Delivering Green Persuasion Strategies with a Conversational Agent: a Pilot Study

Mathyas Giudici
Politecnico di Milano
mathyas.giudici@polimi.it

Giulio Antonio Abbo
IDLab-AIRO - Ghent University - imec
giulioantonio.abbo@ugent.be

Pietro Crovari
Politecnico di Milano
pietro.crovari@polimi.it

Franca Garzotto
Politecnico di Milano
franca.garzotto@polimi.it

Abstract

Climate change is undeniable. The drastic consequences it may have on our lives make a collective effort crucial. Our research explores how Conversational Agents (CAs) can persuade people into environmentally sustainable behaviors, particularly in domestic spaces where these technologies are becoming increasingly popular. In this research work, we conducted an empirical evaluation (N=29) exploring the effectiveness and stance towards the adoption of different persuasive strategies compared to a CA delivering messages referring to just one persuasion strategy. Furthermore, this contribution reports on a custom dialogue manager's implementation, designed to enable the execution of the experiment. Although study results suggested no significant difference in persuasion effectiveness and usability of the conversational agents, participants reported a significant difference in the perceptions of parasocial interactions and dialogue with the CA, preferring the one delivering multiple persuasive strategies.

Keywords: chatbot, environmental sustainability, persuasion, persuasive technology, empirical evaluation

1. Introduction

The Intergovernmental Panel on Climate Change argues that climate change is anthropogenic. The urgent need for a green turnaround has been highlighted; everyone must realize that change requires a collective effort (Masson-Delmotte et al., 2021). Within this discourse, DiSalvo et al. (2010) discussed digital technology implications and the role of Human-Computer Interaction as a facilitator or an interfering factor in achieving sustainable goals. Indeed, technology can convey users' attention and foster – or persuade – them to implement more sustainable behaviors in their lives. To be effective, a *persuasive* technology must be able to intercept users' behavior and suggest improvements accordingly (Oh et al., 2021).

A number of psychological studies pointed out that the process of change is made of different phases and requires multiple persuasive techniques (Davidson, 1992). Such theoretical knowledge informs a design principle for persuasive technology mentioned, for example, in Oinas-Kukkonen and Harjumaa (2018) and applied in the health domain (Dillard & Shen, 2013). Still, the application of multiple strategies in digital persuasion actions exploits graphical-based interfaces (GUI) and no conversational interaction (Oinas-Kukkonen & Harjumaa, 2018). To the best of our knowledge, the unique case involving a conversational interaction paradigm is Beheshtian et al. (2020), which promoted environmental behaviors in a shared living space using a social robot that can sustain a speech-based dialogue with the user, enhancing communication by means of body expressions language and visual elements on an embodied display.

In our research, we explore the use of multiple persuasion strategies using text-based natural language interaction only, therefore eliminating the potentially confounding factors associated with different media that may affect the evaluation results. We have created a chatbot named *Ecobot* that handles textual conversations based on different persuasion strategies. We then evaluate the *effectiveness* and *likability* of *Ecobot* in two different experimental conditions: i) single strategy using the *feedback strategy* only; ii) *combined strategies* - when the chatbot uses a combination of three strategies - i.e., feedback, social comparison, and goal-setting.

In particular, using the *Ecobot* system, we investigate the impact of multiple persuasive strategies promoting sustainable energy consumption on users' attitudes and behaviors toward sustainable energy consumption, assessing both the users' perception and

URI: https://hdl.handle.net/10125/106476
978-0-9981331-7-1

the parasocial interactions. We conducted an empirical evaluation involving 29 participants who interacted with two different versions of the chatbot according to the two conditions of our between-subject design.

The results did not highlight any significant difference in persuasion effectiveness or usability between the two versions of the agent. However, participants reported a significant difference in their perception of parasocial interactions and dialogue, preferring the version that provided multiple persuasive strategies. These preliminary findings highlight the potential of multiple persuasive strategies in engaging and persuading users to adopt more environmentally sustainable practices in text-based systems.

2. Background

Persuasion is the modification of attitudes or behaviors through messages or chats (Conger, 1998) to achieve a final goal. Persuasion can be achieved through spontaneous or controlled procedures (Crano & Prislin, 2006). Spontaneous processes occur when individuals are self-aware and highly motivated for change, while controlled processes involve techniques that influence actions to reach a goal (Crano & Prislin, 2006). In controlled processes, the Theory of Planned Behavior (TPB) by Ajzen (1991) provides a framework for understanding and anticipating individual behaviors in specific contexts. TPB states that intentions and perceived behavioral control, which persuasive strategies can influence, determine behavior. Davidson (1992) reported that the process of change can be divided into different phases, each of which requires different persuasive techniques to be delivered. Such a theoretical concept is widespread and used in the creation of persuasive health campaigns (Dillard & Shen, 2013). For example, Michie et al. (2013) proposed a taxonomy of behavior change techniques that categorizes persuasive strategies into 16 macro areas. This taxonomy offers a systematic guideline with labels, definitions, and examples (health related) of each persuasive strategy.

Psychological theories (like TPB) are the basis of *Persuasive Technologies* (Fogg, 2002), interactive technologies that aim to change attitudes or behaviors through social influence and persuasive techniques. Effective persuasion in digital technologies often involves personalized messages and adaptive strategies. Machine learning algorithms are being explored to customize messages based on user data (Carfora et al., 2020). Still, Oinas-Kukkonen and Harjumaa (2018) presented a framework for designing and evaluating persuasive information technology systems, pinpointing that user persuasion is a multi-phased and complex task, where different factors and strategies need to be applied through the entire journey. In their work, the design principles are exemplified over a traditional GUI with multimodal media (such as text, images, and sound). Effectiveness in the health field has been demonstrated in real use cases like influencing positive eating behaviors by avoiding snacking (Kaptein et al., 2012) and collecting environmental data to assist the elderly (Yared & Abdulrazak, 2016)

Coming to the domain of environmental sustainability, instead, the research landscape is fragmented, and there is no consensus on the most effective persuasion strategies to increase sustainable behavior (Midden et al., 2008). For example, Costanza et al. (2012) proposed real-time visualization of energy consumption to lead to responsible consumption over time. Yun et al. (2013) intervened in workspaces, while Giudici et al. (2023) with gamified activities promote sustainable behaviors.

Conversational Agents – applications designed to interact with users using natural language (Hussain et al., 2019) – are considered a promising technology in sustainability domains (Giudici et al., 2022; Hussain et al., 2019) since they are usually embedded in widespread physical devices in home environments (e.g., Google Home, Alexa) (Sciuto et al., 2018) or people's phones (Jaber & McMillan, 2020), delivering energy feedback (Gnewuch et al., 2018) or suggesting sustainable mobility (Diederich et al., 2019). Many studies tried to understand how to effectively embed persuasion techniques in rule-based conversational agents (e.g. Cacanindin, 2020; Gnewuch et al., 2018). In recent years, with the development of generative conversational agents and Large Language Models (Vaswani et al., 2017), many studies have tried to understand how to embed persuasion in artificially generated conversations. Gunawardane et al. (2019) and JO (2023) are some major examples.

Finally, notable to be mentioned is the work of Beheshtian et al. (2020), who created a social robot to persuade people in shared living spaces towards sustainability, employing multiple strategies like feedback, rewards, and social comparison. To the best of our knowledge, no study focused on how to dynamically integrate multiple strategies in a single-modal text-based interface, as *Ecobot* has been designed to.

3. The *Ecobot* System

Evaluating different communication approaches required the design of *Ecobot*, a custom dialogue manager. Indeed, the existing types of dialogue management – based on finite state machines, frame-based, or generative – created issues during the setup of the experiment. The first approach, using finite state machines, struggles with non-linear conversations: these solutions are limited by the rigidity that the modelization imposes on the conversation. In this experiment, the conversation is not expected to follow a predefined path, rendering it difficult to model the dialogue as a predefined script. Frame-based approaches rely on a predefined set of information that must be elicited from the user and do not allow adopting different strategies. For this evaluation being able to control how the concepts are expressed is essential. Generative end-to-end dialogue engines based on corpora do not currently permit full control over the output.

Ecobot takes an approach that can appear similar to a state-based dialogue manager but instead avoids the definition of an interaction script entirely. Specifically, during the interaction with the user, each utterance is evaluated with an external Natural Language Understanding (NLU) engine to understand which is the most probable underlying intent. The intent is used to decide which action should be performed, following rules decided during the development phase. Differently from dialogue managers based on state machines, all the actions are available, and the algorithm can also access external data from a mock database to determine which one is the most fitting. Finally, the action will provide an answer that is returned to the user. Since the interaction sequence is not explicitly modeled, the system uses a data structure representing a context to achieve a sound conversation. Each action has access to the context, which contains the topics and contents of the conversation.

The benefit of this approach to dialogue management is that it can support use cases where the interlocutor decides to change the conversation's objective or change the topic entirely. In addition, the chatbot can adopt different tones and provide further information based on additional external data. The *Ecobot* is intended to be used to test distinct communication strategies in studies taking place in controlled environments. It is not meant to be used in more complex applications, as it would struggle when scaling to many different actions. In the application presented in this paper, the system is implemented as a backend accessible via a web application. It plays the role of a conversational agent in a smart home, with access to mocked data on energy consumption.

3.1. Architecture and Behaviour

The system represents each persuasion strategy as a *Strategy* instance, that has an identifier to represent its intent and a set of example utterances that are used to train an external NLU engine. In addition, each Strategy has a set of answers that can be displayed to the user and an action that implements the logic of that intent. For example, for the *Feedback* persuasive strategy, we trailed the NLU to find intents where the user asks for details about some appliances. Similarly, the utterances delivered by the chatbot describe the energy consumption of such appliances (e.g., *Today, you have consumed 1kWh for running your washing machine*), taking advantage of the logic to access appliance data.

The central component of the architecture – as represented in Figure 1 – is the *Dialogue Manager*, which receives the user's sentence from the *Interface* and uses the *NLU* to extract the corresponding intent from it. The intent and other information available to the Dialogue Manager are used to select the most fitting Strategy for each case. This choice is coded during the developing process. The Strategy's action is executed, and an answer is returned to the user.

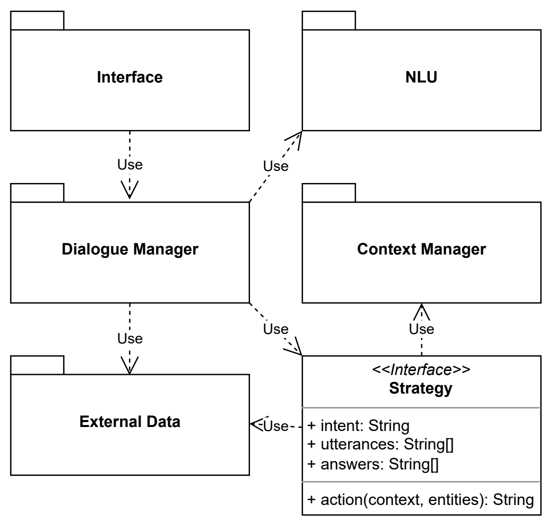

Figure 1: Architecture of the System

In order to perform its operations, the action has access to external data and to the *Context Manager*, which stores data about the current conversation in a simple dictionary structure. In addition, the action can be provided with entities extracted from the user's sentence. Each action implementation that needs additional data will first check the provided entities to evaluate whether the user gave additional information with the sentence. Then, if the information is not found, the action will make a request to the Context Manager to retrieve the information if it was added previously by another action. If even this fails, the system will ask the user for the missing information returning a question that prompts the user. To achieve this, it will save in the context that a question has been asked, and the execution will be continued in the conversation turn that follows. When the data is available, the action completes its operations and returns a response to the user. Thanks to *Ecobot*'s design, the conversation is not blocked if the user decides not to provide the requested data or to change the topic. To show the potential of this system, the conversation in Figure 2 illustrates an initial command from the user with an explicit entity. The system answers with a request to confirm, but the user changes the objective. The system is able to execute the request even though the entity is missing in the request because it was saved in the context.

U: Turn on the washing machine
- *The corresponding action saves in the context that the conversation is about the entity washing machine and that a question was asked*

A: Are you sure? You are consuming 10% more than your neighbors. What about setting a goal to reduce its usage?
U: Let's set an objective!
- *The user does not specify which objective: the corresponding action will look for an entity recently mentioned in the context*

A: I have set an objective for the washing machine.

Figure 2: Example of *Ecobot* Conversation

3.2. Implementation Details

The system is divided into a backend implemented in JavaScript, and a frontend in React. As an NLP engine, NLP.js was chosen for its performance and ease of use. Thanks to this library, the system can be easily configured to support multiple languages for each Strategy. All the software is being released as open source[1].

4. Empirical Study

We carried out a pilot experimental study involving 29 participants to test whether multiple intervention strategies provided by a conversational agent can potentially increase users' intentions to reduce residential energy consumption compared to an agent delivering only the feedback persuasive technique.

The final objective of the study was to investigate the following research questions about a persuasive chatbot for environmental sustainability:

[1]https://gitlab.com/i3lab/ecobot

R1 Is there a significant *effectiveness* difference between using multiple persuasive techniques and using only the feedback technique?

R2 Is there a significant *likability* difference between using multiple persuasive techniques and using only the feedback technique?

We set up two configurations of a web-based persuasive chatbot (i.e., *Ecobot*) and let users engage in conversations concerning energy-saving and environmental issues. The two configurations differed only in the number of strategies delivered, but for the common ones, the messages were the same. In addition, researchers put efforts into keeping the messages sent by the agents homogeneous (with the same tone and expression).

4.1. Research Variables

The experimental design was a between-subject design with *Condition* as the fixed factor. The persuasive strategies adopted in the two experimental conditions were designed according to the most used techniques in the field of persuasive technology for sustainability (Adaji & Adisa, 2022):

- *Combined:* the chatbot delivered three persuasive techniques:
  - *Feedback (FB)*: delivering information about individuals' energy consumption.
  - *Goal setting (GS)*: setting a specific and measurable objective to motivate individuals to minimize their domestic energy use.
  - *Social comparison (SC)*: providing individuals with information about how their behavior compares to other households' energy usage.
- *Feedback*: the chatbot delivered only one persuasion technique, providing different feedback on energy consumption.

The study involved questionnaire-based data collection. Participants responded to questions by giving a score on a 7-point Likert scale. Firstly, the participants' *self-efficacy*[2] was evaluated to determine their confidence in their ability to adopt sustainable practices. Additionally, the *action effect*[2] measure aimed to gauge the participants' perception of how effective their individual actions were in contributing to environmental sustainability. The

[2]Ad hoc questions reported in the Supplement Material (https://doi.org/10.5281/zenodo.8338926)

Future intentions[2], an 11-item ad hoc metric (α = .68), were assessed to understand participants' willingness to engage in sustainable behavior in the future. The *New-environmental paradigm (NEP)* (Dunlap & Van Liere, 1978) scale (α = .75) was used to capture participants' overall worldview and values regarding environmental sustainability. Furthermore, the *Parasocial Interaction (PSI) scale* (Tsai et al., 2021) (α = .84) measured the degree to which participants felt connected to and attached to *Ecobot*. Lastly, the *System Usability Scale (SUS)* (Bangor et al., 2009; Brooke et al., 1996) (α = .91) was employed to gather participants' perceptions of the usability of the chatbot interface. Finally, both the number of interactions with the chatbot and the duration (in seconds) of the interaction were tracked.

4.2. Participants

The study involved 29 subjects (7 females and 22 males) with a mean age of 27 years (range 22-58, M=27.2, SD=8.66). We also collected information about participants' educational backgrounds, such as the number of years enrolled in educational institutions (range 13-21, M=16.2, SD=2.05). A more detailed report on the sample distribution by age and schooling for each condition is represented in Table 1.

All study participants were recruited voluntarily without any financial compensation; they signed a consent form informing them about procedures, goals, and data treatment. Participants belonged to close contacts from the personal community, colleagues, or university students (the latter with a predominantly scientific background). They were all sensitive to environmental sustainability issues and had no past experience in interacting with chatbots able to sustain articulated conversations.

4.3. Procedure

Participants underwent one session in a randomly assigned condition (i.e., *Combined* or *Feedback*). Regardless of the condition, the user's activity during the test is organized in three phases. During the entire session, an observer was always present to take notes.

The first phase is the presentation of the study. Participants were asked to fill out general biographical information and to identify themselves in a hypothetical scenario[3] that depicted excessive electricity consumption in their houses. During the second phase, participants are invited to freely interact with *Ecobot*, talking about their energy consumption. In addition, subjects were asked to *think aloud*, enabling an observer to collect qualitative feedback during the experimentation. In both conditions, the interaction starts with the same message from the chatbot: *"In the last period, your air conditioner consumed 50 kWh. It seems too much, can you do something about it?"* The phase ended when the participant considered the conversation was over and considered they had obtained enough information on their consumption. Finally, in the last phase, participants filled out a questionnaire with all the inquiries to assess the research variables presented in Section 4.1.

4.4. Methodology

We computed scores and conducted two statistical analyses: one for descriptive information and another to explore correlations between scores using JAMOVI software[4]. We performed an independent samples t-test to identify statistical significance in the differences between the two experimental conditions, entitling us to frame better-defined conclusions. Finally, we performed multiple Pearson rank correlations between the variables observed. The Pearson correlation allows us to verify if there is a linear relationship between the variables.

5. Results

In the *Combined* condition, subjects reported positive future intentions toward sustainable behaviors with a mean of 4.49 and a standard deviation of 0.587. Participants declared an average self-efficacy in green domestic behaviors of 4.52 (SD=0.748), while their action effectiveness was attested to 5.19 (SD=1.72). The NEP test presented a mean value of 5.16 (SD=0.661), the PSI test has a mean of 4.18 (SD=0.358), and the SUS score was 82.8 (SD=11.7). Finally, in this condition, participants overall exchanged eighty messages (M=81.4, SD=33.7) for more than fifteen minutes (M=1048 s, SD=338 s).

In the *Feedback* condition, participants showed positive future intentions toward sustainable behaviors with an average value of 4.51 and a standard deviation of 0.621. Participants reported a mean of self-efficacy in green domestic behaviors of 5.02 (SD=1.04), while their action effectiveness was with an average value of 5.18 (SD=1.36). The NEP test presented a mean value of 4.95 (SD=0.881), the PSI test had a mean equal to 3.66 (SD=0.544), and the SUS score was 84.1 (SD=8.41). Finally, in this setting, participants exchanged approximately fifty messages (M=55.1, SD=32.2) for more than ten minutes (M=701 s, SD=325

[3] In the Supplement Material (https://doi.org/10.5281/zenodo.8338926)

[4] https://www.jamovi.org/

Table 1: Demographics of Participants for each Experimental Condition

Condition	Gender		Age Mean	Age SD	Schooling Mean	Schooling SD
	F	M				
Combined	5	9	27.6	8.81	15.9	2.57
Feedback	2	13	26.8	8.80	16.6	1.40

s). A more comprehensive and detailed view of the results is given in the supplementary materials.

Independent Samples T-Test. Considering the Parasocial Interaction scores, the 14 subjects who interacted with the *Combined* chatbot when compared to the 15 subjects in the *Feedback* group demonstrated a statistically significant difference, $t(27)=3.0684$, $p=0.002$ (Figure 3). In addition, Levene's test is significant ($p<0.05$), suggesting a violation also of the assumption of equal variances. We run Welch's Test, finding $t(24.4)=3.1120$, $p=0.002$. In the Interaction Time between the two conditions (Figure 4), results indicated a statistically significant difference, $t(27)=2.8162$, $p=0.004$, as well as a low significant effect for Interaction Number, $t(27)=2.1494$, $p=0.020$ (Figure 5). There were no significant differences for all the other evaluations performed in the two different conditions (as reported in Table 2).

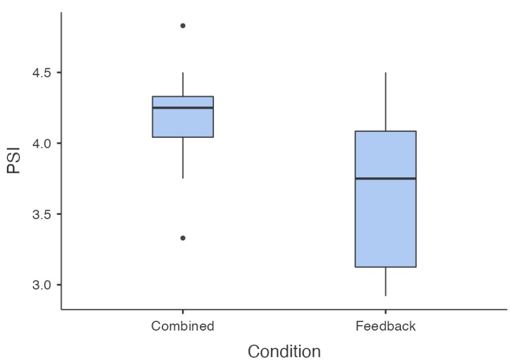

Figure 3: PSI Results Distribution

Correlations. Among all the responses collected by the participants, we found a statistically significant correlation between Interaction Time and Interaction Number. In the *Combined* group, such correlation was $r=0.854$, $p<0.001$, while in the *Feedback* was $r=0.916$, $p<0.001$. In addition, in the *Combined* condition, NEP and SUS were statistically correlated significantly with $r=0.554$, $p=0.040$. While in the *Feedback* condition, there were no additional correlations to be reported.

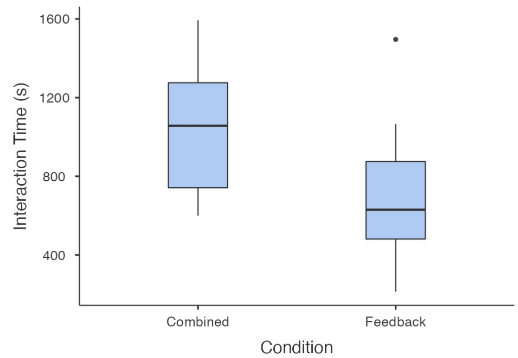

Figure 4: Interaction Time Results Distribution

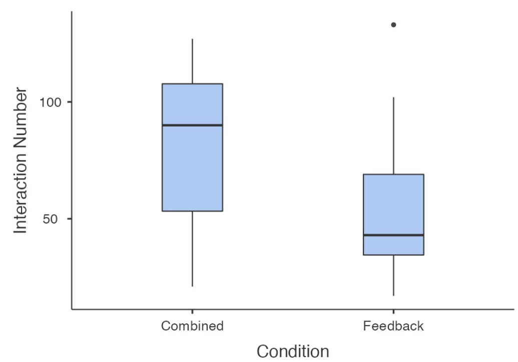

Figure 5: Interaction Number Results Distribution

Finally, looking at possible general correlations, without splitting the data among the two conditions, there was a low statistical correlation between Interaction Time and Self-Efficacy with $r=-0.386$, $p=0.039$.

6. Discussion

This section presents the discussion of the results of the pilot experimental study testing a persuasive conversational agent delivering persuasive intervention strategies to reduce consumption in smart home environments. The goal was to investigate which approach between delivering different intervention strategies and delivering just one intervention strategy was more effective and engaging for users.

Table 2: Independent Samples T-Test Results

	Statistic	df	p
Future Intentions	-0.0671	27.0	0.527
Self-Efficacy	-1.4741	27.0	0.924
Action Effect	0.0233	27.0	0.491
NEP	0.7103	27.0	0.242
PSI	3.0684[a]	27.0	0.002
SUS	-0.3564	27.0	0.638
Interaction Number	2.1494	27.0	0.020
Interaction Time (s)	2.8162	27.0	0.004

Note. $H_a : \mu_{Complete} > \mu_{Feedback}$
[a] Levene's test is significant ($p < .05$)

RQ1 – Effectiveness. Contrary to the initial expectations, no experimental condition of intervention significantly affected users' intentions to perform future sustainable actions. In addition, there is no relevant difference in scores of self-efficacy, action effect, and NEP. In particular, on average, NEP results in the *Combined* version of the chatbot are greater than the ones in the *Feedback*; those results are also comparable and close to the previously reported outcomes by Cruz and Manata (2020). It is reasonable to believe that a single session with the *Ecobot*, which took place in a relatively low ecological experimental setting, was not sufficient to redirect participants' intentions. In addition, it's possible that most participants already regularly engage in low-effort household environmentally sustainable behaviors, including energy conservation. Such daily actions are frequently encouraged through mass media communication stressing the significance of acting in a more sustainable manner. Another explanation can involve the influence of social desirability (Vesely & Klöckner, 2020), for which participants could feel pressured to meet expected social norms. Finally, it is worth reporting the low correlation between Interaction Time with *Ecobot* and users' Self-Efficacy. According to DeVille et al. (2021), the time spent in nature is linked to a better perception of the value of nature, subsequently, greater pro-environmental attitudes and behaviors. Therefore, given our results, we can make two assumptions: (i) users with lower self-efficacy in performing sustainable behaviors interacted more due to being intrigued by the suggested behaviors or motivated by a sense of responsibility towards their low level of green behavior; (ii) users with higher self-efficacy in acting sustainable behaviors interacted less because they were less interested and already aware of green virtuous behaviors. Further future investigations in this area may disclose these hypotheses' validity.

RQ2 – Likability. Our study's Parasocial Interaction scale results are comparable with the previously identified by Tsai et al. (2021), suggesting a positive and effective interaction between subjects and *Ecobot*. In addition, the results pointed out a significant difference in the mean scores of the PSI between the two experimental conditions. Subjects significantly perceived the interaction with the chatbot in the *Combined* condition as more pleasant than the one in the other (i.e., *Feedback*). Results in both conditions pointed out an average SUS score greater than eighty. According to Bangor et al. (2009), the *Ecobot* system is in the acceptable range, with a usability grade of B, giving us an adjective rating between good and excellent. Finally, there is a comparable average total number of both user-chatbot interactions (low significant effect, p=0.020) and total words generated by the bot (low significant effect, p=0.0863). Still, our results show that the users interacted significantly longer with *Ecobot* in its *Combined* version. Although preliminary, these findings seem to indicate that the users in the combined strategy condition interacted for a longer time since they might have liked this version of the chatbot more. Previous work by Cuadrado et al. (2022) has proven the existence of a link between the green attitude and the level of engagement a digital application can provide to users. Future experimentation in more ecological environments (i.e., real home automation context) will prove the real persuasive effectiveness of such a persuasive conversational agent.

Qualitative results. At the end of the entire study, qualitative results were manually extrapolated by reviewing comments and investigating possible patterns among different participants, using thematic analysis (Braun & Clarke, 2012). The patterns highlight particular flaws and strengths of *Ecobot*. Please note that the experiments were conducted in Italian language; the sentences reported in this section were translated into English. The scenario and task provided to participants were very broad, without a specific focus to avoid adding any possible bias. Some users asked very specific questions or tips (e.g., *"If I have to wash a hoodie made with cotton, which program should I choose?"*) that were not included in the *Ecobot* capabilities. This data is very useful for identifying the most unsatisfied requests and adding them for future experimentation in a more ecological setting (i.e., real home automation environment). The majority of the users activated in the early interactions the *help* intent to understand which were the chatbot capabilities. In the future, this intent could be activated automatically during the first interaction with the chatbot as an opening message to the conversation (Moore & Arar, 2019). However, this feature was not implemented to avoid adding bias in the experimentation and to let them freely explore and discover the various *Ecobot* capabilities. The users who used the *Combined* version of the *Ecobot* appeared generally more satisfied with the conversation they took part in than the ones who used the *Feedback* version, as also depicted by quantitative results. This result likely depends on the *Ecobot Combined* version's ability to deal with a greater set of requests, making users understand its purpose and possible daily usage. The context awareness function (Section 3.1) and the ability to completely change the conversation from one message to another were extremely appreciated and heavily used by participants. This appreciation is aligned with previous works in chatbots used in other fields, for example, in the tourism domain (Clarizia et al., 2019) or in the product configuration domain (Gupta et al., 2019). Among the different persuasion strategies implemented by the *Ecobot*, *feedback* was the most delivered, followed by *goal setting* and *social comparison*. This result may depend on the fact that feedback is one of the best-known and most widely used persuasive techniques (even outside the context of environmental sustainability). Some participants highlighted the need for a more user-friendly unit of measurement since average consumption (in terms of kilowatt hours) is difficult to understand (i.e., they cannot understand if it's too much or not). They suggested providing the cost in terms of money (or bill) and comparing their consumption with other users using a percentage (Petkov et al., 2012). Finally, several participants asked the chatbot to provide information on the best time slots to turn on appliances and to retrieve information about consumption of the previous week (or month, or year, etc.). These considerations are aligned with previous results (Costanza et al., 2012).

7. Limitations

Our study presents limitations. First, a sample size bigger than 29 participants (G*Power suggests N=70), more balanced in age and gender, is needed to comprehensively evaluate the conversational agent persuasion toward more environmentally sustainable behaviors. In addition, mixed models can be applied to analyzing complex data and accounting for fixed and random effects. In addition, the study presented in this paper was conducted in a laboratory, using a hypothetical scenario, significantly impacting the ecological validity of the experience. In fact, even though participants interacted with a working conversational agent, the appliances' data was mocked-up, and users' actions did not affect real devices. For all the above reasons, the results of our study are preliminary and insufficient to make definitive claims on the persuasive effectiveness of the *Ecobot*. However, we provided new insight into this emerging topic by addressing our research questions.

8. Conclusion and Future Works

We reported the results of an empirical evaluation involving 29 subjects who interacted with two different versions of *Ecobot*, a dialogue manager system implementing persuasive strategies to push users into more sustainable behaviors in the domestic context. A version of *Ecobot* delivered messages related to the *feedback* persuasion strategy. The other version was *combined* with social comparison and goal-setting strategies. Our investigation pointed out no significant difference in persuasion effectiveness or usability between the two versions of the agent. Conversely, participants reported a significant difference in their opinions of dialogue and parasocial interactions with the agent, preferring the version presenting multiple persuasive methods. These findings have significant implications for designing and implementing future Persuasive Conversational Agents to promote sustainable behaviors, highlighting the potential of combined persuasive strategies in engaging and persuading users to adopt more environmentally sustainable practices. Looking at the future, we have already designed and organized a new study to address the limitations of the exploratory study reported in this paper and to validate and extend the findings. The new study will involve a much wider number of participants, balanced in terms of gender and age. It will be designed as a controlled between-subjects study composed of seven separate experiments designed with a similar experimental protocol as the exploratory study to investigate the effectiveness of each persuasive strategy and all their possible combinations (i.e., 7 conditions - FB, SC, GS, FB + GS, FB + SC, SC + GS, FB + GS + SC). The study will take place in a real home environment enhanced with IoT-based domotic devices integrated with the chatbot to assess in an ecological context the agent's effectiveness in influencing real-life behaviors, supporting the ecological validity of the research. In addition, we are already incorporating a speech-based modality in *Ecobot*, which could make the chatbot more user-friendly and more accessible. We are planning a second future study, in which we will explore the effectiveness of different (combinations of) persuasion strategies based on the speech-based version of *Ecobot* and will compare the results among the two different interaction modalities in all the different persuasive strategy conditions.

Acknowledgments

The authors would like to thank all the participants involved in the study, Simona Sacchi and Giulio Faccenda (Psychological Department - Università degli Studi di Milano-Bicocca, Italy) for their support throughout the project, and Francesco Milano for his valuable work in the development of the solution. This work was supported by the Italian Ministry of University and Research (MUR) and the European Union (EU) under the PON/REACT project.

References

Adaji, I., & Adisa, M. (2022). A review of the use of persuasive technologies to influence sustainable behaviour, 317–325. https://doi.org/10.1145/3511047.3537653

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2).

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, *4*(3), 114–123.

Beheshtian, N., Moradi, S., Ahtinen, A., Väänanen, K., Kähkonen, K., & Laine, M. (2020). Greenlife: A persuasive social robot to enhance the sustainable behavior in shared living spaces. *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 1–12.

Braun, V., & Clarke, V. (2012). *Thematic analysis.* American Psychological Association.

Brooke, J., et al. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*.

Cacanindin, N. M. (2020). Greening Food Consumption Using Chatbots as Behavioral Change Agent. *Journal of Advanced Research in Dynamical and Control Systems*, *12*(01-Special Issue), 204–211. https://doi.org/10.5373/JARDCS/V12SP1/20201064

Carfora, V., Di Massimo, F., Rastelli, R., Catellani, P., & Piastra, M. (2020). Dialogue management in conversational agents through psychology of persuasion and machine learning. *Multimedia Tools and Applications*, *79*(47).

Clarizia, F., Colace, F., De Santo, M., Lombardi, M., Pascale, F., & Santaniello, D. (2019). A context-aware chatbot for tourist destinations. *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, 348–354.

Conger, J. A. (1998). The necessary art of persuasion. *Harvard Business Review*, *76*, 84–97.

Costanza, E., Ramchurn, S. D., & Jennings, N. R. (2012). Understanding domestic energy consumption through interactive visualisation: A field study. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. https://doi.org/10.1145/2370216.2370251

Crano, W. D., & Prislin, R. (2006). Attitudes and persuasion. *Annual review of psychology*, *57*.

Cruz, S. M., & Manata, B. (2020). Measurement of environmental concern: A review and analysis. *Frontiers in Psychology*, *11*, 363.

Cuadrado, E., Macias-Zambrano, L., Carpio, A. J., & Tabernero, C. (2022). The moderating effect of collective efficacy on the relationship between environmental values and ecological behaviors. *Environment, Development and Sustainability*.

Davidson, R. (1992). Prochaska and diclemente's model of change: A case study?

DeVille, N. V., Tomasso, L. P., Stoddard, O. P., Wilt, G. E., Horton, T. H., Wolf, K. L., Brymer, E., Kahn Jr, P. H., & James, P. (2021). Time spent in nature is associated with increased pro-environmental attitudes and behaviors. *International journal of environmental research and public health.*

Diederich, S., Lichtenberg, S., Brendel, A. B., & Trang, S. (2019). Promoting sustainable mobility beliefs with persuasive and anthropomorphic design: Insights from an experiment with a conversational agent.

Dillard, J. P., & Shen, L. (2013). *The sage handbook of persuasion*. Sage.

DiSalvo, C., Sengers, P., & Brynjarsdóttir, H. (2010). Mapping the landscape of sustainable hci. *Proceedings of the SIGCHI conference on human factors in computing systems*.

Dunlap, R. E., & Van Liere, K. D. (1978). The "new environmental paradigm". *The journal of environmental education*, *9*(4).

Fogg, B. J. (2002). Persuasive technology: Using computers to change what we think and do. *Ubiquity*, *2002*(December). https://doi.org/10.1145/764008.763957

Giudici, M., Crovari, P., & Garzotto, F. (2022). Candy: A framework to design conversational agents for domestic sustainability. *4th Conference on*

Conversational User Interfaces. https://doi.org/10.1145/3543829.3544515

Giudici, M., Crovari, P., & Garzotto, F. (2023). Leafy: Enhancing home energy efficiency through gamified experience with a conversational smart mirror. *Proceedings of the 2023 ACM Conference on Information Technology for Social Good*, 128–134. https://doi.org/10.1145/3582515.3609526

Gnewuch, U., Morana, S., Heckmann, C., & Maedche, A. (2018). Designing conversational agents for energy feedback. *International Conference on Design Science Research in Information Systems and Technology*.

Gunawardane, M., Pushpakumara, H., Navarathne, E., Lokuliyana, S., Kelaniyage, K., & Gamage, N. (2019). Zero food waste: Food wastage sustaining mobile application. *2019 International Conference on Advancements in Computing (ICAC)*, 129–132.

Gupta, A., Zhang, P., Lalwani, G., & Diab, M. (2019). Casa-nlu: Context-aware self-attentive natural language understanding for task-oriented chatbots. *arXiv preprint arXiv:1909.08705*.

Hussain, S., Ameri Sianaki, O., & Ababneh, N. (2019). A survey on conversational agents/chatbots classification and design techniques. *Workshops of the International Conference on Advanced Information Networking and Applications*, 946–956.

Jaber, R., & McMillan, D. (2020). Conversational user interfaces on mobile devices: Survey. *Proceedings of the 2nd Conference on Conversational User Interfaces*. https://doi.org/10.1145/3405755.3406130

JO, A. (2023). The promise and peril of generative ai. *Nature*, *614*.

Kaptein, M., De Ruyter, B., Markopoulos, P., & Aarts, E. (2012). Adaptive persuasive systems: A study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *2*(2).

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., et al. (2021). Climate change 2021: The physical science basis. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*, *2*.

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine*.

Midden, C., McCalley, T., Ham, J., & Zaalberg, R. (2008). Using persuasive technology to encourage sustainable behavior. *Sustainability WS Pervasive*, *113*.

Moore, R. J., & Arar, R. (2019). *Conversational ux design: A practitioner's guide to the natural conversation framework*. Association for Computing Machinery.

Oh, Y. J., Zhang, J., Fang, M.-L., & Fukuoka, Y. (2021). A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. *International Journal of Behavioral Nutrition and Physical Activity*, *18*.

Oinas-Kukkonen, H., & Harjumaa, M. (2018). Key issues, process model and system features. *Routledge handbook of policy design*.

Petkov, P., Goswami, S., Köbler, F., & Krcmar, H. (2012). Personalised eco-feedback as a design technique for motivating energy saving behaviour at home. *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, 587–596. https://doi.org/10.1145/2399016.2399106

Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018). " hey alexa, what's up?" a mixed-methods studies of in-home conversational agent usage. *Proceedings of the 2018 designing interactive systems conference*.

Tsai, W.-H. S., Liu, Y., & Chuan, C.-H. (2021). How chatbots' social presence communication enhances consumer engagement: The mediating role of parasocial interaction and dialogue. *Journal of Research in Interactive Marketing*, *15*(3), 460–482.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need.

Vesely, S., & Klöckner, C. A. (2020). Social desirability in environmental psychology research: Three meta-analyses. *Frontiers in Psychology*.

Yared, R., & Abdulrazak, B. (2016). Ambient technology to assist elderly people in indoor risks. *Computers*, *5*(4).

Yun, R., Scupelli, P., Aziz, A., & Loftness, V. (2013). Sustainability in the workplace: Nine intervention techniques for behavior change. *International conference on persuasive technology*.