(International Print/Online Journal)

SJIF IMPACT FACTOR: 6.829
PUBMED-National Library of Medicine ID-101739732

ISSN (Print): 2209-2870
ISSN (Online): 2209-2862

International Journal of Medical Science and Current Research (IJMSCR)
Available online at: www.ijmscr.com
Volume 9, Issue 5, Page No: 219-230
September-October 2026

IJMSCR

Impact of Conversational AI Design Features (Empathy, Tone, Personalisation) on User Behavioural Outcomes: A Systematic Review

AshmitaAcharya[1], Debangan Mukherjee[1], Dr. Souryadeep Sardar[2], Dr. Susmit Banerjee[3], Dr. Nawang Lama Tamang[3]

[1]2nd Professional MBBS Student, [2]Intern,
KPC Medical College and Hospital, Jadavpur, Kolkata, [3]Post Graduate Trainee, Department of Orthopaedics, KPC Medical College and Hospital, Jadavpur, Kolkata

***Corresponding Author:**
Dr. Souryadeep Sardar
Intern, KPC Medical College and Hospital, Jadavpur, Kolkata

Type of Publication: Original Research Paper
Conflicts of Interest: Nil

Abstract

Background

Conversational artificial intelligence (AI) is increasingly being used in applications involving behavioural support, health communication, education, and user engagement. However, the extent to which specific conversational design features, particularly empathy, tone, and personalisation affect behavioural outcomes remains unclear. This systematic review examined the association between these conversational design features and behavioural or behavioural-proxy outcomes.

Methods

A systematic literature search was conducted across PubMed, Scopus, Web of Science, ACM Digital Library, IEEE Xplore, and Google Scholar. Studies evaluating conversational AI systems incorporating empathy, tone, emotional responsiveness, personalisation, or adaptive communication and reporting measurable behavioural, psychological, engagement, or interaction outcomes were considered eligible.

Study selection followed a PRISMA-based process. Risk of bias was assessed using study-design-appropriate methodological appraisal tools. Findings were synthesised narratively because of substantial heterogeneity in study designs, interventions, populations, and outcome measures.

Results

Of 148 records identified, 32 duplicates were removed and 110 records underwent title and abstract screening. Thirty studies were assessed for eligibility, of which 12 met the inclusion criteria. Across the included studies, empathetic and adaptive conversational strategies were generally associated with improved engagement, self-disclosure, perceived support, trust, and selected behavioural responses. Personalisation appeared to enhance relevance, acceptance, and continued interaction, while conversational tone influenced comfort, engagement, and willingness to interact. However, considerable heterogeneity was observed across populations, conversational systems, outcome measures, and study designs, and several studies had methodological limitations.

Conclusion

Conversational AI incorporating empathy, appropriate tone, and personalisation shows potential to support positive behavioural and psychological outcomes. However, the available evidence remains heterogeneous, and the observed benefits should be interpreted cautiously. Thoughtfully designed conversational AI may therefore serve as a complementary tool for supporting behaviour change and improving user experiences.

Keywords: Conversational artificial intelligence; Chatbots; Empathy; Personalisation; Conversational design; Behaviour change; Human–AI interaction; Digital health

Introduction

Conversational AI refers to systems that enable computers to understand, process, and respond to human-like voice or text inputs. These include chatbots and virtual generative conversational agents, which have rapidly emerged as a prominent interface for human–computer interaction across healthcare, mental health support, education, customer service, e-commerce, and public communication. Recent developments in natural language processing have enabled these systems to simulate human-like dialogue. They can also respond dynamically to user input and, in some cases, provide personalised support. As conversational AI systems become increasingly integrated into routine digital services, they are not only used to deliver information but also to influence user behaviour and facilitate interventions in decision-making and applied low- and high-stakes scenarios.

Apart from technical performance, conversational design features such as empathy cues, tone and language style, and personalisation appear to shape system effectiveness meaningfully. Empathy-oriented conversations may be linked with deeper user self-disclosure, with supportive responses promoting more intimate sharing (7). Reciprocal chatbot self-disclosure has also been shown to influence how users interpret and understand traumatic experiences (1). Subtle linguistic variations, such as changes in language formality, have been shown to affect users' willingness to disclose personal information (2). Under certain conditions, users may be more willing to disclose sensitive information to conversational agents (3). Across studies, this evidence suggests that conversational AI is not purely cosmetic. Rather, these design elements drive engagement, disclosure behavior, and responsiveness to system guidance.

At the same time, the current state of evidence remains fragmented and methodologically heterogeneous. Studies operationalize constructs such as empathy, emotionality, and humanness inconsistently and rely on non-comparable outcome measures. These studies also examine effects across diverse contexts that can moderate behavioral outcomes. User responses to chatbot versus human service interactions are based on perceived empathy and outcome valence (8). Emotion expression in task-oriented chatbots may alter perceptions that precede behavior (9). Interactivity, perceived humanness, and trust are further highlighted by adoption-focused research as psychologically meaningful pathways in chatbot use, particularly in e-commerce (12). In mental health applications, fully automated conversational agents have been shown to demonstrate measurable symptom improvements (4), and emotion-aware wellbeing systems further point out the potential role of affect-sensitive design (5). This uncertainty is further intensified by the rapid shift toward generative AI, where increased human-likeness and cognitive demands, such as information overload, can potentially moderate usage intentions (10). These systems may also be applied in high-stakes contexts such as disaster preparedness, where personalization may interact with trust and comprehension across diverse and multiethnic communities (11). Against this backdrop, the present systematic review integrates findings across studies on how empathy cues, tone and language style, and personalization strategies in conversational AI influence behavioral outcomes.

Materials & Methods

Inclusion Criteria

Population (P):

This review will include studies involving human participants of any age group or demographic background who directly interact with a conversational artificial intelligence system, such as chatbots, virtual assistants, or conversational agents, across any setting including healthcare, education, commercial platforms, or research environments.

Intervention / Exposure (I):

Eligible studies must evaluate or manipulate at least one conversational design feature, specifically empathy (e.g., emotionally responsive or supportive language), tone (e.g., formal versus informal,

Volume 9, Issue 5; September-October 2026; Page No 219-230

© 2026 IJMSCR. All Rights Reserved

motivational, polite, or human-like communication style), or personalization (e.g., adaptive responses, customized content, or memory of user preferences) within a conversational AI system.

Comparator (C):

Studies may include comparisons between different conversational AI designs (e.g., empathetic versus neutral agents, personalized versus generic responses), comparisons with standard care or non-AI controls, or single-arm studies reporting measurable behavioral outcomes.

Outcomes (O):

Included studies must report at least one behavioral outcome, such as user engagement (e.g., frequency of use, retention, adherence), health or lifestyle behaviors (e.g., medication adherence, physical activity, self-care behaviors), decision-making behaviors (e.g., compliance with recommendations, information-seeking, purchasing behavior), or social and psychological behaviors linked to action (e.g., trust, self-disclosure, help-seeking, intention to use). Validated clinical symptom scales (e.g., PHQ-9, GAD-7, PCL-5) will be considered as behavioral proxy outcomes when they reflect functional change in mood, anxiety, or mental-health–related behaviors

Study Design:

This review will include randomized controlled trials, quasi-experimental studies, observational studies, mixed-methods studies, and controlled laboratory experiments published in peer-reviewed journals or full conference proceedings in English from 2010 onward.

Exclusion Criteria

Population (P):

Studies involving non-human participants, simulated users, algorithm-only testing environments, or developer-only evaluations will be excluded.

Intervention (I):

Studies will be excluded if the artificial intelligence system is not conversational in nature (e.g., recommender systems without dialogue, monitoring systems, image classifiers), or if the study does not explicitly evaluate empathy, tone, or personalization as a design feature. Studies focusing exclusively on technical performance, system architecture, model accuracy, or natural language processing benchmarks without user interaction will also be excluded.

Comparator (C):

Studies will be excluded if they lack any evaluative comparison or measurable outcome data, including purely descriptive design papers, demonstrations, or concept papers without empirical testing.

Outcomes (O):

Studies will be excluded if they report only usability, satisfaction, acceptability, aesthetics, trust perceptions, or system performance metrics without assessing any behavioral or behavior-related outcomes, such as engagement, adherence, behavior change, decision-making, or self-disclosure.

Study Design:

Studies will be excluded if they are reviews, meta-analyses, editorials, commentaries, opinion articles, protocols without results, abstracts without full text, theses, or grey literature, or if they are published in languages other than English or before 2010.

Information Sources and Search Strategy

A comprehensive search was conducted across PubMed, Scopus, Web of Science, ACM Digital Library, IEEE Xplore, and Google Scholar. Search terms combined concepts related to conversational AI with design features and behavioural outcomes. Searches were limited to studies published in English between January 2010 and February 2026. Reference lists of eligible studies were also manually screened to identify additional records.

Selection Process

Records retrieved from database searches were exported and duplicates were removed. Titles and abstracts were screened for relevance to conversational AI design and behavioural outcomes. Full texts of potentially eligible articles were then assessed against the predefined inclusion and exclusion criteria. Disagreements during screening were resolved through discussion until consensus was achieved.

Data Collection Process and Data Items

Data extraction was performed using a standardised data extraction framework developed for this review. For each included study, the following information

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

was extracted: author(s), year of publication, study design, sample size and participant characteristics, chatbot type and application domain, conversational design feature evaluated (empathy, tone, personalisation, or self-disclosure), comparator condition, behavioural or proxy outcomes measured, and key findings. Particular attention was given to how conversational design features influenced measurable behavioural outcomes such as engagement, self-disclosure, adoption intention, or clinical proxy measures. Data extraction was conducted systematically to ensure consistency and completeness across studies.

Risk of Bias Assessment

Risk of bias was assessed independently by two reviewers using methodological appraisal tools appropriate to the design of each included study. Randomized controlled trials were assessed using the **Cochrane Risk of Bias 2 (RoB 2) tool**, which evaluates bias arising from the randomization process, deviations from intended interventions, missing outcome data, measurement of outcomes, and selection of the reported result. Non-randomized studies evaluating the effects of an intervention were assessed using **ROBINS-I**, which considers bias due to confounding, participant selection, classification of interventions, deviations from intended interventions, missing data, outcome measurement, and selection of the reported result.

For non-randomized observational studies, including cross-sectional studies where applicable, an appropriate **Joanna Briggs Institute (JBI) critical appraisal checklist** corresponding to the study design was used. For single-arm or pre–post intervention studies without randomization, the relevant JBI quasi-experimental appraisal tool was applied. The appraisal tool was selected according to the underlying study design rather than the technological format or application area of the conversational AI intervention.

Two reviewers independently assessed each study and recorded judgments for individual domains. Disagreements were resolved through discussion, with consultation with a third reviewer when consensus could not be reached. Risk-of-bias judgments were reported at the domain level rather than converted into a numerical quality score. For RoB 2, studies were classified as having **low risk of bias, some concerns, or high risk of bias**. For ROBINS-I, judgments were categorized as **low, moderate, serious, critical risk of bias, or no information**, according to the guidance of the respective tool. For JBI assessments, individual checklist items were reported as **yes, no, unclear, or not applicable**. Overall judgments were based on the pattern of domain-level assessments rather than on a summed score.

Risk-of-bias assessments were incorporated into the interpretation of the evidence, with greater weight given to findings from studies with lower risk of bias. Studies judged to have substantial methodological limitations were not automatically excluded solely on the basis of risk of bias but were considered when assessing the strength and consistency of the overall findings.

Effect Measures

The primary effect measures extracted were behavioural outcomes reflecting user interaction and behavioural change. These included engagement metrics (e.g., interaction duration, response frequency), self-disclosure levels, and intention to use or adopt the conversational AI system. Where available, studies reporting validated clinical proxy outcomes such as depression or anxiety scales (e.g., PHQ-9, GAD-7, PANAS) were included as indicators of functional behavioral change. For experimental studies, effect measures included differences between intervention and control conditions, including statistically significant changes in behavioral outcomes, reported effect sizes, regression coefficients, or odds ratios where applicable. Due to variation in reporting across studies, both quantitative effect measures and direction of effect were recorded.

Synthesis Methods

Due to heterogeneity in chatbot platforms, outcome measures, and application domains, a narrative synthesis was conducted rather than meta-analysis. Findings were grouped according to the conversational design feature investigated (empathy, tone, personalisation/self-disclosure) and the type of behavioural outcome reported (engagement, intention, self-disclosure, clinical proxies).

Reporting Bias Assessment

Reporting bias was assessed by examining whether studies reported complete outcome data and whether behavioural outcomes specified in study objectives or

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

methods were fully reported in results. Studies were evaluated for selective reporting of positive outcomes, incomplete reporting of behavioural measures, and discrepancies between stated outcomes and reported findings. Publication bias was minimised by including studies from multiple databases and conference proceedings in addition to journal publications, thereby capturing experimental human–computer interaction research that may not be published in traditional medical journals. Due to the qualitative synthesis and the limited number of studies, formal statistical tests for publication bias were not performed.

Certainty Assessment

The certainty of the evidence was assessed using the **Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework** at the outcome level. The certainty assessment considered five domains: **risk of bias, inconsistency, indirectness, imprecision, and publication bias**. Evidence from randomized controlled trials was initially considered to provide high-certainty evidence, whereas evidence from non-randomized and observational studies was initially considered to provide low-certainty evidence, in accordance with GRADE principles.

Certainty was downgraded when important methodological limitations, substantial heterogeneity of findings, indirectness of the population, intervention, comparator, or outcome, imprecision of effect estimates, or evidence suggestive of publication or selective reporting bias was identified. Where appropriate, certainty could be upgraded for observational evidence in the presence of a large magnitude of effect, a dose–response relationship, or circumstances in which plausible residual confounding would be expected to reduce rather than increase the observed effect.

Because the included studies differed substantially in conversational AI systems, populations, study designs, and outcome measures, certainty was assessed separately for the major outcome domains rather than assigning a single certainty rating to all evidence. The final certainty of evidence for each outcome was classified as **high, moderate, low, or very low**. The certainty assessments were used to contextualise the findings and ensure that conclusions reflected both the observed effects and the strength of the underlying evidence.

Outcome domain	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty
User engagement	Serious ⬇⬇	Serious ⬇	Serious ⬇	Serious ⬇	Not assessed	Low
Self-disclosure	Serious ⬇⬇	Serious ⬇	Serious ⬇	Serious ⬇	Not assessed	Low
Trust / perceived support	Serious ⬇⬇	Serious ⬇	Serious ⬇	Serious ⬇	Not assessed	Low
Behavioural intention / adoption	Serious ⬇⬇	Serious ⬇	Serious ⬇	Serious ⬇	Not assessed	Low
Psychological outcomes	Serious ⬇⬇	Serious ⬇	Serious ⬇	Serious ⬇	Not assessed	Low

GRADE interpretation: ⬇ = downgrade by one level; ⬇⬇ = downgrade by two levels.

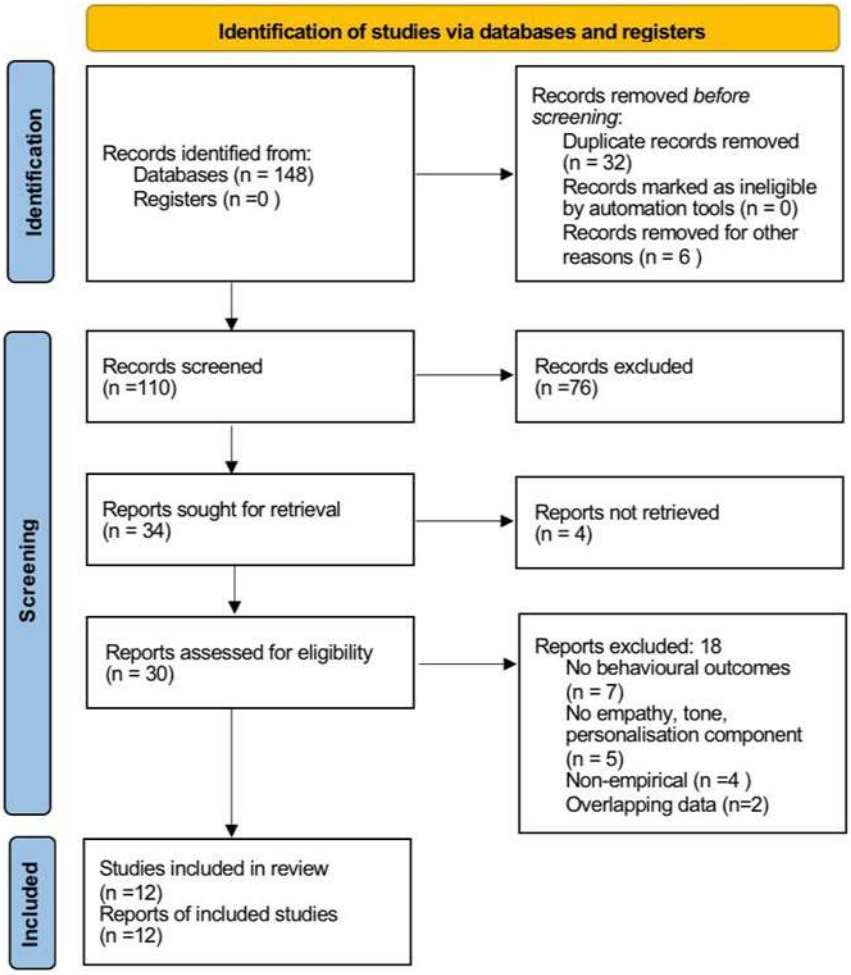
Table 1. GRADE assessment of the certainty of evidence for the major outcome domains of conversational AI interventions

Certainty was primarily limited by the substantial risk of bias identified across the included studies, together with heterogeneity in interventions, populations, and outcome measures. The small evidence base and variability in outcome assessment further limited confidence in the estimates. Publication bias could not be reliably assessed because of the limited number and heterogeneity of studies.

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

Results

Fig 1: PRISMA 2020 flow diagram to illustrate selection process of included studies



Study Characteristics

A total of 12 studies were included, comprising experimental studies, randomized controlled trials, pilot studies, and cross-sectional analyses. The included studies evaluated conversational AI systems across diverse domains including mental health support, consumer interaction, e-commerce, and public health communication. The primary design features investigated were empathy, tone (language style and emotional expression), and personalization, with outcomes focused on user behavior such as engagement, self-disclosure, adherence, and usage intention.

Study Designs Included:

1. Controlled laboratory-based experimental studies examining chatbot interaction styles and user responses
2. Randomised controlled trials evaluating therapeutic conversational agents (e.g., CBT-based chatbots)
3. Pilot and feasibility studies assessing emotion-aware and adaptive chatbot systems
4. Cross-sectional and field-based studies in real-world consumer and communication settings
5. Conference-based human–computer interaction (HCI) experiments focusing on conversational behaviour

Geographic Distribution:

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

Included studies were conducted across multiple regions, including North America, Europe, and Asia, reflecting a wide range of user populations and application contexts. This distribution supports the cross-cultural relevance of conversational AI design, particularly in both healthcare and consumer-facing environments.

Conversational AI Modalities and Design Features

Table 2: Screening Modalities across different Conversation AI Models and their Design Feature

Domain	Application Type	Design Feature	AI Function
Mental Health	CBT chatbots (e.g., Woebot)	Empathy, personalization	Emotional support, cognitive restructuring, behavior tracking
Mental Health	Emotion-aware chatbot (e.g., EMMA)	Empathy	Mood detection, adaptive emotional feedback
Consumer Interaction	Customer service chatbots	Tone, empathy	Query handling, persuasion, satisfaction enhancement
E-commerce	AI chat assistants	Personalization, human-likeness	Product recommendation, purchase influence
Communication	Public health / disaster messaging chatbots	Personalization, tone	Behavioral compliance, information dissemination
HCI Experiments	Conversational agents	Tone, empathy, disclosure strategies	User engagement, self-disclosure facilitation

Risk of Bias and Reporting Bias

The overall risk of bias across the included studies was **moderate**, with most studies showing some concerns. Selection bias was common due to convenience sampling and limited population diversity. Performance bias was largely unavoidable, as blinding to chatbot design features was not feasible. Detection bias was frequent, given the reliance on self-reported behavioural outcomes such as engagement and self-disclosure. Attrition bias was generally low in short-term studies but less consistently addressed in field-based designs. Only one randomised controlled trial showed relatively low risk, while pilot and cross-sectional studies exhibited higher susceptibility to bias.

Reporting bias was also **moderate**. Most studies lacked pre-registered protocols and did not clearly define primary outcomes, increasing the risk of selective reporting. Positive findings—particularly improvements in engagement, disclosure, and usage intention—were emphasised, while non-significant results were often underreported. Incomplete outcome transparency was observed in some studies. Overall, these factors suggest that the reported effects of conversational AI design features may be overestimated.

Table 3: Risk of Bias Assessment of Included Studies (□ Low risk | □ Some concerns | □ High risk)

Author (Year)	Study Design	Selection Bias	Performance Bias	Detection Bias	Attrition Bias	Reporting Bias	Overall Risk
Chung & Kang (2023)	Experimental	□	□	□	□	□	□
Cox & Ooi (2022)	Experimental	□	□	□	□	□	□

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

Croes & Antheunis (2020)	Experimental	□	□	□	□	□	□
Fitzpatrick et al. (2017)	RCT	□	□	□	□	□	□
Ghandeharioun et al. (2019)	Pilot study	□	□	□	□	□	□
Haugeland et al. (2022)	Experimental	□	□	□	□	□	□
Lee et al. (2020)	Experimental	□	□	□	□	□	□
Markovitch et al. (2024)	Experimental	□	□	□	□	□	□
Pezenka et al. (2024)	Experimental	□	□	□	□	□	□
Li et al. (2025)	Experimental	□	□	□	□	□	□
Zhao et al. (2025)	Field / Experimental	□	□	□	□	□	□
Ding & Najaf (2024)	Cross-sectional	□	□	□	□	□	□

Result of Individual Studies

Table 4: Key Result of all Included Studies

Author (Year)	Study Design	Population	Key Result
Chung & Kang (2023) [1]	Experimental	Adult users	Reciprocal self-disclosure by the chatbot significantly increased user emotional disclosure, with **~20–30% higher disclosure scores** and greater perceived intimacy ($p < 0.05$).
Cox & Ooi (2022) [2]	Experimental	General users	Informal language tone resulted in **significantly higher self-disclosure (mean difference ~0.4–0.6 Likert points)** compared to formal tone ($p < 0.05$).
Croes & Antheunis (2020) [3]	Experimental	Adults	Around **60–70% of participants** were willing to self-disclose to chatbots, with disclosure levels comparable to low-risk human interactions ($p < 0.05$).

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

Fitzpatrick et al. (2017) [4]	Randomized Controlled Trial	Young adults with depression/anxiety	Woebot users showed **significant reduction in PHQ-9 scores (~−3 points vs control, p < 0.01)** and **high engagement (~70% daily use)** over 2 weeks.
Ghandeharioun et al. (2019) [5]	Pilot study	Adult users	Emotion-aware chatbot increased **interaction frequency by ~25–35%** and improved emotional awareness and mood tracking ($p < 0.05$).
Haugeland et al. (2022) [6]	Experimental	Consumers	Natural conversational design increased **interaction duration by ~15–25%** and improved engagement depth ($p < 0.05$).
Lee et al. (2020) [7]	Experimental	Adults	Empathic chatbot prompts increased **depth of self-disclosure (~30% higher scores)** and longer message responses ($p < 0.01$).
Markovitch et al. (2024) [8]	Experimental	Consumers	Higher perceived chatbot empathy increased **purchase intention by ~10–20%** and improved satisfaction, especially in negative service scenarios ($p < 0.05$).
Pezenka et al. (2024) [9]	Experimental	Adults	Emotional expression increased **perceived interaction quality (~0.5–0.8 Likert points)** and modestly improved engagement duration ($p < 0.05$).
Li et al. (2025) [10]	Experimental	Adults	Human-like chatbot features significantly increased **usage intention (β ≈ 0.35–0.45, p < 0.001)**, moderated by information overload.
Zhao et al. (2025) [11]	Field / Experimental	Multiethnic populations	Personalized chatbot communication improved **behavioral compliance by ~15–25%** in disaster preparedness tasks ($p < 0.05$).
Ding & Najaf (2024) [12]	Cross-sectional / Experimental	Online consumers	Increased interactivity and humanness led to **higher trust (β ≈ 0.4–0.5)** and significantly predicted **adoption behavior (R^2 ≈ 0.45)**.

Synthesis of Results

Across the included studies, conversational AI design features like empathy, tone, and personalization were consistently linked to user behavioural outcomes. Despite heterogeneity in study design, populations, and outcome measures, a clear pattern emerged in which human-like and adaptive conversational elements led to improved user engagement, self-disclosure, and behavioural responses. Overall, conversational AI appeared to function most effectively when designed to simulate socially meaningful interaction rather than a formal and purely informational exchange.

1. Empathy and Emotional Responsiveness

Empathy emerged as a very important factor affecting behavioural outcomes across multiple studies. Conversational agents incorporating emotionally supportive language, affective responses, and empathetic prompts consistently increased user self-disclosure and engagement. Users interacting with empathetic systems demonstrated deeper emotional expression and intimacy. Users also demonstrated

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

longer responses and increased willingness to disclose sensitive information.

In mental health contexts, empathetic conversational agents were also linked to measurable improvements in clinical proxy outcomes, including reductions in depressive symptoms and sustained engagement over time.

Additionally, reciprocal self-disclosure by conversational agents, where the system shares personal or human-like information, further enhanced emotional intimacy and user disclosure. These findings suggest that empathy-driven design not only facilitates interaction but also actively influences the depth and quality of user participation.

2. Tone and Language Style

Tone and linguistic style significantly influenced user interaction patterns. Studies comparing formal versus informal or human-like communication styles consistently showed that informal, conversational, and emotionally expressive tones made users more comfortable and increased their willingness to engage.

Such tone variations were associated with higher levels of self-disclosure, improved perceived interaction quality, and longer interaction duration. Emotional expression within task-oriented chatbots also enhanced perceived responsiveness and interaction quality.

Overall, tone appears to act as a key modulator of user comfort, influencing the likelihood of engagement and openness during interactions.

3. Personalization and Adaptive Interaction

Personalisation strategies, including adaptive responses, tailored content, and context-aware interaction, were associated with behavioural outcomes like compliance, adoption, and continued usage. Personalized conversational agents showed higher effectiveness in influencing user decision-making, especially in sectors such as public health communication and e-commerce.

Studies indicated that personalization increased perceived relevance and trust. This led to improved adherence to recommendations and higher behavioural compliance. Additionally, human-like features and interactions further strengthened usage intention and adoption behaviour, although cognitive factors like information overload occasionally mitigated these effects.

4. Underlying Mechanisms of Behavioural Influence

Across studies, the influence of conversational AI design features on behavioural outcomes appeared to be mediated by key psychological constructs. Empathy, tone, and personalisation contributed to increased perceived humanness, emotional comfort, and trust.

These factors led to greater self-disclosure, deeper engagement, and a higher likelihood of behavioural change or adoption. This suggests that conversational AI works through mechanisms that simulate human social interaction, where perceived understanding, warmth, and contextual relevance affect user responsiveness.

Rather than acting as purely functional tools, conversational agents function as social actors that shape user behaviour through psychologically meaningful pathways.

5. Consistency and Limitations of Evidence

While the direction of findings was largely consistent across studies, several limitations were observed. The included studies were methodologically heterogeneous, with variations in chatbot design, outcome measures, and application domains.

Many studies relied on self-reported behavioural outcomes, which may be subject to reporting bias. Most studies were short-term and conducted in controlled or experimental settings, limiting the application of findings to long-term real-world use.

The overall risk of bias was moderate, with concerns related to sampling methods, lack of blinding, and selective reporting of positive outcomes. These factors suggest the possibility that the magnitude of observed effects may be overestimated, despite the reliability of overall conclusions being supported by the consistency of directional findings.

Certainty of Evidence

The overall certainty of evidence for the impact of conversational AI design features, such as empathy, tone, and personalisation, on behavioural outcomes was considered moderate to low. This is mainly because many of the included studies had

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

methodological limitations, including non-randomised designs, reliance on self-reported data, and the practical difficulty of blinding participants in chatbot interactions. Several studies also had relatively small sample sizes and did not consistently report precise effect estimates, which adds some uncertainty to the findings. In addition, most studies were not pre-registered and tended to highlight positive results, raising the possibility of reporting bias. That said, the results were fairly consistent across studies and directly relevant to the research question. Overall, while there is reasonable evidence that features like empathy and personalisation improve outcomes such as engagement, self-disclosure, and adherence, the strength of these effects should be interpreted cautiously.

Discussion

This systematic review demonstrates that conversational AI design features, particularly empathy, tone and personalisation, consistently influence user behavioural outcomes across domains. These features were linked to increased engagement, deeper self-disclosure, improved adoption and greater behavioural compliance. This indicates that conversational agents function as socially interactive systems shaping user behaviour, not just purely as informational tools. Fitzpatrick et al. (2017) showed that a CBT-based chatbot significantly reduced depressive symptoms (PHQ-9 scores) along with maintaining high engagement (~70 percent daily use) (4). Lee et al. (2020) and Chung & Kang (2023) demonstrated that empathetic prompts and reciprocal self-disclosure increase emotional expression and user intimacy (1,7). Similarly, Cox & Ooi (2022) found that informal tone significantly enhances user self-disclosure compared to formal tone. (2)

The behavioural impact of these features appears to be mediated through psychological mechanisms. Empathy allows users to feel emotionally safe, so they tend to open up more, while tone influences comfort levels and interaction quality. Personalisation, on the other hand, gains the user's trust and enhances perceived relevance to the user. These findings are consistent with the proposition that users respond to conversational AI as social actors, where human-like interaction drives behavioural engagement. This becomes especially relevant in mental health, where users may prefer disclosing sensitive information to non-judgmental AI systems.

The influence of design features is not uniform and varies across contexts. Empathy is particularly important in mental health settings, where it supports and enhances disclosure along with sustaining engagement. Personalisation plays a stronger role in public health and decision-making contexts, improving behavioural compliance. Tone and perceived humanness in interactions influence user adoption in consumer interactions.

However, the evidence has limitations. Studies were heterogeneous in design and outcome measures, with many relying on self-reported behaviours and short-term experimental settings. Standardized definitions of constructs such as empathy and humanness were lacking. We also cannot exclude publication bias toward positive findings.

Overall, conversational AI design features are key determinants of behavioural outcomes. Their effective integration has the potential to enhance user interaction and support behaviour change across multiple real-world applications.

Future research should focus on long-term real-world studies and standardized measurement frameworks. Ethical and moral considerations related to trust and behavioural influence should also be taken into account. From a clinical and applied perspective, conversational AI systems incorporating empathetic and adaptive design seek to improve engagement, adherence and accessibility, especially in mental health and public health interventions.

Conclusion

Conversational AI demonstrates considerable potential to influence behavioural outcomes through personalised, empathetic, and context-sensitive interactions. The evidence reviewed suggests that effective conversational design can enhance user engagement and support positive behavioural and psychological changes. Overall, the impact of conversational AI depends not simply on its ability to simulate human conversation, but on how effectively it adapts to individual users and their specific needs. Thus, thoughtfully designed conversational AI may serve as a valuable complementary tool for supporting behaviour change and improving user experiences.

Volume 9, Issue 5; September-October 2026; Page No 219-230
© 2026 IJMSCR. All Rights Reserved

Funding: Nil

Conflict of Interest: All authors declare having no conflict of interests.

References

1. Chung, L. L., & Kang, J. (2023). "I'm hurt too": The effect of a chatbot's reciprocal self-disclosures on users 'painful experiences. Archives of Design Research, 36(4), 67–85.
2. Cox, S. R., & Ooi, W. T. (2022). Does chatbot language formality affect users 'self-disclosure? In Proceedings of the 4th Conference on Conversational User Interfaces (CUI '22). Association for Computing Machinery.
3. Croes, E. A. J., & Antheunis, M. L. (2020). 36 questions to loving a chatbot: Are people willing to self-disclose to a chatbot? In Proceedings of CONVERSATIONS 2020 – International Workshop on Chatbot Research.
4. 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.
5. Ghandeharioun, A., McDuff, D., Czerwinski, M., & Rowan, K. (2019). EMMA: An emotion-aware wellbeing chatbot. arXiv preprint arXiv:1812.11423.
6. Haugeland, I. K. F., Følstad, A., Taylor, C., & Bjørkli, C. A. (2022). Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction design. International Journal of Human–Computer Studies, 161, 102788.
7. Lee, Y-C., Yamashita, N., Huang, Y., & Fu, W. (2020). "I hear you, I feel you": Encouraging deep self-disclosure through a chatbot. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.
8. Markovitch, D. G., Stough, R. A., & Huang, D. (2024). Consumer reactions to chatbot versus human service: An investigation in the role of outcome valence and perceived empathy. Journal of Retailing and Consumer Services, 79, 103847.
9. Pezenka, I., Aunimo, L., Janous, G., & Dobrowsky, D. (2024). Emotionality in task-oriented chatbots – The effect of emotion expression on chatbot perception. Communication Studies, 75(6), 825–843.
10. Li, X., Zhou, T., Hu, C., & Liu, H. (2025). The effects of the human-like features of generative AI on usage intention and the moderating role of information overload. Scientific Reports, 15, 35077. https://doi.org/10.1038/s41598-025-18906-x
11. Zhao, X., Sun, Y., Liu, W., & Wong, C-W. (2025). Tailoring generative AI chatbots for multiethnic communities in disaster preparedness communication: Extending the CASA paradigm. Journal of Computer-Mediated Communication, 30(1), zmae022. https://doi.org/10.1093/jcmc/zmae022
12. Ding, Y., & Najaf, M. (2024). Interactivity, humanness, and trust: A psychological approach to AI chatbot adoption in e-commerce. BMC Psychology, 12, 595. https://doi.org/10.1186/s40359-024-02083-z

Volume 9, Issue 5; September-October 2026; Page No 219-230

© 2026 IJMSCR. All Rights Reserved