Personality and habit formation: Is there a link?

Aneta Lisowska[1], Shiri Lavy[2], Szymon Wilk[1], and Mor Peleg[3]

[1] Institute of Computing Science, Poznań University of Technology, Poznań, Poland
{aneta.lisowska, szymon.wilk}@put.poznan.pl
[2] Department of Leadership and Policy in Education University of Haifa, Haifa, Israel
[3] Department of Information Systems, University of Haifa, Haifa, Israel
morpeleg@is.haifa.ac.il

Abstract. We aim to develop a virtual coaching (VC) system which facilitates cancer patients with the formation of positive health habits. The VC should adjust content and delivery of recommendations based on user's preferences. We conducted a survey that elicited potential VC users' perceptions about the appropriate context for notification, aspects driving their motivation, and their ability to perform a selected activity in different circumstances. We analyze 173 survey responses, identify different types of users and suggest how to adjust application content and notification timing to increase their susceptibility to non-medication health interventions.

Keywords: Fogg Behaviour Model, Personality Traits, Digital Behaviour Change Intervention, Survey

1 Introduction

As part of the CAPABLE project [5] we aim to develop a virtual coaching (VC) system which facilitates cancer patients in development of positive health habits. The VC should provide patients with personalized health activity suggestions from the domain of mindfulness, positive psychology, and physical exercise. Personalization of content and delivery of these recommendations is expected to be powered by a machine learning (ML) model. The main challenge of developing ML-based VC systems is that it is hard to evaluate the efficacy of different ML approaches prior to the deployment of the digital intervention. One possibility is to use retrospective data capturing patient's responsiveness to recommendations delivered as a part of similar studies. Nevertheless, access to such data is limited.

An alternative is to use simulated data for the comparison of different ML approaches. Ho *et al.* [2] performed simulations imitating human reaction to notifications given different contexts, based on survey responses. In our prior work[3], we created an artificial user and simulated their habit formation based on Fogg's Behavior Model. According to Fogg, for behavior to occur three factors must be present: motivation, ability and prompt [1]. These factors are person dependent. Therefore, for VC evaluation, several types of user should be simulated.

In this preliminary work, inspired by Ho et al., we conducted a survey that elicited potential VC users' perceptions about the appropriate context for notification, aspects driving their motivation, and their ability to perform a selected activity in different circumstances. The objectives of this analysis are:

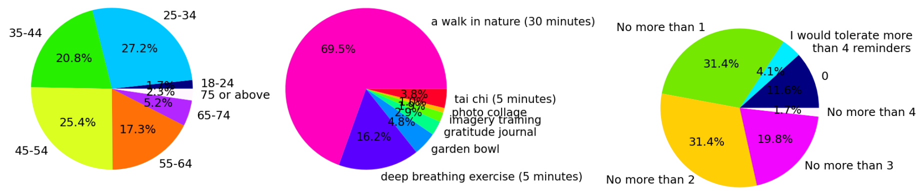
1. To include a variety of user types in our simulation. **We investigate whether there are different types (clusters) of users who might be similarly more receptive to behavior intervention and habit formation.**
2. To understand how to increase user susceptibility to non-medication health interventions. **The VC ideally would adjust recommendations based on users' preferences related to motivation, ability and prompt and to users' personality traits.** We also consider which information about the user is the most informative for adjustment of the VC operations.
3. To map survey responses to personality traits, making a first step toward understanding if personality plays a role in factors impacting habit formation.

2 Survey

We designed a questionnaire consisting of 35 questions: 3 demographic questions (age, gender, nationality), 1 question related to cancer diagnosis, 1 question with a selection from a list of mental well-being exercises that the user would like to perform habitually, 12 closed questions targeting understanding of respondents motivation for performing of health-related activities, 5 closed questions considering factors related to the ability to perform the activity, and 10 closed questions considering the impact of the users' current context on their responsiveness to prompts and their preferred number of prompts. The questionnaire was concluded with 3 open questions relating to motivation, ability and prompt. In the closed questions, respondents were asked to rate using a 5-point Likert scale their agreement to given statements.

We obtained ethical approval from the University of Haifa's Ethics Committee and distributed the survey in two languages: English – through the project's social media and email lists; and Italian – through an Italian Association of Cancer patients (AIMAC). The response to the survey was voluntary and no financial compensation was provided.

173 participants from 19 nationalities responded to our online survey, 98 of whom also replied to the additional BFI-10 [6] personality questionnaire, which consists of 10 questions measuring the Big Five personality traits: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness on 10-point scale. 52.6% of respondents were female, 45.7% male and 1.7% preferred not to specify. Only 17 respondents had been diagnosed with cancer. Fig 1a shows the distribution of respondents by age and Fig 1b shows the distribution of selected mental well-being activities. Interestingly, the vast majority of respondents (69%) selected to walk in nature as an activity they would like to perform for their well-being. The preferred number of daily notifications from



(a) Age (b) Selected mental well-being activity (c) Preferred number of notifications

Fig. 1: Distribution of responses

the VC varies, but only 11% of respondents stated that they would not like to receive any notifications (Fig. 1c).

3 Clustering patients based on their response

For this analysis, we used only the questions from our survey related to Fogg's three elements of motivation, ability, and prompt. We used the Agglomerated Clustering (AC) and Random Forest (RF) algorithms implemented in scikit-learn for analysis of the survey responses.

The AC was used to hierarchically cluster the respondents based on their rating of each question and the RF was used to identify the questions that help to discriminate between different types of respondents. We adopted a four cluster assignment discovered during the initial run of AC as labels for the training of RF classifier, where inputs were responses to each question. We used 75% of responses for training and the remaining 25% for testing.

The RF model trained with balanced class weights and a minimum of 15 samples at a leaf node achieved 72% classification accuracy on the test set. To identify questions that were the most informative in learning to assign the clusters we used two feature importance measures: Mean Decrease in Impurity (MDI) and Mean Decrease in Accuracy (MDA). Using each method, we selected the top 9 questions and then took their intersections, which resulted in 5 questions that were identified by both methods as important.

Finally, we ran AC again with a number of clusters fixed to 4, in order to cluster the responses only based on the top 5 questions. The resulting number of responses in each cluster was: 21,41, 55, and 56 (Fig. 2a). The clustering was run 10 times with 10% of respondents left out to verify the stability of these new clusters. In each run, we used the adjusted Rand index [7] to measure the similarity between clustering obtained from the full responses data set and the set with reduced number of questions. An identical partition would have an index of 1 and a completely different would score 0. The adjusted Rand scores varied from 0.55 to 0.80 between the dropout runs, suggesting that the cluster assignments were not very stable.

To understand how respondent types (clusters) differ and why sometimes cluster assignment may change, we visualized the distribution of responses per cluster for the top 5 questions (Fig. 2). For better interpretability, the responses are grouped into three categories: 1-2 ("no"), 3 ("neutral"), 4-5 ("yes"). It seems that the difference between clusters 0 and 1 vs 2 and 3 is in terms of motivation drivers. Respondents in the former two clusters would not be motivated by the introduction of the social component of competition (Fig. 2b and 2d). Most of the respondents from cluster 0 would also not be driven by their self-improvement (Fig. 2c), whereas including a progress-tracking in notification content or app display might motivate respondents in clusters 2 and 3. The difference between clusters 2 and 3 arises when considering situations in which the users might respond to prompts (Fig 2e and 2f). The users in cluster 3 might be more responsive than those in cluster 2. Interestingly respondents in clusters 1 and 2 were confident that they would not respond to activity recommendations when they are stressed (Fig. 2e). This suggests that just-in-time stress management interventions, such as the recommendation of deep-breathing exercises, might be ineffective for those groups of users. It might be important to know this prior to triggering the intervention, such that these users should not be prompted at times of stress (e.g., as detected from the blood volume pulse measured by a smartwatch [4]), but rather consider suggesting activities at times of lower arousal.

4 Personality trait and factor of habit formation

We investigated whether the discovered clusters differ in terms of the respondents' personality traits. For this analysis, we kept only the respondents who also completed the BFI-10 personality questionnaire. A one-way ANOVA revealed that there was no significant difference in Extraversion, Agreeableness, Conscientiousness, Neuroticism, or Openness between the four clusters (F= [1.13, 1.57, 0.48, 0.61, 0.97], p-value=[0.34, 0.19, 0.69, 0.60, 0.41]). We also trained Extra Tree Regressor to predict each personality trait given the survey question ratings. The model was trained on 75% of responses and tested on 25%. The mean absolute prediction error was 1.6, 1.2, 1.3, 1.7, and 1.5 for Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness, respectively. Responses to questions were generally not linked to personality traits. One question related to mood and motivation for spontaneous action (Fig. 3),was related to neuroticism, suggesting that for a person who scores highly on neurotic trait, recommendation of new activity might be ineffective when they are experiencing lower emotional valence.

5 Discussion

We conducted a survey to understand the difference between factors affecting people's motivation to perform health-related activities and the context in which they are able to respond to VC recommendations. Encouragingly, the majority of the respondents were open to receiving at least one prompt daily and 111 of the 173 respondents' motivation might be spiked by including gamification

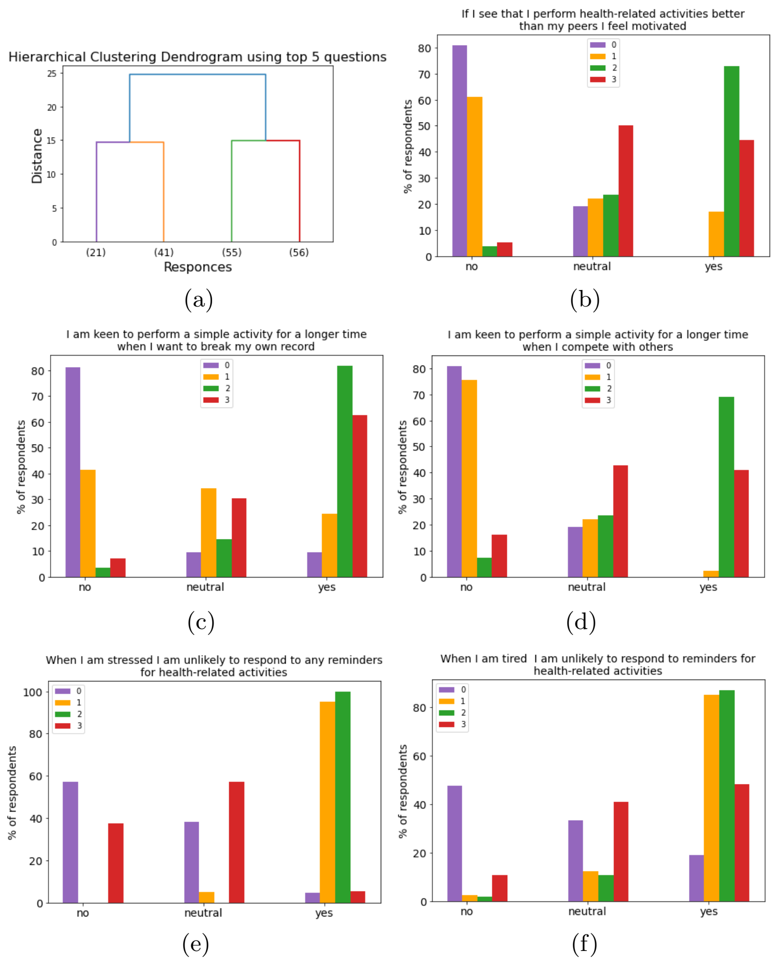


Fig. 2: Clusters and distribution of responses to selected questions as % of all responses in each cluster

components such as score dashboard for comparison with pears or personal score tracking.

Interestingly, the person's internal state of 'stressed', 'tired', 'low mood' results in higher perceived responsiveness to prompts than the person's location or motion. We also investigated if personality traits are related to identified respondents clusters and factors impacting the formation of new habitual behavior. Other than the link between neurotic trait and motivation for novel action in a low mood context, we found no links.

The gathered survey responses and presented user clustering might drive the simulation of people responsiveness to intervention in various contexts. However, the survey results are insufficient to conclude that personality and habit formation are unrelated. The hypothetical self-reported responses to behavior

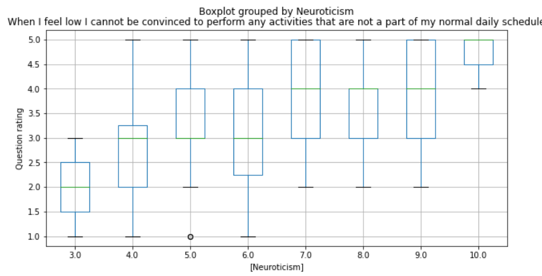


Fig. 3: Distribution of responses to a selected question about motivation across different levels of neurotic trait. X-axis: measurement of the neurotic trait, Y-axis: response to question.

intervention might vary from the actual response and hence the potential link or lack of it must be evaluated experimentally. The limitations of this work include: relatively small number of responses compared to complexity of the habit formation problem and potential selection bias due to distribution of the survey through social and professional network of the authors.

Acknowledgments

The CAPABLE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 875052.

References

1. Fogg, B.J.: A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. pp. 1–7 (2009)
2. Ho, B.J., Balaji, B., Koseoglu, M., Srivastava, M.: Nurture: notifying users at the right time using reinforcement learning. In: Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. pp. 1194–1201 (2018)
3. Lisowska, A., Wilk, S., Peleg, M.: From personalized timely notification to healthy habit formation: A feasibility study of reinforcement learning approaches on synthetic data. In: 2021 AIxIA SMARTERCARE Workshop, AIxIA. CEUR-WS
4. Lisowska, A., Wilk, S., Peleg, M.: Catching patient's attention at the right time to help them undergo behavioural change: Stress classification experiment from blood volume pulse. In: International Conference on Artificial Intelligence in Medicine. pp. 72–82. Springer (2021)
5. Parimbelli, E., Gabetta, M., Lanzola, G., etal: Cancer patients better life experience (capable) first proof-of-concept demonstration. In: Artificial Intelligence in Medicine. pp. 298–303. Springer International Publishing, Cham (2021)
6. Rammstedt, B., John, O.P.: Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. Journal of research in Personality **41**(1), 203–212 (2007)
7. Steinley, D.: Properties of the Hubert-Arable adjusted Rand index. Psychological Methods **9**(3), 386 (2004)