Influence of Animation- Versus Text-Based Delivery of a Web-Based Computer-Tailored Smoking Cessation Intervention on User Perceptions

Jan Mathis Elling, Hein de Vries
Department of Health Promotion, Maastricht University/CAPHRI, the Netherlands

Abstract
Computer-tailored (CT) digital health interventions have shown to be effective in obtaining behaviour change. Yet, user perceptions of these interventions are often unsatisfactory. Traditional CT interventions rely mostly on text-based feedback messages. A way of presenting feedback messages in a more engaging manner may be the use of narrated animations instead of text. The goal of this study was to assess the effect of manipulating the mode of delivery (animation vs. text) in a smoking cessation intervention on user perceptions among smokers and non-smokers. Smokers and non-smokers (N = 181) were randomized into either the animation or text condition. Participants in the animation condition assessed the intervention as more effective ($\eta^2_p = .035$), more trustworthy ($\eta^2_p = .048$), more enjoyable ($\eta^2_p = .022$), more aesthetic ($\eta^2_p = .233$), and more engaging ($\eta^2_p = .043$) compared to participants in the text condition. Participants that received animations compared to text messages also reported to actively trust the intervention more ($\eta^2_p = .039$) and graded the intervention better ($\eta^2_p = .056$). These findings suggest that animation-based interventions are superior to text-based interventions with respect to user perceptions.

Keywords
Digital health, eHealth, computer tailoring, smoking cessation, user experience, user engagement
Digital health stands for the utilisation of digital and mobile technologies to support health and healthcare (World Health Organization, 2018). Digital health can facilitate the transmission of targeted health information to individuals. In computer-tailored (CT) interventions, this health information is matched to the individual behavioural and motivational characteristics through a computerised process (de Vries & Brug, 1999). In order to diagnose the necessary information for the tailoring process, the recipient usually fills out questionnaires which are then used to generate highly individualised information adapted to the needs and characteristics of the recipient. Information that is perceived as personally relevant enhances central processing according to the elaboration likelihood model of persuasion (Petty & Cacioppo, 1986). Information that is processed under the central route is in turn more likely to result in actual and sustained attitudinal changes and is thus more predictive of behaviour. Ruiter et al. (2006) demonstrated that CT health information compared to generic health information motivates people into more attentive information processing. In the area of health promotion, CT interventions have been shown to be effective and cost-effective in motivating people to adopt health promoting behaviour or change health detrimental habits (Krebs et al., 2010; also for Dutch smoking cessation interventions (Cheung et al., 2017).

**Smoking Cessation**

In this study, the digital health program under investigation was aimed at supporting people to quit smoking. The program was built on an earlier intervention that has proven effective in realising smoking cessation (Stanczyk, Bolman, et al., 2014; Stanczyk et al., 2016). Tobacco smoking is a major public health problem in the Netherlands. In 2015, 20,000 deaths were attributed to smoking-related diseases (Rijksinstituut voor Volksgezondheid en Milieu, 2018). In 2017, 23.1% of the Dutch adult population reported to smoke (Nationale Expertisecentrum Tabaksontmoediging, 2018). In the group of smokers, 41.7% tried to quit smoking at least once (Nationale Expertisecentrum Tabaksontmoediging, 2018). However, because of conditioned behaviour and the highly addictive nature of nicotine (Benowitz, 2010), about 95% of smokers who try to quit without treatment fail in their cessation attempts (Hughes et al., 2004). Hence, developing effective smoking cessation interventions is of high societal relevance.

**User Experience and Engagement**

A fundamental problem of digital health interventions, including CT interventions, is attrition, i.e., participants not using the intervention and/or being lost to follow-up (Eysenbach, 2005). In the randomised controlled trial (RCT) of the smoking cessation intervention on which the program presented in this paper is based, 42.9% (591/1,378) of the respondents in the experimental conditions were lost to follow-up after 12 months (Stanczyk et al., 2016). In more open settings than RCTs, even higher attrition rates can be expected (Eysenbach, 2005). For instance, an attrition rate of 78.2% (330/422) was found in a recent tailored digital health intervention for a healthy lifestyle change (van der Mispel et al., 2017).

High attrition may be caused by an unsatisfactory user experience (Crutzen et al., 2011; Crutzen et al., 2009). User experience comprises the user’s cognitive and affective perceptions of a website or a web-based intervention during and after exposure to that service (Crutzen et al., 2011). A
positive user experience is posited to lead to the intention to revisit a web-based intervention and
to the intention to recommend the intervention to others, which is described by the term e-loyalty
(Crutzen et al., 2011). The user experience model of Crutzen et al. (2011) is reported in Figure 1.
Efficiency (ease of searching information), effectiveness (usefulness of the information),
trustworthiness (information perceived as accurate and true), enjoyment (use of intervention elicits
positive feelings), and active trust (feeling able to act purposefully on the information given) are
all posited to have a positive influence on e-loyalty. The positive influence of effectiveness and
trustworthiness is, respectively, partially and fully mediated by active trust. The relations in the
user experience model have been tested in six web-based interventions, in which the importance
of the model’s constructs in predicting e-loyalty has been confirmed (Crutzen et al., 2014; Crutzen
et al., 2011; Crutzen et al., 2012; Nunn et al., 2017).

Another theory proposes that a reason for high attrition may be that users are not engaged with
the intervention (Perski et al., 2019; Perski et al., 2017; Short et al., 2018; Short et al., 2015). A
certain level of engagement is considered a precondition for interventions to be effective (Donkin
et al., 2011). Since engagement has been defined in various ways between disciplines, Perski et al.
(2017) proposed an integrative definition and conceptual framework of engagement with digital
health interventions, called the Digital Behaviour Change Interventions (DBCI) Engagement
Scale. They conceptualised engagement as a multidimensional construct consisting of behavioural
dimensions (e.g., amount of usage) which are underpinned by the user’s subjective experience
(i.e., cognitive and emotional aspects; Perski et al., 2017).

Figure 1. User Experience Model of Crutzen et al. (2011).
User perceptions and user engagement (jointly referred to as user perceptions) are both posited to be influenced by characteristics of the intervention, e.g., aesthetics, user control, mode of delivery (Crutzen et al., 2011; Perski et al., 2017). One way to improve user perceptions is thus to adapt the mode of delivery to the needs of the target population (Smit et al., 2015). It is hypothesised that adaptation of the delivery mode leads to changes in user perceptions, which ultimately result in better adherence.

A mode of delivery for communicating health messages is the use of narrated animation. Narrated animations make use of words (i.e., spoken text) and graphics (i.e., animations that move) and allow thus for more extensive information processing than text-based interventions that rely solely on written words. The cognitive theory of multimedia learning elucidates on the difference between animation and text for information processing (Mayer, 2018). The theory posits that people use separate channels to process visual and auditory information independently and that people can process only a limited amount of information in each channel at one time. An advantage of narrated animation compared to text is thus that recipients are able to process the information using the visual and auditory channel, whereas recipients of text-based messages are limited to the visual channel only. In this way, the cognitive load is balanced between the visual and auditory channel, so neither one is overloaded (Mayer, 2009). Research suggests that spoken animated
videos that are simple, shorter than 5 minutes, positive in tone, and without the use of medical terminology are appreciated by participants in digital health interventions (van het Schip et al., 2020; Vandelanotte & Mummery, 2011). A recent experimental study showed that spoken animations are the most effective way to communicate information on colorectal cancer screening to people with low health literacy, without impairing high health literate people (Meppelink et al., 2015).

Research Goal

The effects of using animations in CT interventions on user perceptions have not been explored yet. Thus, the goal of this study was to assess the effect of manipulating the mode of delivery (animation vs. text) on user perceptions among smokers and non-smokers. Statistically, interaction effects between condition and smoking status need to be tested first. However, we hypothesised that there would be no interaction effects between condition and smoking status, because there is no evidence and no reason to assume that smokers and non-smokers perceive animation and text differently. Therefore, we looked at main effects of condition and smoking status. The hypothesised pattern of the main effects of condition and smoking status on user perceptions is depicted in Figure 2. Regarding main effects of condition, it was hypothesised that user perceptions in the animation-based condition would be assessed significantly better than in the text-based condition (hypothesis 1). Regarding main effects of smoking status, it was hypothesised that user perceptions in the group of smokers would be assessed significantly better than in the group of non-smokers (hypothesis 2). This study was open to smokers and non-smokers in order to test if smokers and non-smokers differ in their evaluation of the two versions. We assumed that feedback messages in the CT intervention will be of higher personal relevance for smokers than for non-smokers, because non-smokers are usually not personally involved in the topic of smoking cessation. Past research has shown that perceived personal relevance significantly predicted higher appreciation of a CT digital health intervention (Kanera et al., 2016).

Method

Design

A between-subjects design with two experimental conditions was used. In one condition, participants received an animation-based version of a smoking cessation intervention. In the other condition, participants received a text-based version of the same intervention. Ethical approval was granted by the Ethical Review Committee Psychology and Neuroscience (ERCPN) at Maastricht University (Master_205_13_03_2019). The study is registered in the Netherlands Trial Register (NL7669, https://www.trialregister.nl/trial/7669).

Intervention

The web-based CT smoking cessation intervention that was used in this study was based on an existing intervention, which was subject to a RCT and has been found to be effective and cost-effective in the Netherlands (Stanczyk, Bolman, et al., 2014; Stanczyk et al., 2016; Stanczyk, Smit,
The I-Change model was used as a theoretical framework for the development of the intervention (de Vries, 2017; de Vries et al., 2008). The present study used a shortened version of the intervention delivered in an animation-based version and a text-based version. The animation-based version made use of narrated animations with little onscreen text, see Appendix A. The text-based version consisted of text-based feedback messages without any graphics, see Appendix A. The content (i.e., messages) was exactly the same in both versions. Whereas the text-based version was readily available from prior research (Stanczyk et al., 2016), the animation videos for the animation-based version had to be developed. All animations were developed using the web-based animated video creation tool Vyond (GoAnimate, Inc., San Mateo, California, U.S.). The intervention website was developed employing responsive web design, implying that the website could be accessed on all common devices including smartphones, tablets, desktop computers, and laptops.

**Recruitment**

A power analysis for analysis of variance (ANOVA) was conducted using G*Power version 3.1 (Faul et al., 2007). Taking into account a modest effect size \( f \) of .22, a power of .80, and an alpha of .05, a minimum total sample size of 165 participants was required. Since we knew from prior research that about 32% of participants have to be excluded after data collection, because they rush through the intervention without actively processing the information, we aimed at recruiting 243 participants (Stanczyk et al., 2013).

Inclusion criterion was that participants were at least 18 years old, as the sale of tobacco to persons under 18 years is illegal in the Netherlands (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2018). Various recruitment strategies were employed including posts on social media, survey exchange websites, posters, flyers, and a research participation credit system within the university. Posters were put up in public places and flyers were distributed door-to-door in the city of Maastricht. Psychology undergraduate students of Maastricht University were recruited through the research participation system of the Faculty of Psychology and Neuroscience. Students taking part in the research participation system received credits for their participation. All other participants could enter a raffle for 10 vouchers of €25 each.

**Participants**

During the study period from April 11, 2019 to December 6, 2019, data from 242 persons were collected. Participants who took less than 5 minutes to complete the study were excluded from data analysis, because in order to actively process the information of the intervention, a minimum amount of 5 minutes was deemed necessary. The original data file consisted of 125 participants in the animation condition and 117 in the text condition. In the animation condition, 37 participants (29.6%) did not take more than 5 minutes and were thus excluded. In the text condition, 24 participants (20.5%) were excluded because of the same reason. The final sample consisted of \( N = 181 \) participants. Seven student participants were recruited through the research participation system of Maastricht University.

People who were interested in taking part in the research were directed to an intervention website on which they could take part in the online study without registration. Participants were
informed that they could leave the study at any time for any reason if they wish to do so without any consequences. Informed consent was obtained online from all participants. Participants were neither informed about the existence of two experimental conditions nor about the randomization process, in order to mitigate the effect of demand characteristics on the results.

Procedure
After giving informed consent, a baseline questionnaire had to be filled in asking for gender, age, educational level, and smoking status. Next, participants were informed that they receive a part of an existing smoking cessation intervention which was originally developed for smokers who are motivated to quit smoking (Stanczyk et al., 2016). Therefore, smokers were asked to fill out the questionnaire as if they wanted to quit smoking. For this purpose, smokers were asked to look back at a time in which they wanted to quit smoking and to answer the questions from that perspective. Non-smokers were asked to immerse in the situation that they are smokers and want to quit smoking.

During the program, participants received tailored feedback on the pros and cons of quitting smoking (i.e., attitude), on preparatory action plans to effectively quit smoking, and on coping plans to deal with situations in which they think it is difficult not to smoke. All items of the tailoring process are reported in Appendix B. For the pros, cons, and coping plans, participants could choose three items each based on their personal preference for which they would like to receive feedback. In total, participants received 14 feedback messages, either as animations or text messages. The personal information for the tailored feedback was gathered by means of questionnaires in between the different sections of the intervention. The computer program TailorBuilder (OverNite Software Europe BV, Geleen, The Netherlands) employed if-then rules to match the personal answers with the relevant feedback messages from a file consisting of all possible feedback messages. After completing the program, participants were asked to fill out an evaluation questionnaire.

Measures
An overview of the complete questionnaire is reported in Appendix C. Demographics were measured by asking for gender (1 = male; 2 = female; 3 = third gender), age of the participant, and education level (1 = low (primary or basic vocational school); 2 = medium (secondary vocational school or high school); 3 = high (higher vocational school or university)). Smoking status was assessed by one item asking whether the participant smokes (1 = not smoking; 2 = smoking).

User experience (Crutzen et al., 2011) was measured by five constructs: effectiveness, trustworthiness, enjoyment, active trust, and design aesthetics. Effectiveness (e.g., “The program gives important information on smoking cessation”) was measured by three items (Cronbach’s alpha = .87) on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree). Trustworthiness (e.g., “The program is trustworthy”) was measured by three items (Cronbach’s alpha = .86) on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree). Enjoyment (e.g., “I found my visit to this program enjoyable”) was measured by three items (Cronbach’s alpha = .92) on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree). Active trust (e.g., “I know now how I can stop smoking”) was measured by three items (Cronbach’s alpha = .88) on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree).
agree). Design aesthetics (e.g., “I think the design of the program is attractive”) was measured by three items (Cronbach’s alpha = .90) on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree).

E-Loyalty (Crutzen et al., 2011) was assessed by two constructs. First, the intention to revisit the intervention (e.g., “It is likely that I will visit the website again in the future”) was assessed by two items (r = .69) on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree). Second, the intention to recommend the intervention to others (e.g., “It is likely that I will recommend this website to others”) was assessed by two items (r = .81) on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree).

The experiential dimensions of the Digital Behaviour Change Interventions (DBCI) Engagement Scale (Perski, 2017) (e.g., “How strongly did you experience interest?”) were measured by five items on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree). The amount of use was measured by one item (“How much time (in minutes) did you spend on the website?”). Participants had to indicate the time in minutes entering free text. The “depth of use”, which is originally part of the DBCI Engagement Scale, was not assessed in this study, as the “depth of use” in a CT intervention is not determined by the individual but by the tailoring process. In this CT intervention, all participants received the same number of sections and health messages. Z-score transformation was applied to all items of the DBCI Engagement Scale. Subsequently, a total sum score was calculated with equal weight given to each item (Cronbach’s alpha = .88).

A grade for the received version of the intervention was measured on a scale ranging from 1 (very bad) to 10 (very good). Time spent on website was measured automatically.

Statistical Analyses
Descriptive analyses and a two-way multivariate analysis of variance (MANOVA) were performed using IBM SPSS Statistics 27. Effectiveness, trustworthiness, enjoyment, active trust, design aesthetics, revisit, recommendation, engagement, and grade have been included as dependent variables. Condition and smoking status have been included as independent variables. Partial η² is reported as effect size with a 90% confidence interval (CI) around it. A 90% CI was chosen instead of a 95% CI, because a 95% CI around partial η² can include 0, even though the test reveals a statistical difference with p < .05 (Steiger, 2004). Partial η² cannot be smaller than zero (Steiger, 2004).

Results
Sample Characteristics
The mean age of the sample was 26.7 years (SD = 11.6; range = 18-82). More non-smokers (57.5%) than smokers participated in the study. The majority of participants was female (69.1%) and had a high level of education (95.6%). Sample characteristics are reported in Table 1.
Table 1. Sample Characteristics and Time Spent on Website

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Full sample</th>
<th>Condition</th>
<th>Smoking status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Total</td>
<td>181</td>
<td>100</td>
<td>88</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>55</td>
<td>30.4</td>
<td>30</td>
</tr>
<tr>
<td>Female</td>
<td>125</td>
<td>69.1</td>
<td>57</td>
</tr>
<tr>
<td>Third gender</td>
<td>1</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>2</td>
<td>1.1</td>
<td>2</td>
</tr>
<tr>
<td>Middle</td>
<td>6</td>
<td>3.3</td>
<td>3</td>
</tr>
<tr>
<td>High</td>
<td>173</td>
<td>95.6</td>
<td>83</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>77</td>
<td>42.5</td>
<td>33</td>
</tr>
<tr>
<td>Not smoking</td>
<td>104</td>
<td>57.5</td>
<td>55</td>
</tr>
<tr>
<td>Age M (SD)</td>
<td>26.7 (11.6)</td>
<td>27.8 (13.5)</td>
<td>25.7 (9.3)</td>
</tr>
<tr>
<td>Time spent on website M (SD)</td>
<td>00:12:20 (00:08:01)</td>
<td>00:15:32 (00:08:54)</td>
<td>00:09:19 (00:05:38)</td>
</tr>
</tbody>
</table>

Differences in User Perceptions between Conditions

Before testing main effects of condition and smoking status on user perceptions, condition × smoking status interaction effects need to be tested. No significant interaction effects between condition and smoking status were found (Pillai’s Trace = .04, $F = 0.84$, $p = .585$, $\eta^2_p = .043$). This implies that smokers and non-smokers were not affected differently by manipulation of the delivery mode. The absence of condition × smoking status interaction effects is reported in Appendix D. Since no interactions were found, main effects of condition and main effects of smoking status on user perceptions are reported in the following paragraphs.

Significant differences between the animation condition and text condition were found in the hypothesised direction (Pillai’s Trace = .29, $F = 7.74$, $p < .001$, $\eta^2_p = .292$). Participants who received the animation version in comparison to the text version reported higher ratings for most constructs. Main effects of condition on user perceptions are reported in Table 2. Participants in the animation condition judged the intervention as more effective, more trustworthy, more enjoyable, and more aesthetic. Participants who received animations in comparison to text messages also reported to actively trust the intervention more. Furthermore, participants receiving animations scored higher on engagement than those receiving text messages. The better evaluation
of the animation condition was confirmed by a higher grade for the animation condition in favour over the text condition. Participants in the animation condition also stayed longer on the website compared to participants in the text condition. The effect sizes of the statistically significant differences in user perceptions between conditions ranged from small to large (range $\eta_p^2 = .022–.233$).

**Differences in User Perceptions between Smoking Status**

Main effects of smoking status on the constructs of user perceptions are reported in Table 3 (Pillai’s Trace = .12, $F = 2.66, p = .007, \eta_p^2 = .124$). Smokers reported to enjoy the intervention more and found the intervention more engaging compared to non-smokers. The effect sizes of the statistically significant differences in user perceptions between smoking status were small (range $\eta_p^2 = .030–.050$).

**Discussion**

The aim of this study was to examine user perceptions toward an animation version and a text version of a CT smoking cessation intervention among smokers and non-smokers. Overall, results showed that participants who received animated videos evaluated the intervention more positively than participants who received text messages regardless of smoking status.

**Table 2.** Main Effects of Condition on Constructs of User Perceptions

<table>
<thead>
<tr>
<th>User perception constructs</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Animation</td>
<td>Text</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td>n = 88</td>
<td>n = 93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>5.47</td>
<td>0.93</td>
<td>5.07</td>
<td>1.15</td>
<td>6.50</td>
<td>.012$^*$</td>
<td>.035 [0.004, 0.090]</td>
<td></td>
</tr>
<tr>
<td>Trustworthiness</td>
<td>5.67</td>
<td>0.84</td>
<td>5.24</td>
<td>0.99</td>
<td>9.01</td>
<td>.003$^*$</td>
<td>.048 [0.010, 0.108]</td>
<td></td>
</tr>
<tr>
<td>Enjoyment</td>
<td>4.88</td>
<td>1.18</td>
<td>4.55</td>
<td>1.27</td>
<td>3.90</td>
<td>.050$^*$</td>
<td>.022 [0.000, 0.068]</td>
<td></td>
</tr>
<tr>
<td>Active trust</td>
<td>5.18</td>
<td>0.93</td>
<td>4.78</td>
<td>1.08</td>
<td>7.11</td>
<td>.008$^*$</td>
<td>.039 [0.006, 0.095]</td>
<td></td>
</tr>
<tr>
<td>Design aesthetics</td>
<td>5.05</td>
<td>1.18</td>
<td>3.76</td>
<td>1.24</td>
<td>53.70 &lt; .001$^{**}$</td>
<td>.233 [0.147, 0.316]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-loyalty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revisit</td>
<td>4.55</td>
<td>1.51</td>
<td>4.26</td>
<td>1.53</td>
<td>1.72</td>
<td>.192</td>
<td>.010 [0.000, 0.047]</td>
<td></td>
</tr>
<tr>
<td>Recommendation</td>
<td>4.94</td>
<td>1.40</td>
<td>4.68</td>
<td>1.32</td>
<td>1.18</td>
<td>.280</td>
<td>.007 [0.000, 0.040]</td>
<td></td>
</tr>
<tr>
<td>Engagement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBCI engagement Scale</td>
<td>0.92</td>
<td>4.82</td>
<td>-0.74</td>
<td>4.46</td>
<td>7.93</td>
<td>.005$^*$</td>
<td>.043 [0.007, 0.101]</td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td>7.25</td>
<td>1.12</td>
<td>6.67</td>
<td>1.34</td>
<td>10.42</td>
<td>.001$^*$</td>
<td>.056 [0.013, 0.118]</td>
<td></td>
</tr>
</tbody>
</table>

*Note. The construct engagement is the sum of six z-scored items. The construct grade was measured on a scale from 1 (very bad) to 10 (very good). All other constructs were measured on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree). $^*$p < .05, $^{**}$p < .001.
Table 3. Main Effects of Smoking Status on Constructs of User Perceptions

<table>
<thead>
<tr>
<th>User perception constructs</th>
<th>Smoking status</th>
<th></th>
<th></th>
<th>F</th>
<th>p</th>
<th>ηp² [90% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive n = 77</td>
<td>Negative n = 104</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>User experience</td>
<td>Effectiveness</td>
<td>5.26</td>
<td>1.07</td>
<td>5.27</td>
<td>1.07</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Trustworthiness</td>
<td>5.47</td>
<td>0.87</td>
<td>5.43</td>
<td>1.00</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>Enjoyment</td>
<td>4.94</td>
<td>1.06</td>
<td>4.54</td>
<td>1.33</td>
<td>5.43</td>
</tr>
<tr>
<td></td>
<td>Active trust</td>
<td>5.06</td>
<td>0.89</td>
<td>4.91</td>
<td>1.12</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>Design aesthetics</td>
<td>4.43</td>
<td>1.38</td>
<td>4.35</td>
<td>1.37</td>
<td>1.52</td>
</tr>
<tr>
<td>E-loyalty</td>
<td>Revisit</td>
<td>4.41</td>
<td>1.52</td>
<td>4.39</td>
<td>1.53</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Recommendation</td>
<td>4.68</td>
<td>1.33</td>
<td>4.89</td>
<td>1.38</td>
<td>0.90</td>
</tr>
<tr>
<td>Engagement</td>
<td>DBCI engagement</td>
<td>1.15</td>
<td>3.89</td>
<td>-0.73</td>
<td>5.09</td>
<td>9.26</td>
</tr>
<tr>
<td></td>
<td>Grade</td>
<td>7.09</td>
<td>1.14</td>
<td>6.85</td>
<td>1.35</td>
<td>2.62</td>
</tr>
</tbody>
</table>

Note. The construct engagement is the sum of six z-scored items. The construct grade was measured on a scale from 1 (very bad) to 10 (very good). All other constructs were measured on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree). *p < .05, **p < .001.

Our first hypothesis was that user perceptions in the animation-based condition would be assessed significantly better than in the text-based condition. Regarding the seven constructs of the framework of Crutzen et al. (2011), the results confirmed the hypothesis for the constructs of user experience but not for e-loyalty (i.e., the intention to revisit and recommend the intervention). In detail, the animation version was assessed as more effective, more aesthetic, more enjoyable, more trustworthy, and people also felt more able to act on the information given. Yet, no differences between conditions were found for the intention to revisit or the intention to recommend the intervention. The results also confirmed our hypothesis regarding the concept engagement of the framework of Perski et al. (2019). Participants in the animation-based condition indeed perceived engagement higher than participants in the text-based condition. Furthermore, we asked participants to give the intervention an overall grade. The animation version was, as hypothesised, rated higher than the text version.

Our second hypothesis was that user perceptions in the group of smokers would be assessed significantly better than in the group of non-smokers, because personal relevance of the intervention topic was deemed higher for smokers. This hypothesis was confirmed only for two out of nine investigated constructs. Statistically significant differences between smokers and non-smokers were only found for the constructs enjoyment and engagement, indicating that smokers found both versions more enjoyable and engaging than non-smokers. No significant differences between smokers and non-smokers were found for the constructs effectiveness, trustworthiness, active trust, design aesthetics, revisit, recommend, and grade. It is important to note that the two constructs that yielded statistical significance are affective constructs (i.e., enjoyment and...
engagement), whereas no differences were found for cognitive (e.g., effectiveness) and cognitive-affective constructs (e.g., trustworthiness). This suggests that smokers and non-smokers evaluate the intervention similarly on a rational level, but are differently affected by the intervention on an emotional level (Crutzen et al., 2011).

Research on the use of animations in CT interventions is scarce, but multiple studies have examined the utilisation of videos in a news-driven format in which, for example, professional presenters or actors read aloud the feedback messages of the intervention. The full intervention, on which the intervention presented in this paper is based, was tested in an RCT (Stanczyk et al., 2011). The original intervention compared a news-driven video version to a text version. The results of the RCT showed that the video version was more effective than the text version in obtaining smoking abstinence after 12 months (Stanczyk et al., 2016). The video condition was also slightly better appreciated, but the difference did not reach statistical significance (Stanczyk, Bolman, et al., 2014). However, another previous study on the same intervention found that manipulation of the delivery mode (video vs. text) had no influence on the processing of the information and on the intention to revisit and recommend the intervention (Stanczyk et al., 2013). Support for the higher evaluation of videos compared to text was in turn found in two experiments comparing a video and text version of a CT intervention for obesity prevention and physical activity (Soetens et al., 2014; Walthouwer et al., 2015). In sum, the results of this study contribute to a body of evidence that suggests that interventions that make use of videos with spoken text are more effective in obtaining smoking abstinence and better appreciated than interventions that rely solely on written feedback messages.

The results of this study have to be interpreted in the context of the fast-changing web. In the recent years, information on the internet is more and more presented in a rich and engaging manner containing videos and interactive features. In particular, there has been an exponential growth of videos on the internet and this trend is expected to keep up in the coming years (Kalogeropoulos et al., 2016). In the year 2017, video traffic accounted for 75% of all web traffic and it is estimated that this figure will increase to 82% by 2022 (Cisco, 2019). A driver for this growth in traffic is the omnipresence of video content on the internet, e.g., in news, ads, and social media. This growth of videos online suggests that internet users have become accustomed to receiving information on the internet in the form of videos. Thus, the expectation of internet users regarding the delivery mode of information may have changed. Whereas in the past, most users expected to receive information in the form of text, nowadays, more and more users may expect to be presented with videos when searching for information on the internet (Kalogeropoulos et al., 2016). This may be one explanation for the remarkable large effect size for the difference in the construct design aesthetics between participants in the animation condition and the text condition.

The results for the construct design aesthetics indicated that participants disapproved of the text version that was solely based on text without any visual elements, probably because solely text-based websites do not meet modern standards. Furthermore, online reading and online information seeking is characterised by mostly browsing and scanning behavior and non-linear reading (Liu, 2005). Contemporary websites with hyperlinks and visuals encourage the user to switch between sections and to just allocate attention to information that seems interesting in that particular
Influence of Delivery Mode on User Perceptions

Elling & De Vries

European Journal of Health Communication 2021, Vol. 2(3) 1-23 CC BY 4.0

moment (Carr, 2011; Ryota & Kep Kee, 2016). The text-based version, however, was built for linear reading. The results of this study indicate that people favor more engaging websites.

Nonetheless, these results must be interpreted with caution and a number of limitations should be borne in mind. The major limitation of this study is the homogenous sample, which does not allow to generalise the findings to the general public. Most of the participants were in their twenties, female, and highly educated. Prior research does not suggest an influence of age or gender on the effectiveness of the smoking cessation intervention on which the intervention under investigation is based (Stanczyk, Bolman, et al., 2014). Yet, there is prior research showing that people with low and high health literacy have different information recall abilities depending on the mode of delivery in which the health information is presented. Meppelink et al. (2015) conducted a two (text format: written vs. spoken) by two (delivery mode: illustration vs. animation) experiment in which participants were randomly allocated to one of the four experimental conditions. The two modes of delivery were medical illustration (i.e., static graphics) and medical animation (i.e., moving graphics), both presenting information on colorectal cancer screening. Next to comparing two modes of delivery, the researchers added a second independent variable “text format”, in the sense that the animations and illustrations were either spoken or written.

Participants’ health literacy was also tested in order to assess whether health information recall and attitudes toward the health messages differed between people with different health literacy. Meppelink et al. (2015) found a three-way interaction effect between text format, delivery mode, and health literacy level on information recall. The results indicated that participants with a low health literacy level recalled significantly more information when confronted with spoken animations compared to spoken illustrations. Since health literacy is associated with education level (van der Heide et al., 2013), one may also expect interaction effects between condition and education level in studies on CT interventions delivered as video or text. However, in two RCTs of a CT smoking cessation intervention and a CT obesity prevention intervention, no differential effects per educational level were found (Stanczyk, Bolman, et al., 2014; Stanczyk et al., 2016; Walthouwer et al., 2015). In sum, there is evidence to suggest that the efficacy of delivery modes may depend on health literacy; however, no moderating effects of education level were found in two CT interventions. More research in this field is needed to examine the role of health literacy and education level in CT interventions.

Another limitation of this study is that smokers were recruited in the light of testing an intervention. Thus, smokers did not necessarily head for the intervention website because they intended to quit smoking. Consequently, revisiting and recommending the intervention for smoking cessation purposes was not necessarily at stake, which may explain why there were no differences in e-loyalty between the two conditions. Moreover, research is needed in which the groups differ to a greater extent, for example one could compare smokers that are willing to quit within one month to a group of never-smokers. In larger samples of smokers, the intention to quit smoking needs to be examined as a moderating variable, because smokers in different stages of change may respond differently to evaluation questions (Prochaska & Velicer, 1997). Pre-contemplating smokers, for instance, tend to be defensive and may thus evaluate the intervention
Influence of Delivery Mode on User Perceptions

Elling & De Vries

European Journal of Health Communication 2021, Vol. 2(3) 1-23 CC BY 4.0

more negatively than smokers in the contemplation or preparation phase (Prochaska & DiClemente, 1983).

In closing, it is noteworthy to mention that information that is perceived as useful, trustworthy, interesting, and attractive is more likely to be processed under the central route of information processing and is in turn more predictive of behaviour, according to the elaboration likelihood model (Petty & Cacioppo, 1986). Consequently, the higher evaluation by participants in the animation condition suggests that more processing of information and behaviour change may occur in the animated version than in the text version. Unfortunately, we are not aware of any research assessing the association between user perceptions and actual use of the intervention or actual behaviour change in the area of smoking cessation. In the context of an alcohol reduction intervention, self-reported engagement did not significantly predict the number of subsequent logins to the intervention program (Perski et al., 2019). Yet, only 3.7% of eligible users answered the questionnaire in this study (Perski et al., 2019). Hence, more research is needed to assess the association between user perceptions and program use as well as behaviour change.

Conclusions

This study has shown that the animation version of the intervention was evaluated better than the text version, regardless of smoking status. In general, therefore, it seems advisable to not only rely on traditional text-based interventions in the future, but to explore the use of feedback messages that are presented as animations. Further experimental research needs to be conducted to establish that animations are superior in obtaining better adherence and actual behaviour change compared to text-based interventions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References


Influence of Delivery Mode on User Perceptions

Elling & De Vries

European Journal of Health Communication
2021, Vol. 2(3) 1-23 CC BY 4.0


Crutzen, R., Cyr, D., & de Vries, N. K. (2012). The role of user control in adherence to and knowledge gained from a website: Randomized comparison between a tunneled version and a freedom-of-choice version. *Journal of Medical Internet Research, 14*(2), Article: e45. https://doi.org/10.2196/jmir.1922


https://doi.org/10.1108/00220410510632040


https://doi.org/10.1002/acp.3482

https://doi.org/10.2196/jmir.3979

https://wetten.overheid.nl/BWBR0004302/2018-11-17

https://assets.trimbos.nl/docs/cdade14e-20dc-462a-9f63-1b71ee995a84.pdf

https://doi.org/10.2196/publichealth.5393

https://osf.io/cj9y7

https://doi.org/10.1093/tbm/ibz039

https://doi.org/10.1007/s13142-016-0453-1

https://doi.org/10.1007/978-1-4612-4964-1_1

https://doi.org/10.1037/0022-006X.51.3.390

https://doi.org/10.4278/0890-1171-12.1.38

https://www.volksgezondheidenzorg.org/info/onderwerp/roken/cijfers-context/gevolgen#node-sterfte-door-roken

https://doi.org/10.1037/0278-6133.25.3.300
Ryota, K., & Kep Kee, L. (2016). How has the internet reshaped human cognition? 


**Author Contributions**

Conceptualisation (main idea, theory): Jan Mathis Elling & Hein de Vries
Funding acquisition: Not applicable
Project administration: Jan Mathis Elling & Hein de Vries
Methodology (design, operationalisation): Jan Mathis Elling & Hein de Vries
Data collection: Jan Mathis Elling
Data analysis: Jan Mathis Elling
Writing – original draft: Jan Mathis Elling
Writing – review & editing: Jan Mathis Elling & Hein de Vries

**Author Biographies**

Jan Mathis Elling is a PhD candidate at the Department of Health Promotion at Maastricht University. His interest lies in behaviour change and digital health. In particular, the effectiveness of behaviour change interventions and factors associated with intervention success.

Hein de Vries is Professor in Health Communication at Maastricht University. He has a special interest in developing health behavior models and intervention development and evaluation.
Appendix A

Screenshots of the Animation-Based Version (left) and Text-Based Version (right), displayed in Apple Safari on an iPhone 6s.
## Appendix B

Items for the Tailoring Process.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Items</th>
</tr>
</thead>
</table>
| **Attitude (pros)**| 1. Improvement of physical fitness  
2. Being an example to others  
3. Improvement of health  
4. Feeling more attractive  
5. Being proud  
6. Health of others  
7. Saving money  
8. Causing less nuisance to others |
| **Attitude (cons)**| 1. Difficulties relaxing  
2. Gaining weight  
3. Being bored more often  
4. Feeling gloomy  
5. Feeling insecure  
6. Feeling stressed  
7. Being less sociable  
8. Getting withdrawal symptoms |
| **Preparatory plans** | I am planning...  
1. ... to stop completely without cutting down on cigarettes first.  
2. ... to dispose all smoking related things from my house.  
3. ... to ask my guests to not smoke in my presence.  
4. ... to tell others that I will stop smoking.  
5. ... to make use of nicotine replacement therapy. |
| **Self-efficacy & coping plans** | I find it difficult not to smoke... & I have made plans to make sure that I will not smoke...  
1. ... if I am stressed.  
2. ... if I am mad.  
3. ... if I am sad.  
4. ... if somebody offers me a cigarette.  
5. ... if I see somebody enjoying a cigarette.  
6. ... if I am at a party.  
7. ... if I am drinking tea or coffee.  
8. ... after I have eaten.  
9. ... if I am having a break.  
10. ... if I get up in the morning.  
11. ... if I feel like needing a cigarette. |
### Appendix C

Overview of the Questionnaire.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Question</th>
<th>Answer options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>What is your gender?</td>
<td>[male] [female] [third gender]</td>
</tr>
<tr>
<td>Age</td>
<td>What is your age in years?</td>
<td>&lt;age&gt;</td>
</tr>
<tr>
<td>Education level</td>
<td>What is your highest completed education?</td>
<td>[Primary education, vmbo, havo-onderbouw, wvo-onderbouw, mbo1] [Havo, vwo, mbo] [Hbo-, wo-bachelor/master]</td>
</tr>
<tr>
<td>Smoking status</td>
<td>Do you smoke tobacco cigarettes?</td>
<td>[yes] [no]</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>The program...</td>
<td>7-point Likert scale ranging from 1 <em>(I totally disagree)</em> to 7 <em>(I totally agree)</em></td>
</tr>
<tr>
<td></td>
<td>1. ... provides important information on smoking cessation.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. ... helps me how to prepare to quit smoking.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. ... helps me to deal with difficult moments when quitting smoking.</td>
<td></td>
</tr>
<tr>
<td>Trustworthiness</td>
<td>The program...</td>
<td>7-point Likert scale ranging from 1 <em>(I totally disagree)</em> to 7 <em>(I totally agree)</em></td>
</tr>
<tr>
<td></td>
<td>1. ... provides trustworthy information.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. ... offers good tips</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. ... is trustworthy.</td>
<td></td>
</tr>
<tr>
<td>Enjoyment</td>
<td>I found my visit to this program...</td>
<td>7-point Likert scale ranging from 1 <em>(I totally disagree)</em> to 7 <em>(I totally agree)</em></td>
</tr>
<tr>
<td></td>
<td>1. ... fun.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. ... enjoyable.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. ... interesting.</td>
<td></td>
</tr>
<tr>
<td>Active trust</td>
<td>To what extent do you agree with the following statements?</td>
<td>7-point Likert scale ranging from 1 <em>(I totally disagree)</em> to 7 <em>(I totally agree)</em></td>
</tr>
<tr>
<td></td>
<td>1. I know now how I can stop smoking.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. The advices are good to use.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. I now know how to prepare my quit attempt.</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Description</td>
<td>Scale</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Design aesthetics</td>
<td>I think the design of the program is... 1. ... attractive. 2. ... beautiful. 3. ... interesting.</td>
<td>7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree)</td>
</tr>
<tr>
<td>Intention to revisit</td>
<td>To what extent do you agree with the following statements? Suppose the program were available. Then it is likely that I... 1. ... will visit the website again in the future. 2. ... will use this website again for information about smoking cessation.</td>
<td>7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree)</td>
</tr>
<tr>
<td>Intention to recommend</td>
<td>To what extent do you agree with the following statements? Suppose the program were available. Then it is likely that I... 1. ... will recommend this website to others. 2. ... will recommend this website to others for information on quitting smoking.</td>
<td>7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree)</td>
</tr>
<tr>
<td>Experiential dimensions (DBCI engagement scale)</td>
<td>How strongly did you experience the following? 1. Interest 2. Curiosity 3. Focus 4. Enjoyment 5. Pleasure</td>
<td>7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree)</td>
</tr>
<tr>
<td>Amount of use (DBCI engagement scale)</td>
<td>How much time (in minutes) did you spend on the website?</td>
<td>&lt;time in minutes&gt;</td>
</tr>
<tr>
<td>Grade</td>
<td>What is the overall score you would award the program?</td>
<td>Scale ranging from 1 (very bad) to 10 (very good)</td>
</tr>
</tbody>
</table>
## Appendix D

Interaction Effects of Condition and Smoking Status on Constructs of User Perceptions.

<table>
<thead>
<tr>
<th>User perception constructs</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Animation</td>
<td>Text</td>
<td>Positive</td>
<td>Negative</td>
<td>F</td>
<td>p</td>
<td>ηp²</td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>5.47</td>
<td>0.93</td>
<td>5.07</td>
<td>1.15</td>
<td>5.26</td>
<td>1.07</td>
<td>5.27</td>
<td>1.07</td>
</tr>
<tr>
<td>Trustworthiness</td>
<td>5.67</td>
<td>0.84</td>
<td>5.24</td>
<td>0.99</td>
<td>5.47</td>
<td>0.87</td>
<td>5.43</td>
<td>1.00</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>4.88</td>
<td>1.18</td>
<td>4.55</td>
<td>1.27</td>
<td>4.94</td>
<td>1.06</td>
<td>4.54</td>
<td>1.33</td>
</tr>
<tr>
<td>Active Trust</td>
<td>5.18</td>
<td>0.93</td>
<td>4.78</td>
<td>1.08</td>
<td>5.06</td>
<td>0.89</td>
<td>4.91</td>
<td>1.12</td>
</tr>
<tr>
<td>Design Aesthetics</td>
<td>5.05</td>
<td>1.18</td>
<td>3.76</td>
<td>1.24</td>
<td>4.43</td>
<td>1.38</td>
<td>4.35</td>
<td>1.37</td>
</tr>
<tr>
<td>Revisit</td>
<td>4.55</td>
<td>1.51</td>
<td>4.26</td>
<td>1.53</td>
<td>4.41</td>
<td>1.52</td>
<td>4.39</td>
<td>1.53</td>
</tr>
<tr>
<td>Recommendation</td>
<td>4.94</td>
<td>1.40</td>
<td>4.68</td>
<td>1.32</td>
<td>4.68</td>
<td>1.33</td>
<td>4.89</td>
<td>1.38</td>
</tr>
<tr>
<td>Engagement</td>
<td>0.92</td>
<td>4.82</td>
<td>-0.74</td>
<td>4.46</td>
<td>1.15</td>
<td>3.89</td>
<td>-0.73</td>
<td>5.09</td>
</tr>
<tr>
<td>Grade</td>
<td>7.25</td>
<td>1.12</td>
<td>6.67</td>
<td>1.34</td>
<td>7.09</td>
<td>1.14</td>
<td>6.85</td>
<td>1.35</td>
</tr>
</tbody>
</table>

*Note.* The construct engagement is the sum of six z-scored items. The construct grade was measured on a scale from 1 (very bad) to 10 (very good). All other constructs were measured on a 7-point Likert scale ranging from 1 (I totally disagree) to 7 (I totally agree).