Multimedia Tools and Applications
https://doi.org/10.1007/s11042-020-09178-w

Dialogue management in conversational agents through psychology of persuasion and machine learning

Valentina Carfora[2] **· Francesca Di Massimo**[1] **· Rebecca Rastelli**[1] **· Patrizia Catellani**[2] **· Marco Piastra**[1]

Received: 20 January 2020 / Revised: 8 May 2020 / Accepted: 4 June 2020 /
Published online: 27 June 2020

Abstract

To be really effective, conversational agents must integrate well with the characteristics of the humans with whom they interact. This exploratory study focuses on a method for integrating well-assessed methods from the field of social psychology in the design of task-oriented conversational agents in which the dialogue management module is developed through machine learning. In particular, the aim is to achieve agents whose policies could take into account the psychological features of the human interactants to deliver personalized and more effective messages. The paper presents the psychological study performed and outlines the overall theoretical architecture of the software framework proposed. On the psychosocial side, we first assessed the effectiveness of differently framed messages aimed to reducing red meat consumption taking the Theory of Planned Behavior (TPB) as the psychosocial model of reference. Turning to the machine learning field, the resulting Structural Equation Model (SEM) was first translated into a probabilistic predictor using Dynamic Bayesian Network (DBN). In turn, such DBN became the fundamental element of a Partially Observable Markov Decision Processes (POMDP) in a reinforcement learning setting. The possibility to elicit complete interaction policies was then studied by applying Neural Monte Carlo Tree Search (Neural MCTS) methods. The results thus obtained introduce the possibility to develop new multidisciplinary and integrated techniques for the development of automated dialogue managing systems.

Keywords Conversational agent · Theory of planned behavior · Psychology of persuasion · Machine learning · Reinforcement learning · Monte carlo tree search

Valentina Carfora and Francesca Di Massimo contributed equally to this work.

✉ Valentina Carfora
valentina.carfora@unicatt.it

Extended author information available on the last page of the article.

1 Introduction

In everyday life, human conversations are more effective when conducted by taking into account the specific characteristics of human beings involved. In the present study, we hypothesized that the design of automatic conversational agents would benefit from the contribution of social psychology, in an attempt to match the type of message being sent and the psycho-social profile of the recipient.

A typical, automatic conversational agent has a three-stage architecture, made up of: spoken language, written language and dialogue management (see [2]). We focused on dialogue management for task-oriented conversational agents. Currently, many *Reinforcement Learning* (RL, [45]) approaches to automating conversations involve end-to-end training of an agent from a set of recorded dialog slates (e.g. [35]). Achieving significant results in this way, however, requires considerable efforts in both collecting sample data and training experiments. The reinforcement learning process can be usefully improved by referring to established models developed by social and persuasion psychology. Such integration might accelerate the training phase and improve the quality of the dialogue achieved through the explicit representation of subtle psychological aspects of the interaction that would be difficult to elicit through end-to-end training alone. An innovative example in this direction is the work by [20], regarding the creation of a user-adaptive persuasive system that can be used in different areas to change behavior, consistent with people's needs and in the direction of health and well-being.

Along this line, in the present study we envisaged the possibility to enhance the reinforcement learning process by harnessing social psychology models related to predicting behavior and behavior change through recommendation messages.

As a first experimental step, we tested the effectiveness of different types of messages aimed to induce a reduction in red meat consumption. This test was performed by analyzing statistical data obtained through the interaction with several participating volunteers. To carry out the analysis, we used multi-group *Structural Equation Modeling* (SEMs, [32]).

In the second step, the resulting SEMs was translated into a *Dynamic Bayesian Network* (DBN, [19]) to obtain a probabilistic predictor trained with the same data collected during the experiments with volunteers. Such DBN formed the basis of a *Partially-Observable Markov Decision Process* (POMDP, [30]). The reinforcement learning algorithm used for finding suitable solutions of the POMDP thus obtained was a single-player version of the neural-net based *Monte Carlo Tree Search* (MCTS, [16]) method used in the *AlphaZero* approach to AI gaming [41]. In this way we elicited specific interaction policies apt to enable a conversational agent to: a) evaluating the psychological profile of the human interactant by asking a limited number of questions; b) identifying the type of message framing that might have the highest probability of inducing the behavior change desired.

The rest of the article is divided in two main parts. The first part describes the experimental assessment of the psychological model and the statistical model obtained. The second part describes the overall reinforcement learning model adopted and the results obtained.

2 Psychological model

Epidemiologic studies have linked high consumption of red/processed meat with the risk of developing various diseases, such as cancer [23] and type 2 diabetes [36]. Despite the increasing awareness of this risk, many individuals eat a quantity of red/processed that is much higher than what suggested by health authorities [25], namely, a maximum of three servings per week [3]. Reducing red/processed meat consumption would therefore be essential to promote health. However, so far there has been a remarkable lack of effective policies, initiatives or campaigns designed to tackle the demand for red/processed meat [7]. This is largely because campaigns aimed at convincing people to change eating habits face several psychological barriers, such as cultural norms [8] or taste preferences [18].

Recent psychological research has investigated how to apply communicative strategies to overcome the aforementioned psychosocial barriers and lead individuals to change their diet [5, 6, 10, 14, 43, 46]. To be persuasive, messages aimed at the reduction of excessive red/processed meat consumption must be formulated in a way that triggers involvement and reduces reactance, that is, receivers' resistance to persuasion. Some message contents can be more effective than others in encouraging receivers to process the text as personally relevant to them, and therefore activate a higher message involvement [34]. In turn, consistent with the principles of the Elaboration Likelihood Model [38], eliciting message involvement may lead individuals to behave in agreement with the recommendation presented in the message [1, 22, 39, 48].

Previous research suggests that exposure to persuasive messages focused on the salient outcomes of an expected behavior enhances receivers' involvement in the message. Along this line, previous research has shown [13] the effectiveness of messages formulated in prefactual (i.e. "If... then...") terms, namely, anticipating the hypothetical consequences of a given behavior. In that study, exposure to prefactual messages regarding the hypothetical consequences of excessive meat consumption led participants to change their attitude towards meat intake. In turn, attitude mediated the effects of messages on meat reduction, and message effects on attitude and behavior persisted one month after the end of the messaging intervention.

The goal of changing behavior through prefactual messages can be pursued using different message frames [17]. For example, the message can stress either the positive or the negative outcomes of the recommended action [40]. In a *positively framed message* the outcome of the action is presented with a positive valence (e.g. "If you eat well, you will improve your health"), whereas in a *negatively framed message* the outcome is presented with a negative valence (e.g. "If you eat badly, you will damage your health"). In the healthy eating domain, prior research showed that the effects of positively and negatively framed messages depends upon the recipient's characteristics, such as motivational orientation and baseline intentions [26].

Positively and negatively framed messages can be further differentiated, however [15, 21]. As regards positively framed messages, a further distinction can be made between messages describing a *gain* (e.g., "If you eat well, you will improve your health") and messages describing a *non-loss* (e.g. "If you eat well, you will avoid damaging your health"). Similarly, as regards negatively framed messages, a further distinction can be made between messages describing a *loss* (e.g., "If you eat badly, you will damage your health") and messages describing a *non-gain* (e.g., "If you eat badly, you will miss the opportunity to improve your health"). So far, research on the effects of persuasive messages on the reduction of red/processed meat consumption has largely ignored the distinction among these four different frames. Most research has focused on the simple gain-loss distinction, for example by labeling the non-loss and non-gain frames as negative frames altogether (see [9] for a review). A notable exception in the healthy eating communication domain is a study by Dijkstra and colleagues [21], who found that gain-framed messages were more effective than non-loss-framed messages in promoting fruit and vegetable intake.

In the present research we exposed participants to gain, non-loss, non-gain and loss messages regarding the reduction of red/processed meat consumption. We tested the effects on message involvement and intention to reduce red/processed meat consumption. We also controlled for the effects of some psychosocial variables that past research has shown to influence intention, regardless of the effect of any message frame. To do so, we referred to the Theory of Planned Behavior (TPB, [1]), according to which intention is the most proximal predictor of the execution of a behavior, and is in turn predicted by attitude towards the behavior, subjective norm and perceived behavioral control. *Attitude* represents the personal evaluation, in positive or negative terms, of performing a specific behavior. *Subjective norm* is the perceived social pressure in relation to performing the behavior. Finally, *perceived behavioral control* is the individual's perception of the internal and external resources possessed to perform the behavior.

Several studies confirmed that the TPB model has a strong predictive power in the case of people's intention to eat red/processed meat. Specifically, a positive attitude towards reduced red/processed meat consumption is the main predictor of the intention to reduce meat consumption [11, 12, 49], followed by perceived behavioral control and subjective norm [27].

Starting from the above, in the present research we tested the effects of differently framed messages (gain, non-loss, non-gain, and loss) on message involvement and intention to eat red/processed meat, controlling for the independent effects of attitude towards reduction, subjective norm, perceived behavioral control, and baseline intention.

2.1 The present study

Through a questionnaire, we first assessed participants' attitude, subjective norm and perceived control behavior regarding the reduction of red/processed meat consumption. We also assessed their baseline intention to eat red/processed meat. One week later, we randomly assigned participants to four different conditions, in which they read a series of messages presenting the outcomes of red/processed meat consumption in terms of gain, non-loss, non-gain or loss. We expected that exposure to different messages would have differential effects on participants' involvement in the message and intention towards red/processed meat consumption. More specifically, we expected exposure to differently framed messages to moderate the effect of message involvement on future intention to eat red/processed meat consumption. We did not make any specific predictions about the higher or lower effect of each type of message, however, given that this was the first attempt in the literature to analyze their differences in the case of the red/processed meat consumption. The psychosocial hypothesized model is illustrated in Fig. 1.

2.2 Method

2.2.1 Participants

A sample of Italian citizens was invited to participate in a university study on public communication. People who agreed to participate received an email with a link to an online questionnaire powered through Qualtrics (Time 1). One week later, participants were invited to read eight messages on the health consequences of eating red/processed meat. They were randomly assigned to four different conditions (gain, non-loss, non-gain, loss messages). After reading the messages, all participants were invited to fill in a second questionnaire (Time 2). The initial sample was made of N = 834 participants. After collecting data, we excluded participants who followed a specific diet (e.g., veganism, vegetarianism or restrictive diets) and participants who did not fully or accurately complete both questionnaires.

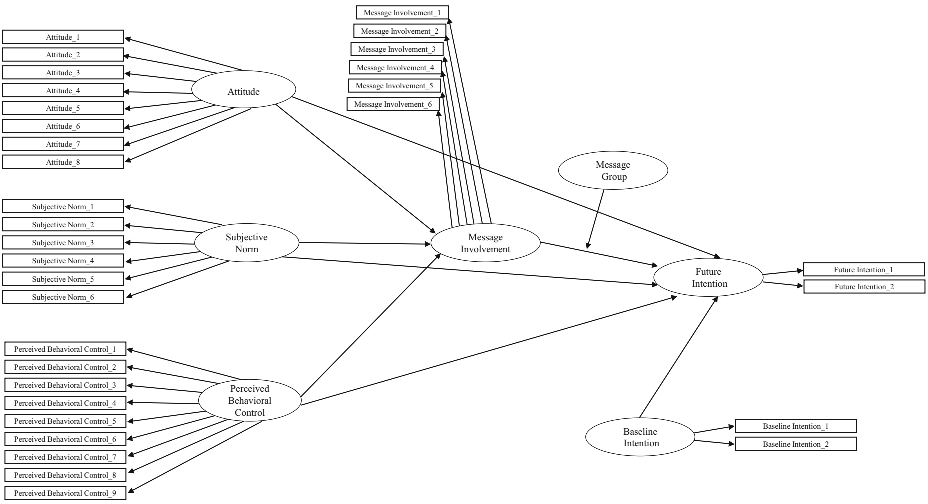

Fig. 1 The Structural Equation Model (SEM) model for the case considered

The final sample consisted of 545 participants (257 males, 288 females; mean age = 39.97, standard deviation = 14.78).

2.2.2 Pre-test measures

At the beginning of the first questionnaire (Time 1), participants reported their age, gender and typical diet (e.g. veganism, vegetarianism or restrictive diets). Then, they read a definition of "red/processed meat consumption" ("Red/processed meat is defined as mammalian meat, that is red when it is raw and dark in colour when cooked. This includes beef, lamb, pork, venison and goat and processed meat, like beef burgers, bacon, sausages etc. One serving is roughly the same size as a deck of cards, that is, at least two servings of vegetable per day"). After that, participants responded to a series of questions aimed at measuring the dimensions described below.

Attitude towards a reduced red/processed meat consumption was measured using a semantic differential scale ranging from (1) to (7) (e.g., "Eating little red/processed meat is. . . bad – good" [11]). Higher values indicated a more positive attitude towards a reduced red/processed meat consumption. Cronbach's alpha was .91.

Subjective norm was assessed with six items (e.g., "Most people who are important to me think that I should eat little red/processed meat"; [11]). Answers were given on a seven-point Likert scale, from (1) "strongly disagree" to (7) "strongly agree". Higher scores indicated a greater perceived social pressure to eat little red/processed meat. Cronbach's alpha was .87.

Perceived behavioural control was assessed using nine items ("If I wanted, I'd be able to avoid eating eat red/processed meat when I am busy"; adapted from [47]). Answers were again given on a seven-point Likert scale, from (1) "strongly disagree" to (7) "strongly agree". Higher scores indicated a greater control over eating little red/processed meat. Cronbach's alpha was 0.89.

Baseline intention to eat red/processed meat in the following month was measured using two items on a 7-point Likert scale (e.g., "In the next month, how often do you intend to eat red/processed meat?”; [13]). Answers were again given on a seven-point Likert scale, from (1) “never” to (7) “very often”. Higher values indicated higher intention to eat red/processed meat. Correlation among the items was $r = .75$, $p < .001$.

2.2.3 Message intervention

One week after completing the first questionnaire all participants were invited to read eight messages (approximately 14 words each) describing the health consequences of eating red/processed meat, and formulated in prefactual terms [13]. They read different messages according to the experimental condition. Participants in the *gain message condition* read eight messages on the positive health outcomes associated to little red/processed meat consumption, such as good functioning of the bowel, arteries and stomach (e.g., “If you eat little red meat and cold cuts, *you will improve* the health of your stomach”). Participants in the *non-loss message condition* read eight messages informing about how eating little red/processed meat is connected to preventing negative health outcomes (e.g., “If you eat little red meat and cold cuts, *you will avoid damaging* the health of your stomach”). Participants in the *non-gain message* condition read eight messages emphasizing how eating excessively red/processed meat is related to missing out positive health consequences (e.g., “If you eat little red meat and cold cuts, *you will miss the chance to improve* the health of your stomach”). Finally, participants in the *loss message condition* read eight messages about the negative health outcomes of eating too much red/processed meat (e.g., “If you eat much red meat and cold cuts, *you will damage* the health of your stomach”). The full list of messages is reported in Appendix.

2.2.4 Post-test measures

After reading the messages, participants were invited to answer to a series of questions measuring the dimensions described below (Time 2). *Message involvement* was measured with six items asking participants to state how involved they were in the messages (e.g., “The message was very interesting”; adapted from [26]). Answers were given on a 7-point Likert scale, from (1) “strongly disagree” to (7) “strongly agree”. Higher values indicated a higher participant’s involvement in the messages. Cronbach’s alpha was .94. *Future intention* to eat red/processed meat was measured with the same two items employed to measure baseline intention in the first questionnaire. Higher values indicated a higher intention to eat red/processed meat. Correlation between the items was $r = .84$, $p < .001$.

2.2.5 Data analysis

Our analysis was aimed to evaluate whether exposure to messages framed in different ways would lead to a different degree of involvement in the messages and, in turn, to a different level of intention to eat red/processed meat in the future.

As a first step of our analysis, we established the adequacy of fit of our theoretical model using a *hybrid SEM*, that is, a model that simultaneously includes latent variables and a mix of path analysis and confirmatory variables [28]. This approach is a full latent variable model consisting of measurement and structural parameters. The measurement parameters are related to the within-construct relationships, that is, the relations among the measured variables (such as the items of a scale) and their respective latent constructs [44]. The structural parameters regard the magnitude and direction of the relations among the latent constructs and are employed to verify the hypothesized relationships in the tested model [44]. In our hybrid model the measurement model was estimated including all study variables (attitude, subjective norm, perceived behavioral control, message involvement, baseline intention, and future intention) and the correlation matrix served as an input to estimate the structural coefficients between constructs and latent variables [32].

The adequacy of fit of the hybrid SEM model was estimated using a chi-square test and recommended incremental goodness-of-fit indices: the root mean square error of approximation (RMSEA), the comparative fit index (CFI) and the Tucker-Lewis Index (TLI). A nonsignificant chi-square test would indicate that the model fits the data well [29]. RMSEA value of 0.05 or less indicates a good fit and values up to 0.08 represent errors that approximate those expected in the population [29]. Finally, CFI and TLI cut-off values of at least 0.90 are generally considered to represent an acceptable fit [29].

After confirming the adequacy of fit of our hybrid SEM model, we used this model as a base model to test the invariance of the message involvement-intention path coefficients across groups. To do so, we applied a *multi-group Structural Equation Modeling*. We constrained the path from message involvement to future intention to be equal in each group, while we left the other path coefficients (among TPB variables, baseline intention and dependent variables) free to vary across groups. By disconfirming the equality (or invariance) of the message involvement-future intention path coefficient across the four groups, we would have been able to establish that the diverse messages read by participants moderated the relationship between message involvement and intention to eat red/processed meat in the future. We evaluated the null hypothesis of the equality of the effect of message involvement on intention at Time 2 across message groups by a Wald test.

2.3 Results and discussion

2.3.1 Preliminary analyses

Table 1 reports the means and standard deviations of all measures and Table 2 shows standardized factor loadings for each item. The items generally showed reasonable variation and were not unduly skewed. To check if randomization was successful, we used multivariate analysis of variance (MANOVA) on baseline intention to eat red/processed meat and age. Results did not show any significant main effect of the message groups on baseline intention and age ($p > .37$). Chi-square did not show any significant differences in gender ($p = .66$) across message groups. Thus, preliminary findings confirmed that randomization was adequate, and that the four message groups were matched on baseline intention to eat red/processed meat.

2.3.2 Main results

The goodness-of-fit statistics for the hybrid model were acceptable. The chi-square test was significant ($\chi^2 = 1744.90, df = 476, p < 0.001$), but all the other indices pointed to a good fit ($RMSEA = 0.07; CFI = 0.90; TLI = 0.90$). As showed in Table 2, the parameter estimates were all significant and presented adequate values (from 0.40 to 0.95). Given that the model fit data well, we tested whether the effect of message involvement on intention changed across groups, controlling for the TPB variables. We did so by computing a multi-group SEM model with message group (gain, loss, non-loss, non-gain and loss) as the grouping variable. The overall model fit of the multi-group SEM was acceptable ($\chi^2 = 3429.00, df = 2050 < 0.001; RMSEA = 0.07; CFI = .90; TLI = .90$; χ^2 from gain message group = 855.68; χ^2 from non-loss message group = 827.950); χ^2 from

Table 1 Means and standard deviations of study variables in each message condition

	Gain Messages (N=134)		Non-loss Messages (N=136)		Non-gain Messages (N=134)		Loss Messages (N=141)	
	M	SD	M	SD	M	SD	M	SD
Attitude Towards Reduced Red/Processed Meat Consumption	4.42	1.47	4.68	1.34	4.60	1.31	4.68	1.25
Subjective Norm	3.32	1.24	3.47	1.11	3.34	1.19	3.44	1.14
Perceived Behavioral Control	4.09	1.20	4.19	1.19	4.15	1.19	4.47	1.15
Baseline Intention to Eat Red/Processed Meat	3.81	0.97	3.89	0.84	3.93	0.83	3.83	0.91
Message Involvement	3.94	1.44	4.22	1.30	4.04	1.35	3.82	1.56
Future Intention to Eat Red/Processed Meat	3.74	0.83	3.83	0.87	3.79	0.79	3.75	0.75

non-gain message group = 895.00; χ^2 from loss message group = 904.29; and indicated that dataset had a good model fit.

As showed by the standardized results, in the *gain message group* participants' future intention to eat red/processed meat and was strongly determined by baseline intention ($\beta = -0.60$; $p = 0.001$) and reduced by message involvement ($\beta = -0.23$; $p = .05$). TPB predictors did not directly explain future intention, but subjective norm predicted message involvement ($\beta = 0.38$; $p = 0.05$) and mediation analyses showed that message involvement fully mediated the effects of subjective norm on future intention ($Ind. = -0.09$; $p = 0.02$). In sum, the perception of a social expectation reduced future intention to eat little red/processed meat when participants were involved by gain messages.

In the case of the *non-loss message group,* participants' intention to eat red/processed meat was again determined by baseline intention ($\beta = -0.51$; $p = 0.001$) and reduced by message involvement ($\beta = -0.15$; $p = .05$) . As in the case of gain messages, TPB predictors did not directly predict future intention but higher levels of positive attitude predicted higher message involvement ($\beta = 0.26$; $p = 0.001$). Therefore, greater positive attitude towards a reduced consumption increased participants' involvement in the non-loss messages, which in turn reduced future intention to consume red/processed meat. Overall, these findings showed that non-loss messages were a very effective intervention.

In the case of the *non-gain message group,* participants' future intention to eat red/processed meat was entirely determined by baseline intention ($\beta = -0.70$; $p = 0.001$) and attitude towards reduction ($\beta = 0.17$; $p = 0.001$), but not reduced by message involvement ($\beta = -0.12$; $p = .13$). The other TPB predictors did not directly explain future intention. Higher message involvement was determined by higher levels of subjective norm ($\beta = 0.20$; $p = 0.04$), but in this group there was not a significant mediation of subjective norm between message involvement and future intention. These findings, therefore, showed that non-gain messages were more involving when people perceived a higher social pressure to eat little red/processed meat. However, in this case message involvement did not reduce participants' intention to eat red/processed meat.

Table 2 Standardized factor loadings for study measures

Measures	Standardized Factor Loading
Attitude	
In your opinion, eating LITTLE red/processed meat is..	
. . . bad – good	0.89
. . . disadvantageous – advantageous	0.88
. . . unpleasant – pleasant	0.47
. . . boring – funny	0.40
. . . negative – positive	0.92
. . . unsatisfying – satisfying	0.89
. . . unwise – wise	0.84
Subjective Norm	
Most people who are important to me think that I should eat little red/processed meat	0.50
Most people who are important to me would approve if I ate little red/processed meat	0.51
Most people who are important to me would prefer that I eat little red/processed meat	0.63
Most people I know eat little meat	0.74
Most people I know believe that eating little red/processed meat is right	0.86
Most people I know would like to eat little red/processed meat	0.85
Perceived Behavioral Control	
If I wanted, I'd be able to avoid eating eat red/processed meat...	
... when I am busy	0.53
... when I am at work/during university semester	0.49
... when it is inconvenient (e.g., when I do not have enough plant-based food alternatives)	0.41
... when I go out to eat (e.g., when I eat at a restaurant)	0.60
... when I eat at the canteen or cafeteria	0.52
... when I am at home	0.85
... when I am at home	0.83
... over the summer	0.78
... during the weekend	0.85
Baseline Intention to Eat Red/Processed Meat	
In the next month, how often do you intend to eat red/processed meat?	0.88
How many portions of red/processed meat do you expect to eat next month?	0.85

Finally, in the case of the *loss message group,* participants' intention to eat red/processed meat was again strongly determined by baseline intention ($\beta = -0.78$; $p = 0.001$), but

Table 2 (continued)

Measures	Standardized Factor Loading
Message Involvement	
The messages that I read...	
... got me involved in what the message had to say	0.89
... seemed relevant to me	0.86
... really made me think	0.87
... were thought-provoking	0.84
... were very interesting	0.88
... caused strong emotions in me	0.73
Future Intention to Eat Red/Processed Meat	
In the next month, how often do you intend to eat red/processed meat?	0.90
How many portions of red/processed meat do you expect to eat next month?	0.92

was also reduced by message involvement ($\beta = -0.24$; $p = .001$). No TPB predictors directly determined participants' future intention. Both subjective norm ($\beta = 0.20$; $p = 0.04$) and attitude towards reducing the red/processed meat consumption ($\beta = 0.20$; $p = 0.02$) influenced message involvement and message involvement fully mediated the effect of attitude on future intention to eat red/processed meat ($Ind. = -0.05$; $p = 0.05$). Overall, these results showed that the more participants had a high level of positive attitude towards reduction and subjective norm, the more they perceived loss messages as involving. In other words, a high positive attitude towards eating little red/processed meat determined low future intention to eat red/processed meat only when people perceived the possible health risks of eating a great amount of red/processed meat as involving.

After considering the links between study variables in each group, we conducted Wald tests to compare the message involvement-future intention path coefficient among groups with the SEM models where this path was constrained to be equal across groups. The Wald tests were significant when comparing both non-loss with gain message group ($\chi^2 = 4.21$, df $= 1$, $p < .05$) and non-loss with non-gain message group ($\chi^2 = 4.01$, df $= 1$, $p < .05$). These findings indicated that there was a degradation in the fit of the model when the path was constrained to be equal across these groups. Therefore, we accepted the first unconstrained model, suggesting that groups differed significantly in the parameters tested by the model.

The aforementioned group comparisons suggested that the persuasive process of the non-loss message was different from those of the gain and non-gain messages. In the case of non-loss messages participants' message involvement was greater when they had a positive attitude towards reducing red/processed meat consumption. However, the messages were involving and reduced the intention to consume red/processed meat even when participants had a low positive attitude towards reducing meat consumption. This was not the case for gain and non-gain messages. Gain messages were effective in reducing future intention only when participants' level of subjective norm increased message involvement. As to non-gain messages, they reduced future intention only when message involvement was supported by the initial attitude.

3 Reinforcement learning

3.1 Dynamic bayesian network

In probability theory and in statistics, *graphical models* [37] are well-known and popular modeling tools. Graphical models are graph descriptions of random variables and dependencies among them in a specific probabilistic model. Besides providing a structural framework for the multi-dimensional, joint probability distribution lying behind, the added value of graphical models resides in their clarity and easiness of visualization. *Bayesian Networks* (BNs, [4]), in particular, are described by *Directed Acyclic Graphs* (DAGs), that is, oriented graphs that do not contain cycles. In a BN, nodes represent random variables and edges represent *conditional* dependencies among them.

In this study, Bayesian Networks were used to translate the SEM described in the previous section. Nonetheless, the SEM in Fig. 1 has a longitudinal temporal structure involving Time 1 (e.g. Baseline Intention) and Time 2 (e.g. Future Intention). To cope with this, we adopted *Dynamic Bayesian Networks* (DBNs, [19]). In a DBN, each time step is represented by a 'slice' of the graphical model, as shown in Fig. 2, and random variables are replicated in each 'slice' whenever the same variables are expected to have different values at different time steps. The main constraint in a DBN is that all edges across subsequent slices must be oriented towards the direction of time.

The purpose of the SEM-translating DBN in Fig. 2 was to get a *predictor*. More precisely, given specific observations at Time 1 - i.e. answers to specific items - the DBN would predict which message group had the highest probability to modify variables at Time 2 in

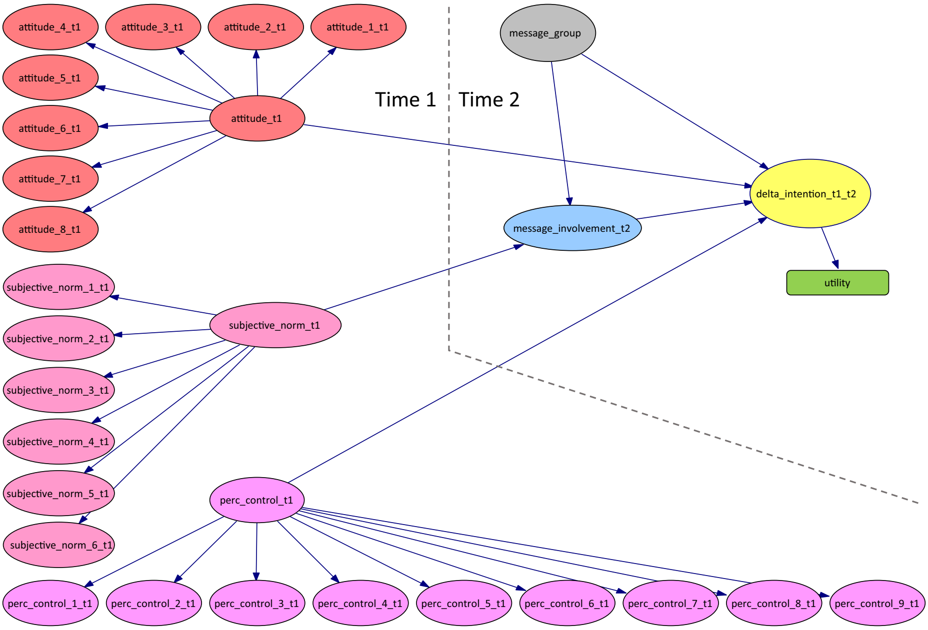

Fig. 2 The Dynamic Bayesian Network obtained form the SEM shown in Fig. 1. The dashed line separates the two slices corresponding to Time 1 and Time 2, respectively. The overall structure reflects the purpose of the DBN: variables at Time 1 were to be assessed, while the relevant Time 2 variables were to be *predicted*

the sense desired. We focused on interactions in which the conversational agent could only make observations at Time 1, whereas the expected value of variables at Time 2 inferred from DBN was to be used by the agent to decide which message group should be sent. Note that, in the DBN in Fig. 2, only the two Time 2 variables relating to involvement and intention change (see below) were mentioned, since those were relevant for assessing the effectiveness of messages.

To ease the learning phase, the DBN was assumed to contain discrete random variables only, ranging on a reduced scale. As described in the previous section, all answers to the items in the questionnaire were on a scale of 7 discrete values, which yields a very large number of combinations (for instance, $7^8 = 5.764.801$ for the conditioning of variable attitude_t1 alone). In addition, all latent SEM variables were continuous averages in the interval $[1, 7]$. We decided to aggregate numerical values in both latent and item variables according to the following reduced scale of 3 values:

- low, if value $\in [1, 3)$;
- medium, if value $\in [3, 5]$;
- high, if value $\in (5, 7]$.

Variable delta_intention_t1_t2 was defined as the difference between the latent SEM variables *Baseline Intention* and *Future Intention*. Its values were aggregated in a scale of 5:

- high_positive, if value ≤ -1;
- positive, if value $\in (-1, -0.2]$;
- neutral, if value $\in (-0.2, 0.2)$.
- negative, if value $\in [0.2, 1)$;
- high_negative, if value ≥ 1).

The utility function was defined as the weighted sum:

$$\begin{aligned} \text{utility} := \; & 2\,P(\text{delta\_intention\_t1\_t2} = \text{high\_positive}) \\ & + P(\text{delta\_intention\_t1\_t2} = \text{positive}) \\ & - P(\text{delta\_intention\_t1\_t2} = \text{negative}) \\ & - 2\,P(\text{delta\_intention\_t1\_t2} = \text{high\_negative}). \end{aligned}$$

Once learnt, the *Joint Probability Distribution* (JPD) of a DBN allows answering, through marginalization and conditionalization, any probabilistic query about the model itself. The conditional probabilities to be learnt for the DBN in Fig. 2 were the following:

- $P(\text{attitude\_}i\text{\_t1})$, for $i \in \{1, \ldots, 8\}$;
- $P(\text{subjective\_norm\_}i\text{\_t1})$, for $i \in \{1, \ldots, 6\}$;
- $P(\text{perc\_control\_}i\text{\_t1})$, for $i \in \{1, \ldots, 9\}$;
- $P(\text{message\_group})$;
- $P(\text{attitude\_}i\text{\_t1} \mid \text{attitude\_t1})$, for $i \in \{1, \ldots, 8\}$;
- $P(\text{subjective\_norm\_}i\text{\_t1} \mid \text{subjective\_norm\_t1})$, for $i \in \{1, \ldots, 6\}$;
- $P(\text{perc\_control\_}i\text{\_t1} \mid \text{perc\_control\_t1}$, for $i \in \{1, \ldots, 9\}$;
- $P(\text{subjective\_norm\_t1} \mid \text{message\_involvement\_t2})$;
- $P(\text{delta\_intention\_t1\_t2} \mid$
  message_group,
  attitude_t1,

message_involvement_t2,
cperc_control_t1).

Note that each subgraph in Fig. 2, composed by one latent variable and its corresponding items, forms a *Naïve Bayesian Classifier* pattern. Although this may seem in contrast with the fact that each latent variable in the SEM is deterministically determined by the average value of the corresponding item variables, such translation is instrumental for taming the complexity of the DBN. For instance, the conditional probability for variable attitude_t1 alone would otherwise include $3^8 = 6.561$ subspaces, which would require an unrealistic amount of training data.

With the above provisions, the JPD of the DBN was learnt through *Maximum Likelihood Estimation* (MLE), meaning that all probabilities required were estimated as relative frequencies in experimental data.

3.2 Partially-observable markov decision process

The method of *Reinforcement Learning* (RL, [45]) is based on the fundamental definition of a *Markov Decision Process* (MDP, [24]). Roughly speaking, in a (discrete) MDP there is a finite number of situations or *states* of a target environment with which the agent is supposed to interact. At each time step, the agent will select an *action* to perform, which will induce a *state* transition and produce a *reward* to the agent. The objective of RL is to find an agent *policy* that generates, for each initial state, the sequence of actions that attains the maximum possible cumulative reward, over a sequence of time steps.

In our specific case, due to the presence of *latent* variables, the agent could not be assumed to have complete knowledge about the state of the environment, as an MDP would require. In such situation, the agent must maintain its own estimates about the current state, by relying on the history of past actions and observations. This entails using the MDP variant called *Partially-Observable Markov Decision Process* (POMDP, [30]).

With reference to the random variables in the DBN (see Fig. 2), in our case the POMDP was defined as follows. The *state space* was defined as the set of the latent SEM variables in the DBN:

$$\begin{aligned}\mathcal{S} := \{&\mathsf{attitude\_t1}, \mathsf{subjective\_norm\_t1},\\ &\mathsf{perc\_control\_t1}, \mathsf{message\_involvement\_t2},\\ &\mathsf{delta\_intention\_t1\_t2}\}.\end{aligned}$$

The *action space* was defined as the set of actions that could be performed by the agent during each stage of the interaction:

$$\begin{aligned}\mathcal{A} := &\{\mathit{ask}\ \mathsf{attitude\_}i\mathsf{\_t1}\}_{i=1}^{8} \cup \{\mathit{ask}\ \mathsf{subjective\_norm\_}i\mathsf{\_t1}\}_{i=1}^{6}\\ \cup\ &\{\mathit{ask}\ \mathsf{perc\_control\_}i\mathsf{\_t1}\}_{i=1}^{9} \cup \{\mathit{send}\ \mathsf{message\_group} = m\},\end{aligned}$$

where $m \in \{\mathsf{gain}, \mathsf{nongain}, \mathsf{loss}, \mathsf{nonloss}\}$. The *observation space* is the set of possible answers to the questions/items:

$$\begin{aligned}\Omega := &\{\mathsf{attitude\_}i\mathsf{\_t1}\}_{i=1}^{8} \cup \{\mathsf{subjective\_norm\_}i\mathsf{\_t1}\}_{i=1}^{6}\\ \cup\ &\{\mathsf{perc\_control\_}i\mathsf{\_t1}\}_{i=1}^{9},\end{aligned}$$

were each answer was in {low, medium, high}.

Each POMDP episode was assumed to represent a complete interaction, from question asking to the choice of a message group to be sent. In general, in a POMDP the agent is assumed to start from an unknown initial state s_0 and to hold a *belief state* described by a probability distribution over the state space. At each time instant t, the agent will be in state s_t and perform an action a_t: in our case that could correspond to either *ask*ing an item or *send*ing a message group. Each agent action produces a transition to state s_{t+1} and an observation o_{t+1}. In the setting considered, observations corresponded to the answers given by the interactant. Each incoming observation o_{t+1} produces a new belief state $b_{t+1}(s_{t+1})$, described by the conditional probability distribution given all observations made thus far in the episode. More formally:

$$b_{t+1}(s_{t+1}) := P(s_{t+1} \mid h_t, a_t, o_{t+1})$$

where

$$h_t := \langle a_0, o_1, \ldots, a_{t-1}, o_t \rangle$$

and a_0 is the first action performed by the agent. In our case, the legitimate actions that the agent could perform were constrained by the past history, since no question/item could be asked more than once. We also deemed the action of *send*ing a message group as terminal, in the sense that it led to a *goal* state in which the interaction episode became complete.

In our setting, the POMDP was assumed to produce *delayed* rewards only, in the sense that the agent could receive a reward at the end of each interaction episode only, once a goal state was reached. The cumulative *reward function* at a goal state s_T was defined as

$$R(h_T) := \mathit{utility}(s_T) - c \cdot n_{\mathit{item}}(h_T)$$

where $\mathit{utility}(s_T)$ was the utility predicted by the DBN given the observations in h_T, $n_{\mathit{item}}(h_T)$ was the number of questions asked during the episode and c was a constant, expressing a cost per item. Such definition of the cumulative reward reflected the design intention to make the agent seeking to send the message that maximizes the probability of inducing a positive effect while trying to minimize the number of questions asked.

In the POMDP framework, the agent's *policy* is defined as:

$$\pi = \pi(b_t(s_t), h_t)$$

whereas the *state-transition* function describes the changes in the environment over time:

$$T(s_{t+1}, s_t, a_t) := P(s_{t+1} \mid s_t, a_t).$$

In our setting, the only relevant state transition was assumed to be the one caused by sending a message group, i.e. from Time 1 to Time 2, whereas actions of question asking could only change the agent belief state. The POMDP *observation function*

$$O(s_{t+1}, s_t, a_t) := P(o_{t+1} \mid h_t, s_t, a_t)$$

was assumed to be the (marginal) DBN probability distribution over the answers to the last asked question, given the history of past observations. As implied by the structure of the

DBN, we assumed that the sequence in which actions were performed was irrelevant. The *expected cumulative reward*, given a specific POMDP policy π, is defined as:

$$R := \mathbb{E}[R(h_T)]$$

where the expectation is taken over all possible histories h_T. The overall objective in a POMDP is finding an optimal policy

$$\pi^*(b_t(s_t), h_t)$$

that maximizes the expected cumulative reward R. As one could guess, the critical problem in a POMDP is finding the optimal policy π^*, which is the topic of the next section.

3.3 Policies from neural monte carlo tree search

Figure 3 represents the tree expansion of the space of all possible policies in the setting described. At each step, the agent performs one action from the action space $\mathcal{A}$. A deterministic policy π must select one action per each expansion step in figure. The interactant, on the other hand, was assumed to answer at random to any question asked, according to probability distribution O. As such, in our setting a specific policy π was represented by a sub-tree of the full expansion in Fig. 3, in which actions were uniquely determined at each decision point and multiple branches followed, one per each possible answer.

Considering the statistics for the expansion tree in figure, we see that at the first step, i.e. $t = 0$, the agent has to select one in $q + m$ legitimate actions, where q is the number of items and m the number of message groups. In the POMDP considered, $q = 8+6+9 = 23$ and $m = 4$, therefore $q + m = 27$. Asking one question, in turn, led to $k = 3$ possible

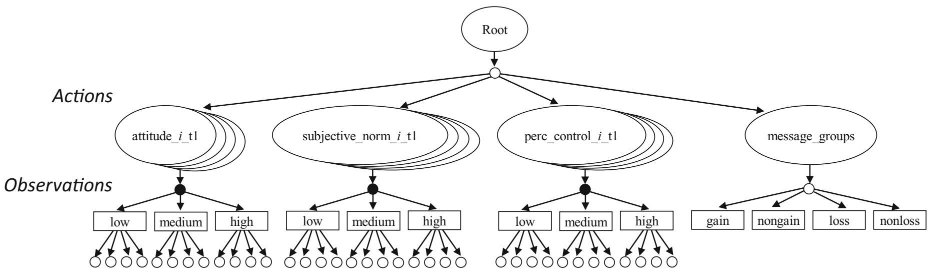

Fig. 3 The POMDP policy space represented as a tree expansion. At each level, an action is performed (ovals - i.e. either asking a question/item or sending a message group) and an observation is made (black dots) of possible answers (square boxes). The expansion proceeds (not shown) until a goal state is reached on all branches

observations. At each subsequent non-terminal step $t > 0$, the number of legitimate actions was $q + m - t$, since no item could be asked more than once. This line of reasoning leads to:

$$N(t) = (q - t)\, k\, N(t+1) + m \; > \; (q - t)\, k\, N(t+1),$$

where $N(t)$ is the number of legitimate actions at level t. Expanding $(q - t)\, k\, N(t+1)$ in the range $t = 0, \ldots, q$ yields $q!\, k^q$, which is a lower bound for the number of action decision points. In the case considered $q!\, k^q \approx 10^{33}$, which makes finding an optimal policy impossible with any brute-force computation method.

The combination of the *Monte Carlo Tree Search* (MCTS) method ([16]) and *Upper Confidence Bound* mechanism (UCB, [33]) was proposed for the approximate optimization of POMDP policies [42]. Figure 4 describes the basic mechanism in the MCTS method. At any given time, one particular selection point is chosen (1) and one possible action is picked, also adding the corresponding node to the ongoing expansion (2). Then a simulation is performed by playing a few *rollouts* (3), namely by completing episodes in the interaction, starting from the newly-added node and selecting actions at random until a goal state is reached and the cumulative reward $R(h_T)$ can be computed. Rewards obtained from multiple rollouts are averaged.

The critical aspect for the effectiveness of the MCTS method is the selection phase (1), which is crucial for taming the combinatorial explosion of the policy space. In MCTS-UCB [33], each selection point p is associated with a function:

$$U_p(a_p) \coloneqq Q_p(a_p) + \eta \cdot \sqrt{\frac{\ln N_p}{N_{a_p}}}$$

where a_p is a legitimate action, Q_p is the current estimate of the (best) cumulative reward attainable, $N_p \geq 1$ is the number of times the node has been selected and N_{a_p} is number of times action a_p has been selected. During the selection phase (1) the node to be expanded is chosen from the root by selecting the action a_p having the largest value of

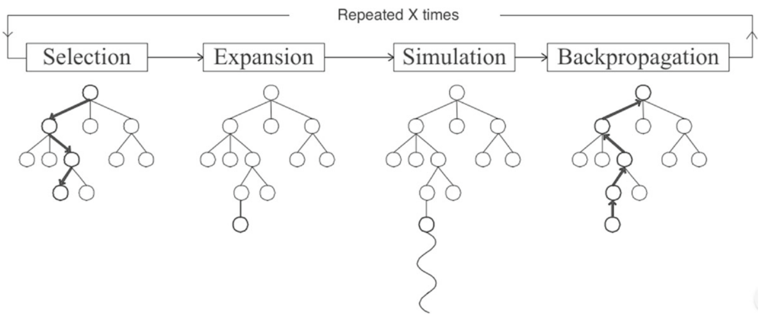

Fig. 4 Main steps in the iterative MCTS method

U_p. After the selection is completed and rollouts are performed, all counters and estimates are updated. This selection function aims to balance the *exploitation* of current estimates Q_p with the *exploration* of alternatives, by adding an extra confidence factor that decreases exponentially with the counter N_{a_p}.

The early experiments performed with MCTS-UCB were not particularly successful with the POMDP discussed here. The reason was that the MCTS-UCB method still requires an expansion, albeit limited, of the policy space. As it happened in our case, such expansion would grow to an unmanageable size even when the agent was constrained to ask no more than 5-6 questions per episode. Worse yet, unless goal states were reached on all best branches, any intermediate results produced by the MCTS-UCB method were unusable since they did not represent a fully-specified policy π. An alternative method could have been adopting *online* MCTS-UCB variant, as proposed in [42] and [31], in which a *local* MCTS-UCB search is performed at each decision point, to select the most promising action. However, such choice would have produced a non-deterministic policy, in that each action selection would have depended on a random search procedure.

The Deep Neural Network (DNN) variant of MCTS proposed in [41], on the other hand, proved to be particularly effective for the *offline* approximate optimization of the POMDP policy in point. In practice, we created a problem-specific, single-player version of the Neural MCTS method in [41]. Reportedly, the latter method had achieved a spectacular success in the realm of two-players, zero-sum adversarial games like go, chess and shogi.

In our experiments, *online* MCTS-UCB search was used for training a DNN whose purpose was keeping a concise and persistent evaluation function of the alternative action options and their expected results. The actual policy π, in fact, was extracted from the DNN alone, after completing the training process: by selecting at each step the action having the highest DNN-estimated probability of effectiveness, we could elicit a deterministic, offline policy π. More in detail, the search process was organized into *episodes*, i.e. full interactions until a goal state was reached. In each episode, *local* MCTS-UCB searches were performed at each decision point by keeping the same MCTS-UCB tree along the entire episode and computing the cumulative reward when reaching the goal state. Then, after each episode, the MCTS-UCB tree was restarted anew. Collected episode histories h_T plus the cumulative rewards $R(h_T)$ were used to train the DNN. The objective was obtaining an effective estimator of the value $Q(h_t, a_t)$ of the cumulative reward of the episode, were h_t is past history and a_t is a legitimate action, together with the probability distribution $P(a_t \mid h_t)$ of the best rewarding action. As suggested in [41], the following function was used for the local MCTS-UCB search:

$$U_p(a_p) \coloneqq Q_p(a_p) + \eta \cdot P(a_p \mid h_p) \cdot \frac{\sqrt{N_p}}{N_{a_p}}$$

Figure 5 shows approximated optimal policies $\tilde{\pi}^*$, in the form of tree expansions, obtained with the Neural MCTS method described. In between the root node, corresponding to the initial state, and the leaves, i.e.the goal states of each tree, intermediate nodes correspond to the observation made (i.e. the answer by the interactant) after an action of

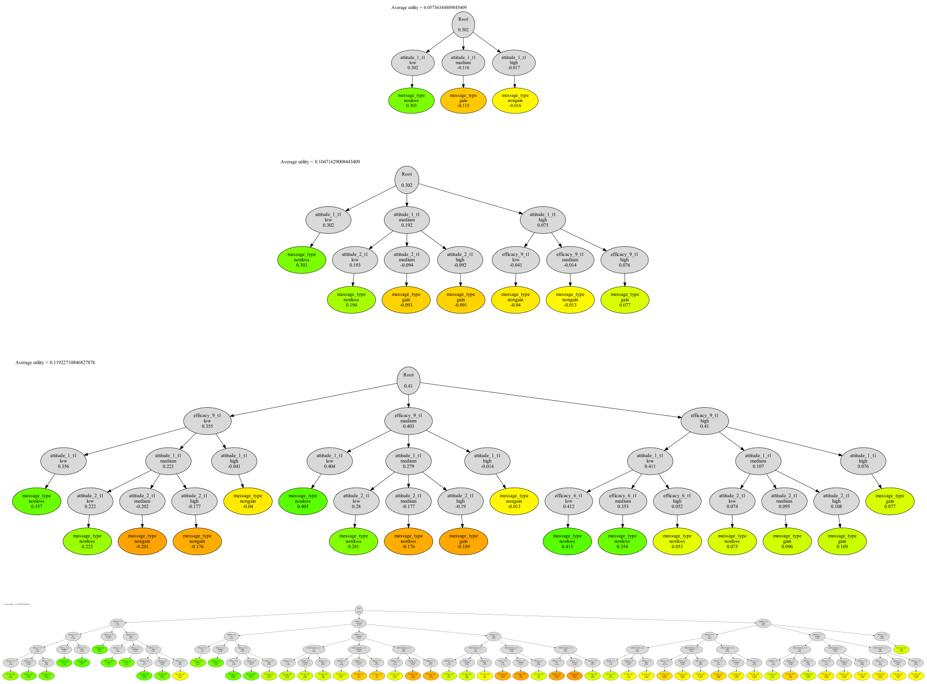

Fig. 5 Four approximately-optimal policies $\tilde{\pi}^*$ obtained with the proposed Neural MCTS method, with a maximum of 1, 2, 3 and 4 questions asked, respectively. Nuances of green (vs. red) correspond to positive (vs. negative) *utility* values attained at goal states

question-asking is performed. Note that all triplets of intermediate nodes having the same parent correspond to the same item. Leaf nodes represent the messages group being sent and the *utility* value attained: positive values lean to the green color while negative ones lean to the red.

Each tree in figure describes a complete, deterministic policy π obtained with the method described. The four trees were obtained by constraining the agents to ask no more than 1, 2, 3 and 4 questions, respectively. The effectiveness of policies obtained was measured in terms of *average utility*, computed by the following weighted sum over the goal states:

$$utility_{\text{avg}} := \sum_{s \,\in\, goal} utility(s) \cdot \frac{1}{3^{N_{\text{ask}}(s)}}$$

where $N_{\text{ask}}(s)$ is the number of items asked to reach the goal and 3 is the number of possible answers. As it can be seen, the green color dominates and in fact all four policies attained positive average utility. The four actual values were 0.057, 0.105, 0.119 and 0.141, respectively. These figures also show that, by allowing the agent to ask more question, the precision of the profiling hence the effectiveness of message selection could be improved. This trend was confirmed by experiments made with looser constraints. The experiments made also demonstrated that the Neural MCTS method could handle the approximate optimization of policies of virtually any length, within the setting described, given sufficient time and computational power.

4 Conclusions

In this pivotal work we explored the possibility of integrating the models and skills of social psychology and machine learning to develop communication strategies useful for reducing the consumption of red/processed meat. We developed a pilot model to create a dialogue manager capable of making a fast profiling of the receivers and selecting the potentially most persuasive messages according to the receivers' profile. Our results, although still to be considered preliminary given the novelty of the area, advance our knowledge both in terms of persuasive communication and in terms of reinforcement learning enriched by expert psychological knowledge.

With regard to persuasive communication, in this study we were able to assess that messages framed in terms of gain, non-loss, non-gain and loss are differentially persuasive in inducing people to reduce their consumption of red and processed meat. Non-loss-framed messages, which inform about the possibility to avoid negative consequences by reducing meat consumption, turned out to be the most effective messages, apt to involve and persuade the majority of receivers, independent on their prior beliefs. The effectiveness of gain- and non-gain-framed messages was instead more related to receivers' psychosocial characteristics. Gain-framed messages were more likely to persuade receivers when they perceived a social expectation to change their eating habit. In turn non-gain-framed messages (related to the possibility of missing out some gains with a negative behaviour) were more likely to persuade receivers when they had a prior positive attitude towards reducing red meat consumption. Finally, loss-framed messages were the least persuasive messages and convinced only receivers who were involved by them. These results add to previous research showing that the effectiveness of persuasive messages on meat intake depends on the way messages are framed and on receivers' prior beliefs (e.g.[26, 46]).

From the perspective of reinforcement learning, we assessed the feasibility of a method that could harness the statistical model deriving from psychosocial experiments to elicit effective dialogue management policies. According to the model-translating probabilistic predictor, policies obtained with our method show that profiling interactions aiming to the selection of the most effective messages could lead to improved persuasion effectiveness. With the metrics adopted, we also observed a consistent trend whereby allowing more questions to be asked led to increasing average utility. Nonetheless, even when the dialogue was constrained to a very limited number of exchanges, the lead of the psychological model could produce more persuasive interaction policies. Finally, while our policy optimization method required highly-sophisticated techniques, involving deep neural networks and Monte Carlo tree search, the policies obtained were deterministic and with a tree-like structure, hence could be implemented in a conversational agent using current technologies.

Overall, the results presented here support the idea that the integration between psychosocial research on the one hand and machine learning on the other could lead to the development of newer and more sophisticated techniques of digital interaction based on personalised and adaptive exchanges. In perspective, such techniques could foster the realization of more humanized automated conversational agents devoted to promoting healthier habits by taking into better account individual differences in terms of motivation and capacity.

Appendix: Persuasive messages

Gain Messages	Non-loss Messages	Non-gain Messages	Loss Messages
If you eat little red meat and cold cuts, *you will improve* the health of your stomach.	If you eat little red meat and and cold cuts, *you will avoid damaging* the health of your stomach.	If you eat much red meat and and cold cuts, *you will miss the chance to improve* the health of your stomach.	If you eat much red meat and and cold cuts, *you will damage* the health of your stomach.
If you eat little red meat and cold cuts, *you will improve* the functioning of your bowel.	If you eat little red meat and and cold cuts, *you will avoid damaging* the functioning of your bowel.	If you eat much red meat and and cold cuts, *you will miss the opportunity to improve* the functioning of your bowel.	If you eat much red meat and and cold cuts, *you will damage* the functioning of your bowel.
If you eat little red meat and cold cuts, *you will improve* the functionality of your heart.	If you eat little red meat and and cold cuts , *you will avoid damaging* the functionality of your heart.	If you eat much red meat and and cold cuts, *you will miss the chance to improve* the functioning of your heart.	If you eat much red meat and and cold cuts, *you will damage* the functionality of your heart.
If you eat little red meat and cold cuts, *you will improve the proper functioning* of your arteries.	If you eat little red meat and and cold cuts, *you will avoid increasing the malfunctioning* of your arteries.	If you eat much red meat and and cold cuts, *you will miss the opportunity to improve the proper functioning* of your arteries.	If you eat much red meat and and cold cuts, *you will increase the malfunctioning* of your arteries.
If you eat little red meat and cold cuts, *you will enhance* the functionality of your kidneys.	If you eat little red meat and and cold cuts, *you will avoid straining* the functionality of your kidneys.	If you eat much red meat and and cold cuts, *you will miss the chance to enhance* the functionality of kidneys.	If you eat much red meat and and cold cuts, *you will strain* the functionality of your kidneys.
If you eat little red meat and cold cuts, *you will enhance* the health of your lungs.	If you eat little red meat and and cold cuts, *you will avoid damaging* the health of your lungs.	If you eat much red meat and and cold cuts, *you will miss the opportunity to enhance* the health of your lungs.	If you eat much red meat and and cold cuts, *you will damage* the health of your lungs.

If you eat little red meat and cold cuts, *you will enhance* the health of your pancreas.	If you eat little red meat and and cold cuts, *you will avoid damaging* the health of your pancreas.	If you eat much red meat and and cold cuts, *you will miss the chance to enhance* the health of your pancreas.	If you eat much red meat and and cold cuts, *you will damage* the health of your pancreas.
If you eat little red meat and cold cuts, *you will improve* the chance of having *an optimal blood pressure*.	If you eat little red meat and and cold cuts, *you will decrease* the chance of having *hypertension*.	If you eat much red meat and and cold cuts, *you will miss* the chance of having *an optimal blood pressure*.	If you eat much red meat and and cold cuts, *you will increase* the chance of having *hypertension*.

References

1. Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50(2):179–211. https://doi.org/10.1016/0749-5978(91)90020-t
2. Allen J, Ferguson G, Stent A (2001) An architecture for more realistic conversational systems. In: Proceedings of the 6th international conference on intelligent user interfaces. ACM, pp 1–8. https://doi.org/10.1145/359784.359822
3. Bach-Faig A, Berry EM, Lairon D, Reguant J, Trichopoulou A, Dernini S, Medina FX, Battino M, Belahsen R, Miranda G, Serra-Majem L (2011) Mediterranean diet pyramid today. Sci Cult Updates Public Health Nutr 14(12A):2274–2284. https://doi.org/10.1017/s1368980011002515
4. Ben GI (2007) Bayesian networks. encyclopedia of statistics in quality and reliability. John Wiley & Sons, New York
5. Bertolotti M, Carfora V, Catellani P (2020) Regulatory focus and the effect of nutritional messages on health and well-being: the case of red meat intake. Appl Psychol Health Well Being 12(1):212–230
6. Bertolotti M, Carfora V, Catellani P (2020) Different frames to reduce red meat intake: the moderating role of self-efficacy. Health Commun 35(4):475–482. https://doi.org/10.1111/aphw.12180
7. Bianchi F, Aveyard P, Astbury NM, Cook B, Cartwright E, Jebb SA (2019) Replacing meat with alternative plant-based products (RE-MAPs): protocol for a randomised controlled trial of a behavioural intervention to reduce meat consumption. BMJ open 9(5):e027016. https://doi.org/10.1136/bmjopen-2018-027016
8. Bohm I, Lindblom C, Åbacka G, Bengs C, Hörnell A (2015) "He just has to like ham" – The centrality of meat in home and consumer studies. Appetite 95:101–112. https://doi.org/10.1016/j.appet.2015.06.015
9. Bosone L, Martinez F (2017) When, how and why is loss-framing more effective than gain-and non-gain-framing in the promotion of detection behaviors? Int Rev Soc Psychol 1(30):184–192. https://doi.org/10.5334/irsp.15
10. Carfora V, Bertolotti M, Catellani P (2019) Informational and emotional daily messages to reduce red and processed meat consumption. Appetite 141:104331. https://doi.org/10.1080/10410236.2019.1567444
11. Carfora V, Caso D, Conner M (2017) Randomised controlled trial of a text messaging intervention for reducing processed meat consumption: the mediating roles of anticipated regret and intention. Appetite 117:152–160. https://doi.org/10.1016/j.appet.2017.06.025
12. Carfora V, Caso D, Palumbo F, Conner M (2018) Promoting water intake. the persuasiveness of a messaging intervention based on anticipated negative affective reactions and self-monitoring. Appetite 130:236–246. https://doi.org/10.1016/j.appet.2018.08.017
13. Carfora V, Catellani P, Caso D, Conner M (2019) How to reduce red and processed meat consumption by daily text messages targeting environment or health benefits. J Environ Psychol 65:101319. https://doi.org/10.1016/j.jenvp.2019.101319
14. Caso D, Carfora V (2017) Messaging intervention for promoting self-monitoring of fruit and vegetable consumption — [Un intervento di messaggistica istantanea per la promozione del monitoraggio del consumo di frutta e verdure]. Psicologia della Salute 1:97–111. https://doi.org/10.3280/PDS2017-001005

15. Cesario J, Corker KS, Jelinek S (2013) A self-regulatory framework for message framing. J Exp Soc Psychol 49:238–249. https://doi.org/10.1016/j.jesp.2012.10.014
16. Chaslot G, Bakkes S, Szita I, Spronck P (2008) Monte-Carlo tree search: a new framework for game AI. Bijdragen
17. Cheng T, Woon DK, Lynes JK (2011) The use of message framing in the promotion of environmentally sustainable behaviors. Soc Mar Q 17(2):48–62. https://doi.org/10.1080/15245004.2011.570859
18. Corrin T, Papadopoulos A (2017) Understanding the attitudes and perceptions of vegetarian and plant-based diets to shape future health promotion programs. Appetite 10:40–47. https://doi.org/10.1016/j.appet.2016.11.018
19. Dagum P, Galper A, Horvitz E (1992) Dynamic Network Models For Forecasting. In: Proceedings of the eighth conference on uncertainty in artificial intelligence
20. de Carolis B, Mazzotta I (2017) A user-adaptive persuasive system based on 'a-rational' theory. Int J Hum-Comput St 108:70–88. https://doi.org/10.1016/j.ijhcs.2017.05.005
21. Dijkstra A, Rothman A, Pietersma S (2011) The persuasive effects of framing messages on fruit and vegetable consumption according to regulatory focus theory. Psychol Health 26(8):1036–1048. https://doi.org/10.1080/08870446.2010.526715
22. Eagly AH, Chaiken S (1993) The psychology of attitudes. Harcourt Brace Jovanovich College Publishers. https://doi.org/10.1002/mar.4220120509
23. Eshel G, Martin PA (2006) Diet, energy, and global warming, Earth Interactions (10)1–17. https://doi.org/10.1175/ei167
24. Fabiani P, Teichteil-Königsbuch F (2010) Markov decision processes in artificial intelligence. Wiley-ISTE, New York
25. Farchi S, De Sario M, Lapucci E, Davoli M, Michelozzi P (2017) Meat consumption reduction in Italian regions: health co-benefits and decreases in GHG emissions. Plos One 12(8):e0182960. https://doi.org/10.1371/journal.pone.0182960
26. Godinho CA, Alvarez MJ, Lima ML (2016) Emphasizing the losses or the gains: comparing situational and individual moderators of framed messages to promote fruit and vegetable intake. Appetite 96:416–425. https://doi.org/10.1016/j.appet.2015.10.001
27. Graça J, Calheiros MM, Oliveira A (2015) Attached to meat? (Un)Willingness and intentions to adopt a more plant-based diet. Appetite 95:113–125. https://doi.org/10.1016/j.appet.2015.06.024
28. Hancock GR, Samuelsen KM (2007) Advances in latent variable mixture models IAP
29. Iacobucci D (2010) Structural equations modeling: Fit indices, sample size, and advanced topics. J Consum Psychol 20:90–98. https://doi.org/10.1016/j.jcps.2009.09.003
30. Kaelbling PL, Littman M, Anthony RC (1998) Planning and acting in partially observable stochastic domains. Artif Intell 101:99–134
31. Katt S, Oliehoek FA, Amato C (2017) Learning in POMDPs with monte carlo tree search proceedings of the 34th international conference on machine learning - Volume 70, ICML 2017
32. Kline R (1988) Principles and practice of structural equation modeling. Guilford Press, New York
33. Kocsis L, Szepesvári C (2006) Bandit based monte-carlo planning. Springer, Berlin, pp 282–293
34. Laczniak RN, Muehling DD (1993) Toward a better understanding of the role of advertising message involvement in ad processing. P&M 10(4):301–319. https://doi.org/10.1002/mar.4220100405
35. Liu B (2018) Learning task-oriented dialog with neural network methods. PhD thesis
36. Misra R, Balagopal P, Raj S, Patel TG (2018) Red meat consumption (heme iron intake) and risk for diabetes and comorbidities? Curr Diab Rep 18(11):100. https://doi.org/10.1007/s11892-018-1071-8
37. Pearl J (1988) Newblock Probabilistic Reasoning in Intelligent systems: Networks of Plausible Inference. Representation and Reasoning Series (2nd printing ed.) Morgan Kaufmann, San Francisco
38. Petty RE, Cacioppo JT (1986) The elaboration likelihood model of persuasion. In: Communication and persuasion. Springer, New York, pp 1–24. https://doi.org/10.1007/978-1-4612-4964-1_1
39. Petty RE, Cacioppo JT (2012) Communication and persuasion: central and peripheral routes to attitude change. Springer Science & Business Media, Berlin
40. Rothman AJ, Bartels RD, Wlaschin J, Salovey P (2006) The strategic use of gain-and loss-framed messages to promote healthy behavior: how theory can inform practice. J Commun 56(suppl_1):S202–S220. https://doi.org/10.1111/j.1460-2466.2006.00290.x
41. Silver D, Schrittwieser J, Simonyan K, et al. (2017) Mastering the game of go without human knowledge. Nature 550(7676):354–359. https://doi.org/10.1038/nature24270
42. Silver David, Veness J (2010) Monte-Carlo planning in large POMDPs in advances in neural information processing systems, 23(NIPS 2010) 2164–2172
43. Stea S, Pickering GL (2019) Optimizing messaging to reduce red meat consumption. Environ Commun 13(5):33–648. https://doi.org/10.1080/17524032.2017.1412994

44. Stephenson MT, Holbert RL (2003) A Monte Carlo simulation of observable versus latent variable structural equation modeling techniques. Commun Res 30(3):332–354. https://doi.org/10.1177/0093650203030003004
45. Sutton R, Barto GA (1998) Reinforcement learning: an introduction. IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council 9:1054
46. Vainio A, Irz X, Hartikainen H (2018) How effective are messages and their characteristics in changing behavioural intentions to substitute plant-based foods for red meat? the mediating role of prior beliefs. Appetite 125:217–224. https://doi.org/10.1016/j.appet.2018.02.002
47. Weller KE, Greene GW, Redding CA, et al. (2014) Development and validation of green eating behaviors, stage of change, decisional balance, and self-efficacy scales in college students. J Nutr Educ Behav 46(5):324–333. https://doi.org/10.1016/j.jneb.2014.01.002
48. Wood W (2000) Attitude change: persuasion and social influence. Annu Review Psychol 51(1):539–570. https://doi.org/10.1146/annurev.psych.51.1.539
49. Zur I, Klöckner CA (2014) Individual motivations for limiting meat consumption. Br Food J 116(4):629642. https://doi.org/10.1108/BFJ-08-2012-0193

Affiliations

Valentina Carfora[2] · Francesca Di Massimo[1] · Rebecca Rastelli[1] · Patrizia Catellani[2] · Marco Piastra[1]

Francesca Di Massimo
francesca.dimassimo01@universitadipavia.it

Rebecca Rastelli
rebecca.rastelli01@universitadipavia.it

Patrizia Catellani
patrizia.catellani@unicatt.it

Marco Piastra
marco.piastra@unipv.it

1 Computer Vision and Multimedia Lab,Università degli Studi di Pavia, Pavia, Italy

2 Dipartimento di Psicologia, Università Cattolica di Milano, Milan, Italy