Comparing Familiar with Inspiring Recommendations to Assist People in Moving More

Ine Coppens[1,*], Luc Martens[1] and Toon De Pessemier[1]

[1] WAVES - imec - Ghent University, iGent - Technologiepark-Zwijnaarde 126, Ghent, Belgium

Abstract

Sufficient physical activity is crucial for people's health and well-being. However, not enough people attain the weekly minimum of 150 minutes. Since current mobile health systems are not optimal to motivate and assist people to move more, this study investigates the effect of personalized suggestions generated by two types of recommender system algorithms: content-based (which provide more familiar recommendations, relevant to existing interests) and user-based collaborative filtering (which deliver more diverse recommendations, allowing inspiration for new interests). By conducting a longitudinal between-subject user study over eight weeks, we will investigate how the two algorithms separately affect motivation and behavior change. We developed two versions of an Android smartphone application to deliver the recommendations, with the only difference being the implemented recommender algorithm. In all other aspects, the apps are identical: Both systems use the same datasets of physical activities and tips to break sedentary behavior, apply the user profile and contextual filter, and integrate the combination of star rating and momentary motivation feedback to provide personalization on preferences and well-being. We will analyze the differences in people's star rating feedback, motivation to move, physical activity, and sedentary behavior. The main hypothesis is that inspiring recommendations from the collaborative algorithm will motivate people more for more physical activity and less sedentary behavior. The results of this study will provide insights for future mobile health recommenders in what type of algorithm and recommendations are most effective in the domain of increasing physical activity and motivating people to move more.

Keywords

health recommender system, physical activity, motivation, behavior change, mobile health, assistive healthcare, sedentary behavior

1. Introduction

Insufficient physical activity (PA) is one of the modifiable underlying causes of chronic diseases, which cause most deaths worldwide [1]. The World Health Organization (WHO) defines evidence-based guidelines for increasing PA and reducing sedentary behavior (SB) [2]. However, in 2016, 27.5% of the adult population did not meet their recommended minimum of 150 minutes of moderate aerobic PA per week [1]. Since sufficient PA is essential for people's health and mental well-being, PA promotion is now more crucial than ever [3].

Electronic health (eHealth) and mobile health (mHealth) interventions use technologies to promote healthy behavior [4]. As such, they can also be used to assist people in moving more by promoting PA and prevent long periods of SB. In previous eHealth and mHealth studies to increase PA, the content of their interventions ranged from activities [5], ideas to break SB [6], more general healthy habits [7], or reminders and tips [8]. Despite their great potential to motivate people, the interventions are often underused [9]. Furthermore, other research suggests that they currently have limited effects on PA and SB, even when implementing behavior change techniques, such as goal setting and self-monitoring [10]. This implies the need for new technologies and more interactive interventions [10].

To increase user engagement and behavior change towards more PA, mHealth systems can implement Recommender System (RS) algorithms to deliver personalized and relevant interventions to the user [9, 11]. RSs generate personalized suggestions based on user preferences to help them with making decisions [12]. They can also be applied in the health domain as Health Recommender Systems (HRSs) to propose healthier suggestions, tailored to the user [13]. Previous work has applied RS techniques to provide personalized well-being recommendations for food and PA [14], for personalized training sessions for marathon running [15], and for health activities [11]. Although providing the most relevant health suggestion to the user would optimize mHealth interventions, application of HRSs for behavior change is still in its infancy [9, 13].

To generate useful recommendations, the RS has to predict what the relevant items are for the user, for which several techniques exist [9, 12]. The content-based tech-

Ine Coppens, Luc Martens and Toon De Pessemier. 2023. Comparing Familiar with Inspiring Recommendations to Assist People in Moving More. In Joint Proceedings of the ACM IUI 2023 Workshops, March 2023, Sydney, Australia, 8 pages.

*Corresponding author.
Ine.Coppens@UGent.be (I. Coppens); Luc1.Martens@UGent.be (L. Martens); Toon.DePessemier@UGent.be (T. De Pessemier)
0000-0002-3051-506X (I. Coppens); 0000-0001-9948-9157 (L. Martens); 0000-0002-3920-7346 (T. De Pessemier)

nique suggests items similar to items liked by the user in the past, based on attributes that describe the items [12]. Alternatively, collaborative filtering uses other users' ratings and assumes that people with the same interests will like the same items [9]. The item-based collaborative filtering technique recommends similar items based on the ratings of other users, while the user-based collaborative filtering technique focuses on recommending items that similar users with similar preferences liked in the past [12]. While the item-based method provides more accurate recommendations as the user's preferences are modeled using similar items, the user-based approach can recommend more diverse and unexpected items [12].

Providing different approaches to predict what a user might like, these RS techniques result in a different selection of recommended items [9]. While content-based RSs succeed in recommending highly relevant items, they often suffer from overspecialization as they suggest items very similar to items the user already knows because the attributes are already defined in the user profile [16]. As such, they fail at recommending more unexpected, surprising, and novel items that could still be relevant to the user [12, 17]. Previous work has addressed this overspecialization problem on the grounds that it leads to lower user satisfaction [18, 17, 19, 16]. Collaborative systems solve this problem because they can recommend items with a very different content when it is liked by similar users [12, 16]. To summarize, there are content-based algorithms that provide familiar recommendations which are highly relevant to existing interests, and collaborative filtering that can deliver more diverse and unexpected suggestions which allow new interests to be explored [12, 19, 17]. Hybrid RS algorithms combine the advantages of the content-based and collaborative approach, providing a balance between relevant and diverse recommendations [12, 18].

In this research, however, we do not want to balance the characteristics of the algorithms by merging them in a hybrid RS. Rather, we want to study the algorithms and the impact of their advantages and disadvantages separately in the domain of PAs. For example, previous research has shown that repetition of the same health behavior makes the behavior easier [20], suggesting that overspecialization may not be a problem in the domain of PA. Similarly, we chose to implement the user-based version of collaborative filtering because it can recommend more diverse items than the item-based version [12], and because we want to emphasize the effect of more diverse recommendations on people's behavior. As such, we investigate the content-based and user-based collaborative RS algorithm separately as two extremes (very relevant versus very diverse) to gain understanding in how they each affect motivation and behavior change.

In this study, concrete PAs and tips to break SB are recommended with the goal to motivate people for healthy behavior change by having more PA and less SB, as recommended by the WHO [2]. Because the two RS techniques will provide different recommendations, we expect a different effect on user motivation and behavior. To the best of our knowledge, this effect of different types of RS algorithms has not been investigated. As such, we examine which RS algorithm will perform best in motivating users for more PA and less SB, responding to the demand of enhancing health interventions with the best personalization approach [9]. By developing two versions of the same Android app, we will conduct a between-subject user study with the following research question:

Which recommender algorithm has the best effect on people's star rating feedback, motivation to move, physical activity, and sedentary behavior?

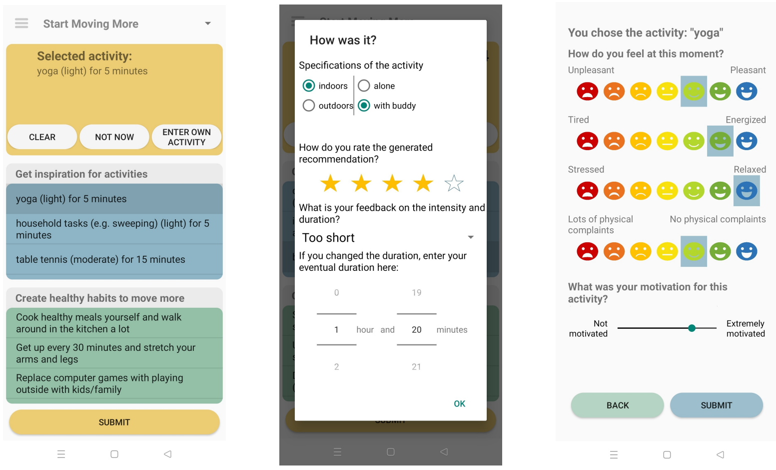
2. Methods

We developed two HRSs that recommend personalized PAs and tips for breaking SB to assist users in their daily life in moving more. For these PA and tip items, our own two datasets were created. The PA dataset was assembled using 354 PAs from the Compendium of Physical Activities [21]. The tip dataset contains ideas from the Belgian website for health (www.gezondleven.be/), resulting in 81 items. The generated recommendations are delivered to the user in an Android app called MoveMoreApp, as shown in Figure 1(a), with its interface similar to our app from a previous study. This app shows three PA and three tip recommendations. When a user executes an item, manual feedback on the recommended items is collected as a rating on five stars, as illustrated in Figure 1(b) with the question *"how do you rate the generated recommendation?"*. Additionally, our system collects users' momentary motivation to move with a slider measured on a 5-point Likert scale (from *"not motivated"* to *"extremely motivated"*), as depicted in Figure 1(c).

2.1. The algorithms

The PA and tip items are recommended to the users with two types of RS algorithms, as illustrated in Figure 2. The initial filter based on the user's profile (available material and maximum impact level) and the contextual filter based on the current weather (obtained using https://openweathermap.org/) and remaining daylight are applied on the PA and tip datasets in both groups to remove unsuitable items.

In the next step, the RSs generate personalized recommendations based on the users' consumption history. This history contains the PAs and tips the user engaged in, together with the provided star rating feedback, momentary motivation, and the user's mood. The star rating feedback and momentary motivation are both



(a) (b) (c)

Figure 1: Three recommended activities and three recommended tips are shown in the user interface of MoveMoreApp (a). After selecting and having engaged in an activity or tip, users can specify more details about it, such as the location, buddy, star rating, and for PAs also feedback on intensity and duration (b). After submitting the item, the app also asks the momentary motivation for the activity or tip, together with the user's current mood (c).

measured on a scale of five and are aggregated with equal weights in the formula: $aggregated\_feedback = (rating + motivation)/2$. In this way, our two RSs optimize their recommendations on both the rating and motivation. The mood is asked at the beginning of every day and after every submit with several emoji, as shown in Figure 1(c). As such, the user's current mood is used to filter the consumption history on previous consumptions with a similar mood, based on the mood micro-profile of [22].

The content-based RS algorithm only needs the user's own consumption history. Calculating the similarity to items consumed in the past relies on attributes that describe the items [12]. As such, our PA and tip dataset were extended with corresponding attributes to describe each item, such as aerobic, flexibility, or balance. The content-based algorithm uses these to represent the user's preferences and match these with all the filtered PA and tip items using the cosine similarity [12]. In the other group, the collaborative filtering searches for similar users who provided similar feedback to the same items using the cosine similarity, and calculates a preference estimation score for all the filtered PAs and tips [12].

At this point, both RS algorithms generated a list of recommended items with corresponding preference estimation scores. The contextual post-filter re-ranks the items based on the current estimated situation [12]. In our study, this situation can be: *free time*, *during work*, *household task*, or *active transport*, and is assigned to every item in the two datasets. In this way, the consumption history also contains the situation history at the corresponding time. To re-rank the items, a value between 0 and 1 that represents how close in time the item's situation is to the situation's occurrences in the history is added to the preference estimation score. As a result, items that match better with the estimated current situation appear higher in the list of recommendations.

Next, the recommended PA items go through the adaptive step. Combined with the user's current PA level and feedback on intensity and duration provided in the app, as shown in Figure 1(b), the system provides a gradual increase in PA intensity and duration, following guidelines of the WHO [2] and the European Society of Cardiology [3]. In the final step, the recommended PAs and tips are shown to the user.

Right at the beginning, when the users did not submit and rate any PAs or tips yet, there is no consumption history present to derive the user preferences from and base the recommendations on, resulting in the *new user cold start problem* [12]. To provide an initial recommendation with the available information, the two algorithms apply the user profile filter and the contextual filter, and then randomly select PAs and tips from this filtered set. As more PAs and tips are chosen, the consumption history will grow over time, resulting in better, more personalized recommendations. It is however possible that users do not like any of the (initial) recommendations and do not select anything. In that case, the app allows users to select their own chosen PAs from the PA dataset with a search functionality when clicking on the "enter own

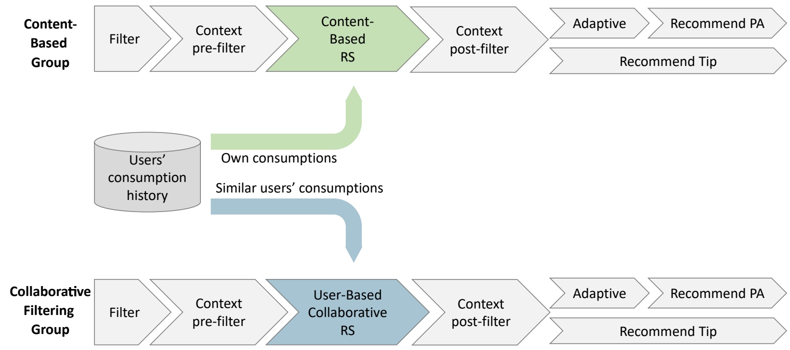


Figure 2: A schematic representation of how the algorithms in both versions work shows that the only difference between the two groups is the applied RS algorithm.

activity" button, as shown in Figure 1(a). These PA consumptions are used in the RS algorithm for subsequent recommendations. The *new item cold start problem*, on the other hand, occurs when no ratings for the unexplored items are available yet [12], which is not a problem for the content-based RS since this algorithm only uses attributes to recommend other items. The collaborative RS however, does depend on item ratings from other users [12]. We address this new item cold start problem by integrating an initial user-item consumption dataset from a previous study, in which items already received ratings from other users and to which no new items are added. Moreover, since our item datasets are relatively small compared to the amount of users (354 PAs and 81 tips) [12], and since we expect that users will engage in daily PA (which is any movement of the body, as defined by the WHO [2]), we estimate a sufficient amount of consumptions after one week to alleviate the cold start problems.

2.2. Participants

The target group of our study are adults who currently do not achieve the 150-minute weekly minimum of moderate PA. An initial screening with questions about age, weekly amount of PA [23], and a PA screening [24] in the app will decide whether or not the participant is eligible to join the study. Aimed at recruiting 50 participants, we promote our study via the Sona research participation system of Ghent University and several Facebook groups for paid studies. The study will run from March until June, 2023.

Participants will receive an incentive of 30 EUR when they used the app for eight weeks and answered all the questionnaires. They are not rewarded for having more PA or for the amount of PAs or tips they submit, because they can also use the app with "not now" and "enter own activity" submits. As such, they are free to choose from the generated recommendations and they are stimulated to only submit items when actually having engaged in them, rather than only rating them for more money.

We designed the processing of data collected by our app together with our ethical committee and data protection officers to be compliant with the General Data Protection Regulation (GDPR) and our study received ethical approval.

2.3. User study design

A longitudinal user study will be conducted following a between-subject study design in which each user is assigned to either the content-based or collaborative filtering method. The advantage of between-subject user studies is the possibility to investigate the long-term effect of one system separately without having to switch between systems, but it also requires more users and more interactions [12]. As illustrated in Figure 2, the only difference between the two groups is the type of RS algorithm. The other steps (user profile filter, contextual pre-filter, contextual post-filter, and adaptive algorithm for PAs) are exactly the same.

Participants are asked to install the Android application on their own smartphone. Immediately after installation, the app randomizes the participants in the content-based or collaborative filtering group. Then, participants are asked to answer the pre-test questionnaire, followed by an eight-week study. During these eight weeks, they can use the app in their daily life to look at the recommendations and choose an item to execute. When an item is selected, as shown in Figure 1(a), this is saved in the app even when the app is closed during the execution of the activity. When the activity or tip is executed, the user goes back to the app to submit and rate it, as depicted in Figure 1(b), in which the eventual duration of the executed PA is also asked. As such, participants are requested to only submit PAs and tips after engaging in them to provide proper feedback on the eventual rating, motivation, and duration. After eight weeks, the app shows a final post-test questionnaire.

Since the goal of our study is to investigate the differences of receiving personalized recommendations from either the content-based or the collaborative RS algorithm, the study duration is dependent on the time it takes for the RSs to succeed in generating personalized recommendations. By providing solutions for the cold start problems as discussed earlier, we expect that the RSs will be able to provide personalization after one week. In total, we decided on a study duration of eight weeks, reasoning that longer durations would result in more user dropout [9]. We expect that users will have submitted sufficient consumptions, and that sufficient PAs and tips will have been recommended in eight weeks to answer our research question.

2.4. Measures and analyses

When the study is finished, statistical analyses will be conducted using IBM SPSS Statistics Version 28 to answer our research question. The research question is divided into four main dependent variables: star rating feedback, motivation to move, amount of PA, and SB. These variables are all measured using the Android app at different points in time. Depending on the timing of measurement of the dependent variable, different types of statistical tests will be conducted on the longitudinal dataset and the pre-post dataset.

Firstly, measurements per individual are repeated over the eight-week study resulting in a longitudinal dataset. The repeated measurements include: star rating feedback on a recommended item, momentary motivation to move, and the daily executed PAs and tips. Because of this longitudinal data, in which the data can be unbalanced (e.g., not every user engages in the same amount of PAs), analyses will be conducted with Generalized Estimating Equations [25] to investigate differences between the groups.

Secondly, motivation and behavior change are also measured in both the pre- and post-test questionnaires to investigate their evolution after the eight-week study. To measure motivation, we chose to utilize the regulation types of motivation as defined by the self-determination theory (SDT), a theory of motivation that distinguishes between autonomous and controlled motivation [26]. Based on the SDT, the motivation for PA (RM4-FM) questionnaire [27] and the Behavioral Regulations for Exercise Questionnaire (BREQ) [28] measure the motivation types for PA and exercise, respectively. By using separate questionnaires, we differentiate between PA, which the WHO defines as any movement of the body [2], and exercise, which is a subset of PA. To measure behavior change, we chose to analyze changes in PA, surveyed with the European Health Interview Survey - Physical Activity Questionnaire (EHIS-PAQ) [23], and SB, surveyed with the Sedentary Behavior Questionnaire (SBQ) [29] because they both allow participants to reflect on their average weekly PA and SB behavior, and they both distinguish between different situations, such as PA or SB at work or as transport. Repeated Measures ANOVA tests will be conducted to investigate the evolution in motivation regulation style and behavior change between the pre- and post-test measurements and between the two groups [30].

A manipulation check will validate whether the manipulation succeeded. The manipulation in our study is generating either familiar recommendations with the content-based RS, or diverse recommendations with the collaborative RS. The user's experience of these recommendations can be measured with the questionnaires of [31]. In these questionnaires, different RS properties are surveyed, such as perceived recommendation accuracy and quality (e.g., *"The recommended items fitted my preference"*), and additional properties that measure beyond accuracy, such as perceived recommendation diversity and variety (e.g., *"The list of recommendations was varied"*) [31, 12]. To keep the app user friendly, the app will not ask these questionnaires every time the user receives a recommendation. Instead, the app will randomly show these questionnaires in 20% of the time after the user chose and submitted a PA or tip recommendation. As a result, these data will also be longitudinal with repeated measures over eight weeks, and Generalized Estimating Equations [25] will be conducted for the analysis of the manipulation check.

As the success and usefulness of an RS algorithm is based on how well it can predict the user's preferences [12], the stability of the preferences determines which algorithm will provide the best recommendations [17]. In some domains, such as movies, user preferences are mostly stable over time, thus eliminating the need for diverse recommendations [17]. On the other hand, some people seek variety in their behavior, indicating the need for novelty and diversity in the recommendations [12]. In this case, RSs should take into account the differences in user preferences, which can be depended on their personality [12] or change over time [32]. For this reason, we also survey the user's preference for variety in the pre-test questionnaire with our own questions, rated on a 5-point Likert scale from "Disagree strongly" to "Agree strongly": *"I like variety in my daily physical activities"* and *"I prefer routine in my daily physical activities"*. This independent variable will serve as a control variable in the aforementioned analyses.

To evaluate the overall performance of all the steps of the algorithms, the "not now" button allows users to provide a reason why now is not a good time for PA. We provided our own feedback sentences to check whether or not the recommendations fit with the weather (e.g., *"It is raining too much"*) or with the current mood (e.g., *"I do not feel good"*), whether or not they are adapted to the user's PA level (e.g., *"The recommendations are too intense"*), and whether or not the situation is suited for the recommendation (e.g., *"I'm still at work/school"*).

3. Expected results

We will first check whether our manipulation succeeded by analyzing the users' experience with the generated recommendations. We expect that participants in the content-based group will assign larger scores for perceived recommendation accuracy and quality [31] because the content-based algorithm will generate recommendations that fit better with the user preferences [12]. Furthermore, we check whether the collaborative algorithm provided more diverse recommendations, as we expect larger scores for perceived recommendation diversity and variety [12, 31].

As content-based RSs generate recommendations that are similar to previously consumed items, and thus, fit better with user preferences [12], we hypothesize that the assigned star rating feedback will be higher in the content-based group. However, content-based RSs do not provide an exploration of new items and expansion of their knowledge [17], and they ignore items with little similarities [18]. Moreover, we expect that integrating more variety and unexpected items in the recommendations with collaborative filtering will enhance their enjoyment [9], inspire them with new interests, and expand their horizon [12, 17]. We hypothesize that increasing inspiration for new ways to move will motivate people more because varied content is important to keep the users engaged [33]. As such, we hypothesize that momentary motivation to move, and thereby also the amount of executed PAs and tips, will be higher in the collaborative filtering group.

Since both groups of participants receive an app aimed at increasing PA, we expect that both groups will have more PA and less SB in the post-test compared to the pretest. By following a between-subject study design, the long-term effect of the applied system can be assessed as a whole [12], allowing us to compare the evolution in motivation regulation style and behavior change between the two groups. Following the SDT, the autonomous motivation regulation types are associated with people's own willingness to engage in the behavior and with more psychological health, while controlled motivation is associated with pressure to behave in a certain way [26]. Because we expect more enjoyment with the inspiring recommendations of the collaborative filtering group [9], we hypothesize that their autonomous motivation for PA will increase. As a result of higher autonomous motivation, we hypothesize that the increase of PA and the decrease of SB will be stronger in the collaborative filtering group because autonomous motivation results in more effective healthy behavior change [26].

Lastly, we hypothesize that the collaborative RS will perform better (e.g., higher star ratings, momentary motivation, and amount of PA) when combined with a user who needs more variety in their behavior because it generates more diverse recommendations [12] and allows exploration of new items and interests [17]. Similarly, we hypothesize that the content-based RS will perform better when combined with a user who prefers routine because it generates recommendations similar to items the user already engaged in and already knows [12, 17, 16]. Moreover, repeating the same behaviors can make them easier [20], mitigating the overspecialization problem of the content-based RS. Since this research examines whether RSs should focus on existing interests or on discovering new interests in the domain of PAs in an eight-week period, we will not investigate whether or not these interests persist as habits, as previous research has indicated that habit formation may take up to 254 days [20].

4. Conclusions and future work

This research investigates whether content-based or collaborative filtering recommendations have a better effect on people's motivation and behavior change for PA when implemented in an HRS that assists people in moving more. The effectiveness of the HRSs will be evaluated with a between-subject eight-week user study and an Android application that randomly assigns each participant to either the content-based or the user-based collaborative filtering RS algorithm. Expecting different effects on motivation and behavior, we hypothesize that collaborative filtering will provide inspiration with new ways to move, and motivate users more than the familiar items suggested by the content-based algorithm.

To the best of our knowledge, the most optimal type of algorithm for an HRS in the domain of PA has not been investigated. Understanding how the algorithms separately affect motivation and behavior change is important before combining them in a hybrid system. As such, this study will contribute to new insights in effective algorithms for developers of future HRSs. For example, future hybrid RS algorithms can assign different weights to content-based and collaborative filtering recommendation outcomes, depending on the degree to which the user prefers a familiar routine or varied inspiration in daily activities.

References

[1] World Health Organization, World health statistics 2020: monitoring health for the SDGs, sustainable development goals, 2020. URL: https://apps.who.int/iris/bitstream/handle/10665/332070/9789240005105-eng.pdf, oCLC: 960910741.

[2] World Health Organization, WHO guidelines on physical activity and sedentary behaviour., 2020. URL: http://www.ncbi.nlm.nih.gov/books/NBK566045/, oCLC: 1237095892.

[3] A. Pelliccia, S. Sharma, S. Gati, et al., 2020 ESC Guidelines on sports cardiology and exercise in patients with cardiovascular disease, European Heart Journal 42 (2021) 17–96. URL: https://academic.oup.com/eurheartj/article/42/1/17/5898937. doi:10.1093/eurheartj/ehaa605.

[4] World Health Organization Global Observatory for eHealth, World Health Organization, MHealth: new horizons for health through mobile technologies., World Health Organization, Geneva, 2011. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf, oCLC: 759446797.

[5] P. Liao, K. Greenewald, P. Klasnja, S. Murphy, Personalized HeartSteps: A Reinforcement Learning Algorithm for Optimizing Physical Activity, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4 (2020) 1–22. URL: https://dl.acm.org/doi/10.1145/3381007. doi:10.1145/3381007.

[6] J. P. Sanders, S. J. Biddle, K. Gokal, L. B. Sherar, M. Skrybant, H. M. Parretti, N. Ives, T. Yates, N. Mutrie, A. J. Daley, ‘Snacktivity™’ to increase physical activity: Time to try something different?, Preventive Medicine 153 (2021) 106851. URL: https://linkinghub.elsevier.com/retrieve/pii/S0091743521004205. doi:10.1016/j.ypmed.2021.106851.

[7] M. Rabbi, A. Pfammatter, M. Zhang, B. Spring, T. Choudhury, Automated personalized feedback for physical activity and dietary behavior change with mobile phones: A randomized controlled trial on adults, JMIR mHealth and uHealth 3 (2015) e42. doi:10.2196/mhealth.4160.

[8] J. Woodward, Y.-P. Chen, K. Jurczyk, K. M. Ross, L. Anthony, J. Ruiz, A Survey of Notification Designs in Commercial mHealth Apps, in: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, Yokohama Japan, 2021, pp. 1–7. URL: https://dl.acm.org/doi/10.1145/3411763.3451639. doi:10.1145/3411763.3451639.

[9] K. L. Cheung, D. Durusu, X. Sui, H. de Vries, How recommender systems could support and enhance computer-tailored digital health programs: A scoping review, DIGITAL HEALTH 5 (2019) 1–19. doi:10.1177/2055207618824727.

[10] A. Direito, E. Carraça, J. Rawstorn, R. Whittaker, R. Maddison, mHealth Technologies to Influence Physical Activity and Sedentary Behaviors: Behavior Change Techniques, Systematic Review and Meta-Analysis of Randomized Controlled Trials, Annals of Behavioral Medicine 51 (2017) 226–239. doi:10.1007/s12160-016-9846-0.

[11] H. Torkamaan, J. Ziegler, Recommendations as Challenges: Estimating Required Effort and User Ability for Health Behavior Change Recommendations, in: 27th International Conference on Intelligent User Interfaces, ACM, Helsinki Finland, 2022, pp. 106–119. URL: https://dl.acm.org/doi/10.1145/3490099.3511118. doi:10.1145/3490099.3511118.

[12] F. Ricci, L. Rokach, B. Shapira (Eds.), Recommender Systems Handbook, Springer US, New York, NY, 2022. doi:10.1007/978-1-0716-2197-4.

[13] S. Hors-Fraile, O. Rivera-Romero, F. Schneider, L. Fernandez-Luque, F. Luna-Perejon, A. Civit-Balcells, H. de Vries, Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: A scoping review, International Journal of Medical Informatics 114 (2018) 143–155. URL: https://linkinghub.elsevier.com/retrieve/pii/S1386505617304690. doi:10.1016/j.ijmedinf.2017.12.018.

[14] H. Alcaraz-Herrera, J. Cartlidge, Z. Toumpakari, M. Western, I. Palomares, Evorecsys: Evolutionary framework for health and well-being recommender systems, User Modeling and User-Adapted Interaction (2022). doi:10.1007/s11257-021-09318-3.

[15] B. Smyth, A. Lawlor, J. Berndsen, C. Feely, Recommendations for marathon runners: on the application of recommender systems and machine learning to support recreational marathon runners, User Modeling and User-Adapted Interaction (2021) 787–838. doi:10.1007/s11257-021-09299-3.

[16] Y. Blanco-Fernandez, J. Pazos-arias, A. Gil-Solla, M. Ramos-Cabrer, M. Lopez-Nores, Providing entertainment by content-based filtering and semantic reasoning in intelligent recommender systems, IEEE Transactions on Consumer Electronics 54 (2008) 727–735. doi:10.1109/TCE.2008.4560154.

[17] R. Saia, L. Boratto, S. Carta, A new perspective on recommender systems: A class path information model, in: 2015 Science and Information Conference (SAI), IEEE, London, United Kingdom, 2015, pp. 578–585. doi:10.1109/SAI.2015.7237201.

[18] Y.-S. Cheng, P.-Y. Hsu, Y.-C. Liu, Identifying and recommending user-interested attributes with values,

Industrial Management & Data Systems 118 (2018) 765–781. doi:10.1108/IMDS-04-2017-0164.

[19] D. Kotkov, J. Veijalainen, S. Wang, How does serendipity affect diversity in recommender systems? A serendipity-oriented greedy algorithm, Computing 102 (2020) 393–411. doi:10.1007/s00607-018-0687-5.

[20] P. Lally, C. H. M. van Jaarsveld, H. W. W. Potts, J. Wardle, How are habits formed: Modelling habit formation in the real world, European Journal of Social Psychology 40 (2010) 998–1009. doi:10.1002/ejsp.674.

[21] B. E. Ainsworth, W. L. Haskell, S. D. Herrmann, N. Meckes, D. R. Bassett, C. Tudor-Locke, J. L. Greer, J. Vezina, M. C. Whitt-Glover, A. S. Leon, 2011 Compendium of Physical Activities: A Second Update of Codes and MET Values, Medicine & Science in Sports & Exercise 43 (2011) 1575–1581. URL: https://journals.lww.com/00005768-201108000-00025. doi:10.1249/MSS.0b013e31821ece12.

[22] L. Baltrunas, X. Amatriain, Towards time-dependant recommendation based on implicit feedback, 2009.

[23] J. D. Finger, J. Tafforeau, L. Gisle, L. Oja, T. Ziese, J. Thelen, G. B. M. Mensink, C. Lange, Development of the European Health Interview Survey - Physical Activity Questionnaire (EHIS-PAQ) to monitor physical activity in the European Union, Archives of Public Health 73 (2015) 59. doi:10.1186/s13690-015-0110-z.

[24] R. Adams, Revised Physical Activity Readiness Questionnaire, Canadian Family Physician Medecin De Famille Canadien 45 (1999) 992, 995, 1004–1005.

[25] S. Pekár, M. Brabec, Generalized estimating equations: A pragmatic and flexible approach to the marginal GLM modelling of correlated data in the behavioural sciences, Ethology 124 (2018) 86–93. doi:10.1111/eth.12713.

[26] E. L. Deci, R. M. Ryan, Self-determination theory: A macrotheory of human motivation, development, and health., Canadian Psychology/Psychologie canadienne 49 (2008) 182–185. doi:10.1037/a0012801.

[27] Deci, Edward L, Richard M. Ryan, Exercise Self-Regulation Questionnaires, Self-Determination Theory: An Approach to Human Motivation and Personality—The Self-Regulation Questionnaires (2004).

[28] P. M. Wilson, W. M. Rodgers, C. C. Loitz, G. Scime, "It's Who I Am … Really!' The Importance of Integrated Regulation in Exercise Contexts1, Journal of Applied Biobehavioral Research 11 (2007) 79–104. doi:10.1111/j.1751-9861.2006.tb00021.x.

[29] D. E. Rosenberg, G. J. Norman, N. Wagner, K. Patrick, K. J. Calfas, J. F. Sallis, Reliability and Validity of the Sedentary Behavior Questionnaire (SBQ) for Adults, Journal of Physical Activity and Health 7 (2010) 697–705. doi:10.1123/jpah.7.6.697.

[30] F. Wan, Statistical analysis of two arm randomized pre-post designs with one post-treatment measurement, BMC Medical Research Methodology 21 (2021) 150. doi:10.1186/s12874-021-01323-9.

[31] B. P. Knijnenburg, M. C. Willemsen, Z. Gantner, H. Soncu, C. Newell, Explaining the user experience of recommender systems, User Modeling and User-Adapted Interaction 22 (2012) 441–504. doi:10.1007/s11257-011-9118-4.

[32] F. Zafari, I. Moser, T. Baarslag, Modelling and analysis of temporal preference drifts using a component-based factorised latent approach, Expert Systems with Applications 116 (2019) 186–208. URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417418305815. doi:10.1016/j.eswa.2018.09.010.

[33] J. Freyne, J. Yin, E. Brindal, G. A. Hendrie, S. Berkovsky, M. Noakes, Push Notifications in Diet Apps: Influencing Engagement Times and Tasks, International Journal of Human–Computer Interaction 33 (2017) 833–845. doi:10.1080/10447318.2017.1289725.