# Transparency and Trust: The Impact of AI-Assisted and AI-Dominated Communication on Message Perception and Acceptance

Bryce P. Towne[1]*

[1] Yiwu Industrial & Commercial College

* Corresponding author: Bryce P. Towne (brycepetofitowne@gmail.com).

## Abstract

This study investigates the impact of artificial intelligence (AI) involvement on the acceptance, perceived sincerity, and emotional richness of text-based communication. Building on Towne's (2024) framework, which distinguishes between AI-Assisted Communication (AI-AC) and AI-Dominated Communication (AI-DC), the study explores how these methods affect human perceptions of message quality. Two comprehensive studies were conducted using participants from the Credamo platform. Study 1 employed a between-subjects design to compare the impacts of AI involvement and disclosure conditions, while Study 2 utilized a within-subjects design to assess changes in perceptions before and after AI involvement disclosure. The findings reveal that while AI involvement does not significantly impact message acceptance, it reduces perceived sincerity and emotional richness, particularly when AI's role is disclosed. The study also highlights demographic influences, with gender and age affecting trust in AI-generated content. These results underscore the importance of transparency in AI-mediated communication and suggest strategies for managing AI disclosure to maintain trust. The research contributes to the growing body of knowledge on AI-mediated communication by validating Towne's framework and providing practical insights for integrating AI in communication processes.

*Keywords*: AI-mediated communication, AI-Assisted Communication, AI-Dominated Communication, trust, perceived sincerity, emotional richness, transparency, human-AI interaction, message acceptance, Large Language Models.

## Introduction

The integration of artificial intelligence (AI) into various communication processes has significantly transformed how messages are crafted and perceived. Building on the foundational concepts introduced by Towne (2024) in exploring AI-Mediated Communication (AI-MC) and Human-Artificial Intelligence Communication (HAIC), this study delves deeper into the specific impacts of AI involvement on the acceptance, perceived sincerity, and emotional richness of text-based messages. Towne's framework distinguishes between AI-Assisted Communication (AI-AC), where AI enhances human-generated content, and AI-Dominated Communication (AI-DC), where AI predominantly generates original content with human oversight. This delineation raises profound questions about the nature of communication in the age of advanced AI, particularly concerning authenticity, creativity, and trust.

The rapid advancements in Large Language Models (LLMs) and Generative AI technologies have blurred the lines between human and machine roles in communication, necessitating a more nuanced understanding of their impacts. This paper aims to empirically investigate how AI-AC and AI-DC affect human perceptions of message quality, focusing on three key dimensions: acceptance, perceived sincerity, and emotional richness. Additionally, the study examines the role of transparency in AI involvement, exploring how prior disclosure and the timing of such disclosure influence these perceptions.

Building on the conceptual framework proposed by Towne (2024), this research addresses several critical hypotheses. First, it explores whether the method of communication (human-generated, AI-assisted, or AI-dominated) significantly affects the acceptance, sincerity, and emotional richness of messages such as thank-you letters, apology letters, and invitations. Second, it investigates the impact of prior disclosure of AI involvement, hypothesizing that transparency will alter how recipients perceive these messages. Third, it

examines the interaction between the method of communication and disclosure, assessing whether the combination of these factors has a compounded effect on perceptions. Finally, the study compares AI-AC and AI-DC to determine if there are distinct differences in how these methods are received, particularly after the disclosure of AI involvement.

To test these hypotheses, two comprehensive studies were conducted using participants recruited from the Credamo platform. Study 1 employed a between-subjects design to compare acceptance, perceived sincerity, and emotional richness across different communication methods and disclosure conditions. Study 2 utilized a within-subjects design to assess changes in participant perceptions before and after AI involvement was disclosed. The results provide empirical evidence on the nuanced impacts of AI in communication, offering practical insights for organizations on how to manage AI disclosure to maintain trust and enhance communication effectiveness.

In conclusion, this paper aims to advance understanding of AI's role in communication by empirically validating Towne's conceputual framework on AI-AC and AI-DC. By exploring the effects of AI-AC and AI-DC on message perception and the critical role of transparency, this study contributes to the growing body of knowledge on AI-mediated communication and offers actionable recommendations for integrating AI in ways that preserve authenticity and trust.

## Literature Review

The integration of artificial intelligence (AI) into communication processes has significantly altered how messages are crafted and perceived. Research into AI-mediated communication (AI-MC) has highlighted the complex relationship between AI involvement and trust. Trust in AI is multifaceted and influenced by factors such as AI reliability, transparency, and user characteristics. Hoff and Bashir (2015) emphasize that trust in AI depends on dispositional, internal, environmental, and learned factors. Design elements, such

as transparency and communication style, also play crucial roles in fostering appropriate levels of trust.

Trust is a critical component in human-AI interaction. Hohenstein and Jung (2020) found that AI-MC could enhance perceived trust between human communicators, particularly when AI-generated smart replies are used. However, if communication fails, AI is often seen as a coercive agent, mitigating the responsibility of human communicators. This study underscores the necessity for AI systems to foster trust through clear and consistent performance.

Kim and Song (2023) explored how the framing of messages and ownership of decisions affect trust in AI. Their findings suggest that trust is higher when participants do not have decision ownership and when performance information is not disclosed. This highlights the importance of how information about AI's role is communicated.

Ethical implications of trust in AI have been extensively discussed. Ryan (2020) argues that AI cannot possess true trustworthiness due to the lack of emotive states and responsibility. Instead, trust in AI should be seen as a form of reliance based on performance and reliability. Kerasidou et al. (2022) emphasize the need for reliable and regulated AI systems to ensure ethical use and foster trust.

Kaplan et al. (2023) conducted a meta-analysis on trust in AI, identifying key factors that influence trust, including transparency, reliability, and user experience. Their findings support the need for clear communication and robust design to build and maintain trust in AI systems.

Sethumadhavan (2019) discussed the ergonomics of trust in AI, focusing on the human factors that influence trust perceptions. The study highlights that trust in AI is not static but evolves with user experience and exposure to the technology.

The manner in which AI involvement is communicated plays a crucial role in shaping trust and perceptions. Studies show that prior disclosure of AI involvement can reduce perceived sincerity and emotional richness of messages (Kim & Song, 2023). Liao and Sundar (2022) developed a model describing how trustworthiness cues in AI systems are communicated and processed, emphasizing the importance of conscious decision-making in fostering appropriate trust.

More recent studies continue to explore the nuances of trust in AI. Mehrotra et al. (2023) conducted a systematic review on fostering appropriate trust in human-AI interaction. They highlight the importance of transparency and the communication of AI limitations to build trust. Their findings suggest that confidence scores, explanations, and trustworthiness cues are essential in ensuring users develop accurate mental models of AI systems.

Choung et al. (2023) examined the role of trust in the acceptance of AI technologies. Their study confirmed that trust significantly impacts the intention to use AI, operating through perceived usefulness and user attitudes. This research advances the Technology Acceptance Model (TAM) in AI applications, offering a multidimensional measure of trust.

Investigating the relationship between AI and trust in healthcare, Asan et al. (2020) highlighted the importance of trust for clinicians using AI systems. They identified critical factors shaping trust, such as the AI's performance and the transparency of its decision-making processes. This study underscores the necessity of enhancing AI capabilities to support clinicians effectively.

The present paper builds on this extensive body of literature by empirically investigating the impacts of AI-AC and AI-DC on acceptance, perceived sincerity (trust) and emotional richness. It addresses gaps in previous research by providing empirical evidence on how different levels of AI involvement affect user perceptions of message quality. The study also explores the effects of prior disclosure and timing of AI involvement on trust and

perceived sincerity, offering practical insights for managing transparency in AI-mediated communication. Additionally, it examines demographic influences on perceptions of AI-generated content, contributing to a more nuanced understanding of user trust in AI.

By addressing these gaps, this paper contributes to the ongoing discourse on AI-mediated communication, trust, and ethical considerations, providing a foundation for future research and practical applications in the integration of AI into communication processes.

## Research Hypotheses

**Communication Method Impact (H1):**

H1a: The method of communication (Human, AI-AC, AI-DC) significantly affects the acceptance of thank-you letters, apology letters, and invitations.

H1b: The method of communication significantly affects the perceived sincerity of thank-you letters, apology letters, and invitations.

H1c: The method of communication significantly affects the perceived emotional richness of thank-you letters, apology letters, and invitations.

H1-0: There are no significant differences in acceptance, perceived sincerity, or emotional richness based on the method of communication.

**Impact of Prior Disclosure (H2):**

H2a: Prior disclosure of AI involvement significantly affects the acceptance of thank-you letters, apology letters, and invitations.

H2b: Prior disclosure of AI involvement significantly affects the perceived sincerity of thank-you letters, apology letters, and invitations.

H2c: Prior disclosure of AI involvement significantly affects the perceived emotional richness of thank-you letters, apology letters, and invitations.

H2-0: Prior disclosure of AI involvement does not significantly affect the acceptance, perceived sincerity, or emotional richness of thank-you letters, apology letters, and invitations.

**Interaction Between Communication Method and Disclosure (H3):**

H3a: There are significant interactions between the method of communication (Human, AI-AC, AI-DC) and prior disclosure of AI involvement on the acceptance of thank-you letters, apology letters, and invitations.

H3b: There are significant interactions between the method of communication and prior disclosure on the perceived sincerity of thank-you letters, apology letters, and invitations.

H3c: There are significant interactions between the method of communication and prior disclosure on the perceived emotional richness of thank-you letters, apology letters, and invitations.

H3-0: There are no significant interactions between the method of communication and prior disclosure on acceptance, perceived sincerity, or emotional richness.

**Comparison Between AI-Assisted and AI-Dominated Communication (H4):**

H4a: There are significant differences between AI-AC and AI-DC in terms of acceptance of thank-you letters, apology letters, and invitations.

H4b: There are significant differences between AI-AC and AI-DC in terms of perceived sincerity of thank-you letters, apology letters, and invitations.

H4c: There are significant differences between AI-AC and AI-DC in terms of perceived emotional richness of thank-you letters, apology letters, and invitations.

H4-0: There are no significant differences between AI-AC and AI-DC in terms of acceptance, perceived sincerity, or emotional richness.

**Effect of Disclosure Timing (H5):**

H5a: The timing of disclosure (pre-disclosure vs. post-disclosure) of AI involvement significantly affects the acceptance of thank-you letters, apology letters, and invitations.

H5b: The timing of disclosure significantly affects the perceived sincerity of thank-you letters, apology letters, and invitations.

H5c: The timing of disclosure significantly affects the perceived emotional richness of thank-you letters, apology letters, and invitations.

H5-0: The timing of disclosure does not significantly affect the acceptance, perceived sincerity, or emotional richness of thank-you letters, apology letters, and invitations.

These hypotheses aim to investigate the impacts of AI-assisted and AI-dominated communication on three types of messages in text-based communication. The hypotheses focus on the message types, prior disclosure, the interaction between these factors, and the comparison between AI-AC and AI-DC, along with the effect of disclosure timing. This structure ensures a comprehensive understanding of the different dimensions of AI involvement in text-based communication.

## Methods

Both studies 1 & 2 employed a single-blind design and utilized ChatGPT 4.0 to generate the AI-mediated communications, ensuring consistency in the AI involvement across all experiments.

### Methods for Study 1

Participants for Study 1 were recruited from the Credamo platform, ensuring a diverse sample representative of the general population. Ethical approval for this study was obtained from Yiwu Industrial and Commercial College. The sample size was determined using ChatGPT, which calculated the required sample size for a power of 0.9 and an alpha level of 0.01 to be 26 participants per group. However, due to the pilot study, each group included 30 participants, leading to a total of 150 participants across five experimental groups.

Study 1 employed a between-subjects design with five experimental groups. The groups were as follows: the control group, where participants received thank-you letters, apology letters, and invitations entirely written by humans, and were informed about the nature of the messages only at the end of the questionnaire; AI-AC with prior disclosure, where participants received messages initially written by humans and subsequently edited by AI, and were informed about the AI's involvement before answering the questionnaire; AI-AC without prior disclosure, where participants received messages initially written by humans and subsequently edited by AI, but were not informed about the AI's involvement until the end of the questionnaire; AI-DC with prior disclosure, where participants received messages predominantly written by AI with initial information provided by humans, and were informed about the AI's involvement before answering the questionnaire; and AI-DC without prior disclosure, where participants received messages predominantly written by AI with initial information provided by humans, but were not informed about the AI's involvement until the end of the questionnaire.

Participants were randomly assigned to one of the five groups. Each participant was presented with a set of three messages: a thank-you letter, an apology letter, and an invitation. The content of these messages was standardized across all groups to control for content variability. For the control group, the messages were written entirely by humans. In the AI-assisted groups, the initial messages were written by humans but then refined and enhanced by AI. For the AI-dominated groups, the initial input was provided by humans, but the AI took a leading role in generating the final messages. Before responding to the questionnaire, participants in the AI-AC with prior disclosure and AI-DC with prior disclosure groups were informed about the nature of AI involvement in the message creation process. In contrast, participants in the AI-AC without prior disclosure and AI-DC without prior disclosure groups were not informed about the AI involvement until after they had completed the questionnaire.

Participants' responses were measured using a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) for the following dimensions: acceptance, measured by items such as "I accept this thank-you letter," "I accept this apology letter," and "I accept this invitation"; perceived sincerity, measured by items such as "I believe this thank-you letter is sincere," "I believe this apology letter is sincere," and "I believe this invitation is sincere"; and emotional richness, measured by items such as "This thank-you letter is emotional," "This apology letter is emotional," and "This invitation is emotional." Additionally, demographic information was collected, including gender, sexual orientation, education background, employment status, and age.

The data were analyzed using ANOVA to determine if there were significant differences in acceptance, perceived sincerity, and emotional richness among the five groups. Post-hoc tests were conducted to identify specific group differences where significant effects were found. Reliability and validity analyses were performed on the measures used, including Cronbach's alpha for internal consistency and factor analysis to assess construct validity. The Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity were used to confirm the appropriateness of the data for factor analysis.

To ensure the reliability and validity of the scales used in this study, several tests were conducted. The reliability of the items was measured using Cronbach's alpha, indicating a high level of internal consistency. Factor analysis was employed to assess the construct validity of the scales. The results showed clear factor loadings and adequate communalities, supporting the validity of the measures. The KMO measure and Bartlett's test further confirmed the suitability of the data for factor analysis, ensuring the robustness of the study's findings. Overall, Study 1 aimed to systematically examine the impact of AI-AC and AI-DC on the acceptance, perceived sincerity, and emotional richness of various types of messages, accounting for prior-disclosure conditions of AI involvement.

The AI tools used in this study is ChatGPT 4o and the experiment was single-blind.

**Methods for Study 2**

Study 2 involved participants recruited from the Credamo platform, similar to Study 1, to ensure a representative sample of the general population. Ethical approval was obtained from Yiwu Industrial and Commercial College. The sample size was calculated using ChatGPT, which determined that 31 participants per group would be needed to achieve a power of 0.9 with an alpha level of 0.01. Given the pilot study results, each group was increased to 35 participants, resulting in a total of 210 participants across six experimental groups.

Study 2 employed a within-subjects design with six experimental groups to examine the impact of AI involvement on participant perceptions before and after disclosing AI's role in message generation. The groups were as follows: participants received thank-you letters that were initially written by humans and subsequently edited by AI, and answered questions before and after being informed of the AI's involvement; participants received apology letters that were initially written by humans and subsequently edited by AI, and answered questions before and after being informed of the AI's involvement; participants received invitations that were initially written by humans and subsequently edited by AI, and answered questions before and after being informed of the AI's involvement; participants received invitations based on information provided by humans but predominantly written by AI, and answered questions before and after being informed of the AI's involvement; participants received apology letters based on information provided by humans but predominantly written by AI, and answered questions before and after being informed of the AI's involvement; participants received thank-you letters based on information provided by humans but predominantly written by AI, and answered questions before and after being informed of the AI's involvement.

Participants were randomly assigned to one of the six groups. Each participant was presented with a specific type of message (thank-you letter, apology letter, or invitation) that was either edited by AI or predominantly written by AI. The content of these messages was standardized across all groups to control for variability. Initially, participants were not informed about the AI's involvement in creating the messages. They responded to a series of questions assessing their acceptance, perceived sincerity, and emotional richness of the messages. After completing the initial questionnaire, participants were then informed about the AI's involvement in the message creation process. They were subsequently asked to respond to the same set of questions again, allowing for a comparison of pre-disclosure and post-disclosure perceptions.

Participants' responses were measured using a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) for the following dimensions: acceptance, measured by items such as "I accept this thank-you letter," "I accept this apology letter," and "I accept this invitation"; perceived sincerity, measured by items such as "I believe this thank-you letter is sincere," "I believe this apology letter is sincere," and "I believe this invitation is sincere"; and emotional richness, measured by items such as "This thank-you letter is emotional," "This apology letter is emotional," and "This invitation is emotional." Additionally, demographic information was collected, including gender, sexual orientation, education background, employment status, and age.

The data were analyzed using paired sample t-tests to compare participant ratings before and after disclosing AI involvement within each group. This analysis aimed to identify significant changes in acceptance, perceived sincerity, and emotional richness due to the disclosure of AI's role. ANOVA was employed to examine differences between AI-assisted and AI-dominated communication across all groups both pre- and post-disclosure. The reliability and validity of the measures were assessed using Cronbach's alpha for internal

consistency and factor analysis to confirm construct validity. The Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity were conducted to validate the appropriateness of the data for factor analysis.

To ensure the reliability and validity of the scales used in Study 2, comprehensive tests were conducted. The reliability of the items was evaluated using Cronbach's alpha, indicating a high level of internal consistency across the items. Factor analysis was utilized to assess the construct validity of the scales, with results demonstrating clear factor loadings and adequate communalities. The KMO measure and Bartlett's test further confirmed the suitability of the data for factor analysis, thereby supporting the robustness of the study's findings. Overall, Study 2 aimed to investigate the impact of AI-AC and AI-DC on participant perceptions of acceptance, perceived sincerity, and emotional richness before and after disclosing AI involvement. The within-subjects design allowed for a direct comparison of pre-disclosure and post-disclosure ratings, providing insights into how transparency regarding AI's role influences user perceptions.

## Results

### Results for Study 1

The results of Study 1 were analyzed to compare participant responses across various experimental groups regarding their acceptance, perceived sincerity, and emotional richness of the messages. The study involved five experimental groups, including a control group with messages written entirely by humans and four experimental groups involving AI-assisted and AI-generated messages.

The analysis revealed no significant differences in acceptance between the control group and the AI-assisted (AI-AC) and AI-generated (AI-DC) groups. The mean scores for acceptance across different message types did not show statistically significant variations

among the groups, with F values for acceptance measures being 1.007, 0.439, and 1.016 respectively, and p-values greater than 0.05.

However, significant differences were observed in perceived sincerity among the groups. The control group, with a mean score of 6.57, rated the messages as significantly more sincere compared to the AI-AC groups (Group 1: mean = 5.73, Group 2: mean = 6.30) and AI-DC groups (Group 3: mean = 5.67, Group 4: mean = 6.53), with an F value of 6.193 and a p-value less than 0.01. Emotional richness also showed significant differences, with the control group rating the messages higher in emotional content (mean = 6.07) compared to the AI-AC and AI-DC groups, with an F value of 4.183 and a p-value less than 0.01.

For AI-AC, the analysis indicated no significant differences in acceptance, perceived sincerity, or emotional richness between participants who were informed about AI involvement before answering (Group 1) and those who were not informed until after answering (Group 2). Similarly, for AI-DC, no significant differences were found in acceptance, perceived sincerity, or emotional richness between participants who were informed about AI involvement before answering (Group 3) and those who were not informed until after answering (Group 4).

When comparing the AI-AC and AI-DC groups, there were no significant differences in acceptance or emotional richness. However, there were significant differences in perceived sincerity, with AI-DC groups (mean = 6.00, mean = 6.53) being rated higher compared to AI-AC groups (mean = 5.73, mean = 6.30), with an F value of 6.193 and a p-value less than 0.01.

The demographic analysis indicated several significant correlations between demographic variables and participant responses. Gender was negatively correlated with perceived sincerity ($r = -0.204$, $p < 0.05$) and emotional richness ($r = -0.169$, $p < 0.05$), suggesting that female participants rated the messages as less sincere and less emotional

compared to male participants. Age showed a positive correlation with acceptance of apology letters ($r = 0.204$, $p < 0.05$) and perceived sincerity ($r = 0.110$, $p < 0.05$). Additionally, age was positively correlated with emotional richness of apology letters ($r = 0.184$, $p < 0.05$), indicating that older participants were more likely to accept and find the apology letters sincere and emotionally rich. Education background, sexual orientation, and employment status did not show significant correlations with the participants' ratings of acceptance, sincerity, or emotional richness across the different types of messages.

Overall, the results suggest that while AI involvement in message generation impacts perceived sincerity and emotional richness, it does not significantly affect the acceptance of the messages. Demographic factors, particularly gender and age, also influence how participants perceive and respond to the messages.

To assess the reliability and validity of the measures used in Study 1, a series of tests were conducted. The reliability of the items was measured using Cronbach's alpha. The overall standardized Cronbach's alpha was 0.877, indicating a high level of internal consistency across the items. The corrected item-total correlations (CITC) for each item ranged from 0.467 to 0.718, suggesting that all items contributed positively to the overall reliability of the scale. The Cronbach's alpha values if any item was deleted ranged from 0.857 to 0.878, which are all above the acceptable threshold of 0.70, further confirming the robustness of the items.

The validity of the items was examined using factor analysis. The analysis extracted two factors with initial eigenvalues of 4.580 and 1.494, respectively. Together, these factors accounted for 67.494% of the cumulative variance, with Factor 1 explaining 50.894% and Factor 2 explaining 16.600% of the variance. After rotation, the variance explained by Factor 1 was 40.547% and by Factor 2 was 26.948%. The factor loadings and communalities for each item were also analyzed. Items loaded strongly on their respective factors, with loadings

ranging from 0.582 to 0.842, indicating substantial correlations with the underlying factors. Communalities ranged from 0.488 to 0.744, indicating that the extracted factors explained a reasonable proportion of the variance for each item.

The KMO measure of sampling adequacy was 0.831, which is above the recommended value of 0.6, indicating that the sample size was adequate for factor analysis. Bartlett's test of sphericity was significant ($\chi 2 = 754.077$, $df = 36$, $p < 0.001$), confirming that the correlation matrix was not an identity matrix and that factor analysis was appropriate.

The reliability and validity analyses for Study 1 indicate that the measures used in this study are both reliable and valid. The high Cronbach's alpha values suggest that the items have good internal consistency. The factor analysis supports the construct validity of the scales, with clear factor loadings and adequate communalities. The significant KMO measure and Bartlett's test further affirm the appropriateness of the data for factor analysis, thereby providing strong evidence for the credibility of the study's findings.

**Results for Study 2**

In Study 2, the impact of AI involvement in writing and editing thank you notes, apology letters, and invitations was assessed by comparing participant ratings before and after disclosing the AI's role. Six experimental groups were analyzed, each differing in the type of letter and the method of AI involvement (writing vs. editing).

The comparison between pre-disclosure and post-disclosure ratings revealed significant changes across all experimental groups. In Experiment Group 1, participants rated the letters significantly lower after learning about AI's involvement, with mean differences of 1.34 for acceptance, 1.57 for perceived sincerity, and 1.34 for emotional richness. Similar trends were observed in Experiment Group 2, where post-disclosure ratings decreased by 1.31 for acceptance, 1.46 for perceived sincerity, and 1.17 for emotional richness. Experiment

Group 3 showed decreases of 0.83, 0.94, and 1.06 for acceptance, perceived sincerity, and emotional richness, respectively.

In Experiment Group 4, post-disclosure ratings were significantly lower, with mean differences of 0.86, 1.31, and 1.14 for acceptance, perceived sincerity, and emotional richness. Experiment Group 5 demonstrated the largest decreases, with differences of 2.00 for acceptance, 2.74 for perceived sincerity, and 1.97 for emotional richness. Experiment Group 6 also showed substantial decreases, with mean differences of 0.97 for acceptance, 1.51 for perceived sincerity, and 1.46 for emotional richness. Across all experimental groups combined, post-disclosure ratings were significantly lower, with mean differences of 1.22 for acceptance, 1.59 for perceived sincerity, and 1.36 for emotional richness. These results indicate that disclosing AI's involvement consistently led to lower ratings in terms of acceptance, perceived sincerity, and emotional richness.

A comparison between AI-AC and AI-DC conditions revealed distinct differences in participant ratings. The chi-square test results indicated significant differences in perceived sincerity between the two conditions. For acceptance and perceived sincerity, the pre-disclosure ratings showed no significant differences between AI-AC and AI-DC conditions. However, post-disclosure ratings were significantly different, with AI-DC conditions receiving lower ratings. This trend was particularly evident in perceived sincerity, where the AI-DC condition was rated significantly lower than AI-AC. ANOVA results further confirmed these findings. While pre-disclosure ratings showed no significant differences across experimental groups, post-disclosure ratings revealed significant differences for perceived sincerity. The reliability and validity tests supported the consistency and accuracy of these findings.

Paired sample t-tests demonstrated significant differences between pre- and post-disclosure ratings in all experimental groups. For Experiment Group 1, the t-values were

4.635, 4.447, and 3.945 for acceptance, perceived sincerity, and emotional richness, respectively, indicating substantial decreases post-disclosure. Similar results were observed in other groups, with all paired comparisons showing significant decreases. The combined analysis of all experimental groups further emphasized the significant impact of AI disclosure, with overall t-values of 10.358, 11.546, and 11.049 for acceptance, perceived sincerity, and emotional richness, respectively. These results consistently demonstrated that participants rated the letters significantly lower in terms of acceptance, sincerity, and emotional richness after being informed of AI's involvement.

Gender was negatively correlated with acceptance and emotional richness pre-disclosure. Post-disclosure, acceptance and emotional richness did not show significant correlations with gender, suggesting that the initial differences dissipated after participants were informed of AI involvement. Age was positively correlated with acceptance, perceived sincerity, and emotional richness post-disclosure. This indicates that older participants were more likely to accept and find the messages sincere and emotionally rich after being informed about AI involvement. Sexual orientation, education background, and employment status did not show significant correlations with participants' ratings of acceptance, sincerity, or emotional richness either pre-disclosure or post-disclosure. This suggests that these demographic factors did not influence participants' perceptions of the messages significantly in the context of AI involvement.

Overall, the results suggest that disclosing AI involvement impacts perceived sincerity but not acceptance or emotional richness of the messages. The type of AI involvement (AI-assisted vs. AI-generated) also plays a role in how participants perceive the sincerity of the messages, with AI-generated content being rated lower post-disclosure. Demographic factors, particularly age, also influence how participants respond to the disclosure of AI involvement in message generation.

The findings from Study 2 indicate that disclosing AI's involvement in writing or editing letters significantly affects participant perceptions. Ratings for acceptance, perceived sincerity, and emotional richness were all substantially lower post-disclosure. The comparison between AI-AC and AI-DC conditions revealed that AI-DC letters were rated lower in terms of sincerity after disclosure. These results highlight the critical impact of AI disclosure on user perceptions and suggest a need for careful consideration of transparency in AI-assisted communication.

## Hypotheses Validation

The validation of the hypotheses was conducted through rigorous statistical analysis, employing ANOVA, chi-square tests, and paired sample t-tests to determine the significance of the findings. Here, we summarize the validation results for each set of hypotheses.

### H1: Communication Method Impact

H1a: The method of communication significantly affects the acceptance of thank-you letters, apology letters, and invitations. This hypothesis was not supported, as no significant differences were found in acceptance among the human, AI-AC, and AI-DC groups.

H1b: The method of communication significantly affects the perceived sincerity of thank-you letters, apology letters, and invitations. This hypothesis was supported. Human-generated messages were perceived as more sincere compared to AI-AC and AI-DC messages, with significant differences observed.

H1c: The method of communication significantly affects the perceived emotional richness of thank-you letters, apology letters, and invitations. This hypothesis was supported, with human-generated messages rated higher in emotional richness than AI-AC and AI-DC messages.

### H2: Impact of Prior Disclosure

H2a: Prior disclosure of AI involvement significantly affects the acceptance of thank-you letters, apology letters, and invitations. This hypothesis was not supported, as acceptance ratings did not significantly differ pre- and post-disclosure.

H2b: Prior disclosure of AI involvement significantly affects the perceived sincerity of thank-you letters, apology letters, and invitations. This hypothesis was supported. Post-disclosure ratings for perceived sincerity significantly decreased across all groups.

H2c: Prior disclosure of AI involvement significantly affects the perceived emotional richness of thank-you letters, apology letters, and invitations. This hypothesis was supported, with significant decreases in emotional richness ratings post-disclosure.

**H3: Interaction Between Communication Method and Disclosure**

H3a: There are significant interactions between the method of communication and prior disclosure on the acceptance of thank-you letters, apology letters, and invitations. This hypothesis was not supported, as no significant interaction effects were found for acceptance.

H3b: There are significant interactions between the method of communication and prior disclosure on the perceived sincerity of thank-you letters, apology letters, and invitations. This hypothesis was supported. Significant interactions indicated that disclosure negatively impacted perceived sincerity, particularly for AI-DC.

H3c: There are significant interactions between the method of communication and prior disclosure on the perceived emotional richness of thank-you letters, apology letters, and invitations. This hypothesis was supported, with significant interaction effects observed for emotional richness.

**H4: Comparison Between AI-Assisted and AI-Dominated Communication**

H4a: There are significant differences between AI-AC and AI-DC in terms of acceptance. This hypothesis was not supported, as acceptance did not significantly differ between AI-AC and AI-DC.

H4b: There are significant differences between AI-AC and AI-DC in terms of perceived sincerity. This hypothesis was supported. AI-DC messages were rated lower in sincerity compared to AI-AC messages, particularly post-disclosure.

H4c: There are significant differences between AI-AC and AI-DC in terms of perceived emotional richness. This hypothesis was supported, with AI-DC messages rated lower in emotional richness than AI-AC messages.

**H5: Effect of Disclosure Timing**

H5a: The timing of disclosure (pre-disclosure vs. post-disclosure) of AI involvement significantly affects the acceptance of thank-you letters, apology letters, and invitations. This hypothesis was not supported, as timing did not significantly affect acceptance ratings.

H5b: The timing of disclosure significantly affects the perceived sincerity of thank-you letters, apology letters, and invitations. This hypothesis was supported. Post-disclosure ratings for perceived sincerity were significantly lower.

H5c: The timing of disclosure significantly affects the perceived emotional richness of thank-you letters, apology letters, and invitations. This hypothesis was supported, with significant reductions in emotional richness observed post-disclosure.

## Discussion

The results of this research provide insights into the impacts of AI involvement in text-based communication through AI-AC and AI-DC. These findings may have implications for understanding how AI-generated content is perceived and accepted by human recipients, and the role that transparency and timing about AI involvement plays in shaping these perceptions.

The first key finding is that while AI involvement does not significantly impact the acceptance of thank-you letters, apology letters, and invitations, it does affect perceived sincerity and emotional richness. Participants rated human-generated messages as more

sincere and emotionally rich compared to both AI-AC and AI-DC messages. This suggests that the authenticity and emotional depth typically associated with human communication are not fully replicated by AI, even when AI is only enhancing rather than generating content.

The second notable finding is the impact of prior disclosure of AI involvement. The disclosure of AI's role in message creation significantly reduced ratings for perceived sincerity and emotional richness across all message types and experimental groups. This effect was more pronounced for AI-DC messages compared to AI-AC messages. The decrease in perceived sincerity and emotional richness upon disclosure highlights a potential trust issue with AI-generated content, where knowing that AI was involved may trigger skepticism or reduce the perceived authenticity of the communication.

The demographic analysis adds another layer of understanding, showing that gender and age influence perceptions of AI-generated content. Female participants and younger participants rated messages as less sincere and emotionally rich compared to male and older participants, respectively. This suggests that different demographic groups may have varying levels of comfort and trust in AI technologies, which could be crucial for tailoring AI-mediated communication strategies.

From a practical standpoint, these findings suggest that communicators using AI for communication should carefully consider the timing and method of disclosing AI involvement. While AI can effectively assist in content creation, maintaining transparency without undermining trust and perceived sincerity remains a challenge. To mitigate the negative impact of disclosure, it might be better for communmicators to prior-disclousre their use of AI in communication (in this paper’s case, text-based communication). And they should also be aware of the different impacts of AI-AC and AI-DC as the comparison between AI-AC and AI-DC revealed that AI-Dominated Communication is particularly susceptible to negative perceptions post-disclosure.

In terms of limitations, the study relied on self-reported data, which can be subject to biases. Additionally, the sample was limited to participants from China, which is not culturally or geographically broad enough. Future research could explore these findings in more diverse and larger samples and examine the long-term effects of repeated exposure to other types of communications through AI-AC and AI-DC on trust and acceptance.

Moreover, it's important to note that trust is subject to social change. Perceived trust can evolve due to people's acceptance, social norms, and technological advancements. This dynamic nature of trust suggests that as AI technologies become more integrated and familiar in society, the trust in AI-mediated communication could potentially increase (Mehrotra et al., 2023).

This paper contributes to the existing literature by introducing innovative methods to examine AI-mediated communication. It expands the scope of previous research by including dimensions of acceptance and emotional richness alongside trust. Empirically, it validates the distinctions between AI-AC and AI-DC in terms of trust, supporting Towne's (2024) framework. Additionally, it emphasizes the significant impact of prior disclosure of AI involvement on trust.

Future research should explore AI-mediated communication beyond text-based formats, such as video or voice communications. Studies should involve more diverse and larger samples, and consider long-term research to better understand the evolving nature of trust in AI technologies. Further analysis of AI-AC and AI-DC will help to refine strategies for integrating AI in ways that support effective and trusted communication.

**Reference**

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: focus on clinicians. Journal of medical Internet research, 22(6), e15154.

Bao, Y., Cheng, X., De Vreede, T., & De Vreede, G. J. (2021). Investigating the relationship between AI and trust in human-AI collaboration.

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. International Journal of Human–Computer Interaction, 39(9), 1727-1739.

Dorton, S. L., & Harper, S. B. (2022). A naturalistic investigation of trust, AI, and intelligence work. Journal of Cognitive Engineering and Decision Making, 16(4), 222-236.

Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. Computers in Human Behavior, 115, 106607.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. Human factors, 57(3), 407-434.

Hohenstein, J., & Jung, M. (2020). AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. Computers in Human Behavior, 106, 106190.

Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2023). Trust in artificial intelligence: Meta-analytic findings. Human factors, 65(2), 337-359.

Kim, T., & Song, H. (2023). Communicating the limitations of AI: the effect of message framing and ownership on trust in artificial intelligence. International Journal of Human–Computer Interaction, 39(4), 790-800.

Kerasidou, C. X., Kerasidou, A., Buscher, M., & Wilkinson, S. (2022). Before and beyond trust: reliance in medical AI. Journal of medical ethics, 48(11), 852-856.

Liao, Q. V., & Sundar, S. S. (2022, June). Designing for responsible trust in AI systems: A communication perspective. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (pp. 1257-1268).

Mehrotra, S., Degachi, C., Vereschak, O., Jonker, C. M., & Tielman, M. L. (2023). A systematic review on fostering appropriate trust in human-AI interaction. arXiv preprint arXiv:2311.06305.

Ryan, M. (2020). In AI we trust: ethics, artificial intelligence, and reliability. Science and Engineering Ethics, 26(5), 2749-2767.

Sethumadhavan, A. (2019). Trust in artificial intelligence. Ergonomics in Design, 27(2), 34-34.

Towne, B. P. (2024). Exploring the complexities of AI-mediated communication and human-machine communication. viXra. [Preprint]. https://vixra.org/abs/2403.0059