AUTHOR VERSION

Evaluation of computer-tailored motivational messaging in a health promotion context

J.E. d'Hondt, R.C.Y. Nuijten, and P.M.E. Van Gorp

Abstract
Persuasive messages have recently been shown to be more effective when tailored to the personality and preferences of the recipient. However, much of the literature on adaptive persuasion has evaluated the effectiveness of persuasive attempts by the direct reactions to those attempts instead of changes on the longer term (e.g. lifestyle changes). Results of this study suggest that adaptive persuasion improves attitudes towards persuasive attempts, but does not necessarily cause a change in longer term behavior. This was found through a randomized controlled trial evaluating the implementation an adaptive persuasive system in a health promotion intervention. This article provides a detailed description of this evaluation and encourages the research community to (1) become more skeptical towards the longer term effectiveness of adaptive persuasive techniques and (2) design more explicitly for longer term changes in behavior.

KEYWORDS
persuasion; individual differences; health promotion

1. Introduction

Disruptive developments in computing power and data collection throughout the last decade have allowed suppliers to obtain and process more information on customer interactions, both relating to the customer itself and the circumstances in which the interactions take place. This information gain does not only aid in designing products better suited to customer needs, it also allows for tailoring promotional efforts to individual customers [1]. Currently, research on personalized promotion has focused most on creating artificial agents and feeding them with information on context variables like individual's valuation of product attributes and prior purchase behavior in order to improve their decisions on product recommendations (also known as recommender systems). However, integration of information on other variables such as price elasticity or personality traits is far less discussed in literature, and is scarcely implemented in situations where effective product promotion is fundamental, such as marketing and e-commerce [11]. The idea of using information on people's personality traits in order to optimize promotional efforts arose from the fact that it was recently found that there exist individual differences in susceptibility to certain principles of persuasion and therefore also in responses to different persuasion strategies [10]. These differences may entail improvements in persuasive approaches when tailored correctly to the recipient, but may also encompass negative results when striking the wrong chord [10]. Tailoring persuasion strategies to the personality of the recipient (adaptive persuasion) has shown to be effective in a variety of contexts, e.g. marketing [9], medication [3] and health care [15]. However, what is remarkable about these studies is that they often only measure the effectiveness of persuasive approaches on short-term behavior and attitudes (e.g. direct feedback or click-through rates [9]) without measuring their effect on long-term behavior. Consequently, while the positive impact of adaptive persuasion on the direct attitudes towards persuasive approached has been shown, it's ability to foster behavioral change still remains unknown. This is, however, important in case adaptive persuasion is desired to be used in contexts like health promotion and medication, where habits are attempted to be changed in order to ensure medication adherence. We have reason to question this ability as a prominent framework on behavioral change, the COM-B framework by Michie et al. [14] suggests that there are three antecedents of behavioral change; Capability, Opportunity and Motivation, and interventions aimed at inducing behavioral change ought to provide support which cultivates compliance to all conditions. This implies that informing artificial agents on the personality of the recipient may help to motivate a person to act, but it may fail to foster behavioral change in case no contextual information is provided regarding the person's capability or opportunity to do so. Therefore, the current study evaluates if adaptive persuasion does not only improve attitudes towards persuasive attempts, but also has the potential to induce behavioral change. This was done by conducting a randomized controlled trial evaluating the implementation of an adaptive persuasive messaging system (APMS) in a health promotion intervention for employees and students of the University of Technology in Eindhoven. In this paper, section 2 will focus on the design of the APMS and methods used to evaluate the system. Section 3 will present the results of the evaluation, subsequently discussing these results and drawing conclusions in the final section.

2. Design process of an APMS

Currently, many of the existing health promotion interventions include some combination of a health tracking device and a mHealth application to log activities and review progress. In this context, the APMS could be integrated into the back-end of such applications, using messages to 'nudge' the user into performing more activities and to remind him/her of certain developments on the platform. Information on in-app behavior and direct feedback could then be used to personalize nudges in order to increase their effectiveness, resulting in greater engagement with the application, and ultimately greater adoption of healthy behaviors. In the case of this study, the APMS was known to be implemented in a mHealth application called GameBus[1]. GameBus was built according to the principles of gamification and provides an environment to host and participate in digital competitions in which points can be earned by performing activities related to living a healthy lifestyle. Since GameBus did not yet contain any form of active messaging towards its users, the platform could be naturally extended with the purpose of studying persuasive messaging personalization. Figure 3 provides a very simplified flow diagram representing the overall user journey of health promotion applications like GameBus, highlighting the aimed position of the APMS in that journey. Similarly, Figure 2 provides a graphical overview of the design of the study.

[1]https://www.gamebus.eu/

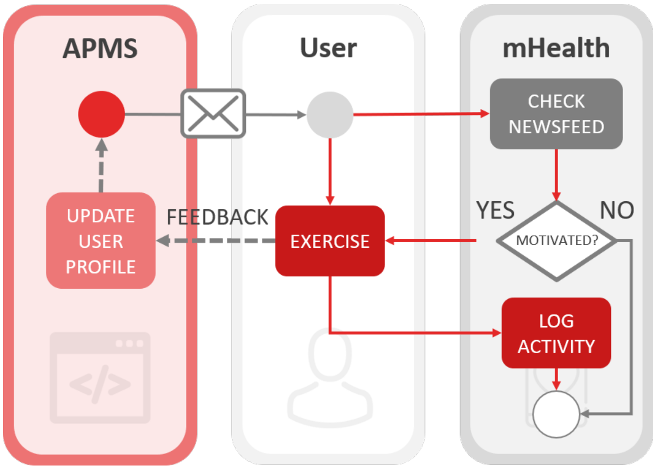

Figure 1. Visual respresentation of general user journey when using mHealth applications

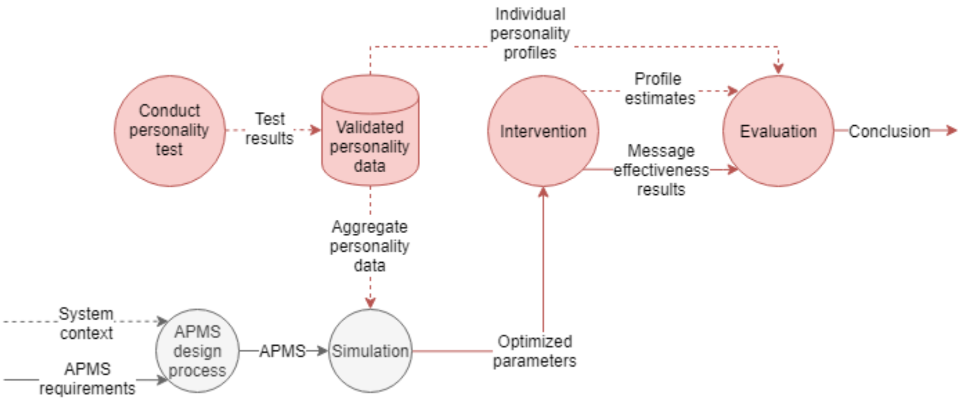

Figure 2. Graphical representation of the study design

2.1. System Requirements

Built on the requirements of *adaptive persuasive technologies* as presented by Kaptein & Van Halteren (2013) [12], the authors suggest five requirement for the APMS to be capable of effectively adapting to individual differences in responses of recipients. Firstly, the system should be able to identify individual users and maintain a specific user profile indicating the probability of success of different influence strategies. Secondly, it should be able to frame messages to be congruent with a specific influence strategy. Thirdly, at time of constructing the message, it should have a clear protocol of choosing the influence strategy used in the message. Fourthly, when the message has been sent, predetermined success measure(s) are needed to assess the effectiveness of the approach. And lastly, after the effectiveness of an approach has been assessed, the system needs specific learning rules in order to update the user profile and optimize long-term message effectiveness.

2.2. System Design

2.2.1. User Profiling

Prior studies on adaptive persuasion report profiling users on different psychological traits such as Need for Cognition [7], susceptibility to different influence principles (e.g. those presented by Cialdini [2]) or the Big Five personality dimensions (which proposes that a person's personality can be modeled by his/her adherence to each of five different traits; extraversion, agreeableness, conscientiousness, neuroticism and openness) [5]. Profiling based on the latter is deemed most appropriate in this context compared to profiling based on Cialdini's principles as we look to induce long-term behavioral change, which reduces the relevance of principles encouraging immediate action like Cialdini's scarcity principle. In addition, there is a greater body of research discussing the variation in motivational systems reflected by each of the Big Five personality traits [9] compared to Need for Cognition. This helps to ensure unambiguous message framing in later stages of the design process. Therefore, users are profiled based on their personality expressed by their estimated adherence to each the Big Five personality traits. In line with this decision, individual user profiles are modeled as a collection of Beta-Binomial models [21] each representing the probability distribution of the value of p_m, indicating the probability that the recipient's response to a message framed congruent with trait m is positive. Note that this implies we are modeling the probability distribution of the value of p_m, i.e. the probability of a probability. This method is suitable here as any persuasive interaction can be seen as a binomial random process, for which prior information (i.e. responses to previous messages) is included to iteratively enhance estimations for the probability of success p_m [12]. See Figure 3 for an example profile. In the figure, every curve corresponds to the probability density function of p_m of a message framed congruent with trait m (being one of the Big 5 personality traits).

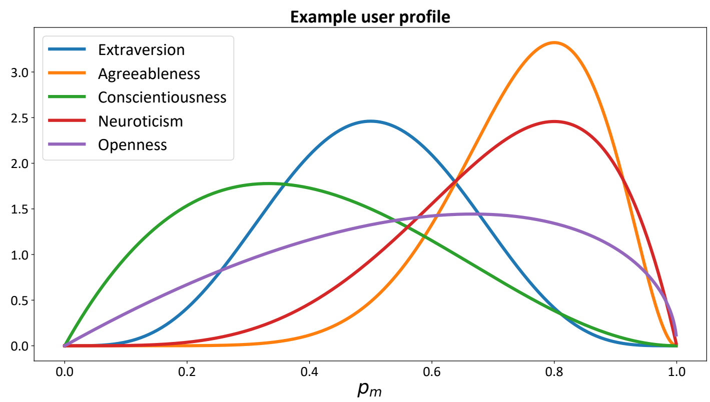

Figure 3. User profile example

2.2.2. Message Framing

In order to preserve significant variation in message content, five distinct motivational message elements were drafted for each of the five personality traits. When constructing an email, a message element was chosen and inserted into the body of an email.

Besides the persuasive element, emails included a personalized preamble and the promotion of a random GameBus activity[2].

2.2.3. Influence Strategy Choice

The algorithm used to determine the optimal persuasion strategy of a given message was derived from the decision rules presented in a study by Kaptein & Parvinen (2015). In the study, they use the concept of *Randomized Probability matching* (RPM) as introduced by Scott [18] to deal with the *explore/exploit dilemma*, and James-Stein shrinkage [20] of individual user profiles to an average profile of the whole sample population to reduce the uncertainty on user profiles of individuals of whom little data has yet been obtained. We refer to the paper of Kaptein and Parvinen [11] for a basic explanation of the algorithm. Both solutions were used in the current study with the modification that α- and β-values corresponding to the Beta-Binomial models (Beta(α, β)) in user profiles were shrunk using the following set of equations. Note that equation 3 was not derived from prior research on RPM and James Stein Shrinkage, but was added to ensure a decrease in the amount of shrinkage over time;

$$\alpha^*_{t,i} = \overline{\alpha}_t + c_i(\alpha_{t,i} - \overline{\alpha}_t) \tag{1}$$

$$\beta^*_{t,i} = \overline{\beta}_t + c_i(\beta_{t,i} - \overline{\beta}_t) \tag{2}$$

$$c_i = 1 - e^{-\lambda * n_i} \tag{3}$$

where:

$\alpha_{t,i}$	α in Beta(α, β) for personality t, participant i
$\overline{\alpha}_t$	α in Beta(α, β) for personality t in average profile
$\beta_{t,i}$	β in Beta(α, β) for personality t, participant i
$\overline{\beta}_t$	β in Beta(α, β) for personality t in average profile
n_i	number of profile updates for participant i
λ	profile maturation constant

This implies that upon strategy selection a profile is used which essentially is a mix between an average profile derived from all reported message feedback and the individual profile of the prospective recipient. The amount of information 'borrowed' from the average profile decreases as a function of the number of profile updates of a participant and the profile maturation constant λ. Following the principles of RPM, strategies are eventually chosen by taking draws from each of the Beta-Binomial distributions in the shrunk profile, choosing the strategy associated with the highest draw.

2.2.4. Success Measure

Message effectiveness was evaluated using two sources of feedback from the recipient. Firstly, users were able to provide direct feedback on the message through a feedback form included in the email. In order to capture additional (contextual) information,

[2]See https://doi.org/10.6084/m9.figshare.8289044.v1 for the list of persuasive elements as used by the APMS

participants could not only indicate if the message content appealed to them, but also why (not). Hence, when asked whether a message suited their preferences, participants could either express their satisfaction with the message (and confirm whether they thought they would soon perform the activity, or not), or express their dissatisfaction with one or multiple (contextual) aspects. In case participants indicated they disliked the message due to it's timing or frequency, they were automatically offered the opportunity to update their message preferences. Only message feedback which includes specified sentiment towards the used influence strategy triggered a model update corresponding to the influence strategy used in the message. Secondly, message effectiveness was indirectly measured using information on the participant's performed activities. To illustrate, upon logging an activity into GameBus, participants were requested to indicate what triggered them into performing that activity. A message was deemed effective in case it's promoted activity was logged within 3 days after the message was sent, indicating it was triggered by a message.

2.2.5. Learning Rule

Consequent to a persuasive approach's effectiveness being reported, the parameters of the recipient's corresponding Beta-Binomial distribution were updated in the following way:

$$\begin{gathered}\alpha_{ti} := \alpha_{ti} + a \\ \beta_{ti} := \beta_{ti} + (1-a) \\ \text{with } a = \left\{ \begin{array}{ll} 1 & \text{if feedback} == \text{positive} \\ 0 & \text{if feedback} == \text{negative} \end{array} \right\}\end{gathered} \tag{4}$$

As the point estimate $\hat{\mu}_t$ of the effectiveness of a message framed congruent with personality t follows $\hat{\mu}_t = \frac{\alpha_t}{\alpha_t+\beta_t}$ with variance $\hat{\sigma}_t^2 = \frac{\alpha_t\beta_t(\alpha_t+\beta_t+1)}{(\alpha_t+\beta_t)^2(\alpha_t+\beta_t+1)}$, the estimated effectiveness was increased/decreased depending on a while the variance of the estimate decreases each time new information is gained through feedback.

2.2.6. System Overview

In the discussed example, the usages of contemporary database, web and messaging middleware allowed the APMS to be deployed on a separate network node other than the system on which activity registrations were taking place (in this case GameBus). Thanks to this approach, the APMS is fully modular and should be able to be build on top of other platforms in a relatively easy manner. Additionally, this also allows the full source code to be shared for further use - see https://github.com/JdHondt/APMS.git.

2.3. Trial Simulation

In order to validate the system's functionalities and to allow for hyper-parameter tweaking (i.e. the profile maturation constant λ), the trial was simulated with conditions similar to the randomized controlled trial through which the system was ought to be evaluated. 150 personality profiles were randomly generated based on the means and standard deviations of personality profiles obtained via a personality test (following the Mini-IPIP design by Donnelan et al. (2006) [4]) included in an intake survey conducted on participants of the randomized controlled trial ($n = 36$). Note that the results of the personality test as part of the intake survey were solely used for evaluation purposes and simulation, and did not serve as direct input for the algorithm's profile development. Reactions to persuasive approaches were simulated by taking a draw from a Bernoulli-distribution with $p = p_m$, m being the trait towards which the message was framed. To evaluate the impact of adaptive persuasion, the APMS updated the profile estimates of only half of the generated profiles, effectively simulating two treatment groups. The profile maturation constant λ and Big Five trait value variance σ_p of the generated profiles (i.e. individual differences in personality) were varied across simulations in order to analyze the ideal value for λ given the length of the intervention and different values of σ_p. 20 simulations were run per λ, σ_p combination to reduce the variability in results.

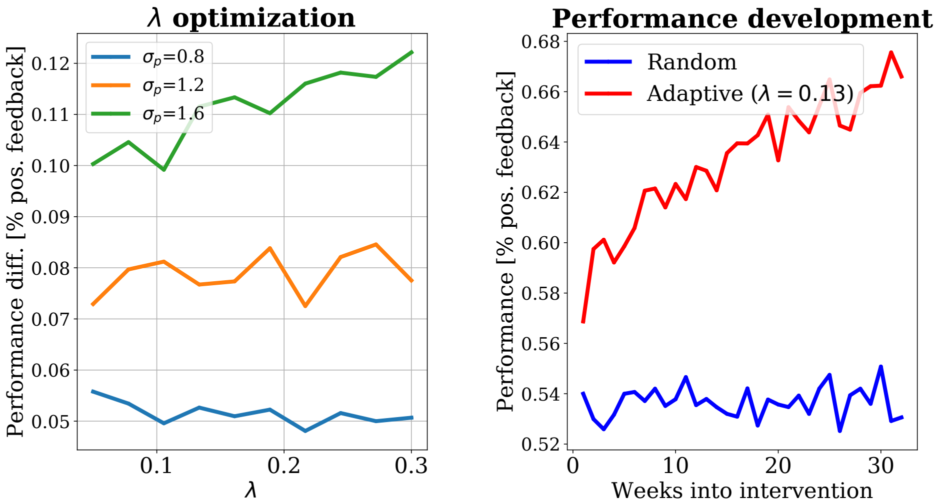

Figure 4. Simulation Results

2.3.1. Simulation Results

The results as presented in Figure 4 verify the effectiveness of RPM in increasing the expected 'performance' of messages. This can be concluded by both the positive differences for all conditions in the left figure and the increasing performance over time of the adaptive condition in the right figure. Furthermore, the λ optimization graph clearly highlight the positive relationship between the σ_p and λ. This relationship is logical as high variability in individual profiles diminishes the usefulness of shrinkage, profiting from earlier influences of individual profile estimates. As the result of the personality survey indicated a personality variance of around $\sigma_p = 0.08$, the authors decided to go with a λ of 0.13, implying strategy selection will be mainly based on the average profile considering the short duration of the intervention.

2.4. Evaluation Methodology - Health Promotion Intervention

The APMS was evaluated by implementing it in a health promotion intervention similar to that of Nuijten et al. (2019), which included a 6-week digital health competition among a sample of students and employees of the Eindhoven University of Technology

($N = 149$). Both an individual and a team-based competition were held among different university departments with prizes being awarded to the entities that ranked best of their competition at the end of the trial. Additionally, weekly rewards were raffled among those who reached a weekly point target. Points were earned by completing photo-based challenges involving engagement in some form of physical activity and taking a photo (so called *FitPic*) as proof of completion[3]. The rationale behind the selection of competition elements discussed above was to provide significant stimuli for all Big Five personality types and their characteristics. To illustrate, both the competition reward structure as GameBus' news feed functionality served as stimuli for participants scoring high on the Extraversion dimension as they are especially sensitive to rewards and social attention [19] or Neuroticism as it is associated with a strong sensitivity to uncertainty and threats [8]. The concept of a team-based competition was introduced to support agreeable individuals as they value communal goals and interpersonal harmony [6]. Similarly, an individual competition with only one winner served as a stimulant for conscientious individuals who value achievement [17]. Lastly, as open individuals value creativity and intellectual stimulation [13], challenges were framed in a flexible manner that offered opportunity for creativity and one's own interpretation.

2.4.1. Conditions

To evaluate the effectiveness of personalized persuasive messaging as opposed to using arbitrary persuasion strategies in messages, participants were distributed over two treatment groups:

(1) *Control:* Random choice of persuasion strategy upon messages creation. Each strategy had equal chance of being chosen.
(2) *Treatment:* Persuasion strategy used in messages chosen based on both feedback received by the prospective recipient on previous messages as well as feedback received by all participants (including participants assigned to control) on messages sent to them.

This between-subjects design was chosen over a within-subjects design to compensate for a learning curve related to GameBus which appeared from previous studies using the application. The first intervention week served as a learning period for participants to get comfortable with the GameBus platform and elements of the competition. In this week both control and treatment received randomized messages and no message feedback was gathered. Adaptive messaging was initialized starting from the second week.

2.5. Statistical Analysis

To evaluate the effectiveness of tailored messages over messages with a random persuasion strategy, multiple generalized linear (mixed effect) models were fitted on the data for different effectiveness metrics and compared to each other using χ^2 tests. Additionally, in case a model reported a significant improvement towards a simpler model, it's fixed effects were analyzed to see what conclusions could be made on the relationship between the fixed effects and target variables. As for the effectiveness metrics, the following variables were chosen:

[3]see https://doi.org/10.6084/m9.figshare.8288882.v1 for the list of challenges.

- *Message Feedback (MF):* Sum of feedback values f_i.
- *Performed Activities (PA):* Number of activities logged in GameBus.
- *Message Success (MS):* Percentage of messages causing an activity being performed[4].

All target variables were measured on an individual level and aggregated by week. f_i was computed following;

$$f_i = \left\{ \begin{array}{rl} 1 & \text{if feedback == positive} \\ 0 & \text{if feedback == negative (not content-related)} \\ -1 & \text{if feedback == negative (content-related)} \end{array} \right\} \tag{5}$$

As for the models, all models are extensions on a so called "null" model. This model describes an overall intercept on the target variable, with the addition of individual-level intercepts as a random effect. Fitting this model basically tests for individual differences in the values for the effectiveness metrics, regardless of the individual's treatment group. All model comparisons were done towards the null model. The models were fitted using Poisson error distributions for both PA and MS target variables, and Gaussian error distributions for target variable MF. Models are presented in Table 1. Both *group* features in Model C and D are binary variables indicating a participant's belonging to the treatment group. The β-values correspond to β_0 (intercept) for Model A and the weights of time, group and $time \times group$ for Models B, C and D, respectively.

3. Results

At the start of the intervention 149 people were registered as eligible participants by direct registration or indirectly as they were ex-participants of Nuijten et. al [16] and did not opt-out for this study. These participants were randomized to either the intervention ($n_i = 75$) or control group ($n_c = 74$). At the end of the intervention, 16 participants had actively withdrawn their participation and 111 were excluded from the data analysis as they did not comply to the requirements of completing the intake survey and providing feedback on at least one message. Of the remaining participants ($n_i = 12$, $n_c = 10$), the majority was male (63.64%) with a mean age of 32.2 ± 6.49 and a baseline BMI of 23.6 ± 2.59 kg/m^2.

	MF		PA		MS	
	χ^2	β	χ^2	β	χ^2	β
Model A: null model	-	0.33**	-	-0.55	-	-1.10
Model B: A + time	2.23	0.07	0.29	-0.03	0	0
Model C: A + group	8.69**	0.54**	0.53	-0.06	0	0
Model D: A + time × group	11.5**	0.07	2.55	0.13	0.35	-0.12

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1. Table showing fixed effects and model comparisons for different target variables

[4]Activity-message causation in MS was determined the same way message effectiveness was measured using information on the performed activities of participants; by flagging messages of which the promoted activity was performed within 3 days, indicating a message as it's trigger.

3.1. System Evaluation

When analyzing the results shown in Table 1, we see a significant positive intercept for model A on MF, indicating a general tendency towards positive message evaluations of participants (regardless of treatment group). Moreover, significant increases in model fit are reported for models C and D on MF, with a significant positive fixed effect for the *group* variable in Model C, but an insignificant β- value for the interaction term on *time* and *group* in Model D. All other model fits are reported to be insignificant. The insignificant β-value for $time \times group$ indicates that there does not seem to be a significant difference in MF development over time between groups, which one would not expect considering the fact that user profile estimates (in theory) should improve over time, resulting increased MF values as illustrated in the simulation results. In view of this discrepancy, Mann-Whitney tests were performed for both week 1 and 6 of the intervention, indicating an insignificant difference in MF in week 1 ($p = 0.119$), but a significant difference in week 6 ($p = 0.035$).

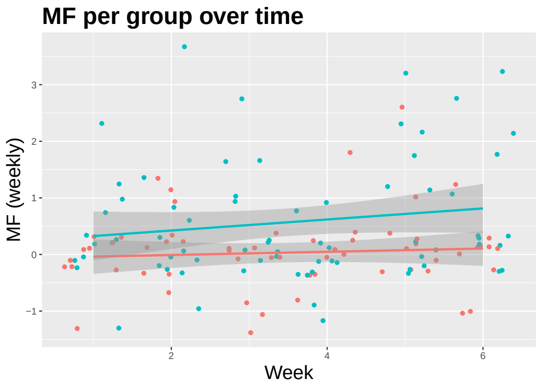

Figure 5. Overview of MF over time

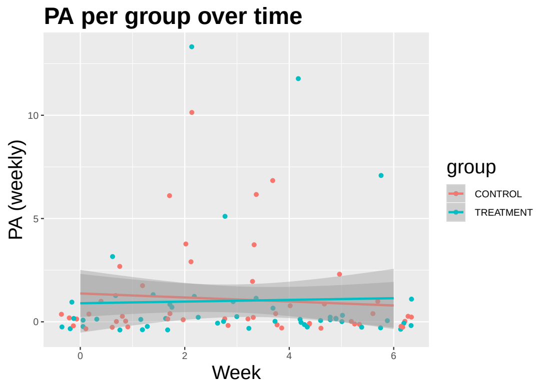

Figure 6. Overview of PA over time

4. Discussion & Conclusions

We conclude that participants receiving personalized messages evaluate messages more positively than those receiving messages framed towards a random personality. These results are in line with those reported in prior research. The insignificant difference in MF in week 1 opposed to the significant difference in week 6 hints towards a between-group MF development difference, but does not support the corresponding hypothesis which might be due to the limited duration of the intervention. From the insignificant fits for models on both PA and MS we can conclude that despite their superior evaluation of messages, participants in the treatment group are not reported to perform more activities. The poor fits and effects of models with MS were mainly a consequence of the fact that only 9 activities were reported indicating a message as activity trigger.
In this study, we presented an extension on the requirements of adaptive persuasive systems which designers can use to create and implement artificial agents which adapt to individual differences in the preferences of users. An important purpose of this study was to evaluate if adaptive persuasive systems have the ability to induce behavioral change in a health promotion context, as the authors felt that this was insufficiently studied by prior research. The reported results suggest that although these systems

may induce elevated attitudes towards persuasive approaches, they do not necessarily cause a change in health behavior. This highlights the multidimensionality in the antecedents of behavioral change, and should encourage researchers and enterprises to create artificial agents which are not only sensitive to context-variables which influence people's motivation to act, but also to those which impact people's capability and opportunity to adopt the desired behavior. In essence, while guidelines are offered to implement adaptive persuasion, we emphasize that effective persuasion alone will not be enough to induce behavioral change in a health promotion context, and sufficient regard has to go out to other support elements that increase capability and opportunity to act.

4.1. Limitations

There are several limitations to this study. One includes it's small sample size, which resulted from high drop-out rates and low participant engagement with the study. According to a post-intervention survey ($n = 12$), the main causes of this inactivity included a substantial barrier to log daily activities, as this was ought to be done manually and annoyance with the email messages, their frequency was felt to be too high. A second limitation involves the Influence Strategy Choice method used in the APMS. Here, considering time constraints, RPM with James-Stein shrinkage was used due to its simplicity and fast implementation. However, while the main attractiveness of this method comes from its short computation time, other more computational intensive methods like Gittins Indices or Multilevel Hierarchical models are estimated to have higher predictive performance. These models could be more suited for this problem as message effectiveness predictions are not required to be computed real-time. Lastly, a problem the authors dealt with during design, was the difficulty in linking activity to message, effectively identifying which approach was successful. This problem is similar to the credit assignment problem discussed in research on Reinforcement Learning, for which currently no definitive solution is known. Unfortunately, this has limited the system's ability to improve user profiles based on direct actions, consequently having to base profiles mainly on message feedback.

4.2. Future Research

In light of the discrepancy between positive evaluation of persuasive approaches and reported behavior found in this study, future research should be conducted evaluating the effectiveness of more complete persuasive systems that combine personalization in terms of promoted product or service with personalization of persuasion strategy to assess if a synergy between methods does bring about behavioral change. Such a system could be implemented in an enlarged replication of the trial discussed this study, with the addition of a more effective recruitment strategy to ensure larger sample sizes and increased participant engagement. Lastly, the current evaluation could be expanded by including an analysis comparing the individual profile estimates which resulted from the intervention with the profiles derived from the personality test conducted pre-intervention (indicated by the top dotted arrow in Figure 2). This could not only provide insights into the system's profile prediction accuracy, it could also shed light on potential differences between reported personality and performed behavior.

References

[1] G. Allenby and P. Rossi. Marketing models of consumer heterogeneity. *Journal of Econometrics*, 89:57–78, 1999.

[2] R. Cialdini, M. Trost, and J. Newsom. The development of a valid measure and discovery of surprising behavioural implications. *Journal of Personality and Social Psychology*, 69(2):318–328, 1995.

[3] A. Dijkstra. Working mechanisms of computer-tailored health education: Evidence from smoking cessation. *Health Education Research*, 20(5):527–539, 2005. ISSN 02681153. .

[4] M. B. Donnellan, F. L. Oswald, B. M. Baird, and R. E. Lucas. The Mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment*, 18(2):192–203, 2006. ISSN 10403590. .

[5] L. R. Goldberg. An Alternative "Description of Personality": The Big-Five Factor Structure. *Journal of Personality and Social Psychology*, 59(6):1216–1229, 1990.

[6] W. G. Graziano and N. Eisenberg. Agreeableness: a Dimension of Personality. 1997. URL `http://cachescan.bcub.ro/e-book/E1/580591/795-870.pdf`.

[7] C. P. Haugtvedt, R. E. Petty, and J. T. Cacioppo. Need for Cognition and Advertising: Understanding the Role of Personality Variables in Consumer Behavior. *Journal of Consumer Psychology*, 1(3):239–260, 1992. ISSN 10577408. .

[8] J. B. Hirsh and M. Inzlicht. The devil you know: Neuroticism predicts neural response to uncertainty. *Psychological Science*, 19(10):962–967, 2008. ISSN 09567976. .

[9] J. B. Hirsh, S. K. Kang, and G. V. Bodenhausen. Personalized Persuasion: Tailoring Persuasive Appeals to Recipients' Personality Traits. *Psychological Science*, 23(6):578–581, 2012. ISSN 14679280. .

[10] M. Kaptein and D. Eckles. Heterogeneity in the Effects of Online Persuasion. *Journal of Interactive Marketing*, 26(3):176–188, 2012. ISSN 10949968. .

[11] M. Kaptein and P. Parvinen. Advancing e-commerce personalization: Process framework and case study. *International Journal of Electronic Commerce*, 19(3): 7–33, 2015. ISSN 15579301. .

[12] M. Kaptein and A. Van Halteren. Adaptive persuasive messaging to increase service retention: Using persuasion profiles to increase the effectiveness of email reminders. *Personal and Ubiquitous Computing*, 17(6):1173–1185, 2013. ISSN 16174909. .

[13] R. R. McCrae, P. T. Costa, and P. T. Costa Jr. Chapter 31 - Conceptions and Correlates of Openness to Experience. *Handbook of Personality Psychology*, (May):825–847, 1997. ISSN 14222795. .

[14] S. Michie, M. M. van Stralen, and R. West. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(1):42, 2011. ISSN 17485908. .

[15] M. Morris and F. Guilak. Mobile Heart Health : Project highlight. *IEEE Pervasive Computing*, pages 57–61, 2009. .

[16] R. Nuijten, P. V. Gorp, U. Kaymak, M. Simons, A. Kemperman, and P. V. D. Berg. Evaluation of the impact of extrinsic rewards on user engagement in a health promotion context. 2019.

[17] B. W. Roberts, O. S. Chernyshenko, S. Stark, and L. R. Goldberg. the Structure of Conscientiousness: an Empirical Investigation Based on Seven Major Personality Questionnaires. *Personnel Psychology*, 58(1):103–139, 2005. ISSN 0031-5826. .

[18] S. L. Scott. A modern Bayesion look at multi-armed bandit. *Applied Stochastic*

Models in Business and Industry, 26:157–164, 2007. ISSN 1524-1904. .

[19] L. Shao, E. Diener, E. M. Suh, R. E. Lucas, and A. Grob. Cross-cultural evidence for the fundamental features of extraversion. *Journal of Personality and Social Psychology*, 79(3):452–468, 2005. ISSN 0022-3514. .

[20] C. Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proceedings of the Third Berkeley symposium on mathematical statistics and probability*, 1(4):197–206, 1955. ISSN 0097-0433.

[21] R. R. Wilcox. A Review of the beta-binomial model and its extensions. *Journal of Educational Statistics*, 6(1):3–32, 1981.