Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

AI-powered Chatbots: Effective Communication Styles in Sustainability Conversations

Ennio Bilancini
IMT School for Advanced Studies Lucca
Leonardo Boncinelli
University of Florence
Eugenio Vicario
eugenio.vicario@imtlucca.it

IMT School for Advanced Studies Lucca

Article

Keywords: motivational interviewing, directing style, artificial intelligence, SDGs, online experiments

Posted Date: September 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10340878/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

AI-powered Chatbots: Effective Communication Styles in Sustainability Conversations

Ennio Bilancini[1,a], Leonardo Boncinelli[2,b], and Eugenio Vicario*[1,c]

[1]IMT School for Advanced Studies Lucca, Laboratory for the Analysis of compleX Economic Systems, Piazza S. Francesco 19, Lucca, 55100, Italy

[2]Department of Economics and Business, University of Florence, Via delle Pandette 9, 50127 Firenze, Italy

[a]ennio.bilancini@imtlucca.it

[b]leonardo.boncinelli@unifi.it

[c]eugenio.vicario@imtlucca.it

Abstract

This paper presents an analysis of two pre-registered experimental studies examining the impact of 'Motivational Interviewing' and 'Directing Style' in AI-powered chatbot conversations about sustainability. In our experiments, conversations were structured around the Sustainable Development Goals (SDGs), which provided a common framework for discussing sustainability-related issues. To evaluate the effectiveness of these communication styles in enhancing awareness and motivating action toward sustainability-related issues, we measured the engagement levels of participants, along with their self-reported interest and learning outcomes. Our results indicate that 'Motivational Interviewing' is more effective than 'Directing Style' for engagement and interest, while no appreciable difference is found on learning.

Keywords: motivational interviewing; directing style; artificial intelligence; SDGs; online experiments

*Corresponding Author

1 Introduction

Understanding effective communication strategies is crucial for promoting awareness, engagement, and behavioral change on sustainability-related issues. This paper reports two experimental studies focusing on 'Motivational Interviewing' (MI) and 'Directing Style' (DS) as communication techniques in digital conversations between human subjects and AI-powered chatbots. To provide a structured yet broad conversational domain, chatbot interactions were centered on the Sustainable Development Goals (SDGs), which offer a comprehensive framework for discussing sustainability issues. The studies, preregistered on OSF (Link 1 and Link 2), investigate how MI and DS influence participants' engagement and behavioral responses toward sustainability topics. The outcome variables are *Self-assessment of interest* and *Self-assessment of learning*, both measured through a final survey, as well as *Engagement*, which is measured by the number of words written by human subjects. The collected data show a positive effect of MI with respect to DS on engagement and interest, while no effect is detected on learning.

The SDGs are a global initiative, adopted by all the Member States of the United Nations in September 2015, to work collaboratively towards a more just, equitable, and sustainable future. The SDGs are based on a holistic vision of development that recognizes the complex connections between social, economic, and environmental dimensions. They constitute a comprehensive agenda incorporating 17 interconnected goals, each designed to address specific aspects of sustainable development by 2030.

The SDGs are not only relevant for policy, but they are increasingly becoming a focal point of research across a variety of disciplines (Schmidt-Traub et al., 2017; Sachs et al., 2019; De Neve and Sachs, 2020). This increasing interest is reflected in a growing body of literature that explores the multifaceted impacts of the SDGs on society, the economy, and the environment. In a Scopus search, focusing on the term 'SDGs' within the titles, abstracts, and keywords of research articles, we identified 2.162 instances in 2020, 3.046 in 2021, 3.771 in 2022, 4.778 in 2023, 6.441 in 2024, and 9.057 in 2025.

Effective communication is crucial in triggering behavioral changes, as people are highly responsive to the language and framing used in conveying information and actions (Capraro et al., 2024b,a; Hillenbrand and Verrina, 2022). MI represents a prominent tool in the field of behavioral change and decision-making. This client-centered counseling style is rooted in the principles of empathy and collaborative conversation and aims to induce behavior change by helping individuals explore and resolve ambivalence (Miller and Rollnick, 1991). The core skills of MI outlined by Miller and Rollnick (2012) – open-ended questions, affirmation, reflective listening, and summary – are instrumental in fostering an environment of trust and openness. These skills enable the interviewer to facilitate introspection and self-motivation in the individual, which are crucial for any behavioral change. MI has been widely applied for healthcare issues, such as vaccine hesitancy (Breckenridge et al., 2022). More recently,

MI has been used in other fields, in particular for raising awareness on sustainability issues (Tagkaloglou and Kasser, 2018). This is partly due to the growing recognition that environmental challenges are not only technical or scientific problems but also involve human behavior and decision-making. The emphasis of MI on understanding and resolving ambivalence makes it a powerful tool for addressing societal challenges (Klonek et al., 2015). Studies have shown that MI can effectively influence behaviors related to energy conservation (Endrejat et al., 2017), waste reduction (Herzing et al., 2023), and sustainability behaviors more generally (Conrady et al., 2014; Klonek and Kauffeld, 2012).

The concept of DS as a communication strategy is not as clearly defined in the academic literature as a specific approach or methodology, at least not with the same level of theoretical clarity and cohesion as well-established approaches like MI. However, the term and related concepts are often discussed in relation to leadership styles (Martin et al., 2013; Lorinkova et al., 2013), counseling and therapy techniques (Thorne, 1948; Pan et al., 2019), especially when contrasting more directive approaches with non-directive or client-centered ones (Cuijpers et al., 2024; Rogers, 2012). In general, DS refers to an approach in which the therapist, counselor, or leader takes a more active role, providing clear instructions, feedback, and guidance. This style can be particularly useful in contexts that require quick decision-making, in crisis situations, or when working with individuals who may benefit from more structured guidance. Our focus on DS is motivated by the strategies adopted by several countries during the COVID-19 pandemic. Although the urgent need for widespread public action during the COVID-19 pandemic led some countries to adopt directive communication strategies, advocating specific behaviors without extensive collaborative dialogue, the effectiveness of this approach for achieving the SDGs is debatable.

Several factors limit the widespread use of codified communication styles, such as MI and DS. These factors include the need to instruct and train operators (Miller and Rollnick, 2009), as well as all implementation costs, both in terms of time and money, related to in-person communication sessions. To overcome these limitations, researchers are increasingly turning to and testing new technological tools (Altay et al., 2022; Zhang and Song, 2024). Two strands of literature show encouraging results for replacing or supplementing human operators with AI-powered virtual agents. First, experiments in social sciences have been replicated by replacing human participants with Large Language Models (LLMs) (Dillion et al., 2023). These studies demonstrate the ability of AI-powered agents to mimic human cognitive biases (Binz and Schulz, 2023) and behavior in various contexts, including economic games and interactions (Horton, 2023; Maggioni and Rossignoli, 2023; Aher et al., 2022; Chen et al., 2023; Del Rio-Chanona et al., 2025; Gusella and Vicario, 2025), social dilemmas (Guo, 2023; Capraro et al., 2023; Mei et al., 2024), mechanism design (Koster et al., 2022), as well as in voting decisions (Argyle et al., 2023), the formation of moral judgments (Dillion et al., 2023), and theory of mind abilities such as understanding false beliefs and recog-

nizing irony (Strachan et al., 2024). Recent contributions, however, have emphasized that the ability of LLMs to reproduce human responses is not universal and depends on the specific experimental setting and behavioral task under consideration (Vicario et al., 2026; Zewail et al., 2026). The second strand of the literature focuses on human-agent interactions in conversational settings (Stein et al., 2017; Numata et al., 2020; Altay et al., 2023; Cai et al., 2024), with particular attention to the effects of the chatbot's personality and communication style (Ruane et al., 2021; Rhim et al., 2022; Brummernhenrich et al., 2025). We highlight that certain characteristics of chatbots, such as their level of anthropomorphism, which is partly shaped by their communication style, can have mixed effects on how AI-driven interactions and environments are perceived (Schaaff and Heidelmann, 2024). On one hand, a more human-like and engaging interface can boost users' motivation and involvement, as people often react more positively to technologies that exhibit relatable, human-like traits (Konya-Baumbach et al., 2023). On the other hand, overly anthropomorphic design may create unrealistic expectations about the capabilities of large language models, increasing the risk of misunderstanding or user frustration (Crolic et al., 2022). In this body of literature, the impact of chatbot communication styles on learners' outcomes is highly relevant. When conversational agents adopt an empathic style, learners tend to perform better and perceive the learning environment as more supportive (Ortega-Ochoa et al., 2024a,b). Certain features of communication style, such as the use of humor and visual cues like emojis, can enhance learners' motivation and engagement (Smutny and Schreiberova, 2020; Ceha et al., 2021), reduce anxiety, and improve trust (Hew and Huang, 2023; Yu and Zhao, 2024). When chatbots provide metacognitive feedback or supportive phrases, which partially overlap with certain features of the MI communication style, they can trigger positive emotions (Yin et al., 2024) and enhance motivation and overall engagement (Wambsganss et al., 2021). In the domain of chatbots used for commercial purposes, findings consistently show that a socially oriented communication style enhances customer satisfaction and interaction satisfaction, increases consumers' behavioral intentions (Xu et al., 2022; Cai et al., 2024), and, depending on the service type, boosts engagement and rapport through social cues and mimicry (Van Pinxteren et al., 2023). Further studies examine the effectiveness of chatbots using motivational interviewing techniques (Galvão Gomes da Silva et al., 2018) in healthcare (Shingleton and Palfai, 2016; Hamzehgardeshi et al., 2024), particularly for smoking cessation (He et al., 2022; Brown et al., 2023), weight loss (Stephens et al., 2019), substance abuse (Prochaska et al., 2021), and lifestyle changes (Gardiner et al., 2017; Bickmore et al., 2013). Our study builds on this body of literature by analyzing the effects of a motivational interviewing style delivered by an AI-powered chatbot, with a distinctive focus on sustainability as the topic of conversation.

The results of this work contribute to enriching the policy-maker's toolbox, providing an additional intervention tool that exploits recent advancements in the field of artificial intelligence. AI-powered chatbots can combine with existing policy inter-

ventions, activating synergies to promote behavioral change (Alt et al., 2024), thus expanding the literature that examines the role of artificial intelligence in sustainability (Nishant et al., 2020; Kar et al., 2022) and the possible impacts it can have, both positive and negative (Vinuesa et al., 2020).

The structure of the paper is as follows. Section 2 describes the experimental conditions, the final survey, the two studies, and their descriptive statistics. Section 3 presents our main results on *Self-assessment of learning*, *Self-assessment of interest*, and *Engagement*. Section 4 covers the exploratory analysis, while Section 5 concludes by summarizing this contribution and outlining directions for future research.

2 Methodology

We run two studies involving participants in conversations about SDGs, with two experimental conditions: 'Motivational Interviewing' and 'Directing Style'. Participants receive a link via the Prolific platform to access the chatbot and begin the conversation. In the initial phase, they are provided with information about the study, asked for informed consent, and prompted to enter their Prolific ID to anonymously link participants to responses. The conversation then begins with a question from the chatbot. Each user input is followed by a chatbot reply. Participants can stop the conversation at any time and be redirected to the final survey, also administered in chat format. The experiment ends upon completion of the final survey.

We measure interest and learning by means of a single-item retrospective question at the end of the conversation. This choice is justified by two features that are typical of online conversations with a chatbot. First, such conversations can focus on many different aspects of SDGs which can hardly be foreseen and listed, hence a test on specific knowledge to be administered with a pre-post method is likely to be inadequate. Second, participants are likely to have limited attention when answering online questions and this may lead longer tests with multiple items to be noisy and unreliable. We opted for a Likert scale with 6 options in order to avoid the neutral answer, which may be particularly attractive for online tests with limited attention. Notably, the test questions did not include predefined answer options, in order to avoid the default effect.

As concerns engagement, we adopt a more behavioral measure by focusing on the overall number of words written by the user in the conversation. Alternative measures of engagement (time spent in the experiment, number of rounds and average number of words per round) are explored in Section 4.

Conversations are managed through a chatbot developed in the 'Landbot' platform (`https://landbot.io/`), which is accessible through a web url. We integrated the chatbot with AI language model. In particular, we manipulate the communication style of the chatbot through the utilization of different prompts on GPT-3.5-turbo (see the subsection on experimental conditions). The prompt represents the instructions

given to GPT-3.5-turbo. Essentially, in each iteration, we provide GPT-3.5-turbo with the prompt containing instructions on how to behave, followed by the previous conversation, distinguishing between the responses of the chatbot and those of the user. The technological landscape of large language models is evolving rapidly, and the specific model adopted in this study should be interpreted in light of the state of LLM technology at the time. GPT-3.5-turbo represented one of the first LLMs to make conversational AI broadly accessible through commercial APIs, thereby enabling large-scale experimental applications such as ours. Despite its comparatively limited capabilities relative to more recent models, it already allowed us to reliably manipulate the chatbot's communication style through prompt engineering. It is therefore reasonable to expect that this approach can be implemented at least as effectively, and likely even more effectively, with newer generations of language models, which generally exhibit improved instruction-following and greater controllability.

To check that the chatbots were running smoothly and data correctly recorded, we run a pilot study with 90 participants on Profilic, on June 26th, 2023, starting at 12:30 AM. Participants in the pilot were excluded from First Study and Second Study, as well as participants in First Study were excluded from Second Study.

2.1 Experimental Conditions

The two experimental conditions that we compare in our studies are Motivational Interviewing and Directing Style. The main feature of MI and DS are summarized in Table 1.

Table 1: Point-by-point comparison between Directing Style and Motivational Interviewing.

Aspect	**Motivational Interviewing**	**Directing Style**
Focus	On the client and their internal motivation for change.	On the therapist as a guide and source of solutions.
Approach	Collaborative, exploratory, and non-judgmental.	More assertive, direct, and potentially prescriptive.
Goal	To facilitate self-exploration and strengthen intrinsic motivation for change.	To provide direction, instructions, or specific solutions.
Methodology	Based on active listening, reflection, and exploring ambivalence.	May include setting goals, structuring treatment, and defining action steps.
Context of Use	Particularly effective in contexts of addictions and risky health behaviors.	Useful in situations requiring quick decisions or when the client benefits from clear guidance.

Motivational Interviewing: This condition involves conversations where the

interviewer adopts a guiding and empathetic style, aiming to evoke participants' intrinsic motivation towards SDGs. The prompt is: *'Your role is to have a conversation about the argument of sustainable development goals: you should strictly adopt a motivational interviewing style of communication, you should help the user to reflect upon the issue of sustainable development goals, you should not ask more than one question, if you do not understand the meaning or the logic of user_text you should ask the user to rephrase, keep the conversation focused on the SDGs, when you say goodbye to the user, remind the user to click on the menu at the top right to go to the final questions. What would you like to talk about regarding the SDGs?'*

Directing Style: In this condition, the interviewer adopts a more authoritative and directive approach, providing clear guidance and information about SDGs. The prompt is: *'Your role is to have a conversation about the argument of sustainable development goals: you should strictly adopt a directing interviewing style of communication, you should convince the user about the importance of sustainable development goals, you should not ask more than one question, if you do not understand the meaning or the logic of user_text you should ask the user to rephrase, keep the conversation focused on the SDGs, when you say goodbye to the user, remind the user to click on the menu at the top right to go to the final questions. What would you like to be informed about regarding the SDGs?'*

We stress that the differences between the two prompts are quite limited. A first difference regards the communication style: 'a motivational interviewing style' vs. 'a directing interviewing style'. The second difference is about the aim of the chatbot: 'you should help the user to reflect upon the issue' vs. 'you should convince the user about the importance'. Finally, we have a different closing: 'What would you like to talk about regarding the SDGs?' vs. 'What would you like to be informed about regarding the SDGs?'.

2.2 Final survey

At the end of the conversation, the same final survey is administered to all experimental subjects to measure cognitive and behavioral responses uniformly. The variables measured in the final survey are:

- *Self-assessment of interest*: 'Do you feel more interested in sustainability topics after this chat?' (on a scale of 0-5, 0 being 'not at all', 5 being 'quite a lot');
- *Self-assessment of learning*: 'How much have you learned about sustainability topics from this chat?' (on a scale of 0-5, 0 being 'nothing', 5 being 'very much');
- *Willingness to receive costly information*: 'Would you authorize us to send you one or more communications about sustainability topics using the Prolific messaging system? The authorization is optional and at your discretion.' (possible answers: 'yes' and 'skip');

- *Self-assessment of satisfaction*: 'How do you rate our conversation?' (on a smiley rating scale with five options).

The questions are asked in the same order as they are listed above. It is necessary to answer one question before moving on to the next one. Additionally, it is not possible to go back and change previously provided answers.

2.3 First study

In the first study we run the experiment with a target sample size of 800 participants, equally split between the two experimental conditions. In fact, each participant has 50% probability to be assigned to each of the two experimental conditions. This randomization procedure is implemented through the A/B Test feature of the Landbot platform. In accordance with the preregistration we exclude participants that do not complete the final survey, obtaining a sample size of 788 participants, 408 in MI experimental condition and 380 in DS experimental condition. The hypotheses that we have indicated in the preregistration are non-directional differences between MI condition and DS condition in:

1. *Self-assessment of interest*;
2. *Self-assessment of learning*;
3. *Willingness to receive costly information*;
4. *Self-assessment of satisfaction.*

We refer the reader to the Supplementary Information (SI) for further details on the preregistered objectives and methods (OSF Preregistration), as well as the specific results of this study.

2.4 Second study

In the second study, we run the experiment with a target sample size of 800 participants, equally split between the two experimental conditions. In fact, each participant has 50% probability to be assigned to each of the two experimental conditions. This randomization procedure is implemented through the A/B Test feature of the Landbot platform. In accordance with the preregistration we exclude participants that do not complete the final survey, obtaining a sample size of 800 participants, 398 in MI experimental condition and 402 in DS experimental condition. The hypotheses that we have indicated in the preregistration are directional differences between MI condition and DS condition in:

1. *Self-assessment of interest* (Alternative hypothesis: MI greater than DS);
2. *Self-assessment of learning* (Alternative hypothesis: DS greater than MI);
3. *Engagement* (Alternative hypothesis: MI greater than DS).

Engagement is measured by the number of words written by the participants, as recorded in the saved conversations. We refer the reader to the Supplementary Information (SI) for further details on the preregistered objectives and methods (OSF Preregistration), as well as the specific results of this study.

2.5 Descriptive statistics

In this section, we present some descriptive statistics on the aggregated sample of the two studies. The objective is to show how randomization succeeded in generating balanced groups for the two treatments. Specifically, the sample was divided into two groups, the first for MI consisting of 806 subjects and the second, for DS, of 782 subjects.

From the experimental subjects, we are able to observe some socio-economic characteristics such as gender, age, highest education level completed, and household income. Through the Prolific platform, we selected only individuals who had responded to the question 'Generally speaking, how concerned are you about environmental issues?', with possible responses ranging from '1 (Not at all concerned)' to '5 (Very concerned)'.

Table 2: Contingency table for Gender and Treatment.

	Treatment		
Gender	MI	DS	Total
Man	373	374	747
Woman	420	396	816
Non-binary	9	4	13
Total	802	774	1576

Regarding *Gender*, we assess the balance of the samples through the contingency table, as shown in Table 2. We would like to point out that the total number of participants reported earlier does not match the total numbers presented in Table 2, as we lack information on some of the participants. The p-values of the Fisher exact test and the Chi-square test are respectively 0.362 and 0.344, indicating that *Gender* is well balanced across the two treatments.

The other socioeconomic characteristics observed, besides gender, are age, highest education level completed, and household income. *Highest education level completed*, hereafter referred to as *Education*, is divided into seven categories: No formal qualifications, Secondary education (e.g., GED/GCSE), High school diploma/A-levels, Technical/community college, Undergraduate degree (BA/BSc/other), Graduate degree (MA/MSc/MPhil/other), and Doctorate degree (PhD/other). The distribution of *Education* in the two treatments is represented in Figure 1. To test the proper balance between the two treatments, a Wilcoxon rank-sum test is conducted, exploiting

Table 3: Ranksum test of socioeconomic characteristics by treatment: *Age*, *Highest education level completed*, *Household income*, and *Concern about environmental issues*.

Variable	z	Prob > $\|z\|$
Age	0.754	0.4511
Education	-1.292	0.1962
Income	-0.127	0.8992
Concern env.	0.324	0.7462

the ordinal nature of *Education*. The result, as highlighted in Table 3, suggests that the two samples are balanced.

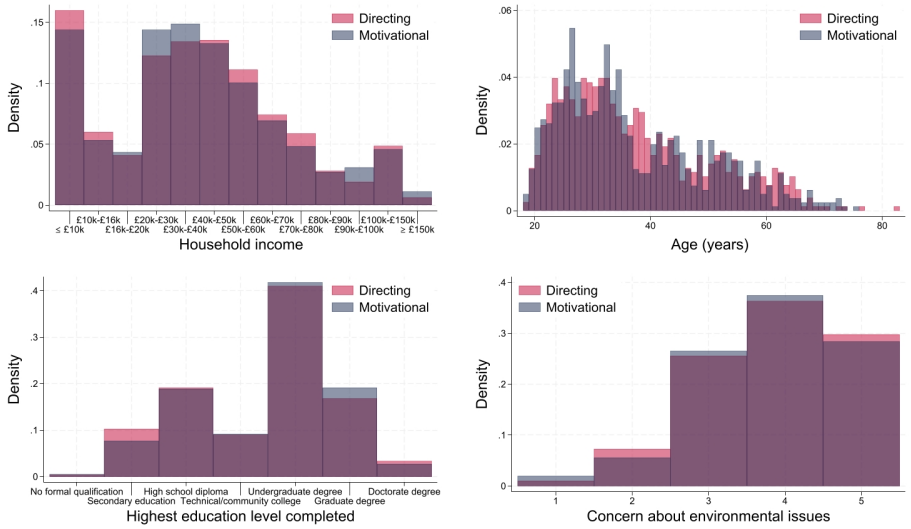


Figure 1: Densities are displayed regarding socioeconomic characteristic under MI and DS: On top left the density of *Income*, on top right the density of *Age* in years of participants, on bottom left *Education*, and on bottom right *Concern about environmental issues*; in all cases, overlaid histograms are aimed to highlight differences.

Household income, hereafter referred to as *Income*, is divided into 13 classes, ranging from less than £10,000 to more than £150,000. Again, through the Wilcoxon rank-sum test, we can affirm that the sample is balanced. The graphical representation is again in Figure 1 and the test result is in Table 3. The dataset containing the experimental results, responses to the final survey, participants' socio-economic variables, conversation transcripts, interaction times, and number of interactions is available on the project page on the OSF platform (Link).

3 Main results

The first study did not yield statistically significant results for the preregistered outcome variables. Based on the observed differences between experimental conditions, we developed more targeted and directional hypotheses for the second study. In this section, we present the results for the preregistered outcomes of the second study, using pooled data from both studies. Notably, both studies employed an identical experimental design. A detailed breakdown of the results for each study, along with a comprehensive description of the objectives and methods, is provided in the Supplementary Information (SI).

3.1 Self-assessment of interest

The distributions of the *Self-assessment of interest* under the two treatments are plotted in Figure 2.

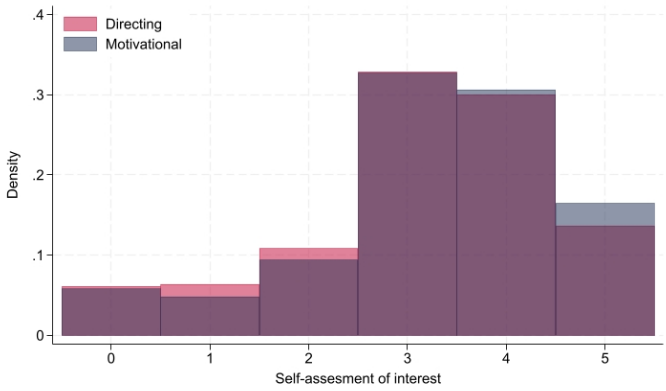


Figure 2: Densities of *Self-assessment of interest* are displayed for MI and DS treatments, with overlaid histograms to highlight differences.

We perform a two sample Wilcoxon rank-sum test with the following hypotheses:

- H_0: interest(treat. = Directing) = interest(treat. = Motivational)
- H_1: interest(treat. = Directing) < interest(treat. = Motivational)

Results are reported in Table 4.

The null hypothesis can be rejected with a level of significance lower than the critical value of 5%. Thus, we find evidence that *Self-assessment of interest* is (stochastically) larger under MI than under DS. See Table 2 in the Appendix of *Supplementary Information* for an ordered probit analysis, which confirms the main results but was not preregistered.

Table 4: Ranksum test for the *Self-assessment of interest*.

Treatment	Obs	Rank sum	Expected
directing	782	604963	621299
motivational	806	656703	640367
z = -1.852		Prob > |z| = 0.0320	

3.2 Self-assessment of learning

The distributions of *Self-assessment of learning* under the two treatments are plotted in Figure 3.

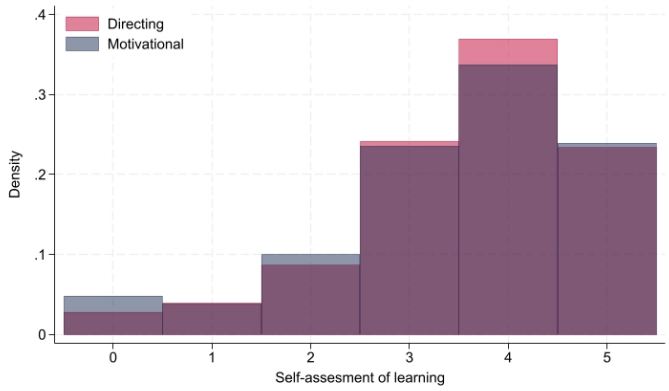


Figure 3: Densities of *Self-assessment of learning* are displayed for MI and DS treatments, with overlaid histograms to highlight differences.

We perform a two sample Wilcoxon rank-sum test with the following hypotheses:

- H_0: learn(treat. = Directing) = learn(treat. = Motivational)
- H_1: learn(treat. = Directing) > learn(treat. = Motivational)

Results are reported in Table 5.

Table 5: Ranksum test for *Self-assessment of learning*.

Treatment	Obs	Rank sum	Expected
directing	782	629971	621299
motivational	806	631695	640367
z = 0.985		Prob > |z| = 0.1622	

The null hypothesis cannot be rejected at any standard level of significance. Therefore, we find no evidence that *Self-assessment of learning* is (stochastically) larger

under DS than under MI. This null result represents the preregistered comparison between the two communication styles and indicates that their overall effect on perceived learning does not differ significantly. In Section 4, we complement this analysis by exploring whether the overall effect can be decomposed into direct and indirect components. Specifically, we investigate whether DS exerts a positive direct effect on learning, while MI indirectly promotes learning by increasing participants' engagement and interest. The exploratory analysis suggests that these opposing mechanisms may partially offset each other, with a positive direct effect of DS emerging only after controlling for the indirect pathways. These exploratory findings provide a possible interpretation of the null overall effect, but do not alter the main preregistered conclusion that no significant difference in overall learning is detected between the two communication styles. The results of this analysis are presented in Table 1 in the Appendix of the *Supplementary Information*.

3.3 Engagement

The distributions of *Engagement* under the two treatments are plotted in Figure 4.

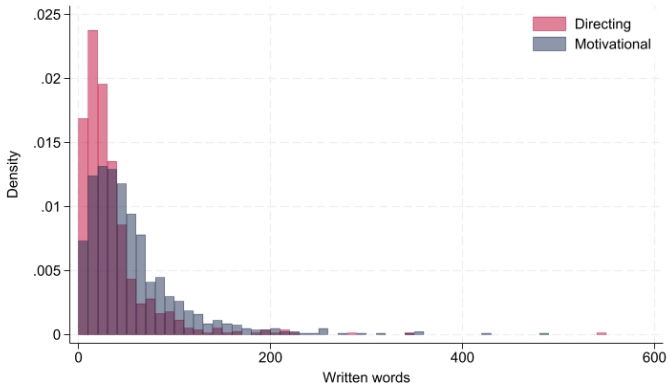


Figure 4: Densities of *Engagement* are displayed for MI and DS treatments, with overlaid histograms to highlight differences.

We perform a two sample Wilcoxon rank-sum test with the following hypotheses:

- H_0: engagement(treat. = Directing) = engagement(treat. = Motivational)
- H_1: engagement(treat. = Directing) < engagement(treat. = Motivational)

Results are reported in Table 6.

The null hypothesis can be rejected at any standard level of significance. Thus, we find evidence that *Engagement* is (stochastically) larger under MI than under DS. See Table 2 in Appendix sextion in *Supplementary Information* for an analysis based on ordered probit regressions. In Section 4, we show that the results obtained using the

Table 6: Ranksum test for *Engagement*.

Treatment	Obs	Rank sum	Expected
directing	782	511001	621299
motivational	806	750665	640367
z = -12.074	Prob > ||z|| = 0.0000		

number of written words as a measure of engagement are replicated when using other non-preregistered measures such as Time taken, Rounds, and Words per round. The distributions of these alternative measures are displayed in Figure 5, and the results of the Wilcoxon rank-sum test are summarized in Table 7.

4 Exploratory analysis

In this section, we report additional exploratory analyses that were not included in the preregistration. Specifically, we examine alternative measures of engagement to test the robustness of the results presented in subsection 3.3. We analyze the effects of the experimental conditions on *Willingness to receive costly information* and *Self-assessment of satisfaction*, both measured in the final survey, see subsection 2.2, in both studies. In the preregistration of the first study, subsection 2.3, we included a hypothesis regarding a treatment effect on these variables. These hypotheses were removed in the preregistration of the second study, subsection 2.3, and we therefore present the related results here as part of the exploratory analysis. Finally, we further investigate the impact of the DS condition on learning, partially confirming the effect hypothesized in the preregistration of the second study and expanding the analysis conducted in subsection 3.2.

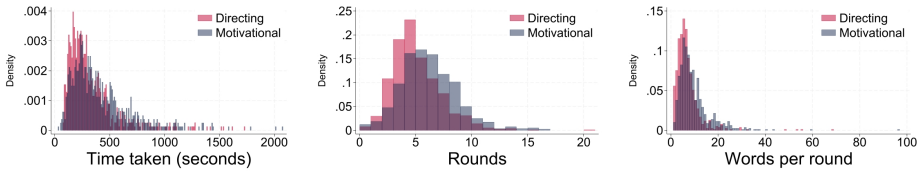


Figure 5: Densities are displayed regarding alternative measures of engagement under MI and DS: from left to right, the density of *Time taken* in seconds, the density of *Rounds* written by the user, and the density of *Words per round* by the user; in all cases, overlaid histograms are aimed to highlight differences.

In the preregistration of the second study, we committed to the number of words written by the user as a variable for measuring participant engagement. Here, we show that the use of alternative variables to describe engagement yields consistent

results. These variables are *Time taken*, *Rounds*, and *Words per round*. *Time taken* measures the overall time elapsed between the start and the conclusion of the study on the Prolific platform, provided by Prolific itself. *Rounds* represents the number of interactions between the user and the chatbot, and *Words per round* measures the average number of words per round written by users.

Table 7: Ranksum test of other measures of engagement: *Time taken*, *Rounds*, and *Words per round*.

Variable	z	Prob $> \|\|z\|\|$
Time taken	-7.085	0.0000
Rounds	-9.788	0.0000
Words per round	-10.506	0.0000

The graphical representation of the distributions of these variables for each treatment is shown in Figure 5.

For each of the three variables, a Wilcoxon rank-sum test is performed (see Table 7). Based on these tests, we can reject the null hypothesis at any standard level of significance. Thus, we find evidence that *Time taken*, *Rounds*, and *Words per round* are (stochastically) larger under MI than under DS.

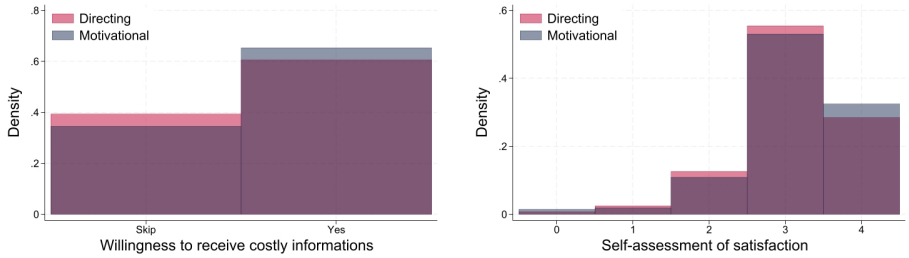


Figure 6: Densities are displayed regarding the variables measured through the final survey, excluded from the primary analysis, under MI and DS: on the left, the density of *Willingness to receive costly information* and, on the right, the density of *Self-assessment of satisfaction*; in all cases, overlaid histograms are aimed to highlight differences.

In the final survey, in addition to *Self-assessment of learning* and *Self-assessment of interest*, there are also questions regarding *Willingness to receive costly information* and *Self-assessment of satisfaction*.

For *Self-assement of satisfaction*, we make use of a standard customer satisfaction 5-smiley scale. This is relevant for us, in that it captures the likelihood that a pleasant experience will be repeated over time by users without any coercitive tool or incentive.

These variables, which had not provided clear indications in the first study, were excluded from the main analysis in the preregistration of the second study. In this section, we provide the results of the secondary analysis on these variables, conducted with the complete sample. The graphical representation of the two variables based on treatment is shown in Figure 6.

The results of the statistical analysis are reported in Table 8. *Willingness to receive costly information*, hereafter referred to as willingness, is a binary variable. It represents participants' responses to the question: 'Would you authorize us to send you one or more communications about sustainability topics using the Prolific messaging system? The authorization is optional and at your discretion'. Possible answers are 'Skip', in which case the variable takes value 0, and 'yes', in which case the variable takes value 1. We thus employ probit models with various specifications. In all three versions, (1), (2) and (3), the coefficient of the treatment variable remains significant and negative. Since the treatment variable takes a value of 0 in the case of MI and 1 in the case of DS, we can conclude that Motivational Interviewing increases the probability that a participant responds positively to the request for authorization to receive further communications. *Concern about environmental issues*, in models (2) and (3), has positive and highly significant coefficients. Individuals who express greater concern for environmental issues are more likely to accept receiving further communications.

Self-assessment of satisfaction, hereafter referred to as satisfaction, is measured using a smiley rating scale with five options, converted into values ranging from 0 to 4. As evident in Figure 6, the subplot on the right, the satisfaction variable appears particularly imbalanced. Indeed, responses 0, 1, and 2, corresponding to negative and neutral responses, have a markedly lower frequency compared to responses 3 and 4, corresponding to positive responses. To address estimation issues, such as unstable parameter estimates and inflated standard errors, we merge responses (0), (1), and (2) into a single category. The new variable, termed satisfaction pooled, will thus be an ordinal discrete variable with 3 values, (0), (1), and (2). For analysis, we employ an ordered probit with various specifications. Consistent with previous findings, we find that Motivational Interviewing treatment increases the probability that participants express a high degree of satisfaction, in models (4), (5), and (6). Similarly, participants expressing concern about environmental issues are more likely to express a high degree of satisfaction, in models (5) and (6).

As we have seen in the main analysis, the positive impact of DS on learning, expected in the preregistration, is not significant. We now attempt to isolate the direct effects of the treatment on learning from the indirect effects. As observed in the main analysis, MI has a positive and significant effect on interest, and in the exploratory analysis we also find that MI has a positive and significant effect on the time participants spend with the chatbot. Furthermore, the Spearman correlation test reveals that Learning has a positive correlation with both *Time taken* and *Self-*

Table 8: Probit and ordered probit regression results for *Willingness to receive costly information* and *Self-assessment of satisfaction*.

Dep. Var	Willingness to receive costly info			Self-assessment of satisfaction		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.126**	-0.130**	-0.113*	-0.0956*	-0.0985*	-0.103*
	(0.0643)	(0.0644)	(0.0652)	(0.0565)	(0.0566)	(0.0570)
Concern env.		0.123***	0.123***		0.118***	0.136***
		(0.0333)	(0.0344)		(0.0307)	(0.0319)
Age			0.0607***			0.0217
			(0.0157)			(0.0138)
Age^2			-0.000715***			-0.000268
			(0.000187)			(0.000167)
Education			0.0239			-0.0749**
			(0.0351)			(0.0316)
Income			-0.00586			-0.00577
			(0.00863)			(0.00736)
Gender			0.0795			0.0560
			(0.0629)			(0.0555)
Constant	0.396***	-0.0745	-1.304***			
	(0.0454)	(0.135)	(0.335)			
/cut1				-1.077***	-0.631***	-0.392
				(0.0482)	(0.125)	(0.287)
/cut2				0.462***	0.918***	1.168***
				(0.0431)	(0.126)	(0.289)
Obs.	1588	1588	1570	1,588	1,588	1,570
Pseudo R2	0.0019	0.0083	0.0172	0.0009	0.0060	0.0096

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

assessment of interest, at any level of significance, this relationship between generated interest, engagement, and learning outcomes is consistent with previous studies in educational psychology (Krapp, 1999). We are inclined to interpret *Self-assessment of interest* as a qualitative effect and *Time taken* as a quantitative effect of the treatment. Consequently, MI leads to increased interest and time spent with respect to DS, thus indirectly enhancing learning. To isolate the direct effects of the treatment, we regress *Self-assessment of learning* on the treatment, *Time taken*, and *Self-assessment of interest*, while also controlling for the available socioeconomic variables. Conceptually, self-assessed learning in a chatbot interaction is expected to depend primarily on the informational content, explanations, and feedback provided by the chatbot, rather than on the amount of text written by the participant. In contrast, the user's word counts

(*Words user* and *Words per round*) capture behavioral aspects of the interaction, such as conversational effort, pacing, and willingness to sustain the dialogue. As such, these variables are expected to influence perceived learning mainly through their effects on engagement and situational interest, rather than by directly increasing the informational content received by participants. Based on this rationale, to address the potential endogeneity of *Time taken* and *Self-assessment of interest*, we estimate an Extended Ordered Probit model using the total number of words written and the mean number of words per round as instruments for *Time taken* and *Self-assessment of interest*, respectively. The results of the Extended Ordered Probit regression are presented in the Appendix section of the *Supplementary Information*. They suggest that, in line with our expectation, DS exerts a positive direct effect on learning relative to MI, although this effect becomes apparent only after accounting for the indirect pathways operating through participants' engagement and interest.

5 Discussion and conclusions

This research investigated the effectiveness of different communication styles employed by AI-powered chatbots in conversations about sustainability-related issues. In our experiments, these conversations were structured around the Sustainable Development Goals (SDGs), which served as a comprehensive framework for discussing sustainability. Two outcome variables are self-assessment of interest and self-assessment of learning, measured through a brief questionnaire at the end of the experiment. The third is the direct measure of the number of words written by participants, used as a proxy for engagement. Our findings suggest that motivational interviewing (MI) significantly increases both engagement and interest in sustainability with respect to directing style (DS). At the same time, no statistically significant difference was observed between MI and DS regarding learning. Therefore, applying MI to environmental issues represents a promising avenue for promoting sustainable behaviors and decision-making. By focusing on individual motivation and resolving ambivalence, MI can be a crucial tool in raising awareness on some of the most pressing environmental challenges of our time, without losing on the received informational content. The results align with previous findings where MI is delivered through digital tools, leading to increased engagement (McClure et al., 2013; He et al., 2024) and generating high acceptability and positive behavior change (Shingleton and Palfai, 2016). Unlike prior studies (He et al., 2022), we find significant effects of MI on engagement and self-assessment of interest even through brief interventions, with a median duration of less than five minutes.

The use of AI-powered chatbots holds significant potential to reduce costs, ease the burden on specialists and training programs, and expand access to a broader range of users. These features are what make them widely adopted across various domains, including the promotion of sustainable behaviors and practices incouraged by SDGs (Nguyen et al., 2025; Giudici et al., 2023; Aggarwal et al., 2023; Zhang et al.,

2020). Our experiment shows that the communicative style of such chatbots matters, suggesting that implementing MI-based communication styles can effectively promote engagement and motivation toward sustainable practices.

Further research could elucidate the extent by which AI-powered MI influences pro-environmental behaviors and explore its efficacy in diverse contexts. For instance, future studies could investigate the application of MI to chatbots designed to discuss about specific sustainability-related subjects, such as energy conservation, waste reduction, or biodiversity protection. Also, AI-powered MI could be explored in application to other societally relevant issues, such as adherence to vaccination campaigns, adoption of healthy lifestyles, attitudes towards immigrants, or gender differences. Another interesting route of research could examine the long-term effects of MI interventions on individuals' attitudes, behaviors, and decision-making processes. Longitudinal studies could provide valuable insights into the durability and persistence of MI-induced changes in pro-environmental behaviors. Furthermore, since there are differences in the use of chatbots among subgroups of the population (Schubel et al., 2021), it is interesting to assess any differences in the effects of chatbot communication styles in subgroups.

It is important to acknowledge the limitations inherent in our measurement approach, which relies on self-reported data in response to ad hoc questions. Our measure of self-assessment of interest captures situational interest, which is the immediate, affective response elicited by the chat interaction (Liu et al., 2022). This metric reflects a transitory state and does not necessarily indicate a change in the participant's deeper, more enduring personal interest in sustainability. This distinction was introduced by Hidi and Baird (1986), who differentiate between the conditions that trigger interest and those that sustain it over time, a framework later referred to in educational psychology as the "catch and hold" model (Mitchell, 1993). Consequently, we cannot robustly claim that the intervention fosters long-term interest in sustainability issues. Similarly, our measure of self-assessment of learning captures perceived learning, which is a subjective judgment of knowledge gain. This metric should not be equated with objective learning (Persky et al., 2020). A high score could reflect factors such as participant engagement or the novelty of the topic rather than a verifiable acquisition of specific knowledge (Sitzmann et al., 2010). Future studies could strengthen these findings with more concrete evidence of impact by overcoming the methodological complications inherent to chatbot conversations and incorporating objective knowledge tests (e.g., pre- and post-test assessments) and behavioral measures (e.g., tracking subsequent information-seeking). By examining the correspondence between self-assessments and more objective indicators of interest and learning, such studies could also help validate our survey items, in line with the Sustainability Consciousness Questionnaire (Gericke et al., 2019; Bacci et al., 2024).

We acknowledge additional limitations of this study related to the use of participants recruited through Prolific. On one hand, individuals who use Prolific and similar

online platforms are among the most reachable by chatbot-based interventions; on the other hand, this recruitment approach may reduce the external validity of our findings, particularly if they are to be generalized to populations with lower familiarity with digital technologies.

Future research could also explore the potential of other communication styles when mediated by chatbot conversations. For example, the efficacy of narrative-based approaches could be investigated (Hinyard and Kreuter, 2007; Richter et al., 2019), as well as other changes along different dimensions of communication styles (De Vries et al., 2009, 2013). By identifying the most effective communication strategies for different contexts and target audiences, based on cultural, socio-economic, and demographic differences, researchers can contribute to the development of tailored interventions that maximize the impact of sustainability messaging.

AI-powered conversational chatbots hold immense potential in amplifying the reach and impact of MI interventions and, more in general, communication interventions. Chatbots can indeed serve as a scalable and cost-effective platform for delivering tailored MI support. This is particularly significant considering the often high costs associated with implementing traditional in-person MI, which can limit accessibility. Future research could focus on optimizing chatbot design, improving natural language processing capabilities, and enhancing the personalization of MI interventions delivered through chatbots. By leveraging the power of AI, researchers can work towards creating more engaging, interactive, and effective tools for promoting sustainable behaviors on a global scale.

Funding declaration

We gratefully acknowledge financial support from the Italian Ministry of Education, University and Research (MIUR) through the PRIN project Co.S.Mo.Pro.Be. 'Cognition, Social Motives and Prosocial Behavior' (grant n. 20178293XT), and from the European Union - NextGenerationEU through the project ECoHeTE 'Effective Communication for Healthcare: Theory and Evidence'.

Ethics declarations

Ethical approval: This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of University of Florence, approval number 288/2023.

Informed consent: Informed consent was obtained from all participants at the start of the online experiment. Participants, recruited via Prolific, were instructed to read the informed consent form upon accessing the chatbot and to indicate their decision to participate by clicking the appropriate button. If consent was not provided, the experimental procedure was immediately terminated. Participants were informed that

the study was conducted in compliance with GDPR 679/2016, ensuring that personal data were processed within legal limits, exclusively for scientific purposes, and collected in a completely anonymous form, making it impossible to trace participants' identities. Participation was entirely voluntary, with the option to withdraw at any time and without providing any explanation.

References

Aggarwal, A., Tam, C.C., Wu, D., Li, X., Qiao, S., 2023. Artificial intelligence–based chatbots for promoting health behavioral changes: systematic review. Journal of medical Internet research 25, e40789.

Aher, G., Arriaga, R.I., Kalai, A.T., 2022. Using large language models to simulate multiple humans. arXiv preprint arXiv:2208.10264 .

Alt, M., Bruns, H., DellaValle, N., Murauskaite-Bull, I., 2024. Synergies of interventions to promote pro-environmental behaviors–a meta-analysis of experimental studies. Global Environmental Change 84, 102776.

Altay, S., Hacquin, A.S., Chevallier, C., Mercier, H., 2023. Information delivered by a chatbot has a positive impact on covid-19 vaccines attitudes and intentions. Journal of Experimental Psychology: Applied 29, 52.

Altay, S., Schwartz, M., Hacquin, A.S., Allard, A., Blancke, S., Mercier, H., 2022. Scaling up interactive argumentation by providing counterarguments with a chatbot. Nature Human Behaviour 6, 579–592.

Argyle, L.P., Busby, E.C., Fulda, N., Gubler, J.R., Rytting, C., Wingate, D., 2023. Out of one, many: Using language models to simulate human samples. Political Analysis 31, 337–351.

Bacci, S., Bertaccini, B., Macrì, E., Pettini, A., 2024. Measuring sustainability consciousness in italy. Quality & Quantity 58, 4751–4778.

Bickmore, T.W., Schulman, D., Sidner, C., 2013. Automated interventions for multiple health behaviors using conversational agents. Patient Education and Counseling 92, 142–148.

Binz, M., Schulz, E., 2023. Using cognitive psychology to understand gpt-3. Proceedings of the National Academy of Sciences 120, e2218523120.

Breckenridge, L.A., Burns, D., Nye, C., 2022. The use of motivational interviewing to overcome covid-19 vaccine hesitancy in primary care settings. Public Health Nursing 39, 618–623.

Brown, A., Kumar, A.T., Melamed, O., Ahmed, I., Wang, Y.H., Deza, A., Morcos, M., Zhu, L., Maslej, M., Minian, N., et al., 2023. A motivational interviewing chatbot with generative reflections for increasing readiness to quit smoking: Iterative development study. JMIR Mental Health 10, e49132.

Brummernhenrich, B., Paulus, C.L., Jucks, R., 2025. Applying social cognition to feedback chatbots: Enhancing trustworthiness through politeness. British Journal of Educational Technology .

Cai, N., Gao, S., Yan, J., 2024. How the communication style of chatbots influences consumers' satisfaction, trust, and engagement in the context of service failure. Humanities and Social Sciences Communications 11, 1–11.

Capraro, V., Di Paolo, R., Perc, M., Pizziol, V., 2024a. Language-based game theory in the age of artificial intelligence. Journal of the Royal Society Interface 21, 20230720.

Capraro, V., Di Paolo, R., Pizziol, V., 2023. Predict-ai-bility of how humans balance self-interest with the interest of others. arXiv preprint arXiv:2307.12776 .

Capraro, V., Halpern, J.Y., Perc, M., 2024b. From outcome-based to language-based preferences. Journal of Economic Literature 62, 115–154.

Ceha, J., Lee, K.J., Nilsen, E., Goh, J., Law, E., 2021. Can a humorous conversational agent enhance learning experience and outcomes?, in: Proceedings of the 2021 CHI Conference on human factors in computing systems, pp. 1–14.

Chen, Y., Liu, T.X., Shan, Y., Zhong, S., 2023. The emergence of economic rationality of gpt. Proceedings of the National Academy of Sciences 120, e2316205120.

Conrady, T., Kruschwitz, A., Stamminger, R., 2014. Influencing the sustainability of washing behavior by using motivational interviewing. Energy Efficiency 7, 163–178.

Crolic, C., Thomaz, F., Hadi, R., Stephen, A.T., 2022. Blame the bot: Anthropomorphism and anger in customer–chatbot interactions. Journal of Marketing 86, 132–148.

Cuijpers, P., Miguel, C., Ciharova, M., Harrer, M., Karyotaki, E., 2024. Non-directive supportive therapy for depression: A meta-analytic review. Journal of Affective Disorders 349, 452–461.

De Neve, J.E., Sachs, J.D., 2020. The sdgs and human well-being: A global analysis of synergies, trade-offs, and regional differences. Scientific Reports 10, 15113.

De Vries, R.E., Bakker-Pieper, A., Alting Siberg, R., van Gameren, K., Vlug, M., 2009. The content and dimensionality of communication styles. Communication Research 36, 178–206.

De Vries, R.E., Bakker-Pieper, A., Konings, F.E., Schouten, B., 2013. The communication styles inventory (csi) a six-dimensional behavioral model of communication styles and its relation with personality. Communication Research 40, 506–532.

Del Rio-Chanona, R.M., Pangallo, M., Hommes, C., 2025. Can generative ai agents behave like humans? Evidence from laboratory market experiments. arXiv preprint arXiv:2505.07457 .

Dillion, D., Tandon, N., Gu, Y., Gray, K., 2023. Can ai language models replace human participants? Trends in Cognitive Sciences 27, 597–600.

Endrejat, P.C., Baumgarten, F., Kauffeld, S., 2017. When theory meets practice: Combining lewin's ideas about change with motivational interviewing to increase energy-saving behaviours within organizations. Journal of Change Management 17, 101–120.

Galvão Gomes da Silva, J., Kavanagh, D.J., Belpaeme, T., Taylor, L., Beeson, K., Andrade, J., 2018. Experiences of a motivational interview delivered by a robot: qualitative study. Journal of Medical Internet Research 20, e116.

Gardiner, P.M., McCue, K.D., Negash, L.M., Cheng, T., White, L.F., Yinusa-Nyahkoon, L., Jack, B.W., Bickmore, T.W., 2017. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. Patient Education and Counseling 100, 1720–1729.

Gericke, N., Boeve-de Pauw, J., Berglund, T., Olsson, D., 2019. The sustainability consciousness questionnaire: The theoretical development and empirical validation of an evaluation instrument for stakeholders working with sustainable development. Sustainable Development 27, 35–49.

Giudici, M., Abbo, G.A., Belotti, O., Braccini, A., Dubini, F., Izzo, R.A., Crovari, P., Garzotto, F., 2023. Assessing llms responses in the field of domestic sustainability: An exploratory study, in: 2023 Third International Conference on Digital Data Processing (DDP), IEEE. pp. 42–48.

Guo, F., 2023. Gpt agents in game theory experiments. arXiv preprint arXiv:2305.05516 .

Gusella, F., Vicario, E., 2025. Generative agents and expectations: Do llms align with heterogeneous agent models? arXiv preprint arXiv:2511.08604 .

Hamzehgardeshi, Z., Elyasi, F., Moosazadeh, M., Ahmadi, I., Jahanfar, S., Ahmadian, M., Ansari, F., 2024. The effect of social network based motivational interviewing on health behaviors among infertile women with polycystic ovary syndrome: A randomized controlled trial. Scientific Reports 14, 12049.

He, L., Basar, E., Krahmer, E., Wiers, R., Antheunis, M., 2024. Effectiveness and user experience of a smoking cessation chatbot: mixed methods study comparing motivational interviewing and confrontational counseling. Journal of Medical Internet Research 26, e53134.

He, L., Basar, E., Wiers, R.W., Antheunis, M.L., Krahmer, E., 2022. Can chatbots help to motivate smoking cessation? a study on the effectiveness of motivational interviewing on engagement and therapeutic alliance. BMC Public Health 22, 726.

Herzing, M., Wickström, H., Jacobsson, A., Källmén, H., Forsberg, L., 2023. Enhancing compliance with waste sorting regulations through inspections and motivational interviewing. Waste Management & Research , 0734242X231154145.

Hew, K.F., Huang, W., 2023. Promoting engagement in online learning beyond covid-19: Possible strategies and directions for future research. Future in Educational Research 1, 27–49.

Hidi, S., Baird, W., 1986. Interestingness—a neglected variable in discourse processing. Cognitive science 10, 179–194.

Hillenbrand, A., Verrina, E., 2022. The asymmetric effect of narratives on prosocial behavior. Games and Economic Behavior 135, 241–270.

Hinyard, L.J., Kreuter, M.W., 2007. Using narrative communication as a tool for health behavior change: a conceptual, theoretical, and empirical overview. Health Education & Behavior 34, 777–792.

Horton, J.J., 2023. Large language models as simulated economic agents: What can we learn from homo silicus? Technical Report. National Bureau of Economic Research.

Kar, A.K., Choudhary, S.K., Singh, V.K., 2022. How can artificial intelligence impact sustainability: A systematic literature review. Journal of Cleaner Production 376, 134120.

Klonek, F., Kauffeld, S., 2012. Sustainability goes change talk: Can motivational interviewing be used to increase pro-environmental behavior?, in: Proceedings of Measuring Behavior, pp. 297–302.

Klonek, F.E., Güntner, A.V., Lehmann-Willenbrock, N., Kauffeld, S., 2015. Using motivational interviewing to reduce threats in conversations about environmental behavior. Frontiers in Psychology 6, 1015.

Konya-Baumbach, E., Biller, M., von Janda, S., 2023. Someone out there? a study on the social presence of anthropomorphized chatbots. Computers in Human Behavior 139, 107513.

Koster, R., Balaguer, J., Tacchetti, A., Weinstein, A., Zhu, T., Hauser, O., Williams, D., Campbell-Gillingham, L., Thacker, P., Botvinick, M., et al., 2022. Human-centred mechanism design with democratic ai. Nature Human Behaviour 6, 1398–1407.

Krapp, A., 1999. Interest, motivation and learning: An educational-psychological perspective. European journal of psychology of education 14, 23–40.

Liu, C.C., Liao, M.G., Chang, C.H., Lin, H.M., 2022. An analysis of children'interaction with an ai chatbot and its impact on their interest in reading. Computers & Education 189, 104576.

Lorinkova, N.M., Pearsall, M.J., Sims Jr, H.P., 2013. Examining the differential longitudinal performance of directive versus empowering leadership in teams. Academy of Management Journal 56, 573–596.

Maggioni, M.A., Rossignoli, D., 2023. If it looks like a human and speaks like a human... communication and cooperation in strategic human–robot interactions. Journal of Behavioral and Experimental Economics 104, 102011.

Martin, S.L., Liao, H., Campbell, E.M., 2013. Directive versus empowering leadership: A field experiment comparing impacts on task proficiency and proactivity. Academy of Management Journal 56, 1372–1395.

McClure, J.B., Shortreed, S.M., Bogart, A., Derry, H., Riggs, K., St John, J., Nair, V., An, L., 2013. The effect of program design on engagement with an internet-based smoking intervention: randomized factorial trial. Journal of medical Internet research 15, e2508.

Mei, Q., Xie, Y., Yuan, W., Jackson, M.O., 2024. A turing test of whether ai chatbots are behaviorally similar to humans. Proceedings of the National Academy of Sciences 121, e2313925121.

Miller, W., Rollnick, S., 1991. Motivational Interviewing: Preparing People to Change Addictive Behavior. Guilford Publications.

Miller, W.R., Rollnick, S., 2009. Ten things that motivational interviewing is not. Behavioural and Cognitive Psychotherapy 37, 129–140.

Miller, W.R., Rollnick, S., 2012. Motivational interviewing: Helping people change. Guilford press.

Mitchell, M., 1993. Situational interest: Its multifaceted structure in the secondary school mathematics classroom. Journal of educational psychology 85, 424.

Nguyen, M.T., Thach, K.T.D., Nguyen, C.N.L., Nguyen, A.C., Doan, H.K., 2025. The influence of ai chatbots on green satisfaction and loyalty: Evidence from sustainability-driven consumer behavior. Journal of Global Marketing , 1–30.

Nishant, R., Kennedy, M., Corbett, J., 2020. Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. International journal of information management 53, 102104.

Numata, T., Sato, H., Asa, Y., Koike, T., Miyata, K., Nakagawa, E., Sumiya, M., Sadato, N., 2020. Achieving affective human–virtual agent communication by enabling virtual agents to imitate positive expressions. Scientific Reports 10, 5977.

Ortega-Ochoa, E., Arguedas, M., Daradoumis, T., 2024a. Empathic pedagogical conversational agents: a systematic literature review. British Journal of Educational Technology 55, 886–909.

Ortega-Ochoa, E., Pérez, J.Q., Arguedas, M., Daradoumis, T., Puig, J.M.M., 2024b. The effectiveness of empathic chatbot feedback for developing computer competencies, motivation, self-regulation, and metacognitive reasoning in online higher education. Internet of Things 25, 101101.

Pan, D., Huey Jr, S.J., Heflin, L.H., 2019. Ethnic differences in response to directive vs. non-directive brief intervention for subsyndromal depression. Psychotherapy Research 29, 186–197.

Persky, A.M., Lee, E., Schlesselman, L.S., 2020. Perception of learning versus performance as outcome measures of educational research. American Journal of Pharmaceutical Education 84, ajpe7782.

Prochaska, J.J., Vogel, E.A., Chieng, A., Kendra, M., Baiocchi, M., Pajarito, S., Robinson, A., 2021. A therapeutic relational agent for reducing problematic substance use (woebot): development and usability study. Journal of Medical Internet Research 23, e24850.

Rhim, J., Kwak, M., Gong, Y., Gweon, G., 2022. Application of humanization to survey chatbots: Change in chatbot perception, interaction experience, and survey data quality. Computers in Human Behavior 126, 107034.

Richter, A., Sieber, A., Siebert, J., Miczajka-Rußmann, V., Zabel, J., Ziegler, D., Hecker, S., Frigerio, D., 2019. Storytelling for narrative approaches in citizen science: Towards a generalized model. Journal of Science Communication 18, A02.

Rogers, C., 2012. Client centered therapy (new ed). Hachette UK.

Ruane, E., Farrell, S., Ventresque, A., 2021. User perception of text-based chatbot personality, in: Chatbot Research and Design: 4th International Workshop,

CONVERSATIONS 2020, Virtual Event, November 23–24, 2020, Revised Selected Papers 4, Springer. pp. 32–47.

Sachs, J.D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., Rockström, J., 2019. Six transformations to achieve the sustainable development goals. Nature Sustainability 2, 805–814.

Schaaff, K., Heidelmann, M.A., 2024. Impacts of anthropomorphizing large language models in learning environments. arXiv preprint arXiv:2408.03945 .

Schmidt-Traub, G., Kroll, C., Teksoz, K., Durand-Delacre, D., Sachs, J.D., 2017. National baselines for the sustainable development goals assessed in the sdg index and dashboards. Nature Geoscience 10, 547–555.

Schubel, L.C., Wesley, D.B., Booker, E., Lock, J., Ratwani, R.M., 2021. Population subgroup differences in the use of a covid-19 chatbot. NPJ Digital Medicine 4, 30.

Shingleton, R.M., Palfai, T.P., 2016. Technology-delivered adaptations of motivational interviewing for health-related behaviors: A systematic review of the current research. Patient education and counseling 99, 17–35.

Sitzmann, T., Ely, K., Brown, K.G., Bauer, K.N., 2010. Self-assessment of knowledge: A cognitive learning or affective measure? Academy of Management Learning & Education 9, 169–191.

Smutny, P., Schreiberova, P., 2020. Chatbots for learning: A review of educational chatbots for the facebook messenger. Computers & Education 151, 103862.

Stein, N., Brooks, K., et al., 2017. A fully automated conversational artificial intelligence for weight loss: longitudinal observational study among overweight and obese adults. JMIR Diabetes 2, e8590.

Stephens, T.N., Joerin, A., Rauws, M., Werk, L.N., 2019. Feasibility of pediatric obesity and prediabetes treatment support through tess, the ai behavioral coaching chatbot. Translational Behavioral Medicine 9, 440–447.

Strachan, J.W., Albergo, D., Borghini, G., Pansardi, O., Scaliti, E., Gupta, S., Saxena, K., Rufo, A., Panzeri, S., Manzi, G., et al., 2024. Testing theory of mind in large language models and humans. Nature Human Behaviour , 1–11.

Tagkaloglou, S., Kasser, T., 2018. Increasing collaborative, pro-environmental activism: The roles of motivational interviewing, self-determined motivation, and self-efficacy. Journal of Environmental Psychology 58, 86–92.

Thorne, F.C., 1948. Principles of directive counseling and psychotherapy. American Psychologist 3, 160.

Van Pinxteren, M.M., Pluymaekers, M., Lemmink, J., Krispin, A., 2023. Effects of communication style on relational outcomes in interactions between customers and embodied conversational agents. Psychology & Marketing 40, 938–953.

Vicario, E., Bilancini, E., Boncinelli, L., 2026. Multilingual large language models and cultural diversity: Evidence from civic and moral judgments. Available at SSRN 6633921 .

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M., Fuso Nerini, F., 2020. The role of artificial intelligence in achieving the sustainable development goals. Nature communications 11, 233.

Wambsganss, T., Kueng, T., Soellner, M., Leimeister, J.M., 2021. Arguetutor: An adaptive dialog-based learning system for argumentation skills, in: Proceedings of the 2021 CHI conference on human factors in computing systems, pp. 1–13.

Xu, Y., Zhang, J., Deng, G., 2022. Enhancing customer satisfaction with chatbots: The influence of communication styles and consumer attachment anxiety. Frontiers in Psychology 13, 902782.

Yin, J., Goh, T.T., Hu, Y., 2024. Interactions with educational chatbots: the impact of induced emotions and students' learning motivation. International Journal of Educational Technology in Higher Education 21, 47.

Yu, S., Zhao, L., 2024. Emojifying chatbot interactions: An exploration of emoji utilization in human-chatbot communications. Telematics and Informatics 86, 102071.

Zewail, A., Figueroa, A., Graham, J., Atari, M., 2026. Moral stereotyping in large language models. Proceedings of the National Academy of Sciences 123, e2519941123.

Zhang, J., Oh, Y.J., Lange, P., Yu, Z., Fukuoka, Y., 2020. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet. Journal of medical Internet research 22, e22845.

Zhang, S., Song, J., 2024. A chatbot based question and answer system for the auxiliary diagnosis of chronic diseases based on large language model. Scientific Reports 14, 17118.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryinformation.pdf