Association for Information Systems
AIS Electronic Library (AISeL)

ECIS 2014 Proceedings

PERSUASIVE MESSAGES: THE EFFECT OF PROFILING

Peter Mayer
Business Engineering Institute AG, St. Gallen, Switzerland, peter.mayer@bei-sg.ch

Frederic Thiesse
University of Würzburg, Wuerzburg, Germany, frederic.thiesse@uni-wuerzburg.de

Elgar Fleisch
University of St. Gallen, St. Gallen, Switzerland, elgar.fleisch@unisg.ch

Peter Mayer, Frederic Thiesse, and Elgar Fleisch, 2014, "PERSUASIVE MESSAGES: THE EFFECT OF PROFILING", Proceedings of the European Conference on Information Systems (ECIS) 2014, Tel Aviv, Israel, June 9-11, 2014, ISBN 978-0-9915567-0-0
http://aisel.aisnet.org/ecis2014/proceedings/track12/3

PERSUASIVE MESSAGES: THE EFFECT OF PROFILING

Completed Research

Mayer, Peter, Business Engineering Institute St. Gallen, Switzerland, peter.mayer@bei-sg.ch

Thiesse, Frédéric, University of Würzburg, Germany, frederic.thiesse@uni-wuerzburg.de

Fleisch, Elgar, University of St. Gallen, Switzerland, elgar.fleisch@unisg.ch

Abstract

In the present study, we investigate to which degree persuasion profiling can increase the effectiveness of adaptive persuasive systems. For this purpose, an experiment was conducted in which subjects were exposed to persuasive SMS messages under three experimental conditions. One group received messages that fit to their personality traits, a second group obtained messages that do not fit, and a third group was exposed to a random selection of messages. Comparing the degree to which the three experimental groups responded to the messages, we could show that well-fitting messages and randomly selected messages perform significantly better than non-fitting messages, whereas the difference between well-fitting and randomly selected messages was not significant.

Keywords: Persuasion profiling, adaptive persuasive technologies, field experiment, linear mixed models.

1 Introduction

Many prototypical implementations of persuasive technologies have been proposed in recent years, which have in common that they apply a static set of persuasive strategies but do not adapt the way a person is influenced to his personality traits. Recently, persuasive systems have come into the focus of IS research that adapt their persuasive strategies to the personality traits of a user. Psychological research has shown that people respond differently to certain persuasive principles. For example one person may rather be influenced by personal goal setting whereas another person may rather be susceptible to social norms. So-called *adaptive persuasive technologies* can implement different persuasive principles and select the one that is most promising for a certain user (Kaptein and Eckles 2010a). Adaptive persuasive technologies must be capable of *persuasion profiling*, i.e. to retrieve "collections of expected effects of different influence strategies for a specific individual" (Kaptein and Eckles 2010a) and thus to apply those persuasion strategies that can be expected to be most effective.

In the present study, we investigate to which degree *persuasion profiling* can increase the effectiveness of adaptive persuasive systems. For this purpose, an experiment was conducted in which subjects were exposed to persuasive SMS messages under different experimental conditions. Comparing the degree to which the experimental groups responded to the messages, we could show that well-fitting messages and randomly selected messages perform significantly better than non-fitting messages, whereas the difference between well-fitting and randomly selected messages was not significant.

The remainder of the paper is organized as follows. In the next section, we summarize work that is related to the profiling of persuasive technologies. Then the research design is described, followed by an explanation of the data collection process. Next, the data analysis method based on linear mixed models is described, followed by a discussion of theoretical and managerial implications.

2 Related Work

Several taxonomies have been developed over the last decades to structure the potential approaches to exert persuasion on a person. Among the first of them, Marwell and Schmitt (1967) identified 16 basic persuasive strategies, which they clustered into the five groups of *rewarding activities*, *punishing activities*, *expertise, activation of impersonal commitment*, and *activation of personal commitments.* Levine and Wheeless (1990) compiled a list of 53 basic persuasive strategies, which were derived from nine earlier taxonomies. Kellermann and Cole (1994) analysed 74 classification systems and developed a taxonomy of 64 persuasive principles. With a focus on persuasive technologies, Fogg (2002) developed a taxonomy of 42 persuasive strategies, which are clustered along six functional aspects of IT-based persuasive systems. Among the many available taxonomies, we decided to apply the rather parsimonious taxonomy proposed by Cialdini (2008), which reduces the vast number of persuasive tactics to six clearly distinguished strategies: (i) authority; (ii) commitment and consistency; (iii) social proof; (iv) liking; (v) reciprocity; and (vi) scarcity.

The studies outlined in the previous section are similar in that they analyse the effectiveness of different persuasive approaches without taking into account individual differences between subjects. Noar et al. (2007) conducted a meta-analytic review of 57 health-related intervention studies to evaluate whether tailored persuasive messages (i.e., messages that are unique to the respective person) are superior to non-tailored ones. They found that tailored messages perform generally better than non-tailored messages. They furthermore found that tailoring is most effective if it is based on theoretical concepts and personality traits like attitudes, self-efficacy, stage of change, processes of change, and social support. With regard to cultural differences, Cialdini et al. (1999) found that cultural conditioning may influence individual susceptibility to certain persuasive strategies. Their results have shown that *social proof* performed better in the more collectivistic culture in Poland, whereas *commitment / consistency* was superior in the more individualistic society of the United States. Analysing individual differences in the processing of persuasive messages, Cacioppo et al. (1986) found that people with a high *Need for Cognition* (NfC) think more intensively about incoming messages than people with low NfC. These findings explain - at least partially - the results found by Kaptein et al. (2009) and Kaptein et al. (2010) that people differ in their general level of susceptibility to persuasion.

Moon (2002) investigated how personality traits influence the effectiveness of different persuasive strategies. She found that dominant personalities are more susceptible to dominant messages, whereas submissive messages show a larger effect when applied to submissive personalities. Similarly, Halko and Kientz (2010) have shown that the perception of differently shaped persuasive approaches is influenced by the so-called Big-5 personality traits, which are *Neuroticism*, *Conscientiousness*, *Agreeableness*, *Extraversion, and Openness* (Goldberg 1993). Cialdini et al. (1995) investigated the effectiveness of the so-called foot-in-the-door tactic. This means that initially a small request is made to a person to create intrinsic commitment. Once the person has agreed, the actually intended larger request is revealed. Cialdini et al. (1995) have shown that this tactic is only effective for individuals with a high *Preference for Consistency* (PFC), i.e. with an intrinsic urge to stay consistent with former actions. They furthermore found that only half of their study participants showed a high PFC level. Guadagno et al. (2001) have demonstrated that an explicit reference to prior commitment increases the compliance of people with a high PFC level, whereas it has a reverse effect on people with low PFC.

To summarize, there is broad empirical evidence that people differ in their general susceptibility to persuasive attempts as well as in their response to certain persuasive principles. Several studies indicate that applying an inappropriate strategy may reverse the intended effect such that a subject might not only deny compliance with a persuasive message, but might even show an adverse change in behaviour. Preferences cannot usually be predicted on the basis of demographic characteristics. Susceptibility for certain strategies is to some degree related to personality traits, but for selecting an optimal persuasive strategy, individual susceptibility must be assessed.

3 Research Design

To investigate whether profiling may increase the effectiveness of persuasion, we designed an experiment in which subjects were exposed to persuasive SMS messages that aimed at motivating them to adopt healthier nutrition behaviour by reducing the number of snacks taken in during a day. Subjects were asked to keep a nutrition diary over two weeks. In the first week, no intervention took place to obtain baseline information about the individual nutrition behaviour of each subject (baseline phase). In the second week, an SMS message with persuasive content was sent daily to each subject (treatment phase). Before the start of the experiment, the subjects' responsiveness to the different persuasive strategies was assessed by a personality questionnaire. As a result of this questionnaire, we obtained a susceptibility score for each persuasive principle so that the best- and worst-fitting persuasive principle could be identified for each subject.

Subjects were assigned randomly to three experimental groups. One group was exposed to messages that fit to their personality traits ("RIGHT" condition), another group received messages that do not fit ("WRONG" condition), and a third group obtained randomly selected messages ("RANDOM" condition). The effectiveness of the different treatment conditions was evaluated by comparing the degree of behaviour change from the baseline to the treatment phase across the different experimental groups.

Nutrition behaviour was monitored via a web-based nutrition diary. Subjects were asked to enter the number of snacks and the number of unhealthy snacks taken in over the day into a web form each evening. As daily data entries could not be enforced, the number of measurements varied between subjects. The independent variables are *condition* (WRONG, RIGHT, RANDOM) and phase (1, 2). The dependent variable is the number of unhealthy snacks (SNACKS) taken in during the day. The resulting experimental design can be represented as a four-level hierarchical data structure. Level 1 data represents the repeated measures of the dependent variable SNACKS. The second level describes the experimental phase (1: baseline; 2: treatment). The third level contains the subjects, which are clustered under the three experimental conditions on the fourth level (WRONG, RIGHT, RANDOM).

The purpose of this study was to assess whether adapting persuasive interventions to the individual's responsiveness to different persuasive principles may increase the effectiveness of the interventions. In the following, hypotheses are formulated that will be tested on the basis of the data obtained from the experiment outlined above.

Before comparing the different experimental conditions, we tested the hypothesis that there is a significant influence of the persuasive messages under each condition:

H1a: The number of unhealthy SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the WRONG condition.

H1b: The number of unhealthy SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the RIGHT condition.

H1c: The number of unhealthy SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the RANDOM condition.

We further theorized that under the WRONG condition, the reduction of the number of snacks is lower than under the RIGHT condition. We further expected that the results for the RANDOM condition lie between the WRONG and RIGHT condition:

H2: Under the WRONG condition, the decrease in the number of unhealthy SNACKS from the baseline to the treatment phase is lower than under the RANDOM condition.

H3: Under the RIGHT condition, the decrease in the number of unhealthy SNACKS from the baseline to the treatment phase is higher than under the RANDOM condition.

H4: Under the WRONG condition, the decrease in the number of unhealthy SNACKS from the baseline to the treatment phase is lower than under the RIGHT condition.

4 Data Collection

The present study relies on a data set collected in a joint research project. It has already been analyzed by Kaptein et al. (2012), who were also part of the project team, by following a different analysis approach. This earlier analysis concluded that non-fitting messages may have no or even an adverse persuasive effect, whereas fitting messages have a stronger positive effect than a random selection of principles. The goal of this study is to validate and further corroborate these findings by applying more rigorous data cleansing and a modified analysis model, which allows for conducting a significance test for the differences between the experimental conditions. This section briefly describes the development and validation of the profiling questionnaire and the persuasive messages. Furthermore, it explains the sample selection process and the experimental data collection. For details of the data collection and validation process, we refer the reader to Kaptein et al. (2012).

Before the start of the actual experiment, subjects were asked to fill in a profiling questionnaire. The questionnaire had been validated by applying Principal Components Analysis on the basis of a sample that was independent from the participants in the experiment (n=215). Results have shown moderate but sufficient reliability measures and an explained variance of 52%. The purpose of this questionnaire was to assess which persuasive principle each subject would be most susceptible to. On the basis of this assessment, the persuasive messages sent to each subject were selected.

In the treatment phase of the experiment, SMS messages were sent to the subjects as persuasive interventions. Depending on the randomly assigned condition, subjects received either a fitting (RIGHT condition), a non-fitting (WRONG condition), or a randomly selected message (RANDOM condition). Since a set of context-specific persuasive messages was not available, an expert group on persuasive technologies developed an initial set of 40 messages. The messages were restricted to four out of the six persuasive principles, namely *authority*, *commitment*, *social proof*, and *scarcity*. For the remaining two principles *reciprocity*, and *liking*, the expert group agreed that appropriate messages cannot be formulated. *Reciprocity* would require doing a favour, which makes its receiver feel obliged to return. To implement the principle of *liking*, a personal relationship would be necessary that is perceived positively. A simulation of these principles via text messages seemed to be too equivocal to be incorporated into this study. Therefore, a restriction to four principles was preferred. To validate that each message implements the intended principle, a web-based card sorting procedure with 10 persons not involved in the study was been applied (Coxon 1999; Harloff 2005).

The profiling questionnaire was completed by 333 participants, which were recruited from a panel by a professional research agency. Out of these respondents, 112 started with the experiment and entered at least one entry in the online diary. Before the start of the experiment, each respondent was assigned to one of the three experimental conditions (WRONG, RIGHT, or RANDOM). Since persuasive treatments usually require several inventions until a behavioural change occurs (Prochaska and DiClemente 1992), we eliminated all subjects from the data set that had not entered at least 2 measures in each of the two experimental phases (baseline and treatment phase). This resulted in a final data set of 476 repeated measures from 55 subjects (16 in the WRONG, 21 in the RIGHT and 18 in the RANDOM group).

5 Data Analysis

5.1 Model building

As analysis methodology, we applied Linear Mixed Models (LMM) to handle the different characteristics of the present longitudinal data set. LMMs do not require balanced samples or an equal number of measures for all subjects. Sphericity, homogeneity of regression slopes, or independence of observations are not assumed. Furthermore, LMMs account for between-subject variability, correlated error terms, heteroscedasticity, and autoregressive correlations within subjects. To find the "best" LMM for our experimental data, a top-down fitting procedure was applied as illustrated in Figure 1.

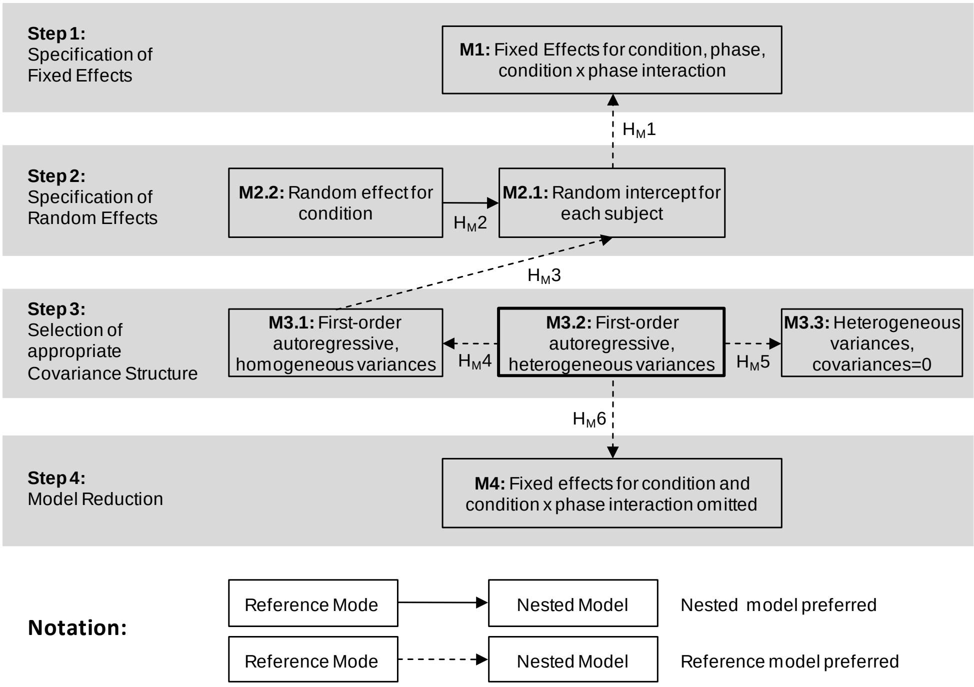

Figure 1. Top-Down LMM Fitting Procedure.

Step 1: Specification of fixed effects

To implement the top-down procedure for the given sample, we first fitted a model, which includes the fixed effects for CONDITION, PHASE and the CONDITION × PHASE interaction. CONDITION is an indicator variable that describes the experimental condition (0: "WRONG"; 1: "RIGHT"; 2: "RANDOM"). PHASE indicates whether a subject has undergone a treatment (0: no treatment; 1: treatment in accordance with condition). The interaction between the two indicator variables is included to test whether the experimental condition does influence the outcome of the treatment. The corresponding linear model can be formulated as follows:

M1: $$\text{SNACKS}_{ti} = \beta_0 + \beta_1 \times \text{CONDITION}_{ti} + \beta_2 \times \text{PHASE}_{ti} + \beta_3 \times \text{CONDITION}_{ti} \times \text{PHASE}_{ti} + \varepsilon_i$$

SNACKS_{ti} denotes the observed number of snacks for subject i at time t. β_0 is the intercept of the linear model, $\beta_{1,\,2,\,3}$ are the regression coefficients, which have to be estimated, and ε_i denotes the residual for subject i.

Step 2: Specification of random effects

In step 2, two random effects were iteratively added and tested. First, a subject-specific random effect was added for the regression intercept (model M2.1), and second, random intercepts for CONDITION were included (M2.2). The hypotheses H_M1 and H_M2 are tested by one-tailed LRTs with REML estimation to decide which of the three models is best-fitting. The two models are specified as follows:

M2.1: $SNACKS_{ti} = \beta_0 + \beta_1 \times CONDITION_{ti} + \beta_2 \times PHASE_{ti} + \beta_3 \times CONDITION_{ti} \times PHASE_{ti} + u_i + \varepsilon_{ti}$

M2.2: $SNACKS_{ti} = \beta_0 + \beta_1 \times CONDITION_{ti} + \beta_2 \times PHASE_{ti} + \beta_3 \times CONDITION_{ti} \times PHASE_{ti} + u_{0i} + u_{1i}\ x\ CONDITION + \varepsilon_{ti}$

The term u_{0i} represents the random intercept associated with subject i. The term u_{1i} denotes the random effect for CONDITION within each subject. For M2.1, a *Variance Component* (VC) matrix is specified as there is only one random effect, and therefore covariances cannot occur. In M2.2, an *unstructured* matrix is applied to allow for any variances and covariances.

Step 3: Selection of an appropriate Covariance Structure for the residuals

A commonly used covariance structure for longitudinal data is the *first-order autoregressive* structure (AR1), which implies that adjacent data points are more correlated than data points which are further apart from each other. Model M3.1 relaxes the model M2.1 such that an AR1 covariance structure is applied to the residuals. Model M3.2 further relaxes model M3.1 by allowing for heterogeneous variances across observations. For this purpose, a *heterogeneous first-order autoregressive* structure (ARH1) was applied. To test the superiority of the ARH1 structure over a more parsimonious structure that assumes uncorrelated observations within each subject, a *diagonal* structure for the **R** matrix was applied in model M3.3. The diagonal matrix allows for differing variances between observations but fixes the covariances between observations to zero. The corresponding hypotheses H_M3, H_M4 and H_M5 were tested by two-tailed LRTs with REML estimation.

Step 4: Model Reduction

As will be shown in the next section, CONDITION is the only effect that is not significant. Since eliminating CONDITION alone would neither change the number of parameters in the model nor the -2Log(L) value, the fixed effects for CONDITION and the PHASE x CONDITION interaction are omitted in this step. Consequently, we test a model in which only the experimental phase influences the outcome variable versus a model in which the outcome depends on the experimental phase, the condition, and the interaction of these two independent variables. As we are comparing models with different fixed effects in this step, ML has to be applied instead of REML estimation. Hypothesis H_M6 is tested by a two-tailed LRT. Once the best-fitting model is found, a final REML estimation is applied to obtain parameter values for further interpretation.

5.2 Model Fitting

Following the model fitting procedure described in the previous section, we first estimated a model that only contained the fixed effects of CONDITION, PHASE, and their interaction (model M1). In step 2, this model was compared to a model which contains random intercepts for each subject. The corresponding hypothesis H_M1 could be confirmed by an LRT, which means that M2.1 better explains the present sample. We can conclude from this confirmation that the intercepts vary significantly across subjects. Hypothesis H_M2 could not be confirmed; therefore a random effect for CONDITION is not included in the final model. This could be expected because a random assignment to the three conditions has been applied. Hence, there should not be a systematic variance across the subjects in the three conditions.

In step 3, we applied three different covariance structures to the residuals of M2.1. A first-order autoregressive covariance structure with heterogeneous variances turned out to best fit the underlying sample data (H_M3 and H_M4 were confirmed, H_M5 was not confirmed). Among the three fixed effects, CONDITION was not significant (p = 0.175), whereas PHASE (p<0.001) and PHASE × CONDITION (p = 0.12) were both significant. In step 4, we therefore omitted the fixed effect of CONDITION. Furthermore, we removed the interaction term from the model because otherwise the number of parameters is not reduced and the -2LL value remains equal (i.e. model fit does not change). An ML-based LRT has shown that this modification reduced the quality of the model. This means that a model, which has PHASE (i.e. treatment vs. no treatment) as the only fixed effect, explains the sample data significantly worse than a model that also takes into account the experimental condition.

To summarize, we conclude from the model fitting procedure that model M3.2 is the best-fitting model among the tested alternatives. The model fitting results for steps two to four are summarized in Table 1.

Hypothesis	**-2LL (nested)**	**Df (nested)**	**-2LL (reference)**	**Df (reference)**	**p**	**Estimation Method**	**Superior Model**
H_M1	1549.304	7	*1449.030*	*8*	<0.001	REML	M2.1
H_M2	*1449.030*	*8*	1446.548	17	0.491	REML	M2.1
H_M3	1449.030	8	*1435.850*	*9*	<0.001	REML	M3.1
H_M4	1435.850	9	*1383.811*	*23*	<0.001	REML	M3.2
H_M5	1396.217	22	*1383.811*	*23*	<0.001	REML	M3.2
H_M6	1387.688	19	*1373.111*	*23*	0.006	ML	M3.2
Values for the superior model in each iteration are marked in italics.							

Table 1. Model Fitting Results.

The final model applied for our further analysis differs in two ways from the model applied by Kaptein et al. (2012). First, a random intercept for each subject was added to account for between-subject variability. Between-subject variability can be assumed as study participants can be expected to have varying baseline levels in their snacking behaviour. Second, an autoregressive covariance structure was added to capture autoregressive correlations within subjects. Our model fitting procedure has shown that these extensions improve the quality of a more restricted model that accounts only for the fixed effects of PHASE, CONDITION, and the CONDITION × PHASE interaction.

5.3 Descriptive Analysis

Before the best-fitting LMM identified in the previous section was applied to the data, we explored the results of our experiment by evaluating graphical illustrations and descriptive statistics. Figure 2 shows the development of the snacking behaviour over time for each experimental group. The time axis in this graph represents the days since the beginning of the experiment. In time point 0, all groups started at almost the same number of snacks per day. Despite some variation, on average the number of snacks remained quite stable in the baseline phase.

Starting with day six (i.e., the seventh day of the experiment), subjects received a persuasive message each day. We can see that the number of snacks has decreased sharply in all three groups on this day. This negative trend was continued for the RIGHT and RANDOM group over the following days. However, the WRONG group showed a different and more constant trend (with a strong outlier on day 12). The graphical evaluations of our data gives strong support for our expectation that persuasive treatment is less effective if an inappropriate persuasive strategy is applied than if an appropriate strategy is adopted. Furthermore, there is strong evidence that there is only a marginal difference between selecting the most appropriate strategy and a random selection of persuasive messages. Finally, descriptive analysis indicates that adopting the most inappropriate strategy might even have a slightly counterproductive effect. In the following section, the observations gained from this graphical analysis will be tested by applying the LMM developed in the previous section.

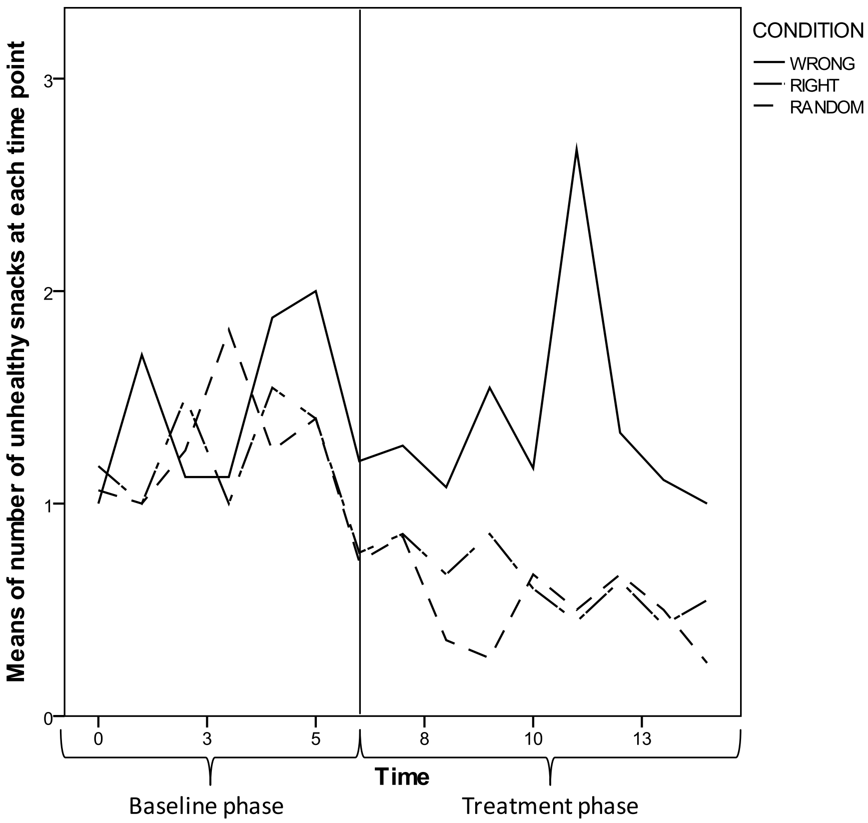

Figure 2. *Number of unhealthy snacks at each time point.*

5.4 Experiment Evaluation

After having identified the best-fitting model, a final REML estimation was conducted on model M3.2. On the basis of the estimation results, the hypotheses formulated above were tested. Table 2 presents the results of the significance tests (Type III F-tests) for the fixed effects of CONDITION, PHASE, and CONDITION × PHASE. We can see that CONDITION alone is not a significant predictor for the outcome variable SNACKS. This could be expected as the experimental condition does not account for treatment. Instead, the effect of PHASE is highly significant, which means that the outcome variable significantly varies between the baseline phase and the treatment phase. Furthermore, the significant interaction term shows that the treatment is significantly different for the different experimental groups.

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	51.376	102.651	.000
CONDITION	2	51.309	1.804	.175
PHASE	1	166.874	13.182	.000
CONDITION x PHASE	2	166.037	4.584	.012

Table 2. *Significance Tests of Fixed Effects.*

With these results, we can neither infer about the direction of the effects nor about the differences between the experimental conditions. Therefore, we next inspected the parameter estimates for the fixed effects (Table 3). As all fixed effect factors are categorical variables, the parameter estimates are represented as contrasts against the reference category, which is RANDOM for CONDITION and TREATMENT for PHASE. Therefore, these parameter levels are set to zero in the contrasts. The contrasts show that the effect of CONDITION alone is significant if we compare WRONG to

RANDOM, but is not significant if we compare RIGHT to RANDOM. However, this finding is not sufficient for the hypotheses to be tested as it does not yet account for the experimental phase.

Parameter	**Estimate**	**Std. Error**	**df**	**t**	**Sig.**	**95% Confidence Interval**	
						Lower bound	**Upper bound**
Intercept	.470015	.174075	57.545	2.700	.009	.121508	.818522
[condition=1]	.799186	.254795	58.328	3.137	.003	.289221	1.309152
[condition=2]	.184098	.235917	56.866	.780	.438	-.288342	.656539
[condition=3]	0	0	.	.	.	.	.
[phase=1]	.671197	.165197	157.774	4.063	.000	.344913	.997480
[phase=2]	0	0	.	.	.	.	.
[condition=1] * [phase=1]	-.731320	.245238	165.584	-2.982	.003	-1.215516	-.247123
[condition=1] * [phase=2]	0	0	.	.	.	.	.
[condition=2] * [phase=1]	-.228387	.226725	164.268	-1.007	.315	-.676058	.219284
[condition=2] * [phase=2]	0	0	.	.	.	.	.
[condition=3] * [phase=1]	0	0	.	.	.	.	.
[condition=3] * [phase=2]	0	0	.	.	.	.	.
Variable encodings: CONDITION = 1 : WRONG; 2 : RIGHT ; 3 : RANDOM PHASE = 1 : BASELINE ; 2 : TREATMENT							

Table 3. Parameter Estimates for Fixed Effects.

Next, we see that PHASE has a significant effect on the dependent variable SNACKS. Across all conditions, a change from phase 1 to phase 2 resulted in a reduction of 0.67 snacks per day. Again, we cannot infer from this result to our hypotheses as it does not account for the different experimental conditions. Therefore we next inspected the results for the CONDITION × PHASE interaction.

The contrasts for the different levels of CONDITION show that under the WRONG condition, the effect of PHASE is 0.731 units weaker than under the RANDOM condition. As this effect is significant, hypothesis **H2** is confirmed. Furthermore, the effect of PHASE under the RIGHT condition is 0.228 units weaker than under the RANDOM condition, but this interaction effect is not significant. Therefore, we reject hypothesis **H3**, which assumes that the RIGHT condition has a positive influence on the effect of PHASE.

To obtain a contrast between the conditions WRONG and RIGHT, the LMM was estimated for a subset of the original data, in which the RANDOM condition was omitted (Table 4). A significant parameter estimate of -0.521 supports our expectation that under the WRONG condition, the treatment effect is lower than under the RIGHT condition. We therefore accept hypothesis **H4**.

Parameter	**Estimate**	**Std. Error**	**df**	**t**	**Sig.**	**95% Confidence Interval**	
						Lower bound	**Upper bound**
Intercept	.653121	.176390	38.699	3.703	.001	.296250	1.009992
[condition=1]	.614281	.271936	40.343	2.259	.029	.064824	1.163738
[condition=2]	0	0	.	.	.	.	.
[phase=1]	.413610	.148377	120.619	2.788	.006	.119850	.707370
[phase=2]	0	0	.	.	.	.	.
[condition=1] * [phase=1]	-.521289	.227117	119.450	-2.295	.023	-.970987	-.071591
[condition=1] * [phase=2]	0	0	.	.	.	.	.
[condition=2] * [phase=1]	0	0	.	.	.	.	.
[condition=2] * [phase=2]	0	0	.	.	.	.	.

Table 4. Parameter Estimates for Fixed Effects (WRONG vs. RIGHT Condition).

Having compared the different conditions against each other, we next assessed whether the treatment with persuasive messages has a significant effect on SNACKS under each condition. For this purpose, we conducted a post-hoc test with LSD adjustment (a more conservative adjustment than LSD is not necessary because the PHASE variable has only two levels). This test estimates the so-called *estimated marginal means*, which are group means estimated from the fitted model. The test then conducts a pairwise comparison of the means for the different factor levels (Table 5).

For the WRONG condition, the estimate for the average number of snacks increases from 1.209 in the baseline phase to 1.269 in the treatment phase (+5%), which is not significant. We therefore reject hypothesis **H1a**, which postulates a decrease of the number of snacks. Under the RIGHT condition, the number of snacks decreases by 0.443 (-40%), which is highly significant with a p-value of 0.005. We therefore accept hypothesis **H1b**. Hypothesis **H1c** is also accepted since the number of snacks under the RANDOM condition decreases by 0.671 (-59%) at a significance level of 0.000.

Condition	**Mean (Baseline)**	**Mean (Treatment)**	**Mean Difference**	**Std. Error**	**df**	**Sig.**	**95% Confidence Intervall**	
							Lower Bound	**Upper Bound**
WRONG	1.209	1.269	-.060	.181	168.46	.741	-.418	.298
RIGHT	1.097	.654	.443	.155	168.08	.005	.136	.749
RANDOM	1.141	.470	.671	.165	157.77	.000	.345	.997

Table 5. Post-Hoc Test.

To summarize, the LMM analysis confirmed the expectations we gained from Figure 2. Whereas under the RIGHT and RANDOM condition we found a significant effect of the persuasive treatment, this effect could not be confirmed for the WRONG group (hypothesis H1a rejected, H1b, H1c accepted). Although not significant, the number of snacks even increased slightly when non-fitting messages were sent to the subjects. Furthermore, we found that there is no significant difference of the treatment effect between the RIGHT and RANDOM (hypotheses H2 and H4 accepted, H3 rejected).

6 Conclusions

The present study has corroborated findings from prior research that people differ in their level of susceptibility towards different persuasive principles. Applying the most appropriate persuasive strategy can be expected to exert a strong persuasive effect. It was expected that this effect will be significantly weaker when an inappropriate principle is applied. In accordance with Kaptein et al. (2012), the study has shown that applying an inappropriate persuasive principle may not only weaken the persuasive effect, but can even reverse the effect with the result that subjects change their behaviour contrary to the intended direction (i.e., the results have shown that under the wrong treatment, the number has slightly increased, but not at a significant level). This indicates that inappropriate treatment causes an aversion towards the desired behaviour, which may lead to defiant reactions to act against the extrinsic motivation attempt. However, our analysis has shown that this effect is not significant. We conclude that a wrong treatment has no reliable effect, neither in the positive nor in the negative sense. Further studies should be focused on the question whether applying a wrong principle may lead to an adverse effect.

In contrast to the findings from Kaptein et al. (2012), a random treatment has not shown to be less effective than the most appropriate treatment. Since a random treatment exerts a balanced set of appropriate and inappropriate stimuli, we would have expected that its persuasive effect lies between the right and the wrong treatment. In fact, we even observed that the random treatment performed slightly better than the right treatment. We can only speculate about the underlying reasons for this finding. We assume that not only the appropriateness of a persuasive principle influences its effectiveness, but also the variety of the messages. Applying the same principle several times may lead to an annoyance or boredom effect so that a basically appropriate principle loses its effectiveness if it is applied more often. If this is true, the relative advantage of the most appropriate principle decreases over time, which makes less appropriate principles relatively more effective. Furthermore, the applied model allowed us to investigate whether the difference between a right and a random treatment is statistically significant. We found that this difference is not significant, so we can conclude that both treatments are equally effective. Further studies should investigate whether the observed finding can be confirmed, since this will have strong practical implications for the implementation of persuasive principles.

Our results have several practical implications for the design of persuasive technologies. As applying an inappropriate principle may exert no or even a reverse effect on the subject, single strategy implementations will be ineffective for a part of its users. Two alternatives are useful to consider. First, one could assess which principle is most appropriate for a subject and then apply it. Second, one could apply a random selection of principles, which will be equally effective over time.

Selecting the right principle may be an impractical approach for many realistic applications. We could confirm that it is possible to determine the susceptibility of a subject to the six persuasive principles. Theoretically, we could ask a user of a persuasive application to fill in such a questionnaire before he is exposed to the optimal persuasive treatment, but practically this will be hardly acceptable in many cases. Therefore more acceptable approaches are necessary to assess the susceptibility of a user to the different persuasive principles. For desktop or mobile applications, it might be possible to achieve an appropriate assessment by designing a game approach. If an application is expected to be used over a longer period of time, it might also be possible to implement a learning algorithm that tests different principles and evaluates their effectiveness before the final principle is selected.

Applying random selection seems to be more applicable in many contexts as it does not require for determining the most appropriate persuasive principle. However, it requires a treatment period that is long enough to apply different principles. If the persuasive treatment is only applied once or twice, many subjects will be exposed solely to inappropriate principles so that the effectiveness may be lower than by applying the most appropriate principle. Instead, if many treatments can be expected, a random selection is preferable as it requires less effort to retrieve a user profile.

In summary our analysis extended the work of Kaptein et al. (2012) with the finding that there is no significant difference between applying a random or a tailored persuasive strategy. From a theoretical point of view, more psychological studies are required to understand the underlying reasons for this result. From a practical point of view, this finding allows designers to mitigate the additional effort for profiling as a random mix of persuasive strategies seems to be equally effective.

References

Cacioppo, J.T., Petty, R. E., Kao, C. F., and Rodriguez, R. 1986. "Central and Peripheral Routes to Persuasion: An Individual Difference Perspective," Journal of Personality and Social Psychology (51:5), pp. 1032–1043.

Cialdini, R. B. 2008. Influence, Science and Practice, (5th ed.) Boston: Prentice Hall.

Cialdini, R. B., Trost, M. R., and Newsom, J. T. 1995. "Preference for Consistency: The Development of a Valid Measure and the Discovery of Surprising Behavioral Implications," Journal of Personality and Social Psychology (69:2), pp. 318–328.

Cialdini, R. B., Wosinska, W., Barrett, D. W., Butner, J., and Gornik-Durose, M. 1999. "Compliance with a Request in Two Cultures: The Differential Influence of Social Proof and Commitment/Consistency on Collectivists and Individualists," Personality and Social Psychology Bulletin (25:10), pp. 1242–1253.

Coxon, A. P. M. 1999. Sorting Data. Collection and Analysis, Thousand Oaks: Sage.

Fogg, B. J. 2002. Persuasive Technology, San Francisco: Morgan Kaufmann.

Goldberg, L. R. 1993. "The Structure of Phenotypic Personality Traits," The American Psychologist (48:1), pp. 26–34.

Guadagno, R. E., Asher, T., Demaine, L. J., and Cialdini, R. B. 2001. “When Saying Yes Leads to Saying No: Preference for Consistency and the Reverse Foot-in-the-Door Effect,” Personality and Social Psychology Bulletin (27:7), pp. 859–867.

Halko, S., and Kientz, J. 2010. “Personality and Persuasive Technology: An Exploratory Study on Health-Promoting Mobile Applications,” In Proceedings of the 5th International Conference on Persuasive Technology, T. Ploug, P. Hasle, and H. Oinas-Kukkonen (eds.), (Vol. 6137)Berlin / Heidelberg: Springer, pp. 150–161.

Harloff, J. 2005. “Multiple Level Weighted Card Sorting,” Methodology: European Journal of Research Methods for the Behavioral and Social Sciences (1:4), pp. 119–128.

Kaptein, M., and Eckles, D. 2010. “Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling,” In Persuasive Technology, T. Ploug, P. Hasle, and H. Oinas-Kukkonen (eds.), Berlin / Heidelberg: Springer, pp. 82–93.

Kaptein, M., Lacroix, J., and Saini, P. 2010. “Individual Differences in Persuadability in the Health Promotion Domain,” In Proceedings of the 5th International Conference on Persuasive Technology, T. Ploug, P. Hasle, and H. Oinas-Kukkonen (eds.), (Vol. 6137)Springer Berlin / Heidelberg, pp. 94–105.

Kaptein, M., Markopoulos, P., De Ruyter, B., and Aarts, E. 2009. “Can You Be Persuaded? Individual Differences in Susceptibility to Persuasion,” In Proceedings of Interact 2009, T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R. Prates, and M. Winckler (eds.), (Vol. 5726)Springer Berlin / Heidelberg, pp. 115–118.

Kaptein, M., De Ruyter, B., Markopoulos, P., and Aarts, E. 2012. “Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking,” ACM Transactions on Interactive Intelligent Systems (2:2).

Kellermann, K., and Cole, T. 1994. “Classifying Compliance Gaining Messages: Taxonomic Disorder and Strategic Confusion,” Communication Theory (4:1), pp. 3–60.

Levine, T., and Wheeless, R. 1990. “Cross-Situational Consistency and Use/Nonuse Tendencies in Compliance-gaining Tactic Selection,” Southern Communication Journal (56:1), pp. 1–11.

Marwell, G., and Schmitt, D. R. 1967. “Dimensions of Compliance-Gaining Behavior: An Empirical Analysis,” Sociometry (30:4), pp. 350–364.

Moon, Y. 2002. “Personalization and Personality: Some Effects of Customizing Message Style Based on Consumer Personality,” Journal of Consumer Psychology (12:4), pp. 313–326.

Noar, S. M., Benac, C. N., and Harris, M. S. 2007. “Does Tailoring Matter? Meta-Analytic Review of Tailored Print Health Behavior Change Interventions,” Psychological Bulletin (133:4), pp. 673–93.

Prochaska, J. O., and DiClemente, C. C. 1992. “In Search of how People Change: Applications to Addictive Behaviors,” American Psychologist (47:9), pp. 1102–14.