Towards Personality-based Assistance in Human-Machine Interaction

Thi-Hai-Ha Dang and Adriana Tapus

Abstract—In HRI, many researches emphasize the impact of the human user’s personality (expressed mainly through the Extroversion dimension) over the perception of the robot’s behavior. In our experiment, where participants interacted/used a novel driving assistance system, we focused on analyzing the role of each BigFive Personality dimension in people’s task performance and in their reaction towards the vocal assistance system. The results show that three of the BigFive Personality dimensions (i.e., Extroversion, Openness, and Agreeableness) present a certain influence towards participants’ performance. We also found that inexperienced users made better progress in using the driving interface than the experienced users. A detailed discussion about the contribution of our current work and future perspectives is provided.

I. INTRODUCTION

An efficient interaction is an interaction that adapts to the user’s preferences and promotes a high satisfaction/enjoyment of the user towards the task performance [3], [6]. User’s preferences may depend on the user’s expertise about the system, and also on his/her personality. In this paper, we discuss about why and how to take into account the user’s personality so as to provide an efficient assistance to the human user.

Inspired from several findings in psychology, user’s personality has become widely studied in Human-Robot Interaction (HRI). HRI researchers are interested in applying psychology knowledge into their works and have demonstrated convincing results, such as the works of [1], [11], [16], [14]. However, most of these works privileged only the Extroversion dimension of the user’s personality trait while somehow avoid discussing on other personality dimensions. We believe that user’s personality should be considered in all of its dimensions in order to establish a consistent relationship between human users and the involved robotic systems. In this paper, we present our first attempt to study the user’s complete personality traits (i.e., considering all the traits from the BigFive personality spectrum) in the context of a novel drone driving system. We also analyze the effect on the novel driving interface of the user’s personal car driving experience (i.e., their car driving experience measured in years).

The paper is organized as follows. We begin by resuming the state-of-the-art of the research attempts to apply psychology findings about human personality into Human-Robot Interaction paradigm (See Section II). We continue by presenting the design of our driving experiment used for measuring the influence of each personality dimension of the user on his/her task performance and how a vocal assistance can help to improve the user’s performance (see Section III). We also provide an intensive discussion on our experimental results (see Section IV). We conclude the paper by summarizing the main points of our contribution (see Section V).

II. PERSONALITY IN HUMAN-MACHINE INTERACTION: STATE OF THE ART

In behavioral psychology, personality refers to the patterns of thoughts, feelings, social adjustments, and behaviors consistently exhibited by an individual over time that strongly influences his/her expectations, self-perceptions, values and attitudes, and predicts his/her reactions to people, problems, and stress [10], [20].

Lots of researches in psychology have been carried out to study the effect of personality in team performance. Works presented in [2], [4] suggest that team performance in creative tasks can be maximized if members’ personality (evaluated in terms of BigFive Personality Traits [8]) meet the optimal pattern consisting of moderate levels of Extroversion, high levels of Openness to Experience, and high levels of Conscientiousness. Moreover, the authors in [12] have studied the effect of group’s personality pattern in cooperative task and found that each dimension of BigFive personality has a different effect on group’s performance in cooperative tasks, such as Extroversion influences tasks that do not enforce very short time constraints, Openness to Experience has impact on search tasks, while Agreeableness was important for tasks where tight collaboration was required.

In HRI domain, researchers focus more and more on user’s personality. Some of these research works are summarized below. The work described in [1] explored the benefits of combining verbal and non-verbal behaviors to generate robot’s personalities appropriately during the interaction with a human peer. The system estimates first the interacting human’s personality traits through a psycho-linguistic analysis of the spoken language, then it uses PERSONAGE natural language generator that tries to generate a corresponding verbal language to the estimated personality traits. Gestures are generated by BEAT toolkit, which performs a linguistic and contextual analysis of the generated language relying on rules derived from extensive research into human conversational behavior. The results showed that individuals preferred to interact more with the robot that had the same personality with theirs. Participants also expressed their preference to the mixed speech-gesture behavior of the robot, saying that the robot’s speech was more engaging and more effective when...
accompanied by appropriate gestures than when no gestures were present.

Lee and collaborators [11] suggest that a robot can motivate people during interaction by using its personality. The authors modeled two kinds of behaviors for the AIBO robot: an extroverted and an introverted behavior (generally modulated in the vocal sound and in the speed of actions during the interaction with the human partner). The participants were asked to play with AIBO robot and evaluate its personality. The obtained results emphasized that participants were more joyful when interacting with the robot that had complementary personality to theirs. Their work also suggested that by changing the robot’s behavior (in this case, its personality), we could change the motivational level of people interacting with the robot.

Another work focusing on socially assistive robotics (SAR), and more precisely on post-stroke rehabilitation therapy [16], examined the effects of robot’s customized behavior on people’s motivation and task performance. The relationship between the extroversion-introversion personality spectrum and the style of encouragement in a rehabilitation task were explored and the role of adapting robot’s behavior to the user’s profile was addressed. The three factor PEN (Psychoticism - Extroversion - Neuroticism) Eysenck Personality model [7] was employed, with a particular focus on the Extroversion dimension. The study showed that users preferred working and interacting with a robot with a similar personality as theirs during the therapy: extroverted users preferred the robot that challenged them during the exercises, while introverted users preferred the robot that praised them.

Furthermore, similar results were also found in Human-Computer Interaction (HCI). In [14], the authors presented an experiment where the influence of personality on human’s task performance was tested. In their experiment, participants were taught to use HyperCard application [13]. By testing several conditions during the experiment (in a 2 x 2 x 2 factorial design as follows: personality of the interface (extroverted/introverted), subjects’ personality (extroverted/introverted), task strength (low and high)), they found that introverted participants made better performance when using introverted interface rather than while employing the extroverted interface. However, this effect was not observed for extroverted participants. In terms of task strength, they found that the extroverted participants realized tasks significantly faster than introverted participants on low task strength, however, no significant difference was found for high task strength. This work also leads to believe that in human-machine interaction, the personality of the machine can influence task performance in a certain manner.

In summary, while behaviour psychology shows effect of personality in all of its dimension, most of the works in computer science are focused on one dimension of the BigFive Personality construct and somehow ignored the possible effects of all the other dimensions (note: BigFive Personality Inventory consists of five dimensions: Extroversion, Openness to Experience, Agreeableness, Conscientiousness, and Neuroticism. Please refer to [9] for further details about Bigfive Personality). We suggest that each dimension of human’s personality has its specific effect in Human-Machine Interaction. To investigate that effect, we conducted an experiment to study the influence of each BigFive dimension towards the users’ performance in a Human-Machine Interaction context.

In the following sections, we present our driving experiment where we investigate the role of a vocal assistance for a driving system and the role of each BigFive Personality dimension on the users’ performance. We also discuss about our experimental results and its possible contribution to promote a better Human-Machine Co-existence.

III. VOCAL ASSISTANCE IN HUMAN-MACHINE INTERACTION

We conducted an experiment where people are invited to drive a PARROT quadrotor drone using the Logitech’s G27 Racing Wheel. In this section, we present the experimental setting, procedures, and results that we obtained.

A. Research questions

As explained before, in this work, we are interested in studying the influence of individual’s BigFive Personality dimensions, including Extroversion, Agreeableness, Conscientiousness, Openness to Experience, and Neuroticism in Human-Machine Interaction. Personality is widely considered as the individual’s patterns of behavior, thoughts, and emotions. In order to build an efficient Human-Machine Interface for a long run, designers should take into account the influence of user’s personality on his/her task performance. Our quest in this experiment is to search for any relationship between the user’s personality and his/her task performance during a designated task given a specific assistance. Additionally, we also want to evaluate how the vocal assistance affect the user’s task performance in order to envision a multi-modal assistance during a Human-Machine Cooperation. Our research questions are stated below:

Research question 1: Does it exist a relationship between user’s personality and his/her task performance in the designated experimental scenario? If yes, how does each BigFive personality dimension relate to the task performance?

Research question 2: Does the addition of the vocal assistance affect the user’s task performance? If yes, how does it affect the user’s performance (i.e., in a positive or a negative way)?

B. Drone driving experiment

1) Experimental Setting: The real-world drone driving system consisted of the Logitech’s G27 Racing Wheel, a computer running Linux and ROS, and a PARROT drone from PARROT company\(^1\). The computer served to map the wheel signals to the drone’s flying commands and to do additional data collection and processing.

The driving environment is a closed room of 4x4 m\(^2\). On the ground there are colored markers serving to define different flying trajectories. Markers are connected by straight

\(^1\)http://www.parrot.com/
lines (serving as flying tracks) (as shown in Fig. 2). We designed two levels of difficulty: easy and difficult. Easy flying trajectory consisted of moving from one marker to another one while staying on the tracks until all marker are visited. Difficult flying trajectory consisted of moving from an initial marker (marker 0) to another marker then going back to the initial marker before moving to another marker. During the flight, the target marker is shown on the screen. This provides information to the driver regarding the next target he/she has to fly to. Images from the below camera of the drone are sent to an image processing program that helps to detect if the drone is on track or not and to detect if the drone is arrived at the target marker or not.

A vocal system has also been implemented and can be activated to additionally assist the driver during the flight. It uses the result from the image processing system to determine the appropriate message. The vocal assistance system is used to warn the driver about the drone position (on/off track) and to announce the next marker when a new targeted marker becomes available. It also helps the driver to keep track of the timing.

During the experiments, we also monitor heart rate and skin conductance of the participants. However, discussion on these physiological signals is not in the scope of this paper.

2) Protocol: Before starting the experiment, the participants were asked to fill-up a pre-study questionnaire about their demographic information. After that, a short introduction about the experiment context and the settings was made. The physiological sensors are then attached to the body of the participant for the real-time measurements of the heart rate and skin conductance signals. The participants are given about 20 minutes to learn to fly the drone with the Logitech’s G27 Racing Wheel and to get used to the driving setting.

The main experiment consisted of 5 minutes of music relaxation and four attempts of driving. Each attempt was associated with a different condition: in terms of difficulty level (easy trajectory vs. difficult trajectory) and vocal assistance (vocal assistance activated vs. vocal assistance not-activated). The order of the four conditions are changed from one participant to another. Trajectories are predefined and announced to the participant before the beginning of each attempt. In each attempt, the participant is asked to drive through the designated trajectory in 5 minutes or less. After each attempt, the participant is asked to answer a questionnaire about his/her emotional impression about the flight before passing to the next attempt.

At the end of the experiment, the participants are asked to fill-up a last questionnaire about their overall impression about the system, which includes how much he/she liked the vocal system (on a 7-point Likert Scale), whether he/she wanted the vocal system to be interactive, and whether he/she wanted to use the vocal assistance if asked to drive the drone one more time. The answers of the two latter questions are either 'yes', 'no', or 'indifferent'.

3) Participants: Twenty-one participants (1 female and 20 males) took part in the experiments; their age varied between 22 and 34, with an average age of 27, all with technical background.

Concerning BigFive Personality, each participant' BigFive Personality Test result consists of five percentile values of the respectively 5 independent dimensions: Extroversion, Conscientiousness, Agreeableness, Neuroticism, and Openness. Each percentile value varies between 0 and 100. In this paper, experimental data is analyzed for each personality dimension independently. For each personality dimension, we divided participants into three groups: (1) the low percentile group in which participants’ percentile value is equal or smaller than 33 in the concerned dimension; (2) the average group in which participants have a percentile between 30 and 66 in the concerned dimension; and (3) the high percentile group in which members’ percentile is equal or greater than 66 in the concerned dimension.

C. Results

In this section, we evaluate the performance of the participants by taking into consideration various variables, such as
as performance and ability, performance and personality, performance and preference towards the vocal assistance. When talking about performance, we are referring to the number of minutes that took the participant to finish a trajectory. For those who could not finish the trajectory in five minutes, the respective performance is fixed to five minutes and is marked as failed for further analysis.

1) Ability vs. Performance: The participants were divided into two groups based on their years of driving: novice drivers who have less than 3 years of driving, and experienced drivers otherwise. In our experiment, we had 10 novice drivers and 11 experienced drivers. Their average performance across the four attempts is presented in Fig. 3.

A repeated-measure ANOVA was used to test the performance difference among the four attempts of each group of drivers. For novice drivers, performance differed significantly across the four attempts, $F(3, 27) = 3.41, p = 0.0316$. Tukey post-hoc analysis indicates that the first attempt has significantly worse performance than the second attempt and the forth attempt with $p < 0.05$. Novice drivers’ performance at the third attempt is not statistically different to other three attempts’ at $p < .05$. For experienced drivers, ANOVA detects no statistical difference in the performance of the four attempts. In summary, novice drivers’ performance improved over time regardless the driving conditions whereas experienced drivers made no significant progress over the four attempts.

2) Vocal Assistance vs. Performance: We now analyze the effectiveness and/or the influence of the vocal assistance on the performance of the participant. Participants’ performance is summarized in Table II. When driving the drone with the vocal assistance activated, participants’ performance was roughly enhanced (i.e., reduced duration for each trajectory). This suggests that the vocal assistance can have a positive impact on participants’ performance. However, no significant difference between performance with vocal assistance and performance without vocal assistance was found with ANOVA.

3) Personality vs. Performance: As mentioned previously (in section III.B.3), experimental data is analyzed for each BigFive personality dimension independently. For each personality dimension, participants are divided into three groups: (1) the low percentile group in which participants’ percentile value is equal or smaller than 33 in the concerned dimension; (2) the average group in which participants have a percentile between 30 and 66 in the concerned dimension; and (3) the high percentile group in which members’ percentile is equal or greater than 66 in the concerned dimension. We used independent measure ANOVA to analyze the between-subjects effects. The statistical test results for our experimental data are presented below:

- **Extroversion vs. Performance**: ANOVA shows that participants with Low Extroversion have significantly the best performance comparing to the participants with Average Extroversion ($F(1, 62) = 5.17, p = .026$) and participants with High Extroversion ($F(1, 62) = 7.51, p = .008$).

- **Openness vs. Performance**: ANOVA shows that the performance of the participants with Low Openness is significantly poorer than that of participants with Average Openness ($F(1, 74) = 5.5, p = .0217$); High Openness group has only 2 members and hence is excluded from the statistical analysis.

- **Agreeableness vs. Performance**: Participants with Low Agreeableness performed significantly better than the participants with High Agreeableness ($F(1, 30) = 8.72, p = .006$). ANOVA detects no statistical difference in the performance of the Average Agreeableness group with respect to the other two groups of Agreeableness.

- **Conscientiousness vs. Performance**: ANOVA detects no statistical difference in the performance between the three groups of Conscientiousness.

- **Neuroticism vs. Performance**: ANOVA detects no statistical difference in the performance between the three groups of Neuroticism.

4) Personality vs. Chance of success: As described in the experimental protocol, each attempt is limited to five minutes. Those who could not finish the designated trajectory were forced to stop and to move on to the next attempt. In this section, we look at how often someone fail to accomplish a trajectory in the five-minute window. The experimental data is presented in Table III, each cell contains the number of failures of each group followed by the size of the group in parentheses.

Average number of failures of each group is graphically presented in Fig. 4. In summary:
TABLE III
NUMBER OF FAILURE BY GROUPS OF PARTICIPANTS IN EACH PERSONALITY DIMENSION

<table>
<thead>
<tr>
<th>Personality Dimension</th>
<th>Low percentile</th>
<th>Average percentile</th>
<th>High percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extroversion</td>
<td>6 (11)</td>
<td>7 (5)</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Openness</td>
<td>10 (9)</td>
<td>3 (10)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>7 (8)</td>
<td>8 (10)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>1 (4)</td>
<td>12 (13)</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>11 (13)</td>
<td>4 (5)</td>
<td>1 (3)</td>
</tr>
</tbody>
</table>

Fig. 4. Average number of failures of each group of personality.

- Average Extroversion participants failed more often than the other groups;
- Low Openness participants have higher rate of failure than the others;
- High Conscientiousness participants have the highest chance of success than other groups (i.e., the lowest rate of failure);
- Low Agreeableness participants have higher chance of success than the other groups;
- High Neuroticism have higher chance of success than the other groups.

5) Effectiveness of the vocal assistance over Personality: We use ANOVA to analyze the significance of the vocal assistance across different personality dimensions (personality groups are conceived as explained earlier in Section III.D.3). We found that the task performance of the participants with High Extroversion was significantly improved when being assisted by the vocal assistance system ($F(1, 9) = 7.58$, $p = .022$). We also found that the participants with Average Openness performed significantly better when being assisted by the vocal assistance system ($F(1, 19) = 7.81$, $p = .012$). No other significant differences were found between the groups of personality on the effect of vocal assistance to performance.

6) Relation between performance and preference of vocal assistance: Correlations between participants’ performance and their preference towards the vocal assistance were also analyzed based on participants’ responses to the final post-experiment questionnaire that we collected. These questions focused basically on how much the participants liked the vocal assistance and their willingness of using the vocal assistance system. To evaluate the correlation between the different factors, we calculated the Pearson Correlation Coefficient between the participants’ answers and their respective performance. We also evaluated the correlation between participants’ answers about the vocal assistance and their driving experience measured in years. To calculate the Pearson Correlation Coefficient, we encode 'yes' as $+1$, 'no' as $-1$, and 'indifferent' as 0. Findings from Pearson Correlation Coefficient test are presented below:

- The longer the duration of the driving, the higher the chance people chose to be assisted by the vocal assistance. The Pearson Correlation Coefficient between duration and respective rating of choosing to be assisted by the vocal assistance is 0.42, which means that there is a weak (positive) correlation between these two criteria (coefficient of determination, $R^2$, is 0.18).
- The longer the duration of the driving, the more people like the vocal assistance. The Pearson Correlation Coefficient between duration and respective rating of liking the vocal assistance is 0.43, which means that there is a weak positive correlation between these two criteria (coefficient of determination, $R^2$, is 0.18).
- More years of driving leads to preference of interactive assistance. The Pearson Correlation Coefficient between years of driving and respective rating of liking an interactive vocal assistance is 0.31, which means that there is a weak (positive) correlation between these two criteria (coefficient of determination, $R^2$, is 0.1).
- Novice drivers like the vocal assistance more than the experienced drivers. The Pearson Correlation Coefficient between years of driving and respective rating of liking the vocal assistance is $-0.34$, which means that there is a weak (negative) correlation between these two criteria (coefficient of determination, $R^2$, is 0.12).
- Novice drivers are more likely to accepted the vocal assistance than the experienced drivers. The Pearson Correlation Coefficient between years of driving and respective rating of liking the vocal assistance is $-0.43$, which means that there is a weak (negative) correlation between these two criteria (coefficient of determination, $R^2$, is 0.18).

IV. DISCUSSIONS

The use of the Logitech wheel (and pedals) as the drone’s controller was completely novel to the participants. Analysis on participant’s performance and their driving experience (measured in years) shows that novice drivers can make progress more easily than experienced drivers. This may be because experienced drivers are used to use the wheel and the pedals in a certain manner, thus having to adapt to a completely new way of using it, was more difficult to them. Less experienced users, on the other hand, can make better progress in using new technological tools.

Moreover, in this experiment, we were also able to test the multi-modal assistance, which combined vision assistance and vocal announcement in the assistance of the system. As explained in the previous section, the combined assistance
(vision and voice) is better than the visual assistance alone. This should be tested on a larger population in order to be statistically proven. Furthermore, future works should also focus on the comparison between the combined assistance and the visual assistance in order to assess importance on each of the assistance means (i.e., visual assistance and vocal assistance).

From the experiment, we also evaluated the effect of participants’ personality on their performance. For instance, Extroversion, Openness, and Agreeableness seem to have a certain influence towards participants’ performance. The interaction between the vocal assistance and the personality was found for Extroversion and Openness dimensions. Hence, we also suggest that studies on Human-Machine Interface/Interaction should take into account the influence of the user’s personality to the task he/she is intended to do in order to optimize the Human-Machine interface/interaction.

Furthermore, Pearson Correlation Coefficient analysis reveals some correlations between participants’ characteristics and their preference towards the vocal assistance. Novice drivers preferred vocal assistance and were eager to use the vocal assistance further. Moreover, the longer people drove in the company of the vocal assistance system, the better they liked it. These results suggest that more assistance is recommended for inexperienced users and for those who have more difficulties doing the designated task. Designers of adaptive Human-Machine assistance should take this information into account.

One of major weaknesses of this experiment is that the population is mostly male, which makes it difficult to generalize to female population. Besides, most participants have technical background, which limits our findings to this specific population. Furthermore, the experiment should be more selective in the sense of having a more balanced distribution on the test population.

V. CONCLUSION

Throughout the paper, we presented our research results on how each of the BigFive Personality dimension affect people’s task performance and their preference towards the multi-modal assistance. We were able to analyze the difference between inexperienced users and experienced users in terms of their performance and their preference towards the assistance system. In particular, our results show that visual plus vocal assistance are more effective than the visual-only assistance. Participant’s personality have a certain effect on the individual performance and on his/her preference towards the vocal assistance, for instance: Low Extroversion or Low Agreeableness affect positively the task performance while Low Openness affects negatively the task performance. Most importantly, our findings show that people with less experience can learn faster and perform better; this is very encouraging for the development of novel HCI/HRI since the use of robotic/automatic systems by novice users is more and more present. However, current results are limited to male and young adults, thus it needs to be investigated on a broader population for further generalization.

ACKNOWLEDGMENT

This work was supported by the French National Research Agency (ANR) through Chaire D’Excellence program 2009 (Human-Robot Interaction for Assistive Applications).

REFERENCES