Jun 25th, 9:00 AM

Reflection during goal setting: An analysis of popular personal informatics apps

Tina Ekhtiar
University of Twente, Netherlands, The

Rúben Gouveia
University of Twente, Netherlands, The

Armağan Karahanoğlu
University of Twente, Netherlands, The

Geke Ludden
University of Twente, Netherlands, The

Citation

Ekhtiar, T., Gouveia, R., Karahanoğlu, A., and Ludden, G. (2022) Reflection during goal setting: An analysis of popular personal informatics apps, in Lockton, D., Lloyd, P., Lenzi, S. (eds.), *DRS2022: Bilbao*, 25 June - 3 July, Bilbao, Spain. https://doi.org/10.21606/drs.2022.787

DRS 2022
Bilbao 25th June — 3rd July

Reflection during goal setting: An analysis of popular personal informatics apps

Tina Ekhtiar*, Rúben Gouveia, Armağan Karahanoğlu, Geke Ludden

Faculty of Engineering Technology, Interaction Design Research Group, University of Twente, the Netherlands

*corresponding e-mail: t.ekhtiar@utwente.nl

doi.org/10.21606/drs.2022.787

Abstract: Setting achievable, realistic goals aligned to people's current abilities and needs is an important part of behavior change. Reflection is a critical part of this process, as it helps people identify and elicit appropriate goals. Commercial applications that support behavior change often use goal setting; however, we know very little about if these implementations are based on theory and how they support reflection. In this paper, we analyze how popular health tracking apps support reflection during goal setting. We found that a majority supported limited or no reflection during the elicitation of goals. We discuss our findings and suggest design considerations for improving how these tools can help people in reflecting and eliciting goals.

Keywords: goal setting; reflection; personal informatics

1. Introduction

Personal informatics (PI) refers to a class of tools that help people track, monitor, and reflect on their behaviors, with the purpose of self-reflection and improving self-knowledge (Epstein et al., 2020). These tools are increasingly prevalent in people's everyday lives, across diverse goals that include behavior change, self-learning, or simply satisfying curiosity.

Goal setting, while being one of the most theoretically researched behavior change techniques, is also one of the most represented techniques used in PI. For example, a recent analysis of top-ranked mobile apps for physical activity found nearly 50% of PI include goal setting in their design (Baretta et al., 2019). As a theory, goal setting is fundamentally focused on achievement motivation (Locke & Latham, 2019), and is based on the premise that conscious goals affect action (Ryan, 1970, as cited in Locke & Latham, 2002). It builds on the idea that human actions are often purposeful and based on conscious goals, and, by setting differing performance goals, individuals can achieve various levels of performance, with varying levels of motivation (Latham et al., 1994). Locke and Latham (2002) state these goals should be specific, challenging, and proximal to be effective. Setting goals is an important aspect of behavior change, as goals affect decision making, clarify expectations, and provide reference for goal achievement (Consolvo, 2012).

Central to goal setting is the concept of reflection, a form of mental processing with a purpose and/or an anticipated outcome (Moon, 1999 as cited in Fleck & Fitzpatrick, 2010). Reflection is a developmental process that takes time and can be motivated by different purposes (Baumer et al., 2014; Fleck & Fitzpatrick, 2010). The process of reflection helps self-learning and understanding different relationships and perspectives, which can lead to challenging personal assumptions and fundamental change (Fleck & Fitzpatrick, 2010). Prior work on PI has highlighted the benefits of reflection for helping people understand their capabilities and set goals that are realistic and achievable (Saksono et al., 2021; Kocielnik et al., 2018). Brief reflective activities during goal setting can influence people in choosing ambitious, yet achievable goals (Lee et al., 2015). In addition, when people take time to envision their goals and ways to overcome possible obstacles, they are more likely to set realistic goals (Avishai et al., 2019; Peng et al., 2021).

In this paper we investigate how highly popular commercial apps help people in eliciting goals and how reflection is facilitated in this process. Our study focuses on a small number of highly popular apps. We focus on highly popular apps as these are at increased risk of offering ineffective care or even causing harm due to their widespread outreach (Powell et al., 2016).

2. Method

We started our study by compiling a list of the most downloaded “Health & Fitness” mobile apps, regardless of the behavioral domain, in the United States’ Google play and Apple stores, as of November 1st, 2021 (Ceci, 2021; Apptopia, 2021). We then filtered this list of apps to provide a tighter focus on personal informatics and goal setting. First, we excluded apps that did not have self-tracking features (i.e., that did not support the tracking of some type of health behavior). Next, we excluded apps that did not allow users to elicit a goal. For instance, some of the apps in our initial list supported tracking but did not ask people what their tracking (or health) goals were, nor allow users to set goals through the app. We also excluded apps that required immediate payment or did not offer a free trial period. Finally, we selected the 10 most downloaded apps from this filtered list (see Table 1). Each of these apps was highly downloaded, with at least 10,000,000 installs.

The selected apps covered a variety of tracking domains. Physical activity made up most of the sample (3 of 10 apps), followed by physical activity and weight and food (3 apps), period and pregnancy (2 apps) and mindfulness (2 apps).

*Table 1. List of analyzed apps. *According to Google Play store, as of November 2021*

App	Tracking Domain	Installs*
Calm (Calm.com, 2012)	Mindfulness	10,000,000+
Headspace (Headspace Inc., 2012)	Mindfulness	10,000,000+
Flo (Flo Health, Inc., 2015)	Period and pregnancy	50,000,000+
Period Tracker Period Calendar (Abishkking Limited, 2014)	Period and pregnancy	100,000,000+
Home Workout (Abishkking Limited, 2017)	Physical activity	100,000,000+
Sweatcoin (Sweatco Ltd, 2016)	Physical activity	10,000,000+
Six Pack in 30 Days (Abishkking Limited, 2018)	Physical activity	100,000,000+
Fitbit (Fitbit, Inc., 2014)	Physical activity, weight, and food	50,000,000+
Lose it! (FitNow, 2008)	Physical activity, weight, and food	10,000,000+
MyFitnessPal (MyFitnessPal, Inc., 2005)	Physical activity, weight, and food	50,000,000+

2.1. Measures

Goal Setting Theory (Locke & Latham, 2002) and Fleck and Fitzpatrick's levels of reflection (2010) were used as a theoretical reference for identifying how the different apps (i) supported key goal properties and sources for goals, and (ii) facilitated reflection during goal setting.

Key properties and sources for goals

Shilts and colleagues (2004) identify three properties for effective goal setting: *proximity* (i.e., the timeframe for achieving a goal), *specificity* (i.e., the explicitness of a goal) and *difficulty* (i.e., how challenging a goal is).

Goal source, another important characteristic of goal setting, clarifies the role of the user in setting the goal (Shilts et al., 2004). Accordingly, goals can either be *self-set* (i.e., people set their own goals)*, assigned* (i.e., goals are assigned to people, without their input), *participatory* (i.e., goals are designed both by the person as well as the app and/or other experts), *guided* (i.e., people choose from a list of goal suggestions) and *group-set* (i.e., goals are designed by - and for a group of people (Shilts et al., 2004).

Levels of reflection during goal setting.

Fleck & Fitzpatrick (2010) describe five levels (or ways) in which reflection can be facilitated. The lowest level, R0, is *description*. This level is described as a "description or statement about events without further elaboration or explanation" (Fleck & Fitzpatrick, 2010, p.2). In the context of personal informatics and goal setting, this level of reflection would occur whenever a user is presented with or asked about nothing more than descriptive data on their own behaviors - or the behaviors of other people - and asked to set a goal. For example, PI might display the average number of steps that a user has walked over the last week and ask them to set a daily step count goal.

The second level of reflection (R1, descriptive reflection) occurs when some explanation or interpretation accompanies the statements and events described in R0. However, these explanations are often descriptive, with limited analysis and focused on only one perspective. In the context of PI and goal setting, this level of reflection could occur whenever a user is provided with a broad interpretation of their data and asked to set a goal. For example, a PI might display a summary and analysis of a users' weekly step count and ask the user what they were doing on average high or low days (i.e., going to work, being with friends, or staying home).

The third level of reflection (R2, dialogic reflection) focuses on exploring relationships between different statements and events, with the goal of facilitating the consideration of different explanations, hypothesis, and points of view. In the context of personal informatics and goal setting, this would occur when users are presented with multiple sources of data on their own behaviors and asked to set a goal. For instance, an activity tracker could help people in understanding their physical activity behaviors by showing them correlations between their activity and sleep data, or how busy their calendar is. This way people are guided through making and reflecting on connections with their behavior while setting a goal.

The fourth level of reflection (R3, transformative reflection) occurs when the revisiting of statements or events is done with the intent to change a behavior. PI tools could support this level by initiating a conversation about the things that users would like to change in their life, while users revisit their data and set goals. For example, asking people what days they are less physically active, what would be different ways they could incorporate more physical activity and making a goal for making this change. PI tools could also suggest explicit, challenging goal suggestions to people while revisiting their data, to encourage change.

The fifth and final level of reflection (R4, critical reflection) occurs when people are supported in thinking of the broader implications of their behaviors within societies and the broader world, rather than solely on their own individual behavior. For example, an activity

tracker could highlight some of the ways in which a person's physical activity and goal setting behaviors could impact health care systems, by sending them the following message: 'Setting a challenging goal could improve your health and ultimately lead to less hospital intakes and avoidable health spending. The whole health care system improves when you take care of yourself!'.

In addition to these five levels, we also consider cases in which reflection does not occur during goal setting (i.e., where none of the previous categories fit in the process of goal setting). An example would include assigning a person to a 10,000-step count goal without any explanation or revisitation of their data.

2.2 Procedure

The first two authors downloaded each of the 10 identified apps and familiarized themselves with how each supported the elicitation of goals. A focus was placed on the onboarding process, i.e. how goal elicitation was supported during initial usage, following app installation (Kapusy & Lógó, 2020). PI tools often ask people to set goals during early interactions with these tools (Middelweerd et al., 2014; Conroy et al., 2014). However, during onboarding, tools often have limited knowledge of people's needs, understandings and capabilities. Further, people might also have little knowledge of their own abilities, especially when new to tracking (Karahanoğlu et al., 2021). This can lead to the misaligment of goals (Munson et al., 2020). Therefore, studying how apps support goal setting from an initial point of usage is imperative.

Both authors started by capturing the total number of goals that each app asked people to set during the app onboarding. Apps often supported multiple goals, as we later discuss in our results. We considered both quantitative (e.g., walking 10,000 steps per day) as well as qualitative (e.g., I want to feel healthier) goals. Each identified goal was then characterized in terms of its properties, source, and level of reflection. During this process, screenshots and a video of the app's interactions were captured to facilitate recalling and discussing the coding or possible disagreements between the authors. Disagreements were resolved by a third researcher, who also installed and analyzed the apps. This was followed by a discussion with the first two authors. The analysis was revised and discussed until consensus was reached.

3. Results

The following section describes how the analyzed apps supported key goal properties and sources and facilitated reflection during goal setting.

3.1 Goal properties

A total of 57 goals were identified across the 10 analyzed apps. On median, each app involved users in the elicitation of 4 goals (IQR=2-8) during onboarding. These goals were of-

ten related to different health behaviors. For example, the onboarding process of MyFitness-Pal included setting an overall goal for using the app followed by a weight, sleep, and physical activity goal.

Approximately half of these goals were quantitative (54.4%, n=31), expressed by numbers. Examples included setting a daily step count goal or a target weight goal. The remaining goals were qualitative (45.6%, n=26). These were often not tracked, serving mostly for customizing the appearance and feedback given by an app. For example, most period and pregnancy apps would ask users to choose a goal for using the app (e.g. track pregnancy, track fertility) and provide customized content for the chosen goal.

Regarding goal proximity, we found that approximately half of the goals had a specific timeframe for being completed (n=30, 53%). The remaining had undefined timeframes. Interestingly, we found this to be strongly dependent on if a goal was quantitative or qualitative. Of the 31 quantitative goals, 27 had a specific timeframe for being completed. While for the 26 that were qualitative, only 1 was associated to a specific timeframe (see Table 2).

Table 2. Type of Goals & Proximity. Goals sometimes had more than one timeframe (e.g., an overall weight loss goal is defined by weekly weight loss goal

Goal Type	**Time Frame**	**# of Goals**
Quantitative	None / Indefinite	5
	Daily	18
	Weekly	9
	Monthly	2
	Yearly	2
Qualitative	None	25
	Weekly	1

3.2 Goal sources

Next, we analyzed the role of users in choosing a goal. Over half of the goals (n=32, 56%) were guided, with the app showing users a list of goals that they could then choose from. For example, all the fitness applications provided users with a list of general health goals which they could elicit (e.g., weight loss or reducing stress). The remaining were either assigned (n=17, 30%) or a minority was self-set by users (n=8, 14%).

We found that the applications differed with respect to the amount of explanation given for the goal that was eventually set, whether it was guided or assigned to the user.

Assigned goals were often found to lack further explanation of why it was set in a particular way. For example, in MyFitnessPal, the weekly weight loss goal would state “Recommended” on losing one pound per week, however it was not stated who or what source the

recommendation came from. Another instance is in Fitbit's water consumption goal. The app assigns a specific water consumption goal that cannot be changed and does not explain how it is set or if it changes depending on the user's activity.

The lack of explanation was found in guided goals as well. For example, in MyFitnessPal, Home Workout, and Lose it! the apps would give options for a goal based on different levels of "intensity" (or how difficult the goal is). No reason would be made for what made the goal difficult or what commitment had to be made to achieve the more challenging goal.

A small number of goals (n=2, 4%) made this process more informative by, for example, using historical data of the user (i.e., the data that is tracked and stored in mobile phones before the app is downloaded) and data of the general population to recommend goals to users. For instance, Sweatcoin fetches and shows the step count data of the user for the past year to guide setting monthly goals. The app can do this through connecting to iPhone's collected step count data. MyFitnessPal uses general health data to give a recommendation of body fat percentage goals based on gender

3.3 How was reflection supported during goal setting?

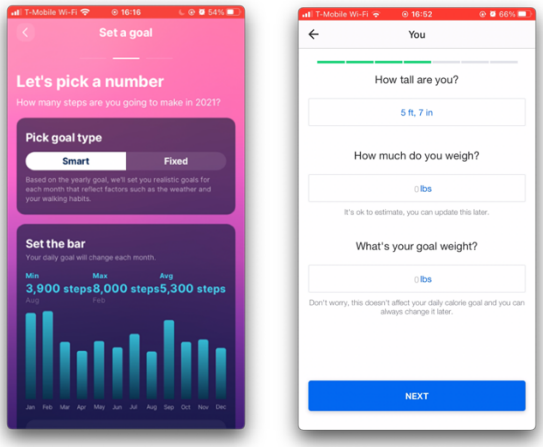
Most apps fostered little or no reflection during the initial elicitation of goals. In fact, only 14 goals (of the identified 57, 24%) supported at least one of Fleck & Fitzpatrick's levels of reflection. Three of the goals supported two levels of reflection.

Table 3. Reflection Levels of Goals Analyzed.

Reflection Level	**# of Goals**
None	43
R0	10
R1	4
R2	1
R3	1
R4	1

The most commonly supported level of reflection was R0 (i.e., description level). In these cases, apps collected and displayed summaries of personal data as users were asked to choose a goal. For example, Sweatcoin started its onboarding process by importing physical activity data from people's phones to the app. Users were shown summaries of their physical activity levels over an entire year and this data was used as a baseline for suggesting monthly goals to users (see Fig. 1.1). Most apps, however, did not import pre-existing data of users. Rather, users were more commonly asked to input different physiological metrics

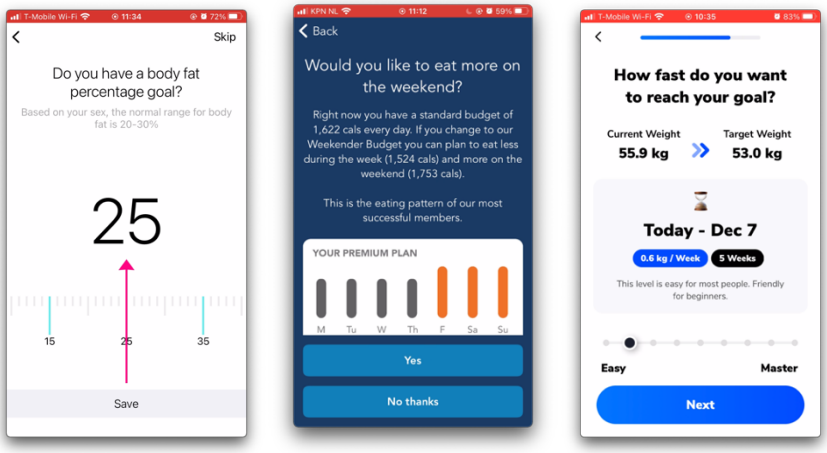
or shown descriptive data from other users before being asked to elicit a goal. For instance, in the case of MyFitnessPal (see Fig. 1.2.), users were asked to state their current weight, followed by their goal weight.



(1) Sweatcoin *(2) MyFitnessPal*

Fig. 1. Apps that supported reflection level R0 (description) during goal elicitation.

The second most supported level of reflection was R1 (i.e., descriptive reflection level). Here, apps provided informational content to users by comparing their goal choices to norms and health recommendations. For example, *Fitbit* highlights how a users' body fat goal choice fits within health recommendations. *Lose it!* compares users' eating plan to those of other members when giving recommendations for goal setting. In the weight management goal setting process, *Home Workout* shows users their current and goal weight and gives options for weight loss through intensity of the timeframe with explanations of the reasonability of goal achievement.



(1) Fitbit *(2) Lose it!* *(3) Home Workout*

Figure 2. Apps that supported reflection level R1 (descriptive) during goal setting.

There was only one instance of R2 reflection during goal setting, in the *Sweatcoin* app (see Figure 3). The app set a step count goal that varied monthly based on the user's history of walking and weather patterns. The app analyzed previous data, created connections based on the user's history and helped put into perspective by providing summaries of the connections between these sources of data and goal suggestions. These analytical representations transcend simple descriptive reflection and begin to suggest more explicit support for dialogic reflection (R2).

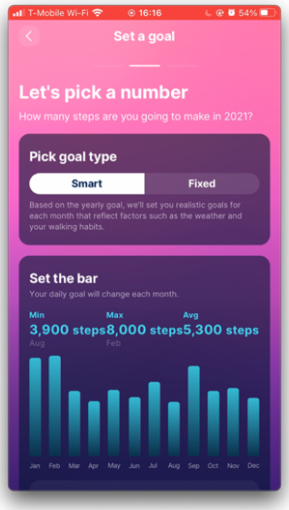


Figure 3. Sweatcoin app that supported reflection level R2 (dialogic reflection) during goal setting.

Only one instance of the level of reflection R3 was found. This was seen in Home Workout (see Figure 4), where when the user was choosing their weight loss goal, they were shown other benefits that achieving this goal would bring to their health (e.g., "Lower blood pressure" or "Reduce the risk of type 2 diabetes"). The user is supported in thinking about their health in a longer term.

Similarly, only one instance of reflection level R4 was found. In *Sweatcoin* (see Figure 5), the user has the option to join races or goals where if enough steps are taken, there will be a donation to a charity. Participating in the goal will make a societal impact.

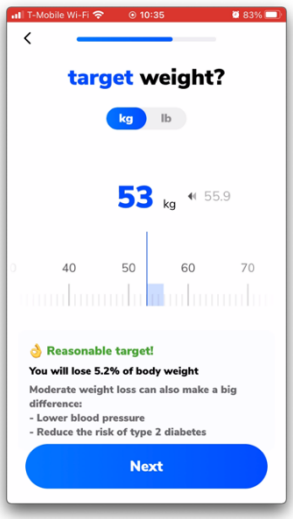
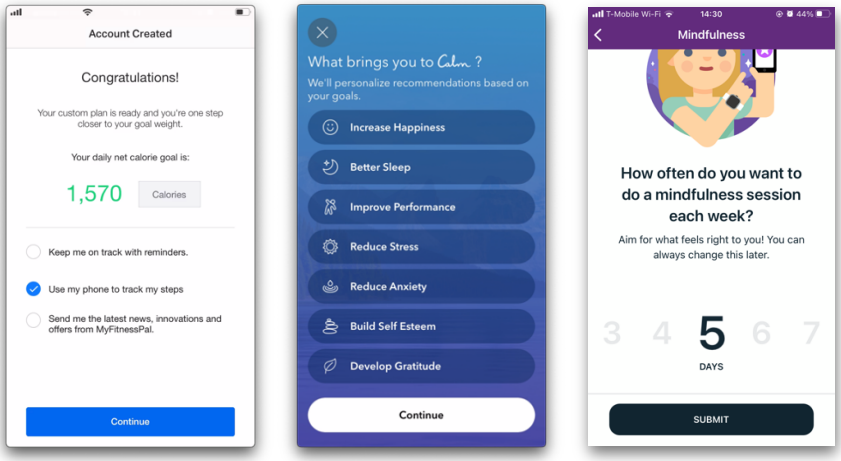


Figure 4. Home Workout app that supported reflection level R3 (transformative reflection) during goal setting.

Figure 5. Sweatcoin app that supported reflection level R4 (critical reflection) during goal setting.

More commonly, we found that the apps did not support any of Fleck & Fitzpatrick's levels of reflection. The apps did not revisit any data, either inputted or collected from the user or from general health information during goal elicitation. For example, in the MyFitnessPal (Figure 6.1), users were automatically assigned to a daily calorie count goal with no accompanying explanation. Calm presented users with a list of goals to choose from, with no accompanying explanation or reflection on the appropriateness of a goal. In a similar way, Fitbit (Figure 6.3) asked users to set the number of days they wished to be mindful, with little or no reflection on what an adequate number would be, nor attempts to understand how mindful the user was or how often they usually felt stress.

We also found some apps (n=3) to lack a connection between reflective questions that were asked to people during the onboarding process and their goals. For example, all of the mindfulness and period tracking apps asked people what their goals were for using the app (e.g., Increase happiness seen in *Calm*, see Figure 6.2) or even ask people how they experienced things like stress. However, after the onboarding, the app would go to the home page without explicitly setting a goal around these factors.



(1) MyFitnessPal *(2) Calm* *(3) Fitbit*

Figure 6. Examples of apps that did not support any reflection level during goal setting.

4. Discussion

In this paper, we analyzed the level of reflection during goal setting in 10 of the current most popular health and fitness personal informatic apps. We did not have a particular focus of type of behavior because we are looking at how the most popular apps integrate reflection during their goal setting process. Regardless of all the different behaviors, there was a lack of reflection during goal setting. We observed that the apps ask questions to help users personalize their goals during onboarding; however, it was unclear how the questions impacted the suggestions of the goals. We believe our findings shed some further light on the reasons why only a minority of PI users engage in goal setting during their use of these tools (Gouveia et al., 2015).

Our findings also show that PI apps often lack proximity (i.e., timeframes), a key aspects of effective goal setting (Locke & Latham, 2002). We found that more than half of the goals have an indefinite timeframe, which leaves the user without a clear future moment to achieve the goal (e.g., 30 days). This contradicts with theory of goal setting which says that effective goal setting should integrate definitive timeframes (Lunenburg, 2011).

In the following section, we discuss our findings to illustrate how goal setting and reflection can be better supported by PI tools' during onboarding processes.

4.1 Black box of goal setting

During the analysis we found that the PI apps often do not explain the reasoning for suggesting a goal. None of the analyzed apps provided explanations for assigned goal suggestions. Assigned goals were often general health recommendations (e.g., drink 64 oz of water per day and 10,000-step goal seen in *Fitbit*) seeming without regard to users current and past abilities. Although most tools collected data about people's abilities, they rarely provided clear reasoning as to how this data was considered when suggesting goals. For example, in the *Lose it!* app, there were over 30 questions asked to collect information from the user (e.g., "What describes your past experiences with weight loss?"). However, when setting a goal, the goal suggestions are not clearly connected to the collected information. Transparency in the setting of a goal improves commitment towards the goal (Woźniak et al., 2020). We believe that PI tools should provide clear and transparent explanations of why information is being collected during the onboarding process and elicitation of goals. Further, PI tools should clearly describe how this data was used to suggest goals.

4.2 A lack of reflective practices for goal setting

Overall, most of the apps did not seem to support reflection during goal setting. The ones that did support reflection often supported lower levels of reflection. On one hand, we believe this is direct result of a lack of knowledge of people's needs and abilities towards a certain behavior. During onboarding, usually very little is known about users. We noticed that most apps tried to better understand users' by asking them specific questions about their needs, habits, and abilities during onboarding. However, these questions were often generic and did not clearly lead to reflection on goals. For example, in *MyFitnessPal* the user is asked to fill in a sentence about how they currently address stress: "I feel better after I ____." After, although the app recognizes that a particular activity supports stress reduction, it does not guide the user in setting a specific goal or use people's current practices during suggestions (i.e., "Next time you feel stressed, you can plan a walk"). Another way to help people reflect without gathered data is to use data of other people who have similar health conditions and goals. For example, Feustel et al. (2018) found that showing users health data of people that have similar health goals, helps trigger reflective activity and form connections on their own behaviors.

Sometimes if the app did ask reflective questions, they were after the goal was set and therefore could not affect the setting of the goal. This could be seen in *Calm* for example, specific reflective questions were asked about stress management but only after the general goal was chosen. Another example can be seen in *Lose it!*, where the user is asked how often they plan to exercise and then after asked how often they currently exercise. If after the reflective questions, another specific goal setting question would be asked (or the questions are asked in reverse), the user could implement the reflection that was inspired from the questions. Li et. al. (2010) commented that personal informatics should support user's reflection with actions. In this case, reflection should be followed by specific goal setting.

4.3 Implications for design

In the following section we give some concrete examples on how each of the different levels of reflection could be implemented by PI tools when supporting people in setting goals.

Description (R0) could be done by asking areas of health that the person wants to focus on and their general habits or health behaviors. This can be done qualitatively (e.g., how do you feel about your eating habits?) or quantitatively (e.g., how many times do you exercise per week). Specific goals could be set related to the past data the user states. For example, if the person states they feel unhappy with their eating habits, the PI could suggest different strategies for improving their diet (e.g., mindful eating, tracking their calories, following a recipe plan) and then setting a goal based on their preferences. General health information can also be stated during the goal setting to help people without collected data.

Descriptive reflection (R1) asks for explanations of R0 reflection, such as asking the users how they feel about an experience (e.g., how did you feel the last time you exercised?). In addition to the guided explanations, open-ended questions asking for input with blank text boxes encourages the user to take time to think about their experience. Both asking about barriers to previous goals and how those challenges were overcome would help users think about past events. PI can also share related data from others with similar experiences.

Dialogic reflection (R2) could be better supported by involving users in making multiple interpretations of their data and from different sources. For example, after asking for a description of past events (R1) (e.g., *When was the last time you felt very stressed?*), people could be asked what caused the event and if there were other ways they could (or did) overcome the challenge. After, the PI can ask users to write down different ways they will address their goal of reducing stress next time they are anxious. If previous data was collected (e.g., step count), then PI can support users by looking for patterns in the data (e.g., *You walk more on weekends*) to help them think about relationships between their health and habits. After reflection, the PI can help the user set a goal incorporating these insights, such as adding an evening walk after work to improve weekday step count.

Transformative reflection (R3) helps users make fundamental changes to their behavior. Building from R2, the PI could support people in changing habits through patterns in their data. People can state their overall health goals and smaller goals (e.g., *I want to get better at rock climbing*) and the PI tool can recommend activities that the person has not thought of yet (e.g., *Incorporating other sports such as swimming will help improve your strength*). The PI can also comment on other health benefits the user would have by changing their behavior (e.g., *Eating less sugar will help reducing chances of heart problems*).

Critical reflection (R4) looks at social and ethical impacts of goal setting. People's goals are often affected by those who surround them. Asking people how others affect their goals and how they can affect other people with their goals, would support deeper level reflection. For example, asking how they eat with their friends and then asking if their friends would be

willing to cook a healthy meal together instead of going out to eat. Goals can also have ethical impacts, such as eating less meat reduces greenhouse gas emissions or walking to work is a more sustainable form of travel. PI can inform users about how their choices can impact the environment and onto their friends and family, then setting a goal related to improving these choices.

We understand that asking questions to the users in onboarding is an important part of reflection and goal setting, however the number of questions that posed to user should also consider the amount of time that user can spend during onboarding. The onboarding goal setting procedure should not be overly burdensome (Kapusy & Lógó, 2020) to users but a positive and empowering experience for their first step in making a health behavior change.

5. Conclusions

This study analyzed how popular PI apps implement reflection in goal setting during onboarding by looking at theory in goal setting and levels of reflection. We found reflection to be missing in a majority of the goals set in first usage of the 10 most downloaded PI apps. There were reflective questions that were asked to users during onboarding, which were not clearly used in goal setting. A clear link between reflection and setting of a goal in PI apps would contribute to supporting reflective goal setting.

Our study has two limitations. The size of the sample (the number of apps was limited to 10) and the analysis of PI apps on a variety of health behaviors. Future studies could analyze more apps and focus on specific health behaviors when selecting the apps. Our focus was not related to a particular health behavior but on overall how PI apps address reflection during onboarding goal setting. We invite the design research and designers to follow up on our findings and implications presented in this paper.

Acknowledgements: This study was supported Pride and Prejudice project by the 4TU federation (www.4tu.nl) under Grant No. 4TU-UIT-346. We also thank Catrin Feron for her help in reviewing the initial list of apps.

6. References

Abishkking Limited. (2014). *Period Tracker Period Calendar* (2.32.0) [Computer software]. https://apps.apple.com/us/app/period-tracker-period-calendar/id896501514

Abishkking Limited. (2017). *Home Workout—No Equipment* (1.10.0) [Computer software]. https://apps.apple.com/us/app/home-workout-no-equipments/id1313192037

Abishkking Limited. (2018). *Six Pack in 30 Days—6 Pack* (1.7.0) [Computer software]. https://apps.apple.com/us/app/six-pack-in-30-days-6-pack/id1338655056

Apptopia. (2021, November 1). *ITunes Connect Top Apps Analysis for Health & Fitness in The United States*. Free App Store Rank Data for ITunes Connect Monday, November 1, 2021. https://apptopia.com/store-insights/top-charts/itunes-connect/health-fitness/united-states

Avishai, A., Conner, M., & Sheeran, P. (2019). Setting Realistic Health Goals: Antecedents and Consequences. *Annals of Behavioral Medicine*, *53*(12), 1020–1031. https://doi.org/10.1093/abm/kaz012

Baretta, D., Bondaronek, P., Direito, A., & Steca, P. (2019). Implementation of the goal-setting components in popular physical activity apps: Review and content analysis. *DIGITAL HEALTH*, *5*, 205520761986270. https://doi.org/10.1177/2055207619862706

Baumer, E. P. S., Khovanskaya, V., Matthews, M., Reynolds, L., Schwanda Sosik, V., & Gay, G. (2014). Reviewing reflection: On the use of reflection in interactive system design. *Proceedings of the 2014 Conference on Designing Interactive Systems*, 93–102. https://doi.org/10.1145/2598510.2598598

Calm.com. (2012). *Calm: Sleep & Meditation* (5.30.1) [Computer software]. https://apps.apple.com/us/app/calm-sleep-meditation/id571800810

Ceci, L. (2021, October 1). *Leading Android health apps in the U.S. 2021, by downloads* [Statisical]. Statistica. https://www.statista.com/statistics/699054/leading-google-play-health-usa-downloads/

Conroy, D. E., Yang, C.-H., & Maher, J. P. (2014). Behavior Change Techniques in Top-Ranked Mobile Apps for Physical Activity. *American Journal of Preventive Medicine*, *46*(6), 649–652. https://doi.org/10.1016/j.amepre.2014.01.010

Consolvo, S. (2012). Designing for Healthy Lifestyles: Design Considerations for Mobile Technologies to Encourage Consumer Health and Wellness. *Foundations and Trends® in Human–Computer Interaction*, *6*(3–4), 167–315. https://doi.org/10.1561/1100000040

Epstein, D. A., Caldeira, C., Figueiredo, M. C., Lu, X., Silva, L. M., Williams, L., Lee, J. H., Li, Q., Ahuja, S., Chen, Q., Dowlatyari, P., Hilby, C., Sultana, S., Eikey, E. V., & Chen, Y. (2020). Mapping and Taking Stock of the Personal Informatics Literature. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(4), 1–38. https://doi.org/10.1145/3432231

Feustel, C., Aggarwal, S., Lee, B., & Wilcox, L. (2018). People Like Me: Designing for Reflection on Aggregate Cohort Data in Personal Informatics Systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(3), 1–21. https://doi.org/10.1145/3264917

Fitbit, Inc. (2014). *Fitbit: Health & Fitness* (3.51) [Computer software]. https://apps.apple.com/us/app/fitbit-health-fitness/id462638897

FitNow. (2008). *Lose it! Calorie Counter* (13.6.302) [Computer software]. https://apps.apple.com/us/app/lose-it-calorie-counter/id297368629

Fleck, R., & Fitzpatrick, G. (2010). Reflecting on reflection: Framing a design landscape. *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction - OZCHI '10*, 216. https://doi.org/10.1145/1952222.1952269

Flo Health, Inc. (2015). *Flo Period & Ovulation Tracker* (8.1) [Computer software]. https://apps.apple.com/us/app/flo-period-ovulation-tracker/id1038369065

Gouveia, R., Karapanos, E., & Hassenzahl, M. (2015). How do we engage with activity trackers?: A longitudinal study of Habito. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, 1305–1316. https://doi.org/10.1145/2750858.2804290

Headspace Inc. (2012). *Headspace: Meditation & Sleep* (3.192.0) [Computer software]. https://apps.apple.com/us/app/headspace-meditation-sleep/id493145008

Kapusy, K., & Lógó, E. (2020). User Experience Evaluation Methodology in the Onboarding Process: Snapchat Case Study. *Ergonomics in Design: The Quarterly of Human Factors Applications*, 106480462096227. https://doi.org/10.1177/1064804620962270

Karahanoğlu, A., Gouveia, R., Reenalda, J., & Ludden, G. (2021). How Are Sports-Trackers Used by Runners? Running-Related Data, Personal Goals, and Self-Tracking in Running. *Sensors*, *21*(11), 3687. https://doi.org/10.3390/s21113687

Kocielnik, R., Xiao, L., Avrahami, D., & Hsieh, G. (2018). Reflection Companion: A Conversational System for Engaging Users in Reflection on Physical Activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(2), 1–26. https://doi.org/10.1145/3214273

Latham, G. P., Winters, D. C., & Locke, E. A. (1994). Cognitive and motivational effects of participation: A mediator study. *Journal of Organizational Behavior*, *15*(1), 49–63. https://doi.org/10.1002/job.4030150106

Lee, M. K., Kim, J., Forlizzi, J., & Kiesler, S. (2015). Personalization revisited: A reflective approach helps people better personalize health services and motivates them to increase physical activity. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, 743–754. https://doi.org/10.1145/2750858.2807552

Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*, 557. https://doi.org/10.1145/1753326.1753409

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Locke, E. A., & Latham, G. P. (2019). The development of goal setting theory: A half century retrospective. *Motivation Science*, *5*(2), 93–105. https://doi.org/10.1037/mot0000127

Lunenburg, F. C. (2011). Goal-Setting Theory of Motivation. *International Journal of Management, Business, and Administration*, *15*(1), 6.

Middelweerd, A., Mollee, J. S., van der Wal, C. N., Brug, J., & te Velde, S. J. (2014). Apps to promote physical activity among adults: A review and content analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 97. https://doi.org/10.1186/s12966-014-0097-9

Munson, S. A., Schroeder, J., Karkar, R., Kientz, J. A., Chung, C.-F., & Fogarty, J. (2020). The Importance of Starting With Goals in N-of-1 Studies. *Frontiers in Digital Health*, *2*, 3. https://doi.org/10.3389/fdgth.2020.00003

MyFitnessPal, Inc. (2005). *MyFitnessPal* (21.23.5) [Computer software]. https://apps.apple.com/us/app/myfitnesspal/id341232718

Peng, W., Li, L., Kononova, A., Cotten, S., Kamp, K., & Bowen, M. (2021). Habit Formation in Wearable Activity Tracker Use Among Older Adults: Qualitative Study. *JMIR MHealth and UHealth*, *9*(1), e22488. https://doi.org/10.2196/22488

Powell, A. C., Torous, J., Chan, S., Raynor, G. S., Shwarts, E., Shanahan, M., & Landman, A. B. (2016). Interrater Reliability of mHealth App Rating Measures: Analysis of Top Depression and Smoking Cessation Apps. *JMIR MHealth and UHealth*, *4*(1), e15. https://doi.org/10.2196/mhealth.5176

Saksono, H., Castaneda-Sceppa, C., Hoffman, J. A., Seif El-Nasr, M., & Parker, A. (2021). StoryMap: Using Social Modeling and Self-Modeling to Support Physical Activity Among Families of Low-SES Backgrounds. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3411764.3445087

Shilts, M. K., Horowitz, M., & Townsend, M. S. (2004). Goal Setting as a Strategy for Dietary and Physical Activity Behavior Change: A Review of the Literature. *American Journal of Health Promotion*, *19*(2), 81–93. https://doi.org/10.4278/0890-1171-19.2.81

Sweatco Ltd. (2016). *Sweatcoin Walking Step Counter* (105.0) [Computer software]. https://apps.apple.com/us/app/sweatcoin-walking-step-counter/id971023427

Woźniak, P. W., Kucharski, P. P., de Graaf, M. M. A., & Niess, J. (2020). Exploring Understandable Algorithms to Suggest Fitness Tracker Goals that Foster Commitment. *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 1–12. https://doi.org/10.1145/3419249.3420131

About the Authors:

Tina Ekhtiar is a PhD candidate at the Faculty of Engineering Technology, in Interaction Design Research Group in the University of Twente. In her research, she focuses on how personal informatics can better support goal setting.

Rúben Gouveia is an Assistant Professor of Interaction Design at the University of Twente, the Netherlands. His research focuses on designing and developing tools that support health and wellbeing and evaluating how these tools augment abilities to reflect and to identify opportunities for change.

Armağan Karahanoğlu is an Assistant Professor of Interaction Design at the University of Twente, the Netherlands. Her research focuses on understanding and designing human-technology relations through creative and alternative research methodologies, especially in sports, health and well-being.

Geke Ludden is a Professor of Interaction Design at the University of Twente, the Netherlands. She studies how the design of products and services influences people's behavior and motivation with a specific interest in supporting healthy behavior and engaging people in therapy at home.