Emotional Features of Interactions with Empathic Agents

Claudia Greco
University of Campania, Italy
claudia.greco@unicampania.it

Carmela Buono
University of Campania, Italy
carmela.buono@unicampania.it

Pau Buch-Cardona
University of Barcelona, Spain
pau.buch@gmail.com

Gennaro Cordasco
University of Campania, Italy
gennaro.cordasco@unicampania.it

Sergio Escalera
University of Barcelona, Spain
sergio.escalera.guerrero@gmail.com

Anna Esposito
University of Campania, Italy
anna.esposito.it

Anais Fernandez
Association E-Seniors, France
anais.fernandez95@gmail.com

Daria Kyslitska
Association E-Seniors, France
daria.kyslitska@gmail.com

Maria Stylianou Kornes
University of Oslo, Norway
m.s.korsnes@psykologi.uio.no

Cristina Palmero
University of Barcelona, Spain
c.palmero.cantarino@gmail.com

Jofre Tenorio Laranga
Osatek, Spain
jtenorio@osatek.eus

Anna Torp Johansen
University of Oslo, Norway
toanna@ous-hf.no

Maria Inés Torres
University of the Basque Country, Spain
manes.torres@ehu.eus

Abstract

The current study is part of the EMPATHIC project, whose aim is to develop an Empathic Virtual Coach (VC) capable of promoting healthy and independent aging. To this end, the VC needs to be capable of perceiving the emotional states of users and adjusting its behaviour during the interactions according to what the users are experiencing in terms of emotions and comfort. Thus, the present work focuses on some sessions where elderly users of three different countries interact with a simulated system. Audio and video information extracted from these sessions were examined by external observers to assess participants' emotional experience with the EMPATHIC-VC in terms of categorical and dimensional assessment of emotions. Analyses were conducted on the emotional labels assigned by the external observers while participants were engaged in two different scenarios: a generic one, where the interaction was carried out with no intention to discuss a specific topic, and a nutrition one, aimed to accomplish a conversation on users' nutritional habits. Results of analyses performed on both audio and video data revealed that the EMPATHIC coach did not elicit negative feelings in the users. Indeed, users from all countries have shown relaxed and positive behavior when interacting with the simulated VC during both scenarios. Overall, the EMPATHIC-VC was capable to offer an enjoyable experience without eliciting negative feelings in the users. This supports the hypothesis that an Empathic Virtual Coach capable of considering users' expectations and emotional states could support elderly people in daily life activities and help them to remain independent.

1. Introduction

The world population will continue to age, and it is expected that one in six people in the world will be over age 65 by 2050 [1]. Along with aging, the likelihood of developing chronic illness also increases, and this means that the elderly, often frail individuals, need external support to cope with daily activities. However, there is a shortage of caregivers available to provide these continuous services. Indeed, Kon and colleagues [2] showed that the number of

caregivers in 2017 was of 7.2 million and it is expected to increase to around 12.9 million by 2035, which is significantly less compared to 101 million of older people care-dependent [3]. An important role is that of family caregivers: family carers are the main source of support of older dependents [4]. The available estimates of the number of informal caregivers ranges from 10% up to 25% of the total population in Europe [4]. Most older adults prefer to live alone instead of staying in a health institution, which does not facilitate the supportive role of family caregivers. Considering that most of the support is provided by family members and that older people prefer to live alone and not in a health institution, it is important to make the elderly independent by promoting active ageing. Active ageing is defined by the World Health Organization as "is the process of optimizing opportunities for health, participation and security in order to enhance quality of life as people age" [5]. Active ageing challenges the common notion that older age is characterized by passivity and dependency. This process aims to extend healthy life expectancy and quality of life for all older people, including those who are frail, and disabled [5]. Principles promoting active ageing are [6]: health (physical, mental and social well-being), participation (activities related to societal spheres such as employment, politics, education, the arts and religion) and security (activities that guarantee the protection, dignity and care). The elderly need physical assistance (eyesight/hearing, mobility, self-care activities) [7], cognitive support (memory loss, early phase dementia and cognitive decline) [8], emotional and social support (isolation, low self-esteem, lack of emotional support and psychological distress also burden the lives of older people) [9]. Information and Communication Technologies (ICT) aid older people to stay active members helping them remain independent [10]. Therefore, technologies play an important role in the well-being and quality of life for elderly people. In particular, intelligent virtual assistants have been considered to provide elderly people independence with daily life assistance [11]. Some studies showed the role of empathy in human-agent interaction: what makes the difference in the assistance provided by a virtual agent is the emotional empathy perceived by the elderly [12]. The EMPATHIC (Empathic, Expressive, Advanced Virtual Coach to Improve Independent Healthy-Life-Years of the Elderly) cite [13] project aims to develop an emotionally expressive EMPATHIC-virtual coach (EMPATHIC-VC) to assist seniors in their daily activities to promote active and independent aging. The EMPATHIC project aims the use of technology aims to improve the quality of life by involving senior users in a healthy lifestyle regarding diet, physical activity and social interactions. By improving the quality of life of the elderly, the virtual agent also reduces the burden of family caregivers and responds to the increasing demand of caregiving figures.

The EMPATHIC-VC is able to perceive and identify users' social and emotional states by multi-modal face, eye gaze and speech analytics modules [13]. In this framework, this article aims to analyze the interaction between EMPATHIC-VC and users (cf Figure 1), focusing mainly on the external observation, provided by expert annotators trained for emotion recognition. In the context of developing an EMPATHIC-VC, it is important to assess how users are perceived by an external observer to the extent that these same users' emotions should be detected by the agent. Hence, the present work is based on the experimental sessions, during which senior users carried out an interaction with a Wizard of Oz (WOZ) driven, simulated system [14]: older people believed they were interacting with an autonomous machine while the system was actually operated by a human being. Data collected from these experimental sessions were analyzed in order to test whether a VC able to perceive the interacting users and adapt to their emotional states would provide an experience characterized by positive feelings and whether users' emotional states would modify in the course of the dialogues or due to conversations' topics.

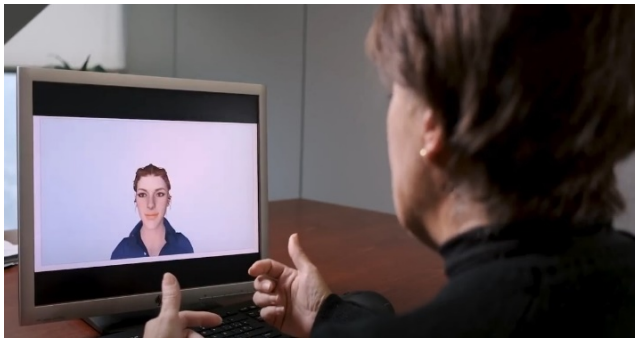

Figure 1. An older adult user while interaction with the EMPATHIC virtual coach.

2. Materials and method

Participants

The inclusion criteria for the study were: 1) female or male older than 65 years; 2) living independently; 3) being able to read, write and speak their mother language. Data were collected for Norway, France and Spain. The total sample is of 131 subjects divided as follows:

- Norway: 31 participants (10 males and 21 females, mean age= 74.98; SD= 5.56)
- France: 33 participants (14 males and 19 females, mean age= 74.88; SD = 8.56)
- Spain: 67 participants (19 males and 48 females, mean age= 69.60; SD = 5.77)

Procedure

Participants signed informed consent (formulated according to the European laws about privacy and data protection) which informed them of the aims and the confidentiality and anonymity nature of the data. As a first step, participants were asked to complete two health questionnaires: Geriatric Depression Scale (GDS) [15] and World Health Organization Quality of Life (WHQOOL-26) [16]. After that, participants were asked to sit in front of a laptop equipped with a webcam, a microphone and mobile connection. Then they were logged into a secure session guaranteeing their anonymity. According to their personal preference, they chose one of five possible virtual agents (3 females and 2 males) as their virtual coach (VC). Details regarding the participants' preferences of virtual agents are provided in [21]. After their choice, participants were required to complete two dialogues with the VC, related to two different scenarios. The first scenario was defined as generic, since the interaction was carried out without any intention to discuss a specific conversational argument and served as an introduction to the system; the second one, named nutrition scenario, aimed to achieve a goal-directed conversation about the users' nutritional habits. The goal could have been eating more vegetables or planning to eat healthier. The system's actions are actually controlled by a human (i.e., the wizard), which is usually located in a separate room connected to the experimental setting through a connection. In this study, the procedure required two experimenters: the wizard and a supervisor that informed and supported the participant with all the required tasks, as the informed consent, the questionnaires, and the debriefing. However, the participants were left alone during the conversations with the EMPATHIC coach in order to avoid any influence of the supervisor on the dialogues. After the conversations with the VC, participants were asked to give her/his feedback of the experience by filling the EMPATHIC Virtual Agent Acceptance Questionnaire (VAAQ) [17], that explores participants' satisfaction in interacting with virtual agents, and the System Usability Scale (SUS) [18] which provides a global view of subjective assessments of usability. Finally, they had to express their perceived feelings while interacting with the VA according to a list of emotional labels identified by the partners of the EMPATHIC project. Both scenarios' conversations were recorded for all participants and audio and video data were extracted from the recordings of the experimental session. In order to assess participants' emotional experience with the EMPATHIC coach, audio and video information gathered from the French, Norwegian and Spanish participants' recorded sessions were separately examined by French, Norwegian or Spanish native external annotators, respectively. The audio information was evaluated by three annotators, while the video information was evaluated by two others. More specifically, French annotators were three males and two females (mean age= 35.6; S.D=15.93), Norwegian annotators were all females (mean age=33; S.D=7.84), and Spanish annotators were four males and one female (mean age=28.8, S.D.=3.63). Annotation was carried out in terms of both categorical and dimensional models of emotions. The model used for the dimensional assessment was the VAD (Valence, Arousal and Dominance [19]), according to which emotional states are defined by three dimensions: 1) valence, which refers to how pleasant or unpleasant the individual feels about something; 2) arousal, which refers to how excited or calm the person feels; 3) dominance, which refers to the sense of control over the situation. The VAD model was used to assess only the speech signal, with no video: annotators had to evaluate participants' voice while conversing with the VC by assigning to their speech one of the following labels in each dimension:

- Arousal (bodily activation): excited/activated, slightly excited, neutral/calm
- Dominance: rather dominant / controlling the situation, neither dominant nor intimidated, rather intimidated /defensive
- Valence (pleasantness): positive, neither positive nor negative, negative

For what concerns the categorical model, annotators had to label the participants' speech with the following emotional labels: "Calm", "Sad", "Amused", "Puzzled" and "Anxious". Only the categorical model was used to analyze the video data, which consisted in the emotional labelling of the video recordings without hearing the audio. In this case, the categorical labels of the emotion dimension assigned to the video information were: "Sad", "Angry", "Surprised", "Amused", "Pensive", "Other", "Neutral".

Data analysis

SPSS 21.0 statistical software was used to analyze audio and video information. Raw data were elaborated by following a procedure consisting in two steps:

1. the entire timeline of each interaction was divided into five temporal segments of equal size to analyze the trend of the features along the time. Temporal segments have different sizes for different participants.
2. the frequency of each emotional label was separately computed for every emotional dimension and temporal segment.

The statistical analyses were firstly performed on the data collected on each country and then on all countries' data aggregated into a single datasheet, with the aim to detect putative differences between Spanish, Norwegian and

	T1		T2		T3		T4		T5	
	Gen.	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**
Arousal										
Excited	0	0	0	0	0	0	0	0	0	0
Slightly excited	29.23	20.87	15.80	20.87	20.74	24.34	19.29	18.51	23.83	26.92
Neutral/ Calm	70.77	73.07	84.20	79.13	79.26	72.64	80.71	75.43	76.19	70.07
Dominance										
Rather dominant/ Controlling the situation	11.48	5.26	4.75	4.50	3.70	4.86	5.11	7.86	7.46	9.17
Neither dominant nor intimidated	88.5	88.3	94.72	95.5	96.3	92.11	94.89	86.08	92.41	87.8
Rather intimidated/ Defensive	0	0.38	0.52	0	0	0	0	0	0.13	0
Valence										
Positive	33.62	22.70	13.94	12.57	19.3	11.37	18.18	7.63	24.51	29.87
Neither positive nor negative	65.74	70.71	83.18	86.44	76.49	84.92	81.14	86.17	70.97	65.92
Negative	0.63	0.54	2.88	0.99	4.2	0.67	0.67	0.13	4.52	1.18
Categorical Emotion										
Calm	92.27	89.45	93.35	96.45	96.86	92.06	94.9	92.17	94.45	88.39
Sad	0	0	0	0	0	0.07	0	0	0	0
Amused	6.23	3.42	1.86	1.05	1.36	1.08	2.45	0.52	2.74	6.04
Puzzled	1.58	1.07	4.93	2.25	1.96	3.76	2.65	1.24	2.8	2.51
Anxious	0	0	0	0.25	0	0	0	0	0	0

Table 1. French participants' occurrence (%) of dimensional and categorical labels perceived in the audio during the users/EMPATHIC coach interaction in the generic (Gen.) and nutrition (Nutr.) scenario.

French samples. Each dimension was separately analysed. For single country's data, the following statistical analyses were considered:

1. Repeated measures ANOVAs to assess differences among annotators' labelling both for categorical and dimensional labels. The within-subjects factors were the annotator and the time.

2. Repeated measures ANOVAs to detect putative changes in the assigned emotional dimensional and categorical emotional dimensions across the five temporal segments, considering time as within-subjects factor and participants' gender as between-subjects factor.

3. Repeated measures ANOVAs to detect whether differences in the assigned labels' occurrences in each temporal segment were affected by the proposed scenarios (generic and nutrition) along which the user/EMPATHIC coach interaction was conducted. The within-subject was the type of scenario and the participants' gender was considered as between-subjects factor.

It should be mentioned that repeated measures ANOVAs have been conducted only on the evaluation provided by the first annotator in all the cases where no significant differences among annotator's labelling were observed. Instead, in those cases where these differences emerged, the repeated measures ANOVAs have been performed both on the first annotator's labelling and on the labelling provided by the annotator differing from the others. However, it must be said that the few cases where significant differences among annotators' labelling were observed did not express contrasting changes in the perceived emotional states of the users, rather a different degree of their positive involvement, as from neutral to calm or vice-versa. For this reason, these differences were not considered relevant changes and the following results will refer to the labelling of participants' emotional status provided by only one of the external annotators. Finally, for cross-countries comparisons, repeated measures ANOVAs analyses were performed on the aggregated data, considering only the labelling provided by the first annotator in each country. The same statistical analysis performed on a single country's data were considered with exception of the repeated measures to assess differences among annotators' labelling. For all the analyses the significant level was set at α = .05 and differences among means were assessed though Bonferroni's post hoc tests. Details of these analyses are provided in the enclosed supplementary material.

3. Results of Audio data analysis

France

Results showed no significant differences in the arousal dimension across the five temporal segments and between generic and nutrition scenarios. As shown in Table 1, participants were mostly perceived as calm. Indeed, the label "Excited" was rarely attributed during both scenarios, meaning that external observers did not perceive that the EMPATHIC coach generated feelings of distress in the users. For what concern the dominance dimension, French

	T1		T2		T3		T4		T5	
	Gen.	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**
Arousal										
Excited	0	0.5	0	0	0	0	0	0	0	0
Slightly excited	5.81	2.96	4.62	1.72	3.94	3.05	1.92	4.4	4.64	3.42
Neutral/ Calm	94.19	96.54	95.39	98.28	96.06	96.95	98.08	95.6	95.35	96.58
Dominance										
Rather dominant/ Controlling the situation	0.73	2.27	0.3	0.42	0.28	2.85	0.62	3.18	0.77	1.24
Neither dominant nor intimidated	99.05	97.71	99.48	99.5	99.72	97.15	99.37	96.82	99.22	98.45
Rather intimidated/ Defensive	0.22	0.02	0.22	0.08	0	0	0	0	0	0.31
Valence										
Positive	3.56	2.73	2.31	2.69	2.94	2.03	0.2	2.16	4.49	5.65
Neither positive nor negative	96.06	95.83	96.27	97.15	97.06	96.97	97.48	95.66	93.63	93.38
Negative	0.37	1.44	1.41	0.16	0	0.99	2.31	2.18	1.87	0.97
Categorical Emotion										
Calm	96.11	95.69	98.32	99.0	95.31	96.12	97.68	94.92	87.82	92.38
Sad	0	0	0	0	0	0	0	0	0	0
Amused	2.9	4.15	0.81	0.74	4.43	3.34	2.32	4.28	12.15	7.4
Puzzled	0.98	0.15	0.88	0.26	0.26	0.54	0	0.81	0	0,22
Anxious	0	0	0	0	0	0	0	0	0	0

Table 2. Norwegian participants' occurrence (%) of dimensional and categorical labels perceived in the audio during the users/EMPATHIC coach interaction in the generic (Gen.) and nutrition (Nutr.) scenario.

participants were mostly perceived as neutral or in control of the situation. The French participants' valence during the conversation with the EMPATHIC coach was mostly perceived as neutral or positive. In detail, some differences also emerged between the type of scenarios: during the first and fourth temporal segments, participants were perceived more frequently as "Positive" in the generic scenario with respect to the nutrition one. Finally, concerning the results of the categorical annotation, they showed the label "Calm" was the most assigned during the generic and nutrition scenarios, followed by the "amused" and "puzzled" ones (see Table 1). No significant differences in labels' occurrence were observed along the temporal segments nor due to the type of scenario.

Norway

Analyses on the Norwegian sample did not reveal significant differences in the occurrences of arousal dimension labels across temporal segments nor between scenarios' type. As occurred for the French participants, Norwegian ones were mostly perceived as calm, suggesting that also for them the EMPATHIC coach aroused a relaxed state during the whole experimental session (see Table 2). The dominance dimension did not significantly change neither across time nor due to the type of scenario. Norwegian participants were mainly perceived as neutral and this indicates that the EMPATHIC Coach was mainly considered as a user-friendly system. Concerning the valence dimension, results of the analysis revealed no significant modification across temporal segments nor due to the type of scenario in the occurrences of the labels. As reported in Table 2, Norwegian participants were mostly perceived as neutral in both scenarios, suggesting that interacting with the EMPATHIC coach was considered a pleasant experience. Finally, regarding the categorical model, the label "Calm" was the most assigned during the generic and nutrition scenarios (see Table 2), though, results showed a decrease of this label occurrence at the end for both generic and nutrition scenarios, and an increase of label "amused" at the end of the generic scenario compared to its beginning.

Spain

Arousal dimension in the Spanish sample did not modify across time and Spanish participants were mostly perceived as "neutral/calm" (see Table 3). The comparison between the generic and nutrition scenarios revealed a slight increase of the "slightly excited" label in the generic with respect to the nutrition scenario. Overall, results indicated that the emotional experience of participants while conversing with the EMPATHIC coach was mainly characterized by a relaxed state. Furthermore, Spanish participants were mostly perceived as "neither dominant nor intimidated" or as "rather dominant/controlling the situation" during both generic and nutrition scenarios. A slight increase in defensiveness was observed at the end of the nutrition scenario, with respect to the generic one, probably because nutrition was a sensitive topic for some participants. Finally, Spanish participants' interaction with the EMPATHIC coach was mainly perceived as "Positive" in both generic and nutrition scenarios (see Table 3). Occurrences of this

	T1		T2		T3		T4		T5	
	Gen.	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**	**Gen.**	**Nutr.**
Arousal										
Excited	0.11	0.13	0	0	0	0	0	0	0	0
Slightly excited	3.14	1.71	3.08	1.12	1.97	1.13	1.67	1.82	2.74	2.59
Neutral/ Calm	96.75	97.95	96.92	98.88	94.53	97.39	95.35	96.68	95.66	92.93
Dominance										
Rather dominant/ Controlling the situation	1.96	2.95	2.24	5.8	5.42	5.05	4.87	5.41	2.79	4.75
Neither dominant nor intimidated	97.61	96.48	97.76	94.06	93.08	92.99	92.14	91.87	95.72	92.2
Rather intimidated/ Defensive	0.43	0.57	0	0.97	0	1.96	0	2.72	0	1.55
Valence										
Positive	48.54	77.59	63.87	78.74	69.73	72.31	71.27	71.49	66.08	63.99
Neither positive nor negative	50.8	21.96	35.84	21.03	29.35	24.29	25.89	26.02	32.14	31.88
Negative	0.57	0.58	0.43	0.16	0.33	1.91	0	0.99	0.16	1.14
Categorical Emotion										
Calm	95.51	96.42	93.16	95.82	93.63	90.86	91.99	91.58	93.27	86.77
Sad	0	0	0.26	0.09	0.44	1.11	0.6	0.21	0	1.7
Amused	1.94	1.78	2.35	2.22	2.09	0.98	1.45	2.64	4.21	5.13
Puzzled	2.19	1.84	3.77	1.87	2.0	5.22	3.51	3.32	1.03	3.38
Anxious	0.36	0.05	0.45	0	0.3	0.34	0	0.75	0	0.03

Table 3. Spanish participants' occurrence (%) of categorical emotion labels perceived in the video during the users/EMPATHIC coach interaction in the generic (Gen.) and nutrition (Nutr.) scenarios.

label tended to increase during the generic scenario, while their frequency remained consistently high during the nutrition one. Concerning the categorical annotation, averaged percentages reported in Table 3 showed an increase at the end of the generic scenario of the label "amused" and a decrease of the label "puzzled". Instead, in the nutrition scenario, the label "amused" increased at the end compared to its beginning. Some differences due to the type of scenario also emerged: Spanish participants were perceived more frequently as "puzzled" in the generic scenario with respect to nutrition one, during the central temporal segments. Probably, this was due to the fact that the generic scenario was always the first one presented to the participants, and they may have not been familiar with the system yet. Overall, these findings suggest that participants were entertained by the EMPATHIC coach.

Cross-countries comparisons

Concerning the arousal dimension, results showed that French participants were perceived more frequently as "slightly excited" and less frequently as "neutral" than Spanish and Norwegian participants during both the generic and nutrition scenarios. According to these findings, French participants were slightly less relaxed and more stimulated by the conversations with the EMPATHIC coach. Concerning the dominance dimension's labels, results showed that French participants were associated with the highest occurrence of the "dominant/controlling the situation" label, followed by Spanish and Norwegian participants. Occurrences of the neutral label followed the opposite direction. Taken together, the results show that, regardless of the type of scenario, the participants were perceived as being able to successfully manage the interaction with the EMPATHIC coach and experienced no feelings of annoyance or discomfort toward the system. Regarding the valence dimension, cross-countries comparisons of emotional labels' occurrences revealed that Spanish participants were associated with the highest occurrence of the "Positive" label, followed by French and Norwegian ones in both generic and nutrition scenarios. Instead, Norwegian participants were mainly perceived as "neutral", probably because they were more cautious in interacting with the EMPATHIC system. Besides these slight differences, these findings suggest that the emotional experience of all users while conversing with the EMPATHIC coach was mainly positive. Finally, regarding the categorical model, results showed that during the generic scenario, Norwegian participants were perceived more frequently as "amused" and less frequently as "puzzled" compared to their Spanish and French counterparts, while in the nutrition scenario, Norwegian participants were perceived less frequently as "puzzled" only with respect to Spanish participants. Besides these slight differences, results suggest that, regardless of the country, the EMPATHIC coach has been perceived as a friendly system and participants' emotional experience while conversing with the EMPATHIC coach was basically never characterized by negative emotions.

	T1		T2		T3		T4		T5	
	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.
Categorical emotion										
Sad	0	0	0	0	0	0	0	0	0.15	0
Angry	0.41	0	0	0	0.04	0	0.49	0	0	0
Surprised	0.43	0	0.15	0.12	0.14	0	0.13	0	0	0
Amused	10.84	4.69	4.23	2.66	6.4	3.1	5.51	2.1	9.91	5.08
Pensive	8.52	8.35	20.74	14.67	19.55	9.47	19.79	9.91	16.54	4.04
Other	0.21	1.6	0.65	3.59	0.76	2.79	0.9	1.76	0.54	2.43
Neutral	79.58	85.3	74.24	78.94	73.1	84.64	73.17	86.22	72.85	88.44

Table 4. French participants' occurrence (%) of categorical emotion labels perceived in the video during the users/EMPATHIC coach interaction in the generic (Gen.) and nutrition (Nutr.) scenarios.

4. Results of video data analysis

France

According to the averaged percentages reported in table 4 the most assigned label was "Neutral" in both generic and nutrition scenarios. However, results revealed some differences between the two scenario types. Participants were perceived more frequently as "Amused" and "Pensive" during the generic scenario with respect to the nutrition one along the five temporal segments in which the interaction develops. On the contrary, occurrences of the label "Neutral" were higher in the nutrition scenario compared to the generic one. These differences may be ascribed to the fact that the nutrition scenario required participants to stay focused on the nutritional topics to achieve the goal of implementing a correct daily regime. For this reason, participants were perceived more emotionally neutral in this type of scenario compared to that with no goal requirements. Even though, the percentages of occurrences of each emotional categorical label indicate that participants were absorbed by the dialogues and seem to enjoy conversing with the EMPATHIC coach. This was supported by the very low occurrences of "Sad" and "Angry" labels during all the temporal segments, which suggest that participants quite never experienced negative emotional states.

Norway

As shown in table 5, Norwegian participants were mostly perceived as "Neutral" while interacting with the EMPATHIC coach. Occurrences of the label "Amused" increased at the end of both the interactional scenarios, compared to the beginning. This increase indicates that participants enjoyed conversing with the system. Furthermore, Norwegian participants were perceived more frequently as "pensive" in the nutrition scenario and more frequently as "neutral" in the generic one, indicating that conversations aiming to accomplish a specific goal led participants to be more attentive to the topic of conversation which, in this case, concerned appropriate nutrition standards. As occurred for French participants, also Norwegian ones were rarely perceived as "sad" or "angry". The lack of perceived negative emotions is in line with results obtained from the speech analysis, suggesting that the interaction with EMPATHIC coach was characterized by positive feelings and the users were absorbed by the conversations with the system.

Spain

Averaged percentages reported in Table 6 showed that the Spanish participants were perceived more frequently as "Neutral". In detail, comparisons between scenarios revealed some differences: Spanish users were perceived more frequently as neutral during the nutrition compared to the generic scenario. Probably, conversations on a specific goal led Spanish participants to remain in a neutral emotional state to stay focused on the goals to be achieved by the conversation's topic. Furthermore, results showed that the "pensive" and "neutral" occurrence labels varied slightly (increasing and decreasing) across temporal segments (see Table 6). In general, participants were mostly perceived to be alert or in a neutral emotional state, indicating that they felt no distress while conversing with the EMPATHIC coach.

Cross-countries comparisons

Results showed some differences due to the country: Spanish participants were perceived more frequently as "neutral" compared to Norwegian and French participants; whereas Norwegian and French participants were perceived more frequently as "amused" compared to their Spanish counterparts. In the nutrition scenario, French participants were perceived less frequently as "pensive" compared to French and Norwegian participants and more frequently as "amused" compared to Spanish participants. These results suggest that the EMPATHIC virtual coach did not elicit negative feelings making all countries' participants comfortable.

	T1		T2		T3		T4		T5	
	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.
Categorical emotion										
Sad	0	0	0	0	0	0	0	0	0	0
Angry	0.63	0	0	0	0	0	0	0	0	0
Surprised	0.38	0.02	0.06	0	0	0	0.41	0	0.15	0.16
Amused	6.53	6.81	4.3	2.89	6.15	4.22	4.12	5.61	11.61	11.43
Pensive	6.76	17.84	15.62	25.9	17.72	22.27	18.58	20.16	11.31	10.89
Other	2.71	0.02	2.81	0	3.07	0.08	3.03	0	2.88	0.22
Neutral	83.56	75.3	77.21	71.21	73.05	73.42	73.85	74.22	74.04	77.29

Table 5. Norwegian participants' occurrence (%) of categorical emotion labels perceived in the video during the users/EMPATHIC coach interaction in the generic (Gen.) and nutrition (Nutr.) scenarios.

	T1		T2		T3		T4		T5	
	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.	Gen.	Nutr.
Categorical emotion										
Sad	0	0	0	0	0	0	0	0	0	0
Angry	0.63	0	0	0	0	0	0	0	0	0
Surprised	0.38	0.02	0.06	0	0	0	0.41	0	0.15	0.16
Amused	6.53	6.81	4.3	2.89	6.15	4.22	4.12	5.61	11.61	11.43
Pensive	6.76	17.84	15.62	25.9	17.72	22.27	18.58	20.16	11.31	10.89
Other	2.71	0.02	2.81	0	3.07	0.08	3.03	0	2.88	0.22
Neutral	83.56	75.3	77.21	71.21	73.05	73.42	73.85	74.22	74.04	77.29

Table 6. Spanish participants' occurrence (%) of categorical emotion labels perceived in the video during the users/EMPATHIC coach interaction in the generic (Gen.) and nutrition (Nutr.) scenarios.

5. Discussion

The present work is part of the EMPATHIC Research & Innovation project, which aims to explore and innovate new interaction paradigms and build human-computer platforms entailing Personalized Virtual Coaches that can support elderly people live independently. In order to do this, the EMPATHIC-VC has to be able to perceive emotional states of users and adjust its behaviour during the interactions as a function of what the users are experiencing in terms of emotions and comfort. The present study attains to this EMPATHIC project objective by analysing the dimensional and categorical emotional assessment carried out by external observers on audio and video information extracted from experimental sessions' recordings. According to the results, interacting with the EMPATHIC coach **did not elicit negative feelings in the users.** Indeed, both dimensional and categorical assessment of participants' emotional experience revealed that they tend to be perceived relaxed and not intimidated by the system. Participants considered the EMPATHIC coach as user-friendly and felt positively involved in the conversations. Supporting this, labels associated with negative emotional states were rarely assigned by annotators during the users interaction, suggesting that participants from all countries mostly experienced positive emotions or they were perceived in a neutral emotional state while conversing with the EMPATHIC coach. Results of the multi-modal analyses on both audio and video data support the hypothesis that an empathic virtual agent could sustain senior users in the maintenance of a healthy lifestyle, and provide elderly people independence with daily life assistance [11]. These results also show that the empathy plays a fundamental role by making the interaction between users and VC more natural and increasing the satisfaction of older people [20]. These promising results suggest that the use of technology plays an important role in the well-being and quality of life for elderly people, by favouring the active ageing process and also reducing the burden of family caregivers. Future directions should persevere in the development of virtual agents able to perceive and identify users' social and emotional states.

Acknowledgments

The research leading to these results has received most of its funding by the project ANDROIDS funded by the program V:ALERE 2019 Università della Campania "Luigi Vanvitelli", D.R. 906 del 4/10/2019, prot. n. 157264,17/10/2019. Other fundings have been received by the EU H2020 research and innovation program under grant agreement N. 769872 (EMPATHIC) and N. 823907 (MENHIR), and the project SIROBOTICS that received funding from Italian MIUR, PNR 2015-2020, D.D. 1735, 13/07/ 2017.

References

[1] United Nations Department of Economic and Social Affairs: Population Division 2019: World

Population Prospects. New York, USA. https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf 1

[2] B. Kon, A. Lam, and J. Chan, "Evolution of Smart Homes for the Elderly". *Proceedings of the 2017 International World Wide Web Conference Committee (IW3C2), 2017.* DOI:10.1145/3041021.3054928. 1

[3] A. Banerjee and R. Sadana, "Integrated care for older people (ICOPE): from guidelines to demonstrating feasibility," *Journal of Frailty & Aging*, 10, 84–85, 2021. DOI:10.14283/jfa.2020.40. 2

[4] V. Zigante, "Informal care in Europe: Exploring formalisation, availability and quality", *European Commission, London School of Economics and Political Science*, 2018. https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8106&furtherPubs=yes 2

[5] World Health Organization. *Active ageing : a policy framework.* World Health Organization, https://apps.who.int/iris/handle/10665/67215, 2002. 2

[6] P. Stenner, T. McFarquhar, and A. Bowling, "Older people and 'active ageing': Subjective aspects of ageing actively," Journal of Health Psychology, vol. 16, no. 3, pp. 467–477, 2011. DOI: 10.1177/1359105310384298. 2

[7] K. Walters, S. Iliffe, S. S. Tai, and M. Orrell, "Assessing needs from patient, carer and professional perspectives: the Camberwell Assessment of Need for Elderly people in primary care," Age and Ageing, vol. 29, no. 6, pp. 505–510, 2000. DOI: 10.1093/ageing/29.6.505. 2

[8] J. E. Graham, K. Rockwood, B. L. Beattie, R. Eastwood, S. Gauthier, H. Tuokko, and I. McDowell, "Prevalence and severity of cognitive impairment with and without dementia in an elderly population," The Lancet 349(9068), 1793–1796, 1997. DOI:10.1016/S0140-6736(97)01007-6. 2

[9] M. Luppa, T. Luck, S. Weyerer, H.-H. Konig, E. Brahler, and S. G. Riedel-Heller, "Prediction of institutionalization in the elderly. A systematic review," Age and Ageing, 39(1), 31–38, 2010. DOI: 10.1093/ageing/afp202. 2

[10] L. Brinkschulte, N. Mariacher, S. Schlögl, M. I. Torres, R. Justo, J. M. Olaso, A. Esposito, G. Cordasco, G. Chollet, C. Glackin, et al. "The EMPATHIC project: building an expressive, advanced virtual coach to improve independent healthy-life-years of the elderly" *Proceedings of the Smarter Lives: SMARTER LIVES 2018: digitalisation and quality of life in the ageing society, Feb 2018, Innsbrück, Austria. pp. 36 - 52*. https://arxiv.org/ftp/arxiv/papers/2104/2104.13836.pdf 2

[11] R. Yaghoubzadeh, M. Kramer, K. Pitsch, and S. Kopp, "Virtual Agents as Daily Assistants for Elderly or Cognitively Impaired People," in Intelligent Virtual Agents: 13th International Conference, IVA 2013, Edinburgh, UK, August 29-31, 2013: proceedings, Berlin: Springer, 2013, pp. 79–91. DOI: 10.1007/978-3-642-40415-3_7 2, 8

[12] H. Prendinger, J. Mori, and M. Ishizuka, "Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game," International Journal of Human-Computer Studies, vol. 62, no. 2, pp. 231–245, 2005. DOI:10.1016/j.ijhcs.2004.11.009 2

[13] Torres M.I., Olaso J.M., Montenegro C., Santana R., Vázquez A., Justo R., Lozano J.A., Schlögl S., Chollet G., Dugan N., Irvine M., Glackin N., Pickard C., Esposito A., Cordasco G., Troncone A., Petrovska-Delacretaz D., Mtibaa A., Hmani M.A., Korsnes M.S., Martinussen L.J., Escalera S., Cantariño C.P., Deroo O., Gordeeva O., Tenorio-Laranga J., Gonzalez-Fraile E., Fernandez-Ruanova B., Gonzalez-Pinto A. "The empathic project: mid-term achievements", *Proceedings of the 12th ACM international conference on pervasive technologies related to assistive environments (PETRA)*, 629–638, 2019. DOI:10.1145/3316782.3322764 2

[14] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, "Wizard of Oz studies — why and how," *Knowledge-Based Systems*, vol. 6(4), 258–266, 1993. DOI:10.1145/169891.169968 2

[15] J. A. Yesavage and J. I. Sheikh, "9/Geriatric Depression Scale (GDS)," Clinical Gerontologist, 5(1-2), 165–173, 1986. DOI:10.1300/J018v05n01_09 3

[16] W. Group, "Development of the WHOQOL: Rationale and Current Status," International Journal of Mental Health, 23(3), 24–56, 1994. DOI:10.1080/00207411.1994.11449286 3

[17] A. Esposito, T. Amorese, M. Cuciniello, I. Pica, M. T. Riviello, A. Troncone, G. Cordasco, and A. M. Esposito, "Elders prefer female robots with a high degree of human likeness," *23rd International Symposium on Consumer Technologies (ISCT)*, 243–246, 2019. doi: 10.1109/ISCE.2019.8900983. 3

[18] J. Brooke, "SUS: A 'Quick and Dirty' Usability Scale," Usability Evaluation In Industry, pp. 207–212, 1996. http://www.tbistafftraining.info/smartphones/documents/b5_during_the_trial_usability_scale_v1_09aug11.pdf 3

[19] H. Gunes and M. Pantic, "Automatic, Dimensional and Continuous Emotion Recognition," International Journal of Synthetic Emotions (IJSE), vol.1, no.1, pp. 68–99, 2010. DOI: 10.4018/jse.2010101605. 3

[20] N. Vannini, S. Enz, M. Sapouna, D. Wolke, S. Watson, S. Woods, K. Dautenhahn, L. Hall, A. Paiva, E. André, R. Aylett, and W. Schneider, "'FearNot!': a computer-based anti-bullying-programme designed to foster peer intervention," European Journal of Psychology of Education, 26(1), 21–44, 2010. DOI: 10.1007/s10212-010-0035-4 8

[21] R. Justo Blanco, L. Ben Letaifa, C. Palmero, E. Gonzalez Fraile, A. Torp Johansen, A. Vazquez, G. Cordasco, S. Scholg, B. Fernandez Ruanova, M. Silva Rodrigues, S. Escalera, M. de Velasco, A. Esposito, M. Korsnes, M.I. Torres, "Analysis of the Interaction between Elderly People and a Simulated Virtual Coach." In Journal of Ambient Intelligence and Humanized Computing, DOI: 10.1007/s12652-020-01983-3, 2020. 3