Scuola IMT Alti Studi Lucca

AI-powered chatbots: effective communication styles for Sustainable Development Goals

Questa è la versione preprint della seguente opera:

Original
AI-powered chatbots: effective communication styles for Sustainable Development Goals / Bilancini, E., Boncinelli, L., Vicario, E.. - (2024). [10.48550/ARXIV.2407.01057]

Availability:
This version is available at: 20.500.11771/37040

Publisher:

Published
DOI:10.48550/ARXIV.2407.01057

Terms of use:

This publication is made accessible in accordance with the terms for deposit in the institutional repository, as defined by the IMT School for Advanced Studies Lucca's Open Access Policy. (https://library.imtlucca.it/sites/default/files/regolamento-policy-open-access-imtlib_0.pdf).

Si prega di consultare le pagine informative dell'editore relative alle politiche di autoarchiviazione.

(Article begins on next page)

05 October 2026

arXiv:2407.01057v1 [econ.GN] 1 Jul 2024

AI-powered Chatbots: Effective Communication Styles for Sustainable Development Goals

Ennio Bilancini[1], Leonardo Boncinelli[2], and Eugenio Vicario[2]

[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

July 2, 2024

Abstract

This paper presents an analysis of two pre-registered experimental studies examining the impact of 'Motivational Interviewing' and 'Directing Style' on discussions about Sustainable Development Goals. To evaluate the effectiveness of these communication styles in enhancing awareness and motivating action toward the Sustainable Development Goals, 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

1 Introduction

Understanding effective communication strategies is crucial in promoting the Sustainable Development Goals (SDGs). 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. The studies, preregistered on OSF (Bilancini et al., 2023a,b), investigate how MI and DS influence participants' engagement and behavioral responses toward SDGs. 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 2,030 instances in 2020, 2,872 in 2021, 3,527 in 2022, and 4,384 in 2023.

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). 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 power-

ful 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. 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 (Horton, 2023; Aher et al., 2022), social dilemmas (Guo, 2023; Capraro et al., 2023), as well as in voting decisions (Argyle et al., 2023) and the formation of moral judgments (Dillion et al., 2023). The second strand of the literature focuses on human-agent interactions in conversational settings (Stein et al., 2017; Numata et al., 2020). Studies here examine the impact of the chatbot's communication style, focusing on the effectiveness of chatbots using motivational interviewing techniques (Da Silva et al., 2018) in healthcare (Shingleton and Palfai, 2016), 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).

The results of this work contribute to enriching the policy-maker's toolbox, pro-

viding an additional intervention tool that exploits recent advancements in the field of artificial intelligence. AI-powered chatbots can combine with existing policy interventions, activating synergies to promote behavioral change (Alt et al., 2024).

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'.

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.

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.

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*

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.

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.*

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.

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.

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

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.

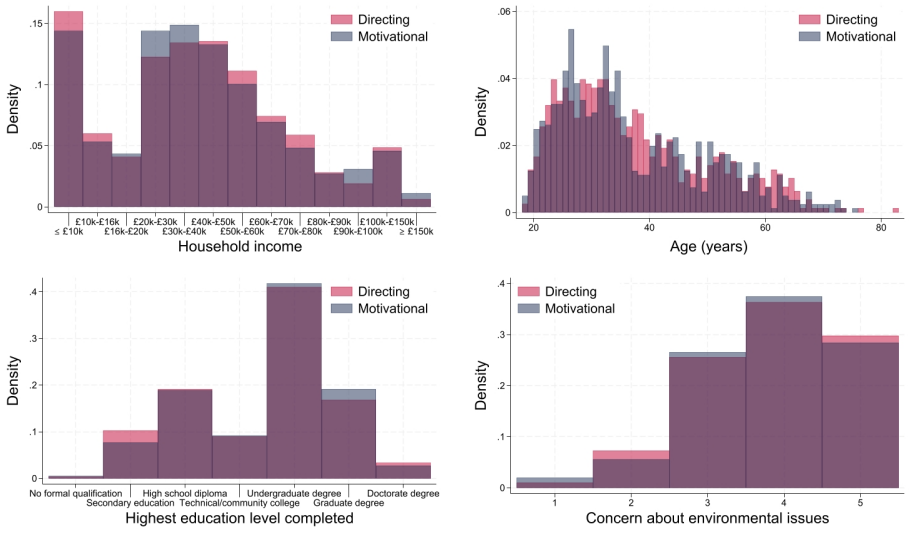

Figure 1: Densities are displayed regarding socioeconomic characteristic under MI and DS: On top left the density of *Income*, on top right the density of *Age* in years of participants, on bottom left *Education*, and on bottom right *Concern about environmental issues*; in all cases, overlaid histograms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

3 Main results

The first study did not yield statistically significant findings for the pre-registered outcome variables. Building on the observed differences between experimental conditions,

we formulated more focused hypotheses for the second study. Here, we present the results for the pre-registered outcomes in the second study, utilizing pooled data from both studies. Importantly, both studies used an identical experimental design. The supplementary material provides a breakdown of the results for each study.

3.1 Self-assessment of interest

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

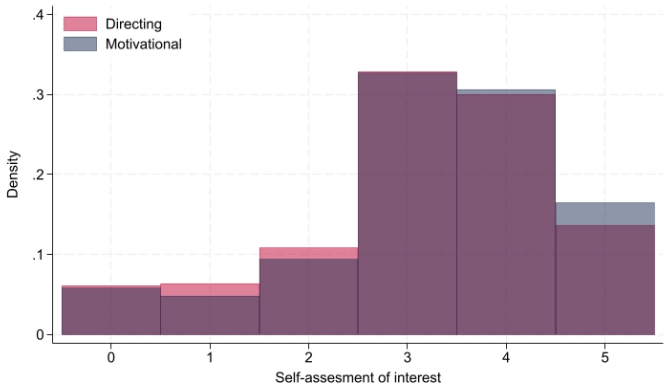

Figure 2: Densities of *Self-assessment of interest* are displayed for MI and DS treatments, with overlaid histograms to highlight differences.

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.

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		

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 10 in Appendix A for an analysis based on ordered probit regressions.

3.2 Self-assessment of learning

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

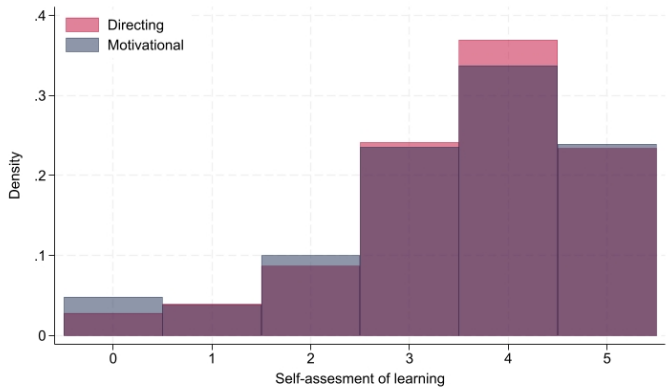

Figure 3: Densities of *Self-assessment of learning* are displayed for MI and DS treatments, with overlaid histograms to highlight differences.

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 can not 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. See the discussion in Section 4 and Table 9 in Appendix A for an analysis based on an instrumental variable regression.

3.3 Engagement

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

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

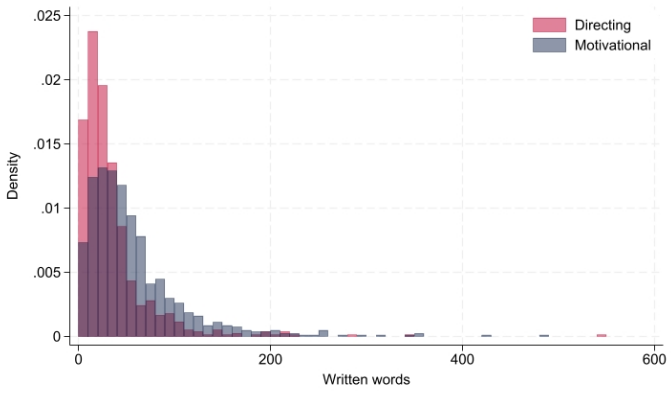

Figure 4: Densities of *Engagement* are displayed for MI and DS treatments, with overlaid histograms to highlight differences.

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

Results are reported in table 6.

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	

The null hypothesis can be rejected at any standard level of significance. Thus, we find evidence that *Self-assessment of interest* is (stochastically) larger under MI than under DS. See Table 10 in Appendix A for an analysis based on ordered probit regressions.

4 Exploratory analysis

In this section we conduct further exploratory analyses that were not preregistered. More specifically, we consider different measures of engagement, we look at the effects of the experimental conditions on *Willingness to receive costly information* and *Self-assessment of satisfaction*, and we elaborate on the impact of DS on learning.

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

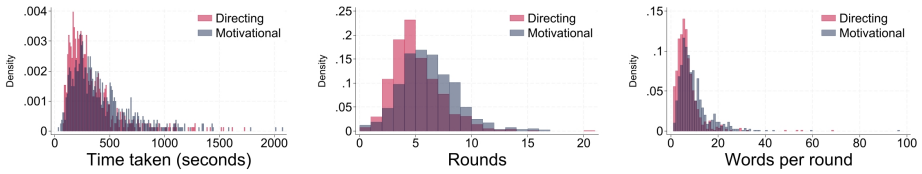

Figure 5: Densities are displayed regarding alternative measures of engagement under MI and DS: from left to right, the density of *Time taken* in seconds, the density of *Rounds* written by the user, and the density of *Words per round* by the user; in all cases, overlaid histograms are aimed to highlight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

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.

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*.

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

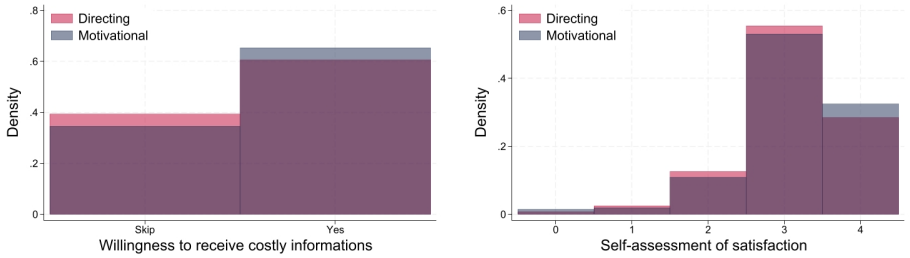

Figure 6: Densities are displayed regarding the variables measured through the final survey, excluded from the primary analysis, under MI and DS: on the left, the density of *Willingness to receive costly information* and, on the right, the density of *Self-assessment of satisfaction*; in all cases,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

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 5, 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

Table 8: Columns (1)-(3) report the results of probit regressions with *Willingness to receive costly information* as the dependent variable; Columns (4)-(6) report the results of ordered probit regressions with *Self-assessment of satisfaction* as the dependent variable.

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

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-assessment of interest*, at any level of significance. 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. To avoid endogeneity issues with the covariates *Time Taken* and *Self-assessment of interest*, an Extended Ordered Probit regression is employed. The total number of words written and the mean number of words per round are used as instruments for *Time Taken* and *Self-assessment of interest*, respectively. The results of the Extended Ordered Probit regression are presented in Appendix A. This analysis suggests that, in line with our expectation, DS does have a positive impact on learning with respect to MI, which is however detected only after controlling for *Time Taken* and *Self-assessment of interest*.

5 Conclusions

This research investigated the effectiveness of communication styles employed by a chatbot designed to conversate with users about sustainable development goals (SDGs). All outcome variables are self-reported measures in a brief questionnaire at the end of the conversation. 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.

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.

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, 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.

Aknowledgments

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'.

Declarations

E.B., L.B., and E.V. designed research, performed research, and wrote the paper.

Declaration of generative AI and AI-assisted technologies in the writing process: During the preparation of this work the authors used ChatGPT for grammar checking. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Competing Interests: The authors declare no competing interests.

Ethical approval: Ethical Committee of the University of Florence, approval number: 288/2023.

References

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.

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.

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

Bilancini, E., Boncinelli, L., Vicario, E., 2023a. Green chatbot. URL: osf.io/brhzy, doi:10.17605/OSF.IO/BRHZY.

Bilancini, E., Boncinelli, L., Vicario, E., 2023b. Green chatbot 2.0. URL: osf.io/e4bzp, doi:10.17605/OSF.IO/E4BZP.

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.

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.

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

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.

Da Silva, J., Kavanagh, D., 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, Article–number.

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.

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

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.

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.

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

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.

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.

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.

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.

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.

Miller, W., Rollnick, S., 1991. Motivational Interviewing: Preparing People to Change Addictive Behavior. Guilford Publications. URL: https://books.google.it/books?id=h16_QgAACAAJ.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

A Appendix

Table 9: The subtables report the results of an ordered probit regression having endogenous covariates and with *Self-assessment of learning* as dependent variable.

Principal regression

Learn	Ordered probit
Treatment	0.0685**
	(0.0324)
Concern env.	-0.0248
	(.0175)
Age	-0.0115
	(0.00751)
Age^2	0.000130
	(0.00009)
Education	-0.0176
	(0.0114)
Income	0.000792
	(0.00468)
Gender	-0.016
	(0.0284)
Time taken	0.0000442
	(0.000147)
Interest	
1	-0.2
	(0.14)
2	-0.347**
	(0.168)
3	-0.513**
	(0.245)
4	-0.953***
	(0.32)
5	-1.412***
	(0.423)
/cut1	-2.4
/cut2	-2.095
/cut3	-1.725
/cut4	-1.216
/cut5	-0.59

IV regression

Interest	Ordered probit
Words per round	0.0116***
	(0.00445)
/cut1	-1.477
/cut2	-1.127
/cut3	-0.705
/cut4	0.189
/cut5	1.115

IV regression

Time taken	OLS
Words user	3.993***
	(0.289)
Words per round	-5.93***
	(2.158)
Const.	235.0713***
	(10.617)

Correlations

corr(e.Interest,e.Learn)	0.883***
	(0.0483)
corr(e.Time taken,e.Learn)	0.134***
	(0.0422)
corr(e.Time taken,e.Interest))	0.168***
	(0.0277)

Obs.	1536

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

Table 10: In columns (1)-(3), the results of ordered probit with *Self-assessment of interest* as the dependent variable; In columns (4)-(6), the results of OLS with *Engagement* as the dependent variable.

Dep. Var.	Interest			Engagement		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.0965*	-0.103**	-0.108**	-22.54***	-22.57***	-22.47***
	(0.0524)	(0.0525)	(0.0528)	(2.374)	(2.371)	(2.397)
Concern env.		0.259***	0.252***		1.537	1.293
		(0.0307)	(0.0311)		(1.351)	(1.383)
Age			0.0325**			-0.744
			(0.0129)			(0.645)
Age^2			-0.000362**			0.0125
			(0.000154)			(0.00804)
Education			0.0130			0.964
			(0.0210)			(0.941)
Income			-0.0125			-0.152
			(0.00888)			(0.413)
Gender			0.125**			-4.463*
			(0.0524)			(2.336)
/cut1	-1.606***	-0.659***	-0.000632			
	(0.0578)	(0.125)	(0.269)			
/cut2	-1.245***	-0.286**	0.374			
	(0.0494)	(0.123)	(0.269)			
/cut3	-0.830***	0.143	0.807***			
	(0.0441)	(0.123)	(0.269)			
/cut4	0.0665	1.066***	1.737***			
	(0.0408)	(0.126)	(0.272)			
/cut5	0.985***	2.012***	2.688***			
	(0.0461)	(0.132)	(0.276)			
Constant				56.79***	50.87***	60.06***
				(1.897)	(5.664)	(13.73)
Observations	1,588	1,588	1,570	1,588	1,588	1,570
R2				0.053	0.054	0.063
Pseudo R2	0.0007	0.0182	0.0212			

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