

## RESEARCH ARTICLE

# A Comprehensive User Modeling Framework and a Recommender System for Personalizing Well-Being Related Behavior Change Interventions: Development and Evaluation

ANITA M. HONKA<sup>1</sup>, HANNU NIEMINEN<sup>2</sup>, (Member, IEEE), HEIDI SIMILÄ<sup>3</sup>,  
JOUNI KAARTINEN<sup>3</sup>, AND MARK VAN GILS<sup>2</sup>, (Member, IEEE)

<sup>1</sup>Department of Smart Health, VTT Technical Research Centre of Finland Ltd., 33101 Tampere, Finland

<sup>2</sup>Faculty of Medicine and Health Technology, Tampere University, 33100 Tampere, Finland

<sup>3</sup>Department of Smart Health, VTT Technical Research Centre of Finland Ltd., 90570 Oulu, Finland

Corresponding author: Jouni Kaartinen (jouni.kaartinen@vtt.fi)

This work was supported in part by the Advanced Research and Technology for Embedded Intelligent Systems Industry Association (ARTEMIS-IA) within Europe and the Finnish Agency for Technology and Innovation (WITH-ME), Tekes, under Grant 332885; in part by the Orion Research Foundation sr.; and in part by the Finnish Cultural Foundation.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Human Sciences of University of Oulu.

**ABSTRACT** Health recommender systems (HRSs) have the potential to effectively personalize well-being related behavior change interventions to the needs of individuals. However, personalization is often conducted with a narrow perspective, and the underlying user features are inconsistent across HRSs. Particularly, theory-based determinants of behavior and the variety of lifestyle domains influencing well-being are poorly addressed. We propose a comprehensive theory-based framework of user features, the virtual individual (VI) model, to support the extensive personalization of digital well-being interventions. We introduce a prototype HRS (With-Me HRS) with knowledge-based filtering, which recommends behavior change objectives and activities from several lifestyle domains. With-Me HRS realizes a minimum set of important VI model features related to well-being, lifestyle, and behavioral intention. We report the preliminary validity and usefulness of the HRS, evaluated in a real-life health-coaching program with 50 participants. The recommendations were used in decision-making for half of the participants and were hidden for others. For 73% of the participants (85% with visible vs. 62% with hidden recommendations), at least one of the recommended activities was included into their coaching plans. The HRS reduced coaches' perceived effort in identifying appropriate coaching tasks for the participants (effect size: Vargha-Delaney  $\hat{A} = 0.71$ , 95% CI 0.59-0.84) but not in identifying behavior change objectives. From the participants' perspective, the quality of coaching improved (effect size for one of three quality metrics:  $\hat{A} = 0.71$ , 95% CI 0.57-0.83). These results provide a baseline for testing the influence of additional user model features on the validity of recommendations generated by knowledge-based multi-domain HRSs.

**INDEX TERMS** Behavioral sciences, digital health behavior change interventions, disease prevention, eHealth, filtering algorithms, knowledge based systems, recommender systems, user evaluation, user modeling.

The associate editor coordinating the review of this manuscript and approving it for publication was Fabrizio Messina<sup>1</sup>.

## I. INTRODUCTION

In Europe, nearly 90% of the disease burden is attributed to chronic diseases, such as cardiovascular disease, cancer, and diabetes. Most of these diseases can be avoided or at least delayed with healthy behaviors. [1] Digital health behavior change interventions (DHBCIs), personalized to the needs and capabilities of individuals, have the potential to offer cost-effective solutions for empowering individuals to take care of their well-being [2], [3], [4]. Personalization can increase user engagement with digital interventions [3], [5], which is imperative for positive health outcomes. We consider personalized DHBCIs as adaptive interventions [6], [7] that aim to modify intervention content, dose, timing, or approach according to the characteristics of an individual users in order to achieve favorable behavioral or health outcomes.

Well-being is a broad concept comprising behavioral, mental, physical, and social dimensions. When attempting to improve well-being with the goal of preventing lifestyle-related diseases, several behavioral domains need to be taken into account, such as physical activity, dietary habits, sleep, smoking, alcohol consumption, stress management, work-life balance, and the cultivation of social relationships [8], [9]. Furthermore, the personal, social, and environmental factors that determine behavior [10], [11], [12] should be considered when personalizing health behavior change interventions. Individuals differ, for instance, in their behavior change needs, readiness to change behavior (intention), preferences, capabilities, and life situations, and their environmental and social circumstances vary. Each of these behavioral determinants either support or hinder change. In addition, the opportune moments to engage in behavior change activities differ between individuals, which calls for just-in-time adaptive interventions [7], [13].

Consequently, there are several aspects to consider when personalizing DHBCIs, including the a) identification of appropriate behavior change objectives and activities and the behavioral determinants to be targeted (i.e., personalization of the behavior change plan); b) adaptation of the selected objectives and activities based on individuals' adherence and the effectiveness of the activities; c) identification of appropriate educational, motivational, or feedback messages; d) identification of the opportune moments to deliver messages and prompts; and e) adaptation of the tone or style of interaction according to individuals' preferences and personalities.

As proposed earlier by Honka et al. [14], this type of extensive personalization requires the instantiation of a comprehensive user model, the so-called *virtual individual (VI) model*, which defines all the relevant knowledge constituents for intervention personalization. The VI model should cover the theoretical constructs of behavior change [10], [11], [12], since they define the behavioral determinants to be considered when personalizing DHBCIs, and thus, facilitate the identification of appropriate behavior change techniques (BCTs) [15], [16]. In this study, the VI model concept is further developed

by proposing a theory-based framework of user features that support the implementation of extensively personalized DHBCIs.

In addition to the VI model development, we introduce a prototype health recommender system (HRS), using a standard recommendation approach, which recommends behavior change objectives and activities from various behavioral domains with the aim of promoting well-being and preventing lifestyle-related diseases. HRSs have been introduced as a promising solution for personalizing DHBCIs [5], [17], [18], but the existing applications rarely consider well-being from a multi-domain perspective (see Section II. Related work). In addition, the current HRS realizes a selection of the VI model features that we consider sufficient for serving the minimum requirements for personalizing multi-domain DHBCIs. The selected user model features are related to well-being, lifestyle, and behavioral intention. We study the impact of this minimum set of features on the performance of the implemented multi-domain HRS. Typically, HRS research focuses on the development of recommendation methods, although both the underlying user model and the applied recommendation method contribute to the suitability of recommendations. This study focuses on the user modelling aspect by providing baseline results for finding the most effective user features for personalization. Disease management systems are beyond the scope of the study.

## II. RELATED WORK

A majority of HRSs that focus on promoting well-being and healthy lifestyle provide recommendations for physical activity (PA) or healthy diet or deliver tailored motivational messages for smoking cessation [17], [18], [19], [20]. Some have also addressed alcohol consumption [21], mental well-being [18], or sleep [22]. HRSs have been used for recommending personalized goals, healthy activities, peer support, and reliable health information and for selecting educational and motivational health messages or the appropriate timing for message delivery [5], [18], [20], [23]. For example, PA-related HRSs have recommended PA modes (e.g., running, walking, gym), intensities, or durations [24], [25], [26], [27]; personalized goals in terms of weight loss and calorie expenditure [24], [25]; [24], [25], [26], [27]; exercise buddies [27]; and suitable timings or places for exercise sessions [24], [26], [27]. Diet-related HRSs have recommended recipes, meal plans, restaurants, and healthy items from restaurant menus, and they have provided also nutritional advice [18]. For stress management, HRSs have recommended mental activities such as mindfulness, breathing, and cognitive exercises [28], [29] or activities related to PA, social engagement, and enjoyment [30].

Typical recommendation methods include content-based, collaborative, demographic, and knowledge-based filtering as well as hybrid approaches [23], [31], [32]. All of these methods have been employed also in HRSs that promote well-being and healthy lifestyle [5], [17], [18]. In content-based filtering, items that are similar to those rated positively by

the user are recommended. In collaborative filtering, items that have been evaluated highly by other users sharing similar item preferences with the target user are recommended, whereas in demographic filtering, items preferred by other users sharing a similar demographic profile with the target user are recommended. In the knowledge-based approach, explicit knowledge about the user, derived, for example, from questionnaires or wearable devices, is used to filter suitable items. Cheung et al. [5] consider knowledge-based filtering especially appropriate for HRSs, and many of the implementations to date are based on this method [18]. In hybrid approaches, different recommendation methods are used. The majority of HRSs utilize hybrid methods [5], [18] such as in [30], [33], and [34]. In addition, both supervised and unsupervised machine learning have been utilized in HRSs [5], including random forests [28], reinforcement learning [24], [35], and neural networks [26].

Utilizing the methods of recommender systems for personalizing DHBCIs is appealing: Content- and knowledge-based filtering can efficiently generalize to a high number of user features compared to the traditional rule-based tailoring without considerably increasing the complexity of the system [36]. Furthermore, the combination of collaborative and demographic filtering can be used to collect the preferences of a group of people who share similar well-being issues and life situations, which can be used to recommend novel intervention items to a specific individual [5]. Hence, in terms of personalization, HRSs have the potential to consider well-being from a multidimensional viewpoint and harness the multitude of individual-specific factors that determine behavior for personalization.

However, to the best of our knowledge, HRSs based on such comprehensive user models have not been implemented. Typically, the user models have focused on a limited set of behavioral domains, often PA or dietary habits [17], [18], [20], and they do not cover any of the theory-based determinants of behavior [5], [37]. Some HRSs address one or two behavioral determinants. For instance, in [33], [38], and [39], smoking cessation messages are personalized according to the readiness to change construct. In [30], users' self-efficacy (i.e., belief in one's capability to perform the behavior under different circumstances [40]) and skills are leveraged to personalize stress management activities. A rare example of extensive theory-based personalization is provided by the smoking cessation application, Quit and Return [41], which addresses several constructs of the Integrated-Change Model (attitude, readiness to quit, self-efficacy, social support, action planning, and skills) [42]. Overall, examples of HRSs that are firmly grounded on behavioral theories are limited. The lack of multi-domain interventions and the insufficient consideration of behavioral determinants are major shortcomings for HRSs that aim to engage individuals in healthy lifestyle changes.

Furthermore, the user model features vary considerably across different HRSs, indicating a lack of common understanding of the important features. The typical

(non-theoretical) feature types that have been used for personalization include basic demographics, such as age and gender (e.g., [25], [27], [33], [35], [39], [43], [44], [45]), health risks (e.g., [25], [27], [33], [43], [46]), and health behaviors (e.g., [20], [22], [24], [25], [27], [33], [35], [39], [43], [44], [45]). In addition, many HRSs utilize context-related features to determine opportune moments for delivering recommendations, of which location and the time of day are the most prevalent (e.g., [24], [26], [27], [28], [33]), but calendar availability [27], [28] and users' momentary activities [24], [26] have also been used in some examples. Considering user preferences (e.g., preferred PA modalities and time slots, dietary restrictions) [20], [25], [26], [27], [44], [45] and the usefulness or effectiveness of recommendations (either user-evaluated or inferred) [24], [28], [30], [35], [38], [43] are also quite common. Some HRSs consider mental states (e.g., stress level, mood) [28], [29], [47], social ties [27], [33], environmental conditions [22], [47], or personality traits [28].

### III. OBJECTIVES

This study contributes to the development of personalized DHBCIs that promote well-being and prevent lifestyle-related diseases by guiding and empowering individuals to make healthy lifestyle changes. First, a comprehensive, theory-based VI model framework is introduced with practical user feature examples. The framework includes features that represent the psychological, social, and environmental factors determining behavior in the context of everyday life, and it considers well-being and healthy lifestyle from a multi-domain viewpoint. After defining the VI model, we describe the development of a prototype web-based HRS, called With-Me HRS, which implements a subset of the VI model features for personalizing the recommendation of behavior change objectives and activities. When generating the recommendations, several behavioral domains are considered, as opposed to most HRSs that have a restricted focus. Finally, we evaluate the preliminary validity and usefulness of With-Me HRS in a real-life remote health-coaching program.

The present work aims to advance a common understanding of the user features required for the extensive personalization of DHBCIs, which is currently lacking especially in the HRS research field. Furthermore, an example of a HRS that considers well-being and healthy lifestyle comprehensively, beyond only PA and dietary habits, is introduced. This kind of multi-domain interventions are novel in the HRS literature, and the current study provides baseline results regarding the personalization of such interventions.

### IV. METHODS: IMPLEMENTATION AND EVALUATION

#### A. VIRTUAL INDIVIDUAL MODEL

To define the key constituents of the comprehensive VI model, we sought to identify various factors governing behavior and behavior change from the theories explaining health behavior. Many of the theories have overlapping constructs, but behavioral scientists have attempted to reach a

consensus about the most important ones [10], [40]. Based on the comparisons of theories conducted by behavioral scientists [10], [40], [48] and a review into the fields of psychology, behavioral economics, and social marketing (e.g., [49], [50], [51], [52], [53]), Honka et al. [14] formed a synthesis of the key determinants of behavior. We utilized this synthesis to define the VI model constituents. In addition, the stage of change construct defined by the Transtheoretical Model (TTM) of behavior change [49] was included into the VI model, as it is widely used to explain the multistage process of change [54]. The stage of change construct describes one's readiness to change behavior (i.e., the behavioral intention or motivation). Finally, the principles of evidence-based intervention planning for health promotion [10], [55], [56] were considered when designing the VI model. Specifically, the following questions guided the selection of VI model constituents:

- 1) What are the risk behaviors to be addressed (e.g., unhealthy eating rhythm, insufficient sleep, lack of exercise)?
- 2) How motivated a person is to modify these behaviors (e.g., based on TTM [49])? Are they aware of the need to change behavior?
- 3) Which determinants of behavior should be addressed for increasing motivation and eliciting behavior change (e.g., outcome expectations/attitude, self-efficacy, social influence, perceived barriers, environmental context)? [11], [14]
- 4) What are the factors that facilitate or impede behavior change (e.g., time and monetary resources, personal skills, environmental or social factors)? [11], [14]
- 5) What motivates and interests the person? How should intervention materials and messages be framed to increase motivation towards behavior change (e.g., elicit emotions vs. stick to facts, negative vs. positive framing [57], [58], [59])?
- 6) What are the opportune moments to provide support? [7]
- 7) What kind of behavior change techniques [16], [60] and activities are effective for the person?

To provide answers to the open questions, we identified four key, high-level elements that form the core of the VI model: *Health & well-being*, *Resources*, *Motives & preferences*, and *Behavior change needs and determinants*. These factors determine one's behavior change needs, the type of support needed, and personal interests and preferences, and they should be used to personalize the intervention content. Furthermore, we included an element describing the *Momentary context* to facilitate the identification of opportune moments for providing support. We also included *Intervention items* and *Progress evaluation* elements; the former describes the content of the personalized intervention, and the latter tracks the person's adherence to the intervention and the effectiveness of the intervention. Progress evaluation is important for identifying whether the intervention should be updated. Fig. 1 presents the VI model elements and the related

**TABLE 1. Behavioral domains supported by With-Me HRS.**

| Health behavior and well-being domains |                         |
|--|-------------------------|
| Sleep sufficiency                      | Workload management     |
| Sleep quality                          | Recovery from stress    |
| Eating rhythm                          | Anxiety                 |
| Diet quality                           | Personal values         |
| Emotional eating                       | Quality of relationship |
| Physical activity                      | Self-esteem             |
| Alcohol consumption                    | Smoking                 |

feature types. In the figure, two additional blocks are visible: intervention items appropriate for other individuals similar to the target person and an intervention library defining the available items to select from. These blocks are not part of the VI model, though closely related, as data from similar individuals can provide added value for intervention personalization (via collaborative and demographic filtering) and the intervention library defines the space for personalization.

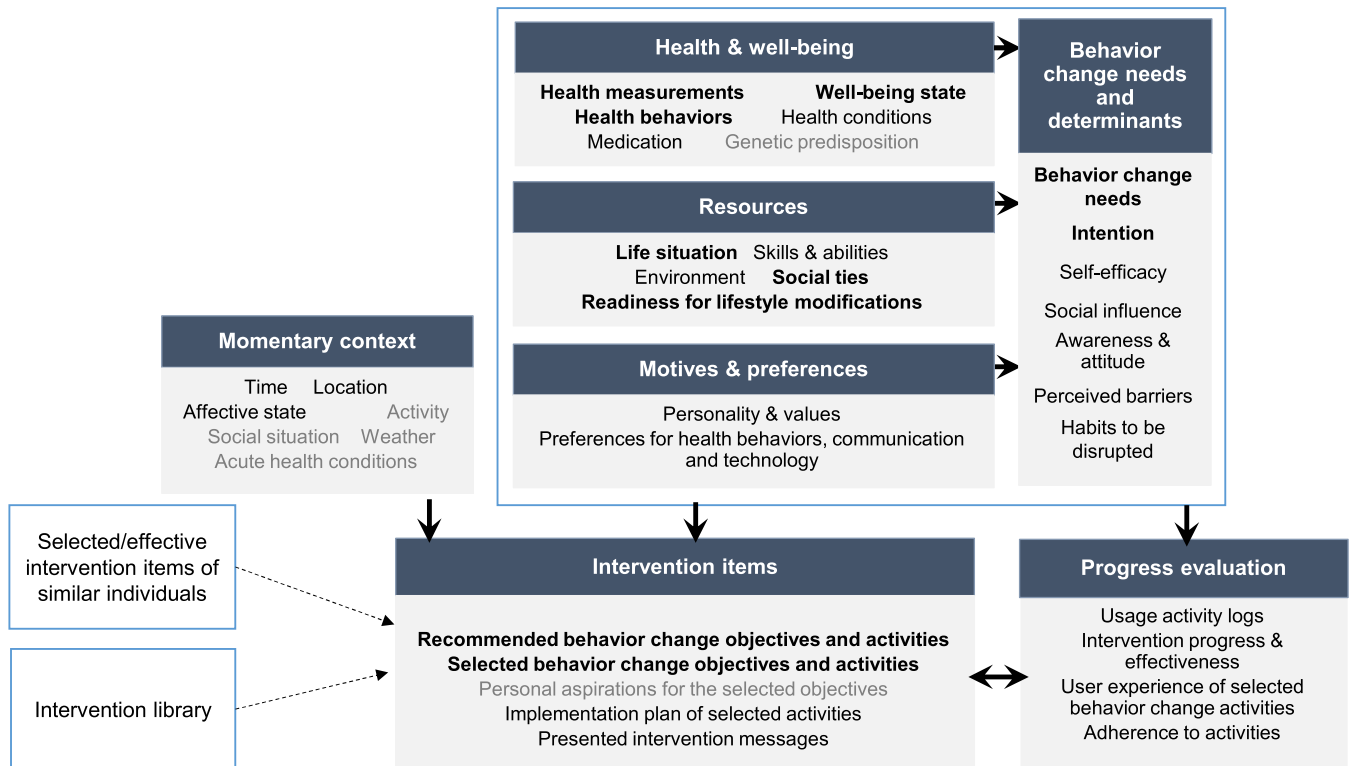
When populating the VI model elements with an individual's data, a digital representation of the individual, the *personal profile*, is formed. Detailed descriptions of the proposed VI model elements, the proposed feature types along with concrete feature examples, and the interrelations between the features are provided in Appendix 1.

## B. WITH-ME HEALTH RECOMMENDER SYSTEM

### 1) OVERVIEW

With-Me HRS was developed to provide support in identifying appropriate coaching plans for the participants of an occupational stress management program that involved human coaching. It was designed to evaluate individuals' behavior change needs in terms of 14 behavioral domains related to well-being and healthy lifestyle (see Table 1) as well as to recommend suitable behavior change activities based on the identified needs and domain-specific readiness to change. It assisted health coaches in identifying suitable behavior change objectives and activities (i.e., coaching tasks) for the participants by providing a comprehensive overview of the analyzed behavioral domains and recommending activities accordingly.

The VI model feature types relevant to behavior change needs and readiness to change were selected for implementation in the With-Me user model (bolded in Fig. 1). We consider these aspects as the two most important feature types for the user models of multi-domain DHBCIs, since the first, obvious step in such interventions is to identify the appropriate behavior change objectives for an individual [60], and readiness to change appears to be the single best predictor for behavior [11]. In addition, feature types from the VI model's Resources element (Fig. 1) were implemented to reflect the characteristics of the stress management program's target population, consisting of individuals who were active in work-life and lived with a family.



**FIGURE 1.** The high-level elements of the proposed VI model including the related feature types. The feature types that were included in the implementation of With-Me HRS are bolded, and the features that we consider promising for personalization but lacking a solid theoretical background or proven practical value are grayed out. The intervention item components are closely associated with the VI model but not part of it.

With-Me HRS was implemented as a web tool. It was integrated with the Movendos web-based health-coaching service (v1.27, Movendos Ltd.) [61] and the LimeSurvey online survey tool.<sup>1</sup> Together, these modules formed a digital health-coaching system. The content of the stress management program and the functionalities of the overall coaching system are described in [62]. In this study, we focus on the implementation of the HRS module only. Fig. 2 depicts the technical architecture of With-Me HRS and its connections to the other modules of the overall coaching system. The HRS was composed of *Personal profile*, *Profiler*, *Recommendation engine*, and *Intervention library* components.

The Personal profile included a user model that was associated with a database that populated the model’s features with an individual’s past and current data. The user model specified the features utilized for personalization and the structure of user data. The Profiler component analyzed the available data and created and maintained the Personal profile according to the data structure specified by the user model (see subsection Profiler below for details). The Personal profile provided a user-interface for coaches, which allowed coaches to examine the analysis results and to correct possible mistakes in the results. The data used for profiling were mostly collected with the online survey tool. In addition, objective indicators of physiological well-being and physical activity

were provided by Firstbeat lifestyle assessment (Firstbeat Technologies Ltd.),<sup>2</sup> and they were manually entered into the HRS. Firstbeat lifestyle assessment is based on the analysis of heart rate variability and movement that are measured via chest electrodes.

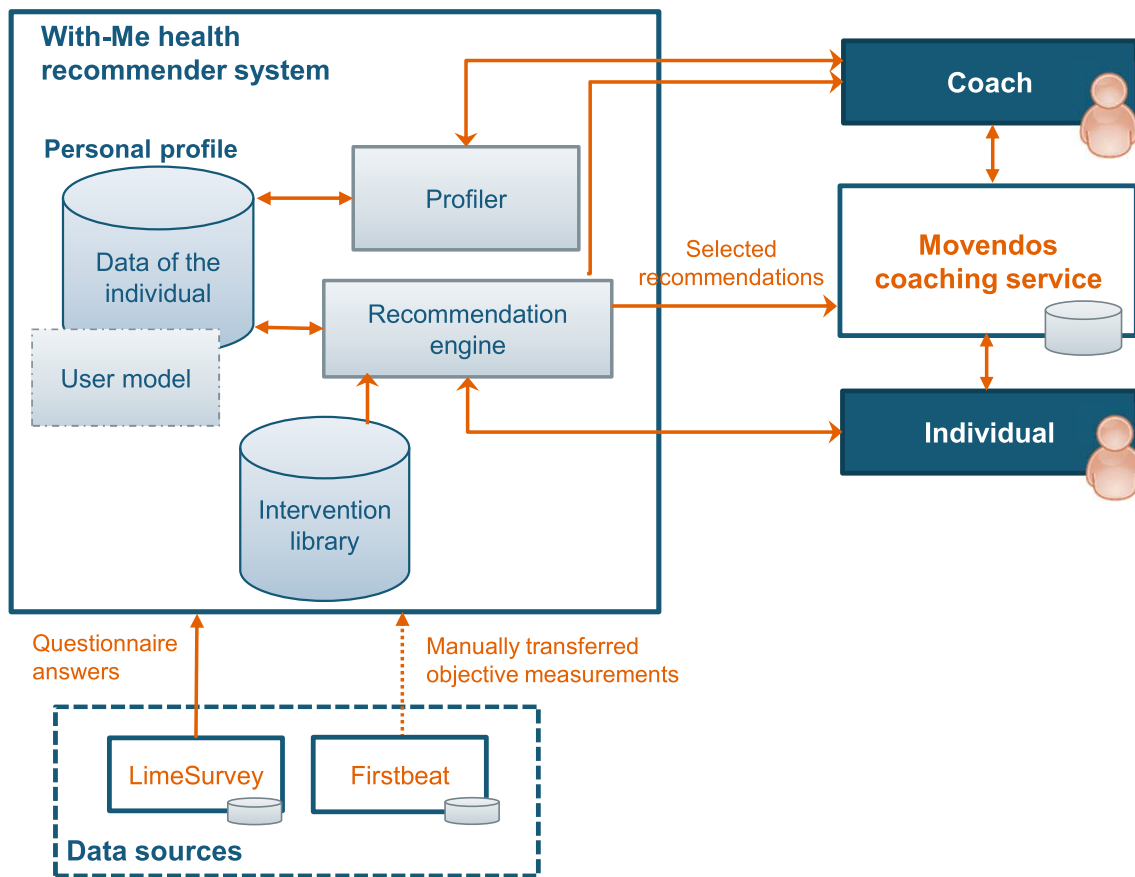
Based on the constructed Personal profile, the Recommendation engine suggested behavior change activities from the Intervention library (see subsection Recommendation engine below for details). Only the most recent user data were used for recommendations. The Recommendation engine provided a user-interface for both coaches and individuals for presenting the recommended activities and enabling coaching task selection. The reference ids of the recommended and selected activities were stored in the Personal profile. Information about the selected activities was also transferred to the Movendos health-coaching service.

2) PROFILER  
a: PERSONAL PROFILE

The Profiler populated the user model underlying With-Me HRS, thus forming the Personal profile. The user model covered a subset of the feature types included in the envisioned comprehensive VI model (Fig. 1): well-being state, health behaviors, health measurements, behavior change

<sup>1</sup>www.limesurvey.org

<sup>2</sup><https://www.firstbeat.com/en/wellness-services/wellness-professionals/individual-wellbeing/>



**FIGURE 2.** The architecture of With-Me HRS and its connections to the other modules of the overall digital health-coaching system that was utilized in the occupational stress management program described in [62].

needs, readiness to change (intention), life situation, social ties, and the reference ids of the recommended and selected items (see Appendix 1 for details). The Profiler analyzed data acquired via questionnaires (e.g., WorkOptimum for occupational health [63], Cognitive Fusion Questionnaire for anxiety assessment [64], and a modified version of the stages of change survey [65]) and, when available, via the Firstbeat lifestyle assessment conducted based on a 3-day measurement period. Based on the available data, the Profiler interpreted participant’s behavior change needs and readiness to change regarding each of the 14 behavioral domains listed in Table 1.

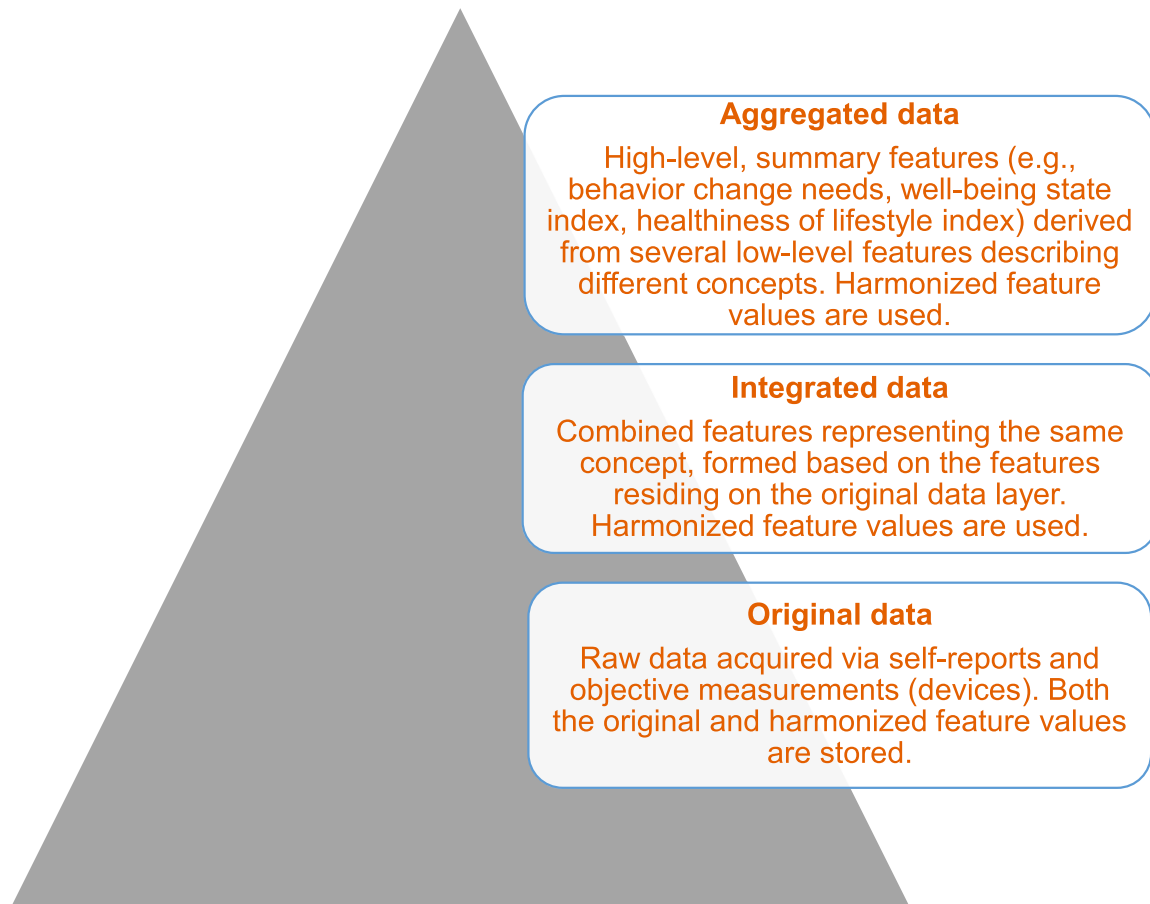
The coaches could review the results of the Profiler’s behavior change needs analysis via its user-interface. For each behavioral domain, the individual’s need for change (5-point scale: 1 = no need, 5 = strong need) and the readiness to change were presented. Readiness to change was categorized according to the TTM’s stage of change construct (pre-contemplation, contemplation, preparation, action, maintenance [49]). The behavioral domains were presented in the order of importance by ranking them according to the behavior change need. In addition, the user-interface revealed per behavioral domain the original user data that were processed by the Profiler, i.e., the participants’ self-reported

values and Firstbeat indicators. The domains for which the Profiler was not able to assess the change need with high confidence were denoted with a warning sign to urge the coach to check the data behind the analysis and to modify the results if needed. Low confidence could be caused, for instance, by conflicting self-report and Firstbeat indicator values (see Appendix 2 for details).

*b: USER MODEL’S DATA STRUCTURE*

In addition to defining the features of the Personal profile, the With-Me user model specified the hierarchical structure of the features, their interrelations, and the common properties used to describe them. We implemented the hierarchical structure via three data layers: *original*, *integrated*, and *aggregated data*. The original data layer included original measures, provided directly by the available data sources (the participant or measurement device) and formed the bottom level of the hierarchy. The integrated data layer combined features representing similar concepts, and the aggregated data layer combined features describing different concepts into high-level summary features (Fig. 3).

The following properties were used to describe the features residing on the different layers: *timestamp* indicating when the value of a feature was acquired, *original value*



**FIGURE 3.** The hierarchical data layers of the With-Me user model.

of the feature (available only on the original data layer), *harmonized value* transforming the original feature value to a unified 5-point scale, *confidence* indicating the reliability of the feature value via a continuous scale from 0 to 1 (1 = highest reliability), and *source* denoting the origin of the feature value (participant's self-report, Firstbeat assessment, Profiler's analysis, or coach's modification). We used harmonized values to simplify the feature computations at the higher layers of data hierarchy and confidence values to determine the reliability of the Profiler's analysis results. In Appendix 2, we describe the data layers in more detail and the related data-processing algorithms executed by the Profiler.

### 3) RECOMMENDATION ENGINE

#### *a: USER-INTERFACE*

The Recommendation engine recommended behavior change activities (items) from the Intervention library based on the identified behavior change needs and readiness to change behavior, which were analyzed by the Profiler (but could be modified by the coach). The user-interface of the Recommendation engine presented the recommended activities and the items of the Intervention library to both the coaches and participants. In addition, the participants could propose

at most three activities to their coaches to be included in their coaching plans, either from the recommended list of items or from the Intervention library, or alternatively, they could create custom activities. The coaches were able to view the proposed activities through the user-interface. The number of activities that could be proposed was limited to three, since for multi-domain behavioral interventions, including 2-3 behavior change objectives seems to be optimal in terms of intervention efficacy [66].

#### *b: INTERVENTION LIBRARY*

The Intervention library included over 100 items related to different behavior change activities. Each item was labelled by the behavioral domains it was supposed to target and the TTM's stages of change it was applicable to. A professional health coach was involved in designing the Intervention library. The activities were based on different behavior change techniques (BCTs) [16], and activities of varied difficulty or effort levels were included. Many of the activities utilized the Oiva web portal,<sup>3</sup> developed to promote mental well-being, which included short exercises based on the acceptance and commitment therapy [67]. Examples of the

<sup>3</sup><https://oivamieli.fi/>

Intervention library items (and the related BCTs) include: “Read an online article about the symptoms of stress and good practices for stress management” (information about health consequences), “Take a quiz for evaluating your alcohol consumption patterns” (feedback on behavior), “Get an exercise buddy” (social support), “Use Oiva to ponder the reasons that are keeping you from fulfilling your personal values in everyday life” (pros and cons), “Make a realistic list of work tasks for the upcoming work day” (action planning), “Keep a diary about eating habits for three days” (self-monitoring), “Keep fruits in sight and vegetables easily accessible at home” (restructuring the physical environment), “Wake up at the same time every day” (habit formation), and “Practice mindfulness skills with Oiva exercises” (behavioral practice).

### c: RECOMMENDATION LOGIC

With-Me HRS utilized the case-based recommendation technique of knowledge-based filtering [68], where items (or cases) that matched the behavior change needs and readiness to change of a participant (i.e., the target case) were retrieved from the Intervention library. Participants’ behavior change objectives were determined by the behaviors that they had at least a moderate need for change. Only activities relevant to the objectives were considered for recommendation as explained in the following paragraphs.

Let us denote  $B$  as the set of 14 behaviors supported by With-Me HRS (see Table 2 in Appendix 2),  $T = \{1, 2, 3, 4, 5\}$  as the set of the TTM’s stages of change (1 = pre-contemplation, 5 = action), and  $I$  as the set of items included in the Intervention library. Each behavior was described in the Personal profile with a vector  $\mathbf{b}^i = [b_{str}^i \ b_{stg}^i]$ , where for a behavior  $i \in B$ ,  $b_{str}^i \in [0, 1]$  denotes the strength of the behavior change need (0 = no need, 1 = strong need) and  $b_{stg}^i \in T$  denotes the stage of change for the behavior. Furthermore, each activity  $j \in I$  was described in the Intervention library with the set of properties  $A^j = \{A_{beh}^j, A_{stg}^j, opr\_A^j\}$ , where

- 1)  $(A_{beh}^j \subset B, \leq)$  is a partially ordered subset of  $B$  including only those behaviors that are in the focus of activity  $j$ , ordered based on relevancy,
- 2)  $A_{stg}^j \subset T$  is a subset of  $T$ , indicating the stages of change to which activity  $j$  is applicable,
- 3)  $opr\_A^j \in \{\max, \min, \text{weighted}\}$  is an operator determining how to evaluate the combined relevancy of the set of behaviors  $A_{beh}^j$  in terms of the Personal profile: either all the behaviors  $a_{beh,n}^j \in A_{beh}^j, n \in \{1, \dots, 14\}$  need to match the Personal profile (max); at least one of them needs to match (min); or the relevancy of each behavior is weighted as such that the first item  $a_{beh,1}^j$  in the set matters the most and the other behaviors have a supporting role only (weighted).

For example, let us assume that the HRS supports only three behaviors: 1) physical activity, 2) relaxation, and 3) sleep. Let the Intervention library include an activity  $j$  that is suitable for individuals who have challenges regarding

relaxation or sleep and who are in the pre-contemplation, contemplation, or preparation stage. Thus, the activity  $j$  has the properties  $A_{beh}^j = (2, 3)$ ,  $A_{stg}^j = \{1, 2, 3\}$ , and  $opr\_A^j = \min$ . Furthermore, we introduce an individual  $\mathbf{P}$  with the profile

$$\mathbf{P} = \begin{bmatrix} \mathbf{b}^1 \\ \mathbf{b}^2 \\ \mathbf{b}^3 \end{bmatrix} = \begin{bmatrix} 0.76 & 2 \\ 0.9 & 5 \\ 0.25 & 4 \end{bmatrix}, \quad (1)$$

which we use as an example for demonstrating the recommendation logic.

The recommendation logic was based on two similarity metrics,  $\text{sim\_need}^j \in [0, 1]$  and  $\text{sim\_stage}^j \in [0, 1]$ , which together determined the suitability,  $\text{sim\_total}^j \in [0, 1]$ , of an activity  $j$  for recommendation (0 = low, 1 = high similarity). The  $\text{sim\_need}^j$  metric described the similarity between the behaviors related to activity  $j$  and the behavior change needs identified in the Personal profile. The metric was based on the Manhattan distance between value pairs  $\{(b_{str}^i, 1) \mid i \in A_{beh}^j\}$ . Thus, only the behaviors relevant to activity  $j$  were considered. The general formula for the metric is

$$\text{sim\_need}^j = 1 - \left( opr\_A^j_{i \in A_{beh}^j} \left| b_{str}^i - 1 \right| \right), \quad (2)$$

where the operator  $opr\_A^j$  determines how to combine the distances. If  $opr\_A^j = \text{weighted}$ , a weighted normalized Manhattan distance was computed with weights set according to the relevance of the behaviors denoted by the ordered set  $(A_{beh}^j, \leq)$ . In our example, the similarity between the behavior change needs of profile  $\mathbf{P}$  and the behaviors relevant to activity  $j$  is computed as

$$\text{sim\_need}^j = 1 - \min(|0.9 - 1|, |0.25 - 1|) = 0.9. \quad (3)$$

The  $\text{sim\_stage}^j$  metric described the similarity between the stages of change that activity  $j$  was applicable to and the set  $T^j = \{b_{stg}^i \mid i \in A_{beh}^j\}$ .  $T^j$  included the Personal profile’s stages of change that corresponded to the behaviors relevant to activity  $j$ . If  $\exists t^j \in T^j$  such that  $t^j \in A_{stg}^j$ , then  $\text{sim\_stage}^j = 1$  (i.e., at least one matching stage was found). Otherwise, the closest values in both sets  $t^{j*} \in T^j$  and  $a_{stg}^{j*} \in A_{stg}^j$  were identified, and the similarity between them was computed as

$$\text{sim\_stage}^j = 1 - \frac{1}{4} \left| t^{j*} - a_{stg}^{j*} \right|. \quad (4)$$

In our example case,  $T^j = \{4, 5\}$ . Hence,  $T^j$  does not share any common elements with  $A_{stg}^j$ . The closest values in the two sets are  $t^{j*} = 4$  and  $a_{stg}^{j*} = 3$ . Now,

$$\text{sim\_stage}^j = 1 - 0.25 \times |4 - 3| = 0.75. \quad (5)$$

Finally, the overall suitability for activity  $j$  was computed as

$$\begin{aligned} \text{sim\_total}^j &= 0.5 \times (\text{sim\_need}^j + \text{sim\_stage}^j), \\ &\text{if } \text{sim\_need}^j \geq 0.5, \\ &\text{otherwise } \text{sim\_total}^j = 0. \end{aligned} \quad (6)$$



**TABLE 2.** Key steps of the recommendation logic.

| Step # | Action   |
|--------|--|
| 1      | For each activity $j$ , regarding the behaviors it focuses on, compute <ol style="list-style-type: none"> <li>the similarity between the behaviors and the Personal profile's behavior change needs <math>\rightarrow \text{sim\_need}^j \in [0,1]</math>,</li> <li>the similarity between the stages of change relevant to activity <math>j</math> and the Personal profile's readiness to change <math>\rightarrow \text{sim\_stage}^j \in [0,1]</math>,</li> <li>total similarity: if <math>\text{sim\_need}^j \geq 0.5</math>, <math>\text{sim\_total}^j = \text{mean}(\text{sim\_need}^j, \text{sim\_stage}^j)</math>, otherwise <math>\text{sim\_total}^j = 0</math>.</li> </ol> |
| 2      | Preselect activities for which $\text{sim\_total} \geq 0.5$ .  |
| 3      | Randomize the order of the preselected activities, and sort them according to $\text{sim\_total}$ .  |
| 4      | Recommend the top-20 items from the ordered list of activities.  |

Thus, only the activities that targeted the behaviors for which the individual had at least a moderate need for change were considered potentially suitable for recommendation. In our example,

$$\text{sim\_total}^j = 0.5 \times (0.9 + 0.75) = 0.825. \quad (7)$$

The  $\text{sim\_total}^j$  metric was computed for each activity  $j \in I$ , and the activities for which  $\text{sim\_total}^j \geq 0.5$  were preselected for recommendation. The order of the preselected items was randomized, after which they were sorted in descending order based on  $\text{sim\_total}$ . Mixing the items ensured that activities addressing different behaviors were included at the top of the ordered list. Finally, the top-20 activities were selected for recommendation. The key steps of the recommendation logic are summarized in Table 2.

### C. EVALUATION STUDY

The validity and usefulness of With-Me HRS were studied as a secondary objective of a pilot randomized controlled trial (RCT) [62], where technology-assisted and traditional telephone coaching for occupational stress management were compared in terms of intervention effectiveness and the time use of health coaches. The study was approved by the Ethics Committee of Human Sciences at the University of Oulu, Finland. Informed consent was obtained by regular mail from the individuals interested to participate in the study.

#### 1) PARTICIPANTS AND PROCEDURES

Altogether 50 participants were recruited, who worked full-time (in the areas of information technology; education; culture; social, health, and customer services), reported a decreased state of well-being, lived in a relationship, and were motivated to enhance their well-being by making lifestyle changes or doing exercises related to mental well-being. The participants were recruited among the employees of the City of Oulu, Finland, most of whom worked in

female-dominant occupations (e.g., teachers, nurses, social workers, etc.). Nearly all eligible participants were female (96.0%, 48/50), and their mean age was 46.40 years (SD 9.67). The participants were randomly allocated to two groups: one receiving technology-assisted health coaching via telephone ( $N = 25$ ) and the other receiving traditional telephone coaching ( $N = 25$ ). In terms of the scope of this paper, the relevant difference between the two groups was related to the usage of With-Me HRS in supporting the first two coaching calls. In technology-assisted coaching, health coaches utilized the HRS to define participants' initial coaching plans (*group with visible recommendations*), whereas in traditional coaching, the HRS generated recommendations, but they were not utilized in decision-making (*group with hidden recommendations*). Three health coaches were involved in the study, each having an equal number of participants from both groups. Further details regarding the participants and the study design are presented in [62].

At the beginning of the intervention, both groups answered an online questionnaire regarding well-being, health behaviors, and readiness to modify behaviors. The WorkOptimum assessment for occupational health [63] was part of the questionnaire. In addition, the group with visible recommendations conducted the 3-day measurements related to the Firstbeat lifestyle assessment. Based on the questionnaire answers and the selected Firstbeat indicators (available only for the group with visible recommendations), the HRS's Profiler component analyzed participants' behavior change needs and readiness to change (as described in Section IV.B).

For the group with visible recommendations, the coaches prepared for the first coaching call by exploring participants' results regarding Profiler's behavior change needs analysis (via its user-interface) and Firstbeat lifestyle assessment (in a portable document format, PDF). The Firstbeat assessment results were provided also to the participants before the first coaching call. During the call, participants' behavior change needs were discussed, and a high-level behavior change objective was agreed upon (e.g., sleep better, manage workload, eat healthier). The coaches also instructed the participants to preselect one to three behavior change activities from the HRS as their preferred coaching tasks before the next coaching call, which was scheduled after two weeks. The activities could be selected either from the recommended list of items or the Intervention library, or the participants could create custom activities. The coaches were asked to make corrections to the Profiler's needs analysis immediately after the first coaching call, in case they found any inconsistencies between the analysis results and their discussions with the participants to ensure that the HRS's recommendations were up to date before the participants were exposed to them. During the second coaching call, the coaching tasks preselected by a participant were either confirmed by the coach or adjusted in mutual agreement. The agreed tasks formed the initial coaching plan for the participant.

For the group with hidden recommendations, the coaches did not utilize Profiler's needs analysis when preparing for

the first coaching call. Instead, they received the results of the WorkOptimum questionnaire in a PDF report. The report was also provided to the participants before the first coaching call. During the call, participants' behavior change needs were discussed. In addition, on the contrary to the other group, the initial coaching plan, including the behavior change objectives and coaching tasks, was already set during the first call. With-Me HRS did not influence the decision-making, as neither the coaches nor the participants examined its outputs when making the coaching plan. However, immediately after the coaching call (and after the coaching plan was set), the coaches were asked to review the results of the Profiler's needs analysis so that the generated recommendations could be validated with all the participants, not limited only to the group with visible recommendations.

## 2) MATERIALS AND OUTCOME MEASURES

The evaluation study aimed to assess the preliminary validity and usefulness of With-Me HRS. The primary outcome for validity was the proportion of participants for whom recommended activities were included in the coaching plan. We also examined the proportion of participants (for the group with visible recommendations) who preselected activities from the recommended list of items as their preferred coaching tasks. In addition, we examined the number and type of changes made by the coaches to the results of the Profiler's behavior change needs analysis to understand whether the employed profiling algorithms included systematic flaws. The usefulness of With-Me HRS was studied by assessing the ease of coaching from the perspective of coaches and the quality of coaching from the participant viewpoint.

The validity of the HRS was evaluated based on coaches' self-reports and the information stored in the Personal profile database. For each of the 14 behavioral domains (Table 1), immediately after the (first) coaching call, the coaches were asked to record on a paper form a) whether they were able to evaluate the domain (need and readiness to change) based on the discussion they had with a participant, b) whether they made modifications to the Personal profile regarding the domain, and c) justifications for the modifications. In addition, the coaches were asked to write down the coaching tasks included in the participant's coaching plan (after the first or second coaching call depending on the group). From the database, metrics were retrieved regarding the changes made by the coaches to the results of the Profiler's needs analysis, the activities recommended by the Recommendation engine, and for the group with visible recommendations, the activities preselected by the participants.

The usefulness of the HRS from coaches' perspective was evaluated with the following two questionnaire items: (1) *"During the coaching call, it was easy to identify the behavior change needs and objectives for the client."* (ease of identifying participants' needs) and (2) *"During the coaching call, it was easy to identify suitable coaching tasks for the client."* (ease of identifying coaching tasks). The participants' opinions were collected with the following items: (1) *"My coach*

*understood my well-being related needs with ease."* (ease of explaining needs), (2) *"My coach helped me realize new areas for improvement that are important for my well-being."* (improved self-awareness of needs), and (3) *"I am satisfied with the coaching call(s)."* (satisfaction with coaching calls). Each item was measured on a 5-point Likert-scale (1 = completely disagree, 5 = completely agree). For the group with visible recommendations, the coaches assessed the ease of identifying participants' needs and coaching tasks immediately after the first and second coaching calls, respectively, whereas the participants provided their assessments after the second coaching call. For the group with hidden recommendations, all the assessments were conducted after the first coaching call.

## 3) DATA ANALYSIS

For assessing the validity of the recommendations, we considered only those participants for whom the coaches had reviewed the results of the Profiler's needs analysis and recorded the selected coaching tasks. We compared the recommendations to the selected coaching tasks but did not expect exact word-to-word matches, since coaches typically used much shorter names for the tasks than was used in the Intervention library's item descriptions. Therefore, for instance, "zumba two times a week" (coaching task) was matched with "I will start an exercise hobby" (recommended item), or "walking" (coaching task) was matched with "I will take 7000 steps per day" (recommended item). Furthermore, we excluded from the comparison five coaching tasks that were not part of the Intervention library, as our aim was to validate the recommendation algorithm, not the content of the Intervention library. To evaluate the changes made to Profiler's analysis results, we categorized them into three groups to describe the reasoning behind the changes: a) the participant's situation had changed after answering the online questionnaire utilized by the Profiler, b) the Profiler's profiling logic was suboptimal in terms of the input features or their weights (see Appendix 2), or c) the reason was unclear. The categorization was conducted based on the justifications provided by the coaches for the changes, the selected coaching tasks, participants' answers to the online questionnaire, and Firstbeat indicators (when available).

The usefulness of the HRS was evaluated by comparing the group-level medians of the coaches' and participants' self-assessments (coaches' ease of identifying participants' needs and coaching tasks; participants' ease of explaining needs, improved self-awareness of needs, and satisfaction with coaching calls) between the groups with visible and hidden recommendations. In addition to medians, the first (Q1) and fourth (Q4) quartiles of the self-assessments are reported. The Mann-Whitney U test was conducted to determine the statistical significance of between-group differences. The differences were considered statistically significant at an alpha level of 0.05. The Vargha-Delaney  $\hat{A}$  measure of stochastic superiority [69] is reported as an indicator of the between-group effect size coupled with the 95% confidence

interval (CI). The effect size computations were performed with the *r*companion package of the free R statistical software (version 4.0.5). The 95% CIs were computed using the bootstrap procedure (see e.g., [70]).

## V. EVALUATION RESULTS

### A. VALIDITY

Complete and valid data were available for 41 (out of 50) participants for assessing the validity of the recommendations. For 73% (30/41) of the participants, at least one of the recommended activities was included into the coaching plan. The proportion of participants with a recommended activity selected as a coaching task was higher for the group with visible recommendations (85% or 17/20) than for the group with hidden recommendations (62% or 13/21). However, also the number of coaching tasks was higher for the group with visible recommendations (median 3.0 tasks [Q1 2.8; Q4 3.0] vs. median 1.0 task [Q1 1.0; Q4 2.0]). Of the participants for whom two or more coaching tasks were defined, 53% (10/19) of the group with visible recommendations and 43% (3/7) of the group with hidden recommendations had at least two of the tasks selected from the recommended activities. Furthermore, the recommendations appeared highly suitable for the participants of the group with visible recommendations, as 90% (18/20) of them suggested to their coach to include at least one of the recommended activities in their coaching plans, and 50% (10/20) proposed to include three recommended items (the maximum number of items).

Regarding Profiler's behavior change needs analysis, the coaches reported modification needs for 21 (out of 50) participants in terms of 1 to 3 (out of 14) behavioral domains per participant. For 16 participants, modifications were required because of a changed life situation. For seven participants, some of the modification needs were due to faults in the profiling logic, and for five participants the reasons for the modifications were unclear. Most of the modifications due to participants' changed situations were related to increased readiness to change behavior (reported for 14 participants), and some were related to behavior change needs (reported for 7 participants). According to the coaches' notes, the coaching call had had a positive influence on the motivation to change behavior for many participants, which explains the modification needs regarding the readiness levels. In addition, a delay of one to two months took place between the participants' profiling questionnaire answers and in scheduling the first coaching call, which may have made part of the Profiler's analysis results outdated.

The coaches' notes revealed also some improvement needs for the profiling logic regarding physical activity (PA) and sleep: It appeared that the profiling logic gave too much weight on the short-term (3-day) PA levels, assessed via Firstbeat indicators, compared to the self-reported levels (evaluated for the past month). This resulted in incorrect inference about the PA needs of the participants who were usually inactive but temporarily increased their activity levels

during the Firstbeat measurement period. To infer the behavior change needs regarding sleep, separating sleep quality and sufficiency from each other was not sensible, as poor sleep quality had a direct impact on sleep sufficiency.

### B. USEFULNESS

For the coaches, it was considerably easier ( $\hat{A} = 0.71$ , 95% CI 0.59-0.84) to identify appropriate coaching tasks for the group with visible recommendations than for the group with hidden recommendations. However, the coaches' perceived effort for identifying participants' behavior change needs was similar for the two groups. According to participants' self-assessments, the group with visible recommendations was considerably more satisfied with coaches' abilities to understand their well-being related needs ( $\hat{A} = 0.71$ , 95% CI 0.57-0.83) and moderately more satisfied with the coaching call(s) ( $\hat{A} = 0.67$ , 95% CI 0.53-0.80) and coaches' abilities to make them realize new, personally relevant behavior change needs ( $\hat{A} = 0.69$ , 95% CI 0.55-0.80) than the group with hidden recommendations. Hence, With-Me HRS appeared to be useful in improving coaching quality from the participants' perspective. The details of the between-group differences regarding the usefulness of the HRS are provided in Table 3.

## VI. DISCUSSION

### A. PRINCIPAL FINDINGS

We proposed a comprehensive, theory-based framework, the virtual individual (VI) model, to support the extensive personalization of digital health behavior change interventions (DHBCIs) for promoting well-being. In addition, we implemented a prototype health recommender system, With-Me HRS, which recommended a personalized set of behavior change activities. The user model underlying the HRS implemented a subset of the VI model feature types, of which health behaviors, well-being state, health measurements, behavior change needs, and readiness to change were utilized for personalization. The HRS supported a multi-domain intervention by considering various behavioral domains related to well-being and healthy lifestyle, namely sleep, physical activity, eating habits, alcohol consumption, smoking, workload management, recovery from stress, anxiety, self-esteem, personal values, and quality of relationships.

According to the conducted evaluation study in the health-coaching context, the recommendations were suitable for the participants, and at least one of the recommended activities was included into the personal coaching plans (from a maximum of three activities) for more than 70% of the participants. The results regarding the usefulness of With-Me HRS in supporting coaches' work were mixed, as the HRS reduced coaches' perceived effort in identifying appropriate coaching tasks for participants, but not in identifying their behavior change needs. From the participants' perspective, the usefulness of the HRS was clear, as the participants for whom coaches could utilize the HRS in decision-making

**TABLE 3.** Between-group differences regarding the usefulness of With-Me HRS.

| Outcome                                 | Visible recommendations<br>Mdn (Q1;Q4) | Hidden recommendations<br>Mdn (Q1;Q4) | $U$    | $\hat{A}^a$ (95% CI) | $P$    |
|---|--|---------------------------------------|--------|----------------------|--------|
| <b>Coach</b>                            |  |                                       |        |                      |        |
| Ease of identifying participant's needs | 4.0 (3.0; 4.0)                         | 4.0 (3.50; 4.0)                       | 340.0  | 0.57 (0.41-0.72)     | .390   |
| Ease of identifying coaching tasks      | 4.0 <sup>b</sup> (4.0; 5.0)            | 4.0 (3.0; 4.0)                        | 182.0  | 0.71 (0.59-0.84)     | .005** |
| <b>Participant</b>                      |  |                                       |        |                      |        |
| Ease of explaining needs                | 5.0 (4.25; 5.0)                        | 4.0 (4.0; 5.0)                        | 174.0  | 0.71 (0.57-0.83)     | .004** |
| Improved self-awareness of needs        | 4.0 (4.0; 4.0)                         | 4.0 (3.0; 4.0)                        | 185.0  | 0.69 (0.55-0.80)     | .01*   |
| Satisfaction with coaching call(s)      | 5.0 (4.0; 5.0)                         | 4.0 (3.5; 5.0)                        | 195.50 | 0.67 (0.53-0.80)     | .023*  |

The items were assessed by either coaches or participants and measured using a 5-point Likert-scale (1 = completely disagree, 5 = completely agree). Unless otherwise stated,  $N=24$  and  $N=25$  for the groups with visible and hidden recommendations, respectively.

<sup>a</sup>Vargha-Delancy  $\hat{A}$  measure of stochastic superiority for effect size estimation. Limits for interpretation: 0.56 (small), 0.64 (medium), 0.71 (large) [69]

<sup>b</sup> $N=25$

\* $p < .05$ , \*\* $p < .01$

were more satisfied with the quality of coaching than the participants with hidden recommendations.

### B. RELEVANCY OF USER MODEL FEATURES IN PERSONALIZATION

In the past HRS research, the attempts to improve the performance of HRSs have mostly been focused on finding accurate recommendation techniques (e.g., [27], [30], [34], [44]), while user models have attracted less research interest, even though wisely chosen user features can increase the suitability of recommendations significantly, which is important for improved user engagement and a positive health impact. The VI model provides a common user model framework that serves different personalization goals by considering not only the health and behavior change needs of individuals, which are the most widely used features for personalization and, beyond doubt, the most important ones in terms of the expected health impact, but also various other factors that influence user engagement and intervention adherence. These factors enable to a) identify the right kind of support to be provided while considering users' preferences regarding alternative behavior change activities; b) identify the opportune moments for delivering support; c) associate the recommended behavior change activities with personally meaningful goals; and d) use persuasive message framing and the tone of communication that is perceived as pleasant and credible by the user.

Some of the proposed VI features that influence user engagement and adherence have been utilized in earlier HRSs. Of the context-related features, location and the time of day are the most widely used for determining the appropriate content to recommend and the opportune moments for recommendations, although other interesting features have also been used (e.g., current activity, affective state, weather, calendar availability) [24], [26], [27], [29], [33], [47]. The

time lag between receiving and reading messages has been used to infer the best time to disrupt a user [39]. In addition, user preferences regarding physical activity (PA) modes and food items have been used to personalize recommendations [25], [26], [27], [44], [45]. However, we could not find examples that attempted to make behavior change objectives personally meaningful or which personalized the tone of messages. Value-based personal aspirations and personality traits were included as features to the VI model to serve these purposes. Values are personal beliefs of desired end states that guide behavior and choices [71]. Therefore, aligning health behavior change objectives with one's values may be motivating. Furthermore, framing messages based on personality or values has been shown to increase the persuasiveness of messages [59], [72], [73].

The VI model features describing personal resources and the determinants of behavior change are highly relevant for determining the type of support to be provided, as these summarize the key constructs found in different behavioral theories [10], [11], [14]. These features enable the adaptation of recommendations to individuals' readiness and capabilities to change behavior. When a person is not motivated to change behavior despite a clear health need, the features can be used to select intervention items that raise awareness of one's behavior change needs and strengthen one's capabilities. In a few HRSs, readiness to change has been used for personalization [38], [39]. In addition, knowledge of the factors influencing readiness to change (self-efficacy, attitudes/outcome expectations, social influence) and the possible barriers preventing good intentions from translating into actions (e.g., environmental constraints, lack of skills, old habits to be disrupted) are required to increase motivation and provide appropriate support [14], [40], [51]. The Quit and Return mobile application for smoking cessation [41] is a rare example of a HRS

where various behavioral determinants are considered for personalization.

The progress evaluation features of the VI model facilitate the monitoring of individuals' adherence to recommendations and the effectiveness of the intervention. These features are useful for providing feedback to users, and more importantly, for the dynamic adaptation of intervention content. For instance, HRSs that focus on PA have adapted recommendations based on monitoring the effectiveness of past recommendations. In [35], users' PA levels were monitored after sending motivational messages, and effective message types were learned for each individual. In [43], the effectiveness of activities was determined by monitoring changes in health outcomes (blood pressure, body mass index, and waist circumference) across users. Then, those activities were recommended which appeared effective for the users sharing a similar demographic and health profile with the target individual.

We propose to include in the VI model features related to genetic predisposition as an experimental component. The idea of utilizing genetics for the personalization of health interventions is intriguing, as it might reveal which behavior change activities (e.g., dietary habits, exercise modes, sleep patterns) are most effective in reducing personal health risks of an individual. Some computer-tailored interventions already utilize genetic information, for instance, for personalizing exercise regimes or nutritional intake [74], [75]. However, genetic testing needs to become mainstream before the value of genetics in personalization can be appropriately studied.

### C. STRENGTHS AND LIMITATIONS OF WITH-ME HRS

Most of the earlier HRSs have focused on only a few health behaviors (PA or diet), which is insufficient for the prevention of lifestyle-related diseases. With-Me HRS took a comprehensive approach by acknowledging various behavioral domains that contribute to well-being and healthy lifestyle. However, modifying all the possible unhealthy habits at once is unrealistic [66], and the unhealthiness of behavior varies between different domains across individuals. Hence, before recommending actual behavior change activities, a high-level assessment of behavior change needs across different domains should be conducted, which ideally should result in a few selected behavior change objectives. With-Me HRS provides an example developed towards this direction. However, the HRS did not recommend activities in a similar detail as some previous examples have recommended, such as specific PA intensities or durations or certain food items and proportions to be included in meals [18], [25], [27]. Detailed recommendations were not crucial, since the usage context of the HRS involved human experts who could provide personal guidance during the coaching calls for performing the recommended activities. For a fully stand-alone HRS, recommending detailed activities would become more important.

In the HRS research, employed recommendation methods are often described, whereas details of the underlying user

models and the available items to be recommended are rarely provided, although the user model, recommendation method, and intervention library together determine the accuracy and suitability of the recommendations. Therefore, to accumulate knowledge of the most effective personalization techniques, details regarding all these three aspects should be reported. In the present work, we provide information on the user model features used for personalization, how the features are measured, and the algorithms used to process raw measurements into features (Appendix 2), in addition to describing the recommendation method and intervention library items. Other examples of detailed user model descriptions are provided in [25] and [43]. Regarding the intervention library, it is important to ensure that the available items are varied enough for catering to the needs of different individuals. In our case, a professional health coach was involved in designing the content of the intervention library, who ensured that the activities typically used in human-delivered health-coaching for the behavioral domains supported by With-Me HRS were included.

In the With-Me user model, a harmonized value scale was used to describe the values of well-being and health-related features. Using harmonized values, when possible, can simplify aggregated feature computations (e.g., in terms of behavior change needs) and add flexibility to the resulting personal profile by allowing data source independent analysis. For instance, when several alternative devices or questionnaires can be used to measure the same concept of interest, such as PA level or sleep duration, harmonized values allow switching the data source without having to modify the computation logic of higher-level features. We chose to use a 5-point scale for the harmonized values, since the behavior change needs, which was the most relevant aggregated feature type in With-Me HRS, were described with such a scale. The use of a more fine-grained scale was not considered to provide additional information value for recommending activities. However, for some other use cases, using a 5-point scale for harmonized values may compress the original data too much, and using a 7- or 10-point scale may be more appropriate.

With-Me HRS was designed to be used only at the beginning of the health-coaching program, which limits its usefulness as a stand-alone HRS for the long-term support of health behavior change. With-Me HRS was incapable of collecting user data actively on a regular basis and updating the recommendations accordingly. Although data updates were supported and they resulted in a new set of recommendations, the user model was unable to identify trends in the data, and past values were not considered in the recommendations. For stand-alone HRSs, the capability to dynamically adapt to individuals' evolving situations while also monitoring the effectiveness of the recommendations is imperative.

The With-Me user model implemented the proposed VI model only to a limited extent. Only the features that we considered the most important were implemented, namely, behavior change needs derived from the features describing well-being and health behaviors, and the readiness to

change different behaviors. Particularly, features related to self-efficacy and skills were not implemented, although they are among the important predictors of behavior change [40], and the intervention library items were not graded by effort level. In the usage context of With-Me HRS, this limitation did not pose problems, as coaches were available to guide the participants in performing the activities. However, for a stand-alone HRS designed for long-term behavior change support, recommending activities with gradually increasing effort levels that match individuals' self-efficacy and skills could be useful.

#### D. INTERPRETATION OF EVALUATION RESULTS

Rather than seeking for the most accurate recommendation method, which is common in the HRS research, the purpose of the present study was to examine how well a standard recommendation method, which utilizes a minimum set of user features for personalization that we consider important (behavior change needs and readiness to change), performs in the novel context of a multi-domain, real-life intervention. Behavior change activities recommended by With-Me HRS were included into the health-coaching plans of more than 70% of the participants, which we consider as a reasonably good result achieved with the limited user model, especially when only half of the participants (and their coaches) were exposed to the recommendations. This result provides a reference baseline for testing the influence of additional user model features on the validity of recommendations.

We wish to raise awareness of the importance of conducting empirical studies that focus on finding the most effective user features for personalization. It may be wise to conduct these studies with standard recommendation methods for better comparison. We chose to utilize knowledge-based filtering, since it allows to personalize recommendations based on the specific characteristics of an individual [5], [18], which is especially important in health and well-being applications [5], [36]. Indeed, knowledge-based filtering has been widely used in HRSs before [18], and it can be considered as one of the standard approaches to which more complex, hybrid recommendation methods are compared.

We cannot compare our evaluation results directly to previous work as the methods and study settings used to validate HRSs are versatile. In Table 4, the validation approaches of some recent HRSs are summarized. According to the review by De Croon et al. [18], the majority of validation studies have been conducted offline without the involvement of real users (e.g., via simulated or existing datasets), or via single-session user studies and surveys. Studies involving users who use HRSs "in the wild" are less common, which is considered a major challenge in the field [18]. In offline studies, standard error metrics (precision, accuracy, recall, F1-score, etc.) are commonly used to measure the performance of recommendation algorithms [18]. In real-life studies, these metrics are inconvenient because requiring users to rate all the available items for identifying the true negatives and positives would significantly increase user burden and

hamper the real-life setting. Instead, user satisfaction with the recommended items (e.g., [33], [38], [44]), self-reported or observed compliance to recommendations (e.g., [24], [30], [35]), and changes in health outcomes (e.g., [24], [38], [46]) have been reported for assessing the suitability of recommendations. In addition, user experience, perceived usefulness, and usability of HRSs are typically assessed, but with varying self-report scales or interview questions [18].

In the present study, we assessed the suitability of recommendations by monitoring the number of recommended behavior change activities that were selected to the participants' coaching plans. While this is a stronger indicator for suitability rather than merely measuring user satisfaction with recommendations, the most reliable approach for validation, however, would be to assess the impact of recommendations on participants' behavior, i.e., evaluate participants' adherence to the selected activities. As continuous monitoring of behavior was not implemented in With-Me HRS, we were not able to evaluate participants' actual adherence to the recommendations. Nevertheless, some indications of adherence may be inferred from the results of the related pilot RCT [62], which describes the outcomes of the health-coaching intervention where With-Me HRS was utilized as a technological component. According to the results, participants' self-reported diligence in performing the selected coaching tasks at the beginning of the intervention was slightly better in the group receiving (visible) recommendations compared to the participants who were not provided the opportunity to examine the recommendations (group with hidden recommendations).

Even though the coaches considered With-Me HRS useful for identifying suitable coaching tasks for the participants, it did not seem helpful for identifying behavior change needs. Perhaps, participants' behavior change needs were straightforward to identify during the coaching calls per se, as the individuals participating voluntarily in the health-coaching program likely had a good idea of the areas they wished to improve already beforehand. Hence, it may seem from the coach's perspective that additional support for identifying participants' behavior change needs was not needed. However, the participants for whom coaches utilized With-Me HRS for decision-making evaluated coaches' abilities to understand their behavior change needs and make them realize new, important areas for change higher than the group with hidden recommendations. Thus, it seems that the Profiler's user-interface encouraged coaches to analyze participants' behavior change needs systematically across different behavioral domains when making decisions on coaching objectives, which resulted in improved participant satisfaction. The coaches may have even tried to convince participants about their most important behavior change needs indicated by the Profiler. However, we do not know how much of the improved participant satisfaction was mediated by the heart rate variability based Firstbeat lifestyle assessment, which was provided only for the group with visible recommendations.

**TABLE 4. Recommendation and evaluation approaches in recent HRS studies.**

| Authors                                   | Recommendation items  | Algorithm  | Comparative algorithms  | Study setting                                | Evaluation metrics   |
|---|---|--|---|--|--|
| Alcaraz-Herrera <i>et al.</i> (2022) [44] | Meals, physical activities  | Hybrid: collaborative filtering inspired genetic algorithm                           | Simulations with and without collaborative filtering and with varying genetic operators | Offline simulations & user study (N=205)     | Aptitude, user satisfaction  |
| Torkamaan <i>et al.</i> (2022) [30]       | Stress management activities  | Hybrid: content + knowledge (case) based filtering                                   | Content-based filtering   | Real-life (N=49)                             | User-reported compliance to recommended items, user experience                               |
| Hors-Fraile <i>et al.</i> (2022) [33]     | Messages for smoking cessation  | Hybrid: knowledge (case) based filtering + demographic-based collaborative filtering | Knowledge (case) based filtering  | Real-life (N=371)                            | User satisfaction, user engagement, user-reported smoking abstinence                         |
| Ribeiro <i>et al.</i> (2022) [45]         | Weekly meal plans (recipes)   | Knowledge (constraint) based filtering   | None  | Offline (simulated data)                     | User satisfaction, recommendations meeting nutritional goals, variability in recommendations |
| Starke <i>et al.</i> (2021) [34]          | Recipes   | Hybrid: content + collaborative filtering  | Content-based and collaborative filtering, random item ranking (baseline)               | Offline (existing dataset)                   | Precision, recall, mean absolute error   |
| Ferretto <i>et al.</i> (2020) [43]        | Physical activities   | Hybrid: demographic-based collaborative filtering                                    | None  | User study with simulated data               | User satisfaction  |
| Dharia <i>et al.</i> (2018) [27]          | Physical activity schedules (mode, duration, location, timing), fitness buddies | Hybrid: collaborative filtering + knowledge (constraint) based filtering             | Simulations with varied similarity metrics applied in collaborative filtering           | Offline (existing dataset), real-life (N=24) | Offline: root mean squared error, mean absolute error; Real-life: user experience            |

Finally, when interpreting the evaluation results, it is important to bear in mind that nearly all the study participants were women, and the results may not hold for men.

#### E. FUTURE WORK

We wish to call attention towards systematic, experimental research that seeks to identify the most relevant user model features for personalizing DHBCIs in terms of improving user engagement and delivering health impact. In addition, best practices for developing multi-domain interventions are needed. The introduced conceptual virtual individual model provides ideas of features to be experimented with, and the evaluation of the multi-domain With-Me HRS provides reference results for testing the influence of user features, beyond behavior change needs and readiness to change, on the validity of recommendations. The implemented With-Me user model should be expanded at least with features describing self-efficacy, skills, and the momentary context. Features related to self-efficacy and skills enable to recommend behavior change activities that are helpful but not too challenging, and knowledge about momentary context is required for providing support at opportune moments.

With-Me HRS was based on knowledge-based filtering, which is a straightforward approach for testing the impact of user features on the validity of recommendations. However, once the most impactful user features are identified, a hybrid method combining knowledge-based filtering with demographic-based collaborative filtering would be more appropriate. Such a hybrid method has also been suggested in [5]. Knowledge-based filtering could be used as a first step to identify the subset of recommendable items that

match the most critical user features (e.g., behavior change needs, motivation and capabilities to change behavior, personal restrictions), whereas demographic-based collaborative filtering could be used as the second step to recommend items from the identified subset that were preferred by or effective for other users sharing a similar life situation with the target user. Knowledge-based filtering ensures that inappropriate or irrelevant items are not recommended, and demographic-based collaborative filtering reduces the risk of excluding highly suitable, novel items from the recommendations that may be missed by knowledge-based filtering, as it relies solely on expert knowledge (i.e., on the defined user features and the corresponding item labels). Therefore, this type of a hybrid method could facilitate extensive personalization even with a subset of the features proposed by the VI model framework.

With-Me HRS supported a multi-domain intervention, but the recommended behavior change activities were not very specific. In the future, it may be wise to implement multi-domain HRSs with two hierarchical layers to be able to provide domain-specific detailed recommendations efficiently. The top layer would be in charge of recommending behavior change objectives. The second layer could be built from domain-specific HRS submodules, which comprise specific user models and intervention libraries relevant to the domain in question. This approach would enable the modular development and usage of multi-domain HRSs. Submodules could be activated as need arises according to the identified behavior change objectives.

Finally, With-Me HRS was not designed as a standalone system, and it did not support dynamic recommendations that

adapt to individuals' evolving situations and to the effectiveness of past recommendations. For standalone HRSs, it is imperative to monitor individuals' adherence to recommendations along with changes in well-being, behavior, and behavioral determinants. Part of the monitoring could be conducted via questionnaires, especially regarding psychological factors but, when possible, unobtrusive monitoring should be used (e.g., via wearable devices, smartphones, environmental sensors) to reduce user burden and subjective bias in self-reporting.

## APPENDIX

### Appendix 1: Elements of the virtual individual model

Detailed description of the proposed VI model elements, including the proposed feature types with concrete feature examples and the interrelations between the features.

### Appendix 2: Data layers and profiling logic

Profiler's data layers and the data-processing algorithms executed by the Profiler are described in detail, including information on a) the user model features used for personalization, b) how the features are measured, and c) the algorithms used to process raw measurements into features.

## ACKNOWLEDGMENT

Various people have contributed to this work. Special thanks goes to Juha Leppänen for implementing the With-Me HRS software, Ulla-Maija Junno for participating in the development of the intervention library, and Mikko Lindholm for assisting in preprocessing the coaches' notes. The authors would like to thank Hannu Mikkola, Tero Myllymäki, Salla Muuraiskangas, Juho Merilahti, and the entire WITH-ME consortium for supporting this work. They also thank all the health coaches and participants who participated in the evaluation study. Finally, they warmly thank Elina Mattila for providing valuable comments to the first draft of this manuscript.

## REFERENCES

- [1] (Oct. 23, 2022). European Commission. *Health Promotion and Disease Prevention Knowledge Gateway: EU Burden From Non-Communicable Diseases and Key Risk Factors*. [https://knowledge4policy.ec.europa.eu/health-promotion-knowledge-gateway/eu-burden-non-communicable-diseases-key-risk-factors\\_en](https://knowledge4policy.ec.europa.eu/health-promotion-knowledge-gateway/eu-burden-non-communicable-diseases-key-risk-factors_en)
- [2] P. Krebs, J. O. Prochaska, and J. S. Rossi, "A meta-analysis of computer-tailored interventions for health behavior change," *Preventive Med.*, vol. 51, nos. 3–4, pp. 214–221, Sep. 2010, doi: [10.1016/j.ypmed.2010.06.004](https://doi.org/10.1016/j.ypmed.2010.06.004).
- [3] M. P. Couper, G. L. Alexander, N. Maddy, N. Zhang, M. A. Nowak, J. B. McClure, J. J. Calvi, S. J. Rolnick, M. A. Stopponi, R. J. A. Little, and C. C. Johnson, "Engagement and retention: Measuring breadth and depth of participant use of an online intervention," *J. Med. Internet Res.*, vol. 12, no. 4, pp. 1–19, Dec. 2010, doi: [10.2196/jmir.1430](https://doi.org/10.2196/jmir.1430).
- [4] E. G. Spanakis, S. Santana, M. Tsiknakis, K. Marias, V. Sakkalis, A. Teixeira, J. H. Janssen, H. de Jong, and C. Tziraki, "Technology-based innovations to foster personalized healthy lifestyles and well-being: A targeted review," *J. Med. Internet Res.*, vol. 18, no. 6, p. e128, Jun. 2016, doi: [10.2196/jmir.4863](https://doi.org/10.2196/jmir.4863).
- [5] K. L. Cheung, D. Durusu, X. Sui, and H. de Vries, "How recommender systems could support and enhance computer-tailored digital health programs: A scoping review," *Digit. Health*, vol. 5, Jan. 2019, Art. no. 205520761882472, doi: [10.1177/2055207618824727](https://doi.org/10.1177/2055207618824727).
- [6] L. M. Collins and K. C. Kugler, *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: Advanced Topics*, 1st ed. Cham, Switzerland: Springer, 2018, doi: [10.1007/978-3-319-91776-4](https://doi.org/10.1007/978-3-319-91776-4).
- [7] I. Nahum-Shani, S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, and S. A. Murphy, "Just-in-time adaptive interventions (JITAs) in mobile health: Key components and design principles for ongoing health behavior support," *Ann. Behav. Med.*, vol. 52, no. 6, pp. 446–462, 2018, doi: [10.1007/s12160-016-9830-8](https://doi.org/10.1007/s12160-016-9830-8).
- [8] F. Sassi and J. Hurst. (Mar. 25, 2008). *The Prevention of Lifestyle-Related Chronic Diseases: An Economic Framework*. OECD. Paris, France. [Online]. Available: <http://www.oecd.org/dataoecd/57/14/40324263.pdf>
- [9] E. Diener, S. J. Heintzelman, K. Kushlev, L. Tay, D. Wirtz, L. D. Lutes, and S. Oishi, "Findings all psychologists should know from the new science on subjective well-being," *Can. Psychol. Psychologie Canadienne*, vol. 58, no. 2, pp. 87–104, May 2017, doi: [10.1037/cap0000063](https://doi.org/10.1037/cap0000063).
- [10] S. Michie, "Making psychological theory useful for implementing evidence based practice: A consensus approach," *Qual. Saf. Health Care*, vol. 14, no. 1, pp. 26–33, Feb. 2005, doi: [10.1136/qshc.2004.011155](https://doi.org/10.1136/qshc.2004.011155).
- [11] M. Fishbein, "A reasoned action approach to health promotion," *Med. Decis. Making*, vol. 28, no. 6, pp. 834–844, 2008, doi: [10.1177/0272989X08326092](https://doi.org/10.1177/0272989X08326092).
- [12] R. Davis, R. Campbell, Z. Hildon, L. Hobbs, and S. Michie, "Theories of behaviour and behaviour change across the social and behavioural sciences: A scoping review," *Health Psychol. Rev.*, vol. 9, no. 3, pp. 323–344, Aug. 2015, doi: [10.1080/17437199.2014.941722](https://doi.org/10.1080/17437199.2014.941722).
- [13] D. Spruijt-Metz and W. Nilsen, "Dynamic models of behavior for Just-in-Time adaptive interventions," *IEEE Pervasive Comput.*, vol. 13, no. 3, pp. 13–17, Jul. 2014, doi: [10.1109/MPRV.2014.46](https://doi.org/10.1109/MPRV.2014.46).
- [14] A. Honka, K. Kaipainen, H. Hietala, and N. Saranummi, "Rethinking health: ICT-enabled services to empower people to manage their health," *IEEE Rev. Biomed. Eng.*, vol. 4, pp. 119–139, 2011.
- [15] T. L. Webb, J. Joseph, L. Yardley, and S. Michie, "Using the Internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy," *J. Med. Internet Res.*, vol. 12, no. 1, p. e4, Feb. 2010, doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376).
- [16] S. Michie, M. Richardson, M. Johnston, C. Abraham, J. Francis, W. Hardeman, M. P. Eccles, J. Cane, and C. E. Wood, "The behavior change technique taxonomy (V1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions," *Ann. Behav. Med.*, vol. 46, no. 1, pp. 81–95, Aug. 2013, doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6).
- [17] J. Pincay, L. Teran, and E. Portmann, "Health recommender systems: A State-of-the-art review," in *Proc. 6th Int. Conf. eDemocracy eGovernment (ICEDEG)*, Apr. 2019, pp. 47–55, doi: [10.1109/ICEDEG.2019.8734362](https://doi.org/10.1109/ICEDEG.2019.8734362).
- [18] R. De Croon, L. Van Houdt, N. N. Htun, G. Štiglic, V. Vanden Abeele, and K. Verbert, "Health recommender systems: Systematic review," *J. Med. Internet Res.*, vol. 23, no. 6, Jun. 2021, Art. no. e18035, doi: [10.2196/18035](https://doi.org/10.2196/18035).
- [19] C. Anderson, "A survey of food recommenders," 2018, *arXiv:1809.02862*.
- [20] S. Ghanvatkar, A. Kankanhalli, and V. Rajan, "User models for personalized physical activity interventions: Scoping review," *JMIR Mhealth Uhealth*, vol. 7, no. 1, Jan. 2019, Art. no. e11098, doi: [10.2196/11098](https://doi.org/10.2196/11098).
- [21] S. Colantonio, G. Coppini, D. Germanese, D. Giorgi, M. Magrini, P. Marraccini, M. Martinelli, M. A. Morales, M. A. Pascali, G. Raccichini, M. Righi, and O. Salvetti, "A smart mirror to promote a healthy lifestyle," *Biosyst. Eng.*, vol. 138, pp. 33–43, Oct. 2015, doi: [10.1016/j.biosystemseng.2015.06.008](https://doi.org/10.1016/j.biosystemseng.2015.06.008).
- [22] V. Pandey, D. D. Upadhyay, N. Nag, and R. Jain, "Personalized user modelling for context-aware lifestyle recommendations to improve sleep," in *Proc. HealthRecSys*, 2020, p. 7.
- [23] W. Yue, Z. Wang, J. Zhang, and X. Liu, "An overview of recommendation techniques and their applications in healthcare," *IEEE/CAA J. Automat. Sinica*, vol. 8, no. 4, pp. 701–717, Apr. 2021, doi: [10.1109/JAS.2021.1003919](https://doi.org/10.1109/JAS.2021.1003919).
- [24] M. Rabbi, M. H. Aung, M. Zhang, and T. Choudhury, "MyBehavior: Automatic personalized health feedback from user behaviors and preferences using smartphones," in *Proc. UbiComp*, Osaka, Japan, 2015, pp. 707–718, doi: [10.1145/2750858.2805840](https://doi.org/10.1145/2750858.2805840).
- [25] R. Ali, M. Afzal, M. Hussain, M. Ali, M. H. Siddiqi, S. Lee, and B. H. Kang, "Multimodal hybrid reasoning methodology for personalized wellbeing services," *Comput. Biol. Med.*, vol. 69, pp. 10–28, Feb. 2016, doi: [10.1016/j.combiomed.2015.11.013](https://doi.org/10.1016/j.combiomed.2015.11.013).



- [26] C. G. Lim, Z. M. Kim, and H. J. Choi, "Developing a mobile wellness management system for healthy lifestyle by analyzing daily living activities," *Stud Health Technol. Inf.*, vol. 245, pp. 146–150, Jan. 2017, doi: [10.3233/978-1-61499-830-3-146](https://doi.org/10.3233/978-1-61499-830-3-146).
- [27] S. Dharía, M. Eirinaki, V. Jain, J. Patel, I. Varlamis, J. Vora, and R. Yamauchi, "Social recommendations for personalized fitness assistance," *Pers. Ubiquitous Comput.*, vol. 22, no. 2, pp. 245–257, 2018, doi: [10.1007/s00779-017-1039-8](https://doi.org/10.1007/s00779-017-1039-8).
- [28] P. Paredes, R. Giald-Bachrach, M. Czerwinski, A. Roseway, K. Rowan, and J. Hernandez, "PopTherapy: Coping with stress through pop-culture," in *Proc. 8th Int. Conf. Pervasive Comput. Technol. Healthcare*, 2014, pp. 109–117.
- [29] S. Clarke, L. G. Jaimes, and M. A. Labrador, "MStress: A mobile recommender system for just-in-time interventions for stress," in *Proc. CCNC*, Las Vegas, NV, USA, 2017, pp. 1–5, 2017, doi: [10.1109/CCNC.2017.8015367](https://doi.org/10.1109/CCNC.2017.8015367).
- [30] H. Torkamaan and J. Ziegler, "Recommendations as challenges: Estimating required effort and user ability for health behavior change recommendations," in *Proc. 27th Int. Conf. Intell. User Interfaces*, Mar. 2022, pp. 106–119, doi: [10.1145/3490099.3511118](https://doi.org/10.1145/3490099.3511118).
- [31] H. Lattar, A. B. Salem, H. H. B. Ghézala, and F. Boufares, "Health recommender systems: A survey," in *Proc. SETIT*, Maghreb, Tunisia, 2018, pp. 182–191, doi: [10.1007/978-3-030-21005-2](https://doi.org/10.1007/978-3-030-21005-2).
- [32] T. N. T. Tran, A. Felfernig, C. Trattner, and A. Holzinger, "Recommender systems in the healthcare domain: State-of-the-art and research issues," *J. Intell. Inf. Syst.*, vol. 57, pp. 171–201, Aug. 2021, doi: [10.1007/s10844-020-00633-6](https://doi.org/10.1007/s10844-020-00633-6).
- [33] S. Hors-Fraile, M. J. J. M. Candel, F. Schneider, S. Malwade, F. J. Nunez-Benjumea, S. Syed-Abdul, L. Fernandez-Luque, and H. de Vries, "Applying collective intelligence in health recommender systems for smoking cessation: A comparison trial," *Electronics*, vol. 11, no. 8, p. 1219, 2022, doi: [10.3390/electronics11081219](https://doi.org/10.3390/electronics11081219).
- [34] A. Starke, C. Trattner, H. Bakken, M. Johannessen, and V. Solberg, "The cholesterol factor: Balancing accuracy and health in recipe recommendation through a nutrient-specific metric," in *Proc. RecSys*, Amsterdam, The Netherlands, 2021, pp. 1–11.
- [35] E. Yom-Tov, G. Ferraru, M. Kozdoba, S. Mannor, M. Tennenholtz, and I. Hochberg, "Encouraging physical activity in patients with diabetes: Intervention using a reinforcement learning system," *J. Med. Internet Res.*, vol. 19, no. 10, p. e7994, Oct. 2017, doi: [10.2196/JMIR.7994](https://doi.org/10.2196/JMIR.7994).
- [36] R. S. Sadasivam, S. L. Cutrona, R. L. Kinney, B. M. Marlin, K. M. Mazor, S. C. Lemon, and T. K. Houston, "Collective-intelligence recommender systems: Advancing computer tailoring for health behavior change into the 21st century," *J. Med. Internet Res.*, vol. 18, no. 3, pp. 1–13, Mar. 2016, doi: [10.2196/jmir.4448](https://doi.org/10.2196/jmir.4448).
- [37] S. Hors-Fraile, O. Rivera-Romero, F. Schneider, L. Fernandez-Luque, F. Luna-Perejon, A. Civit-Balcells, and H. de Vries, "Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: A scoping review," *Int. J. Med. Inf.*, vol. 114, pp. 143–155, Jun. 2018, doi: [10.1016/j.ijmedinf.2017.12.018](https://doi.org/10.1016/j.ijmedinf.2017.12.018).
- [38] R. S. Sadasivam, E. M. Borglund, R. Adams, B. M. Marlin, and T. K. Houston, "Impact of a collective intelligence tailored messaging system on smoking cessation: The Perspect randomized experiment," *J. Med. Internet Res.*, vol. 18, no. 11, p. e6465, Nov. 2016, doi: [10.2196/jmir.6465](https://doi.org/10.2196/jmir.6465).
- [39] S. Hors-Fraile, F. J. Núñez-Benjumea, L. C. Hernández, F. O. Ruiz, and L. Fernandez-Luque, "Design of two combined health recommender systems for tailoring messages in a smoking cessation app," in *Proc. RecSys*, Boston, MA, USA, 2016, pp. 1–4.
- [40] M. Fishbein, H. C. Triandis, F. H. Kanfer, M. Becker, S. E. Middlestadt, and A. Eichler, "Factors influencing behavior and behavior change," in *Handbook of Health Psychology*, A. Baum, T. R. Revenson, J. E. Singer, Eds. Hillsdale, NJ, USA: Lawrence Erlbaum, 2001, pp. 3–17.
- [41] S. Hors-Fraile, S. Malwade, F. Luna-Perejon, C. Amaya, A. Civit, F. Schneider, P. Bamidis, S. Syed-Abdul, Y.-C. Li, and H. de Vries, "Opening the black box: Explaining the process of basing a health recommender system on the I-Change behavioral change model," *IEEE Access*, vol. 7, pp. 176525–176540, 2019, doi: [10.1109/ACCESS.2019.2957696](https://doi.org/10.1109/ACCESS.2019.2957696).
- [42] H. De Vries, "An integrated approach for understanding health behavior: the I-change Model as an example," *Psychol. Behav. Sci. Int. J.*, vol. 2, no. 2, pp. 555–585, 2017, doi: [10.19080/pbsij.2017.02.555585](https://doi.org/10.19080/pbsij.2017.02.555585).
- [43] L. R. Ferretto, E. A. Bellei, D. Biduski, L. C. P. Bin, M. M. Moro, C. R. Cervi, and A. C. B. De Marchi, "A physical activity recommender system for patients with arterial hypertension," *IEEE Access*, vol. 8, pp. 61656–61664, 2020, doi: [10.1109/ACCESS.2020.2983564](https://doi.org/10.1109/ACCESS.2020.2983564).
- [44] H. Alcaraz-Herrera, J. Cartlidge, Z. Toumpakari, M. Western, and I. Palomares, "EvoRecSys: Evolutionary framework for health and wellbeing recommender systems," *User Modeling User-Adapt Interact.*, vol. 2022, pp. 1–39, Jan. 2022, doi: [10.1007/s11257-021-09318-3](https://doi.org/10.1007/s11257-021-09318-3).
- [45] D. Ribeiro, T. Barbosa, J. Ribeiro, F. Sousa, E. F. Vieira, M. Silva, and A. Silva, "SousChef system for personalized meal recommendations: A validation study," *Appl. Sci.*, vol. 12, no. 2, p. 702, Jan. 2022, doi: [10.3390/app12020702](https://doi.org/10.3390/app12020702).
- [46] S. Hales, "Social networks for improving healthy weight loss behaviors for overweight and obese adults: A randomized clinical trial of the social pounds off digitally (social POD) mobile app," *Int. J. Med. Informat.*, vol. 94, pp. 81–90, Oct. 2016, doi: [10.1016/j.ijmedinf.2016.07.003](https://doi.org/10.1016/j.ijmedinf.2016.07.003).
- [47] T. Rist, A. Seiderer, S. Hammer, M. Mayr, and E. Andre, "CARE-Extending a digital picture frame with a recommender model to enhance well-being of elderly people," in *Proc. PervasiveHealth*, Istanbul, Turkey, 2015, pp. 112–120, doi: [10.4108/icst.pervasivehealth.2015.259255](https://doi.org/10.4108/icst.pervasivehealth.2015.259255).
- [48] S. Lippke and J. P. Ziegelmann, "Theory-based health behavior change: Developing, testing, and applying theories for evidence-based interventions," *Appl. Psychol.*, vol. 57, no. 4, pp. 698–716, Oct. 2008, doi: [10.1111/j.1464-0597.2008.00339.x](https://doi.org/10.1111/j.1464-0597.2008.00339.x).
- [49] J. O. Prochaska and W. F. Velicer, "The transtheoretical model of health behavior change," *Amer. J. Health Promotion*, vol. 12, no. 1, pp. 38–48, 1997, doi: [10.4278/0890-1171-12.1.38](https://doi.org/10.4278/0890-1171-12.1.38).
- [50] L. K. Bartholomew, G. S. Parcel, G. Kok, and N. H. Gottlieb, *Planning Health Promotion Programs: An Intervention Mapping Approach*, 2nd ed. San Francisco, CA, USA: Jossey-Bass, 2006.
- [51] B. Verplanken and W. Wood, "Interventions to break and create consumer habits," *J. Public Policy Marketing*, vol. 25, no. 1, pp. 90–103, 2006.
- [52] M. Siegel and L. D. Lotenberg, *Marketing Public Health: Strategies to Promote Social Change*, 2nd ed. Sudbury, MA, USA: Jones & Bartlett, 2007.
- [53] D. Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York, NY, USA: HarperCollins, 2008.
- [54] K. Glanz and D. B. Bishop, "The role of behavioral science theory in development and implementation of public health interventions," *Annu. Rev. Public Health*, vol. 31, no. 1, pp. 399–418, 2010, doi: [10.1146/annurev.publhealth.012809.103604](https://doi.org/10.1146/annurev.publhealth.012809.103604).
- [55] M. E. Fernandez, R. A. C. Ruiter, C. M. Markham, and G. Kok, "Intervention mapping: Theory- and evidence-based health promotion program planning: Perspective and examples," *Frontiers Public Health*, vol. 7, Aug. 2019, doi: [10.3389/fpubh.2019.00209](https://doi.org/10.3389/fpubh.2019.00209).
- [56] T. Willmott and S. Rundle-Thiele, "Are we speaking the same language? Call for action to improve theory application and reporting in behaviour change research," *BMC Public Health*, vol. 21, no. 479, pp. 1–8, Mar. 2021, doi: [10.1186/s12889-021-10541-1](https://doi.org/10.1186/s12889-021-10541-1).
- [57] P. Salovey, "Promoting prevention and detection: Psychologically tailoring and framing messages about health," in *Science and Medicine in Dialogue: Thinking Through Particulars and Universals*, R. Bibace, J. D. Laird, K. L. Noller, J. Valsiner, Eds. Westport, CT, USA: Praeger, 2005, pp. 19–36.
- [58] J. Hornik, C. Ofir, and M. Rachamim, "Quantitative evaluation of persuasive appeals using comparative meta-analysis," *Commun. Rev.*, vol. 19, no. 3, pp. 192–222, Jul. 2016, doi: [10.1080/10714421.2016.1195204](https://doi.org/10.1080/10714421.2016.1195204).
- [59] R. Josekutty Thomas, J. Masthoff, and N. Oren, "Adapting healthy eating messages to personality," in *Proc. PERSUASIVE*, Amsterdam, The Netherlands, 2017, pp. 119–132, doi: [10.1007/978-3-319-55134-0](https://doi.org/10.1007/978-3-319-55134-0).
- [60] G. Kok, N. H. Gottlieb, G.-J.-Y. Peters, P. D. Mullen, G. S. Parcel, R. A. C. Ruiter, M. E. Fernández, C. Markham, and L. K. Bartholomew, "A taxonomy of behaviour change methods: An intervention mapping approach," *Health Psychol. Rev.*, vol. 10, no. 3, pp. 297–312, Jul. 2016, doi: [10.1080/17437199.2015.1077155](https://doi.org/10.1080/17437199.2015.1077155).
- [61] H. Nieminen, S. Baggström, H. Lehtonen, A. Leppisaari, H. Mikkola, and I. K. Korhonen, "Mobile solution for increasing the efficiency of health coaching for preventative health care and chronic disease management," *Trans. Japanese Soc. Med. Biol. Eng.*, vol. 51, p. 153, Jan. 2013, doi: [10.11239/jsmbe.51.M-153](https://doi.org/10.11239/jsmbe.51.M-153).
- [62] S. T. Muuraiskangas, A. M. Honka, U.-M. Junno, H. O. Nieminen, and J. K. Kaartinen, "A technology-assisted telephone intervention for work-related stress management: Pilot randomized controlled trial," *J. Med. Internet Res.*, vol. 24, no. 7, Jul. 2022, Art. no. e26569, doi: [10.2196/26569](https://doi.org/10.2196/26569).
- [63] H. Ahveninen, J. Rintala, J. Ollikainen, J. Suhonen, and H. Arola, "TyöOptimi-kysely antaa tunnistamaan työssäkäyvän kognitiivisia ja muistioireita," *Suomen Lääkärelehti*, vol. 69, no. 4, pp. 207–214, 2014.

- [64] D. T. Gillanders, H. Bolderston, F. W. Bond, M. Dempster, P. E. Flaxman, L. Campbell, S. Kerr, L. Tansey, P. Noel, C. Ferenbach, S. Masley, L. Roach, J. Lloyd, L. May, S. Clarke, and B. Remington, "The development and initial validation of the cognitive fusion questionnaire," *Behav. Therapy*, vol. 45, no. 1, pp. 83–101, Jan. 2014, doi: [10.1016/j.beth.2013.09.001](https://doi.org/10.1016/j.beth.2013.09.001).
- [65] C. R. Nigg, P. M. Burbank, C. Padula, R. Dufresne, J. S. Rossi, W. F. Velicer, R. G. Laforge, and J. O. Prochaska, "Stages of change across ten health risk behaviors for older adults," *Gerontologist*, vol. 39, no. 4, pp. 473–482, Aug. 1999.
- [66] K. Wilson, I. Senay, M. Durantini, F. Sánchez, M. Hennessy, B. Spring, D. Albarracín, "When it comes to lifestyle recommendations, more is sometimes less: A meta-analysis of theoretical assumptions underlying the effectiveness of interventions promoting multiple behavior domain change," *Psychol. Bull.*, vol. 141, no. 2, pp. 474–509, 2015, doi: [10.1037/a0038295](https://doi.org/10.1037/a0038295).
- [67] E. Mattila, R. Lappalainen, P. Väikkynen, E. Sairanen, P. Lappalainen, L. Karhunen, K. Peuhkuri, R. Korpela, M. Kolehmainen, and M. Ermes, "Usage and dose response of a mobile acceptance and commitment therapy app: Secondary analysis of the intervention arm of a randomized controlled trial," *JMIR mHealth uHealth*, vol. 4, no. 3, p. e90, Jul. 2016, doi: [10.2196/mhealth.5241](https://doi.org/10.2196/mhealth.5241).
- [68] C. C. Aggarwal, *Recommender Systems: The Textbook*. Cham, Switzerland: Springer, 2016.
- [69] A. Vargha and H. D. Delaney, "A critique and improvement of the CL common language effect size statistics of McGraw and wong," *J. Educ. Behav. Statist.*, vol. 25, no. 2, pp. 101–132, Jun. 2000, doi: [10.3102/10769986025002101](https://doi.org/10.3102/10769986025002101).
- [70] J. C.-H. Li, "Effect size measures in a two-independent-samples case with nonnormal and nonhomogeneous data," *Behav. Res. Methods*, vol. 48, no. 4, pp. 1560–1574, Dec. 2016, doi: [10.3758/s13428-015-0667-z](https://doi.org/10.3758/s13428-015-0667-z).
- [71] S. H. Schwartz and W. Bilsky, "Toward a universal psychological structure of human values," *J. Personality Social Psychol.*, vol. 53, no. 3, pp. 550–562, Sep. 1987, doi: [10.1037/0022-3514.53.3.550](https://doi.org/10.1037/0022-3514.53.3.550).
- [72] S. C. Matz, M. Kosinski, G. Nave, and D. J. Stillwell, "Psychological targeting as an effective approach to digital mass persuasion," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 48, pp. 12714–12719, Nov. 2017, doi: [10.1073/pnas.1710966114](https://doi.org/10.1073/pnas.1710966114).
- [73] T. Graham and W. Abrahamse, "Communicating the climate impacts of meat consumption: The effect of values and message framing," *Global Environ. Change*, vol. 44, pp. 98–108, May 2017, doi: [10.1016/j.gloenvcha.2017.03.004](https://doi.org/10.1016/j.gloenvcha.2017.03.004).
- [74] A. Coletta and R. B. Kreider, "Genetic profiling for weight loss: Potential candidate genes," *Bioenergetics*, vol. 4, no. 2, pp. 1–14, Sep. 2015, doi: [10.4172/2167-7662.1000126](https://doi.org/10.4172/2167-7662.1000126).
- [75] C. Celis-Morales et al., "Design and baseline characteristics of the Food4Me study: A web-based randomised controlled trial of personalised nutrition in seven European countries," *Genes Nutrition*, vol. 10, no. 1, pp. 1–13, 2015, doi: [10.1007/s12263-014-0450-2](https://doi.org/10.1007/s12263-014-0450-2).



**ANITA M. HONKA** received the M.Sc. degree in mathematics from Tampere University, Tampere, Finland, in 2009, where she is currently pursuing the Ph.D. degree in biomedical sciences and engineering.

From 2009 to 2022, she worked as a Research Scientist at VTT Technical Research Centre of Finland Ltd., Tampere, Finland, in the field of digital health and well-being. In 2018, she spent six months as an Honorary Visiting Scholar with the Digital Psychiatry Laboratory, Flinders University of South Australia, Adelaide, Australia. She is interested in the development of personalized behavior change support systems for promoting healthy lifestyle and well-being, particularly from the viewpoint of predicting human behavior and well-being. She has expertise in health data analytics and visualization, as well as in algorithm development for user profiling and intervention recommendation.



**HANNU NIEMINEN** (Member, IEEE) was born in Iitti, Finland, in 1960. He received the B.S. and M.S. degrees in biomedical engineering from the Technical University of Tampere, in 1985, and the Ph.D. degree in signal processing from the Tampere University of Technology, in 1995.

From 1985 to 1994, he worked as a Research Scientist at the Biomedical Signal Processing Team, Technical Research Center of Finland. From 1994 to 2011, he held several managerial positions at Nokia Corporation. Since 2011, he has been a Senior Research Fellow with the Faculty of Medicine and Health Technology, Tampere University. He is the author or more than 40 articles and holds seven patents. His research interests include decision support for medicine, data visualization, and digital solutions for preventive health care.



**HEIDI SIMILÄ** received the M.Sc. degree in biophysics and the D.Sc. (Tech.) degree in computer science and engineering from the University of Oulu, Oulu, Finland, in 2006 and 2017, respectively.

From 2006 to 2019, she worked as a Research Scientist and from 2019 to 2021, as a Senior Scientist at VTT Technical Research Centre of Finland, Oulu. From 2014 to 2016, she also worked part-time as a Researcher at the Center for Life Course Health Research, University of Oulu. She specializes in signal processing, algorithm development, and predictive analytics. Her research interest includes deriving meaningful health information out of heterogeneous personal data.



**JOUNI KAARTINEN** received the B.Sc. degree in automation technology from North Karelia Polytechnics, Joensuu, Finland, in 1994, and the M.Sc. degree in information technology from the University of Oulu, Oulu, Finland, in 1996. He has been working as a Research Scientist at VTT, since 1999, and later as a Senior Scientist. He was worked as a Team Leader of health information systems, well-being technology, and situation awareness and sensing teams at VTT,

from 2004 to 2012. His research interests include machine reasoning methods, architectures, data management, and sensor networks in several domain areas. Last years, he has researched and developed eHealth solutions for the personal well-being in several contract research and European joint projects called AmIE (ITEA2), A2E2 (AAL), V2me (AAL), AiB (AAL), With-Me (Artemis), and MoMaMo (Business Finland).



**MARK VAN GILS** (Member, IEEE) received the M.Sc. degree in applied physics and the Ph.D. degree in medical electrical engineering from the Eindhoven University of Technology, The Netherlands, in 1990 and 1995, respectively.

From 1996 to 2020, he was working at VTT Technical Research Centre of Finland Ltd., in roles progressing from Research Scientist via Research Team Leader to Research Professor. Since 2021, he has been a Full Professor at Tampere University, Tampere, Finland. He is the author of more than 160 articles and four patents. His research interests include decision support for health and wellness applications, biomedical signal processing, artificial intelligence, and machine learning in healthcare.

Prof. van Gils was a recipient of two conference awards and two awards for scientific publishing at VTT, in 2018 and 2019. He obtained the title of Docent (Adjunct Professor) at Aalto University, Espoo, Finland, in 2000. He has been a Theme Editor of the theme "Biomedical Health Informatics" of IEEE EMBC Annual Conferences, since 2015.