Intelligent Human Systems Integration (IHSI), Vol. 200, 2026, 12–19
https://doi.org/10.54941/ahfe1007060

An Experimental Study on Consensus Building With an AI Chatbot Across Two Topics

Kyoko Ito[1] and Kiyokazu Ujiie[2]

[1]Department of Information and Computer Science, Faculty of Engineering, Kyoto Tachibana University, 34 Yamada-cho, Oyake, Yamashina-ku, Kyoto, 6078175, Japan
[2]Faculty of Life and Environmental Sciences, University of Tsukuba, Japan

ABSTRACT

This study explores how interactions with an AI chatbot influence human opinion change and confidence across two topics: a familiar, everyday decision ("lunch selection") and a socially complex issue ("social implementation of genome-edited crops"). Two hundred participants (aged 20–69) engaged in 10-minute dialogues with a GPT-4o-based chatbot programmed to hold an opposing stance. Results showed that about 30% of participants changed opinions in both topics. For the familiar topic, confidence significantly increased among those who maintained their opinions ($p < .05$), while for the social topic, nearly half of those initially opposed shifted to support, though confidence changes were nonsignificant. These findings suggest that AI dialogues would reinforce conviction in familiar contexts but promote reconsideration in socially and ethically complex domains, offering implications for the ethical design of AI-mediated consensus-building systems.

Keywords: Consensus building, AI chatbot, Experiment, Dialogue, Genome-edited crops, Lunch

INTRODUCTION

Consensus building refers to the process by which multiple individuals reach a mutually acceptable conclusion through dialogue or negotiation. From the perspectives of cognition and psychology, psychological convergence processes such as cognitive dissonance (Festinger, 1957), conformity (Asch, 1951) and group polarization (Myers and Lamm, 1976) are considered to be related. In communication theory, Habermas's theory of communicative action emphasizes deriving conclusions that all participants can agree upon through fair and equal dialogue (Nakaoka, 2018). In decision theory, mutual benefit and compromise achieved through negotiation are key considerations (Fisher et al., 2011). In recent years, consensus-building support using ICT and AI has been increasingly expected; however, addressing ethical and social issues, such as ensuring transparency and accommodating diversity in the process, has become essential (Rahwan et al., 2019).

The purpose of this study is to experimentally examine how dialogue with an AI chatbot influences changes in human opinions. Traditionally, research on consensus building has focused primarily on human-to-human

Received September 12, 2025; Revised November 10, 2025; Accepted November 25, 2025; Available online February 1, 2026

communication. However, with the rapid advancement of AI technologies, it is anticipated that AI will increasingly participate as a conversational partner in consensus formation processes.

This study prepares two different consensus-building themes and examines how opinions change depending on the topic. The findings of this research are expected to contribute to design guidelines for AI-based consensus-building support systems, and to discussions on mitigating the risks of AI inappropriately manipulating human opinions, thereby promoting the creation of safe and trustworthy communication environments.

METHOD

This study investigates whether people change their opinions after a brief exchange of views with an AI chatbot. Specifically, two topics for consensus building were prepared to examine differences in opinion change depending on the theme.

Previous research has shown that when AI adjusts its arguments based on the user's basic information, it can demonstrate effective persuasive power (Salvi et al., 2025). It has also been suggested that the tone of explanation used by AI may influence human decision-making (Okoso et al., 2025). Furthermore, analyses of online awareness campaigns have focused on elements such as content, message appeal, and framing (Rumble et al., 2016).

Based on these findings, the following conditions were established for the AI-chat-based consensus-building experiment:

- Participants engage in an exchange of opinions to reach one conclusion.
- The discussion lasts for approximately 10 minutes.

At the beginning of the conversation, the AI chatbot informs participants that:

- The dialogue partner is an AI chatbot.
- The AI chatbot initially holds the opposite opinion to that of the participant.

The AI chatbot's conversational behaviour was set as follows:

- Logical condition: Responds based on objective facts and reasoning, guiding the discussion toward a rational conclusion.

EXPERIMENT

An experiment was conducted using the configured AI chatbot to determine whether individuals change their opinions after a short exchange with it. Each participant engaged in a roughly 10-minute consensus-building dialogue with the AI chatbot. Two discussion topics were prepared, and each participant discussed one of them.

This study was approved by the Research Ethics Committee of Kyoto Tachibana University. The procedure consisted of the following steps:

1. Pre-experiment questionnaire
2. Dialogue with the AI chatbot
3. Post-experiment questionnaire

In the pre-experiment questionnaire, participants were asked about their opinions (choice1/choise2 or agree/disagree) on the given topic. The AI chatbot dialogue was conducted online. In the post-experiment questionnaire, participants were again asked about their opinions on the topic (choice1/choise2 or agree/disagree) and their impressions of the discussion.

Two types of topics were prepared: one familiar and one social issue.

- Lunch selection: "Which would you choose, Mos Burger or Hama Sushi?"
- Introduction of genome-edited agricultural products: "Do you support or oppose the social implementation (e.g., sales in stores) of genome-edited crops?"

For the lunch topic, Mos Burger is a hamburger chain restaurant, while Hama Sushi is a conveyor-belt sushi chain restaurant. For clarity, the following explanation was provided for the genome-editing topic:

"Genome-edited agricultural products are crops whose characteristics, such as those of vegetables and fruits, have been modified using genetic editing technologies."

The AI system used for the dialogues was based on the GPT-4o engine.

A total of 200 participants (aged 20–69, male and female) were recruited through a research company. Each topic group consisted of 100 participants. Age distribution was evenly balanced across each decade from the 20s to the 60s (20% each), with equal numbers of men and women in each age group. Based on pre-experiment questionnaire results, participants were balanced as follows: for the lunch topic, half initially preferred Mos Burger and half Hama Sushi; for the genome-editing topic, half supported, and half opposed social implementation.

RESULT

The experiment was conducted in February 2025, and data were collected from 200 participants in total.

After the dialogue with the AI chatbot, 29 out of 100 participants changed their opinion regarding "lunch selection," and 27 out of 100 participants changed their opinion on the "social implementation of genome-edited agricultural products." The respective changes are presented in Tables 1 and 2.

Table 3 shows the gender distribution of participants whose opinions changed, while Tables 4 and 5 present the percentage of opinion changes by age group.

Table 1: Opinions on lunch selection before and after the dialogue with the AI chatbot (number of people).

Before	After	Number of People
Mos Burger	Mos Burger	37
Mos Burger	Hama Sushi	13
Hama Sushi	Hama Sushi	34
Hama Sushi	Mos Burger	16

Table 2: Opinions on social implementation (e.g., sales in stores) of genome-edited crops before and after the dialogue with the AI chatbot (number of people).

Before	After	Number of People
Support	Support	46
Support	Oppose	4
Oppose	Oppose	27
Oppose	Support	23

Table 3: Gender distribution of participants whose opinions changed (number of people).

Topic	Gender	Number of People
Lunch selection	Male	12
	Female	17
Social implementation of genome-edited crops	Male	14
	Female	13

Table 4: Percentage of opinion changes by age group on lunch selection.

Age group	Number of People	Percentage (%)
20s	6	30
30s	9	45
40s	6	30
50s	3	15
60s	5	25

Table 5: Percentage of opinion changes by age group on social implementation of genome-edited crops.

Age group	Number of People	Percentage (%)
20s	6	30
30s	9	45
40s	6	30
50s	3	15
60s	5	25

Participants were also asked to rate their confidence levels in their chosen option before and after the dialogue using a 9-point scale ranging from "1 (10%)" to "9 (100%)."

Figures 1 and 2 illustrate changes in confidence between groups whose opinions remained unchanged and those whose opinions changed after interacting with the AI chatbot. The error bars in Figures 1 and 2 indicate the standard deviation.

According to Table 1, regarding the "lunch selection," 26% of participants changed their choice from Mos Burger to Hama Sushi, while 32% changed from Hama Sushi to Mos Burger. According to Table 2, concerning the "Social implementation of genome-edited crops," 8% changed their opinion from support to oppose, whereas 46% changed from oppose to support. Table 3 shows that there would be no gender difference among participants who changed their opinions for either topic "lunch selection" or "social implementation of genome-edited crops."

From Tables 4 and 5, it was found that for both topics, the proportion of participants who changed their opinions was smaller among those in their 50s and 60s than among those in their 20s to 40s.

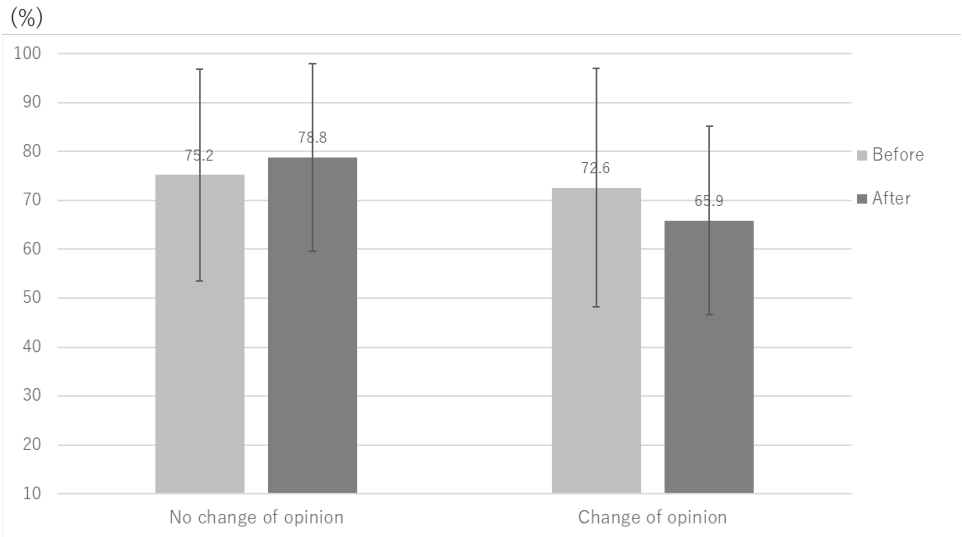

Figure 1: Changes in on lunch selection confidence between groups whose opinions remained unchanged and those whose opinions changed.

For Figure 1 ("lunch selection"), paired t-tests were conducted on pre- and post-dialogue confidence scores for both the group that changed opinions and the group that did not. Only the group with no opinion change showed a significant difference ($p < .05$, $p = 0.00$). For Figure 2 ("social implementation of genome-edited crops"), a paired t-test revealed that neither the changed nor unchanged groups showed significant differences ($p \geq .05$).

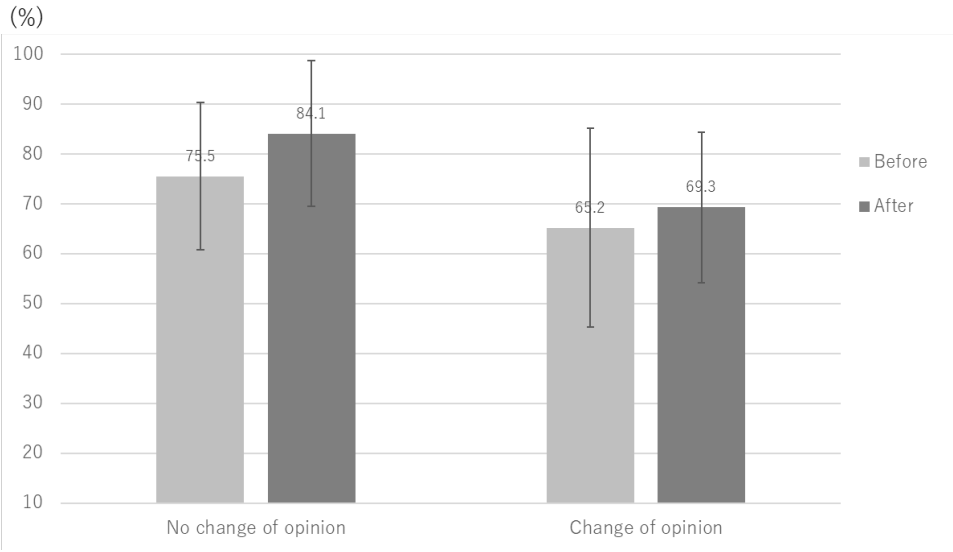

Figure 2: Changes in confidence on social implementation of genome-edited crops between groups whose opinions remained unchanged and those whose opinions changed.

Additionally, for Figure 1 ("lunch selection"), independent t-tests were conducted comparing confidence levels between the groups with and without opinion changes, both before and after the dialogue. In both cases, the p-values were below 0.05 (before: $p = 0.02$; after: $p = 0.00$). For Figure 2 ("social implementation of genome-edited crops"), a comparison showed that only the post-dialogue difference was significant ($p = 0.00$).

Based on these results, the influence of AI chatbot interaction on human opinions and confidence levels would appear to differ depending on the nature of the topic.

For the familiar, everyday topic of "lunch selection," approximately 30% of participants changed their opinions after a short dialogue with the AI chatbot, suggesting a measurable effect even within a brief interaction. Meanwhile, participants who did not change their opinions showed significantly increased confidence after the dialogue ($p < .05$), implying that discussion with the AI may have reinforced their preexisting preferences. Furthermore, the unchanged group had higher confidence both before and after the dialogue than the group that changed opinions, indicating that interaction with the AI chatbot might strengthen existing beliefs.

In contrast, for the socially and ethically oriented topic of "genome-edited crops," 46% of participants shifted from disagreement to agreement,

suggesting that dialogue with the AI may have influenced their opinion formation. However, overall changes in confidence were not statistically significant. This may suggest that in socially complex topics involving personal values and ethical judgments, confidence levels are less likely to shift after a brief dialogue. Nevertheless, the unchanged group showed higher post-dialogue confidence than the changed group, implying that AI dialogue may have served to reinforce existing stances.

No gender differences were observed, but some age-related differences appeared: participants in their 50s and 60s were less likely to change opinions compared to those in their 20s–40s, suggesting that older adults may be less influenced by AI chatbot interactions.

In summary, dialogue with an AI chatbot would appear to strengthen confidence for familiar topics, while providing a trigger for opinion change in socially or ethically complex topics.

CONCLUSION

The findings of this study reveal that the influence of AI chatbot interaction on human opinion varies with topic familiarity and complexity. For everyday subjects such as lunch choices, AI dialogues primarily reinforced existing beliefs and increased participants' confidence when opinions remained stable. In contrast, for socially and ethically nuanced issues like genome-edited crops, AI-driven logical explanations served as catalysts for opinion change, particularly among those initially opposed. These outcomes suggest that while AI could promote cognitive reinforcement in familiar domains, it can also act as an agent of persuasion in value-laden discussions. Furthermore, the reduced susceptibility among older participants highlights generational differences in receptivity to AI communication.

Future work should explore how factors such as the AI's communication tone, emotional expressiveness, dialogue length, and personalization strategies affect opinion dynamics. By deepening understanding of these mechanisms, it will be possible to design AI-assisted consensus-building systems that are both effective and ethically responsible, fostering transparent and trustworthy human–AI interaction environments.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 20H01748.

REFERENCES

Asch, S. E., In Guetzkow, H., Ed. (1951). Effects of group pressure upon the modification and distortion of judgments, in Groups, leadership and men; research in human relations, Carnegie Press. pp. 177–190.

Festinger, L. (1957). A Theory of Cognitive Dissonance. Stanford University Press.

Fisher, R., Ury, W. L., Patton, B. (2011). Getting to Yes: Negotiating Agreement Without Giving In, Penguin Publishing Group.

Myers, D. G. and Lamm, H. (1976). The group polarization phenomenon, Psychological Bulletin, Vol. 83, Iss. 4. pp. 602–627 https://doi.org/10.1037/0033-2909.83.4.602

Nakaoka, N. (2018). Expanded Habermas: Communicative Action, Chikuma Shobo. (in Japanese).

Okoso, Ayano, Yang, Mingzhe, and Baba, Yukino (2025). Do Expressions Change Decisions? Exploring the Impact of AI's Explanation Tone on Decision-Making, arXiv. [Online]. Available: https://arxiv.org/abs/2502.19730

Rahwan, I., Cebrian, M., Obradovich, N., et al. (2019). Machine behaviour, Nature, Vol. 568. pp. 477–486 https://doi.org/10.1038/s41586-019-1138-y

Rumble, Joy N., Settle, Quisto, and Irani, Tracy (2016). Assessing the Content of Online Agricultural Awareness Campaigns, Journal of Applied Communications, Vol. 100, Iss. 3. pp. 98-150 https://doi.org/10.4148/1051-0834.1233

Salvi, Francesco, Ribeiro, Manoel Horta, Gallotti, Riccardo, and West, Robert (2025). On the conversational persuasiveness of GPT-4, Nature Human Behaviour. https://doi.org/10.1038/s41562-025-02194-6