





Article

Applying Collective Intelligence in Health Recommender Systems for Smoking Cessation: A Comparison Trial

Santiago Hors-Fraile ^{1,2,3}, Math J. J. M. Candel ⁴ , Francine Schneider ³ , Shwetambara Malwade ⁵,
Francisco J. Nunez-Benjumea ¹, Shabbir Syed-Abdul ^{5,6,7,*} , Luis Fernandez-Luque ¹ and Hein de Vries ³ 

- ¹ Research, Development and Innovation, Adhera Health, 08025 Barcelona, Spain; shors@adherahealth.com (S.H.-F.); knunez@adherahealth.com (F.J.N.-B.); luis@adherahealth.com (L.F.-L.)
- ² ATC Department, ETSII, Universidad de Sevilla, Avenida Reina Mercedes SN, 41012 Sevilla, Spain
- ³ Department of Health Promotion, Care and Public Health Research Institute (CAPHRI), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; francine.schneider@maastrichtuniversity.nl (F.S.); hein.devries@maastrichtuniversity.nl (H.d.V.)
- ⁴ Department of Methodology and Statistics, Care and Public Health Research Institute (CAPHRI), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; math.candel@maastrichtuniversity.nl
- ⁵ International Center for Health Information Technology (ICHIT), College of Medical Science and Technology, Taipei Medical University, 172-1 Keelung Rd., Section 2, Da'an District, Taipei 106, Taiwan; sv14kekade@tmu.edu.tw
- ⁶ Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, 172-1 Keelung Rd., Section 2, Da'an District, Taipei 106, Taiwan
- ⁷ School of Gerontology Health Management, College of Nursing, Taipei Medical University, 172-1 Keelung Rd., Section 2, Da'an District, Taipei 106, Taiwan
- * Correspondence: drshabbir@tmu.edu.tw; Tel.: +886-2-6638-2736 (ext. 1514)



Citation: Hors-Fraile, S.; Candel, M.J.J.M.; Schneider, F.; Malwade, S.; Nunez-Benjumea, F.J.; Syed-Abdul, S.; Fernandez-Luque, L.; de Vries, H. Applying Collective Intelligence in Health Recommender Systems for Smoking Cessation: A Comparison Trial. *Electronics* **2022**, *11*, 1219. <https://doi.org/10.3390/electronics11081219>

Academic Editor: Prasan Kumar Sahoo

Received: 22 February 2022

Accepted: 6 April 2022

Published: 12 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Background: Health recommender systems (HRSs) are intelligent systems that can be used to tailor digital health interventions. We compared two HRSs to assess their impact providing smoking cessation support messages. Methods: Smokers who downloaded a mobile app to support smoking abstinence were randomly assigned to two interventions. They received personalized, ratable motivational messages on the app. The first intervention had a knowledge-based HRS ($n = 181$): it selected random messages from a subset matching the users' demographics and smoking habits. The second intervention had a hybrid HRS using collective intelligence ($n = 190$): it selected messages applying the knowledge-based filter first, and then chose the ones with higher ratings provided by other similar users in the system. Both interventions were compared on: (a) message appreciation, (b) engagement with the system, and (c) one's own self-reported smoking cessation status, as indicated by the last seven-day point prevalence report in different time intervals during a period of six months. Results: Both interventions had similar message appreciation, number of rated messages, and abstinence results. The knowledge-based HRS achieved a significantly higher number of active days, number of abstinence reports, and better abstinence results. The hybrid algorithm led to more quitting attempts in participants who completed their user profiles.

Keywords: health recommender systems; smoking cessation; demographic filtering; message appreciation; engagement; behavior change

1. Introduction

For decades, computers have been used to generate health recommendations [1,2]. This usage of computers to adjust health materials to each person, in order to make them relevant and credible for their situation, replicating what an actual human counselor would do, is called computer tailoring [3,4]. Computer tailoring involves the generation of participant-specific recommendations, typically in the form of messages, by computers, which is done after an assessment of each person to match their characteristics, needs, and

interests [5–8]. Traditionally, these systems were designed with if-then-else rules, using users' input to some questionnaires to determine the correct feedback from the system, in a fashion similar to following a decision tree. They were designed to be used in single or multiple sessions, in which participants had to sit in front of a computer for a long period of time to conduct the intervention, answering many questionnaires and reflecting on the elaborate generated content derived from their answers. One of the areas where computer-tailoring was effective is in supporting health behavioral change [9,10]. Recurring areas that apply these computer-tailored behavioral changes are dietary improvements, physical activity promotion, smoking cessation, and mammography screening, as reviewed by Krebs and Neuhauser [11]. Moreover, a review by Cheung et al. (2017) [12] showed that computer-tailored interventions specifically for smoking cessation were the most successful of all eHealth interventions. The impact was higher when there were multiple tailored recommendations [1,13]. However, public health impacts were limited [14,15] when there was low usage [16], as users need to be engaged with the intervention to acquire the desired behavior change.

Health recommender systems (HRSs) are a newfangled way of providing personalized health information using artificial intelligence (AI) [17], which go beyond the traditional if-then programming paradigm used in computer-tailored systems [18]. HRSs can be conceived of as adaptive persuasive technologies, capable of learning the most relevant strategy to provide behavior change support [19]. In most common scenarios, they predict the relevance that each user would assign to a given potentially recommendable message, learning, and adapting over time from feedback. HRSs which use a collaborative filtering approach can take advantage of the 'collective intelligence' of all users' profiles and their feedback, as opposed to traditional computer-tailoring. This can potentially allow the user to avoid having to answer lengthy questionnaires. Instead, users can provide a reduced set of details on their profiles, hence reducing entry barriers to receiving personalized support. Consequently, HRSs may be used to offer higher personalization levels at low-effort cost for users across interventions, potentially leading to increased improvements in health outcomes [20,21] and higher user engagement rates [22–24].

However, Schäfer et al. (2017) previously explained that advanced HRSs need to achieve high levels of personalization, which classic recommender systems algorithms (like collaborative filtering) might not have, as they do not take into account key user information, such as patient gender, age, comorbidities, context, or ethnicity [25]. Also, other elements, such as being aware of the user context and grounding the recommendations on a behavioral model, are desired [26] and cannot be covered by the collaborative filtering algorithm approach on its own. User profile information is obtained by requesting users to fill in their profile (sometimes with lengthy questionnaires), thereby reverting to the necessary step in traditional computer tailoring. Despite completion being less time consuming than in a traditional-tailoring approach, and not mandatory, even a task of an additional few minutes may be undesirable and lead to lower engagement levels, or even increase attrition risks [27], which are common problems with HRSs [28].

As advanced health recommender systems may offer the necessary level of personalization in healthcare, we wanted to compare two HRSs in the context of smoking cessation support. The HRSs would select the most relevant motivational messages to be sent to smokers who are making an attempt to cease smoking through a mobile app, aiming to increase their abstinence rates. The first HRS used knowledge-based algorithms (KBA). The second one used hybrid algorithms (HA), which performed a knowledge-based step just like the KBA, but this output was then passed to a second step which performed collaborative filtering; more specifically, demographic filtering. Demographic filtering is a variant of the collaborative filtering algorithm and uses users' profiles to compute similarities. This demographic filtering step had embedded, in its design, the principles of the Integrated Model for explaining motivational and behavioral change [26], also known as I-Change. This model was chosen because it was useful for supporting smoking cessation in previous interventions [29]. The I-Change model assumes that people transition

through three phases (awareness, motivation, and action) to develop a behavior, which are conditioned by information and predisposing factors. Each of these phases has relevant determinants of change (e.g., attitudes and skills), which are the elements we included as part of the algorithm design.

After reading each message, all participants were asked to rate them with a score ranging from one to five stars for message relevance and usefulness. This information was registered in the database, keeping track of the relationship among users, messages and the given ratings. Although the ratings were irrelevant to the KBA algorithm, they were useful for improving collective intelligence of the demographic filtering algorithm used in other participant groups. Also, this provided users of both groups with a feeling of control over the system, and gave the impression that they could provide feedback, thereby making their experiences more homogeneous and comparable.

The tested HRS approaches were chosen to assess the impact of collaborative intelligence. The KBA approach was the most similar approach to traditional computer tailoring studies, as the knowledge-based algorithm works with constraints to match user needs. The HA adds the collaborative component in the selection process, which was not used in traditional computer tailoring. However, we did not use a pure collaborative filtering algorithm and used the hybrid algorithm instead, with a first knowledge-based step, because we wanted to ensure that users would never receive incorrect messages for their profiles, and, at most, would receive messages with low relevance for them. This feature was provided by the first filtering done by the knowledge-based step. We expected that the HA would be able to generate more relevant and personalized recommendations than its simpler knowledge-based-only version, leading to less dropouts, higher message appreciation, engagement, and effective behavioral change. In addition, we expected that those participants who fully completed their profiles in the HA group would have better results than those who did not, because the collaborative filtering algorithm could use that information to provide more relevant recommendations. This evaluation goes beyond the traditional way of assessing recommender systems, based on how well a given recommendation matches previous interactions between users and items [30–32]. However, some authors have already proposed that HRSs need to include some kind of utility function that takes into account both user satisfaction (message appreciation and engagement) and health impacts (smoking cessation behavioral changes) [25,33], as we do in this study.

At the same time, these higher levels of personalization and relevance in messages might not be needed; for instance, for smokers already highly motivated to quit. Hence, information on specific subgroup effects is also needed. Currently, it is not clear how to best develop more advanced systems, and whether such integration will yield better outcomes, in terms of users' message appreciation, engagement levels, and health outcomes, compared to simpler HRS algorithms, which address a reduced number of user profile elements. This consideration is in line with two recommended future challenges for HRSs identified by Schäfer et al. (2017) [25]: (1) selecting dataset sources and ensuring their quality, and (2) exploring different models to personalize intervention contents, based on users' health contexts, histories, and goals.

The first goal of the present study was to compare two smoking cessation support HRSs: one only knowledge-based (KBA), and the other a hybrid approach, using knowledge-based filtering with collective intelligence using demographic filtering (HA). Both systems were compared on: (a) message appreciation, (b) engagement with the system, and (c) one's own self-reported smoking cessation status, as indicated by the last seven-day point prevalence (7D-PP) report in different time intervals, by looking at their evolution during the six months, and the overall main effects. Additionally, we analyzed potential subgroup differences in these three outcomes by gender, age (categorized as either an older or younger generation), profile completion, motivation to quit levels, and nicotine dependence levels.

In the following sections we explain how this study was performed, including its materials and methodology, and provide an overview of the demographics of the recruited participants, the results of each metric, and a discussion on the findings.

Related Works

Traditional computer tailoring interventions were effective in different contexts, such as reducing alcohol drinking [34–36], nutrition [4,37,38], cancer screening [39,40], physical activity [9,10,41], sunscreen usage [42], and smoking cessation [2,43–47]. Personalized tailored interventions were delivered in the form of messages and reminders through emails and letters [34–38,46,47]. The possibility of emails arriving in spam folders resulted in low adherence rates, and limiting of the participants' involvement. Thus, these interventions indicated the need for exploring mechanisms to deliver messages other than emails and letters [36,38].

Moreover, traditional computer tailoring was based on 'static' scores for each individual's answers. Researchers started considering the usage of recommender systems in a number of health interventions. A recent well-studied therapeutic area is nutrition [48–52]. For example, Elahi et al. (2015) introduced and evaluated a food recommender system, obtaining user preferences through an interaction design, which showed positive user feedback [53]. Musto et al. (2020) [50] presented a strategy using knowledge-aware recommender systems to recommend an appropriate diet. Gómez-del-Río et al. (2020) described an activity recommender system to promote healthy habits in obese children, while using gamification as a mechanism to enhance personalization and increase user motivation [54]. In addition, several other studies have investigated the usage of recommender systems to help people nourish themselves more healthily [55–58].

However, there has been a need to optimize content not only towards individual past preferences, but also to support healthy content, thus ensuring optimization to a user's health goals [48].

Other therapeutic areas, such as cervical cancer [59], diabetes [60], and mental health [61] have also been explored. However, they did not implement an intelligent artificial intelligence-based framework [60], and some even lacked a filtering functionality on the recommended activities [61].

For the specific case of smoking cessation support, a previous study using an HRS which featured a hybrid algorithm (a combination of demographic filtering, content-based, and utility-based approaches) showed positive effects of the HRS that supported behavioral changes [62,63]. This hybrid HRS was embedded into a mobile app incorporated into the routine care workflow of a hospital-based smoking cessation unit, and the app was offered to patients referred from other specialized care units of the hospital. The PERSPeCT study [64] compared rule-based computer tailoring to a hybrid recommender system. However, they did not assess a six month point smoking cessation, being limited to shorter period outcomes. Another study, Smoker2Smoker [65], used a recommender system that combined content-based ranking and collaborative filtering methods to select a message for an individual participant. However, being an observational study, there were limitations in ascertaining causality.

As recommender systems have grown in interest, multiple literature reviews have been conducted in recent years around the HRS topic [28,66–71]. Hors Fraile et al. (2018) suggested improvements in HRSs by utilizing relevant behavior change theories and applying important features of tailored interventions [28,67]. A review by Ferretto et al. (2017) indicated limited use of recommendation systems in mobile healthcare applications, although the use of mobile phones and connection to the internet is ever increasing [69]. Another scoping review by Cheung et al. (2019) illustrated the potential benefits and need for incorporating a collaborative filtering method with demographic filtering as a second step to knowledge-based filtering in HRSs [67]. The authors also demonstrated the potential of this type of hybrid HRS towards enhancing user experience. To the best of our

knowledge, there has been no randomized study comparing two smoking cessation HRSs and further analyzing the sub-group differences.

2. Materials and Methods

2.1. Design

A trial was approved by the Ethical Committee of Taipei Medical University Joint Institutional Review Board (TMU-JIRB), and conducted from 10 November 2017 to 15 January 2020. After voluntarily downloading the mobile app, without any prior recommendation of any clinician, and filling out a baseline measurement, respondents were randomly assigned to either the KBA smoking cessation intervention or the HA smoking cessation intervention. Baseline measurements were taken between 10 November 2017 and 15 July 2019, and post-tests ended 6 months after the baseline. Between the baseline and final follow-up, respondents could freely interact with the system to obtain more motivational message support, as well as provide evaluations on the various messages they received.

The Android version of the app used in this study was launched on 10 November 2017, and the iOS version of the app was launched on 6 August 2018. Different time intervals were defined to perform in-depth analyses to measure the evolution of the two user groups: 0~7, 8~14, 15~21, 22~30, 31~60, 61~120, and 121~180 days. These time intervals were defined to have a good understanding of participants' evolution throughout the study, aiming to distribute the collected data in time periods where none would be without data. As time passes, participants are likely to stop using the app as intensively as in the initial days. The registered information would then be sparser across later months, and therefore the time periods should increase their duration. Grouping the data in longer time intervals allowed more effective analysis of each period, and the intervals are commonly assessed time points for smoking cessation assessments (7 days, 15 days, 1 month, 2 months, and 6 months). The initial time point of these time intervals was the first day with a valid quitting attempt.

2.2. Interventions

Both HRSs selected smoking cessation motivational messages and sent them to a mobile app used in the 3M4Chan study [72], performed within the H2020 Project Smoke-FreeBrain [73]. The mobile app interface was the same in both groups (Figure 1), as were the personalized elements included in the motivational messages, such as referencing the name of the user in the message and the initial message delivery frequency between participants.

According to Abroms et al. (2015) [74], five messages were sent on the quitting date, one message per day during the first week after the quitting day, and three messages a week after that. This frequency could be changed every 14 days after the second week of the quitting attempt by participants themselves, regardless of their group. Users were given the choice of changing the frequency, i.e., increasing it up to one message per day or decreasing it down to one message per week, by answering a weekly question within the app. The time to send each message was set at random, given an allotted time range previously configured by each participant.

In total, the system had 311 different messages. The same message could only be sent a maximum of three times to a participant, as explained in the system design description by Hors-Fraile et al. (2019) [26]. Also, users could report their abstinence status in the app by answering the following weekly question: "Are you still resisting the temptation or have you smoked? Please, be honest." And possible answers were (a) I have not smoked; (b) Only one cigarette; (c) Two or three cigarettes; and (d) Four or more cigarettes.

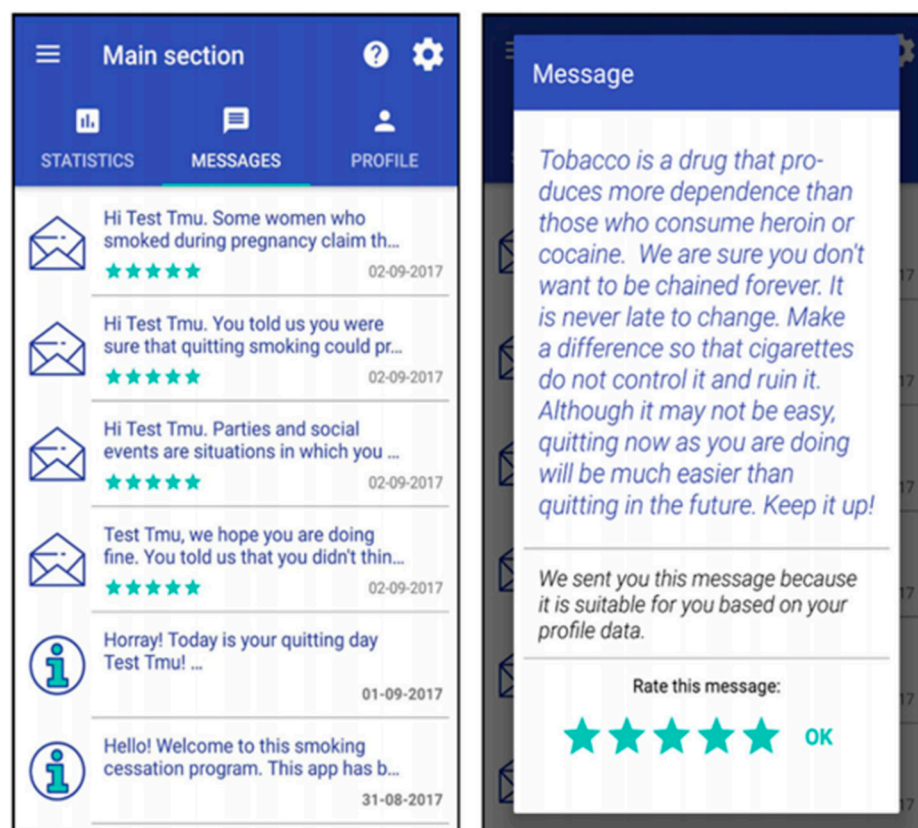


Figure 1. English version of the mobile app interface.

2.2.1. Messages

The messages used in the study were created by a behavioral science researcher in English. They were then translated into Mandarin Chinese, and validated by two Taiwanese doctors specialized in smoking cessation. The messages followed the tailoring recommendations for smoking cessation support made by the World Health Organization (WHO) [75]. These guidelines included case scenarios of behavioral support sessions, reasons, strategies, and tips for people to stop smoking. We reflected that knowledge in the sentences by elaborating on recurrent topics and following suggested approaches to support abstinence.

In addition, we also introduced determinants of change psychological constructs which influence a behavior proposed in the I-Change model [29], which was previously used for smoking cessation studies [76–79]. The chosen determinants were used in previous studies to increase awareness, raise motivation, and change behaviors [78,80–83]. The included determinants were: attitudes towards stopping smoking (the perceived advantages and disadvantages of quitting), social support to quit, skills to manage situations when feeling tempted to smoke, self-efficacy to quit, as in the perception of the smoker's ability to achieve cessation, and action planning of the tasks to be successful (e.g., throwing away all ashtrays at home). All messages were enounced from a positive point of view, addressing the reader with the pronoun 'you', using the active voice and easy to understand vocabulary, avoiding technical terms and complicated words, and with a maximum word count in English of 200, with an average of 85.5 words per message. Ten of the behavior change techniques proposed by Abraham et al. (2011) [84] were included across the messages. The messages also considered health communication methods, such as repeating answers, creating empathy, adding new knowledge, and changing existing misconceptions. Additional examples of the applications of these techniques can be found in Appendices C and D of the study by Hors-Fraile et al. (2019) [26]. As an illustrative example, the following sentence presents a motivational message intended to enhance the social skills of a participant called 'John',

who reported struggling with social pressure to smoke when completing his profile during the enrollment process in the mobile app: “Hi John. You told us that you cannot refuse a cigarette when someone offers it to you. Well, we understand that you may struggle with it. However, you can practice how to kindly refuse a cigarette. A good reinforcing strategy is to add to the sentence one of the reasons you have to quit smoking. ‘No. I am quitting because I want to keep my teeth clean and white’, or, ‘No. I am quitting to have a longer life and enjoy it with my grandchildren’. In this way, you say no and also remind yourself why you are doing it. If you prefer not to say the reason, you can just think it. To get a natural and almost immediate reaction, you can ask someone you know to role play a person inviting you to smoke, and you have to reject the invitation. Even if you think this exercise is not worth doing, it can help you succeed in a real situation.”

2.2.2. Knowledge-Based Algorithm (KBA) System

The KBA system computed an adapted KBA approach over a pool of smoking cessation motivational messages. The adaptation consisted of pre-selecting message characteristics based on five user profile features: age, gender, quitting date, number of smoked cigarettes, and weekly expenditure on tobacco products. If the message dealt with time or date-specific content (e.g., things to do early in the morning, related to the weekend, etc.), then the context when the message was going to be sent (the time and week day) was also considered, and it was calculated before starting this KBA. Thus, the user requests which are necessary to run knowledge-based algorithms [85] were fixed by design. This was done so that participants would receive only one message according to the message frequency delivery pattern of Abroms et al. (2015) [74]. In addition, the intended intervention behavior was not to give participants the possibility of choosing from different messages, but to send them one relevant message per delivery. In this way, participants could focus their attention on a single message concept, not being flooded with different concepts at the same time, which might not be good for remembering them, and which could even overwhelm participants, producing anxiety and becoming a trigger to smoke.

Consequently, the KBA retrieved all messages compatible with the given participant profile. A message was compatible with the participant profile if all the meta-features (the defining attributes) of the messages were also found in the user profile. If a message did not have a specific meta-feature value it was considered a valid match. For example, a message that did not have a gender meta-feature was assumed to be compatible for both male and female users. Also, we enforced the algorithm to ensure a strict matching process. This means that if a message has a single meta-feature that is not present in the user profile, that message cannot be sent to that user. A simple illustrative case is shown in Table 1 below. The last column refers to whether the different messages are compatible for the user profile of a participant with the following meta-features: middle-aged male, less than 10 cigarettes per day, whose quitting date was 9 days ago, with a low weekly expenditure, and who set no time limit to receive motivational messages during the day. The intermediate columns of the table represent the meta-features of each message.

Table 1. Example of how the KBA processes message compatibility for a given user.

Potentially Compatible Messages	Message Meta-Features					
	Age Group	Gender	Number of Daily Smoked Cigarettes	Weekly Expenditure	Quitting Date	Context
Message A	Young	Male		High		Only for mornings
Message B			<10		Less than 15 days ago	Only for weekends
						No (Not matching in age group & weekly expenditure)
						Yes

Table 1. Cont.

Potentially Compatible Messages	Message Meta-Features					
	Age Group	Gender	Number of Daily Smoked Cigarettes	Weekly Expenditure	Quitting Date	Context
Message C	Old	Female	>20			Only for evenings No (Not matching in age group, gender, and number of smoked cigarettes)
Message D	Middle-aged			High		Only for mornings Yes
Message E		Male	10–20	Low		No (Not matching in number of daily smoked cigarettes)
Message F					More than 30 days ago	No (Not matching in quitting date)
Message G		Female			Less than 7 days ago	No (Not matching in gender, and quitting date)
Specific meta-features for the target user profile	Middle-aged	Male	<10	Low	9 days ago	No restrictions

The result of applying the KBA algorithm was a list of compatible and potentially relevant messages for that user. Then, as only one message could be sent at a time, the system selected messages from the list that had been sent fewer times to that user and picked one at random in case of a draw.

2.2.3. Hybrid Algorithm (HA) System

The HA system computed a two-step algorithm approach, also known as cascade [86]. The first step performed the same knowledge-based algorithm as in the KBA described above to filter potential non-compatible messages with the target participant. In the second step, a user-based demographic filtering algorithm was applied to the output of the first step (as opposed to random selection used in the KBA). This second step selected a motivational message based on the premise of prioritizing those which were found relevant by other similar participants (called neighbors) [87]. The selection was done using a score that was calculated for each message. This score represented the probability that the message was relevant for the given user. The algorithm used for this calculation worked as follows:

First, all users and their message ratings were represented in a matrix (not including the user and ratings for which we are going to send the message). This means that we did not limit the neighbor size limit (the number of users used to compute the similarity). Second, a neighbor similarity index was computed for each user in the matrix following the equation below.

$$sim(A, B) = \frac{\sum_{i=1}^n \left(\delta_{F_A(i), F_B(i)} \right) + \sum_{j=1}^m \left(\frac{\delta_{F_A(j), F_B(j)}}{\max(|F_A(j)|, |F_B(j)|)} \right)}{\sum_{k=1}^{n+m} \left(\delta_{F_A(k), F_B(k)} \right)} \quad (1)$$

where:

A and B are two users;

F_u represents all meta-features completed by user ' u ';

$F_u(x)$ represents the value of the meta-feature ' x ' of user ' u ';

$|F_u(x)|$ represents the number of potential valid values of meta-feature ' x ' of user ' u ';

n is the total number of meta-features which can have only two values (e.g., yes or no, male or female, etc.);

m is the total number of meta-features which can have more than two values (e.g., high, medium, or low motivation level);

$\delta_{y,z}$ represents a function that sums the number of matching meta-features between the lists of meta-features ' y ' and ' z '.

The system used all available information about user meta-features in the neighbor (with a minimum of nine, i.e., the core questions, and a maximum of 60 variables, after adding 51 voluntary extended-profile questions). The more meta-features in common between participants, the more similar their profiles were and the higher the likelihood of rating messages similarly. Hence, this similarity calculation approach could cover both participants who completed the extended profile and those who did not.

Third, the ratings for a message were multiplied by the corresponding user similarity index and added together following the equation below.

$$message_relevance_score_{m,u} = \sum_{i=1}^n rating_{m,i} * sim(i, u)$$

where:

$rating_{m,i}$ is the rating given to message m by user i ,

u is the user that is going to get the recommendation

i is one of the neighbors of u

n is the total number of neighbors

If a message had been re-rated by the same user, only the last rating was considered.

Fourth, the message relevance scores were normalized to a range between 0 and 1. Messages which were not rated previously were assigned a 0.5 final score. This represented a 50% chance of being relevant or not. Fifth, the algorithm selected only the subset of messages which had been sent to the user a lower number of times to maximize message diversity and minimize repetitions. Messages which had been sent three times already to that user were discarded. We chose three as the maximum number because we wanted users to realize their meanings, and the implications of the message arguments, following the recommendations of Cacioppo et al. (1989) [88]. Messages sent once or twice previously were sent with a complementary sentence in the same message acknowledging the fact that the message had, at some point, been sent already to the participant, but stressing the importance of the message content as to why the participant was receiving it again. Finally, a weighted random selection, based on the relevance score, was run to select the message to be sent.

Figure 2 summarizes the message selection process for the two groups, and the detailed selection process, as previously explained by Hors-Fraile et al. (2019) [26].

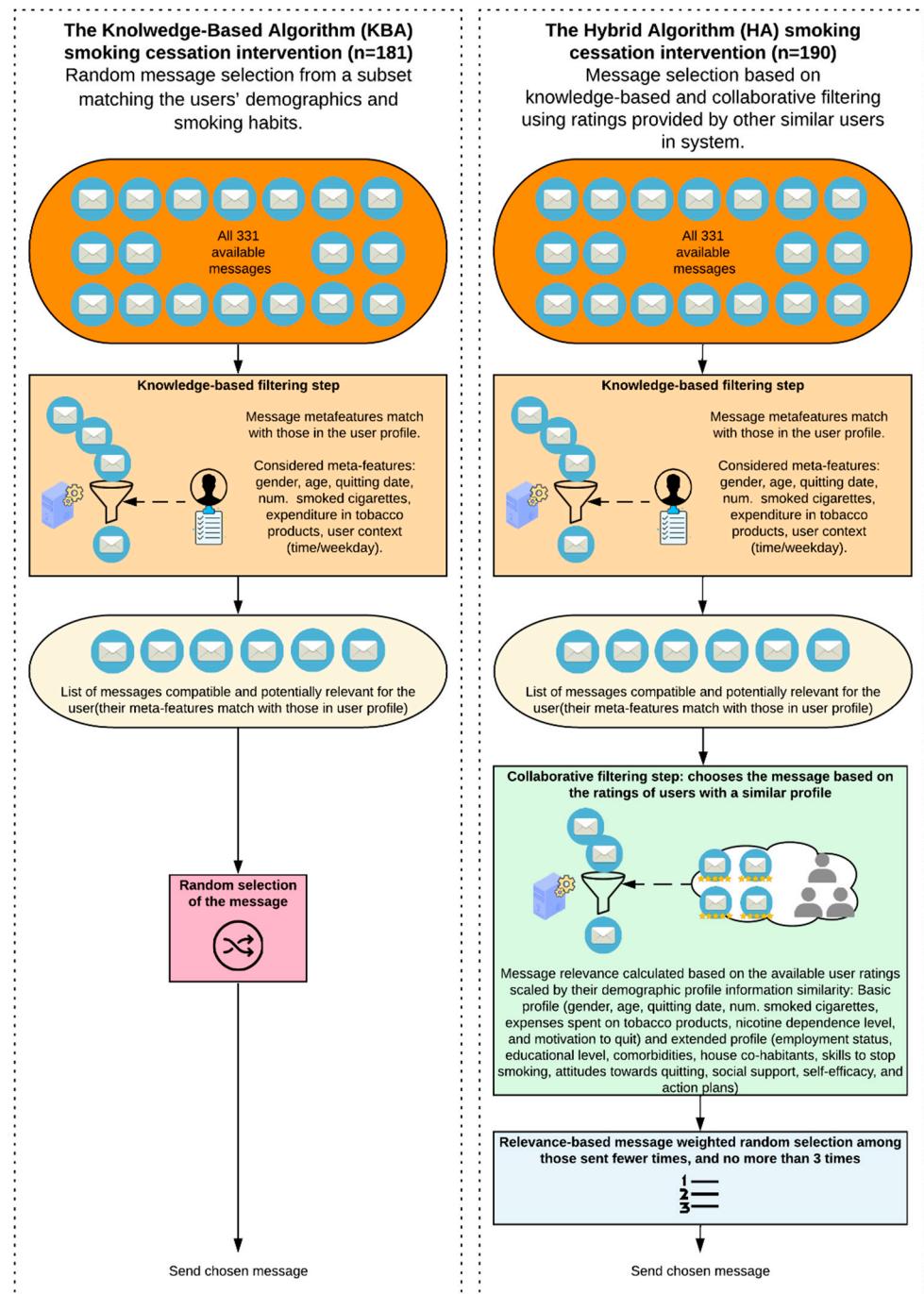


Figure 2. Message selection summary for the two HRSs used in the study.

2.3. Participants and Recruitment

Participants considered for this study were smokers over 20 years of age, the legal adult age in Taiwan, willing to quit, who owned a compatible smartphone (Android or iPhone), and who downloaded the smoking cessation “Quit and Return” app, registered on 15 July 2019 or earlier, accepted the terms and conditions of the mobile app, made a valid quitting attempt, and who were able to understand any of the two languages in which the mobile app was offered: Mandarin Chinese or English. A quitting attempt was considered valid if it was not cancelled in the first 24 h after it was initiated, when the quitting day was not set in the past, and when the quitting day was set within a maximum of 90 days in the future with respect to the user registration date in the system. The system used a

JavaScript random function to randomly allocate registered users between the two groups to receive motivational messages to stop smoking using a different personalization strategy. Participants were blinded to this randomization.

2.4. Measurements

2.4.1. Demographics

All participants had to answer nine core questions on their profile to use the app. These included their gender, age, employment status, date on which they began smoking, quitting date, number of cigarettes smoked weekly, and amount of money spent weekly on tobacco. In addition, they had to complete two standardized questionnaires to determine their nicotine dependence [89] and motivation to quit [90]. They could also voluntarily complete a questionnaire with 51 additional extended-profile features about comorbidities, people they share their house with, educational level, physical activity routines, and questions based on the I-Change behavioral change model [29], such as skills for stopping smoking, attitudes towards quitting, social support, self-efficacy, and action plans. This extended-profile questionnaire was used in the intervention by the HA to create a more comprehensive user profile and calculate similarities in the demographic filtering step.

We considered five user attributes as indicators of potential differences in the study outcome. Four of them were direct variables from the previously described core questions, and the fifth attribute was derived from answering the voluntary questions.

First, we assessed gender (male = 0, female = 1).

Second, we assessed nicotine dependence based on the Fagerström test [89] (low = 0, high = 10). Its result was used to classify participants following the categorization included in the Fagerström test: participants with scores of ≤ 4 were included in the “low-dependence” sub-group (recorded as 0), those with scores of 5 or 6 were included in the “medium-dependence” sub-group (recorded as 1), and those with scores of ≥ 7 were included in the “high-dependence” sub-group (recorded as 2).

Third, we assessed the level of motivation to quit based on the Richmond test [90] on a scale from 0 (low) to 10 (high). These test scores were used to classify participants into four motivation sub-groups proposed in the Richmond test using specified cutoff points [90]. We merged the low and medium-low sub-groups to avoid having two groups with too few participants. Thus, smokers with scores of ≤ 5 were assigned to the “low and medium-low motivated” (recorded as 0) sub-group, those with scores of 6 or 7 were included in the “medium-highly motivated” sub-group (recorded as 1), and those with scores of ≥ 8 were included in the “highly motivated” sub-group (recorded as 2).

Fourth, we assessed their age on the starting day of their first valid quitting attempt. To do so, we created two participant age sub-groups (generations), based on their potential familiarity with technology, as proposed by Berkup [91]: “baby boomers or older” (born before 1964) and “generation X or younger” (born after 1965) (baby boomers or older = 0, generation X or younger = 1).

Finally, we assessed participants’ completion of the extended-profile questionnaire by checking their profile data registry (incomplete = 0, complete = 1). The latter measurement was only relevant for participants receiving recommendations generated by the HA, because the intervention with the simpler knowledge-based algorithm (KBA group) was not influenced by the extended-profile questionnaire.

2.4.2. Outcomes

Message Appreciation

Message appreciation was assessed by comparing the average rating provided to messages rated by users of the KBA and HA groups in each time interval. Rating options ranged from one to five stars. Users could self-determine whether to rate messages or not, and so the frequency of message appreciation differed between persons.

Engagement with the System

We assessed three metrics in each time interval: (1) the number of active days in the system, with an active day defined as any day when a participant opened the app at least once; (2) the number of rated messages that were sent and rated in a given time interval, and (3) the number of abstinence reports submitted in each time interval. A fourth metric, the number of registered quitting attempts that were not cancelled within the following 24 h after being set, was assessed for the entire intervention period, as it did not make sense to subdivide it. We considered those attempts cancelled within 24 h as people just exploring the app with no real commitment to quit smoking.

Smoking Behaviors

The 7D-PP smoking status was assessed using self-reported abstinence reports by participants in each time interval. The 7D-PP metric was defined as a self-report of smoking no cigarettes (not even a puff) in the last 7 days. Smokers had to report whether they were currently smoking (1) or non-smoking (0), based on having relapsed or being abstinent at the time of the report (regardless of the number of cigarettes smoked). A maximum of one abstinence report could be sent each week. As most time intervals lasted several weeks, several self-reported status reports could be sent during a given time interval; we always considered the last one submitted within each time interval. Thus, this metric provides the 7D-PP for the last self-reported smoking status submitted in each time interval. The 7D-PP was identified as an important metric to assess self-reported smoking cessation outcomes [92].

To compare smoking cessation outcomes with other studies, smoking behavior changes were calculated as the proportion of eligible participants that sent positive and negative self-reported abstinence reports. Three analyses were conducted. First, an analysis on available data, averaging the values of the last 7D-PP abstinence reports of the intermediate time intervals. This was done taking the last report within each time interval, and then using those values to calculate the average across the study for each participant. Second, an analysis on the 7D-PP abstinence taking the very last available abstinence report value ever reported in the study by each participant. Third, a pessimistic analysis of the previous one where non-respondents of the abstinence reports within each time interval were considered as non-abstinent (penalized imputation). For this, we took the very last abstinence report value ever reported, but we considered relapsers; those participants who did not submit an abstinence report. We used this latter analysis as a conservative approach to avoid optimistic comparison effects of the intervention [93].

2.5. Statistical Analysis

To analyze the data, we first performed a descriptive analysis of the sample demographics. Secondly, three kinds of dropout were defined:

- dropout in terms of no longer sending message ratings,
- dropout in terms of no longer sending abstinence reports, and
- dropout in terms of no longer being active on the app.

For each of these, logistic regression was performed to identify potential determinants of dropout, which were subsequently used as covariates in the primary analyses on the effects of the type of HRS. The independent variables used in the dropout analysis were:

- gender,
- nicotine dependence levels (low, medium, high),
- motivation level (low and medium-low, medium-high, and high),
- age (born after 1965 versus born on, or before, 1965),
- employment situation (employed versus unemployed),
- and completion of the extended profile (yes versus no).

To take care of the dependencies within observations of each subject, mixed models were used to examine the effects of the app. Depending on the scale level and distribution

of the outcome, a different model was used. For message appreciation (after categorizing this variable into 4 levels), an ordinal mixed model, for engagement metrics (number of active days, number of ratings, number of abstinence reports, and number of quitting attempts), a negative binomial regression mixed model, and for smoking cessation metric, a logistic mixed model. For each outcome, a suitable model for the (co)variances of the random effects was chosen to adequately capture the dependencies in the outcome across time. Since for smoking cessation, none of the examined covariance structures of random effects convergence was obtained, the analysis was done by standard logistic regression.

To compare smoking cessation outcomes with other studies, smoking behavior changes were calculated as the proportion of eligible participants that sent abstinence reports in each time interval. Next, both an analysis on available data, according to logistic regression, and a sensitivity analysis were done, assuming a pessimistic scenario (penalized imputation) in which non-respondents within each time interval were considered to be non-abstinent.

For all metrics, we also performed an in-depth analysis of the impact of having completed the extended profile questionnaire, using the same type of regression model for each metric. In statistical testing, a significance level of $\alpha = 0.05$ was used.

3. Results

3.1. Description of the Sample and Involvement Level

In total, 844 participants downloaded the app and registered. Among these, 371 had a valid quitting attempt and were eligible for our study; 290 (78.16%) were male and 81 (21.83%) were female with a mean age of 36.90 (standard deviation (SD) 10.21, coefficient of variation (CV) 0.28) years. The mean nicotine dependence-level score was 5.13 (SD 2.59, CV 0.50), and the mean motivation-to-quit score was 7.54 (SD 1.91, CV 0.25).

In total, 181 users were allocated to the KBA group and 190 users were allocated to the HA group; these numbers slightly differed because of differences in making a valid quitting attempt. No statistically significant differences between the two groups were found in age, gender, nicotine dependence, motivation to quit, employment status, or completion of the extended-profile questionnaire. Table 2 shows a more detailed analysis. In total, 843 rated messages and 373 abstinence reports were provided by the 371 users of both systems during the 6-month period after making their first valid quitting attempt.

Table 2. Participants' demographic and smoking characteristics distribution between groups.

Variable	Total (<i>n</i> = 371)	KBA (<i>n</i> = 181)	HA (<i>n</i> = 190)	Test Statistics <i>p</i> -Value
Mean age, years (SD, CV)	36.90 (10.21, 0.28)	37.71 (10.78, 0.29)	36.13 (9.60, 0.27)	F(1369) = 2.229 <i>p</i> = 0.136
Generation distribution: Percentage younger (no.)	92.7% (344)	90.6% (64)	94.7% (180)	$\chi^2 = 2.34$ * <i>p</i> = 0.13/ <i>p</i> = 0.16
Gender: Percentage male (no.)	78.2% (290)	76.8% (139)	79.5% (151)	$\chi^2 = 0.39$ * <i>p</i> = 0.53/ <i>p</i> = 0.62
Mean nicotine dependence score (SD, CV)	5.13 (2.59, 0.50)	5.08 (2.69, 0.53)	5.17 (2.50, 0.48)	F(1386) = 0.13 <i>p</i> = 0.77
Mean motivation to quit score (SD, CV)	7.54 (1.91, 0.25)	7.54 (2.03, 0.27)	7.55 (1.80, 0.24)	F(1369) = 0.01 <i>p</i> = 0.93
Employment status: Percentage employed (no.)	83.0% (308)	83.4% (151)	82.6% (157)	$\chi^2 = 0.04$ * <i>p</i> = 0.83/ <i>p</i> = 0.89
Profile completion: Percentage completed (no.)	43.9% (163)	42.0% (76)	45.8% (87)	$\chi^2 = 0.54$ * <i>p</i> = 0.46/ <i>p</i> = 0.47

Notes: KBA, knowledge-based algorithm group; HA, hybrid algorithm group; SD, standard deviation; CV, coefficient of variation; * next to the *p*-value of the asymptotic chi-square test (left), also the *p*-value of the Fisher exact test is given (right).

3.2. Dropout Analysis

The dropout analysis for message ratings showed that there was no interaction effect between the app and time period on the dropout rate ($p = 0.394$), so a potential interaction between the group and time period on the average ratings, or the number of rated messages, could not be due to differences in dropout rate at the different time intervals. There was also no main effect of the app for this type of dropout ($p = 0.375$). The variables of gender, nicotine dependence, motivation level and completed extended profile turned out to be predictors of this type of dropout and were subsequently used as covariates in the analysis of the outcome variables of message ratings and number of rated messages.

For the number of active days there was also no interaction effect of HRS group and period on dropout rate ($p = 0.910$). So, the difference between the HRS groups in terms of dropout rate did not differ across time. Averaged across time, there was also no difference in dropout rate between the two HRS groups ($p = 0.583$). The variables gender, employment situation, nicotine dependence and completed extended profile were significant predictors of this type of dropout and were subsequently used as covariates in the analysis of the outcome variable number of days active.

For dropout based on smoking cessation reports, no significant HRS group by period interaction ($p = 0.682$), and also no main effect of HRS group ($p = 0.158$), was found. So, the difference between the groups, in terms of dropout rate, did not differ across time, and there was also, averaged over the time intervals, no difference in dropout rates between the groups. Only completion of the extended profile was a significant predictor of dropout ($p = 0.002$).

3.3. Overview of Outcomes

All outcome parameters, excepting that of the number of quitting attempts, were assessed from four perspectives: their evolution across time, the main effect of the type of app, the potential interaction effect between time and type of app, and the possible interaction effect between type of app and having completed the profile or not. The number of quitting attempts, measured over the whole study period, was only studied regarding the main effect of the type of app, the interaction of the type of app, and having completed the profile or not.

3.3.1. Message Appreciation Results

First, we assessed the evolution of the mean message appreciation (range 0–5) by each of the 2 HRS groups across the different time intervals (see Figure 3, where the Y axis is extended beyond the maximum mean level to avoid cutting the top part of some error bars). Next to the means, in each figure the 95% confidence intervals (CI) are also displayed, represented by the two error bars around each of the sample means.

First, the ordinal mixed model analysis showed that the difference between the two groups, in terms of appreciation, did not develop differently across time ($p = 0.897$). Also, there was no statistically significant difference between the two groups in terms of average ratings across time (KBA: Mean = 3.835, SD = 1.218, CV = 0.318 vs. HA: Mean = 4.185, SD = 1.205, CV = 0.288, $p = 0.353$). Hence, there was no difference between the KBA and HA groups in terms of the level of message appreciation.

Third, concerning potential interaction effects with other factors, the analysis on the role of completing the extended profile showed that the effect of having completed the extended profile did not differ between the KBA and HA groups ($p = 0.980$). Also, for both groups, completers of the extended profile showed the same level of message appreciation as those who did not complete it ($p = 0.976$). However, as emerged from the dropout analysis, for both HRS groups completers of the extended profile had a higher probability to stay in the study ($p < 0.001$). No interaction effects were found for the other factors (gender, age, nicotine dependence level, and motivation to quit).

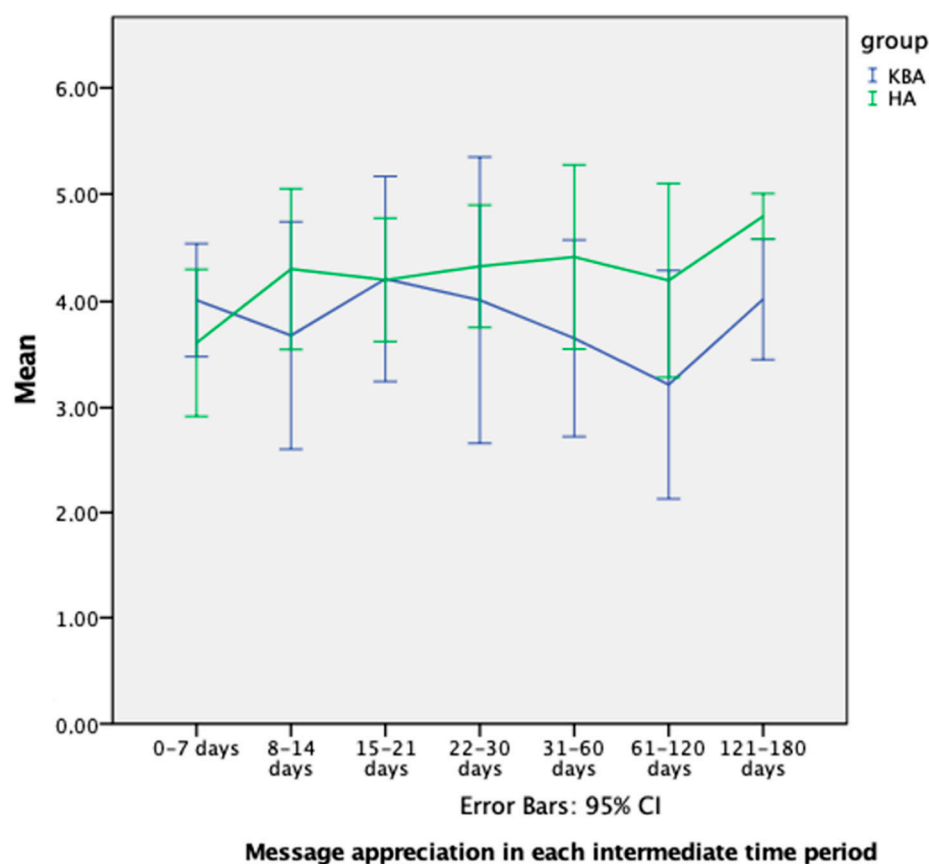


Figure 3. Evolution of the message appreciation (a higher score is a higher appreciation).

3.3.2. Engagement Results

Number of Rated Messages

First, concerning the evolution of the number of rated messages for each of the HRS groups across the different time intervals (as shown in Figure 4), the results of the negative binomial regression mixed model showed that the number of rated messages did not develop differently for the two HRS groups across time ($p = 0.920$). We chose a negative binomial regression because the variance of the outcome variable was much larger than its mean, resulting in overdispersion. Thus, negative binomial regression was preferred over Poisson regression. Second, we did not find an overall main effect of the type of app when considered across the whole 0–180 days' period. Both groups had a similar number of rated messages (KBA: Mean = 3.59, SD = 5.877, CV = 1.637 vs. HA: Mean = 3.31, SD = 5.677, CV = 1.715; $p = 0.884$). Third, concerning interaction terms, the analysis of the impact of completing the extended profile questionnaire showed that its effect on the number of rated messages did not differ between the KBA and HA groups ($p = 0.894$) and also, when averaged across both groups, had no significant effect on the number of rated messages ($p = 0.156$). However, there was an effect on dropout: completers of the extended profile had a significantly lower probability to dropout from the study ($p < 0.001$).

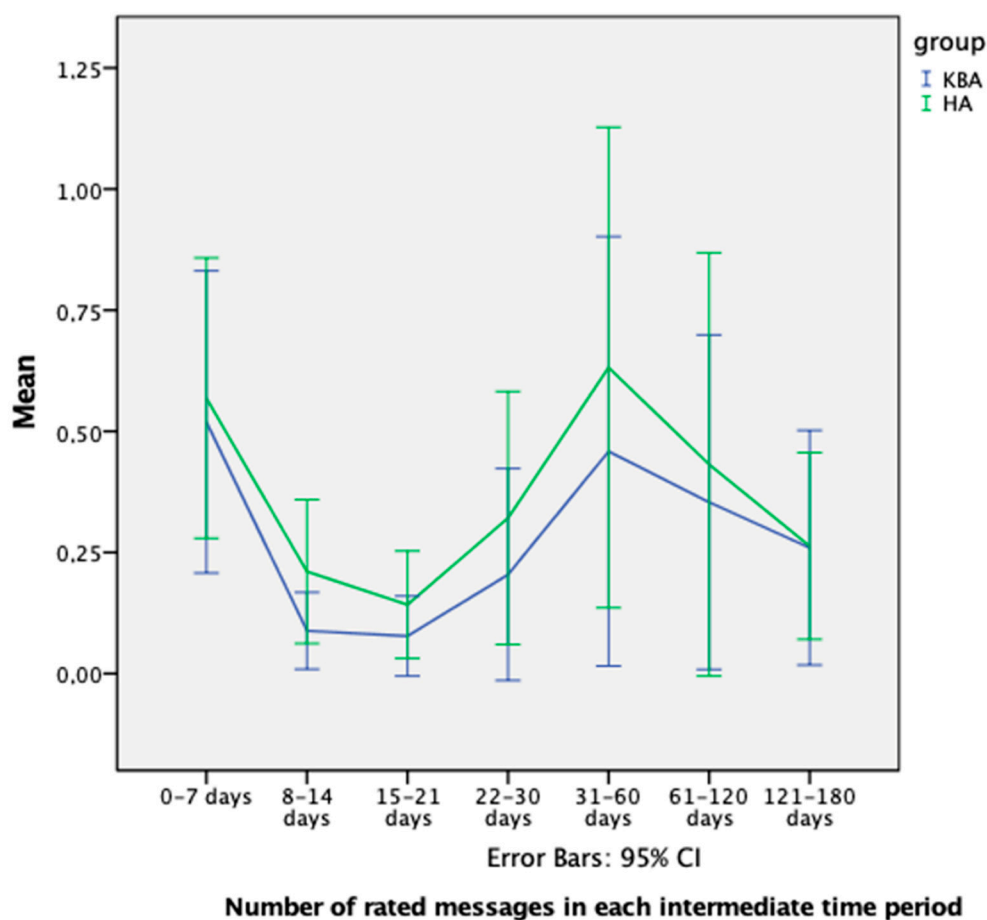


Figure 4. Evolution of the number of rated messages (higher is better).

Number of Active Days

First, the evolution of the mean number of active days on the mobile app for each of the HRS groups across the different time intervals (see Figure 5) developed differently across time for HA and KBA, although it was marginally significant ($p = 0.051$).

We also found that the covariate employment status of the participants had a marginally significant interaction with the HRS group ($p = 0.070$). Consequently, the examination of the differences between the two HRS groups at each time interval was done separately for employed and unemployed participants. To control the type I error rate, for each of these groups, Holm correction was applied when testing the app difference at each of the 7 time intervals. Unemployed participants showed no difference between the two app groups for any time interval in terms of active days. However, employed participants had a significant difference in active days between the two app groups for the last two time intervals of 61–120 days ($p = 0.007$), and 121–180 days ($p = 0.008$), where the HA led to less active days than the KBA.

Examining the main effect of type of app over the 0–180 days' time period also showed that the KBA led to a higher number of active days (KBA: Mean: 2.74, SD: 4.372, CV = 1.596 vs. HA: Mean = 1.99, SD = 2.336, CV = 1.174; $p = 0.043$).

There was no interaction effect of HRS group and completion of the extended profile on the number of active days ($p = 0.815$). The difference between the app groups in the number of active days was the same for the completers and non-completers of the extended profile. There was a marginally significant effect of having completed the profile on the number of active days ($p = 0.068$), in that completers were, on average, active for more days. Also, completers of the extended profile had a significantly lower probability of dropping out of the study ($p < 0.001$).

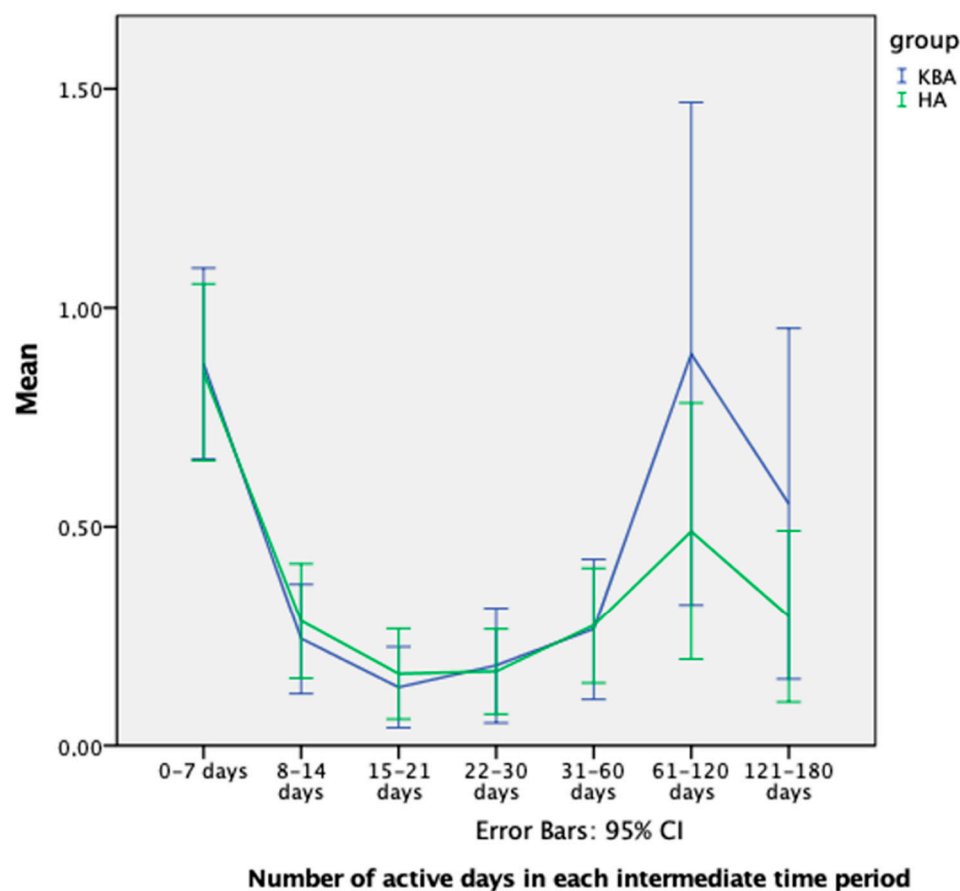


Figure 5. Evolution of the number of active days (higher is better).

Number of Quitting Attempts

An analysis, by the negative binomial regression model, of the number of quitting attempts over the whole study period (0–180 days) showed that there was no significant difference between the two groups in terms of the number of quitting attempts (KBA: Mean = 1.43, SD = 1.291, CV = 0.903 vs. HA: Mean = 1.71, SD = 2.431, CV = 1.369; $p = 0.275$).

Testing the interaction between HRS group and having completed the extended profile on the number of quitting attempts showed that there was a significant interaction effect ($p = 0.042$). The effect is such that the effect of having completed the extended profile on the number of quitting attempts is significantly larger in the HA group. In the HA group there was a very significant effect of having completed the extended profile on the number of quitting attempts ($p < 0.001$), with the incidence rate ratio of completers versus non-completers being 1.869. In the KBA group there also was a significant effect of having completed the extended profile on the number of quitting attempts ($p = 0.020$), the incidence rate ratio of completers versus non-completers being, however, lower, 1.342.

Number of Abstinence Reports

The evolution of the number of sent abstinence reports for each of the HRS groups across the different time intervals is shown in Figure 6.

The difference between the two HRS groups in terms of the number of smoking cessation reports did not develop differently across time ($p = 0.339$). However, when testing the difference between the two HRS groups over the 0–180 days' period, the average amount of cessation reports (KBA: Mean = 1.11, SD = 1.806, CV = 1.627; HA: Mean = 0.7, SD = 1.118, CV = 1.597) was significantly higher for the KBA group ($p = 0.001$). Note that this cannot be explained by the dropout rate, as there were no significant differences in dropout rate between the two app groups.

There was no significant effect of having completed the extended profile ($p = 0.870$), implying that completers of the extended profile had the same incidence rate of the number of cessation reports as the non-completers (and that was the case for both app groups). However, completers of the extended profile in both groups had a higher probability to stay in the study ($p = 0.002$).

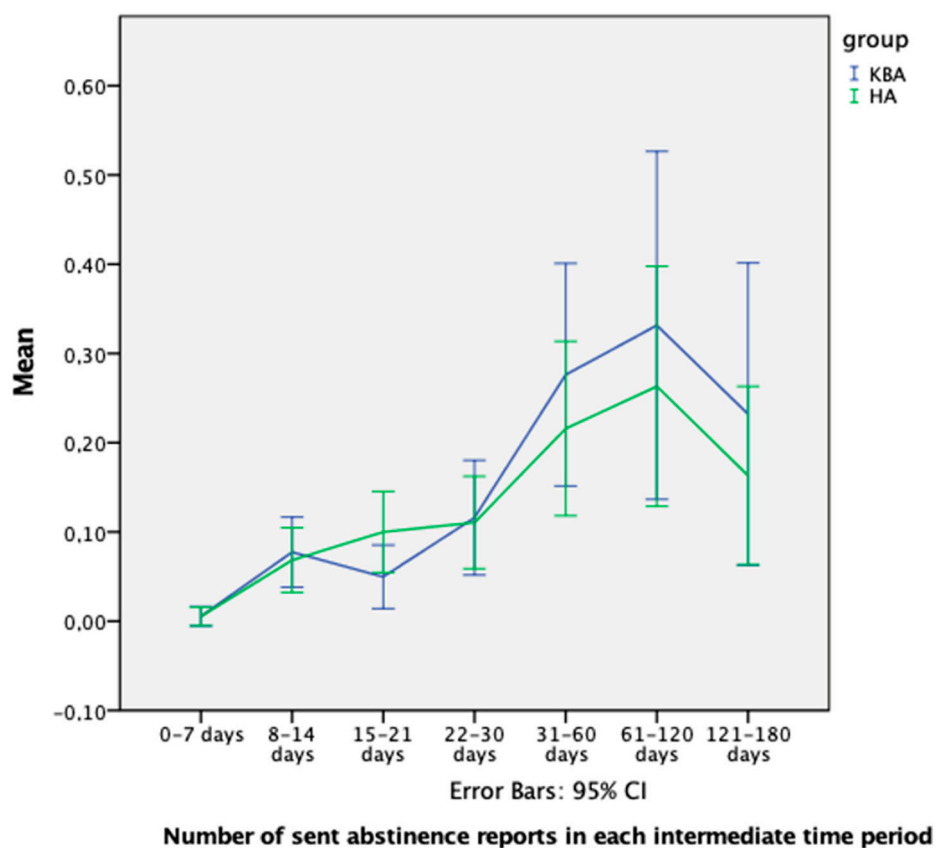


Figure 6. Evolution of the number of sent abstinence reports (higher is better).

3.3.3. Smoking Cessation Results

7D-PP Abstinence: Analysis on Available Data

The evolution of the 7D-PP abstinence for each of the HRS groups across the different time intervals is shown in Figure 7 (the mean was calculated by averaging the last abstinence report of each participant in each group in each period, coded as 0 for being abstinent and 1 for relapse).

The analysis based on the standard logistic regression model revealed no statistically significant interaction between group and time period ($p = 0.737$). Also, the two HRS groups did not differ significantly ($p = 0.123$) on 7D-PP, averaged across 0–180 days, with abstinence rates of 48.8% in KBA vs. 37.9% in HA.

An overall analysis, using logistic regression only, on 7-day point prevalence between 0 and 180 days for the effect of the type of HRS revealed that belonging to the HA group led to a lower 7D-PP of abstinence (OR = 0.364; KBA: 54% vs. HA: 30.2%, $p = 0.023$). The difference of this analysis with the previous analysis of 7D-PP averaged across 0–180 days is that, in this analysis, only each participant's last available abstinence report over the whole time span of 0–180 days was considered, and not for each time interval. In the previous analysis, all participants' intermediate last available abstinence reports were averaged in determining whether there was a difference between the HRS groups.

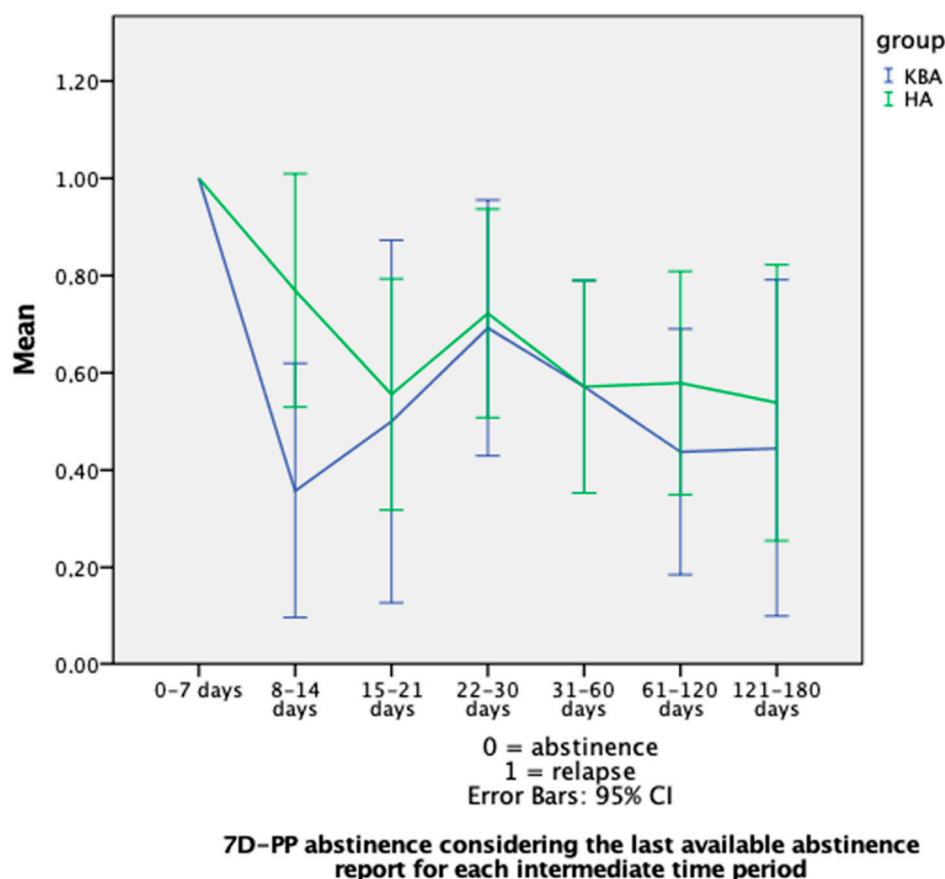


Figure 7. Evolution of the 7D-PP abstinence (based on the last abstinence report) (a lower mean in the figure means more abstinence).

The in-depth analysis of the impact of completing the extended profile for the two groups showed a marginally significant interaction between app group and completing the extended profile ($p = 0.065$). In the HA group, the completers had a lower probability of being 7 days abstinent ($OR = 0.566$), whereas in the KBA group the completers had a higher probability of being 7 days abstinent ($OR = 1.872$), but in none of these groups were these differences significant ($p = 0.209$ and $p = 0.175$ respectively).

Also, having completed the extended profile, was a significant predictor of dropout ($p = 0.002$), such that for both groups those who completed the extended profile, had a lower probability to drop out.

7D-PP Abstinence: Sensitivity Analysis under a Pessimistic Scenario

The standard logistic regression analysis showed no interaction between group and time period ($p = 0.607$). Also, no main effect of HRS group on 7-day point prevalence was found (KBA: 3.4% vs. HA: 3.3%, $p = 0.847$). So, there was no significant difference between the two groups in terms of 7-day point prevalence under a pessimistic scenario. Also, the logistic regression analysis for the 0–180 days' time period, which considered only the last available cessation report, showed no significant effect on the probability of 7-day cessation ($p = 0.392$).

The results of the main effects for each measure comparing both systems are summarized in Tables 3–5.

Table 3. Comparison between the two HRS groups of their results for main effects on the appreciation variable for non-dropouts over the 0–180 days' time period.

Appreciation Variable	Regression Coefficient of HRS (Coding: KBA = 0, HA = 1)	Test Statistic	<i>p</i> -Value	Odds Ratio	95% Confidence Interval for Effect Size
Message ratings	0.818	F(1129) = 0.869	0.353	2.265	(0.399, 12.85)

Table 4. Comparison between the two HRS groups of their results for main effects on engagement variables for non-dropouts over the time period of 0–180 days.

Engagement Variable	Regression Coefficient of HRS (Coding: KBA = 0, HA = 1)	Test Statistic	<i>p</i> -Value	Incidence Rate Ratio	95% Confidence Interval for Effect Size
Number of rated messages	−0.036	F(1233) = 0.021	0.884	0.965	(0.595, 1.565)
Number of active days	−0.258	F(1440) = 4.112	0.043	0.773	(0.602, 0.992)
Number of abstinence reports	−0.403	F(1421) = 11.702	0.001	0.710	(0.530, 0.843)
Number of quitting attempts	0.105	$\chi^2(1) = 1.190$	0.275	1.111	(0.920, 1.340)

Table 5. Comparison between the two HRS groups of their results for main effects on smoking cessation variables for non-dropouts over the time period of 0–180 days.

Smoking Cessation Variable	Regression Coefficient of HRS (Coding: KBA = 0, HA = 1)	Test Statistic	<i>p</i> -Value	Odds Ratio	95% Confidence Interval for Effect Size
7D-PP reports: available data (on last report available in each time interval)	−0.474	$\chi^2(1) = 2.383$	0.123	0.623	(0.341, 1.138)
7D-PP reports: available data (on last report available from 0–180 days period)	−1.010	$\chi^2(1) = 5.162$	0.023	0.364	(0.151, 0.880)
7D-PP reports: pessimistic scenario (on last report available in each time interval)	−0.044	$\chi^2(1) = 0.037$	0.847	0.957	(0.610, 1.501)
7D-PP reports: pessimistic scenario (on last report available from 0–180 days period)	−0.301	$\chi^2(1) = 0.732$	0.392	0.740	(0.371, 1.478)

4. Discussion

4.1. Main Findings

The aim of this study was to compare two different health recommender systems for smoking cessation. The first system used a knowledge-based (KBA) algorithm, whereas the second used a hybrid algorithm (HA), employing KBA and demographic filtering. Effects were studied concerning participants' message appreciation, engagement with the system, and self-reported smoking cessation outcomes.

4.1.1. Message Appreciation

Message appreciation was almost always higher across the intermediate time periods for the HA group, as can be seen in the message appreciation evolution chart (see Figure 3). However, this difference was neither significantly different for the intermediate time periods nor averaged across the whole 6-month period. This result suggests that, in our study, although the demographic filtering step in the HA was delivering slightly more relevant recommendations, it did not suffice to be statistically significant compared to a random selection of the filtered messages resulting from the knowledge-based step. This lack of statistical significance contradicts the potential advantage of collaborative filtering that several previous studies showed [94,95], where messages had higher relevance after going through that step, leading to more and better ratings, which would reflect higher message appreciation [96,97]. In our case, the lack of statistical significance could be related to the cold-start problem [98]. We estimated, during the study design phase, that we would recruit a higher number of participants, and those participants would provide more ratings, quickly reducing the impact of the cold-start problem. Another explanation for the lack of statistical significance is the limited sample size: 843 ratings provided by 57 participants, of which 30 were in the HA group. A higher sample size would yield more power to show that the apparent trend of higher appreciation and number of rated messages in the HA is, indeed, due to the demographic filtering step [99,100].

4.1.2. Engagement

Concerning engagement, we found mixed results. For the number of rated messages, similar to message appreciation, we had a higher number of rated messages for all the time intervals in the HA group, but this was not statistically significant, and the same rationale elaborated for message appreciation is applicable to this measure. Finding no statistical difference in the number of rated messages than for message appreciation can be explained by the same reasons (cold start, and sample size). Lack of power, due to small sample size, may especially be an issue here, as, for the engagement metrics, the coefficient of variation in the large majority of cases is larger than 1, reflecting large variability of the metrics in the sample. In addition, as we did not find statistically significant differences in message appreciation previously, which was the metric related to some predecessors of behavioral intentions of use [101], it follows that no statistical difference on the number of rated messages was found either.

Yet, we found differences in other types of engagement metrics: number of active days and number of abstinence reports, both being higher in the KBA group. These were unexpected results because we had expected, at least, the same level of engagement between KBA and HA groups, since the HA added an extra filtering layer, aiming to provide motivational messages tailored to the preference of the participant rather than at random as in the KBA group. Therefore, in the worst-case scenario, where the demographic filtering step was not working, message appreciation was to be expected to be similar but not worse, and more in-depth studies may be needed to further understand this finding. One explanation could be that the factor which made participants engage with the mobile app in terms of active days, and submit abstinence reports, was not related to the algorithms, but with other variables that we did not consider in our study and which were more favorable for the KBA group. For example, perceived trust in the app could be one of these factors, as it played a relevant role in other health apps before [102,103]. This explanation is also in line with the study by Dovaliene et al. (2016) [104], who showed that user satisfaction was not relevant for mobile app engagement, which is what we were aiming to maximize through higher message appreciation, as the more a participant appreciates messages, the more he or she may be satisfied with the app. However, as the mobile app interface was the same for both groups, the perceived trust, or any other non-considered variable, was more prevalent in those messages sent by the KBA. In the case of trust in the recommendations, there are approaches, such as the application of deep-learning, that could be applied to enhance trust [105].

Concerning interaction effects, we found that employed participants in the KBA group had a higher number of active days in the last two time periods. A potential explanation for this result could be found in other participant variables. Although we did not measure income and educational levels, they could explain a higher usage of a cessation app in terms of socioeconomic status. Socioeconomic status is a known predictor for smoking prevalence and knowledge of smoking effects [106,107] (higher educational level relates to higher knowledge), and could also be linked to mobile app engagement, as has been shown for other health-related app studies [108–110].

Another interaction effect revealed that completers of the extended profile in the HA group made more quitting attempts. This suggested that, when the HA was provided with more information to personalize the recommendations, it may have influenced determination to make a quitting attempt after a relapse. The evidence of the rationale for this effect would be more solid if we also found similar effects on message appreciation for the completion of the extended profile in the HA group, because participants pleased with received messages would be willing to re-engage with the app after a relapse. However, we previously identified the non-significant difference of having completed the extended profile in the appreciation metric. Therefore, although the effect on quitting attempts was significant, it will need to be further explored in future studies to reinforce its evidence.

When testing different approaches to calculate similarity, varying the elements used to compute the similarity also varied the actual outcomes [111,112]. In our case, during the algorithm design phase, we did not know what the optimal number and type of variables to use were, as each dataset would have a different optimal point [113]. It was unclear which are the most important ones and how to deal with dissimilarities in information provided. Ultimately, to calculate the similarity we decided to use all the available information of the 60 questions defining the user profile. However, future studies should carefully reassess what variables may better define user similarities. A lower number of variables may improve user experience because they will complete shorter, but more meaningful, questionnaires. Also, we did not know how the number of user variables for computing similarity impacted on other metrics, such as smoking abstinence. Increasing the number of questions might begin reducing effective similarity among neighbors. This can happen if variables do not really impact smoking and diminish the effect of those which do. Pu et al. (2012) [27] recommended minimizing preference elicitation in profile initialization. Yet, as extended profiles may improve smoking cessation attempts, further experimental research is needed to assess which core questions are needed. We recommend future researchers willing to use demographic filtering in digital health interventions explore approaches that maximize the number of participants completing user profiles. The trade-off between a mandatory questionnaire and total freedom to complete it should be considered to avoid participants having a poor user experience, but also retain the potential engagement benefits found in this study. Similarly, regarding demographic user profile variables, users with insufficient ratings may not be getting good enough recommendations, as they were also part of the similarity calculation in our case. This problem could be enhanced in the future by using implicit ratings, as proposed in the study by Ahmaidan et al. [114].

4.1.3. Smoking Cessation

We found that the HA algorithm was unable to provide significantly better smoking cessation outcomes than the KBA in the 7D-PP analysis averaging the intermediate time period abstinence report results. The abstinence measurements averaged across time intervals can be considered an approximation of continuous abstinence, which was a relevant measure included in the Russell 2.0 standard for smoking cessation outcomes [93]. However, the lack of abstinence reports at specific time points (e.g., 6 months) reduces the reliability of this approximation, as we may be considering the continuous prevalence of participants based on only one report made in the first month, for instance.

The analysis for 7D-PP considering the last abstinence report of each participant as the value for the whole study showed that the KBA algorithm performed better. This

7D-PP can detect delayed quitting, and was also included in the Russell 2.0 standard as a relevant smoking outcome measure [92]. However, as our trial was done under real-world conditions, where we could not consistently collect abstinence reports at specific time points, it was the decision of the participant as to whether to voluntarily submit an abstinence report or not. This analysis implicitly assumes that the participants remained within their last reported abstinence status until the end of the study. This assumption may have introduced bias, as we cannot know the actual abstinence status of participants at the end of the study if they did not report it.

To further examine the sensitivity of these results, we also considered a pessimistic scenario analysis in which participants who dropped out were considered as smokers (penalized imputation). In this case, we found no differences between groups. This result differs from the corresponding analysis on available data results because in both groups there were participants who never sent an abstinence report and were assumed to be relapses in this analysis, hence reducing the differences between groups. Although some authors showed that penalized imputation does not necessarily lead to less biased effect, estimates, or more conservative effect estimates than complete case analysis [115], suggested that it may be too overly critical [116]. Considering all the previous measures, it seems that the overall impact on smoking cessation outcomes for the participants was not significant between groups.

In addition, it is conceivable that certain messages could result in effects different to those intended, and thus lower the efficacy of the HA system. For example, for a participant called John, one of the messages could read “Hi John. Did you know that as you are no longer a smoker, you are no longer part of a chain that favors child exploitation? Yes, you’ve read right. In countries like Pakistan, USA, and Indonesia child labor is used in the tobacco industry! Children and teenagers are exposed to toxic substances and hard work conditions. Stay smoke-free for their sake, and for yours!”. This message, which was intended to provide knowledge, may not be positively appreciated because, even phrased in positive terms, thinking about the fact of having contributed to child labor is unpleasant. However, the message content might cast a deep impact on the participants, leading to non-perceived higher abstinence motivation. Therefore, if users did not like the message, the HA would be unlikely to send it again in favor of other better rated messages, whilst the random selection done by the KBA algorithm would give the same probability of sending the message again to a new user. Unfortunately, in this study we could not perform traceability back to what specific message was rated with what score after the study. Consequently, we cannot validate whether this type of situation happened to be the reason for the unexpected engagement results.

Although the findings of the 7D-PP abstinence, considering only the last abstinence report available, contradict our initial assumptions, the pessimistic scenario and 7D-PP abstinence, considering the averaged abstinence reports, were in line with a study by Westmaas et al. (2018) [117], who compared two tailoring levels in e-mails to support smoking cessation, and found no differences between the basic and advanced versions. This may mean that having a more complex system for tailoring smoking cessation support does not necessarily provide better abstinence results.

To compare our abstinence results to other similar previous studies, a recent review [118] examined SmartQuit [119] and SmokeFree28 [120], which managed to achieve abstinence rates of 13% at 2-month and 21% at 28-day follow-ups, respectively. Mobile-based interventions for smoking cessation, such as Clickotine [121], reported a 7-day abstinence rate of 45.2% and a 30-day abstinence rate of 26.2% in an 8-week study following an intention-to-treat analysis. In more traditional computer-tailored interventions that had follow-up assessments at 6 months, as in our study, abstinence rates were up to 18.3% (10.2% intention to treat) [122] and 20.4% (8.5% following intention to treat) [123]. These results are similar to the results of our study (54.1% in the KBA group and 30.2% in the HA group, following analysis on available data for 7-day point prevalence abstinence,

considering the last available abstinence report for each intermediate time period, and 8.4% in the KBA group and 11% in the HA group in the pessimistic scenario).

4.2. Additional Considerations

Across all metrics, we identified that the completion of the extended profile was associated with less dropouts. As this was consistent for both the HA and KBA groups (in which completion had no effect on the algorithm), we propose that it could be that the personality of participants who filled out the extended questionnaire is such that they are more curious, or willing, to interact with the system. This was consistent with the work of Karumur et al. (2018), who explored how user personality can influence user engagement and activity in recommender systems, based on a collaborative filtering approach [124].

The external validity of mobile apps for health behavior change has been previously criticized and it is currently being studied in other domains, such as physical activity [125]. We conducted this trial under real-world conditions, where we did not reward participants with money, nor stimulate them to complete the abstinence reports; additionally, researchers did not recruit, or follow-up, participants in any way. Hence, the present results of real-world effectiveness, in terms of the effects of such an HRS, increase the external validity of our study [126], compared to efficacy results in previously mentioned studies. Thus, the results of this trial study have to be understood within this real-world context and might not be directly comparable to other studies, in which recruitment and follow-up methods may have positively impacted the commitment of participants.

Regarding the recommender system algorithms used in this study, we could not compare their theoretical performances with others found in the literature before running the study because of their differences in design, that makes them incompatible with existing datasets, which were not focused on health behavior. For instance, in other contexts (e.g., movie recommendations) recommender systems could use existing datasets as benchmarks for performance assessment, such as MovieLens [127]. Hence, commonly used metrics for recommendation systems, such as the root mean square error (RMSE), mean average error (MAE), normalized discount cumulative gain (NDCG), and precision could not be applied to assess the algorithms before the study, as we lacked an initially rated, compatible dataset.

This study focused on metrics which assessed the impacts of the recommendations on participants' behavior, as Sahoo et al. (2019) [128] proposed for HRSs, not on the technical performance of the recommender system itself. Schäfer et al. (2017) [25] also concluded that HRSs needed multidimensional user satisfaction measures, which covered message appreciation and engagement metrics. Further, message appreciation can be seen as a proxy for Mean Average Precision at 1. This sets our study in line with previous experimental studies for HRS performance analysis based on hits and total numbers of recommendations or users. This was the case in a study by Rivero-Rodriguez et al. (2013) [129], which followed a similar approach to assess their HRS, using the hit rate (no. of hits/no. of users), as a performance metric. Another example can be found in a study by Bocanegra et al. (2017) [130], in which they used precision to assess their recommender system. Despite applying similar approaches, the direct comparison of performance metrics and reflection on the conclusions generated in other studies are still limited in our case because they are totally different study designs and research questions. A recent scoping review [28] backs the relative scarcity of studies applying HRSs in the health domain and their diversity in therapeutic areas and reported outcomes.

The most similar study we found was the SoloMo study which presented results of another hybrid recommender system for smoking cessation [131]. That algorithm was tested in a clinical context where patients were followed up by healthcare professionals for one year [63]. Patients referred to the smoking cessation unit from other specialized care units of the hospital, were invited to the study. The precision of that system (which was directly related to appreciation, as previously mentioned) achieved a high score (which would mean high appreciation), with a minimum value of 0.96 over a total maximum of one. There were key differences with our present study, which explain this difference.

First, the end of their messages included the name of the patient's doctor, and this may have increased the perceived quality of the messages, leading to high precision (and we can therefore assume, also appreciation). Second, patients may have thought that their doctors were going to check their ratings and wanted to please them with higher scores. Third, the potential rating options were only positive, negative, or neutral, as opposed to the five-star scale included in our study, which provided greater rating granularity. Fourth, the participants enrolled in SoLoMo were referred from a specialized care unit, which eventually contributed to enhancing their motivation and appreciation of the system, potentially driven by a positive framing bias effect. Despite the apparently better results in the SoLoMo study [63], a smaller range in the rating values (only three options compared to the five we had) limits the opportunity of HRSs to learn from users' opinions in digital interventions for smoking cessation.

Further, in a recent scoping review [28], several gaps in the reviewed HRS studies were found. Our study covered many of them, including: (1) reporting the results of our study based on a large user cohort size ($n = 371$), compared to previous ones found in the literature; (2) using an HRS which was grounded in a behavioral change theory (the I-Change model [29]), recommending messages with behavioral change techniques (following the guide by Abraham et al. (2011) [84]; (3) using advanced profile adaptation and (4) having a clear explicit feedback system (in the HA group). Table 6 presents the final classification of this health recommender system following the proposed taxonomy by Hors-Fraile et al. [28].

Table 6. Classification for both health recommender systems' classification.

Taxonomy aspect	Element	KBA	HA
Domain	Therapeutic area	Smoking cessation	Smoking cessation
	Target population	Current smokers willing to quit, speaking Mandarin Chinese	Current smokers willing to quit, speaking Mandarin Chinese
	Type of recommendation (items)	Text messages	Text messages
	Device interface	Android and iPhone mobile app	Android and iPhone mobile app
	Tailoring	Yes	Yes
	Country	Taiwan	Taiwan
Methodology and procedures	Used metrics to assess performance	Self-reported smoking abstinence (7-day point prevalence), message appreciation, engagement (number of quitting attempts, number of rated messages, number of abstinence reports, number of days active in the app)	Self-reported smoking abstinence (7-day point prevalence), message appreciation, engagement (number of quitting attempts, number of rated messages, number of abstinence reports, number of days active in the app)
	Number of test users	181	190
	Effectiveness on patients	Yes	Yes
	Success percentage	7-day point prevalence: 3.4%- of 48.8% Message appreciation: 3.835/5 Mean number of: quitting attempts = 1.43, rated message = 3.59, abstinence reports = 1.1, days active in the app = 2.74	7-day point prevalence: 3.3–37.9% Message appreciation: 4.185/5 Mean number of: quitting attempts = 1.71, rated messages = 4.31, abstinence reports = 0.7, days active in the app: = 1.99
	Duration of total intervention	6 months	6 months
	Number of sessions	Minimum: 1 (Maximum: 50 (estimated))	Minimum: 1 (Maximum: 50 (estimated))
	Electronic Health Record connection	No	No
	Cost-effectiveness	Not analyzed	Not analyzed

Table 6. Cont.

Taxonomy aspect	Element	KBA	HA
Health promotion theoretical factors and behavior change theories	Attitude	Yes	Yes
	Social influence	Yes	Yes
	Self-efficacy	Yes	Yes
	Action and Coping planning	Yes	Yes
	Supporting identity change	Yes	Yes
	Rewarding	Yes	Yes
	Advising on changing routines	Yes	Yes
	Advising on coping	Yes	Yes
	Advising on medication use	No	No
Technical aspects	Recommendation interface	Top-N (N = 1)	Top-N (N = 1)
	Recommendation technology	Attribute-based recommendations	Attribute-based recommendations + People-to-People correlation
	Finding recommendations	Selection options + request recommendation list	Selection options + request recommendation list
	Initial profile representation techniques	Manual	Manual
	Profile presentation technique	User-item feature matrix	User-item feature matrix, history-based model, user-item rating matrix, demographic features
	Profile learning technique	Not necessary	Not necessary
	Relevant feedback	Not necessary	Explicit feedback
	Profile adaptation technique	Manual	Manual
	Information filtering method	Knowledge-based	Hybrid: knowledge-based + demographic filtering
	User-profile item matching technique	Not applicable	Nearest neighbor (Pearson), Find similar users

4.3. Limitations

Despite strengths of this study, such as comparing two different HRSs and their participants' age, gender, nicotine dependence, employment status, motivation to quit, or completion of the extended profile questionnaire, our study was subject to some limitations.

1. The HRSs considered all users' feedback for computing recommendations. This implies that the feedback provided by one group affected the generation of recommendations for the other. This design decision was taken to reduce the cold start effect.
2. Between 22 May and 6 June 2018, and between 1 and 6 August 2018, there was a server service interruption that prevented users from registering the app and receiving messages.
3. We could not verify the smoking status self-reports. Self-reports may provide a valid estimation of cessation rates as they were used in several previous studies [132]. The Society for Research on Nicotine and Tobacco Subcommittee on Biochemical Verification considered the use of biochemical validation unnecessary in studies with limited face-to-face contact [133]. However, the use of bogus-pipeline procedures [134], or some biochemical verification methods, would have improved the validity of the smoking status reports [135]. Also, the pessimistic scenario analysis we conducted, intending to follow a conservative approach, may not have accurately reflected the actual behavior of the participants. Also, this study may have suffered from errors derived from some users' difficulties in accurately recalling their behaviors, as Ramo et al. previously reflected on for an anonymous survey about smoking behaviors [136].
4. We considered the last status for the abstinence report as the value for each time interval. This way of measuring smoking cessation results hampered direct comparisons with previous studies.

5. In this effectiveness study, smokers could report by self-chosen times, resulting in the fact that we could not assess all data for all participants at one specific time (e.g., smoking cessation status after 1 month).
6. It is conceivable that specific subgroup effects could have occurred in our analyses, requiring more sophisticated models with more two-way (or even three-way) interactions to explain our results. However, due to the sparsity of the collected data, these more sophisticated models could not be applied.
7. User-experience metrics, such as perceived quality and satisfaction, which are commonly evaluated nowadays in the field of recommender systems [27,137], were not included in this study.
8. Persuasion profile meta-features to determine what recommendation style (e.g., authority shown in the message, the reflected consensus stated in the message, the message sender liking perception, etc.) [138,139] would persuade participants the most, were not considered. Such types of meta-features could have added extra personalization power to the HRS without needing the participants to complete additional questions in their user profile.

4.4. Recommendations

Based on our results, we recommend future studies to keep exploring the usage of different types of HRS to support smoking abstinence, as the results of both algorithms improved the unassisted cessation success rates for six months which is around 3–5% [140,141], and were above some nicotine-replacement therapy rates, whose six-month abstinence rate is around 7% [142,143]. More research is needed to explore relationships between message appreciation, engagement, and health outcomes. When using HRS with collaborative filtering, new means to determine recommendations' relevance, other than participants' appreciation, should be considered. For instance, including the achieved health outcomes as a complementary rating (e.g., asking how useful previously sent messages were when rating a message about abstinence status). In this way, messages that are well appreciated but do not contribute towards supporting abstinence would not be recommended in the future, and messages which are both contributing to support abstinence and highly appreciated would be prioritized to be sent by the HRS.

As we did not find differences between HRS for gender, age, nicotine dependence level, and motivation to quit subgroups, we encourage future research to be conducted considering other variables, such as trust and socioeconomic status, which may help better understanding of smokers' behaviors. Still, gender, age, nicotine dependence level, and motivation to quit could be good meta-features for HRS similarity computation and we suggest keeping them as part of the HRS and message design processes.

In addition, future research should consider larger sample sizes with more than six months of follow-up time, as results between the KBA and HA, in terms of message appreciation and number of rated messages, suggest that these differences could become significant. Alternatively, this may be compensated for by increasing the frequency of the sent messages, as the HRS would have more information to process new recommendations faster, and/or have larger sample sizes to be better able to detect such differences. To avoid overloading participants with too many messages, which may be bothersome and negatively impact their user experience, we consider that a suitable solution could be to offer 'on-demand' messages. This larger sample would facilitate applying more complex statistical models, which may help us explain some results about which we could only guess in this study. Also, we suggest pro-actively persuading users to complete their profiles if they do not do so voluntarily during the enrollment phase, to maximize the impact of collaborative filtering, giving participants more probability to make a new quitting attempt in case of relapse. This could be achieved, for instance, by making the digital solution ask the participants to complete one or two unanswered questions of their user profile every day. It would yield a low entry barrier to start using the solution, whilst it would allow computation recommendations with increasing user profile information over time.

5. Conclusions

The first goal of this study was to compare the two presented HRSs. We found the KBA-led participants had more active days on the mobile app, completed more abstinence reports, and had better 7D-PP abstinence results, despite being non-significant when averaged across time, and also being non-significant in the pessimistic scenario. The additional step of demographic filtering in the HA only improved the number of quitting attempts. However, the HA group seemed to rate a higher number of messages and gave better ratings to the messages, based on trends shown on the evolution line charts; yet both measures did not statistically differ between groups. The second goal was to identify potential subgroup differences, and we found that participants who completed their extended profiles were more likely to stay in the study and, among the employed, those who were in the KBA group had higher engagement, in terms of active days, than those in the HA group. No other differences were found for any other subgroups (gender, age, nicotine dependence level, motivation to quit).

Our findings provide insights into the usage of health recommender systems in a real-world setting. We conclude that collaborative intelligence provided mixed results, some of them unexpected, and more research is needed to fully take advantage of it in the context of smoking cessation support. However, this study showed a promising future for health recommender systems. Combining behavioral change techniques and models in recommender systems, even with simple algorithms, can lead to higher smoking cessation rates than unsupported quitting and also more than some nicotine replacement therapies.

Author Contributions: Conceptualization, S.H.-F., L.F.-L., F.S. and H.d.V.; methodology, S.H.-F., M.J.J.M.C., F.S. and H.d.V.; software, S.H.-F.; validation, M.J.J.M.C., S.H.-F. and H.d.V.; formal analysis, M.J.J.M.C. and S.H.-F.; investigation, S.H.-F., S.M. and S.S.-A.; resources, S.M. and S.S.-A.; data curation, S.H.-F.; writing—original draft preparation, S.H.-F., M.J.J.M.C., S.M. and F.J.N.-B.; writing—review and editing F.S., L.F.-L., S.H.-F. and H.d.V.; visualization, S.H.-F. and M.J.J.M.C.; supervision, S.H.-F.; project administration S.H.-F. and S.S.-A.; funding acquisition, L.F.-L., S.S.-A. and F.J.N.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme under Grant 681120, and in part by the Taipei Medical University was funded by the Ministry of Science and Technology, Taiwan (grant numbers 106-2923-E-038-001-MY2, 108-2221-E-038-013 and 110-2923-E-038-001-MY3); and the Taipei Medical University, Taiwan (grant number 108-3805-009-110 and 109-3800-020-400).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Taipei Medical University-Joint Institutional Review Board (TMU-JIRB) (protocol code N201702069, 17 March 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: We want to thank Panagiotis Bamidis from Aristotle University of Thessaloniki for his commitment and leadership coordinating the SmokeFreeBrain project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brug, J.; Glanz, K.; Van Assema, P.; Kok, G.; Van Breukelen, G.J. The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. *Health Educ. Behav.* **1998**, *25*, 517–531. [[CrossRef](#)] [[PubMed](#)]
2. Strecher, V.J. Computer-tailored smoking cessation materials: A review and discussion. *Patient Educ. Couns.* **1999**, *36*, 107–117. [[CrossRef](#)]
3. De Vries, H.; Brug, J. Computer-tailored interventions motivating people to adopt health promoting behaviours: Introduction to a new approach. *Patient Educ. Couns.* **1999**, *36*, 99–195. [[PubMed](#)]
4. Brug, J.; Oenema, A.; Campbell, M. Past, present, and future of computer-tailored nutrition education. *Am. J. Clin. Nutr.* **2003**, *77*, 1028S–1034S. [[CrossRef](#)] [[PubMed](#)]

5. Dijkstra, A. The psychology of tailoring-ingredients in computer-tailored persuasion. *Soc. Personal. Psychol. Compass* **2008**, *2*, 765–784. [\[CrossRef\]](#)
6. Enwald, H.P.K.; Huotari, M.-L.A. Preventing the obesity epidemic by second generation tailored health communication: An interdisciplinary review. *J. Med. Internet Res.* **2010**, *12*, e24. [\[CrossRef\]](#)
7. Ghalibaf, A.K.; Nazari, E.; Gholian-Aval, M.; Tara, M. Comprehensive overview of computer-based health information tailoring: A systematic scoping review. *BMJ Open* **2019**, *9*, e021022. [\[CrossRef\]](#)
8. Kreuter, M.; Farrell, D.; Olevitch, L.; Brennan, L. *Tailored Health Messages: Customizing Communication with Computer Technology*; Lawrence Erlbaum: Mahwah, NJ, USA, 1999.
9. Kroeze, W.; Werkman, A.; Brug, J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. *Ann. Behav. Med.* **2006**, *31*, 205–223. [\[CrossRef\]](#)
10. Neville, L.M.; O'Hara, B.; Milat, A. Computer-tailored physical activity behavior change interventions targeting adults: A systematic review. *Int. J. Behav. Nutr. Phys. Act.* **2009**, *6*, 30. [\[CrossRef\]](#)
11. Krebs, P.; Prochaska, J.O.; Rossi, J.S. A meta-analysis of computer-tailored interventions for health behavior change. *Prev. Med.* **2010**, *51*, 214–221. [\[CrossRef\]](#)
12. Cheung, K.L.; Wijnen, B.; de Vries, H. A review of the theoretical basis, effects, and cost effectiveness of online smoking cessation interventions in the Netherlands: A mixed-methods approach. *J. Med. Internet Res.* **2017**, *19*, e230. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Dijkstra, A.; De Vries, H.; Roijackers, J.; van Breukelen, G. Tailoring information to enhance quitting in smokers with low motivation to quit: Three basic efficacy questions. *Health Psychol.* **1998**, *17*, 513. [\[CrossRef\]](#)
14. Glasgow, R.; Christiansen, S.M.; Kurz, D.; King, D.K.; Woolley, T.; Faber, A.J.; Estabrooks, P.A.; Strycker, L.; Toobert, D.; Dickman, J. Engagement in a diabetes self-management website: Usage patterns and generalizability of program use. *J. Med. Internet Res.* **2011**, *13*, e9. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Schneider, F.; van Osch, L.; Schulz, D.N.; Kremers, S.P.; de Vries, H. The influence of user characteristics and a periodic email prompt on exposure to an internet-delivered computer-tailored lifestyle program. *J. Med. Internet Res.* **2012**, *14*, e40. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Bennett, G.G.; Glasgow, R.E. The delivery of public health interventions via the Internet: Actualizing their potential. *Annu. Rev. Public Health* **2009**, *30*, 273–292. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Kreps, G.L.; Neuhauser, L. New directions in eHealth communication: Opportunities and challenges. *Patient Educ. Couns.* **2010**, *78*, 329–336. [\[CrossRef\]](#)
18. Sadasivam, R.S.; Cutrona, S.L.; Kinney, R.L.; Marlin, B.M.; Mazor, K.M.; Lemon, S.C.; Houston, T.K. Collective-intelligence recommender systems: Advancing computer tailoring for health behavior change into the 21st century. *J. Med. Internet Res.* **2016**, *18*, e42. [\[CrossRef\]](#)
19. Kaptein, M.; Eckles, D. Selecting effective means to any end: Futures and ethics of persuasion profiling. In Proceedings of the International Conference on Persuasive Technology, 2010, Copenhagen, Denmark, 7–10 June 2010; pp. 82–93.
20. Noar, S.M.; Benac, C.N.; Harris, M.S. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol. Bull.* **2007**, *133*, 673. [\[CrossRef\]](#)
21. Lederman, R.; Smith, S.; Dreyfus, S.; Alzougool, B. Tailoring patient information to encourage patient engagement. *Int. J. Netw. Virtual Organ.* **2016**, *16*, 191–203. [\[CrossRef\]](#)
22. Strecher, V.J.; McClure, J.B.; Alexander, G.L.; Chakraborty, B.; Nair, V.N.; Konkel, J.M.; Greene, S.M.; Collins, L.M.; Carlier, C.C.; Wiese, C.J. Web-based smoking-cessation programs: Results of a randomized trial. *Am. J. Prev. Med.* **2008**, *34*, 373–381. [\[CrossRef\]](#)
23. Spittaels, H.; De Bourdeaudhuij, I.; Brug, J.; Vandelanotte, C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. *Health Educ. Res.* **2007**, *22*, 385–396. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Morrison, L.; Moss-Morris, R.; Michie, S.; Yardley, L. Optimizing engagement with Internet-based health behaviour change interventions: Comparison of self-assessment with and without tailored feedback using a mixed methods approach. *Br. J. Health Psychol.* **2014**, *19*, 839–855. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Schäfer, H.; Hors-Fraile, S.; Karumur, R.P.; Calero Valdez, A.; Said, A.; Torkamaan, H.; Ulmer, T.; Trattner, C. Towards health (aware) recommender systems. In Proceedings of the 2017 International Conference on Digital Health, London, UK, 2–5 July 2017; pp. 157–161.
26. Hors-Fraile, S.; Malwade, S.; Luna-Perejon, F.; Amaya, C.; Civit, A.; Schneider, F.; Bamidis, P.; Syed-Abdul, S.; Li, Y.-C.; De Vries, H. Opening the Black Box: Explaining the Process of Basing a Health Recommender System on the I-Change Behavioral Change Model. *IEEE Access* **2019**, *7*, 176525–176540. [\[CrossRef\]](#)
27. Pu, P.; Chen, L.; Hu, R. Evaluating recommender systems from the user's perspective: Survey of the state of the art. *User Modeling User-Adapt. Interact.* **2012**, *22*, 317–355. [\[CrossRef\]](#)
28. Hors-Fraile, S.; Rivera-Romero, O.; Schneider, F.; Fernandez-Luque, L.; Luna-Perejon, F.; Civit-Balcells, A.; de Vries, H. Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: A scoping review. *Int. J. Med. Inform.* **2018**, *114*, 143–155. [\[CrossRef\]](#)
29. De Vries, H. An integrated approach for understanding health behavior; the I-change model as an example. *Psychol. Behav. Sci. Int. J.* **2017**, *2*, 555–585. [\[CrossRef\]](#)
30. Gunawardana, A.; Shani, G. Evaluating recommender systems. In *Recommender Systems Handbook*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 265–308.

31. Said, A.; Tikk, D.; Stumpf, K.; Shi, Y.; Larson, M.A.; Cremonesi, P. Recommender Systems Evaluation: A 3D Benchmark. In Proceedings of the RUE@ RecSys, Dublin, Ireland, 9–13 September 2012; pp. 21–23.
32. Herlocker, J.L.; Konstan, J.A.; Terveen, L.G.; Riedl, J.T. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst. (TOIS)* **2004**, *22*, 5–53. [\[CrossRef\]](#)
33. Valdez, A.C.; Ziefle, M.; Verbert, K.; Felfernig, A.; Holzinger, A. Recommender systems for health informatics: State-of-the-art and future perspectives. In *Machine Learning for Health Informatics*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 391–414.
34. Jander, A.; Crutzen, R.; Mercken, L.; Candel, M.; de Vries, H. Effects of a web-based computer-tailored game to reduce binge drinking among Dutch adolescents: A cluster randomized controlled trial. *J. Med. Internet Res.* **2016**, *18*, e29. [\[CrossRef\]](#)
35. Bingham, C.R.; Barretto, A.I.; Walton, M.A.; Bryant, C.M.; Shope, J.T.; Raghunathan, T.E. Efficacy of a web-based, tailored, alcohol prevention/intervention program for college students: Initial findings. *J. Am. Coll. Health* **2010**, *58*, 349–356. [\[CrossRef\]](#)
36. Martinez-Montilla, J.M.; Mercken, L.; de Vries, H.; Candel, M.; Lima-Rodríguez, J.S.; Lima-Serrano, M. A web-based, computer-tailored intervention to reduce alcohol consumption and binge drinking among Spanish adolescents: Cluster randomized controlled trial. *J. Med. Internet Res.* **2020**, *22*, e15438. [\[CrossRef\]](#)
37. Brug, J.; Steenhuis, I.; van Assema, P.; de Vries, H. The impact of a computer-tailored nutrition intervention. *Prev. Med.* **1996**, *25*, 236–242. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Springvloed, L.; Lechner, L.; de Vries, H.; Candel, M.J.; Oenema, A. Short-and medium-term efficacy of a Web-based computer-tailored nutrition education intervention for adults including cognitive and environmental feedback: Randomized controlled trial. *J. Med. Internet Res.* **2015**, *17*, e23. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Rawl, S.M.; Skinner, C.S.; Perkins, S.M.; Springston, J.; Wang, H.-L.; Russell, K.M.; Tong, Y.; Gebregziabher, N.; Krier, C.; Smith-Howell, E. Computer-delivered tailored intervention improves colon cancer screening knowledge and health beliefs of African-Americans. *Health Educ. Res.* **2012**, *27*, 868–885. [\[CrossRef\]](#) [\[PubMed\]](#)
40. van der Aalst, C.M.; de Koning, H.J.; van den Bergh, K.A.; Willemsen, M.C.; van Klaveren, R.J. The effectiveness of a computer-tailored smoking cessation intervention for participants in lung cancer screening: A randomised controlled trial. *Lung Cancer* **2012**, *76*, 204–210. [\[CrossRef\]](#)
41. Smeets, T.; Brug, J.; de Vries, H. Effects of tailoring health messages on physical activity. *Health Educ. Res.* **2008**, *23*, 402–413. [\[CrossRef\]](#) [\[PubMed\]](#)
42. de Vries, H.; Logister, M.; Krekels, G.; Klaasse, F.; Servranckx, V.; van Osch, L. Internet-based computer tailored feedback on sunscreen use. *J. Med. Internet Res.* **2012**, *14*, e48. [\[CrossRef\]](#)
43. Smit, E.S.; Evers, S.M.; de Vries, H.; Hoving, C. Cost-effectiveness and cost-utility of Internet-based computer tailoring for smoking cessation. *J. Med. Internet Res.* **2013**, *15*, e57. [\[CrossRef\]](#)
44. Smit, E.S.; de Vries, H.; Hoving, C. Effectiveness of a Web-based multiple tailored smoking cessation program: A randomized controlled trial among Dutch adult smokers. *J. Med. Internet Res.* **2012**, *14*, e82. [\[CrossRef\]](#)
45. Stanczyk, N.E.; Crutzen, R.; Bolman, C.; Muris, J.; de Vries, H. Influence of delivery strategy on message-processing mechanisms and future adherence to a Dutch computer-tailored smoking cessation intervention. *J. Med. Internet Res.* **2013**, *15*, e28. [\[CrossRef\]](#)
46. Elfeddali, I.; de Vries, H.; Bolman, C.; Pronk, T.; Wiers, R.W. A randomized controlled trial of Web-based Attentional Bias Modification to help smokers quit. *Health Psychol.* **2016**, *35*, 870. [\[CrossRef\]](#)
47. Etter, J.-F.; Perneger, T.V. Effectiveness of a computer-tailored smoking cessation program: A randomized trial. *Arch. Intern. Med.* **2001**, *161*, 2596–2601. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Starke, A.D.; Trattner, C. Promoting Healthy Food Choices Online: A Case for Multi-List Recommender Systems. In Proceedings of the ACM IUI 2021 Workshops, College Station, TX, USA, 14–17 April 2021.
49. Starke, A.; Asotic, E.; Trattner, C. “Serving Each User”: Supporting Different Eating Goals Through a Multi-List Recommender Interface. In Proceedings of the Fifteenth ACM Conference on Recommender Systems, New York, NY, USA, 27 September–1 October 2021; pp. 124–132.
50. Musto, C.; Trattner, C.; Starke, A.; Semeraro, G. Towards a knowledge-aware food recommender system exploiting holistic user models. In Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization, Genoa, Italy, 12–18 July 2020; pp. 333–337.
51. Pecune, F.; Callebert, L.; Marsella, S. A Recommender System for Healthy and Personalized Recipes Recommendations. In Proceedings of the HealthRecSys@ RecSys, New York, NY, USA, 27 September–1 October 2020; pp. 15–20.
52. Chavan, P.; Thoms, B.; Isaacs, J. A Recommender System for Healthy Food Choices: Building a Hybrid Model for Recipe Recommendations using Big Data Sets; AIS Electronic Library. Available online: https://aisel.aisnet.org/hicss-54/hc/wellness_management/2/ (accessed on 5 January 2021).
53. Elahi, M.; Ge, M.; Ricci, F.; Fernández-Tobías, I.; Berkovsky, S.; Massimo, D. Interaction Design in a Mobile Food Recommender System. In Proceedings of the IntRS@ RecSys, Vienna, Austria, 16–20 September 2015; pp. 49–52.
54. Gómez-del-Río, N.; González-González, C.S.; Toledo-Delgado, P.A.; Muñoz-Cruz, V.; García-Peñalvo, F. Health promotion for childhood obesity: An approach based on self-tracking of data. *Sensors* **2020**, *20*, 3778. [\[CrossRef\]](#)
55. Elswailer, D.; Harvey, M.; Ludwig, B.; Said, A. Bringing the “healthy” into Food Recommenders. In Proceedings of the DMRS, 2015, Bolzano, Italy, 22–23 October 2015; pp. 33–36.
56. Tran, T.N.T.; Atas, M.; Felfernig, A.; Stettinger, M. An overview of recommender systems in the healthy food domain. *J. Intell. Inf. Syst.* **2018**, *50*, 501–526. [\[CrossRef\]](#)

57. Zenun Franco, R. Online recommender system for personalized nutrition advice. In Proceedings of the Eleventh ACM Conference on Recommender Systems, 2017, Como, Italy, 31 August 2017; pp. 411–415.
58. Ribeiro, D.; Machado, J.; Ribeiro, J.; Vasconcelos, M.J.M.; Vieira, E.F.; de Barros, A.C. SousChef: Mobile Meal Recommender System for Older Adults. In Proceedings of the ICT4AgeingWell, Porto, Portugal, 28–29 April 2017; pp. 36–45.
59. Kuanr, M.; Mohapatra, P.; Piri, J. Health recommender system for cervical cancer prognosis in women. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 20–22 July 2021; pp. 673–679.
60. Ammar, N.; Bailey, J.E.; Davis, R.L.; Shaban-Nejad, A. Using a Personal Health Library-Enabled mHealth Recommender System for Self-Management of Diabetes Among Underserved Populations: Use Case for Knowledge Graphs and Linked Data. *JMIR Form. Res.* **2021**, *5*, e24738. [[CrossRef](#)] [[PubMed](#)]
61. Rohani, D.A.; Quemada Lopategui, A.; Tuxen, N.; Faurholt-Jepsen, M.; Kessing, L.V.; Bardram, J.E. MUBS: A personalized recommender system for behavioral activation in mental health. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13.
62. Hors-Fraile, S.; Benjumea, F.J.N.; Hernández, L.C.; Ruiz, F.O.; Fernandez-Luque, L. Design of two combined health recommender systems for tailoring messages in a smoking cessation app. *arXiv* **2016**, arXiv:1608.07192.
63. Carrasco-Hernandez, L.; Jódar-Sánchez, F.; Núñez-Benjumea, F.; Conde, J.M.; González, M.M.; Civit-Balcells, A.; Hors-Fraile, S.; Parra-Calderón, C.L.; Bamidis, P.D.; Ortega-Ruiz, F. A Mobile Health Solution Complementing Psychopharmacology-Supported Smoking Cessation: Randomized Controlled Trial. *JMIR mHealth uHealth* **2020**, *8*, e17530. [[CrossRef](#)]
64. Sadasivam, R.S.; Borglund, E.M.; Adams, R.; Marlin, B.M.; Houston, T.K. Impact of a collective intelligence tailored messaging system on smoking cessation: The Perspect randomized experiment. *J. Med. Internet Res.* **2016**, *18*, e285. [[CrossRef](#)]
65. Chen, J.; Houston, T.K.; Faro, J.M.; Nagawa, C.S.; Orvek, E.A.; Blok, A.C.; Allison, J.J.; Person, S.D.; Smith, B.M.; Sadasivam, R.S. Evaluating the use of a recommender system for selecting optimal messages for smoking cessation: Patterns and effects of user-system engagement. *BMC Public Health* **2021**, *21*, 1749. [[CrossRef](#)]
66. Pincay, J.; Terán, L.; Portmann, E. Health recommender systems: A state-of-the-art review. In Proceedings of the 2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG), Quito, Ecuador, 24–26 April 2019; pp. 47–55.
67. Cheung, K.L.; Durusu, D.; Sui, X.; de Vries, H. How recommender systems could support and enhance computer-tailored digital health programs: A scoping review. *Digital Health* **2019**, *5*, 1–19. [[CrossRef](#)]
68. Su, J.; Guan, Y.; Li, Y.; Chen, W.; Lv, H.; Yan, Y. Do recommender systems function in the health domain: A system review. *arXiv* **2020**, arXiv:2007.13058.
69. Ferretto, L.R.; Cervi, C.R.; de Marchi, A.C.B. Recommender systems in mobile apps for health a systematic review. In Proceedings of the 2017 12th Iberian Conference on Information Systems and Technologies (CISTI), Lisbon, Portugal, 14–17 June 2017; pp. 1–6.
70. Prajapati, J.; Brahmabhatt, K.N. A Systematic Literature Review on Health Recommender Systems. In Proceedings of the International e-Conference on Intelligent Systems and Signal Processing, Hanoi, Vietnam, 19–20 March 2022; pp. 557–572.
71. De Croon, R.; Van Houdt, L.; Htun, N.N.; Štiglic, G.; Abeele, V.V.; Verbert, K. Health Recommender Systems: Systematic Review. *J. Med. Internet Res.* **2021**, *23*, e18035. [[CrossRef](#)] [[PubMed](#)]
72. Hors-Fraile, S.; Malwade, S.; Spachos, D.; Fernandez-Luque, L.; Su, C.-T.; Jeng, W.-L.; Syed-Abdul, S.; Bamidis, P.; Li, Y.-C.J. A recommender system to quit smoking with mobile motivational messages: Study protocol for a randomized controlled trial. *Trials* **2018**, *19*, 618. [[CrossRef](#)] [[PubMed](#)]
73. Bamidis, P.D.; Paraskevopoulos, E.; Konstantinidis, E.; Spachos, D.; Billis, A. Multimodal e-health services for smoking cessation and public health: The smokefreebrain project approach. In Proceedings of the MEDINFO 2017: Precision Healthcare through Informatics: Proceedings of the 16th World Congress on Medical and Health Informatics, Hangzhou, China, 31 January 2018; p. 5.
74. Abroms, L.C.; Whittaker, R.; Free, C.; Van Alstyne, J.M.; Schindler-Ruwisch, J.M. Developing and pretesting a text messaging program for health behavior change: Recommended steps. *JMIR mHealth uHealth* **2015**, *3*, e107. [[CrossRef](#)] [[PubMed](#)]
75. Walsh, R.A.; Sanson-Fisher, R.W. *Encouraging People to Stop Smoking*; World Health Organisation: Geneva, Switzerland, 2001.
76. Panday, S.; Reddy, S.P.; Ruiter, R.A.; Bergström, E.; de Vries, H. Determinants of smoking cessation among adolescents in South Africa. *Health Educ. Res.* **2005**, *20*, 586–599. [[CrossRef](#)] [[PubMed](#)]
77. Hoving, E.; Mudde, A.; De Vries, H. Predictors of smoking relapse in a sample of Dutch adult smokers; the roles of gender and action plans. *Addict. Behav.* **2006**, *31*, 1177–1189. [[CrossRef](#)] [[PubMed](#)]
78. De Vries, H.; Eggers, S.M.; Bolman, C. The role of action planning and plan enactment for smoking cessation. *BMC Public Health* **2013**, *13*, 393. [[CrossRef](#)]
79. Elling, J.M.; Crutzen, R.; Talhout, R.; De Vries, H. Tobacco smoking and smoking cessation in times of COVID-19. *Tob. Prev. Cessat.* **2020**, *6*, 39. [[CrossRef](#)]
80. Kasten, S.; Eggers, S.; de Vries, H. The importance of self-efficacy and action planning for smoking cessation. *Eur. Health Psychol.* **2015**, *17*, 694.
81. Dijkstra, A.; Vries, H.D. Self-efficacy expectations with regard to different tasks in smoking cessation. *Psychol. Health* **2000**, *15*, 501–511. [[CrossRef](#)]
82. Vries, H.D.; Mudde, A.N. Predicting stage transitions for smoking cessation applying the attitude-social influence-efficacy model. *Psychol. Health* **1998**, *13*, 369–385. [[CrossRef](#)]
83. De Vries, H.; Backbier, E.; Kok, G.; Dijkstra, M. The Impact of Social Influences in the Context of Attitude, Self-Efficacy, Intention, and Previous Behavior as Predictors of Smoking Onset. *J. Appl. Soc. Psychol.* **1995**, *25*, 237–257. [[CrossRef](#)]

84. Abraham, C.; Kools, M. *Writing Health Communication: An Evidence-Based Guide*; Sage: Ventura County, CA, USA, 2011.
85. Burke, R. Knowledge-based recommender systems. *Encycl. Libr. Inf. Syst.* **2000**, *69*, 175–186.
86. Çano, E.; Morisio, M. Hybrid recommender systems: A systematic literature review. *Intell. Data Anal.* **2017**, *21*, 1487–1524. [[CrossRef](#)]
87. Herlocker, J.L.; Konstan, J.A.; Riedl, J. Explaining collaborative filtering recommendations. In Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, Philadelphia, PA, USA, 2–6 December 2000; pp. 241–250.
88. Cacioppo, J.T.; Petty, R.E.J.B.; Psychology, A.S. Effects of message repetition on argument processing, recall, and persuasion. *Basic Appl. Soc. Psychol.* **1989**, *10*, 3–12. [[CrossRef](#)]
89. Heatherston, T.F.; Kozlowski, L.T.; Frecker, R.C.; Fagerstrom, K.O. The Fagerström test for nicotine dependence: A revision of the Fagerstrom Tolerance Questionnaire. *Br. J. Addict.* **1991**, *86*, 1119–1127. [[CrossRef](#)] [[PubMed](#)]
90. Richmond, R.L.; Kehoe, L.A.; Webster, I.W. Multivariate models for predicting abstinence following intervention to stop smoking by general practitioners. *Addiction* **1993**, *88*, 1127–1135. [[CrossRef](#)] [[PubMed](#)]
91. Berkup, S.B. Working with generations X and Y in generation Z period: Management of different generations in business life. *Mediterr. J. Soc. Sci.* **2014**, *5*, 218. [[CrossRef](#)]
92. Cheung, K.L.; de Ruijter, D.; Hiligsmann, M.; Elfeddali, I.; Hoving, C.; Evers, S.M.; de Vries, H. Exploring consensus on how to measure smoking cessation. A Delphi study. *BMC Public Health* **2017**, *17*, 890. [[CrossRef](#)]
93. Gupta, S.K. Intention-to-treat concept: A review. *Perspect. Clin. Res.* **2011**, *2*, 109. [[CrossRef](#)]
94. Cacheda, F.; Carneiro, V.; Fernández, D.; Formoso, V. Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems. *ACM Trans. Web (TWEB)* **2011**, *5*, 1–33. [[CrossRef](#)]
95. Candillier, L.; Jack, K.; Fessant, F.; Meyer, F. State-of-the-art recommender systems. In *Collaborative and Social Information Retrieval and Access: Techniques for Improved User Modeling*; IGI Global: Hershey, PA, USA, 2009; pp. 1–22.
96. Desrosiers, C.; Karypis, G. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender Systems Handbook*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 107–144.
97. Ziegler, C.-N.; McNee, S.M.; Konstan, J.A.; Lausen, G. Improving recommendation lists through topic diversification. In Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, 10–14 May 2005; pp. 22–32.
98. Schafer, J.B.; Frankowski, D.; Herlocker, J.; Sen, S. Collaborative filtering recommender systems. In *The Adaptive Web*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 291–324.
99. Serrano-Guerrero, J.; Herrera-Viedma, E.; Olivas, J.A.; Cerezo, A.; Romero, F.P. A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2.0. *Inf. Sci.* **2011**, *181*, 1503–1516. [[CrossRef](#)]
100. Porcel, C.; Tejeda-Lorente, A.; Martínez, M.; Herrera-Viedma, E. A hybrid recommender system for the selective dissemination of research resources in a technology transfer office. *Inf. Sci.* **2012**, *184*, 1–19. [[CrossRef](#)]
101. Pu, P.; Chen, L.; Hu, R. A user-centric evaluation framework for recommender systems. In Proceedings of the Fifth ACM Conference on Recommender Systems, Chicago, IL, USA, 23–27 October 2011; pp. 157–164.
102. Torous, J.; Nicholas, J.; Larsen, M.E.; Firth, J.; Christensen, H. Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements. *Evid.-Based Ment. Health* **2018**, *21*, 116–119. [[CrossRef](#)] [[PubMed](#)]
103. Sandborg, J.; Henriksson, P.; Larsen, E.; Lindqvist, A.-K.; Rutberg, S.; Söderström, E.; Maddison, R.; Löf, M. Participants' Engagement and Satisfaction With a Smartphone App Intended to Support Healthy Weight Gain, Diet, and Physical Activity During Pregnancy: Qualitative Study Within the HealthyMoms Trial. *JMIR mHealth uHealth* **2021**, *9*, e26159. [[CrossRef](#)] [[PubMed](#)]
104. Dovaliene, A.; Piligrimiene, Z.; Masiulyte, A. Factors influencing customer engagement in mobile applications. *Eng. Econ.* **2016**, *27*, 205–212. [[CrossRef](#)]
105. Ahmadian, S.; Ahmadian, M.; Jalili, M. A deep learning based trust-and tag-aware recommender system. *Neurocomputing* **2021**, in press. [[CrossRef](#)]
106. Arnold, C.L.; Davis, T.C.; Berkel, H.J.; Jackson, R.H.; Nandy, I.; London, S. Smoking status, reading level, and knowledge of tobacco effects among low-income pregnant women. *Prev. Med.* **2001**, *32*, 313–320. [[CrossRef](#)]
107. Fakir, S.E.; Serhier, Z.; Berraho, M.; Elrhazi, K.; Tachfouti, N.; Slama, K.; Nejari, C. Knowledge and perceptions of smoking according to income level in Morocco. *Am. J. Health Promot.* **2011**, *25*, 387–391. [[CrossRef](#)]
108. Kaye, L.; Gondalia, R.; Rummelink, E.; Su, J.; Thompson, A.; Barrett, M. Impact of Socioeconomic Status on the Relationship Between Smartphone App Engagement and Adherence Among Patients with Asthma. In *B102. Caring for Lung Diseases in Underserved and Vulnerable Populations*; American Thoracic Society: New York, NY, USA, 2020; p. A4281.
109. Crafoord, M.-T.; Fjell, M.; Sundberg, K.; Nilsson, M.; Langius-Eklöf, A. Engagement in an Interactive App for Symptom Self-Management during Treatment in Patients With Breast or Prostate Cancer: Mixed Methods Study. *J. Med. Internet Res.* **2020**, *22*, e17058. [[CrossRef](#)]
110. Garnett, C.; Perski, O.; Tombor, I.; West, R.; Michie, S.; Brown, J. Predictors of engagement, response to follow up, and extent of alcohol reduction in users of a smartphone app (Drink Less): Secondary analysis of a factorial randomized controlled trial. *JMIR mHealth uHealth* **2018**, *6*, e11175. [[CrossRef](#)]
111. Zheng, Y.; Mobasher, B.; Burke, R. Similarity-based context-aware recommendation. In Proceedings of the International Conference on Web Information Systems Engineering, Amsterdam, The Netherlands, 20–24 October 2015; pp. 431–447.
112. Guo, G.; Zhang, J.; Yorke-Smith, N. A novel Bayesian similarity measure for recommender systems. In Proceedings of the IJCAI, Beijing, China, 3–9 August 2013; pp. 2619–2625.

113. Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J. Analysis of recommendation algorithms for e-commerce. In Proceedings of the 2nd ACM Conference on Electronic Commerce, Minneapolis, MN, USA, 17–20 October 2000; pp. 158–167.
114. Ahmadian, M.; Ahmadi, M.; Ahmadian, S. A Reliable Deep Representation Learning to Improve Trust-aware Recommendation Systems. *Expert Syst. Appl.* **2022**, *197*, 116697. [\[CrossRef\]](#)
115. Nelson, D.B.; Partin, M.R.; Fu, S.S.; Joseph, A.M.; An, L.C. Why assigning ongoing tobacco use is not necessarily a conservative approach to handling missing tobacco cessation outcomes. *Nicotine Tob. Res.* **2009**, *11*, 77–83. [\[CrossRef\]](#)
116. Blankers, M.; Smit, E.S.; van der Pol, P.; de Vries, H.; Hoving, C.; van Laar, M. The missing= smoking assumption: A fallacy in internet-based smoking cessation trials? *Nicotine Tob. Res.* **2016**, *18*, 25–33. [\[CrossRef\]](#) [\[PubMed\]](#)
117. Westmaas, J.L.; Bontemps-Jones, J.; Hendricks, P.S.; Kim, J.; Abrams, L.C. Randomised controlled trial of stand-alone tailored emails for smoking cessation. *Tob. Control* **2018**, *27*, 136–146. [\[CrossRef\]](#) [\[PubMed\]](#)
118. Regmi, K.; Kassim, N.; Ahmad, N.; Tuah, N.A. Effectiveness of mobile apps for smoking cessation: A review. *Tob. Prev. Cessat.* **2017**, *3*, 12. [\[CrossRef\]](#)
119. Bricker, J.B.; Mull, K.E.; Kientz, J.A.; Vilardaga, R.; Mercer, L.D.; Akioka, K.J.; Heffner, J.L. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. *Drug Alcohol Depend.* **2014**, *143*, 87–94. [\[CrossRef\]](#) [\[PubMed\]](#)
120. Ubhi, H.K.; Michie, S.; Kotz, D.; Wong, W.C.; West, R. A mobile app to aid smoking cessation: Preliminary evaluation of SmokeFree28. *J. Med. Internet Res.* **2015**, *17*, e17. [\[CrossRef\]](#)
121. Iacoviello, B.M.; Steinerman, J.R.; Klein, D.B.; Silver, T.L.; Berger, A.G.; Luo, S.X.; Schork, N.J. Clickotine, a personalized smartphone app for smoking cessation: Initial evaluation. *JMIR mHealth uHealth* **2017**, *5*, e56. [\[CrossRef\]](#)
122. Meyer, C.; Ulbricht, S.; Baumeister, S.E.; Schumann, A.; Rüge, J.; Bischof, G.; Rumpf, H.J.; John, U. Proactive interventions for smoking cessation in general medical practice: A quasi-randomized controlled trial to examine the efficacy of computer-tailored letters and physician-delivered brief advice. *Addiction* **2008**, *103*, 294–304. [\[CrossRef\]](#)
123. Te Poel, F.; Bolman, C.; Reubsaet, A.; de Vries, H. Efficacy of a single computer-tailored e-mail for smoking cessation: Results after 6 months. *Health Educ. Res.* **2009**, *24*, 930–940. [\[CrossRef\]](#)
124. Karumur, R.P.; Nguyen, T.T.; Konstan, J.A. Personality, user preferences and behavior in recommender systems. *Inf. Syst. Front.* **2018**, *20*, 1241–1265. [\[CrossRef\]](#)
125. Willemse, B.J.P.C.; Kaptein, M.C.; Hasaart, F. Developing Effective Methods for Electronic Health Personalization: Protocol for Health Telescope, a Prospective Interventional Study. *JMIR Res. Protoc.* **2020**, *9*, e16471. [\[CrossRef\]](#) [\[PubMed\]](#)
126. Kennedy-Martin, T.; Curtis, S.; Faries, D.; Robinson, S.; Johnston, J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* **2015**, *16*, 495. [\[CrossRef\]](#)
127. Harper, F.M.; Konstan, J.A. The movielens datasets: History and context. *Acm Trans. Interact. Intell. Syst. (TiiS)* **2015**, *5*, 1–19. [\[CrossRef\]](#)
128. Sahoo, A.K.; Pradhan, C.; Barik, R.K.; Dubey, H. DeepReco: Deep learning based health recommender system using collaborative filtering. *Computation* **2019**, *7*, 25. [\[CrossRef\]](#)
129. Rivero-Rodriguez, A.; Konstantinidis, S.T.; Sánchez-Bocanegra, C.L.; Fernández-Luque, L. A health information recommender system: Enriching YouTube health videos with Medline Plus information by the use of SnomedCT terms. In Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, Porto, Portugal, 20–22 June 2013; pp. 257–261.
130. Bocanegra, C.L.S.; Ramos, J.L.S.; Rizo, C.; Civit, A.; Fernandez-Luque, L. HealthRecSys: A semantic content-based recommender system to complement health videos. *BMC Med. Inform. Decis. Mak.* **2017**, *17*, 63.
131. Hors-Fraile, S.; Schneider, F.; Fernandez-Luque, L.; Luna-Perejon, F.; Civit, A.; Spachos, D.; Bamidis, P.; de Vries, H. Tailoring motivational health messages for smoking cessation using an mHealth recommender system integrated with an electronic health record: A study protocol. *BMC Public Health* **2018**, *18*, 698. [\[CrossRef\]](#)
132. Hutton, H.E.; Wilson, L.M.; Apelberg, B.J.; Avila Tang, E.; Odelola, O.; Bass, E.B.; Chander, G. A systematic review of randomized controlled trials: Web-based interventions for smoking cessation among adolescents, college students, and adults. *Nicotine Tob. Res.* **2011**, *13*, 227–238. [\[CrossRef\]](#)
133. Benowitz, N.; Jacob, P.; Ahijevych, K.; Jarvis, M.; Hall, S.; LeHouezec, J.; Hansson, A.; Lichtenstein, E.; Henningfield, J.; Tsoh, J. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. *Nicotine Tob. Res.* **2002**, *4*, 149–159. [\[CrossRef\]](#)
134. Stanczyk, N.; de Vries, H.; Candel, M.; Muris, J.; Bolman, C. Effectiveness of video-versus text-based computer-tailored smoking cessation interventions among smokers after one year. *Prev. Med.* **2016**, *82*, 42–50. [\[CrossRef\]](#)
135. Patrick, D.L.; Cheadle, A.; Thompson, D.C.; Diehr, P.; Koepsell, T.; Kinne, S. The validity of self-reported smoking: A review and meta-analysis. *Am. J. Public Health* **1994**, *84*, 1086–1093. [\[CrossRef\]](#) [\[PubMed\]](#)
136. Ramo, D.E.; Hall, S.M.; Prochaska, J.J. Reliability and validity of self-reported smoking in an anonymous online survey with young adults. *Health Psychol.* **2011**, *30*, 693. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Knijnenburg, B.P.; Willemssen, M.C.; Gantner, Z.; Soncu, H.; Newell, C. Explaining the user experience of recommender systems. *User Modeling User-Adapt. Interact.* **2012**, *22*, 441–504. [\[CrossRef\]](#)
138. Kaptein, M.; Eckles, D. Heterogeneity in the effects of online persuasion. *J. Interact. Mark.* **2012**, *26*, 176–188. [\[CrossRef\]](#)
139. Kaptein, M.; Markopoulos, P.; De Ruyter, B.; Aarts, E. Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles. *Int. J. Hum.-Comput. Stud.* **2015**, *77*, 38–51. [\[CrossRef\]](#)

-
140. Smith, A.L.; Chapman, S.; Dunlop, S.M. What do we know about unassisted smoking cessation in Australia? A systematic review, 2005–2012. *Tob. Control* **2015**, *24*, 18–27. [[CrossRef](#)]
 141. Hughes, J.R.; Keely, J.; Naud, S. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction* **2004**, *99*, 29–38. [[CrossRef](#)]
 142. Moore, D.; Aveyard, P.; Connock, M.; Wang, D.; Fry-Smith, A.; Barton, P. Effectiveness and safety of nicotine replacement therapy assisted reduction to stop smoking: Systematic review and meta-analysis. *BMJ* **2009**, *338*, b1024. [[CrossRef](#)]
 143. Shiffman, S.; Rolf, C.N.; Hellebusch, S.J.; Gorsline, J.; Gorodetzky, C.W.; Chiang, Y.K.; Schleusener, D.S.; Marino, M.E.D. Real-world efficacy of prescription and over-the-counter nicotine replacement therapy. *Addiction* **2002**, *97*, 505–516. [[CrossRef](#)]