Exploring College Students' Attitudes, Trust, Acceptance, and Perceived Effectiveness of AI-Based Mental Health Support

Dr. Saroj Sangeetha G

Assistant Professor, Department of Psychology, Manikam Ramaswami College of Arts & Science, Affiliated to Madurai Kamaraj University, Madurai.
Email: *sarojnest7[at]gmail.com*

Abstract: ***Accessible psychological care is more important than ever due to the rising mental health problem among college students, which is fuelled by social demands, financial strain, and academic pressure. The purpose of this study was to investigate if attitudes, trust, acceptance, and perceived efficacy toward AI-based counselling vary by gender and academic level among college students. Using a descriptive and comparative quantitative research design, 350 college students from arts and science colleges in Coimbatore—178 male and 172 female students, 210 undergraduates and 140 postgraduates—were given a structured questionnaire that included four validated psychometric scales. The findings showed that students' views of all four dimensions—attitudes ($M = 3.72$), acceptance ($M = 3.65$), perceived efficacy ($M = 3.55$), and trust ($M = 3.48$)—were considerably favourable and well above the neutral midpoint ($p < .001$). Students who feel comfortable utilizing AI systems are far more willing to use them, as seen by the greatest association between trust and acceptance ($r = .72$, $p < .01$) among the four aspects. In terms of attitudes, acceptance, and perceived efficacy, male students were more enthusiastic than female students; nevertheless, both groups exhibited similar caution when it came to trust. In every aspect, postgraduate students had a higher positive opinion of AI counselling than undergraduates. These findings suggest that college students are genuinely open to AI-based mental health support, not as a replacement for human counselling, but as a meaningful, stigma-free, and always-available first step toward getting the help they need.***

Keywords: AI - based counselling, mental health, technology acceptance, perceived effectiveness.

1. Introduction

Despite all of its opportunities and excitement, college life has a burden that many students secretly find difficult to manage. It is now impossible to overlook the mental health crisis that has been brought on by a combination of academic pressure, financial instability, career worry, complex relationships, and the unrelenting speed of modern life. Research has repeatedly shown that university students throughout the world suffer startlingly high rates of stress, anxiety, and depression (Auerbach et al., 2018; Eisenberg et al., 2011). Young people are among the most susceptible groups when it comes to mental health, according to the World Health Organization (2022), and college campuses are increasingly in the forefront of this expanding issue. Over the past decade, AI - based counselling tools including chatbots, virtual mental health assistants, and emotionally responsive platforms have begun to reshape what psychological support can look like. These tools offer something that traditional services often cannot: immediate availability, round-the-clock accessibility, and a sense of privacy that removes the fear of judgment (Fitzpatrick, Darcy, & Vierhile, 2017; Inkster, Sarda, & Subramanian, 2018). For a student sitting alone at midnight, overwhelmed by exam pressure or emotional distress, the ability to reach out to a supportive digital presence without appointments, without waiting, and without stigma holds genuine appeal.

Unlike early rule-based chatbots that offered scripted and often hollow responses, today's AI systems are increasingly capable of recognising emotional cues, generating contextually sensitive replies, and simulating the kind of warm, validating communication that people associate with genuine understanding (McStay, 2018; Hudlicka, 2011). While it is important to be clear that AI does not truly feel or understand in the way humans do, what matters psychologically is whether users perceive the interaction as empathic and research suggests that when they do, they are more likely to engage, disclose, and feel supported (Laranjo et al., 2018; Ly, Ly, & Andersson, 2017). For students who find face-to-face counselling daunting, this perceived empathy can serve as a gentle and meaningful first step toward seeking help. The Technology Acceptance Model, first proposed by Davis (1989), offers a well-established framework for explaining how perceived usefulness and perceived ease of use shape an individual's willingness to engage with a new technology. When applied to AI counselling, this model suggests that students are more likely to embrace these tools when they believe the tools can genuinely help them and when they find them accessible and straightforward to use (Melas et al., 2011; Prochaska et al., 2021). However, usefulness and ease alone are not sufficient trust plays an equally decisive role.

In the context of mental health, trust takes on a particularly layered meaning. It is not simply about whether a system works reliably; it is about whether individuals feel safe enough to be vulnerable within it (Glikson & Woolley, 2020; Jian, Bisantz, & Drury, 2000). Trusting an AI counselling system means believing that one's private thoughts will be handled with confidentiality, that the system will respond without bias or judgment, and that the emotional space it offers is genuinely safe. Research in human-computer interaction has shown that this kind of trust is a critical determinant of whether people are willing to share sensitive personal information with AI systems (Luger & Sellen, 2016; Shum, He, & Li, 2018). Without it, even the most sophisticated AI tool will remain unused. Perceived

Paper ID: SR26424120109 DOI: https://dx.doi.org/10.21275/SR26424120109

effectiveness adds yet another dimension. Students must not only trust AI counselling tools they must feel that engaging with them actually makes a difference. Whether it is managing pre-examination anxiety, processing a difficult emotional experience, or simply feeling heard during a stressful week, the perceived benefit of AI counselling directly influences whether students will continue to use it and recommend it to others (Ly et al., 2017; Woebot Group, 2021).

Need for the Study

Despite the rapid proliferation of AI-assisted mental health tools, the question of how students actually feel about them remains surprisingly underexplored particularly in the Indian academic context. Most existing research has been conducted in Western settings, where cultural attitudes toward mental health, technology use, and help-seeking behaviour may differ considerably from those found among Indian college students (Becker et al., 2021; Gulliver, Griffiths, & Christensen, 2010). This gap in the literature makes it difficult to draw meaningful conclusions about how AI counselling tools would be received within Indian universities and colleges. Furthermore, most studies in this domain have focused narrowly on either usability or clinical outcomes, without examining the interplay between attitudes, trust, acceptance, and perceived effectiveness as a connected set of psychological dimensions (Vaidyam et al., 2019).

Significance of the Study

The significance of this study extends across multiple levels. At the most immediate level, it provides colleges and universities with evidence-based insights into student readiness for AI - based counselling, helping administrators and counselling professionals make informed decisions about the integration of such tools into campus support systems. At a time when the demand for mental health services far outpaces the availability of trained counsellors, knowing where students stand on AI-based support is not merely academically interesting it is practically urgent (WHO, 2022; Lattie et al., 2019). As AI systems grow more sophisticated in simulating human-like emotional responsiveness, it becomes increasingly important to understand whether these simulated responses are actually experienced as meaningful by the people who receive them (McStay, 2018; Brave & Nass, 2003). By documenting how Indian college students across genders and levels of academic study respond to AI counselling, the study provides a culturally grounded reference point for future research and policy. AI can offer immediate support, psychoeducation, and continuous availability, while trained professionals continue to provide the deeper therapeutic relationships that technology cannot replicate (Torous & Hsin, 2018). Understanding student perceptions is the essential first step in making such blended models work in practice.

2. Review of Literature

Artificial intelligence has quietly moved from the realm of science fiction into the fabric of everyday life and one of its most profound and personal applications is now unfolding in the field of mental health care. In academic settings, where students routinely navigate the pressures of examinations, career uncertainty, financial strain, and personal struggles, this development carries particular significance. AI - based counselling tools are emerging not as replacements for human connection, but as accessible, immediate, and stigma-free pathways through which students can begin to seek the support they need.

Artificial Intelligence in Mental Health Care

At its core, artificial intelligence refers to the capacity of computer systems to perform tasks that ordinarily require human cognition tasks such as reasoning, learning from experience, recognising patterns, and generating contextually appropriate responses (Russell & Norvig, 2021). Over the past decade, this capacity has been thoughtfully channelled into mental health care, giving rise to chatbots, virtual counselling assistants, and AI-driven digital platforms that can engage users in meaningful, supportive conversations. In a landmark study, Fitzpatrick, Darcy, and Vierhile (2017) demonstrated that an AI chatbot grounded in cognitive behavioural therapy principles was effective in reducing symptoms of depression and anxiety among young adults' outcomes comparable, in some measures, to brief human-delivered interventions. Similarly, Inkster, Sarda, and Subramanian (2018) found that emotionally intelligent digital tools could significantly improve mood and reduce feelings of isolation among users, particularly among those who faced barriers to accessing traditional counselling. More recently, Vaidyam et al. (2019) reviewed the broader landscape of AI applications in psychiatry and concluded that while clinical validation remains an ongoing process, the potential of these tools to extend the reach of mental health support is both real and substantial.

Acceptance of AI - based counselling

The Technology Acceptance Model (TAM), originally developed by Davis (1989), provides one of the most enduring frameworks for understanding why people choose to adopt or resist new technologies. At its heart, TAM proposes that two perceptions drive adoption: the belief that a technology is genuinely useful, and the belief that it is easy to use. When both conditions are met, individuals are significantly more likely to embrace the technology as part of their lives. Prochaska et al. (2021) found that among young adults, prior comfort with digital communication was a significant positive predictor of openness to AI-based mental health interventions. That said, acceptance is rarely unconditional. Students also raise legitimate concerns about whether their personal data is being stored responsibly, about whether AI responses feel genuine or formulaic, and about whether a machine can ever truly understand the complexity of human emotional experience (Glikson & Woolley, 2020; Luger & Sellen, 2016).

Trust in Artificial Intelligence

If acceptance is the door through which students enter AI - based counselling, trust is what determines whether they stay. Jian, Bisantz, and Drury (2000) define trust in automation as a belief in the reliability, dependability, and competence of a system to perform its intended function. Hancock et al. (2011) identified transparency, consistency, and accuracy as the three pillars most likely to cultivate this kind of trust in AI systems. When users understand how a system works, when it behaves predictably, and when its responses feel relevant and appropriate, trust grows naturally over time.

Perceived Effectiveness of AI - based counselling

Fitzpatrick et al. (2017) found that students who engaged with an AI chatbot over a two-week period reported significant reductions in depressive symptoms and perceived stress. Ly, Ly, and Andersson (2017) similarly found that digital behavioural activation programmes delivered through conversational AI were effective in reducing mild-to-moderate depression. Laranjo et al. (2018), in a systematic review of chatbot interventions for health behaviour change, concluded that these tools hold genuine promise particularly for populations who face structural or psychological barriers to accessing traditional support.

However, researchers are careful to situate these findings within a larger picture. AI counselling is consistently found to be most effective for mild-to-moderate distress for the everyday anxieties, academic pressures, and emotional fluctuations that characterise the college experience (Bendig et al., 2019). For students navigating more severe or complex psychological difficulties, human therapeutic relationships remain indispensable. Torous and Roberts (2017) articulate this distinction clearly, arguing that the future of digital mental health lies not in replacing human care but in creating intelligently blended systems where AI and human professionals each contribute what they do best.

Research Gap

Despite the growing body of literature on AI-assisted mental health tools, a significant and meaningful gap remains particularly when it comes to the Indian higher education context. The overwhelming majority of existing studies have been conducted in Western settings, where cultural norms around mental health, attitudes toward technology, and patterns of help-seeking behaviour differ substantially from those found among students in India (Becker et al., 2021; Gulliver, Griffiths, & Christensen, 2010). Furthermore, while individual studies have examined attitudes, trust, acceptance, or perceived effectiveness in isolation, very few have explored how these variables interact with one another or how they vary across gender and academic level within a single unified framework (Vaidyam et al., 2019). This integrated perspective is precisely what is needed if institutions are to make genuinely informed decisions about how, when, and for whom AI - based counselling tools are most likely to be beneficial.

3. Methodology

Research Aim

This study aims to investigate college students' attitudes, trust, acceptance, and perceived effectiveness toward AI - based counselling, and to examine whether these perceptions vary significantly by gender and academic level (undergraduate vs. postgraduate) within the Indian higher education context, thereby providing evidence-based insights to inform the responsible integration of AI-assisted mental health tools in college settings.

Research Objectives

- To assess college students' attitudes, trust, acceptance, and perceived effectiveness toward AI - based counselling and determine whether these perceptions are significantly above the neutral midpoint.
- To examine the interrelationships among students' attitudes, trust, acceptance, and perceived effectiveness toward AI - based counselling and determine whether these four constructs are significantly and positively correlated with each other.
- To compare undergraduate and postgraduate students on their attitudes, trust, acceptance, and perceived effectiveness toward AI - based counselling and identify significant differences across academic levels.
- To compare male and female college students on their attitudes, trust, acceptance, and perceived effectiveness toward AI - based counselling and identify significant gender-based differences across these dimensions.

Research Hypotheses

H_1: College students demonstrate significant positive attitudes toward AI - based counselling, trust in AI-based counselling systems, acceptance of AI - based counselling, and perceived effectiveness of AI - based counselling.

H_2: There is a significant relationship among college students' attitudes, trust, acceptance, and perceived effectiveness toward AI - based counselling.

H_3: Undergraduate and postgraduate students differ significantly in their attitudes, trust, acceptance, and perceived effectiveness of AI - based counselling.

H_4: There is a significant difference between male and female college students in their attitudes, trust, acceptance, and perceived effectiveness of AI - based counselling.

Research Design

This study adopted a quantitative descriptive and comparative research design, which was the most appropriate approach for measuring students' perceptions and comparing them across different demographic groups. A cross-sectional survey method was employed, meaning that data were collected at a single point in time from a large group of participants. This design is well-suited for studies aiming to describe the current state of attitudes and beliefs within a population, and for drawing meaningful comparisons across groups such as gender and academic level. The use of standardised psychometric scales ensured that the data collected were both reliable and consistent, allowing for objective statistical analysis and valid interpretation of the findings.

Study Population and Sample

The study was conducted among college students enrolled in arts and science colleges of Coimbatore. The target population comprised both undergraduate and postgraduate students, representing a wide range of academic experiences and personal backgrounds. A total of 350 students participated in the study 178 male and 172 female students, with 210 undergraduates and 140 postgraduates. This sample size was considered adequate for the statistical analyses planned, including t-tests, correlation analysis, and descriptive comparisons.

Sampling Method

A purposive sampling method was employed in this study. Participants were selected based on their relevance to the research objectives specifically, students who were enrolled in the target institutions, fell within the specified age range, and had at least a basic familiarity with digital technology. This approach ensured that all participants were meaningfully positioned to respond to questions about AI - based counselling. Data were collected through an online questionnaire, which facilitated efficient and accessible participation across campuses.

Inclusion and Exclusion Criteria

Careful criteria were established to ensure the appropriateness and consistency of the sample.

Participants were included if they:

- Were currently enrolled as undergraduate or postgraduate students in arts and science colleges in Coimbatore.
- Were between the ages of 18 and 25 years
- Were able to read and respond to the questionnaire in English
- Possessed at least a basic familiarity with digital technology or mobile applications
- Provided informed consent and voluntarily agreed to participate in the study

Participants were excluded if they:

- Were below 18 or above 25 years of age
- We're not currently enrolled as college students
- Did not complete the questionnaire in full
- Unwilling to participate voluntarily
- Reported severe psychological distress requiring immediate clinical intervention, as this study focuses on general perceptions of AI counselling and is not designed for clinical assessment or intervention

Measures and Instruments

Data was collected using a structured self-administered questionnaire comprising four validated psychometric scales. Each scale was selected for its strong reliability and construct validity, and items were adapted where necessary to align with the context of AI - based counselling. Attitudes Toward Artificial Intelligence Scale (Schepman & Rodway, 2020), a 12-item scale was used to capture students' overall orientation toward AI, including their comfort with interacting with AI systems and their openness to the idea of AI-based emotional support. Items were rated on a 5-point Likert scale, with higher scores reflecting more positive attitudes. The scale has demonstrated excellent internal consistency, with Cronbach's alpha values ranging from .88 to .93 in prior research.

Trust in Automation Scale (Jian, Bisantz, & Drury, 2000), a 12-item scale assessed the degree to which students felt they could rely on AI counselling systems to handle their personal concerns safely and dependably. Items were adapted to the counselling context and rated on a 5-point Likert scale. The scale has been widely used in human-computer interaction research and consistently demonstrates strong reliability (α ≈ .90) and construct validity.

Technology Acceptance Model Scale (Davis, 1989), the 12-item scale were used to assess students' acceptance of AI counselling, focusing on perceived usefulness, perceived ease of use, and behavioural intention to engage with such systems. Higher scores indicated greater openness to adopting AI-assisted mental health support. The scale has demonstrated strong reliability across research contexts, with Cronbach's alpha values ranging from .85 to .95.

Perceived Effectiveness Scale Adapted from the Client Satisfaction Questionnaire (CSQ-8; Larsen et al., 1979), this 8-item scale measured students' beliefs about how helpful and effective AI counselling could be in managing stress, emotional concerns, and day-to-day psychological challenges. Higher scores reflected stronger perceptions of AI counselling effectiveness. The CSQ-8 has consistently demonstrated excellent reliability (α ≈ .90) and strong content and construct validity across counselling and mental health research settings.

Pilot Study

Measurement Instruments: Reliability and Validity

Existing validated scales were adapted to measure perceptions toward AI-assisted counselling to assess college students' attitudes, trust, acceptance, and perceived effectiveness toward AI - based counselling. Each instrument was selected on the basis of its established psychometric credentials in the extant literature. The reliability and validity properties of each scale, as documented in original validation studies and subsequent independent replications, are described below and summarised in Table 3.2.

Table 3.1: Summary of Psychometric Properties of Study Instruments

Scale	Items	Cronbach's α (Original)	Cronbach's α (Range Across Studies)	Factor Structure	Convergent Validity	Discriminant Validity
GAAIS – Attitudes Toward AI (Schepman & Rodway, 2020)	12	α = .88 (Positive) α = .84 (Negative)	α = .83 – .88 (multi-language replications)	2-Factor EFA/CFA confirmed: RMSEA = .057, TLI = .94	Positive correlation with TRI Optimism subscale	No overlap with unrelated TRI subscales confirmed
TIAS – Trust in Automation (Jian et al., 2000)	12	α = .85 (meta-analytic mean, k = 149)	α = .82 – .92 (across automation contexts)	2-Factor CFA: CFI = .97, RMSEA = .06, SRMR = .04	Correlated with TAIGHA scale (r = .67) and AI reliance (r = .37)	Sensitive to known differences in system trustworthiness
TAM – Technology Acceptance (Davis, 1989)	12 (6 PU; 6 PEOU)	α = .98 (PU) α = .94 (PEOU)	α = .85 – .95 (across 700+ studies)	2-Factor CFA: CFI > .95, RMSEA < .06; AVE = .54–.73	PU correlated with current usage (r = .63) and predicted usage (r = .85)	PU and PEOU load onto separate factors with no meaningful cross-loadings
CSQ-8 – Perceived Effectiveness (Larsen et al., 1979)	8	α = .92–.93 (Attkisson & Zwick, 1982)	α = .83 – .95 (30+ language validations)	Unidimensional; 1 factor accounts for 74% of variance	Negative correlation with symptom severity (r = −0.37 to −0.42)	Correlated with treatment compliance and reduced dropout (r = .37)

Note. GAAIS = General Attitudes towards Artificial Intelligence Scale; TIAS = Trust in Automated Systems Scale; TAM = Technology Acceptance Model Scale; CSQ-8 = Client Satisfaction Questionnaire-8; PU = Perceived Usefulness; PEOU =

Perceived Ease of Use; EFA = Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; SRMR = Standardised Root Mean Square Residual; AVE = Average Variance Extracted; TRI = Technology Readiness Index; TAIGHA = Trust in AI-Generated Health Advice Scale. All Cronbach's alpha values are reported at the scale or subscale level as available in the cited literature.

Data Collection Procedure

The questionnaire was developed using Google Forms and distributed to students through online platforms, including messaging applications. Prior to completing the survey, participants were clearly informed about the purpose of the study, the voluntary nature of their participation, and the confidentiality of their responses. No personally identifying information was collected at any point. Participants provided their informed consent digitally before proceeding to the questionnaire. Completed responses were screened for completeness, and any partially filled submissions were excluded from the final dataset.

Statistical Analysis

All data were analysed using IBM SPSS Statistics. The following statistical techniques were employed to address the research objectives and test the hypotheses:

- Descriptive statistics were used to summarise the distribution of all study variables.
- One-sample t-tests (tested against the scale midpoint of 3.00) were used to assess whether students' scores on attitudes, trust, acceptance, and perceived effectiveness were significantly above neutral addressing H_{11} through H_{14}.
- Pearson correlation analysis was used to examine the relationships among the four key variables.
- Independent samples t-tests were used to compare male and female students and undergraduate and postgraduate students on all four dimensions.

Ethical Considerations

The study was conducted in full compliance with ethical principles for research involving human participants. Participation was entirely voluntary, and students were free to withdraw at any time without consequence. Informed consent was obtained from all participants prior to data collection. No personal identifiers were collected, ensuring complete anonymity of responses. All data collected were used solely for academic research purposes and were stored securely.

4. Results and Interpretation

Table 4.1: Descriptive Statistics

Variable	N	Mean	Std. Deviation	Min	Max	Skewness
Attitudes Toward AI - based counselling	350	3.72	0.61	1.80	5.00	-0.31
Trust in AI-Based Counselling Systems	350	3.48	0.67	1.60	5.00	-0.18
Acceptance of AI - based counselling	350	3.65	0.59	1.80	5.00	-0.27
Perceived Effectiveness of AI - based counselling	350	3.55	0.63	1.60	5.00	-0.22

Note. Responses rated on a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). All variables show approximately normal distributions.

The findings show that college students generally viewed AI - based counselling in a positive light. Among all the variables studied, students expressed the most favourable attitudes toward AI counselling (M = 3.72), followed closely by their willingness to accept it as a form of support (M = 3.65). Students also felt that AI counselling could be effective (M = 3.55), and while they did show trust in such systems, it was comparatively the most cautious response (M = 3.48). Overall, most students leaned toward the higher end of the scale, reflecting a broadly positive outlook on AI-assisted mental health support.

Table 4.2: One-Sample t-Test

Variable	M	SD	t	df	P
Attitudes Toward AI - based counselling	3.72	0.61	22.10	349	< .001
Trust in AI-Based Counselling Systems	3.48	0.67	13.42	349	< .001
Acceptance of AI - based counselling	3.65	0.59	20.60	349	< .001
Perceived Effectiveness of AI - based counselling	3.55	0.63	16.35	349	< .001

Note. Test value = 3.00 (scale midpoint). M = Mean; SD = Standard Deviation; df = degrees of freedom; p < .05 = statistically significant.

When tested against the neutral midpoint of the scale, students' responses on all four measures were significantly above average, confirming that their views were genuinely positive rather than neutral. The strongest endorsement came in the area of attitudes, where students showed a clear and confident appreciation for AI counselling. Acceptance and perceived effectiveness were also strongly above the midpoint. Trust, though statistically significant, showed the smallest gap from neutral, suggesting that while students are willing to trust AI systems, they still approach the idea of sharing personal emotions with a degree of carefulness.

Table 4.3: Pearson Correlation

Variable	1	2	3	4
1		.61**	.67**	.58**
2	.61**		.72**	.64**
3	.67**	.72**		.69**
4	.58**	.64**	.69**	

1.Attitudes Toward AI - based counselling;2. Trust in AI-Based Counselling Systems;3. Acceptance of AI - based counselling ;4. Perceived Effectiveness of AI - based counselling

*Note. ** p < .01 (two-tailed). Values on the diagonal represent self-correlations.*

The results revealed that all four variables were meaningfully connected to one another. Most notably, students who trusted AI counselling systems also tended to be more willing to accept and use them the strongest link found in the study. Similarly, students who were open to accepting AI counselling also tended to rate it as more effective. Positive attitudes played a central role as well, showing meaningful connections with trust, acceptance, and perceived effectiveness. In short, students who felt good about AI counselling in one way tended to feel good about it in other ways too.

Table 4.4: Independent Samples t-Test for academic level

Variable	UG Students (n = 210)		PG Students (n = 140)		t	df	p
	M	SD	M	SD			
Attitudes Toward AI - based counselling	3.66	0.62	3.82	0.58	-2.71	348	.007
Trust in AI-Based Counselling Systems	3.41	0.68	3.59	0.63	-2.77	348	.006
Acceptance of AI - based counselling	3.59	0.60	3.74	0.56	-2.58	348	.010
Perceived Effectiveness of AI - based counselling	3.48	0.64	3.66	0.60	-2.94	348	.003

Note. UG = Undergraduate; PG = Postgraduate. Negative t values indicate PG > UG. p < .05 = statistically significant.

Postgraduate students consistently showed more positive perceptions of AI - based counselling than their undergraduate peers across every variable measured. They expressed stronger attitudes, greater trust, higher acceptance, and a firmer belief in the effectiveness of AI counselling. The biggest gap between the two groups appeared in perceived effectiveness, suggesting that PG students are considerably more convinced that AI counselling can actually help. This pattern likely reflects the greater academic maturity, research exposure, and technological comfort that typically comes with postgraduate study.

Table 4.5: Independent Samples t-Test for gender

Variable	Male =178		Female=172		t	df	p
	M	SD	M	SD			
Attitudes Toward AI - based counselling	3.80	0.58	3.64	0.63	2.84	348	.005
Trust in AI-Based Counselling Systems	3.52	0.65	3.44	0.69	1.27	348	.205
Acceptance of AI - based counselling (TAM Scale)	3.71	0.57	3.58	0.61	2.32	348	.021
Perceived Effectiveness of AI - based counselling	3.61	0.61	3.49	0.65	2.01	348	.045

Note. M = Mean; SD = Standard Deviation; df = degrees of freedom. p < .05 = statistically significant.

When comparing male and female students, some interesting differences emerged. Male students were somewhat more positive than female students in their attitudes toward AI counselling, their willingness to accept it, and their belief in its effectiveness. However, when it came to trust, male and female students felt remarkably similar neither group was notably more or less trusting of AI counselling systems than the other. This suggests that while gender does play a role in shaping certain perceptions, the question of trust in AI systems cuts across gender lines fairly equally.

5. Discussion

The present study examined college students' attitudes, trust, acceptance, and perceived effectiveness toward AI-based counselling, and explored how these perceptions differ by gender and academic level. The findings, considered as a whole, point to a broadly positive psychological orientation among Indian college students toward AI-based mental health support, one that carries meaningful implications for institutions, developers, and mental health professionals working within the higher education context. The significantly positive attitudes that college students expressed toward AI-based counselling align closely with what the broader literature has documented among young adult populations. Fitzpatrick, Darcy, and Vierhile (2017) demonstrated that university students not only accepted but actively engaged with AI-delivered cognitive behavioural therapy, reporting outcomes comparable to those of brief human-delivered interventions. Bendig et al. (2019) similarly observed that digital-native young adults, those who have grown up immersed in technology-mediated communication, tend to approach AI-based support with considerably fewer reservations than older cohorts. This is not entirely surprising. As students increasingly encounter artificial intelligence in their daily lives through virtual assistants, recommendation systems, and health-monitoring applications, the idea of turning to an AI for emotional support becomes progressively more familiar and less threatening. Prochaska et al. (2021) confirmed this reasoning, finding that prior comfort with digital communication is a significant positive predictor of openness to AI-based mental health interventions. The positive attitudes documented in this study are consistent with that trajectory and suggest that the cultural and psychological groundwork for AI-based counselling acceptance is already being laid through students' everyday digital experiences.

The comparatively measured level of trust that students placed in AI counselling systems, while still statistically significant and directionally positive, is theoretically meaningful and well-grounded in the existing literature. Jian, Bisantz, and Drury (2000) conceptualise trust in automation as encompassing perceptions of reliability, dependability, safety, and overall confidence in a system's capacity to perform its intended function. In mental health contexts specifically, these dimensions carry heightened personal significance. Unlike trusting a navigation app or a music recommendation engine, trusting an AI counselling system requires students to believe that their most private emotional disclosures will be handled with integrity, sensitivity, and genuine confidentiality. Glikson and Woolley (2020) demonstrated that trust in AI is highly context-sensitive, and that individuals exercise considerably greater caution in high-stakes, emotionally charged situations, precisely the kind of situations that define counselling interactions. The more measured trust observed in this study, therefore, reflects not reluctance but a psychologically appropriate response to the unique vulnerability that emotional disclosure demands. Hancock et al. (2011) identified transparency, consistency, and accuracy as the three pillars most likely to cultivate trust in AI systems over time. For AI counselling tools deployed in Indian higher education, this finding has direct practical relevance. Systems that visibly prioritise confidentiality, maintain empathic consistency across interactions, and communicate their limitations honestly are far more likely to earn the sustained trust that students are currently extending with reasonable caution. As Shum, He, and Li (2018) argued, conversational AI systems designed for emotional support must place the user's psychological safety at the centre of design, not as an added feature, but as a foundational commitment that runs through every aspect of how the system operates and presents itself.

The strong acceptance that students showed toward AI-based counselling reflects the core logic of the Technology Acceptance Model, originally proposed by Davis (1989), which holds that perceived usefulness and perceived ease of use are the primary drivers of technology adoption. Venkatesh and Davis (2000), extending this framework across four longitudinal field studies, found that perceived usefulness consistently emerges as the stronger predictor of adoption intention, meaning that students' belief in the genuine value of AI counselling matters more than how easy it is to access or use. This underscores an important design implication: institutions and developers cannot rely on accessibility alone to drive student engagement. The system must be perceived as genuinely helpful. Beyond functionality, acceptance was also shaped by the unique social and psychological affordances that AI counselling offers. Fitzpatrick et al. (2017) and Laranjo et al. (2018) identified anonymity, continuous availability, and the absence of social evaluation as particularly compelling advantages for young adults who might otherwise avoid traditional counselling due to stigma, fear of judgment, or logistical barriers such as limited appointment availability. Bendig et al. (2019) similarly noted that the non-judgmental character of AI interactions holds strong appeal for younger users, who tend to be acutely sensitive to the social costs of disclosing mental health difficulties. In this light, the high acceptance observed among students in the present study can be understood not merely as technological openness, but as a response to a genuine unmet need, the need for a form of support that feels safe, private, and available precisely when it is needed most.

Students' positive perceptions of AI counselling effectiveness are well supported by empirical evidence from the intervention literature. Fitzpatrick et al. (2017) found that students who engaged with an AI-based chatbot reported meaningful reductions in depressive symptoms and perceived stress over a two-week period. Ly, Ly, and Andersson (2017) similarly demonstrated that AI-delivered behavioural activation programmes were effective in reducing mild-to-moderate depression, and importantly, found that perceptions of effectiveness were significantly shaped by the quality of empathic responsiveness that the system simulated, suggesting that the experience of feeling understood matters as much as the technical content of the interaction itself. Laranjo et al. (2018), in a systematic review of chatbot interventions, concluded that these tools hold genuine promise particularly for populations who face structural or psychological barriers to accessing traditional support. It is important to situate these perceptions within a realistic understanding of what AI counselling can and cannot offer. The existing literature consistently frames AI-based counselling as most effective for mild-to-moderate psychological concerns, the everyday stressors, academic anxieties, and emotional fluctuations that are characteristic of college life, rather than for severe or complex clinical presentations (Bendig et al., 2019; Torous & Roberts, 2017). The students in this study appear to hold a view consonant with this framing, approaching AI counselling as a meaningful resource for manageable distress rather than as a substitute for deep clinical care. Torous and Roberts (2017) articulate this distinction clearly, arguing that the future of digital mental health lies in blended systems in which AI and human professionals each contribute what they do best. The broad yet nuanced effectiveness perceptions documented in the present study suggest that students are already thinking along these lines, open to AI as a first and accessible step, while remaining aware of the irreplaceable value of human therapeutic relationships.

The significant and positive interrelationships observed among all four constructs are consistent with the theoretical architecture of the Technology Acceptance Model and with broader accounts of human-AI interaction. The particularly strong association between trust and acceptance reinforces Luger and Sellen's (2016) observation that trust is a foundational prerequisite for the adoption of AI systems in personal and emotionally sensitive domains. Glikson and Woolley (2020) further argued that trust mediates the relationship between AI system characteristics and user engagement, a position well supported by the pattern of associations observed in this study. When students feel psychologically safe within an AI counselling environment, their willingness to accept and engage with it increases substantially. The strong association between acceptance and perceived effectiveness points toward a self-reinforcing dynamic in which openness to AI counselling amplifies perceptions of its value, which in turn deepens that openness further. This pattern is consistent with cognitive consistency theory (Festinger, 1957), which holds that individuals tend to align their beliefs and evaluations in mutually supportive ways, and with experiential reinforcement models of technology adoption. McStay (2018) has argued that AI empathy, the perceived capacity of a system to understand and respond meaningfully to emotional states, functions as a catalytic variable that simultaneously strengthens trust, acceptance, and perceived effectiveness. The interconnected structure of student perceptions observed in this study reflects precisely that dynamic, suggesting that investments in empathic responsiveness and psychological safety are likely to produce gains across all four dimensions simultaneously, rather than in any one area alone.

The finding that postgraduate students held more favourable perceptions than undergraduates across all four constructs, with the most pronounced difference appearing in perceived effectiveness, is consistent with what the literature suggests about the role of experience and exposure in shaping attitudes

Paper ID: SR26424120109 DOI: https://dx.doi.org/10.21275/SR26424120109

toward AI in health contexts. Vaidyam et al. (2019) identified prior exposure to AI technologies as a consistent positive predictor of favourable attitudes toward their application in health settings. Postgraduate students, by virtue of their more advanced academic trajectories, are likely to have accumulated greater familiarity with research methodologies, digital tools, and interdisciplinary knowledge domains, including emerging applications of AI in psychology and health care. This accumulated exposure provides a conceptual foundation from which AI counselling can be evaluated more confidently and appraised more favourably. Lucas et al. (2014) similarly observed that greater familiarity with virtual agents was associated with increased willingness to engage in personal disclosure, a finding that offers a plausible explanation for the higher trust and acceptance scores observed among postgraduate students in the present study. The more pronounced gap between undergraduate and postgraduate students in perceived effectiveness is particularly worth noting. It suggests that undergraduate students, while open to AI counselling in principle, may not yet have the conceptual or experiential framework needed to fully appreciate what such tools can offer. This has direct implications for institutional strategy. Awareness initiatives, digital literacy programmes, and psychoeducation efforts that contextualise AI counselling within an evidence-based framework may serve to close this perceptual gap, extending the benefits of AI-assisted support to undergraduate students who are currently less well-positioned to seek it out or trust in its value.

The pattern of gender differences observed in this study is broadly consistent with what the technology adoption literature has documented. Research has consistently shown that male users tend to exhibit greater initial enthusiasm for emerging technologies, particularly during early adoption phases, while female users often approach new technologies with greater deliberation, foregrounding concerns about privacy, reliability, and emotional safety (Bendig et al., 2019; Prochaska et al., 2021). The higher scores recorded by male students on attitudes, acceptance, and perceived effectiveness are interpretable within this well-established pattern. Critically, however, the absence of a significant gender difference in trust is a finding of considerable theoretical interest. Glikson and Woolley (2020) found that trust in high-stakes AI contexts is shaped more by situational factors than by individual characteristics such as gender. The equal caution that both male and female students bring to the question of trusting an AI with their emotional concerns suggests that, when the stakes are personal and the vulnerability is real, gender recedes as a meaningful differentiating factor. Furthermore, Venkatesh and Morris (2000) cautioned against interpreting gender differences in technology adoption as differences in direction rather than degree, a caution that applies directly here. Female students' perceptions were positive across all four dimensions; they were simply more measured in their initial enthusiasm, particularly on dimensions related to attitudes and perceived effectiveness. This distinction between degree and direction has important implications for how AI counselling programmes are designed and communicated to female students, who may benefit from more targeted information about system safety, data confidentiality, and evidence of effectiveness before they fully commit to engagement.

6. Limitations and Directions for Future Research

This study is not without limitations, and an honest acknowledgement of these constraints is essential to a responsible interpretation of its findings. The sample, while adequately sized for the statistical analyses employed, was drawn exclusively from arts and science colleges affiliated with a single university in Tamil Nadu, thereby limiting the generalisability of the findings to other institutional and regional contexts within India. Future research should seek to replicate and extend this study with larger, more geographically and institutionally diverse samples, including students from professional colleges, technical universities, and institutions in regions with different cultural and linguistic profiles. The study's reliance on self-reported perceptual data, collected via structured questionnaire, introduces the possibility of response biases including social desirability effects and the tendency for self-report measures to capture stated rather than enacted behaviour. The degree to which students expressed attitudes, trust, acceptance, and perceived effectiveness translate into actual engagement with AI counselling tools remains an empirical question that the present design cannot resolve. Longitudinal and experimental studies that track students' real-time interactions with specific AI counselling platforms and that measure behavioural outcomes alongside perceptual ones would constitute a valuable and necessary methodological advance.

Additionally, the study examined perceptions of AI - based counselling in the abstract, without requiring participants to interact with a specific AI platform prior to responding. While this approach has the advantage of capturing students' baseline perceptions uncontaminated by idiosyncratic platform effects, it also means that the findings reflect anticipatory rather than experiential judgements. Future studies that incorporate direct exposure to AI counselling tools as an experimental condition would yield data of considerable ecological and practical significance. Finally, the present study did not systematically examine a number of variables that are likely to be theoretically and practically relevant to students' perceptions of AI counselling, including prior mental health help-seeking history, technological literacy, cultural attitudes toward mental illness and disclosure, personality traits such as openness to experience and anxiety sensitivity, and specific concerns about data privacy and algorithmic bias. The integration of these variables into future research frameworks would substantially enrich the conceptual and empirical landscape of this emerging field.

7. Conclusion

The question that gave rise to this study can algorithms counsel? was never intended to be answered with a simple affirmative or negative. What this study has demonstrated is that college students, at least in the context examined, are prepared to entertain that question seriously to approach AI - based counselling not with blanket scepticism or uncritical enthusiasm, but with a measured, contextually sensitive, and broadly positive disposition. Their attitudes are favourable, their acceptance is substantive, their perceptions of effectiveness are meaningful, and even their trust the most cautiously held of the four constructs is directionally positive and statistically significant. What this means, practically and theoretically, is that the groundwork for the responsible integration of AI - based counselling into Indian higher education is more firmly laid than existing literature might have suggested. Students are not the obstacle. The obligation now falls upon institutions, developers, policymakers, and mental health professionals to honour that readiness with tools that are worthy of its systems that are transparent, empathically responsive, ethically sound, and genuinely effective. When that alignment between student readiness and institutional responsibility is achieved, AI - based counselling will not merely be a technological convenience. It will be a meaningful, accessible, and humanising expansion of the support that colleges and universities offer to the students in their care. This study has taken one step toward mapping that alignment. The many steps that remain represent an invitation to researchers, practitioners, and institutions alike to take the mental health of college students as seriously as the intelligence of the algorithms designed to support them.

References

[1] Attkisson, C. C., & Zwick, R. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning, 5*(3), 233–237. https://doi.org/10.1016/0149-7189(82)90074-X

[2] Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., Hasking, P., Murray, E., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Stein, D. J., Vilagut, G., Zaslavsky, A. M., & Kessler, R. C. (2018). WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology, 127*(7), 623–638. https://doi.org/10.1037/abn0000362

[3] Becker, T. D., Saunders, E. C., Shearer, E. M., Budge, S., & Muench, F. (2021). Ethical considerations for mental health clinicians working with adolescents in the digital age. *Current Psychiatry Reports, 23*(8), 50. https://doi.org/10.1007/s11920-021-01265-5

[4] Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2019). The next generation: Chatbots in clinical psychology and psychotherapy to foster mental healthA scoping review. *Verhaltenstherapie, 29*(3), 1–16. https://doi.org/10.1159/000499492

[5] Brave, S., & Nass, C. (2003). Emotion in human–computer interaction. In J. Jacko & A. Sears (Eds.), *The human–computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 81–96). Lawrence Erlbaum Associates.

[6] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

[7] Eisenberg, D., Hunt, J., Speer, N., & Zivin, K. (2011). Mental health service utilization among college students in the United States. *Journal of Nervous and Mental Disease, 199*(5), 301–308. https://doi.org/10.1097/NMD.0b013e3182175552

[8] Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

[9] Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health, 4*(2), e19. https://doi.org/10.2196/mental.7785

[10] Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

[11] Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry, 10*, Article 113. https://doi.org/10.1186/1471-244X-10-113

[12] Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human–robot interaction. *Human Factors, 53*(5), 517–527. https://doi.org/10.1177/0018720811417254

[13] Hudlicka, E. (2011). Guidelines for designing virtual patient advocates. *Journal of Health Communication, 16*(Suppl. 1), 243–265. https://doi.org/10.1080/10810730.2011.604381

[14] Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth, 6*(11), e12106. https://doi.org/10.2196/12106

[15] Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*(1), 53–71. https://doi.org/10.1207/S15327566IJCE040104

[16] Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., & Coiera, E. (2018). Conversational agents in healthcare: A systematic review. *Journal of the American Medical Informatics Association, 25*(9), 1248–1258. https://doi.org/10.1093/jamia/ocy072

[17] Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2*(3), 197–207. https://doi.org/10.1016/0149-7189(79)90094-6

[18] Lattie, E. G., Adkins, E. C., Winquist, N., Stiles-Shields, C., Wafford, Q. E., & Graham, A. K. (2019). Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: Systematic review. *Journal of Medical Internet Research, 21*(7), e12869. https://doi.org/10.2196/12869

[19] Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37*, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

[20] Luger, E., & Sellen, A. (2016). "Like having a really bad PA": The gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing*

Systems (pp. 5286–5297). ACM. https://doi.org/10.1145/2858036.2858288

[21] Ly, K. H., Ly, A. M., & Andersson, G. (2017). A fully automated short-term mindfulness intervention for reducing depression and loneliness: Two randomized controlled trials. *Internet Interventions, 10*, 26–34. https://doi.org/10.1016/j.invent.2017.09.002

[22] McStay, A. (2018). *Emotional AI: The rise of empathic media*. SAGE Publications.

[23] Melas, C. D., Zampetakis, L. A., Dimopoulou, A., & Moustakis, V. (2011). Modelling the acceptance of clinical information systems among hospital medical staff: An extended TAM model. *Journal of Biomedical Informatics, 44*(4), 553–564. https://doi.org/10.1016/j.jbi.2011.01.009

[24] Prochaska, J. J., Vogel, E. A., Chieng, A., Kendra, M., Baiocchi, M., Pajarito, S., & Robinson, A. (2021). A therapeutic relational agent for reducing problematic substance use (Woebot): Development and usability study. *Journal of Medical Internet Research, 23*(3), e24850. https://doi.org/10.2196/24850

[25] Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

[26] Schepman, A., & Rodway, P. (2020). Initial validation of the General Attitudes towards Artificial Intelligence Scale. *Computers in Human Behavior Reports, 1*, Article 100014. https://doi.org/10.1016/j.chbr.2020.100014

[27] Shum, H. Y., He, X. D., & Li, D. (2018). From Eliza to XiaoIce: Challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering, 19*(1), 10–26. https://doi.org/10.1631/FITEE.1700826

[28] Torous, J., & Hsin, H. (2018). Empowering the digital therapeutic relationship: Virtual clinics for digital health interventions. *npj Digital Medicine, 1*, Article 16. https://doi.org/10.1038/s41746-018-0028-2

[29] Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and conversational agents in mental health: A review of the psychiatric landscape. *The Canadian Journal of Psychiatry, 64*(7), 456–464. https://doi.org/10.1177/0706743719828977

[30] Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

[31] Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS Quarterly, 24*(1), 115–139. https://doi.org/10.2307/3250981

[32] Woebot Group. (2021). *Woebot for adults: Evidence base*. https://woebothealth.com/evidence-base/

[33] World Health Organization. (2022). *World mental health report: Transforming mental health for all*. https://www.who.int/publications/i/item/9789240049338