ResearchGate

See discussions, stats, and author profiles for this publication at: http://www.researchgate.net/publication/254461605

Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking

ARTICLE *in* TRANSACTIONS ON INTERACTIVE INTELLIGENT SYSTEMS · JUNE 2012

DOI: 10.1145/2209310.2209313

CITATIONS
17

READS
171

4 AUTHORS, INCLUDING:

Maurits Kaptein
Radboud University, Donders Centre for Cog…
42 PUBLICATIONS **214** CITATIONS
SEE PROFILE

Boris de Ruyter
Philips
79 PUBLICATIONS **935** CITATIONS
SEE PROFILE

Panos Markopoulos
Technische Universiteit Eindhoven
236 PUBLICATIONS **1,897** CITATIONS
SEE PROFILE

Available from: Panos Markopoulos
Retrieved on: 29 November 2015

Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking

MAURITS KAPTEIN, Eindhoven University of Technology/Philips Research
BORIS DE RUYTER, Philips Research
PANOS MARKOPOULOS, Eindhoven University of Technology
EMILE AARTS, Eindhoven University of Technology/Philips Research

10

This article describes the use of personalized short text messages (SMS) to reduce snacking. First, we describe the development and validation (N = 215) of a questionnaire to measure individual susceptibility to different social influence strategies. To evaluate the external validity of this Susceptibility to Persuasion Scale (STPS) we set up a two week text-messaging intervention that used text messages implementing social influence strategies as prompts to reduce snacking behavior. In this experiment (N = 73) we show that messages that are personalized (tailored) to the individual based on their scores on the STPS, lead to a higher decrease in snacking consumption than randomized messages or messages that are not tailored (contra-tailored) to the individual. We discuss the importance of this finding for the design of persuasive systems and detail how designers can use tailoring at the level of social influence strategies to increase the effects of their persuasive technologies.

Categories and Subject Descriptors: H.1.2 [**Models and Principles**]: User/Machine Systems—*Human factors*

General Terms: Design, Measurement

Additional Key Words and Phrases: Persuasion, social influence, personalization, tailoring, persuasion profiling

ACM Reference Format:
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 Trans. Interact. Intell. Syst. 2, 2, Article 10 (June 2012), 25 pages.
DOI = 10.1145/2209310.2209313 http://doi.acm.org/10.1145/2209310.2209313

1. INTRODUCTION

Persuasive technologies—technologies that are intentionally designed to influence the attitudes or behavior of users [Fogg 2003]—are entering the public domain [Oinas-Kukkonen and Harjumaa 2009]. Lately, a number of technologically mediated or initiated interventions have focused on supporting people in maintaining a healthy lifestyle [Consolvo et al. 2009]. This is facilitated by the fact that the problems associated with an unhealthy lifestyle are becoming more and more evident [Patrick et al. 2009], the costs associated with overall health care are increasing [Brug et al. 2005], and technological means to reach people are becoming widespread. In this article we evaluate the use of personalized text messages on a mobile device to encourage people

Author's address: M. Kaptein, Eindhoven University of Technology, Den Dolech 2, 5600 MB Eindhoven, the Netherlands. HG 2.54; email: maurits@mauritskaptein.com.
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from the Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, USA, fax +1 (212) 869-0481, or permissions@acm.org.
© 2012 ACM 2160-6455/2012/06-ART10 $10.00
DOI 10.1145/2209310.2209313 http://doi.acm.org/10.1145/2209310.2209313

to adopt a healthier dietary regime. Novel in our approach is a focus on individual differences in people's responses to the different types of social influence strategies used in the text messages.

1.1. Healthy Lifestyle

A healthy lifestyle is largely dependent on people's own choices of behaviors [Lacroix et al. 2009]. Healthy behaviors such as frequent physical activity, abstention from smoking, and a healthy dietary regime all are major determinants of general health [World Health Organization 2003]. Both physical activity and dietary habits are strongly related to the growing trend of obesity [Kromhout et al. 2001]. Obesity has increased significantly in recent years [Ogden et al. 2006] and has led to increased risks of several chronic diseases such as hypertension and diabetes [He et al. 2000]. Furthermore, obesity increases the risks for cardiovascular diseases [Must et al. 1999] and overall mortality [Muennig et al. 2000]. These serious consequences underline the relevance of research efforts that help people make the right behavioral choices.

While people generally have favorable attitudes towards healthy behaviors they often encounter difficulties when trying to maintain a workout schedule or keep to a dietary regime. For those motivated but unable to adopt a healthy lifestyle, external interventions can be very effective. Active coaching by experts, goal setting, and counseling can be beneficial for people when trying to adopt healthy behaviors [Lacroix et al. 2009].

1.2. Persuasive Technologies for For a Healthy Lifestyle

In response to health problems due to a unhealthy lifestyle, public awareness of the importance of a healthy lifestyle has increased considerably. The field of health promotion is quickly expanding and offers several commercially available health solutions. Starting from public campaigns by governments and health professionals, the field nowadays employs a multidisciplinary approach, integrating insights and methods from multiple domains, to optimize the effectiveness of health interventions. Within this context, persuasive technologies can play a key role. Persuasive technologies can provide a cost-effective means to employ large scale, personalized interventions [Fogg 2003].

Researchers and designers of persuasive technologies have already focused on bringing about socially desirable changes in attitudes and behaviors [Lockton et al. 2008]. Applications have been designed and tested, which influence people to (among others) smoke less [Räisänen et al. 2008], lose weight [Maheshwari et al. 2008], or maintain a healthy workout regime [Lacroix et al. 2009]. One of the benefits of employing persuasive technologies, as opposed to public campaigns or human interventions, is that the interventions can, theoretically, easily be tailored for specific individuals. Using interactive persuasive technologies, one can both alter the content of the interventions, as well as the timing. Several behavior change theories have made this tailoring of messages a central theme in health intervention programs (e.g., Kroeze et al. [2006], Neville et al. [2009], Johnson et al. [2008]).

Health-related persuasive technologies already exist in commercial form. Services, or product-service combinations, like Philips DirectLife, MiLife and FitBug ([Hurling et al. 2008; Lacroix et al. 2008]) make an attempt at influencing people to adopt a healthier lifestyle. The majority of these services use implementations of influence strategies and other theories from motivation and persuasion research to gain compliance and change attitudes or behaviors. Typically, these products make life easier by automatically monitoring the user's behavior through wearable sensors. Next, Web sites or mobile applications allow for the presentation of feedback and coaching.

Through the use of strategies like goal setting, tailored encouraging feedback, and social proof, these services support users to make positive changes in their physical activity behavior or nutritional intake [Lacroix et al. 2009].

1.3. Text Messages for a Healthy Diet

In this article we focus on influencing snacking behavior using Short Text Messages (SMS's) on a mobile phone. Mobile phones have frequently been used as platforms, to employ persuasive technologies because of their pervasiveness in everyday life, and their ability to be at the right place, at the right time [Kass 2007]. Mobile applications have been developed and tested in all areas of the persuasive technology field, from applications that help chronic patients manage their disease [Franklin et al. 2008], to persuasive services that promote the sexual health of teenagers [Parkes et al. 2005]. In all instances researchers as well as users recognize the power of mobile devices for pervasive persuasion attempts.

Mobile interventions for a healthy diet have not only been part of the design research space, but have also made their way into randomized controlled clinical trials. Patrick et al. [2009] show a larger effect on weight loss of an SMS-based intervention program versus a paper-based one over a four month time period. They conclude that "text messages might prove to be a productive channel of communication to promote behaviors that support weight loss in overweight adults." In another paper [Consolvo et al. 2009], the same group of researchers argue in favor of the mobile phone as a general carrier for health related interventions and underline the impact such technologies can have.

Despite the convincing results by Patrick et al. [2009] and the overwhelming interest in mobile applications by persuasive technology researchers and developers, not all health-related interventions using mobile phones have proven successful. An experimental study by McGraa [2010] showed little effect of SMS messages on people's dietary or workout regimes over a five week period—this even though the intervention was clearly designed incorporating state-of-the-art social science findings on influence and persuasion. In this article we argue that the absence of this effect might be primarily due to the absence of tailoring: adaptation to individual preferences or traits.

1.4. Usage of Influence Strategies

McGraa [2010] used a number of so called *social influence strategies* as identified by Cialdini [2001] to increase the possible effectiveness of the SMS messages used in their trial. While these social influence strategies are not the only means that can be employed by persuasive technologies to influence users, we believe they provide a powerful means for tailoring. In the remainder of this article we focus on the use of implementations of these social influence strategies to influence snacking behavior.

Among scholars, there is debate on the number of influence strategies that exist: estimates range from six [Cialdini 2001] to more than 100 [Rhoads 2007]. These differences originate from differences in definitions and granularity. Some scholars identify each influence attempt separately, while others group multiple influence attempts that function through similar psychological processes under a common heading. We take the latter approach by making a clear distinction between an *influence strategy*—the general description of the psychological process that produces the persuasion—and its *implementation(s)*. A single influence strategy can be implemented in a number of different ways. In accordance with McGraa [2010], we subscribe to the taxonomy of

6 general social influence strategies that all function through different psychological principles, as proposed by Cialdini [2001].

In his elaborate book about the topic Cialdini [2001] describes and explains the effectiveness of implementations of *reciprocity*, people feel obligated to return a favor [James and Bolstein 1992], *scarcity*, people will value scarce products [West 1975], *authority*, people value the opinion of experts [Milgram 1974], *consistency*, people do as they said they would [Cialdini 2001; Deutsch and Gerard 1955], *consensus*, people do as other people do [Cialdini 2004; Ajzen and Fishbein 1980], and *liking*, we say yes to people we like. Cialdini [2001] has shown the positive average effects of implementations of these social influence strategies on people's compliance to persuasive requests.

Social influence strategies can be regarded as means by which a request for an eventual end (e.g. to decrease snacking frequency) is made [Kaptein and Eckles 2010]. To illustrate: an SMS message could prompt a user to do a workout that evening (the end goal) by, (a) sending a suggestion from a fitness expert, or (b) showing that a number of the users' friends are working out that evening. While in both cases the end is the same, the means supporting the request are different.

Despite numerous studies that show the effectiveness of implementations of any of these influence strategies, there are also accounts of unexpected failures (e.g. McGraa [2010]). This, and similar, results [Gerber et al. 2009] make it hard for designers of persuasive systems to estimate the effectiveness of implementations of influence strategies in the interventions they design.

1.5. Individual Differences in Responses to Influence Strategies

Social psychologists, while trying to explain the sometimes mixed findings originating from studies examining the effects of influence strategies, have focused on possible individual differences in responses to these strategies. Early on, Cacioppo et al. [1986] identified the Need for Cognition (NfC) as a construct that "people's tendency to think." They showed through a number of experimental evaluations that individuals high in NfC are more likely to scrutinize arguments based on their content than react to cues associated with, but peripheral to, the central arguments of, the advocacy. Thus while authority arguments seem effective overall, the effect on compliance of adding an authority argument will be less for those high in NfC than for those low in NfC. By showing that NfC is a stable trait, Cacioppo et al. [1986] suggests the efficacy of using a number of social influence strategies tailored for those low in NfC.

In addition to NfC—which can be regarded as an individual difference in people's response to implementations of any influence strategy—investigators have also focused on people's responses to specific influence strategies. Cialdini et al. [1995] has shown that people's preference for consistency positively correlates with their behavioral responses to implementations of the Commitment and Consistency strategy. For example, people who score high on preference for consistency are more likely to comply to a request made using the "Foot in the Door" technique [Guadagno et al. 2001]—an implementation of the consistency strategy.

Kaptein et al. [2009] extended the idea of Cialdini by attempting to measure the susceptibility of different individuals to each of the six social influence strategies. In an initial study using a 12 item questionnaire, 2 items per strategy, they show a strong correlation between measured susceptibility to consensus, the strategy and behavioral responses to a request accompanied by this strategy. There was a clear linear relationship between susceptibility to consensus and the number of email addresses participants provided when this request was supported by a message stating that "80% of all participants provided one or more email addresses to us." These and other similar results [Kaptein et al. 2010] lead us to believe that tailoring based on individual

susceptibility to influence strategies employed by a persuasive system might be beneficial in improving its effectiveness.

1.6. Outline

In this article we design and implement a personalized persuasive application and employ the system for a period of two weeks. During this period we explicitly test the effects of tailoring of social influence strategies by, (a) measuring self-reported susceptibility to distinct social influence strategies using a susceptibility to persuasion scale, and (b) adopting the SMS messages that participants receive accordingly. We first describe the development of the susceptibility to persuasion scale (STPS), which elaborates on the 12-item version presented by Kaptein et al. [2009]. We then validate the scale by administering it to 215 participants and examining both its internal and external validity.

Next, we describe the design and employment of an SMS-based health intervention geared at reducing snacking behavior. We describe the development and selection of the persuasive messages that implement different social influence strategies. To estimate the effects of tailoring we compare a group of participants who receive messages that are specifically selected based on their *susceptibility profile* to groups that receive, (a) messages with random implementations of influence strategies, or (b) messages with implementations of strategies that they are least susceptible to according to their scores on the STPS. We examine the effects of this tailoring or contratailoring on compliance: a reduction in self-reported snack intake. Finally, we provide a number of guidelines for designers intending to build systems that use personalized persuasion based on social influence strategies.

2. MEASURING SUSCEPTIBILITY TO INFLUENCE STRATEGIES

Kaptein et al. [2009] describe a 12-item susceptibility to persuasion scale that they use to predict responses to a behavioral request implementing the consensus strategy. We develop this scale further by adding a number of items for each of the six latent variables of interest. These items measure the susceptibility of people to implementations of the six social influence strategies identified by Cialdini [2001]. In this section we describe which items are added to the original scale and how these came about. The internal reliability of the scale is evaluated by administering the scale to $N = 215$ participants.

2.1. Item Construction

Starting from the 12-item questionnaire presented in Kaptein et al. [2009] (see Table I) we created additional items for each of the six latent variables of interest. In a session with a group of five persuasive technology researchers, eight or more items per variable were composed. Items were constructed to fit the underlying latent variable as much as possible and to appeal both to specific instances of the influence strategy (e.g. "I always follow advice from my general practitioner") as well as to broad statements of the latent variable of interest (e.g. "I am very inclined to listen to authority figures"). In a pretest with $N = 9$ participants, the understandability and clarity of each of the items was evaluated and per variable 5–6 items were selected for further testing. The complete item-set used for evaluation is presented in Table I.

2.2. Scale Validation

To determine the internal validity of the *Susceptibility To Persuasive Strategies scale* (STPS) we administered the scale to $N = 215$ participants. All participants filled out

Table I. Items Used to Measure Susceptibility to Persuasion

Items marked with * are selected after the factor analysis and used in the final trial. The italicized items were used in the initial evaluation of the STPS [Kaptein et al. 2009].

Principle	Abbreviation	Susceptibility item
Reciprocity	Recip_1∗	*When a family member does me a favor, I am very inclined to return this favor.*
	Recip_2∗	*I always pay back a favor.*
	Recip_3∗	If someone does something for me, I try to do something of similar value to repay the favor.
	Recip_4∗	When I receive a gift, I feel obliged to return a gift.
	Recip_5∗	When someone helps me with my work, I try to pay them back.
Scarcity	Scarce_6∗	*I believe rare products (scarce) are more valuable than mass products.*
	Scarce_7∗	*When my favorite shop is about to close, I would visit it since it is my last chance.*
	Scarce_8∗	I would feel good if I was the last person to be able to buy something.
	Scarce_9∗	When my favorite shampoo is almost out of stock I buy two bottles.
	Scarce_10∗	Products that are hard to get represent a special value.
Authority	Auth_11	*I always follow advice from my general practitioner.*
	Auth_12	*When a professor tells me something I tend to believe it is true.*
	Auth_13∗	I am very inclined to listen to authority figures.
	Auth_14∗	I always obey directions from my superiors.
	Auth_15∗	I am more inclined to listen to an authority figure than a peer.
	Auth_16∗	I am more likely to do something if told, than when asked.
Commitment	Commit_17∗	*Whenever I commit to an appointment I always follow through.*
	Commit_18∗	*I try to do everything I have promised to do.*
	Commit_19∗	When I make plans I commit to them by writing them down.
	Commit_20	Telling friends about my future plans helps me to carry them out.
	Commit_21∗	Once I have committed to do something I will surely do it.
	Commit_22∗	If I miss an appointment, I always make it up.
Consensus	Consens_23∗	*If someone from my social network notifies me about a good book, I tend to read it.*
	Consens_24∗	*When I am in a new situation I look at others to see what I should do.*
	Consens_25	I will do something as long as I know there are others doing it too.
	Consens_26∗	I often rely on other people to know what I should do.
	Consens_27∗	It is important to me to fit in.
Liking	Like_28	*I accept advice from my social network.*
	Like_28	*When I like someone, I am more inclined to believe him or her.*
	Like_28∗	I will do a favor for people that I like.
	Like_28∗	The opinions of friends are more important than the opinions of others.
	Like_28∗	If I am unsure, I will usually side with someone I like.

both the 32 items of the susceptibility scale as well as the 18-item Need for Cognition scale [Cacioppo et al. 1986]. Need For Cognition (NfC) was included to asses the external validity of the overall scale by comparison to a known construct.

2.2.1. Method. The items were administered using a 7-point scale. The endpoints of the scale were labeled with "Completely Disagree" and "Completely Agree", where a score of 7 marked "Completely Agree". The midpoints of the scale were not labeled. Participants were provided with a "Don't Know" option for each item.

All 215 participants fully completed the questionnaire. The full questionnaire, including demographics and NfC, took around 25 minutes to complete. The average age of the participants was 23.0 years (SD = 7.4). Of the participants, 167 were female (75.9%). All participants were US college students enrolled in an introductory research methods class and participated for course credit. Participants filled out the questionnaire online using their own PCs after receiving an email with a link to

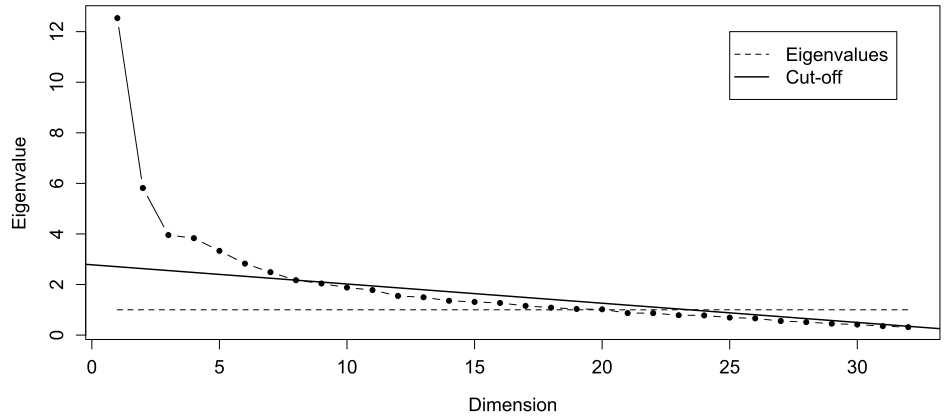

Fig. 1. Scree plot showing the eigenvalues (y-axis) of each of the extracted components (x-axis). A clear increase is visible from component 6 upwards indicating a proper fit of a six factor solution.

the study. Participants were given one week to participate at a time of their own choosing. Approval by the Institutional Review Board was obtained prior to conducting the study.

2.2.2. Results. To check whether the collected data on the 32-item scale is suited for factor analysis we first performed several well-known diagnostic checks [Mulaik 1972].

(1) 30 of the 32 items correlated at least .3 with one other item.
(2) The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.81, well above the recommended value of .6.
(3) Bartlett's test of sphericity was statistically significant ($\chi^2(496) = 2353.8$, $p < 0.001$).
(4) All of the communalities were above .3.

Given these positive indications an exploratory factor analysis was conducted including all 32-items that were administered. To determine the appropriate number of factors we first inspected the scree plot. Figure 1 shows the eigenvalues of all the factors with a line superimposed to indicate the clearest cutoff. It is clear that a six factor solution—as expected based on the latent variables used to construct the items—seems like an appropriate fit for this data.

Principal components analysis (PCA) was used because the primary purpose was to identify and compute composite susceptibility scores for each of the latent variables. The total cumulative variance explained by all six components was 52%, which is in a common range for multidimensional constructs. The six component solution was further examined using Oblimin rotation. Table II gives an overview of the loadings of each of the items on the components. Because of relatively low loadings, or high cross-loadings, we decided to remove six items from the scale: *Auth_11*, *Auth_12*, *Commit_20*, *Consens_25*, *Like_28*, and *Like_32*. Refitting the six-component solution to the 26-item scale led to a cumulative variance of 56%, and no cross-loadings of factors over 0.3. In Table I the 26 items used to compute scores on the six latent variables of the STPS are marked with an *.

For each of the six factors we computed a composite score by averaging over the 3–5 items in each subscale. Table III presents each of the subscales with their appropriate descriptives. Overall, this analysis indicates that the six factors underlying the STPS are moderately internally consistent. The correlations between the scores on

Table II. Factor Loadings Based on a Principle Components Analysis with Oblimin Rotation for 32 Items of the STPS

Loadings smaller than .3 are suppressed.

	Comp. 2	Comp. 3	Comp. 1	Comp. 4	Comp. 6	Comp. 5
Recip_1			0.34			
Recip_2			0.72			
Recip_3			0.73			
Recip_4			0.67			
Recip_5			0.67			
Scarce_6					0.77	
Scarce_7					0.33	
Scarce_8					0.44	
Scarce_9					0.34	
Scarce_10					0.84	
Auth_11	0.37	0.45				
Auth_12		0.44	0.40			
Auth_13		0.73				
Auth_14		0.71				
Auth_15		0.75				
Auth_16		0.59				
Commit_17	0.84					
Commit_18	0.63					
Commit_19	0.51					
Commit_20						
Commit_21	0.80					
Commit_22	0.76					
Consens_23				0.54		
Consens_24				0.53		
Consens_25		0.40		0.42		
Consens_26				0.63		
Consens_27				0.53		
Like_28				0.64		
Like_29						0.58
Like_30						0.70
Like_31						0.51
Like_32				0.53		

Table III. Overview of the Composite Scores of the STPS

Presented are the mean, standard deviation, and Cronbach's α of each of its subscales.

Subscale	# items	Mean (SD)	Cronbach's α
Reciprocity	5	5.3 (0.83)	0.75
Scarcity	5	4.7 (0.98)	0.63
Authority	4	4.3 (1.10)	0.75
Commitment	5	5.1 (0.97)	0.81
Consensus	4	4.1 (0.98)	0.60
Liking	3	5.1 (0.91)	0.61

the subscales range from 0.2 to 0.4. As a check of the external validity of the scale we examined the correlation between the total composite score of participants on the STPS and their score on NfC (Cronbach's $\alpha = 0.89$). The correlation between these two scales was -0.14, $p < 0.05$.

2.3. Conclusions

Based on the validation of the STPS scale we decide to proceed with our influence strategy, performing specific measurements using the 26-item STPS scale. The subscales are moderately internally reliable, and the correlations between the

separate subscales are relatively low, indicating that the STPS indeed measures people's susceptibility to 6 distinct strategies. The negative direction of the correlation with NfC shows that—as expected—those high in overall susceptibility to persuasion score low on NfC. The relative size of this correlation shows that the STPS does not merely measure the inverse of the NfC but rather a (set of) distinct trait(s).

In our further evaluation of the use of tailored text messages to reduce snacking behavior we will use the 26-item STPS as put forward in Table I to determine users' a priori susceptibility to different social influence strategies. In the next section we describe the generation and evaluation of text messages that are used in the final intervention.

3. DESIGNING PERSUASIVE MESSAGES

To create a text message intervention to help reduce snacking behavior it is necessary to design the specific messages that will be used in the trial. Since we aim to evaluate the benefits of tailoring the messages to the susceptibility of the user as measured by the STPS, we designed a set of text messages that could be employed on a mobile device (e.g. # chars. $<$ 140) and which each implemented a distinct social influence strategy. Similar design efforts have been undertaken for example by Patrick et al. [2009]. They describe their use of over 3000 different MMS messages in a randomized clinical trial. They distinguish between *topic* messages (e.g. "Control your portions by setting aside a large snack package into smaller bags or buy 100 calorie snack packs!"), *questions* (e.g. "How often do you meal plan?"), and tips (e.g. "In a rush? Buy pre-cut vegetables like carrots, celery, and mushrooms for a quick, easy, and low calorie snack!"). These messages however were not designed specifically to implement different social influence strategies. To support future message design efforts we detail our design process.

3.1. Designing Implementations of Persuasive Strategies

To design the messages, two persuasive technology researchers independently tried to generate as many messages as possible for each of the six strategies identified by Cialdini [2001] that would be usable given the context. In total, 42 messages where created. After combining the lists, it became evident that both the *Liking* as well as the *Reciprocity* strategies were hard to implement properly in the context of mobile text messaging to reduce snacking behavior. For the *Liking* strategy to be successfully implemented there is need for a bond between the persuader and the receiver of the message. Given that there is no clear social actor in play that receivers of the message could relate to, this strategy was hard to implement. For the *Reciprocity* strategy to be most effective, a favor has to be done to the persuadee prior to the persuasive request. This strategy was proven hard to implement in technology mediated contexts before Kaptein and Eckles [2010] and was as such omitted.

We ended up with a number of implementations (40 in total), of four social influence strategies, for use in our desired context. In this set of messages, there were 13 messages that aimed to implement the *Authority* strategy, 11 that tried to implement the *Consensus* strategy, nine that aimed to implement the *Scarcity* strategy, and seven for the *Commitment* strategy.

While each of the messages was created by experienced researchers to implement a specific strategy, it is unclear whether they succeeded to appeal to the type of reasoning associated with the specific strategy. Do people really perceive, for example, the message, *"According to Weightwatchers snacking can seriously increase obesity"* as an appeal to authority? While Weightwatchers is an obvious authority in this context it is less clear how much a reader of this message actually perceives the strategy to be

implemented. Even more, given a number of empirical results that show the effectiveness of these social influence strategies to be stronger when not processed consciously, it is not evident that participants would be able to reliably indicate the expected effectiveness and its reason(s) when this information is elicited in, for example, a questionnaire study.

To enable evaluation of the created messages and select a number of messages to be used in the final intervention, we set up an evaluation with ten researchers with expertise in the field of human computer interaction. Each researcher received a small description of each of the four social influence strategies that we tried to implement to become familiar with these terms. Next, researchers were shown the 40 messages one by one, and were asked to categorize them into the following categories: (a) implements the Authority strategy, (b) implements the Consensus strategy, (c) implements the Scarcity strategy, (d) implements the Commitment strategy, or (e) Other / Don't know.

After the ten researchers rated each of the messages, we analyzed the ratings by looking at agreement matrices between the researcher-ascribed categories. For the *Authority* strategy there was a general high agreement: for all of the messages that were intended to implement this strategy, at least 70% of the researchers ascribed the message to this strategy. We proceeded by selecting three of the implementations designed to implement the Authority strategy, one which was ascribed to this strategy by all raters, the other two by 90% of the raters. Similarly we picked three implementations of the Consensus strategy, three of the Scarcity strategy, and three of the Commitment strategy. It has to be noted that the messages that implemented the Scarcity strategy were the least identifiable, with one of the three selected messages only ascribed to this strategy by 60% of the raters.

Table IV gives an overview of the messages that were selected, the strategies they aim to implement, and the agreement between raters. With three implementations of each strategy we are able to, (a) select messages based on the respective strategy that they implement, and (b) reduce repetition of sent messages.

3.2. Conclusions

We chose to design text messages for four of the six persuasive strategies identified by Cialdini. For each strategy, three messages were selected. Each message that was selected could unambiguously be assigned as implementing the respective strategy by a panel of experts who independently rated each of the messages. The final messages that were are selected and used in the evaluation of the adaptive text messaging for snacking reduction trial are presented in Table IV. The next section describes the setup and evaluation of a system that implements tailored messages.

4. ADAPTIVE PERSUASIVE MESSAGES *IN SITU*

To evaluate whether the STPS can aid in message selection and whether tailoring of influence strategy has the desired effects, we set up a two week text-messaging intervention that used messages to implement persuasive strategies as prompts to reduce snacking behavior. As described in the introduction, text messaging has had mixed effects in previous attempts toward this end. In this study, we tried to evaluate whether messages that are tailored to the susceptibility of individual users are more effective than messages that are not tailored.

Since our primary aim was to examine the feasibility of tailoring on the level of influence strategies, the current study compares the following three conditions. (1) *The tailored condition* (TC), in which messages that implement the strategy that the current user is most susceptible to, given their scores on the STPS, are presented, (2) *the contra-tailored condition* (CTC), in which implementations of strategies that the

Table IV. The Messages Used in the Intervention

For each of the four social influence strategies used in this trial three implementations are used.

Strategy	Message	Agree
Authority	Try not to snack today. According to the College of Physicians this is an easy way to lead a healthier life.	100%
Authority	Dietitians advise having 3 meals a day without snacking. Try to reduce snacking.	90%
Authority	The World Health Organization advices not to snack. Snacking is not good for you.	90%
Consensus	90% of people benefit from reducing snacking between meals. It will boost your energy and you will live a healthier life.	90%
Consensus	Everybody agrees: not snacking between meals helps you to stay healthy.	90%
Consensus	Reduce snacking. You are not on your own: 95% of participants have already reduced snacking.	90%
Commitment	The aim of this study is to live healthier. Reducing snacking is a way to achieve that.	100%
Commitment	Try to obtain your goal for living a healthier life by not snacking. You are committed!	90%
Commitment	You have to continue what you've started: you are participating in this test to lead a healthier life. Reduce snacking.	90%
Scarcity	There is only one chance a day to reduce snacking. Take that chance today!	90%
Scarcity	This test lasts only 3 weeks: you have the unique opportunity to enhance your health by reducing snacking.	70%
Scarcity	Today is a unique opportunity to lead a healthy life. Reduce snacking.	60%

current user is least susceptible to are sent to the user, and finally (3) *the random condition* (RC) in which users receive a random selection out of all the created messages. We hypothesize—based on previous evidence of lowered compliance to contra-tailored messages in the health promotion domain [Kaptein et al. 2010]—that tailored messages will be more effective than random messages or contra-tailored messages, to reduce snack consumption.

4.1. Method

A two-week trial to evaluate the different message conditions between subjects was set up. Since snacking behavior varies substantially between people, we chose to include a one-week baseline assessment of individual snacking behavior before introducing the different messaging conditions (between-subject) and establishing their effects on the snacking behavior within a single user.

4.1.1. Participants. Participants for this second Study were recruited via a professional recruitment agency. A call for participation was sent out via email to potential Dutch participants between 18 and 65 years of age, with fluent understanding of English, and in possession of a mobile phone. The call for participation detailed that the study would take two full weeks and would entail filling out several questionnaires and receiving daily text messages on their mobile phones. In total, 334 potential participants clicked on the link that took them to a designated Web site and started the introduction questionnaire.

At the end of the introduction questionnaire, participants were asked to provide their mobile phone numbers. After providing their phone number a text message with an activation code to log in at the designated Web site was sent to participants. In total 162 participants fully completed this signup process and activated their study participation.

After signing up, participants received text messages for a period of two weeks (2×5 days, workdays only). Participants were instructed—both prior to the study as well as via the text messages—to go to the designated Web site every evening to fill out a short diary. The first week was used to establish a baseline snacking frequency for each participant, while the intervention was employed in the second week. We included for our final analysis only those participants who filled in at least one diary during each of the two weeks (e.g. during the baseline measurement and during the intervention). Our final sample was composed of 73 participants. The average age of the participants was 34.9 years (SD = 11.1). Of our final sample 32 (43.8%) were females. Upon completion, participants were awarded research participation credits with a monetary value of two Euros (default amount provided by the research agency).

4.1.2. Procedure & Measurents. After going to the designated study Web site all participants filled out a small questionnaire regarding their snacking behavior, their shopping behavior, and their motivation to decrease snacking. Next, participants were administered the STPS and they provided their mobile phone numbers to sign up for the text messaging part of the study. Participants then received one text message a day (on workdays) for a period of two weeks, and subsequently filled out a small online diary every day.

The introduction questionnaire included all 32-items of the initial STPS scale, although only the 26 items that scored consistently on the factors of interest were used to allocate participants to conditions (See 4.1.3). In addition to the STPS, the questionnaire included the following questions.

— How many times week do you usually visit a supermarket to buy ingredients to prepare a dinner?
— How many times a week do you prepare your own meals?
— Would you like to eat healthier? (Scored *Yes*, *No*)
— Do you feel you generally eat healthy dinners? (Scored *Yes*, *No*)

Finally, participants were asked for their age and gender and proceeded to the sign-up procedure.

During the sign-up procedure participants provided their mobile phone number and received a text message with an eight Digit authorization code. Participants filled in their authorization code on the study Web site to create a personal profile and supplied a user name and password for later logins. After logging in with their user name and password, participants were asked to fill in the first diary. The diary—in the first week—consisted of the following questions.

— How many snacks did you have today? (Open ended)
— How many *unhealthy* snacks did you have today? (Open ended)
— How healthy was your nutrition today? (Five-point scale, *Very unhealthy* to *Very healthy*)

For the first week, participants received one text message a day, which asked them to fill in their diary for that day. Messages were automatically sent to all participants between 5 and 6 pm. This phase of the research was the baseline period (Phase 1).

After Phase 1, participants entered the second week (Phase 2), in which they received the persuasive messages according to the experimental condition to which they were allocated (See 4.1.3 for details). The persuasive message contained an implementation of either the Authority, Consensus, Commitment, or Scarcity strategy, as described in 3.1. After receiving the persuasive message, participants were again asked to fill out their online diaries. In Phase 2, one additional question was added to the

daily diary: "How useful was the text message you received?" scored (1) *Not at all useful* to (6) *Very useful*.

After receiving five persuasive messages during Phase 2 and filling out their last diary, participants received a "thank you for your participation" message on their screen. On this final screen, was a link to a Web site with more information about snacking. On this final Web site, participants found four different types of information about snacking: (a) information provided by experts, (b) "Read what others are doing," (c) "How to stick to your dietary goals," and (d) "This is your last chance to learn more about snacking." These different types of information were presented in random order to participants and implemented the Authority, Consensus, Commitment, and Scarcity strategies respectively. Even though exploration of this closing Web site was completely voluntary, we monitored participants' click-through behavior as a possible indicator of their persuasion preferences.

4.1.3. Conditions. Based on participants' answers to the STPS a mean score on each of the four variables of interest—their self-reported susceptibility to each of the four strategies—was computed. Next, participants were randomly assigned to one of the following three conditions.

(1) *The tailored condition* (TC). Participants assigned to this condition received, during Phase 2, messages that were randomly selected implementations of the two strategies (of the four implemented in this study) that they had the highest susceptibility scores on. Hence, the messages were tailored to their personal profile to be most effective.
(2) *The contratailored condition* (CTC). Participants assigned to this condition received random implementations of the two strategies they had the lowest mean scores on as judged from the STPS. Hence, the messages were tailored to be the least effective.
(3) *The random condition* (RC). Participants in this condition received randomly selected messages out of the full set of messages presented in Table IV.

The study setup, including this allocation over conditions, enabled us to study the change in snacking behavior over the course of two weeks (within subjects) between the three different conditions (between subjects).

4.2. Results

4.2.1. Overview. In total, 506 diaries were filled out by the 73 participants included in our analysis. Figure 2 shows the frequency distribution of the number of diaries filled out per participant. It is clear that a large proportion of the participants filled out all ten diaries and thus were very engaged in the study during the full study period. As stated before, only participants that filled out at least one diary during both of the experimental phases were included.

Of our included participants a large majority (80.8%) indicated that they were motivated to eat healthier. This was true even though most participants stated that they already had healthy eating habbits (89.0%). All of our participants indicated that they prepare (cook) their own meals at least once a week, and 45.2% indicated that they prepare their own meals more than five times a week. 95.9% of our participants visited a supermarket to purchase food at least once a week. These figures indicate that our participants were largely individually responsible for their own food consumption and shopping habits and thus personal text-messages could possible influence their behavior. Based on the scores on the STPS, 22.8% of our respondents indicated that they were most susceptible to the scarcity strategy, 14.9% by the authority strategy, 53.1% by the commitment strategy, and 9.1% by the consensus strategy.

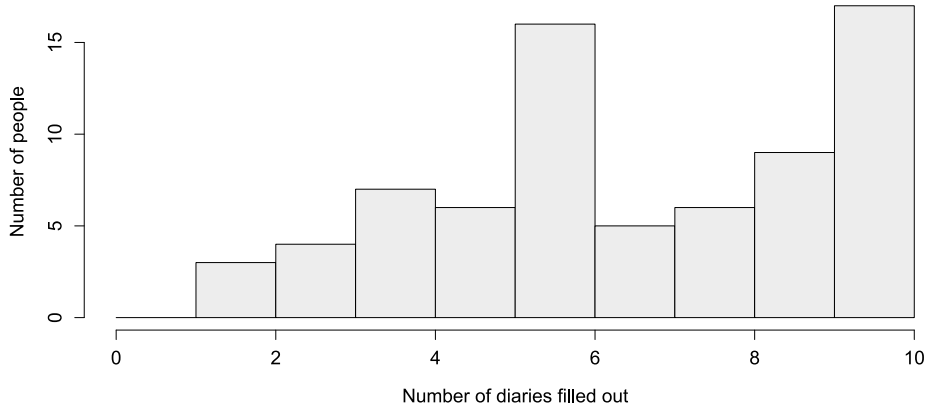

Fig. 2. Overview of the number of diaries filled out per participant.

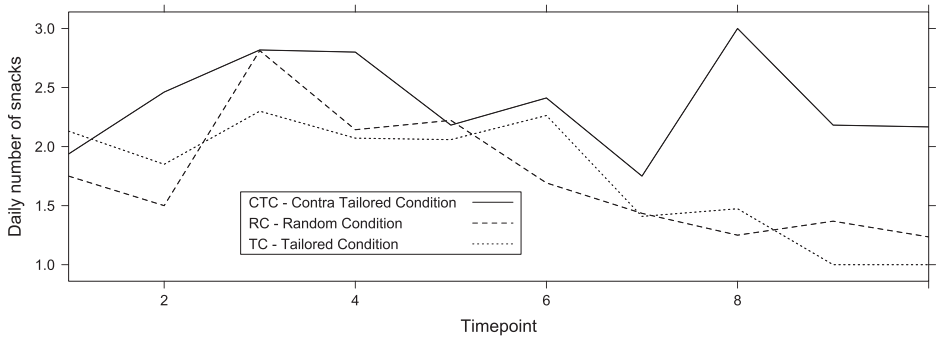

Fig. 3. Overview of the average number of snacks eaten each day by our participants separated for the three experimental conditions.

4.2.2. Snacking Behavior. The primary test to see whether messages tailored to participants' scores on the STPS can be effective in reducing snacking behavior is provided by an examination of the progression of the self reported snacking behavior during the experiment between the three experimental groups. The daily diary contained three questions that are indicative of the effect of tailored or contratailored messages on snacking behavior. We examined each of these separately.

The primary measurement—the self-reported number of snacks eaten by participants each day—is graphically represented in Figure 3. It is clear from this figure that—while variable over the days—the snacking consumption decreased over time in both the *RC* and the *TC* conditions, while it did not decrease in the *CTC*. It is evident that the decrease is largest from time-point 6 onwards: this corresponds to the entry into Phase 2 of the experiment, and thus the actual separation of messages between the three conditions.

To statistically test the effects of our conditions over time—days in the experiment—we fit a multilevel model with varying intercepts for participants. This allows us to reliably estimate the effects of time on the number of snacks consumed in each of the conditions despite the missing data on some of the time points and the differences in baseline snacking behavior between participants. We start by fitting a null model [Snijders and Bosker 1999], which can formally be written as

$$y_{ij} \sim \mathcal{N}(\mu_j, \sigma^2_{err}), \tag{1}$$

Table V. Comparing the Null Model with Models Including a Time Effect and Different Time Effects for Each Condition

Model	Df	AIC	$logLik$	χ^2	$Pr(>Chi^2)$
A: Null model	3	1857.70	−925.85		
B: Model including time	4	1849.64	−920.82	10.07	<.01
C: Time, split for phase 1 and phase 2	5	1843.39	−916.70	8.24	<.01
D: Time and Time P2 × Condition	7	1837.79	−911.89	9.60	<0.01

Table VI. Overview of the Fixed Effects of the Model D Including an Interaction Between Time and Condition to Predict Snacking Behavior

Empirical p-values are based on MCMC simulations.

	Estimate	Std. Error	t-value	p
Intercept	2.16	0.19	11.33	<.001
Time Phase 1	0.05	0.04	1.07	.28
Phase 2: Time:CTC (contra tailored)	−0.01	0.08	−0.16	.85
Phase 2: Time:RC (random)	−0.22	0.07	−3.30	<.001
Phase 2: Time:TC (tailored)	−0.30	0.07	−4.29	<.001

with $\mu_j \sim \mathcal{N}(\bar{\mu}, \sigma_\mu^2)$ for $j = 1, ..., N = 73$ subjects. Interpretation of this model is straightforward. The number of snacks consumed, y_{ij}, by each subject on each day is explained by per-participant intercepts μ_j, which are modeled as drawn from a normal distribution with a grand mean $\bar{\mu}$ and a variance of σ_μ^2. A restricted likelihood ratio test on our data shows that σ_μ^2 deviates significantly from zero, justifying use of separate intercepts for each participant ($p < 0.001$).

From the null model, which does not include any information about time or phase of the trial nor information about the experimental conditions, we build a model that includes both time as well as condition to explain the snacking behavior of our participants. Adding time—centered around the phase shift—as a fixed factor to the model (Thus $y_{ij} \sim \mathcal{N}(\mu_j + \beta_{time} X_{time}, \sigma_{err}^2)$) significantly increased model fit (see Table V, model B). In this model the coefficient of time is significantly negative, $\beta time = -0.06$, $t = -3.18$, $p < 0.01$, indicating an overall decrease in the number of snacks consumed by our participants during our trial.

After examining the effects of time, a phase × time interaction was added to the model to allow for different effects in the baseline and treatment phase. This significantly increases model fit, see Table V model C, which indicates that the effects of the messages over time differs for the two phases. Finally, to test whether the conditions significantly influence the snacking behavior of our respondents, we fit a model in which time during phase two interacts with condition—essentially fitting separate time effects for the different conditions during phase two. Here again, model comparisons show a significant increase of model fit; see Table V model D.

Table VI shows the fixed effects of model D—the model allowing for different slopes for the different conditions during phase 2 of the trial. These fixed effects show that even though averaged over all conditions snack intake decreased significantly during our intervention, reality is more granular: During the baseline phase the number of snacks does not decrease significantly. During the treatment phase the number of snacks consumed by our participants decreases significantly only for RC and TC participants. The fixed effects table also indicates that the decrease in snack consumption is higher for those in the TC condition than those in the RC condition.

Figure 4 shows the number of snacks per time-point—day in the experiment—for each participant. Both the time-point as well as the number of snacks are jittered to prevent overlap. Imposed on these raw data points, are the overall time trend (light

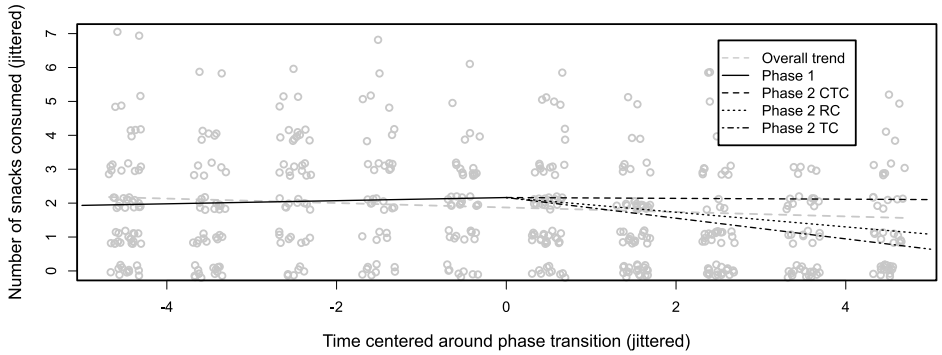

Fig. 4. Overview of the number of snacks eaten each day by our participants with the regression lines of the effects over time of the different conditions superimposed. The *CTC* line represents the predicted number of snacks eaten during the intervention period in the Contra Tailored Condition. *RC* and *TC* represent the Random Condition and the Tailored Condition respectively.

Table VII. Overview of the Fixed Effects of the Model Including an Interaction Between Time and Condition Predicting Perceived Healthiness of Nutrition

	Estimate	Std. Error	*t*-value	*p*
Intercept	3.44	0.08	41.43	<.001
Time Phase 1	0.01	0.03	0.51	.61
Phase 2: Time:CTC (contra-tailored)	−0.03	0.05	−0.58	.52
Phase 2: Time:RC (random)	0.05	0.04	1.35	.13
Phase 2: Time:TC (tailored)	0.08	0.04	1.88	<.05

gray), the trend in during the first phase of the trial (black solid line), and the three time trajectories during the treatment phase. Of these latter three, the solid line shows the *CTC* fit, the dashed line shows the *RC* fit, and the dotted line shows the *TC* fit. Participants who received tailored messages decreased their snacking intake during phase two of the experiment. Participants who received random messages also decreased their snack intake but did so to a smaller extent. Finally, there is no evidence that participants who received messages that where contra-tailored decreased their snack intake.

Results were very similar for the number of unhealthy snacks that participants ate. Again the model including an interaction between condition and time was favored by likelihood ratio tests. Snacking significantly decreased in the random condition, $\beta_{Time:RC} = -0.13$, $t = -2.49$, $p < .01$. The effect was even larger in the tailored condition, $\beta_{Time:TC} = -0.21$, $t = -3.34$, $p < .01$. The effect of time in phase 2 for the CTC was not significantly different from zero, $\beta_{Time:CTC} = -0.02$, $t = -0.31$, $p = .68$.

Next to the self-reported number of consumed snacks the perceived healthiness of the participant's diet as a function of the text-messages (and the progression over time) is an interesting indicator of the success of the intervention. Analyzing the progression over time of participants' response to the question "How healthy was your nutrition today" shows that the perceived healthiness of participants' nutrition is relatively stable over time (See Table VII). There is a significant increase during the treatment phase only for the *TC* participants. Figure 5 shows the ratings of our participants at each time-point (again jittered to prevent overlap). It is clear that the healthiness ratings are, overall, relatively stable over time. However, the *CTC* participants perceived their diet during the treatment phase as slightly less healthy than in the baseline phase, while for *CT* participants the opposite was the case.

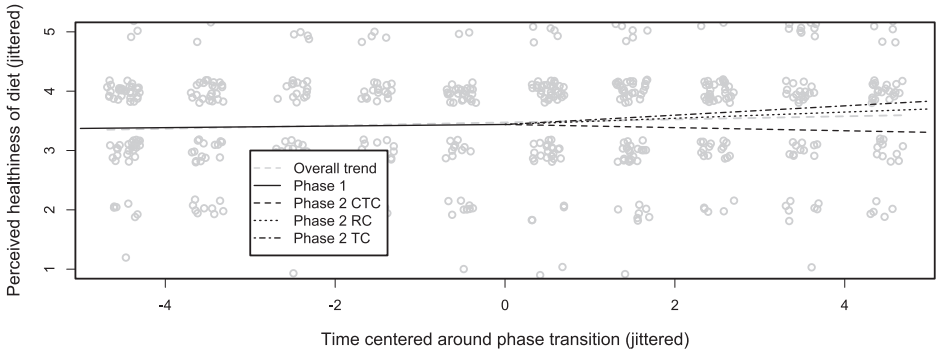

Fig. 5. Overview of the perceived healthiness of daily nutrition score for participants with the regression lines of the effects over time of the different conditions superimposed. The *CTC* line represents the predicted scores in the Contra Tailored Condition. *RC* and *TC* represent the Random Condition and the Tailored Condition respectively.

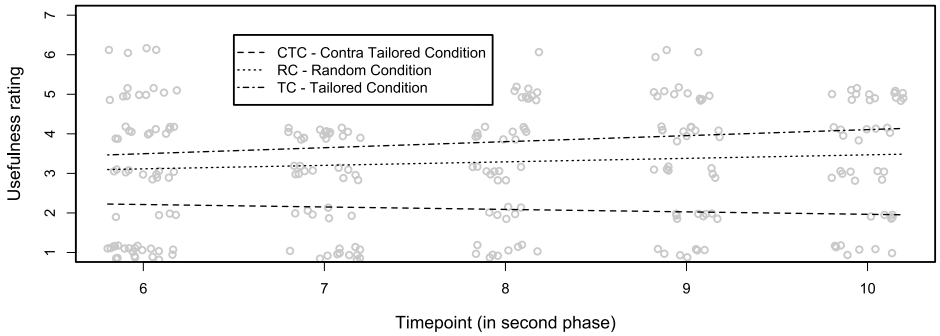

Fig. 6. Overview of the usefulness ratings over time. Added, are the estimated regression lines of the effect of time. The *CTC* line represents the predicted scores in the Contra Tailored Condition. *RC* and *TC* represent the Random Condition and the Tailored Condition respectively.

Table VIII. Overview of the Fixed Effects of the Model Including an Interaction Between Time and Condition Predicting Perceived Usefulness of the Received Message

	Estimate	Std. Error	*t*-value	*p*
Intercept	2.58	0.31	8.32	<.001
Time:CTC (contra-tailored)	−0.06	0.05	−1.36	.12
Time:RC (random)	0.09	0.04	2.17	<0.01
Time:TC (tailored)	0.15	0.04	3.50	<0.001

4.2.3. Additional Analysis. During phase 2 of the experiment (time-points 6 to 10) participants were asked to evaluate the usefulness of the messages they received. Figure 6 shows the usefulness evaluations per time-point with the fixed effect line of the effect of time for each of the three experimental conditions superimposed. The fixed effects of the interaction between time and condition on the perceived usefulness ratings are presented in Table VIII. It is clear that the text messages are perceived as more useful in the tailored condition than in the contra-tailored condition, with the random condition scoring in between.

A number of respondents dropped out of the study so we decided to look at the dropout rates in more detail. Since only from the start of Phase 2, do the messages that are received by participants differ for the three different experimental conditions,

we decided to look closely at the dropouts caused in the second phase of the study as a function of the experimental condition. Participants who started phase 2 by filling out their diary at either time-point 6 or 7 but who did not finish by filling out diary number 10 were marked as dropouts. In total 61.1% of our participants dropped out during phase 2. When looking at the different conditions it is clear that the dropout rate was higher among CTC participants (63.8%) than among CT participants (56.6%). The effect of condition on dropout rate in phase 2 is however not statistically significant, $\chi^2(2, 72) = 4.83$, $p = 0.08$. No significant differences with the RC participants were observed.

Finally, we analyzed whether the scores on the STPS predict information-seeking after the experiment. As mentioned before, participants arrived—after completing the full study—at a Web site giving more information about snacking. The information was presented in four categories, in a random order. The categories reflected the four persuasive strategies used in the experiment and previously assessed by the STPS. Only 13 participants out of the full participant pool ended up searching for more information. Of these 13 participants who decided to look for more information 10 (76.9%) chose a category that was ranked among their top two susceptibility scores as judged from the STPS. This is significantly higher than one would expect if a random choice was made, $Z = 3.026$, < 0.01.

4.3. Conclusions

Our text-messaging trial, which was an attempt to test the external validity of measurements obtained using the STPS in an in situ intervention, showed that while persuasive text-messages can be effective in changing people's behavior (snacking) and attitude (perceived healthiness and usefulness), these changes depend on the right choice of influence strategy for the right participant. Using contratailored persuasive messages, and hence messages participants thought they would not be susceptible to, did not decrease snacking behavior, and led to a low estimation of the usefulness of the text messaging intervention. Conversely, the tailored persuasive messages, those adapted to fit with the profile derived using the STPS, lead to a decrease in snacking behavior and an increase in the perceived usefulness of the intervention. Not surprisingly, the randomly selected messages had an intermediate effect: while these messages where perceived as useful and decreased snacking behavior they were less effective than the tailored messages. This result, which is consistent for both the healthiness measure, and for the usefulness measure, shows that persuasive text-messaging interventions are likely to be more effective if they are tailored towards the receiver of the message.

The type of influence strategy that is used, proved to be a useful variable for tailoring, and the STPS was used successfully to determine the favorable strategies for each participant. This latter claim is also strengthened by the power of the STPS scores to predict information seeking after the trial. Finally, it is important to note that the effects of contra-tailoring, in this study, seem larger than the effects of tailoring. Compared to a random message selection explicitly picking the wrong strategy leads to a strongly adverse reaction which led to an ineffective intervention for CTC participants.

5. DISCUSSION

In this article we described the development of a text based intervention to decrease snacking behavior. To tailor the use of different influence strategies to a specific user we first designed and evaluated the STPS: a scale which measures the susceptibility of a user to each of the six social influence strategies identified by Cialdini [Cialdini

2001]. After internally validating this scale we created and evaluated different persuasive messages which could be used during a text-based intervention to decrease snacking and we confirmed that each message belonged to one of four selected persuasive strategies. In our final evaluation we showed the external validity of the STPS: users who received messages tailored to their susceptibility as measured using the STPS were more likely to decrease snacking than those who received messages which they indicated to be least susceptible to.

We hope that these results are valuable to both researchers and practitioners. The STPS, and its composite scores, as presented in this article can be readily used to tailor persuasive interventions to individuals. Furthermore, the emphasis on influence-strategies as a level of analysis important for tailoring, provides new research and design directions. While most research efforts have focused on adapting the timing, or personalizing the type of request that is made by a persuasive application, we believe that the usage influence strategies as a level of personalization can prove beneficial in future personalizing attempts.

5.1. Persuasion Profiles

Using the composite scores on the STPS we created a profile for each of our participants. This profile, indicating the estimated effectiveness of different influence strategies for a given individual is called a Persuasion Profile. During the Persuasive 2010 Conference, Kaptein and Eckles [2010] elaborated on the impact such profiles might have and how they can be created. We think it is worth mentioning here that the profile based on the STPS proved useful in this applied setting. We showed the internal consistency of the scale, and we demonstrated its predictive validity during the two-week trial. A dynamic persuasion profile, in which a persuasive application monitors the behavioral response from a user to adapt its estimates—and hence updates the persuasion profile—can likely benefit from an initialization based on an individuals score on the STPS.

The large effect of the contra-tailored condition in our trial showed that perhaps the most important emphasis of the profile should be on the prevention of the use of badly chosen influence strategies. It is worth future research effort to determine whether the STPS is suited to uniquely identify the majority of contraproductive influence strategies for each user—given a specific context—and can thus be used to prevent backfiring in case of wrongly selected influence attempts. This identification is especially important when, as in this study, the influence attempt is geared towards a behavior that is of importance for the health or wellbeing of the user. Badly chosen influence attempts could in these situations lead to an increase of the unhealthy behavior which the intervention intends to decrease.

Persuasion profiles, which are demonstrated in the trial presented in this article, are distinct from other profiles that are already in use in (e.g.) marketing. While online stores most probably keep a profile of the product preferences of consumers, this profile is likely, (a) disclosed to you, and (b) primarily useful for that specific online store. The disclosure is eminent on sites that state: "This product has been selected especially for you." While disclosing this ends-oriented tailoring seems to increase buying behavior, there is evidence [Kaptein et al. 2011a] that similar disclosure for means-oriented tailoring—as is done in when disclosing one's persuasion profile—has the adverse effect. Saying that you receive the authority argument because that will persuade you most decreases compliance to the authority argument. Researchers and practitioners will have to find ways in which both privacy and autonomy of users (or consumers) can be assured while not undermining the effects of using persuasion profiles for socially desirable ends.

The difference between specific—mostly product preference oriented—profiles and general profiles that can be shared across contexts is also clearly indicated in our presented trial. The STPS questions participants about, for instance, "when my favorite shampoo is almost out of stock I buy two new bottles." While this question is totally outside of the domain of the snacking behavior trial the information obtained proved useful. Thus, profiles on the level of influence strategy effectiveness seem stable across multiple contexts. This implies that the profile created by an online store—based on sales behavior—can be of use for text-messaging interventions like the one described here.

We believe that the most powerful forms of digital persuasion will emerge from a combination of the customization of both the ends of a request—as in a sales setting, is done by classical recommender systems—and the means, by using a persuasion profile. Customization of not merely the type of influence strategy that is used, but also the goals that are presented to individual users and the timing of messages will make digital technologies more persuasive and more pervasive. As a first step, the current work could be extended to include snacking advice that relates to user's current individual snacking behavior, or that is presented at a time just prior the decision to eat snacks. Such additional customization is, however promising, currently hardly guided by theoretical frameworks that would inform designers about the exact timing of such messages or the exact content.

5.2. Ethical Implications

We have shown that with personalization of the influence strategies, one uses the effectiveness of a text-messaging trial to decrease snacking behavior can be improved. The persuasion profile—based on the STPS—that was used for this purpose seems to be a reliable indicator for which strategies one should employ for one specific users, and which strategies one should not employ.

While the target behavior in this setting is generally regarded as positive this might not always be the case. Persuasion profiles might also be used for marketing purposes or political ends. Here, the two distinct features of these profiles, a lack of transparency and the possibility of using the profile across multiple contexts, raise serious questions about its use in these settings. On the other hand, as shown in our trial, a failure to adapt to individuals can lead to persuasive technologies not delivering their promise. If an application designed to help you, for example, stop smoking, increases your smoking behavior by choosing the wrong influence strategies, the failure to adapt could be considered unethical. Instances of backfiring due to wrong influence strategy selection are already reported in the literature [e.g., Kaptein et al. 2010].

As far as we are aware there is currently no regulation regarding persuasion profiles. While there are guidelines posed by the Federal Trade Commission (USA) on the use of behavioral targeting in online stores, we feel that persuasion profiles are sufficiently distinct to warrant separate attention (See also Kaptein et al. [2011b]). We hope that by highlighting both the possible positive as well as negative consequences, policy makers and researchers can engage in a meaningful discussion on the topic. We believe that a key characteristic of persuasion profiles that has not been empirically settled and relates heavily to its ethical implications, is the size of the expected effect of proper adaptation. Will tailored persuasive applications be able to steer human behavior into undesired directions or is a mere nudge in an already feasible direction the best designers can do? In the latter case the nudge should be made as effective as possible. The former however presents a dark scenario, which would speak in favor of abandoning profiling efforts.

5.3. Design Implications

In this article we successfully used a profile created using the STPS to differentiate between different text messaging interventions. In this section we discuss three core features a system should have to be able to implement the use of persuasion profiles.

— *Identification*. To create, maintain, and use persuasion profiles, it is first of all important for individual users to be identified. Identification here does not imply keeping a record of personally identifiable information but does entail the coupling of the persuasion profile to a specific individual. In our trial the mobile phone number provided a unique key to identify our participants. In different settings different keys (e.g. cookies, bluetooth keys, RFID) can be used for the same purpose.
— *Estimation*. A second necessary feature to build adaptive persuasive systems is a means of estimating the effectiveness of the use of implementations of distinct persuasive strategies. In our trial, we used the STPS for this purpose but this approach might not always be feasible. In some contexts it will be easier to observe behavioral responses to implementations of persuasive strategies than to administer a questionnaire. In both cases however it is important that the strategy effects are estimated reliably, maintained (and possibly updated dynamically), and tied to an identifier.
— *Presentation*. To build systems that adapt their persuasive interventions at the level of influence strategies, is the ability to present different strategies in a meaningful way in the given context. In our trial we were only able to implement four of the six strategies targeted by the STPS in a way that was perceived to be consistent and credible given the delivery medium and the relationship to the technology. Strategies like *Reciprocity* are hard to implement on single occasions and not every context or medium allows for the interchangeable usage of several strategies.

A final requirement for building dynamic adaptive persuasive systems—systems that update the persuasion profile over time—is the ability to monitor the effectiveness of different persuasive strategies. In our trial only self-reported snacking behavior was available as a measure of effectiveness during the trial and this measure was not used to update the strategy effectiveness estimates. However, when behavioral measures can be obtained and linked to the personal identifier a dynamic profile can be employed.

5.4. Limitations

The text-messaging trial showed a positive effect on snacking behavior of tailored messages as opposed to random or contratailored messages. This study however has several limitations that are important for an accurate assessment of the validity of the presented results.

A first limitation is the relatively short time between administering the STPS and testing the effectiveness of the persuasive messages. Since only one week separated the two events, an alternative explanation for the results obtained in this study is a consistency effect. Participants tried to respond to the text messages consistent with the way they filled out the STPS. Although we believe this effect to be small since the STPS does not directly address snacking behavior but rather a general susceptibility to the different social influence strategies, and because it is unlikely for participants to have recognized the text-messages as implementations of the influence strategies addressed by the STPS, we cannot rule out this explanation. Such a consistency effect would likely wear off over longer periods of study or more long-term interventions. To test such a hypothesis and more thoroughly examine the dynamics of long term

behavioral change, researchers should conduct studies over longer periods of time than reported here.

Second, we believe that the question about the usefulness of the message during the second stage of the experiment—in retrospect—might have increased elaboration about the received messages. By consciously asking people to evaluate the message, and not just their own snacking behavior, the implementations of the influence strategies could have become salient. Based on the Elaboration Likelihood Model [Petty and Cacioppo 1986] we would expect the effects of the messages to be smaller in states of high salience. Thus, the results presented here might underestimate the actual effects of the messages when processed in a lower elaboration state.

Next, the number of dropouts in the second phase of the experiment was relatively large: more than half of the participants did not fill out the final diary. While the models presented in the analysis section rely on all data points, and thus also include participants who did not finish all diaries, the high number of dropouts might still have influenced the conclusions presented in this article. However, we feel the explicit analysis of the dropouts over conditions does not provide substantial evidence of an effect of the experimental condition on the dropouts. Thus, dropout can be considered at random and of limited impact on conclusions that are based on comparisons of the conditions.

Finally, the use of self-report measures presents a limitation of the current work. However, we focused explicitly on snacking as opposed to healthy eating in general to make the self-report task relatively easy to carry out. Also, we provided participants with a clear definition of snacking prior to filling out their diary to make sure a common understanding was created. This however does not rule out that participants—be it conscious or unconscious—erroneously reported their snack intake. Replication of the presented experiment using behavioral measures would be feasible.

5.5. Future Work

While we showed the practical implications of adaptation to measurements obtained by the STPS we feel that several aspects of persuasion profiles still need attention. First of all test-retest reliability of both the STPS as well as the persuasion profile itself (according to behavioral responses) should be assessed to estimate the need for dynamic vs. static profiles. Second, the promise of persuasion profiles to be effective in multiple contexts needs to be tested explicitly. Third, the current research should be conducted on different and more diverse samples of users. Our selection of participants who are motivated to eat healthy imposes a threat to the external validity of our findings for nonmotivated users. Finally, we feel that we should work towards profiles that can be disclosed to users without decreasing the beneficial effects of compliance. Hence, it's a design challenge to provide transparent persuasion profiles without defeating their positive purpose.

5.6. Final Remarks

In this article, we built and evaluated a text-messaging intervention to influence snacking behavior. We have shown that text messages that were personalized—tailored—on the level of influence strategy susceptibility were more effective in reducing snacking consumption than messages that were not tailored. This effect was not only evident at the behavioral level, but the personalized messages were also regarded as more useful by the receivers. Contrary to other attempts in this direction, we were able to model the effect of the personalization over multiple days and were able to show that personalization was effective beyond the first exposure. We have as

such presented a clear case in favor of personalized persuasion over generic persuasive applications.

This article contributes to the current state of the art of designing persuasive technologies not just by highlighting the importance of a personalized approach. The article also presents a validated measurement instrument that proved useful for the estimation of individual differences in susceptibility to different social influence strategies. This measurement instrument offers designers of persuasive applications the ability to measure individual differences at the level of influence strategy preference.

Next to measuring individual differences we have also shown how designers can create personalized interventions based on differences in influence strategy susceptibility. We have detailed how, for the context of snacking reduction, different implementations of the different social influence strategies were designed and evaluated. Our approach—in which the designers create a set of implementations that are later evaluated by experts—can be used by designers of persuasive interventions in other domains.

In our trial, personalized persuasion proved beneficial over nonpersonalized versions in reducing snacking behavior. By presenting the STPS—a scale for reliable assessment of influence strategy preference—and by presenting the design and evaluation of a personalized persuasive intervention, we have provided designers with a clear-cut example of implementing personalization at the level of social influence strategies in persuasive applications.

REFERENCES

AJZEN, I. AND FISHBEIN, M. 1980. *Understanding Attitudes and Predicting Social Behavior*. Prentice-Hall, Englewood Cliffs, NJ.

BRUG, J., OENEMA, A., KROEZE, W., AND RAAT, H. 2005. The Internet and nutrition education: Challenges and opportunities. *Euro. J. Clin. Nutrition 59*, 130–139.

CACIOPPO, J. T., PETTY, R. E., KAO, C. F., AND RODRIGUEZ, R. 1986. Central and peripheral routes to persuasion: An individual difference perspective. *J. Pers. Social Psychol. 51*, 5.

CIALDINI, R. 2001. *Influence, Science and Practice*. Allyn & Bacon, Boston.

CIALDINI, R. 2004. The science of persuasion. *Sci. Amer. Mind 284*, 76–84.

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. *J. Pers. Social Psychol. 69*, 318–328.

CONSOLVO, S., MARKLE, K., PATRICK, K., AND CHANASYK, K. 2009. Designing for persuasion: Mobile services for health behavior change. In *Proceedings of the 4th International Conference on Persuasive Technology (PERSUASIVE)*.

DEUTSCH, M. AND GERARD, H. B. 1955. A study of normative and informational social influences upon individual judgment. *J. Abnorm. Social Psychol. 51*, 629–636.

FOGG, B. J. 2003. *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann.

FRANKLIN, V. L., GREENE, A., WALLER, A., GREENE, S. A., AND PAGLIARI, C. 2008. Patients' engagement with "sweet talk": A text messaging support system for young people with diabetes. *J. Med. Internet Res. 10*, e20.

GERBER, B. S., STOLLEY, M. R., THOMPSON, A. L., SHARP, L. K., AND FITZGIBBON, M. L. 2009. Mobile phone text messaging to promote healthy behaviors and weight loss maintenance: A feasibility study. *Health Inform. J.* 15, 1, 17–25.

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. *Pers. Social Psychol. Bul. 27*, 7, 859–867.

HE, J., WHELTON, P. K., APPEL, L. J., CHARLESTON, J., AND KLAG, M. 2000. Long-term effects of weight loss and dietary sodium reduction on incidence of hypertension. *Hypertension. 35*, 544–549.

HURLING, R., CATT, M., DE BONI, M., FAIRLEY, B. W., HURST, T., MURRAY, P., RICHARDSON, A., AND SODHI, J. S. 2008. Using Internet and mobile phone technology to deliver an automated physical activity program: Randomized controlled trial. *J. Med. Internet Res. 9*.

JAMES, J. M. AND BOLSTEIN, R. 1992. Effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys. *Publ. Opinion Quarterly 54*, 442–453.

JOHNSON, S. S., PAIVA, A. L., CUMMINS, C. O., JOHNSON, J. L., DYMENT, S. J., WRIGHT, J. A., PROCHASKA, J. O., PROCHASKA, J. M., AND SHERMAN, K. 2008. Transtheoretical model-based multiple behavior intervention for weight management: Effectiveness on a population basis. *Prevent. Med. 46,* 3, 238–246.

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., Springer Berlin, 82–93.

KAPTEIN, M., MARKOPOULOS, P., DE RUYTER, B., AND AARTS, E. 2009. Can you be persuaded? Individual differences in susceptibility to persuasion. In *Proceedings of International Conference on Human-Computer Interaction (Interact)*.

KAPTEIN, M., LACROIX, J., AND SAINI, P. 2010. Individual differences in persuadability in the health promotion domain. In *Persuasive Technology*, T. Ploug, P. Hasle, and H. Oinas-Kukkonen Eds., Springer Berlin, 94–105.

KAPTEIN, M., DUPLINSKY, S., AND MARKOPOULOS, P. 2011a. Means based adaptive persuasive systems. In *Proceedings of the Annual Conference on Human Factors in Computing Systems (CHI)*. ACM, New York, NY, 335–344.

KAPTEIN, M., ECKLES, D., AND DAVIS, J. 2011b. Envisioning persuasion profiles: Challenges for public policy and ethical practice. *Interactions 18*, 66–69.

KASS, A. 2007. Transforming the mobile phone into a personal performance coach. In *Mobile Persuasion: 20 Perspectives on the Future of Behavior Change*, B. J. Fogg and D. Eckles Eds., Stanford Captology Media.

KROEZE, W., WERKMAN, A., AND BRUG, J. 2006. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. *Ann. Behav. Med. 31*, 205–223.

KROMHOUT, D., BLOEMBERG, B., SEIDELL, J. C., NISSINEN, A., AND MENOTTI, A. 2001. Physical activity and dietary fiber determine population body fat levels: The seven countries study. *Internat. J. Obesity Related Metabolic Disorders 25*, 301–306.

LACROIX, J., SAINI, P., AND HOLMES, R. 2008. The relationship between goal difficulty and performance in the context of a physical activity intervention program. In *Proceedings of Mobile Human-Computer Interaction Conference (MobileHCI)*.

LACROIX, J., SAINI, P., AND GORIS, A. 2009. Understanding user cognitions to guide the tailoring of persuasive technology-based physical activity interventions. In *Proceedings of the International Conference on Persuasive Technology (Persuasive)*.

LOCKTON, D., HARRISON, D., AND STANTON, N. A. 2008. Design with intent: Persuasive technology in a wider context. In *Proceedings of the International Conference on Persuasive Technology (PERSUASIVE)*.

MAHESHWARI, M., CHATTERJEE, S., AND DREW, D. 2008. Exploring the persuasiveness of "just-in-time" motivational messages for obesity management. In *Proceedings of the International Conference on Persuasive Technology (PERSUASIVE)*.

MCGRAA, K. L. 2010. The effects of persuasive motivational text messaging on adherence to diet and exercise programs across different personality traits. Ph.D. thesis, Fielding Graduate University.

MILGRAM, S. 1974. *Obedience to Authority*. Tavistock, London.

MUENNIG, P., LUBETKIN, E., JIA, H., AND FRANKS, P. 2000. Gender and the burden of disease attributable to obesity. *Amer. J. Pub. Health. 9*, 1662–1668.

MULAIK, S. 1972. *Foundations of Factor Analysis*. McGraw-Hill.

MUST, A., SPADANO, J., COAKLEY, E. H., FIELD, A., COLDITZ, G., AND DIETZ, W. 1999. The disease burden associated with overweight and obesity. *J. Amer. Med. Assn. 282*, 16, 1523–1529.

NEVILLE, L. M., O'HARA, B., AND MILAT, A. J. 2009. Computer-tailored dietary behaviour change interventions: A systematic review. *Health Ed. Res. 24*, 699–720.

OGDEN, C. L., CARROLL, M. D., CURTIN, L. R., MCDOWELL, M. A., TABAK, C. J., AND FLEGAL, K. M. 2006. Prevalence of overweight and obesity in the United States, 1999-2004. *J. Amer. Med. Assn. 295*, 13, 1549–1555.

OINAS-KUKKONEN, H. AND HARJUMAA, M. 2009. Persuasive systems design: Key issues, process model, and system features. *Comm. Assoc. Inform. Syst. 24*, 28.

PARKES, A., HENDERSON, M., AND WIGHT, D. 2005. Do sexual health services encourage teenagers to use condoms? A longitudinal study. *J. Family Planning Reprod. Health Care. 31*, 271–280.

PATRICK, K., RAAB, F., ADAMS, M. A., DILLON, L., ZABINSKI, M., ROCK, C. L., GRISWOLD, W. G., AND NORMAN, G. J. 2009. A text message based intervention for weight loss: Randomized controlled trial. *J. Med. Internet Res. 11*, 1–9.

PETTY, E. P. AND CACIOPPO, J. T. 1986. The elaboration likelihood model of persuasion. *Advances Exper. Social Psychol. 19*.

RÄISÄNEN, T., OINAS-KUKKONEN, H., AND PAHNILA, S. 2008. Finding kairos in quitting smoking: Smokers' perceptions of warning pictures. In *Proceedings of the International Conference on Persuasive Technology (PERSUASIVE)*.

RHOADS, K. 2007. How many influence, persuasion, compliance tactics & strategies are there? http://www.workingpsychology.com/numbertactics.html.

SNIJDERS, T. AND BOSKER, R. 1999. Multilevel analysis: *An Introduction To Basic And Advanced Multilevel Modeling*. Sage Publications Ltd.

WEST, S. G. 1975. Increasing the attractiveness of college cafeteria food: A reactance theory perspective. *J. Appl. Psychol. 60*, 656–658.

WORLD HEALTH ORGANIZATION. 2003. Diet, nutrition and the prevention of chronic diseases. Tech. rep., World Health Organization.

Received February 2011; revised June 2011, September 2011; accepted September 2011