# Investigating Mechanisms for User Integration in the Activity Goal Recommendation Process by Interface Design

Katja Herrmanny
katja.herrmanny@uni-due.de
University of Duisburg-Essen
Duisburg, Germany

Simone Löppenberg
University of Duisburg-Essen
Duisburg, Germany

Michael Schwarz
michael.schwarz@uni-due.de
University of Duisburg-Essen
Duisburg, Germany

## ABSTRACT

In the field of physical activity recommendation, we have to deal with many confounding variables that lead to high result uncertainty. Assuming that users' competence is an essential factor for reduction of the problem of inaccurate recommendations, we present and evaluate an approach on how to integrate users in the recommendation process. We investigate if and how interface element design can contribute to understanding, reflection and modification of the recommendation result. In the work described here, we use interface elements that allow for planning of physical activity goal striving. Results show that such interface elements can principally empower users, support recommendation reflection and stimulate user interaction with the recommendation.

## KEYWORDS

user integration, recommendation, goal setting, user empowerment, user interface, activity tracking

## 1 INTRODUCTION AND SCOPE OF THE PAPER

Besides algorithm accuracy, design of user interfaces is an important component of recommender systems, that has gained more and more interest in the last years [10, 13, 21]. Especially health-related personalised recommendations have to deal with many confounding variables, that are unknown to the algorithm or not quantifiable and thus difficult or even impossible to be considered in the recommender's reasoning [8]. Due to this and other aspects like autonomy issues, integrating the user in the process is an essential part of health-related recommender systems [8]. However, in this field little research has been done to investigate how the user could be integrated. In this paper, we investigate how to design a user interface to integrate the user in the recommendation process of physical activity goals by:

- empowering users to understand the recommendation and its implications,
- reflecting and evaluating the recommendation, and
- providing the opportunity to actively manipulate it.

In our approach, empowerment and reflection are mainly achieved by the planning of recommendation realisation, which helps the user to assess whether the recommended goal is realistic or not. It also supports the user in appropriate modifications. We present two interface alternatives and an evaluation regarding their potential to integrate the user in the process in the way described above. We further investigated understanding, usability, and aesthetics which are relevant factors for user engagement.



*IntRS '19: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, 19 Sept 2019, Copenhagen, DK*

## 2 RELATED WORK

Recommender systems intend to support users in the decision making process based on their preferences and needs [5]. There has been a lot of research focusing on the prediction accuracy [5]. However, recently it has been shown that accuracy of the algorithm influences the user experience only partially [13] and that the key to success are the functions provided by the user interface of recommender systems [10]. User interface design and dialogs affect usability, acceptance, item rating behaviour, selection behaviour, trust, and willingness to buy and reuse the system [5]. In order to improve recommender systems and user satisfaction, it is beneficial to provide users with the opportunity to interact with recommendations and to make adjustments if needed [12]. However, it is often not possible to provide feedback to the system, which is important to adapt its assumptions about the user [12]. In the context of rating based recommender systems (e.g. movies, music), it has also been shown that interactive recommender systems are advantageous since they can factor in changed user interests over time or corrections to previously made mistaken ratings [11]. Yet most recommender systems consider user ratings as always correct [11]. The authors [11] therefore suggest to support user interaction with the recommendation by allowing an adjustment of previous ratings. This would provide explicit feedback to the system, instead of implicit feedback, which is typically done by monitoring users' behavior [3].

In some contexts, giving explicit user feedback requires an understanding of the recommendation. This can be supported by explanation. Explanation interfaces are used in different fields - such as expert systems, medical decision support systems, intelligent tutoring systems, data exploration systems, and recommender systems [18]. By explaining the recommendation result, they aim at providing transparency and, in consequence, trust and user acceptance [18]. Such explanation is also termed user empowerment. Empowerment can also be found in the health sector, especially under the concept of patient empowerment [1, 7, 14]. In this domain, empowerment includes knowledge transfer and persuasion. Following Kondylakis et al. [14], patient empowerment is achieved through the accessibility of information (e. g. the opportunity to get information on the internet). This is in line with Alpay et al. [1] who state that the term empowerment "is frequently used to describe a situation where patients are encouraged to be active in their own health management". Regarding the effect of empowerment, it has

been shown that empowerment is provided by e-health applications and can positively influence patients' long-term health status [7]. Specifically for goal recommendations in activity tracking applications, it has been shown that users wish to get explaining or illustrating information [9]. One approach to providing such information was to illustrate the impact of the activity goal by reference routes [9]. Doing so, users should get a better sense of the amount of activity of the selected goal. In summary, in the health sector, empowerment includes knowledge transfer and persuasion. [19] distinguish between empowerment and persuasion. They use the term empowerment to describe explanation interfaces in health recommender systems. In a literature review they found that "empowering the user by interactively guiding his decision, and creating trust, using [...] explanatory interfaces" are relevant concepts in health recommender research. Following Schäfer et al. [19], in health recommender systems, empowerment is achieved through explanation of the internal logic of the recommender system, which they claim to be one of the key challenges for health-related recommender systems. Our understanding of empowerment goes even beyond. Beside explanation of the recommender's reasoning, empowerment as we define it also includes transferring domain knowledge and explaining of the recommendation's implications.

## 3 CONTRIBUTION

To overcome the shortcomings of algorithmic recommendations, we (as well as other researchers mentioned in the previous section) propose to integrate the user in the recommendation process to achieve better results.

In contrast to previous approaches, in this work we don't focus on textual explaining elements. Instead, we investigate, if and how interface design can be used to integrate users in the recommendation process.

Our idea is to

- support user empowerment by interface elements that explain the impact of the recommendation,
- support recommendation reflection by interface elements that indicate inaccuracy as well es elements that explain the impact of the recommendation and ,
- support user engagement by interface elements that allow for manipulation of the recommendation and its implementation planning.

In order to investigate the general potential of our idea to foster user integration through interface design, we designed exemplary interface elements, which are described in detail in the following section. This work does not focus on the evaluation of the interface elements itself or the app they are framed in. They are just used as tools to investigate our research questions.

In summary, this paper contributes by investigating the potential of interface design to support user integration in the recommendation process for the purpose of improving the recommendation.

## 4 APPROACH

The scenario of our work is a physical activity tracking app, providing recommendations for an appropriate (i.e. challenging, but not overburdening), numerical weekly activity goal based on user and context data of the specific user. As described above, recommendations in this field are very error-prone because of large variations over time and a large number of confounding variables indeterminable for the algorithm. To meet this challenge we propose to integrate the user in the recommendation process. Therefore, it is necessary (a) to empower users to understand the recommendation and its implications, (b) to reflect and evaluate it and (c) to provide adequate opportunities to actively manipulate it. To reach this aim, we used two main strategies. The first one is transparency. By showing the uncertainty of the algorithm, we want to make the user aware of the necessity of his/her influence. The second strategy is implementation planning to illustrate the impact of the recommendation. The specific (exemplary) user interface elements we designed, are described in the following.

Firstly, we presented the recommendation - which is a numerical value on a continuous scale, in contrast to discrete items like in conventional recommender systems - as a range on a modifiable slider. The recommendation with the highest probability to be the best fitting one, is used as default value. We further added a colour gradient to the slider indicating uncertainty regarding the system's recommendation (for different values of the range). Presenting a range instead of a single value and indicating uncertainty of the algorithm should support users in recommendation interpretation by understanding that the recommendation is a non-exact one that needs to be reflected and probably to be adjusted. The slider is intended to encourage and enable users to adjust it.

Kilocalories (kcal) were used to present the recommendation, which is, besides the number of steps, one of the most common units for measuring physical activity. The advantage of that approach is that all kind of physical activity can be subsumed in this unit. However, it is problematic that the recommendation may be very abstract for the user, especially as kcal are more common in the field of nutrition. So secondly, we converted the recommendation to a unit that is more intuitive and better to interpret by the user. Therefore, it was converted to the time needed to be active in three different intensity levels (low, moderate, high) in order to achieve the weekly goal. The reasons for dividing the goal into different activity levels are explained in section 5. The relation between the three activity levels could be modified by the user (which leads to more time needed to achieve the goal if the user reduces the more intensive activity and increases the low-intensive one and vice versa). By giving the user a sense of the amount of time needed for goal striving and thus for the difficulty of the goal, the user should be empowered to understand the implications of the recommendation and be encouraged to reflect it.

In addition to this, we thirdly allowed for a more detailed planning to give an even better sense of how difficult it is to achieve the chosen goal in daily life. The interface therefore provides the user with the opportunity to plan, how to distribute the required time to different activity units. This also aims at fostering reflection of the goal and, as a consequence, encouraging users to modify it.

## 5 MOCK-UPS

We designed mock-ups with three main interface elements which we implemented for an android application. The targeted application should be an activity tracking app, which recommends a

physical activity goal considering user and context parameters. The goal unit should be consumption of kilocalories per week. The three interface elements are:

(1) Goal selection element (see figure 1 (a)): The first element is a slider. The slider range represents the recommendation range and the default slider position represents the most recommended value, i.e. the value with the highest probability to be the most suitable one. For the study, the goal is initially equipped with a default value of 1400 kcal. As this study focuses on the interface and not the algorithm, this recommendation is the same for all participants to exclude this as a confounding variable. (The recommendation algorithm is not part of this paper and thus described elsewhere.) A colour gradient at the slider range indicates the probabilities resulting from the recommendation algorithm. The user has the opportunity to modify the goal within the given range.
(2) Conversion into time element (see figure 1 (b)): The second element converts the kcal goal into amount of time, which is less abstract. As time needed to consume a certain amount of kilocalories strongly depends on the intensity of the performed activity, it is necessary to select which ratio of different activity levels will be performed. Otherwise the conversion into time would be far too inexact and thus not helpful. To get an approximation without cognitively overburdening the user, we provided three intensity categories used in the health sector, which can be seen in table 1. For calculation we used an average MET value of each category, which is also presented in table 1. MET (= metabolic equivalent of task) is a unit to indicate the intensity of an activity and can be converted to calories and vice versa. Depending on the activity category, more or less active time is needed to achieve the goal. Different exhausting activities or sports fall into different categories (see table 1). We enable users to specify, which amount of each intensity level they plan to do to achieve the goal. This is done by an adjustable circular seekbar (circle), which allows for modification of the intensity types portions within the circle. The default state in our study is 1/3 for each part of the circle, i.e. for each intensity level. Within the circle, the selected goal is shown and how much of the goal has already been distributed among the intensities. If the circle is filled, all kilocalories of the goal are planned. Below the circle, the derived duration in minutes of each intensity level is shown. Three different icons represent the different intensity levels (walking for low intensity; cycling for moderate intensity; running for high intensity).
(3) Activity unit planning element (see figure 1 (c)): The third interface element consists of input elements for planning units of activity for each intensity level. For example, if the user has chosen to do 60 minutes of high-intense activity (in interface element 2), those could be distributed to one exercise or activity unit of 40 minutes and one of 20 minutes. 10 minutes is the minimum duration. By interacting with the circular seekbar or by tapping on the icons below, it is possible to switch between the activity units of low, moderate and high intensity. Below the heading it is presented how much time remains to be distributed to activity units for the regarding intensity level. Below there is a table in which each row represents an activity unit (for each intensity level) and each column represents 10 minutes. By tapping on the buttons (+ -) on the left and on the right, the duration of the activity unit can be modified, whereby the respective cell is coloured. By using the buttons (+ -) below the table, a new activity unit (row) can be added.

**Table 1: Intensity categories with corresponding MET values**

| Intensity | MET | Example Activity |
|---|---|---|
| low | 3 | slow walking, walking downstairs, golf |
| moderate | 5 | walking, weight lifting, dancing |
| high | 7 | running, playing soccer |

The separation of the activity into low, moderate and high intensity is done in order to convert the goal into time and thus better estimate the calorie consumption and check if the selected goal is realistic.

The distribution into activity units should help the users to get a feeling for their own activity behaviour and to integrate activities into their everyday life through precise planning. This also serves as a supporting element to check whether the selected goal is realistic or not.

The three elements of the page (weekly goal, conversion into time, activity units) are interdependent. When the weekly goal is increased or decreased, the amount of time to be distributed among the intensity levels also increased or decreased. This has an immediate effect on the minutes of activity units to be planned, which are adapted to the intensity distribution. The application can provide helping information for each area if required.

Especially for reasons of reflection, we assumed that it might be helpful to show all interface elements on one single page. When users modify their goal, they immediately see the corresponding implications on the other planning steps, and thus can better evaluate which amount of adjustment is adequate. On the other hand, space is limited on smart-phone displays. Layouts are quickly overloaded which could result in information overload for the users. This is presumably not helpful in the reflection process. Thus, we decided to design two interfaces with mainly the same elements. The first one is scrollable and presents all described interface elements on a single page. The second one presents them on successive pages with the opportunity to navigate between the pages.

## 5.1 Interface 1: Single Page

Figure 1 shows the single page interface. At the top of the page, the goal selection slider is presented. The conversion into time is presented below the goal selection element with an adjustable circle as described above. Below the circle, the resulting duration in minutes of each intensity unit is shown. The third interface element, which is the planning of specific activity units, is presented on the bottom of the page.

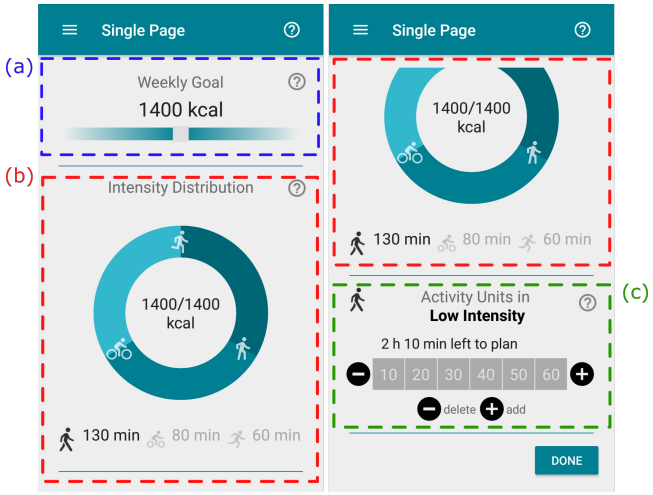


**Figure 1: Single Page Interface (translated from German) with (a) goal selection, (b) conversion into time, (c) activity unit planning**

## 5.2 Interface 2: Multiple Pages

In addition to the single page interface a second one was designed and implemented. This interface presents the three areas on successive pages (Figure 2). The first page contains the weekly goal with the slider for goal adjustment. The heading of this section has been changed and formulated as a request ("Choose your goal for the next week"). In addition, a visualisation below the slider converts the calorie goal to minutes of activity for each intensity level. Above the button "Intensity Distribution", which announces the next page, there is also a short explanation for the next page's content and why it is needed. This page's content is identical to the content of the intensity distribution area in the single page interface. The circle can be used to increase and decrease the intensities. The kcal unit is converted into a temporal unit accordingly. Here, too, the next step of the display is briefly explained before the button "Activity Units", which leads to the following page. The last page contains the input elements for the planning of specific activity units. Due to the increased amount of space compared to the single page interface, all intensity types can be listed at the same time, and don't have to be switched. Besides that, the interaction opportunities with the elements remain the same.

## 6 EVALUATION

The main aim of the evaluation was to investigate if the above described mechanisms and interface elements help to:

- empower users to understand the recommendation and its implications,
- reflect and evaluate the recommended goal, and
- provide the opportunity to actively manipulate it.

Moreover, we investigated if these supportive functions depend on the apps presentation mode (single or multiple pages). Further aspects of interest concerned user experience, usability, and aesthetics. All of these aspects can influence whether the user is willing to interact with the interface or not.

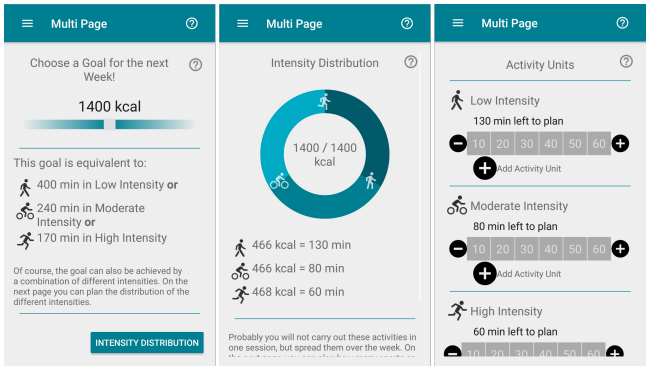


**Figure 2: Multiple Pages Interface (translated from German)**

We evaluated both interfaces with respect to our research questions in a laboratory setting with a mixed-method design. A between-subject design was chosen, in which the participants were randomly assigned to one of two groups. Depending on the group, one of the two interfaces was presented to the participants and they were given the following tasks:

First, we asked the participants to imagine to use an activity tracking app and presented them the interface in order to explore it. They were asked to select an activity goal for the next week based on the (simulated) system recommendation. Meanwhile, they should describe their thoughts and impressions with the think aloud method [6], which was recorded. Interaction behaviour was observed and documented. Afterwards, an online questionnaire was presented to the participants to collect quantitative data regarding the perception of the interface. In the online questionnaire we assessed usability, user experience and aesthetics. Therefore, the User-Experience questionnaire (UEQ) [15], the System Usability Scale (SUS) [4] , the Visual Aesthetics of Websites Inventory (VisAWI) [20] and the After-Scenario Questionnaire (ASQ) [16] were used. Finally, a semi-structured interview was conducted to get the participants' opinion about the interface and its supportive potential. The study took approximately 30 - 45 minutes.

## 7 RESULTS AND IMPLICATIONS

In total, 27 persons (group 1: $n$=13; group 2: $n$=14) participated in the study. They were aged 19 to 67 years ($M$=30.32, $SD$=17.19). 23 of them stated that they were interested in apps for physical activity support and 12 already had experience in using them.

## 7.1 Behaviour Observation and Think Aloud

We used behaviour observation and the think aloud method to objectively assess empowerment, reflection and appropriateness of manipulation elements. Therefore, interaction behaviour and the comments made during the interaction were documented, transcribed and analysed.

The think aloud results indicate empowerment and reflection processes. When initially interacting with the goal selection element, as expected, some participants (n = 9) mentioned problems in estimating, which number of kcal would be an adequate goal.

For example, one participant said:

> "I have no idea. 1000 kcal sounds good." (participant 22; multiple pages interface)

Later, when seeing the conversion into a temporal unit, she commented:

> "It's quite convenient, that it is given how much [time] that would be." (participant 22; multiple pages interface)

Another participant first commented:

> "No idea. I don't know what's average." (participant 24; multiple pages interface)

Later on when interacting with the activity units, she reflected:

> "I guess I could have chosen more in total. As it is, I have to do very little per day." (participant 24; multiple pages interface)

Transferring kcal into a temporal unit (intensity distribution element), helped to understand the recommended goal and reflect whether it is realistic or not. The aspect of interpretation of the kcal unit and the supportive potential of the conversion into time is also explicitly addressed in the interview (see below).

Also the order of interaction with the interface elements was analysed to objectively assess reflection. A non-linear order of interaction and revision of the choices previously made indicates that following elements fostered reflection of the selected goal planning. A linear order does not provide information about whether the recommendation was reflected (but intendedly not modified) or not. The observed order of interaction with the interface elements was the same with both interfaces. The majority of the participants completed the task in linear order (n=21). Seven persons (single page: $n$ = 4; multiple pages: $n$ = 3) operated with the elements in non-linear order and made adjustments to elements already used during processing. The order of interaction with the elements 1 (goal selection element), 2 (conversion into time element) and 3 (activity unit planning element) can be seen in table 2.

**Table 2: Order of Interaction with Interface Elements**

| Participant | Interface | Interaction Order |
|---|---|---|
| 4 | multiple | 1 → 2 → 3 → 1 → 2 → 3 → 2 → 3 |
| 7 | single | 1 → 2 → 3 → 2 → 3 → 2 → 3 |
| 8 | multiple | 1 → 2 → 3 → 2 → 1 → 2 → 3 |
| 11 | single | 1 → 2 → 3 → 2 → 3 |
| 12 | multiple | 1 → 2 → 3 → 2 → 3 |
| 24 | multiple | 1 → 2 → 3 → 2 → 3 |
| 25 | single | 1 → 2 → 1 → 2 → 3 |

To evaluate if the provided opportunities to manipulate the recommendation are appropriate, we analysed correctness of handling as well as reported and observed difficulties: During the interaction with both interfaces, the first interface element (goal selection through slider) was clear for most participants. Two participants had general questions of understanding and stated that it could not be set precisely. The time unit was clear, also. The only interpretation problem that occurred, was that one participant asked for a conversion from minutes into hours. Some participants were confused by the different activity levels. However, the majority of the participants ($n$ = 19) seemed to understand the intensity distribution without help of the supervisor, eight asked for help. Regarding the activity units (interface element 3), 15 participants asked for help. After hinting at the help texts integrated in the interface, help was provided, if still necessary. The goal selection element was explained to two participants, the conversion into time with the intensity distribution to seven participants and the activity units to 14 participants. Despite problems with usability and understanding, the elements were commended in their function and meaning.

## 7.2 Interview

For the interviews, we used the qualitative content analysis [17]. The interviews were transcribed, afterwards the statements were clustered and analysed. Due to recording problems, one data set is missing, so that the interview section is based on 26 data sets. In the following, the results are presented and enriched by a few exemplary translated statements. Since the two interface versions do primarily differ in the distribution of the elements on different pages, the results are mainly presented for both interfaces simultaneously.

*7.2.1 Overall Satisfaction.* The overall evaluation of the participants was positive for both interfaces. In general, the interfaces were rated positively.

> "It's quiet good. I usually use a pedometer. There it [goal planning] is just implicit. Here I can exactly plan how much I can make a day, [...] do I just want low intensity or modify and increase it." (participant 7; single page interface)

*7.2.2 Understanding and Supportive Potential.* As an important factor for the supportive potential of the interface elements, we asked participants if they understood the dependencies of the three interface elements. In order to objectively evaluate the understanding, we further asked them to explain these dependencies. In most cases ($n$ = 25), the dependencies between weekly goal, conversion into time and activity units were explained correctly. However, some participants reported initial comprehension problems.

> "At the beginning you are thrown in a bit, but that is actually okay, because it is so clearly structured." (participant 13; single page interface)

Although we did not explicitly ask for it, the main idea that the goal could be adjusted was mentioned positively.

Regarding the first element, we asked the participants if they found the unit kilocalories meaningful. As stated above, we expected that kilocalories might be too abstract for the users. But, since it is the common unit for activity tracking, it was known by all participants. However, eight of them (single page interface: n = 5; multiple pages interface: n = 3) reported problems in interpreting it as they were not used to the unit. For instance, they asked for reference values that would have been helpful. Although there were also many participants who did not report problems with the kilocalories unit, all except one participant agreed that the conversion into time (interface element 2) helped to interpret and estimate the goal. The separation of activities into different intensities was not easy to understand for 6 persons. Others liked this aspect:

> "I liked that you could divide it into intensive, moderate and low activity, because you could split it up even if you were not an athletic kind of person. You can see that you can achieve a lot even by just climbing stairs or going for a walk." (participant 20; multiple pages interface)

With regard to the third interface element - the planning of activity units - most participants (n = 20) found it helpful. For example one participant said:

> "[...] It allows a bit of back and forth planning." (participant 11; single page interface)

Three of them even wished to have further functionalities: One participant wished to assign the selected units to concrete activities, two would like to assign them to specific days. A calender function was also desired by further participants when directly asked for additional functionality in another interview question. Three of the participants who found the planning of activity units helpful, reported (initial) problems in understanding. Four participants (tendentially) did not find the activity units helpful. Three of them reported problems in understanding. Two participants did not make a precise statement whether the unit planning element was supportive for them or not.

*7.2.3 Usability.* Despite positive usability ratings (see section 7.3) some usability problems that occurred when operating with the interface elements, could be identified in the interview. In general, difficulties in operation were often ($n$ = 13) mentioned.

> "What bothered me was that sometimes it was hard to select the things". (participant 11; single page interface)

Although there were many positive responses to the circular seekbar that was used for intensity distribution, four participants had operating problems when moving the pointers. Regarding the planning of activity units (third interface element), it was still desired that the addition and removal of activity units should not only be possible by using the buttons, but also by touching the respective units.

*7.2.4 Aesthetics.* Also the aesthetics were rated predominantly positive for both interfaces. The colour and graphical presentation of the page was particularly rated as outstanding.

> "The colours fit well together. It is not obtrusive or boring". (participant 19; single page interface)

However, other participants found the colour design to be too uniform ($n$ = 3). Especially the circular seekbar was criticised for being too similar or boring.

> "The colours were too similar that you had to look exactly what you had just selected." (participant 12; multiple pages interface)

## 7.3 Online Questionnaire

With an online questionnaire we assessed user experience, aesthetics and usability. Due to extreme values, identified by box plot diagrams, some participants were excluded from the analysis of the questionnaires. 22 participants (group 1: $n$=11; group 2: $n$=11) remained. 91% of them stated that they were interested in apps for physical activity support and 50% also had experience in using them.

**Table 3: Descriptive Results of the Single Page Interface**

| | Single Page | | Multi Page | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| UEQ attractiveness | 1.58 | .66 | 1.80 | .60 |
| UEQ perspicuity | 1.91 | .94 | 2.08 | .49 |
| UEQ efficiency | 1.92 | .86 | 1.36 | .56 |
| UEQ dependability | 1.34 | .71 | 1.52 | .75 |
| UEQ stimulation | 1.11 | .62 | 1.30 | .44 |
| UEQ novelty | 1.11 | .73 | 1.17 | .59 |
| VisAWI | 5.85 | .38 | 5.78 | .63 |
| VisAWI simplicity | 6.00 | .51 | 5.69 | .63 |
| VisAWI diversity | 5.07 | .71 | 5.45 | .77 |
| VisAWI colourfulness | 6.41 | .45 | 6.05 | .99 |
| VisAWI craftsmanship | 6.09 | .38 | 6.05 | .66 |
| ASQ | 3.00 | 1.38 | 2.45 | 1.66 |
| SUS | 77.27 | 13.80 | 80.69 | 10.43 |

*7.3.1 UEQ.* Both interfaces have a good UEQ score (see table 3). A value higher than .80 is an indicator for a positive rating, while a value lower than -.80 is an indicator for a negative rating. Results of the multiple pages interface are descriptively higher in all sub scales, except for the result of efficiency. We tested for significance with a t-test for the UEQ perspicuity scale and - due to missing prerequisites (normal distribution or variance homogeneity) - a Mann-Whitney-U-test for the remaining subscales. There were no significant results.

*7.3.2 VisAWI.* For the VisAWI there are no reference values for interpretation given by the authors. They are stating, that lower values imply a negative rating and higher values a positive one. In this case, ratings from 1 to 3.5 are interpreted negative and ratings from 3.5 to 7 positive, since the scale is from 1 to 7. In general, the single page interface ($M$=5.85, $SD$=.38) as well as the multiple page interface ($M$=5.78, $SD$=.63) were rated positive (table 3) and do not differ significantly from each other (Mann-Whitney-U-Test: U = 57.700, $p$ = .847). The same applies for the subscales.

*7.3.3 ASQ.* For the ASQ questionnaire, there were also no reference values given for interpretation. The scale ranges from 1 (very positive) to 7 (very negative). As before, for interpretation we split the scale. Values from 1 to 3.5 are interpreted positively and values from 3.5 to 7 negatively. As presented in table 3, the single page interface has an average of 3.0 ($SD$=1.38) and the multiple pages interface of 2.45 ($SD$=1.66). Both values are therefore evaluated positive, whereby the multiple pages interface performed better. However, the differences are not significant (Mann-Whitney-U-Test: U = 41.000, $p$ = .217).

*7.3.4 SUS.* The SUS score can have a range from 0 to 100. According to Bangor et al. [2] the single page interface was rated “good” and the multiple pages interface “excellent” (see table 3). Therefore, the multiple pages interface had a better result, what however did not appear significant (Mann-Whitney-U-Test: U = 54.500, $p$=.699).

# 8 DISCUSSION

Three main interface elements were designed and evaluated in two slightly different interfaces. The three elements are 1) an adjustable slider showing a recommendation range together with an indication of uncertainty, 2) conversion into time depending on activity level, and 3) activity unit planning. They were designed to empower users to understand the recommendation and its implications, to reflect and evaluate it, and to provide the opportunity to actively manipulate it. In the following, we discuss if and how the interface elements reached those aims.

## 8.1 Comparison between the Interfaces

Results do not show any significant differences between the two interfaces in interaction behaviour, feedback regarding the perceived support of the interface elements, usability, user experience or aesthetics. Returning to previous interface elements to make adjustments, was assumed to be more difficult for navigation on different pages in the multiple pages interface. However, since the order of interaction does not vary between the interfaces, both arrangements of elements (single page interface: on one page, multiple pages interface: on more consecutive pages) seem possible. Consequently, whether those interrelated elements are arranged on single or on multiple pages doesn't seem to be an influencing factor for recommendation reflection and modification. As all results are independent from presentation mode, we will subsequently discuss them without separation between the two interface variants.

## 8.2 Overall Interface

All participants interacted with all interface elements. This indicates that they were perceived as helpful and have the potential to integrate users in the process. As a limitation, it can not surely be said, which amount of interaction was fostered by the study situation and which by the interface design itself. However, participants were not explicitly told to interact with each interface element. On the one hand, the study situation could have fostered interaction with the elements. On the other hand, it is likely, that users of an activity tracking app in a real-world setting (compared to a study setting) are more intrinsically motivated to interact with the system and choose the most adequate goal. Moreover, in real-world settings there is much more time to get used to the system and its interface elements. Exploration of and interaction with the interface could be enhanced week by week.

It has been shown that most people interacted with the elements in a sequential order. For the single page interface this means that the elements were operated from top to bottom and for the multiple pages interface that the elements were operated page by page. Possible reason are (1) that they were satisfied with the initially chosen value, or (2) that they considered their study task to be completed when testing and understanding all elements and did not see the need to actually find an appropriate goal in the study setting, or (3) that for those people reflection was not stimulated sufficiently to lead to interaction. However, more than 20 percent of the participants returned to previous elements during the interaction process to make adjustments. Stepping back from one interface element to another, indicates that exploration and interaction has not only been done for the study, but actually stimulated reflection.

Some understanding problems occurred. As explanation was given during the study, this should not confound the results. Probably, these problems would partially be solved through extended, unobserved exploration in real-world settings. As already suggested above, functionality and exploration of the interface could be enhanced week by week. Additionally, we addressed understanding problems in the re-design (see section 9) as far as possible.

The questionnaires UEQ and VisAWI addressed the aesthetics of the interfaces and showed positive ratings. Regarding usability, ASQ ratings were positive in tendency and very positive values were achieved in the SUS. In the interview, the colours were also predominantly commended, whereas a few people felt that the colour design was too uniform. However, the interview and interaction showed that some participants had problems with the operation, e.g. with the exact setting of the goal selection element or with the selection of elements, which can be corrected by technical adjustments. In total, overall ratings for usability, user experience and aesthetics are good. This is an important precondition for whether the user is willing to interact with the interface or not. This precondition can be seen as fulfilled for the study and the interfaces.

## 8.3 Interface Element 1: Goal Selection

All participants made use of the slider, which indicates that a modifiable value within a range seems more appropriate than one single recommended value. Think aloud comments show, that they reflected, what would be an appropriate goal. However, as expected and in line with the interview results, the limited range and the default value alone do not seem to sufficiently support users in evaluating, what an appropriate value would be. Additional empowerment is needed here (see 8.4). One participant suggested a reference value to better interpret the recommended goal. This is surprising as we intended the goal range and the default value (and the colour gradient) to serve as such a reference. At least this one participant does not seem to interpret it in the intended way. As participants did not seem to pay attention to the colour gradient indicating the certainty of the result, this could not help in terms of empowerment and reflection, which contradicts former research [8]. One reason might be that the interface contained more elements than the interfaces used in the cited literature and therefore the focus of participants was different. Another reason might be, that in the presented work the colour gradient was from light blue to dark blue. In the cited study, the gradient had a colour coding with the colours red, yellow and green.

## 8.4 Interface Element 2: Conversion into Time

Think aloud, behaviour observation and interview results all indicate, that conversion into a temporal unit strongly empowered users to better estimate and reflect what is an appropriate goal value on the recommended range. For some participants this reflection lead to revision of the initially chosen goal. Conversion into time goes along with separation into different activity levels, which was difficult to understand for some participants and thus lowered the intended empowerment. Unfortunately, this separation is unavoidable, as the conversion otherwise would have been far too inexact and not meaningful anymore. However, results show that nevertheless the interface element can have the intended supportive

potential. Although, there were some problems in handling, which need to be resolved, all participants interacted with the interface element.

### 8.5 Interface Element 3: Activity Unit Planning

Also the activity unit was used by all participants and the majority found it helpful. Some wished for modified elements or opportunities that allow for an even more detailed planning. Reconsidering the initially selected goal and stepping back to the first interface element demonstrates that such a planning element can in principal empower users to reflect the recommendation. The usability and understanding problems reported by some participants are addressed in the re-design of the interface presented below.

## 9 SUMMARY, RE-DESIGN AND CONCLUSION

We investigated if user interface design can in principal support users of an activity tracking system in understanding and reflecting the system's goal recommendation as a basis for appropriately exerting influence on the recommendation result. This kind of user integration is important as recommendations in this field are very error-prone because of large variations over time and a large number of confounding variables that are indeterminable for the system. In our presented approach, we pursue three main aims, which are (a) to empower users to understand the recommendation and its implications, (b) to reflect and evaluate it and (c) to provide the opportunity to actively manipulate it. In our approach, these aims are pursued through transparency of the algorithm uncertainty and by providing activity planning elements, which are intended to have explanatory function regarding the impact of the recommendation and to stimulate reflection of the recommended goal. Further, these elements should enable and support users in appropriate modifications of the recommended goal. We designed two different interfaces with three elements: A modifiable slider for goal selection showing a recommendation range, default value and colour indicator for probability of suitability of the recommendation; a conversion of the goal unit (kcal) into a temporal unit (minutes) in conjunction with different activity levels; an element to plan concrete activity units, i.e. how often and how long users plan to be active to achieve their goal. The two interface variants differed in the arrangement of the elements (single page or multiple pages). We evaluated the interface with regard to the three main aims presented above. Results were the same for both interface variants. They show that there is a need of user empowerment and that empowerment can be reached by interface elements that explain the impact of the recommendation. In this case, the second interface element achieved this by converting the goal from an abstract unit to a unit, participants are more used to and which is more conceivable. The third interface element works by further illustrating what is necessary in the daily life to achieve the goal. The study shows, that both of these interface elements can support reflection of the recommendation. Exerting active influence on the recommendation was initially stimulated by just providing the opportunity to do so, with the first interface element. Additionally, results show that as a consequence of reflection, stimulated by interface elements 2 and 3, further manipulation of the recommended goal was fostered.

There are some limitations and potential for improvement. Although the work described focuses on investigating the research questions, it is part of the user centered development of a broader behavior change application. Therefore, we re-designed the interface (Figure 3). We changed the slider's colour gradient and added colour coding from green to orange. As some few participants encountered problems in distinguishing the intensities in the circular display, we increased the contrast of the colours so that they are easier to distinguish. Addressing the understanding problems, we revised the help texts for all elements. Regarding the activity unit planning element, we redesigned the element to make it more understandable and meet the user demand for a more detailed planning, such as a calendar function. Instead of the fields of 10 minute intervals per activity unit, we now provide one field per day of the week. Duration of the daily activity time can be modified via plus or minus buttons. We further modified the label.

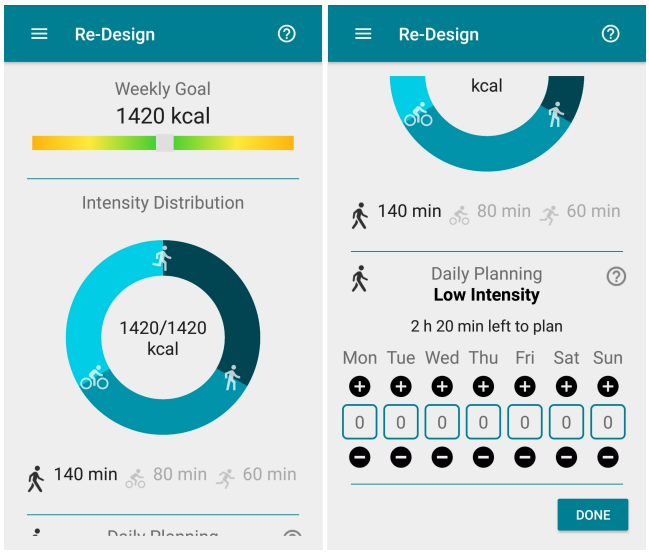


**Figure 3: Interface Re-Design (translated from German)**

Those improvements refer to the specific design of the specific interface elements used in this study. As they are exemplary implementations for interface elements, the revealed limitations do not affect the gain of knowledge regarding the research question. It can be concluded that implementation planning elements in particular and interface elements in general have the potential to empower users, support recommendation reflection and foster user interaction with the recommendation.

## ACKNOWLEDGMENTS

This work is part of the research project Personal Analytics, funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF), reference number: 16SV7110, aquired and headed by Aysegül Dogangün.

# REFERENCES


[1] Laurence L. Alpay, Olivier Blanson Henkemans, Wilma Otten, Ton A. J. M. Rövekamp, and Adrie C. M. Dumay. 2010. E-health applications and services for patient empowerment: directions for best practices in The Netherlands. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association* 16, 7 (2010), 787–791. https://doi.org/10.1089/tmj.2009.0156

[2] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123.

[3] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and A. Gutiérrez. 2013. Recommender systems survey. *Knowledge-Based Systems* 46 (2013), 109–132. https://doi.org/10.1016/j.knosys.2013.03.012

[4] John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[5] Alexander Felfernig, Robin Burke, and Pearl Pu. 2012. Preface to the special issue on user interfaces for recommender systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 313–316. https://doi.org/10.1007/s11257-012-9120-5

[6] Marsha E Fonteyn, Benjamin Kuipers, and Susan J Grobe. 1993. A description of think aloud method and protocol analysis. *Qualitative health research* 3, 4 (1993), 430–441.

[7] Juan M. García-Gómez, Isabel de La Torre-Díez, Javier Vicente, Montserrat Robles, Miguel López-Coronado, and Joel J. Rodrigues. 2014. Analysis of mobile health applications for a broad spectrum of consumers: a user experience approach. *Health informatics journal* 20, 1 (2014), 74–84. https://doi.org/10.1177/1460458213479598

[8] Katja Herrmanny and Aysegül Dogangün. 2017. The Impact of Prediction Uncertainty in Recommendations for Health-Related Behavior. In *Proceedings of the 2nd International Workshop on Health Recommender Systems co-located with the 11th International Conference on Recommender Systems (RecSys 2017)*, David Elsweiler, Santiago Hors-Fraile, Bernd Ludwig, Alan Said, Hanna Schäfer, Christoph Trattner, Helma Torkamaan, and André Calero Valdez (Eds.). RWTH, 14–17. http://ceur-ws.org/Vol-1953/healthRecSys17_paper_8.pdf

[9] Katja Herrmanny, Jürgen Ziegler, and Aysegül Dogangün. 2016. Supporting Users in Setting Effective Goals in Activity Tracking. In *Persuasive Technology*, Alexander Meschtscherjakov, Boris de Ruyter, Verena Fuchsberger, Martin Murer, and Manfred Tscheligi (Eds.). Springer International Publishing, Cham, 15–26.

[10] Jamil Hussain, Wajahat Ali Khan, Muhammad Afzal, Maqbool Hussain, Byeong Ho Kang, and Sungyoung Lee. 2014. Adaptive User Interface and User Experience Based Authoring Tool for Recommendation Systems. In *Ubiquitous Computing and Ambient Intelligence. Personalisation and User Adapted Services*, Ramón Hervás, Sungyoung Lee, Chris Nugent, and José Bravo (Eds.). Lecture Notes in Computer Science, Vol. 8867. Springer International Publishing, Cham, 136–142. https://doi.org/10.1007/978-3-319-13102-3{_}24

[11] Dosam Hwang, Xuan Hau Pham, and Jason J. Jung. 2011. Preference-based user rate correction process for interactive recommendation systems. In *iiWAS 2011*, David Taniar, Eric Pardede, Hong-Quang Nguyen, Wenny Rhayu, and Ismail Khall (Eds.). ACM Press, New York, New York, USA, 412. https://doi.org/10.1145/2095536.2095619

[12] Dietmar Jannach, Ingrid Nunes, and Michael Jugovac. 2017. Interacting with Recommender Systems. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion - IUI '17 Companion*, George A. Papadopoulos, Tsvi Kuflik, Fang Chen, Carlos Duarte, and Wai-Tat Fu (Eds.). ACM Press, New York, New York, USA, 25–27. https://doi.org/10.1145/3030024.3030027

[13] Bart P. Knijnenburg, Martijn C. Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 441–504. https://doi.org/10.1007/s11257-011-9118-4

[14] Haridimos Kondylakis, Lefteris Koumakis, Eleni Kazantzaki, Maria Chatzimina, Maria Psaraki, Kostas Marias, and Manolis Tsiknakis. 2015. Patient Empowerment through Personal Medical Recommendations. In *MEDINFO 2015: eHealth-enabled health*, Indra Neil Sarkar, Andrew Georgiou, and Paulo Mazzoncini de Azevedo Marques (Eds.). IOS Press, Amsterdam and Berlin and Tokyo and Waschington, DC, 1117. https://doi.org/10.3233/978-1-61499-564-7-1117

[15] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*, Andreas Holzinger (Ed.). Lecture Notes in Computer Science, Vol. 5298. Springer Berlin Heidelberg, Berlin, Heidelberg, 63–76. https://doi.org/10.1007/978-3-540-89350-9{_}6

[16] James R. Lewis. 1991. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM Sigchi Bulletin* 23, 1 (1991), 78–81.

[17] Philipp Mayring. 2010. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Beltz. https://books.google.de/books?id=BJlxSQAACAAJ

[18] Pearl Pu and Li Chen. 2007. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems* 20, 6 (2007), 542–556.

[19] Hanna Schäfer, Santiago Hors-Fraile, Raghav Pavan Karumur, André Calero Valdez, Alan Said, Helma Torkamaan, Tom Ulmer, and Christoph Trattner. 2017. Towards Health (Aware) Recommender Systems. In *Proceedings of the 2017 International Conference on Digital Health - DH '17*, Patty Kostkova, Floriana Grasso, Carlos Castillo, Yelena Mejova, and Arnold Bosman (Eds.). ACM Press, New York, New York, USA, 157–161. https://doi.org/10.1145/3079452.3079499

[20] Meinald T Thielsch and Morten Moshagen. 2011. Erfassung visueller Ästhetik mit dem VisAWI. *Tagungsband UP11* (2011).

[21] Bo Xiao and Izak Benbasat. 2007. E-commerce product recommendation agents: use, characteristics, and impact. *MIS quarterly* 31, 1 (2007), 137–209.