conducting inferential analyses.

Third, Pearson correlation analysis was used to examine the relationships among trust, self-disclosure, and therapeutic alliance. Fourth, multiple regression analysis was conducted to determine the predictive role of trust and self-disclosure in students' perceptions of therapeutic alliance in AI-assisted counseling. Additional analyses, including independent-samples t-tests and one-way ANOVA, were also used to explore differences based on gender, year of study, and experience using AI when necessary. The level of statistical significance was set at  $p < 0.05$ . In addition to statistical significance, effect sizes were also considered to provide more substantively meaningful interpretations.

Table 4. Data Analysis Framework

Research Objective	Variables	Analysis Technique
To Describe Students' Trust, Self-Disclosure, and Therapeutic Alliance	All main variables	Descriptive statistics
To Test the Internal Consistency of the Instruments	All instrument scales	Cronbach's alpha, item-total correlation
To Examine Relationships Among the Main Variables	Trust, Self-Disclosure, Therapeutic Alliance	Pearson correlation
To Test the Effect of Trust and Self-Disclosure on Therapeutic Alliance	Trust and Self-Disclosure → Therapeutic Alliance	Multiple regression
To Examine Differences Based on Gender, Year of Study, and AI Use	Demographic variables and main variables	t-test, ANOVA

Source: Developed by the authors based on the research objectives and statistical analysis plan.

### Ethical Considerations

This study was conducted in accordance with ethical principles for research involving human participants (Association, 2025). Ethical approval was obtained from the Research Ethics Committee of Universitas Islam Negeri Raden Fatah Palembang under approval number B-501/Un.09/V.1/KS.02/03/2026. The study was carried out based on the principles of voluntary participation, confidentiality, anonymity, and respect for participant autonomy.

Before participating, students were provided with sufficient information regarding the purpose of the study, the duration of completion, data protection procedures, and their right to withdraw at any time without consequence. No directly identifying personal information was collected in the final dataset. All data were stored securely and were accessible only to the research team. Because the topic of the study involved perceptions of counseling and emotional disclosure, participants were also informed that the questionnaire did not constitute a psychological intervention or formal counseling service. Information about student counseling services was made available in case participation caused discomfort or prompted the need for further assistance.



Figure 1. Research Methods Framework for AI-Assisted Counseling

## RESULTS AND DISCUSSION

### Results

A total of 1,284 valid responses were analyzed to describe students' trust in AI-assisted counseling, self-disclosure in the context of AI-assisted counseling, and perceptions of therapeutic alliance. Overall, the findings indicate that students tended to hold moderately positive attitudes toward the use of AI in counseling, although variation remained across the three variables (Almaskari et al., 2025).

### *Students' Trust in AI-Assisted Counseling*

Descriptive analysis showed that students' level of trust in AI-assisted counseling fell within the moderately high category. The mean trust score was 3.68 (SD = 0.64), indicating that most students viewed AI as a potentially useful initial means of supporting the counseling process (Hawke et al., 2025; Li, Lei, et al., 2023)

### *Self-Disclosure in the Context of AI-Assisted Counseling*

Compared with trust, students' self-disclosure in AI-assisted counseling showed a more moderate mean score of 3.39 (SD = 0.71). This indicates that students tended to be fairly open, yet still remained cautious when revealing personal problems, emotional experiences, or psychological vulnerabilities through AI-mediated systems (Wang et al., 2017).

### *Therapeutic Alliance in AI-Assisted Counseling*

Students' perceptions of therapeutic alliance in AI-assisted counseling were moderately positive, with a mean score of 3.57 (SD = 0.62). These results indicate that students are beginning to accept the possibility of forming a meaningful helping relationship in services involving AI, particularly in the early stages of (Alqurashi, 2025; Amin et al., 2024).

Table 5. Descriptive Statistics of the Study Variables

Variable	n	Min	Max	Mean	SD	General Category
Trust in AI-Assisted Counseling	1,284	1.63	5.00	3.68	0.64	Moderately high
Self-Disclosure in AI-Assisted Counseling	1,284	1.29	5.00	3.39	0.71	Moderate to moderately high
Therapeutic Alliance in AI-Assisted Counseling	1,284	1.56	5.00	3.57	0.62	Moderately high

Source: Authors' own calculations based on survey data analyzed using IBM SPSS Statistics.

To provide a more detailed picture, indicator-level analysis showed that the highest score on the trust variable appeared in the aspect of accessibility and rapid response, whereas the lowest score appeared in confidence in AI's emotional sensitivity. For the self-disclosure variable, students tended to be more willing to discuss academic problems and mild stress than highly personal or traumatic issues. For the therapeutic alliance variable, the highest score was found in initial support and conversational structure, whereas the lowest score appeared in feeling deeply understood (Almaskari et al., 2025).

Table 6. Mean Scores by Main Indicators

Variable	Indicator	Mean	SD
Trust	Response reliability	3.74	0.69
	Accessibility and ease of use	3.88	0.66
	Psychological safety	3.61	0.72
	Sensitivity to emotional conditions	3.49	0.75
Self-Disclosure	Willingness to share academic problems	3.65	0.73
	Willingness to share mild stress and anxiety	3.52	0.76
	Willingness to share deeply personal problems	3.18	0.81
	Comfort in expressing emotions	3.23	0.78
Therapeutic Alliance	Initial support	3.71	0.67
	Clarity of goals and direction of help	3.63	0.65
	Feeling understood	3.41	0.73
	Connectedness in the helping relationship	3.52	0.69

Source: Authors' own calculations based on indicator-level analysis of the study variables.

**Relationships Among Trust, Self-Disclosure, and Therapeutic Alliance**

Pearson correlation analysis showed that all main variables were positively and significantly related. Trust in AI-assisted counseling was positively correlated with self-disclosure ( $r = 0.58, p < 0.001$ ) and therapeutic alliance ( $r = 0.72, p < 0.001$ ). Self-disclosure was also positively correlated with therapeutic alliance ( $r = 0.64, p < 0.001$ ).

Table 7. Correlations Among the Main Variables

Variable	1	2	3
Trust in AI-Assisted Counseling	1		
Self-Disclosure in AI-Assisted Counseling	0.58***	1	
Therapeutic Alliance in AI-Assisted Counseling	0.72***	0.64***	1

Source: Authors' own calculations based on Pearson correlation analysis of the survey data.

To further examine the contribution of trust and self-disclosure to therapeutic alliance, a multiple regression analysis was conducted. The results showed that the regression model was

significant overall,  $F(2, 1281) = 885.42, p < 0.001$ , with  $R = 0.76$  and  $R^2 = 0.58$ . This indicates that trust and self-disclosure jointly explained 58% of the variance in therapeutic alliance in AI-assisted counseling. Partially, trust made a stronger predictive contribution ( $\beta = 0.53, p < 0.001$ ) than self-disclosure ( $\beta = 0.33, p < 0.001$ ).

Table 8. Multiple Regression Results Predicting Therapeutic Alliance

Predictor	B	SE B	Beta	t	p
Constant	0.74	0.09	—	8.22	<0.001
Trust in AI-Assisted Counseling	0.49	0.02	0.53	21.84	<0.001
Self-Disclosure in AI-Assisted Counseling	0.29	0.02	0.33	13.78	<0.001

Model summary:  $R = 0.76; R^2 = 0.58; \text{Adjusted } R^2 = 0.58; F(2, 1281) = 885.42; p < 0.001$

Source: Authors' own calculations based on multiple regression analysis of the survey data.

As an additional analysis, students who had previous experience using AI applications showed higher mean scores on trust, self-disclosure, and therapeutic alliance than students with no or only limited experience (Peng & Wan, 2025; Tong et al., 2022).

Table 9. Differences in Scores by Experience Using AI

Variables	Used AI Before (N = 1,109) Mean ± SD	No/Limited Experience (N = 175) Mean ± SD	t	p
Trust in AI-Assisted Counseling	3.74 ± 0.61	3.29 ± 0.72	8.41	<0.001
Self-Disclosure in AI-Assisted Counseling	3.44 ± 0.69	3.08 ± 0.77	6.31	<0.001
Therapeutic Alliance in AI-Assisted Counseling	3.62 ± 0.59	3.21 ± 0.70	7.48	<0.001

Source: Authors' own calculations based on t-test analysis of the survey data by prior experience using AI.

## Discussion

The findings of this study show that trust, self-disclosure, and therapeutic alliance are positively interconnected in the context of AI-assisted counseling among university students. Students generally reported moderately positive levels across the three variables, suggesting that AI-assisted counseling is beginning to be accepted as a meaningful early support option rather than merely a technical tool. Among the three variables, trust showed the highest mean score and emerged as the strongest predictor of therapeutic alliance. This pattern suggests that students' willingness to view AI as part of a helping relationship depends first on whether they perceive the system as reliable, safe, and responsive. In other words, before students can feel supported by AI in a counseling context, they must first believe that the system is worthy of confidence (Amin et al., 2024; Nazir, 2026).

The findings also show that self-disclosure remains more restrained than trust. Although students may perceive AI-assisted counseling as accessible and nonjudgmental, they still appear cautious when sharing deeply personal concerns. This is an important point because openness in counseling depends not only on convenience, but also on emotional safety, perceived understanding, and confidence that sensitive disclosures will be handled appropriately (Buğa et al., 2025; Wang et al., 2017). The lower score on feeling deeply understood also supports this interpretation. Students may accept AI as helpful in the early stages of help-seeking (Adams et al., 2024), but they do not yet appear to regard it as fully capable of reproducing the depth of human understanding typically associated with counseling relationships (Richardson, 2012).

The strong association between trust and therapeutic alliance reinforces the view that relational acceptance is central to the use of AI in counseling. In this study, students were more likely to report a positive sense of alliance when they trusted the AI system. This suggests that therapeutic alliance in AI-assisted settings may be built less through emotional reciprocity in the traditional human sense and more through perceptions of consistency, responsiveness, clarity, and psychological safety. Even so, the findings do not support the idea that AI can replace human counselors. Rather, they suggest that AI may function most effectively as an entry-level support resource within a broader helping system (Amin et al., 2024; Barzkar et al., 2025; Nazir, 2026).

### **Implications**

These findings suggest that AI-assisted counseling may be most useful in higher education as an initial support tool for screening, psychoeducation, emotional reflection, and early help-seeking, especially in campuses with limited counseling resources (Chen, 2024). However, because students' self-disclosure remains more moderate than their trust, AI should be positioned not as a substitute for human counselors but as a transitional and complementary resource within a blended counseling model (Amin et al., 2024; Chen, 2024). To strengthen its acceptance, the design and implementation of AI-assisted counseling should prioritize transparency, ethical safeguards, data security, response sensitivity, and clear communication about the limits of AI (Nazir, 2026), while also improving students' digital psychological literacy so they can better understand when AI is helpful and when human counseling is still necessary (Li et al., 2026; Ni & Cao, 2025). More broadly, these findings support a blended counseling model that combines AI-based initial support with more in-depth human counseling interventions (Almogren et al., 2024; Li et al., 2026).

### **Limitations**

This study has several limitations that should be considered when interpreting the findings. First, the cross-sectional design does not allow for definitive causal conclusions regarding the relationships among trust, self-disclosure, and therapeutic alliance. Second, the data were collected through participants' self-reports, which may be subject to perceptual bias, social desirability bias, or a tendency to provide normative responses. Third, although the sample was relatively large and involved multiple universities across Indonesia, the representation of each region and type of higher education institution was not fully balanced. Fourth, this study measured students' perceptions of AI-assisted counseling rather than actual clinical experiences in fully implemented and sustained AI-based counseling sessions. Therefore, the findings are more appropriately understood as reflecting an initial pattern of psychological acceptance of AI integration in counseling services, rather than as conclusive evidence of AI's therapeutic effectiveness. Fifth, the study did not distinguish in detail among the types of AI platforms imagined or used by participants, even though differences in interface, response quality, and context of use may substantially influence perceptions of trust, disclosure, and therapeutic alliance. Future studies are recommended to employ longitudinal, experimental, or mixed-methods designs to better explain the dynamics of therapeutic relationships in real-world AI use contexts.

### **CONCLUSION**

This study demonstrates that students' trust, self-disclosure, and therapeutic alliance are positively interconnected in the context of AI-assisted counseling. Trust emerged as the strongest predictor of therapeutic alliance, while self-disclosure also contributed significantly to students' perceptions of a meaningful helping relationship. These findings indicate that the relational quality of AI-assisted counseling depends not only on technological efficiency, but also on whether students perceive the system as safe, reliable, and supportive enough to encourage openness. For counseling practice, AI-assisted counseling may function best as an initial support tool that facilitates help-

seeking, emotional reflection, and early engagement before deeper interaction with a human counselor. Rather than replacing human counselors, AI should be positioned as a complementary resource within blended counseling models. Future research should examine these dynamics longitudinally and explore how different AI platforms shape trust, disclosure, and therapeutic alliance over time.

## ACKNOWLEDGMENTS

The author would like to thank the participating universities and all students who generously contributed their time and responses to this study. Appreciation is also extended to colleagues and research assistants who supported the data collection process and provided valuable administrative and technical assistance. The author is grateful to the institutional stakeholders who facilitated access to participants across multiple campuses in Indonesia.

## AUTHOR CONTRIBUTIONS STATEMENT

N.N. conceived the study, designed the research, developed the instruments, collected the data, performed the statistical analysis, interpreted the findings, and wrote the manuscript.

## FUNDING

The authors declare that no funding was received for this paper.

## REFERENCES

- Adams, D., Chuah, K.-M., Devadason, E., & Azzis, M. S. A. (2024). From novice to navigator: Students' academic help-seeking behaviour, readiness, and perceived usefulness of ChatGPT in learning. *Education and Information Technologies*, 29(11), 13617–13634. <https://doi.org/10.1007/s10639-023-12427-8>
- Almaskari, A. M., Al-, T., & Lawati, A. Al. (2025). *Students' perceptions of AI mental health chatbots : an exploratory qualitative study at Sultan Qaboos University*. 1–9. <https://doi.org/10.1136/bmjopen-2025-103893>
- Almogren, A. S., Al-rahmi, W. M., & Dahri, N. A. (2024). Integrated Technological Approaches to Academic Success : Mobile Learning , Social Media , and AI in Visual Art Education. *IEEE Access*, 12(October), 175391–175413. <https://doi.org/10.1109/ACCESS.2024.3498047>
- Alqurashi, T. S. (2025). The Impact of Artificial Intelligence on Intimate Partner Relationships. *The American Journal of Family Therapy*, 0(0), 1–20. <https://doi.org/10.1080/01926187.2025.2591600>
- Amin, M., Alturki, N., Thomas, J., & Alkhalifa, A. K. (2024). Human vs . AI counseling : College students ' perspectives. *Computers in Human Behavior Reports*, 16(June), 100534. <https://doi.org/10.1016/j.chbr.2024.100534>
- Association, W. M. (2025). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human participants. *Jama*, 333(1), 71–74.
- Barzkar, F., Zaribaf, A., Mirfazeli, F. S., & Keshavarz-Akhlaghi, A.-A. (2025). The Machine as Therapist: Unpacking Transference and Emotional Healing in AI-Assisted Therapy. *Journal of Contemporary Psychotherapy*, 55(4), 361–368. <https://doi.org/10.1007/s10879-025-09677-7>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health*, 6(June), 1–18. <https://doi.org/10.3389/fpubh.2018.00149>
- Buğa, A., Özdemir, M., & Bozbayındır, F. (2025). Exploring psychological counselors' perspectives on the characteristics and dynamics of online counseling. *Frontiers in Psychology*, 16, 1708309. <https://doi.org/10.3389/fpsyg.2025.1708309>

- Chaudhry, B. M., & Debi, H. R. (2024). *User perceptions and experiences of an AI-driven conversational agent for mental health support*. <https://doi.org/10.21037/mhealth-23-55>
- Chen, J. (2024). *Artificial intelligence significantly facilitates development in the mental health of college students : a bibliometric analysis*. March. <https://doi.org/10.3389/fpsyg.2024.1375294>
- Chen, L. (2024). Development of an AI-Based System to Enhance School Counseling Models for Asian Elementary Students With Emotional Disorders. *IEEE Access*, 12(August), 159121–159136. <https://doi.org/10.1109/ACCESS.2024.3483456>
- Haq, M. D. R., & Rubya, S. (2023). *An Overview of Chatbot-Based Mobile Mental Health Apps : Insights From App Description and User Reviews Corresponding Author : 11*. <https://doi.org/10.2196/44838>
- Hawke, L. D., Hou, J., Nguyen, A. T. P., Phi, T., Gibson, J., Ritchie, B., Strudwick, G., Rodak, T., & Gallagher, L. (2025). Digital Conversational Agents for the Mental Health of Treatment-Seeking Youth : Scoping Review Corresponding Author : *JMIR Mental Health*, 12, 1–14. <https://doi.org/10.2196/77098>
- Healthtech, I., Limited, P., Healthtech, I., Limited, P., Healthtech, I., Limited, P., Healthtech, I., Limited, P., Healthtech, I., & Limited, P. (n.d.). *Chatbots as Tools for Psychoeducation and Self-Help in Mental Health*. <https://doi.org/10.4018/979-8-3693-3112-5.ch003>
- Hoffman, B. D., Oppert, M. L., & Owen, M. (2024). Computers in Human Behavior : Artificial Humans Understanding young adults' attitudes towards using AI chatbots for psychotherapy : The role of self-stigma. *Computers in Human Behavior: Artificial Humans*, 2(2), 100086. <https://doi.org/10.1016/j.chbah.2024.100086>
- Hunt, D. F., Bailey, J., Lennox, B. R., Crofts, M., & Vincent, C. (2021). Enhancing psychological safety in mental health services. *International Journal of Mental Health Systems*, 15(1), 33. <https://doi.org/org/10.1186/s13033-021-00439-1>
- Jung, H. W., Park, J. Y., Holoubek, T., Jung, W., & Park, J. (2025). *Socially Assistive Robots in Mental Healthcare : Principles and Conceptual Framework for User-Centered Design*. 2827–2851.
- Koulouri, T., Macredie, R. D., & Olakitan, D. (2022). *Chatbots to Support Young Adults' Mental Health : An Exploratory Study of Acceptability*. 12(2). <https://doi.org/10.1145/3485874>
- Krause, M., Altimir, C., & Horvath, A. (2011). *Deconstructing the Therapeutic Alliance : Reflections on the Underlying Dimensions of the Concept Una Deconstrucción de la Alianza Terapéutica : Reflexiones sobre las Dimensiones Subyacentes del Concepto*. 22, 267–283.
- Lawati, A. Al, Al-mahrouqi, T., Maskari, A. Al, Riyami, Q. Al, Aufi, H. Al, Jose, S., & Sinawi, H. Al. (2026). *Perceptions and attitudes of undergraduate university students toward artificial intelligence-powered mental health chatbots : a cross-sectional study*.
- Li, B., Paracha, S., & Limniou, M. (2026). Investigating the use of AI in psychological counseling : *Mental Health and Digital Technologies*, 3(1), 3–31. <https://doi.org/10.1108/MHDT-04-2025-0027>
- Li, H., Lei, Y., Zhou, Q., & Yuan, H. (2023). Can you sense without being human ? Comparing virtual and human influencers endorsement effectiveness. *Journal of Retailing and Consumer Services*, 75(May), 103456. <https://doi.org/10.1016/j.jretconser.2023.103456>
- Li, H., Zhang, R., Lee, Y., Kraut, R. E., & Mohr, D. C. (2023). *Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being*. <https://doi.org/10.1038/s41746-023-00979-5>
- Lotringen, C. M. Van, Jeken, L., Westerhof, G. J., & Klooster, P. M. (2021). *Responsible Relations : A Systematic Scoping Review of the Therapeutic Alliance in Text-Based Digital Psychotherapy*. 3(July). <https://doi.org/10.3389/fdgth.2021.689750>
- Miqdadi, A. I., Nayef, M., & Majdi, A. (2025). Attitudes Toward Seeking Professional Psychological Help , Self-Stigma of seeking help , and Preferences for AI Mental Health Chatbots vs . Human

- Counsellors : A Mixed Methods Study. *International Journal of Mental Health and Addiction*, 0123456789. <https://doi.org/10.1007/s11469-025-01595-y>
- Mitsea, E., & Drigas, A. (2023). *behavioral sciences Digitally Assisted Mindfulness in Training Self-Regulation Skills for Sustainable Mental Health : A Systematic Review*.
- Nazir, T. (2026). Therapists and Technology : A Qualitative Study on AI ' s Role in Counseling. *Human Behavior and Emerging Technologies*, 2026. <https://doi.org/10.1155/hbe2/3616773>
- Ni, Y., & Cao, Y. (2025). *Exploring ChatGPT ' s capabilities , stability , potential and risks in conducting psychological counseling through simulations in school counseling*. 2(3), 213–239. <https://doi.org/10.1108/MHDT-02-2025-0013>
- Peng, Z., & Wan, Y. (2025). Human vs. AI: What makes students prefer to self-disclosure to AI teaching assistant? The effect of psychological safety and relationship norms. *Education and Information Technologies*, 30(14), 20851–20878. <https://doi.org/org/10.1007/s10639-025-13586-6>
- Ping, Y. (2024). *Experience in psychological counseling supported by artificial intelligence technology*. 32, 3871–3888. <https://doi.org/10.3233/THC-230809>
- Podolan, M., & Gelo, O. C. G. (2024). The Role of Safety in Change-Promoting Therapeutic Relationships: An Integrative Relational Approach. *Clinical Neuropsychiatry*, 21(5), 403–417.
- Rackoff, G. N., Zhang, Z. Z., & Newman, M. G. (2025). *Chatbot-delivered mental health support : Attitudes and utilization in a sample of U . S . college students*. <https://doi.org/10.1177/20552076241313401>
- Richardson, M. S. (2012). Counseling for Work and Relationship. *The Counseling Psychologist*, 40(2), 190–242. <https://doi.org/10.1177/0011000011406452>
- Seager, M. (2007). *The Concept of 'Psychological Safety' – A Psychoanalytically - Informed Contribution Towards 'Safe, Sound & Supportive' Mental Health Services*. 8734. <https://doi.org/10.1080/02668730601020291>
- Shen, J., Dipaola, D., Ali, S., Sap, M., & Park, H. W. (2024). *Empathy Toward Artificial Intelligence Versus Human Experiences and the Role of Transparency in Mental Health and Social Support Chatbot Design : Comparative Study Corresponding Author : 11, 1–13*. <https://doi.org/10.2196/62679>
- Spector, P. E. (2019). *Do Not Cross Me : Optimizing the Use of Cross-Sectional Designs*. 125–137.
- Tong, F., Lederman, R., Alfonso, S. D., Berry, K., & Bucci, S. (2022). *Digital Therapeutic Alliance With Fully Automated Mental Health Smartphone Apps : A Narrative Review*. 13(June), 1–12. <https://doi.org/10.3389/fpsy.2022.819623>
- Tremain, H., Mcenery, C., Fletcher, K., Murray, G., & Tremain, H. (2020). *The Therapeutic Alliance in Digital Mental Health Interventions for Serious Mental Illnesses : Narrative Review Corresponding Author : 7*. <https://doi.org/10.2196/17204>
- Wampold, B. E., & Flückiger, C. (2023). The alliance in mental health care : conceptualization , evidence and clinical applications. *World Psychiatry*, 21(1), 25–41. <https://doi.org/10.1002/wps.21035>
- Wang, L., Yan, J., Lin, J., & Cui, W. (2017). International Journal of Information Management Let the users tell the truth : Self-disclosure intention and self-disclosure honesty in mobile social networking. *International Journal of Information Management*, 37(1), 1428–1440. <https://doi.org/10.1016/j.ijinfomgt.2016.10.006>
- Wang, Y. (2025). *Effect of anthropomorphism and perceived intelligence in chatbot avatars of visual design on user experience : accounting for perceived empathy and trust*. May. <https://doi.org/10.3389/fcomp.2025.1531976>