Scaffolding for an OLM for Long-Term Physical Activity Goals

Lie Ming Tang
University Of Sydney
Darlington, NSW, Australia
ltan8012@uni.sydney.edu.au

Judy Kay
University Of Sydney
Darlington, NSW, Australia
judy.kay@sydney.edu.au

ABSTRACT

An important role of open learner models (OLMs) is to support *self-reflection*. We explore how to do this for an OLM based on fine-grained long term physical activity tracker data that many people are accumulating. We aim to tackle two well-documented challenges that people face, in making effective use of an OLM for reflection. 1. We created a tutorial to scaffold *sense-making* needed to understand the meaning of the OLM. 2. We integrated an interface scaffold to help users *consider key questions* for effective reflection. We report the results of a qualitative think-aloud lab study with 21 participants viewing their own long term OLM. To evaluate the *tutorial scaffolding*, we split participants into an experimental group, who did a tutorial before exploring the OLM and a control group which explored the interface without the tutorial. To evaluate the *reflection scaffolding*, all participants first explored the interface as they wished. We then provided *goal prompts* to scaffold reflection. Our study revealed that, under lab conditions, the tutorial scaffolding was not needed – all participants in both groups could readily understand the OLM. However, we found that several of the goal prompts were important to help participants consider key questions for effective reflection. Our key contribution is insights into the *design of scaffolding for reflection in a life-long learning context* of gaining insights and setting goals for physical activity.

CCS CONCEPTS

• **Human-centered computing → User studies**; **Walkthrough evaluations**; **Interaction design**; *User models*;

KEYWORDS

OLM, Scaffolding, User Interface, Long Term Physical Activity Data

ACM Reference Format:
Lie Ming Tang and Judy Kay. 2018. Scaffolding for an OLM for Long-Term Physical Activity Goals. In *UMAP '18: 26th Conference on User Modeling, Adaptation and Personalization, July 8–11, 2018, Singapore, Singapore.* ACM, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.1145/3209219.3209220

UMAP '18, July 8–11, 2018, Singapore, Singapore

ACM ISBN 978-1-4503-5589-6/18/07…$15.00
https://doi.org/10.1145/3209219.3209220

1 INTRODUCTION

There is a growing body of work that aims to create Open Learner Models (OLMs) to support meta-cognition [10, 17, 28, 38, 51]. Much OLM research has been in the context of formal education. However, this important role of user models is also important in *life-long* and *life-wide* learning. In particular, one key role of an OLM is to support *self-reflection* [5, 11, 13, 16, 17]. This is especially important for achieving very long term goals, such as achieving and maintaining healthy levels of physical activity [6, 32, 39]. OLMs can play several important roles, including support users' curiosity about their data, allowing for playfulness in tracking style [44], learner trust [1] and several broader important meta-cognitive activities [12, 13, 18, 25].

Our work aims to take insights from OLM research into the area of *personal informatics*, where emerging sensor technologies enable people to collect considerable personal data. We focus on the goal to harness data from worn sensors about physical activity. Such sensors are becoming ubiquitous with the emergence of dedicated devices such as the Fitbit [1] as well as through ambient tracking via non-dedicated devices such as smartphones [2] [3]. Our work takes an OLM perspective, first to transform long term physical activity data into a user model, then to create an OLM interface, called *iStuckWithIt*, to support self-reflection on the user model.

In this paper, we consider two research questions about the scaffolding for self-reflection using *iStuckWithIt*:

- Do users need a scaffolding introductory tutorial to self-reflect using *iStuckWithIt*?
- Do users benefit from a reflection scaffold to systematically self-reflect on core long term goals represented in the OLM?

To explore this, we conducted a lab study where 21 existing long term physical activity trackers were asked to use *iStuckWithIt* [53], with 2 additional scaffolding elements: *tutorial* introduction and *goal prompts* for reflection. The tutorial scaffolding asks users to review the data of 2 hypothetical users, with data that highlights critical features of the dashboard. The goal prompt scaffolding is a side panel (pop-up) that asks users to answer 5 questions about their goals and their behaviour including whether they are achieving their goal, whether they should change their goal and to consider differences between when they are at work and not at work, weekend and weekdays or on holidays. These questions prompt users to considering their goal setting as well as how environmental and temporal factors that are known to affect physical activity, as documented in health literature [14].

The next section reviews related work followed by the study design and results. We conclude with a discussion of the findings and lessons learned for future designs.

[1] fitbit.com
[2] https://www.apple.com/au/ios/health/
[3] https://support.apple.com/en-au/HT203037

2 BACKGROUND

This section positions our work in relation to three bodies of previous research. First, we build on OLM research, where a user model is made available to the user to support goals such as self-reflection. Then we introduce the largely independent work on personal informatics, including an overview of the design of the basic *iStuckWithIt* interface. The third key strand is on meta-cognitive scaffolding for OLMs. We then introduce the main *iStuckWithIt* interface and explain the goals of our work in terms of the new contributions we aim to achieve.

Open Learner Modelling has a long history, beginning with the recognition that a user model (also called student or learner model) could be made available to the user [9, 31]. An OLM could serve several roles, including the learner interacting to negotiate or argue about the user model [9], supporting user control over their personal data [31] and for metacognitive processes of self-reflection, self-monitoring and planning [11, 13, 17]. There has been considerable research on the ways to present learner models, comparing various forms [13, 29]. There is also a body of studies that have demonstrated their effectiveness for learning in formal educational contexts [17, 38, 41, 42].

While OLM research has been largely concerned with formal educational settings, emerging sensor and mobile technologies have led to Personal Informatics research [34] and the similar Quantified Self movement in the broader community[4]. These communities also aim to create useful representations of users, available in interfaces that have similar goals to OLMs. This community has demonstrated that, while people see the potential value of such data for self-reflection, current tools fail to support this well [15, 27, 35, 45, 47, 53]. Indeed, there is a growing body of evidence that points to a lack of perceived usefulness of long term tracker data [26, 45, 46]. In personal informatics, the user models need to be designed to represent aspects of user's goals, linking the available sensor data to those goals. A key problem in creating the model, and associated OLM interfaces, relates to problems in the accuracy of the data due to incompleteness. For example, a worn activity tracker only gives reliable data when the user wears it and this should be considered in reasoning about the user's activity level and modelling their goal achievement. Incomplete data compromises the usefulness of tracking. People can lose confidence when they are confronted with gaps or incorrect reports due to gaps [8, 45]. Failure to account for, or recognise, incomplete data can mean that people consider the data not to be useful which has been reported in recent years [19, 26, 30, 33, 45, 48, 54]. A similar problem has been identified for OLMs for formal learning, with the need to represent the uncertainty in the model [2, 21].

While the ideal OLM interface would be readily understood by the user, in practice this may be difficult to achieve. Even with a quite simple skillometer, consisting of just seven bars [37], there were challenges in both understanding the model and the meaning of the components display as well as in reflection. Self-evaluation is especially important for achieving the very long term goals, such as achieving and maintaining healthy levels of physical activity [6, 32, 39]. OLMs can play several important roles, including supporting users' curiosity about their data, allowing for playfulness in tracking style [44], facilitating learner trust [1] and several other important meta-cognitive activities [12, 13, 18, 25].

[4]http://quantifiedself.com/

This work explores the role of scaffolding for the *iStuckWithIt* interface. The design of this interface and the nature of insights people made when using it have been previously reported [53]. We now briefly introduce that version of the interface, as shown at the left of Figure 1. Broadly, the design is based on a calendar visualisation. The labels A-H illustrate key features. A marks the drop-down menu to select the class of goal the user wants to see; those in the study are steps per day, count of active minutes per day and distance walked per day. The main interface is marked B for a period when the user has data about their activity levels (with all cells either white or shades of blue) and C when there was a period with no data because the user did not wear their tracker in that period (grey cells). The figure shows the interface configured for a goal target of 10,000 steps a day and only cells exceeding this are bright blue. The configuration in the figure sets a 50%, or 5,000 step threshold for the lighter blue and then white indicates days that have data (>= 1 step) but less than 5,000 steps. The bars marked with D were designed to help the user take account of the impact of their actual wearing behaviour on the results shows. The bars show the average number of hours per day the user wore the tracker in each week. When this is low, as in the case of weeks nearest the D, the results are based on just the limited data that is available. In the figure, the user has clicked on the cell near E to see more details for that day. The upper middle is the configuration section, labeled F. This enables the user to change the thresholds for the goals. The right part of the figure, G, is the reflection scaffolding that is the focus of this paper. H enables the user to alter the display from goal oriented, as in this figure, to a gradient.

In summary, there has been considerable research in OLMs, especially in formal educational settings. There is a growing body of work in personal informatics and broader community interest in Quantified Self. Both have identified a key challenge for effective use of personal data – although they have not used the term, OLM, they highlight the need for scaffolding to help people make sense of complex collections of personal data, user models, so as to support their self-reflection. Our work tackles this problem. This paper goes beyond our previous report of *iStuckWithIt*'s design [53] as we now describe the study of the two forms of scaffolding we explored for the *iStuckWithIt* OLM interface: the tutorial scaffolding to introduce the interface and the reflection scaffolding.

3 STUDY DESIGN

Our two research questions were:

- RQ1: Tutorial scaffold: Do users need a scaffolding introductory tutorial to self-reflect using *iStuckWithIt*?
- RQ2: Goal prompts scaffold: Do users benefit from a reflection scaffold to systematically self-reflect on core long term goals represented in the OLM?

This section first describes the overall design of the study and then the detailed design for each research question.

We recruited 21 long term Fitbit users, people who had collected at least 6 months of personal physical activity data. We then conducted a between-subjects lab study in terms of the first research question. This study session had the following stages

(1) Nine participants (9) worked through the scaffolding tutorial, described below.
(2) All participants were asked to explore the main interface in *iStuckWithIt*. We asked them to *think-aloud* [43], explaining what they saw, understood and their insights.
(3) We then asked all participants to consider the reflective questions, labeled (G) in Figure 1.
(4) Finally, we interviewed them on their experiences of viewing their own data and what insights they learned.

In the next two sub-sections, we present the motivation and design of the tutorial and goal prompt scaffolding. We also explain the study design to evaluate each of them.

3.1 RQ1: Tutorial Scaffolding

While some long term physical activity trackers do use their tracker over the long term, most fail to make use of their own long term data [26, 45, 53]. This means that our study design should account for the likelihood that the *iStuckWithIt* interface would provide the first opportunity for participants to see their own *long term* goal performance for physical activity, in terms of steps, active minutes and distance. We anticipated that participants would benefit from a tutorial that introduced them to *iStuckWithIt*.

To evaluate this, we prepared a tutorial, based on a set of exercises to explore the *iStuckWithIt* OLM for two hypothetical users, Alice and Alex. These provided carefully designed datasets which highlighted key aspects that the interface was intended to enable people to understand about their own long term physical activity model. We asked 9 of our participants to complete this prior to seeing *iStuckWithIt* with their own data. The steps in this tutorial were:

(1) Participants were told that Alex and Alice's each had a goal of "at least 30 very active minutes per day".
(2) Participants were asked to consider whether Alice achieved her goal or not.
(3) They then did this for Alex.
(4) In each case, the experimenter allowed the participant to explore the OLM, thinking aloud to explain how they interpreted it.
(5) At the end, if participants had failed to see and understand key features, the experimenter explained them.

Figure 2 shows the data for Alice who started using her Fitbit in August 2015. Key features are:

(1) The first month of tracking had quite high tracker use - few grey cells;
(2) In mid-September, there is a 2-week gap in tracking - all grey cells;
(3) After this, there are many grey cells, reflecting days Alice did not wear her tracker, especially on weekends.
(4) Consistently higher wear time in August and September and then consistently lower hourly wear-time after September.
(5) Scaffold users to reflect or speculate on the *potential causes* for this change from Aug/Sept to afterwards.

Figure 1 shows Alex's data which highlights the following:

(1) Low physical activity levels during weekend compared with weekdays.
(2) Large gaps of several months (blocks of grey cells) between wearing activity tracker.
(3) Overall inconsistent hourly wear-time and some periods with low wear-time.

We recorded observations and participant comments in their think-aloud exploration of the interface. If, after some exploration, the participant did not notice key aspects, they were prompted about them. We also recorded whether the user commented on how missing data could have affected accuracy of the step counts as well as comments around wearing behaviour of Alice and Alex.

The aim of the tutorial scaffolding study is two-fold. First, we wanted to discover which features of the long term physical activity tracker data are easily understood and which are not. We also wanted to see the impact of the tutorial, to learn whether participants who had the extra learning scaffolding were better able to make sense of their own data than those who did not do the tutorial.

3.2 RQ2: Goal prompts scaffolding

After participants had finished exploring the main *iStuckWithIt* interface with their own data, we asked them to open the goal prompt scaffolding, labelled G at the right of Figure 1. This was designed around two core forms of reflection:

(1) Reflect on *goal achievement* and consider reviewing the *goal setting* – the first two questions, marked +.
(2) Reflect on *factors affecting goal achievement* – the last 3 prompts, also marked with +, about weekend (versus weekdays), holidays and work versus non-work.

The benefits of scaffolding or support for self-regulation skills such as self-monitoring, goals and goal setting are well documented [3, 4]. Also, previous work [49, 50] using goal setting as a strategy to promote health and physical activity behaviour change has demonstrated the potential of this approach. The first category of our questions aims to remind users to reflect on their goals and goal setting. The second category of questions called for the participant to consider factors that are known to be important for people's physical activity levels. Health studies have consistently indicated differences between activity levels on weekends versus weekdays [7, 24, 52]. Moreover, previous studies have shown that by helping users consider such questions can support reflection on activity tracker data [22]. Therefore, our scaffolding design was based on literature indicating the benefits of reminding users to consider the context of their activity levels and behaviour is likely to be useful.

In addition to our prompt questions, we did consider others. For example, studies of existing users of physical activity trackers have reported that such users are quite interested in peaks and lows [26, 36]. On the matter of influencing people to change tracking behaviour, Epstein et al. [23] found potential value in using visualisations for encouraging users to return after a long gap in tracker use. While these are potentially useful, we did not include them as our focus was of studying whether people could make sense of their data at the OLM interface and reflect on important features associated with learning about their own long-term goal achievement.

Figure 1: OLM to support reflection on achievement of long term physical activity goals. (A): Drop down menu to select between datasets (i.e., steps, active minutes, distance). (B): Calendar visualisation, with colour showing activity level on each day – dark blue means 10,000+ steps, light blue >5,000 but <10,000 steps, white has >0 but <5,000 steps. (C): Grey striped cells are days with no data. (D): Bar graph showing average wear-time in hours per day for each week. (E): Pop-up showing additional details of a particular day / cell. (F): Configuration – to adjust the thresholds for colouring of the cells. (G): Goal prompt scaffolding questions.

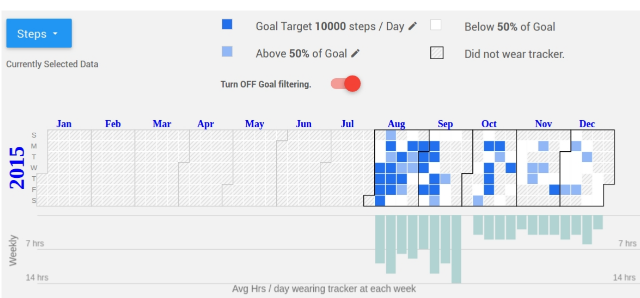

Figure 2: Data for hypothetical user Alice where she started using data in August until Dec. Her daily wear-time declined after September especially on weekends and wear-time is also lower but consistent.

4 RESULTS

In this section, we first introduce the participants of our study then we discuss the results around the two research questions.

4.1 Participants

Table 1 presents details of our participants, ordered by gender in Col. 2, then by scaffolding condition Col. 3. This shows 9 participants did tutorial (Y) and 12 did not (N). In terms of background, many participants were highly educated and several worked or studied in IT, shown in Col. 5 and Col. 6. This group may have higher literacy and skills levels in data analysis than a more general population. This is similar to participants in other qualitative studies of existing long term personal trackers [26]. There were more male participants (14) than female (7). Our participants ages are spread across the age groups shown in Col. 4 with 25-34 being the largest group at 6, the lowest 18-24 (3) and 4 participants in the others. Our demographic is similar, in terms of age and gender, to the population of personal activity tracker users and wearable technology adopters [20]. The duration of tracker use varied from 6 to 38 months shown in Col. 7 (average: 23 months).

Col. 8 shows the %-age of days with at least 1 step. This varied widely (min: 15%, max: 100%, average: 68%, std: 30%). Col. 9 is the wear-time (number of hours per day users with >= 1 step recorded). Our participants generally had high wear-time (min: 9 hours, max: 20, average: 15, std: 3) and Col. 10 shows the standard deviation (min: 2, max: 6, average: 4, std: 1). Col. 9 and Col. 10 show that while overall wear-time is high there is large variation both between and within individuals.

Overall, our participants had very wide differences in consistency of days with tracking – the %-age of days with any data. For example, 6 participants (P8, P9, P10, P14, P15 and P20) had 100% of

Table 1: Participant profiles grouped by gender. Col. 1 participant identifier. Col. 2 gender. Col. 3 whether participant is in the tutorial scaffolding condition. Col. 4 to Col. 6 participant age, occupation and education. Col. 7 duration in months of tracking data (first to last day with data). Col. 8 %-age of days with at least 1 step Col. 7). Col. 9 and Col. 10 the median and standard deviation of wear-time (number of hours per day with at least 1 step). The last 4 rows are summary statistics over all participants (average, standard deviation, min and max) of Col. 8, Col. 9 and Col. 10. N=21

1	2	3	4	5	6	7	8	9	10
P#	M/F	Tut	Age	Occupation	Education	Dur	Daily	Hours (med)	Hours (std)
P1	F	N	25 - 34	Part-time	U-Grad	7	86%	13	6
P2	F	N	25 - 34	Property Admin	U-Grad	9	96%	17	5
P3	F	N	45 - 54	Manager	P-Grad	38	39%	14	3
P4	F	N	35 - 44	IT Support	P-Grad	8	88%	16	4
P5	F	Y	55 - 64	Academic	Prof	15	15%	14	6
P6	F	Y	25 - 34	IT Developer	U-Grad	27	62%	9	4
P7	F	Y	45 - 54	Academic	P-Grad	37	32%	13	4
P8	M	N	55 - 64	IT Developer	U-Grad	6	100%	16	3
P9	M	N	55 - 64	Retired Military	U-Grad	13	100%	17	5
P10	M	N	45 - 54	Dir of IT	P-Grad	21	100%	18	2
P11	M	N	35 - 44	Engineer	Prof	23	60%	16	4
P12	M	N	18 - 24	Student	U-Grad	27	31%	13	4
P13	M	N	25 - 34	Researcher	P-Grad	29	70%	13	3
P14	M	N	35 - 44	Self Employed	U-Grad	36	100%	20	4
P15	M	N	45 - 54	Manager	U-Grad	38	100%	17	2
P16	M	Y	25 - 34	Student	P-Grad	18	29%	N/A	N/A
P17	M	Y	25 - 34	Student	P-Grad	18	38%	11	5
P18	M	Y	18 - 24	Student	Highschool	22	44%	13	3
P19	M	Y	55 - 64	Professor	P-Grad	26	41%	12	4
P20	M	Y	18 - 24	Student	P-Grad	27	100%	18	3
P21	M	Y	35 - 44	Professor	P-Grad	29	91%	13	4
Summary Stats						Avg	68%	15	4
						Std	30%	3	1
						Min	15%	9	2
						Max	100%	20	6

days with data - meaning they wore their tracker every day in the period from the first day to the last in the dataset. These participants also averaged higher wear-time within each day, recording between 16 and 20 hours per day.

In contrast, the 6 participants with lowest consistency in wearing the tracker had between 15% and 39% of days that had any data (P6, P10, P2, P17, P20 and P18) and their median wear-time was 9 to 16 hours per day. Since in medical research 10 hours of wear-time is considered sufficient for meaningful data [40], even those who averaged 9 hours (P6) may have acceptable data quality. Interestingly, these 2 groups of users had similar wear-time variation indicated by the standard deviation (i.e., variation in the number of hours per day with at least 1 step) between averaging 4 hours for lower daily adherence users and 3 for highly daily adherent users.

These participant statistics indicate that while there are large variations in the *number of days with data* (%-age of days with at least 1 step) all participants had high wear-time.

Table 2: Table showing whether users identified each of the notable items as part of their tutorial scaffolding. Row. 1 participant ID. Row. 2 %-age of days with at least 1 step. Row. 3 median hours per day with >1 step. Row. 4 to Row. 6 3 notable items in the Alice tutorial. Row. 7 to Row. 9 notable items in the Alex tutorial Row. 10 notable items identified by each participant as %-age of all 6 such items. N=9.

1		Participant ID	P20	P21	P6	P18	P19	P17	P7	P16	P5	
2		Daily Adherence %	100	91	62	44	41	38	32	29	15	
3		Wear-time (median hours / day)	18	13	9	13	12	11	13	N/A	14	
4	Alice	Lower daily adherence after September, during 2 weeks gap during September to October especially on weekends.	x	x	x	x	x	x	x	x	x	100%
5		Lower wear-time after September.	x	x	x		x			x		56%
6		Reflect or speculate on the causes for this change from Aug/Sept to afterwards.	x	x	x	x	x	x	x	x	x	100%
7	Alex	Low physical activity levels during weekend versus weekdays.	x		x						x	33%
8		Large gaps between wearing activity tracker.	x	x	x	x	x	x	x	x	x	100%
9		Inconsistent and lower wear-time.		x						x		22%
10		Items identified (%)	83%	83%	83%	50%	67%	50%	50%	83%	67%	

4.2 RQ1: Tutorial Scaffolding

Table 2 shows which notable item each participant identified for hypothetical users, Alice from Row. 4 to Row. 6 and Alex, from Row. 7 to Row. 9 during the tutorial scaffolding condition. All participants identified at least 50% of these but none could identify all items.

Our participants readily identified whole day wearing patterns, distinguishing days with any data (appearing as white or light or dark blue cells) against days with no data (grey cells). In the Alice condition, all participants observed that she had more missing days after September (Row. 4) and all commented or speculated on the causes for this (Row. 6). For example, P20 commented on Alice's gap in data during September, *"Stopped wearing in September, October returned, maybe had gone on holiday"*. In the Alex condition, all participants commented on the large gaps between periods of more consistent tracking over the 2 year period (Row. 8). They also commented on Alex being more consistent in wearing his tracker on more days in the second year.

The notion of the wear-time in terms of *number of hours with any data in a day*, was harder for our participants to discover. Row. 5 shows that 4 participants (44%) did not comment on the low wear-time for Alice after September. Moreover, only 2 participants (22%) noticed the drop in Alice's wear-time (Row. 5) as well as the inconsistent wear-time across different months and years in Alex's data (Row. 9). P16 commented when viewing Alice's data, *"I see she is not always using her tracker, especially in the last few months. Only 4 - 5 hours per day"* and when viewing Alex's data *"He used to wear the tracker more longer than recently"*. P21 when viewing Alice's data, *"Towards the end, not only was [Alice] less active but also wearing it for shorter periods"*. When viewing Alex's data, he commented, *"Alex is also more consistently wearing the tracker in the second year."*.

Interestingly, most participants (3 of 4) who commented or reflected on wear-time also commented on their own wear-time when viewing their own data during the think-aloud. (The outlier is P16 the sole participant who did not have wear-time data).

Notably, several participants in the *control condition* – who did *not* do the tutorial scaffolding, also commented on wear-time in their own data. For example, P12 commented on how the hours of wear (wear-time) may affect his activity levels, *"If I wear my tracker longer, it will track more of my activity. I take off my tracker when I get home. So maybe I should just leave my tracker on myself at home as well to get more steps, little steps I do walking around home"*. P4 was surprised at the number of hours she wore her tracker, *"I was surprised I was wearing it for 19 hours"*. When investigating a day that had very low steps as well as hours of wear, P5 commented, *"I thought I wore it for longer that day, it could be that I bumped it and turned it off. I think it happens sometimes when I bump it"*.

Row. 7 indicates that only 3 participants (33%) noticed the substantially lower steps and lower days with any data on weekends. One of the three who did, P7 reflected on the differences between Alice and Alex, commenting, *"Alex didn't wear it on weekends, which I didn't notice on Alice"*.

Row. 2 and Row. 3 of Table 2 shows each participant's own daily %-age of days with data (Daily Adherence) and wear-time against the number of notable items in each tutorial scenario. This shows the completeness of their own data, so we can see this against the notable items they identified. For example, P6 had the lowest median wear-time (9 hours per day) and P5 had the lowest daily adherence (with just 15% of days with data). During the scaffolding tutorial, we found several cases where users related their own experiences and thinking when viewing these 2 hypothetical user's data. For example, P5 explained their low, 15% of days with data, was due to significant medical conditions, and commented *"Sometimes she doesn't wear her tracker, so she's like me"*. P17 who had 38% daily adherence commented, *"I can see Alex is achieving his goal. But as time passes, he started to drop in his wear. In the beginning, I think he is more motivated, like myself as well."*. When looking at Alice's data, P17 commented, *"I think people forget to wear or perhaps it doesn't match the fashion"*. Interestingly, when he explored his own long term data, after the tutorial, he explained that he stopped wearing his Fitbit tracker because his did a sport that required a wrist based apparel that prevented him from wearing his Fitbit. His earlier speculation on the reason for Alice's missing days could be due to fashion concerns, is similar to this reasoning. P21 commented, *"I can see there are some days where Alex forgot to wear it or forgot to track properly – maybe she was on holiday or something"*. When analysing his own long term tracker data, P21 who has relatively high daily adherence (91%) commented on the days he did not wear the tracker due to holidays and work travels. P18 who has daily adherence of 44% also commented on the motivational effect of adherence and missing data, *"She was very consistent but she stopped. I think it's discouraging to have so much missing data."*.

This section summarised results for the first research question which explored whether the tutorial scaffold was needed and useful. Comparing the participants' understanding of their own *iStuckWithIt* OLM, we observed that participants in both the tutorial and control conditions could understand the main features in terms of *days the goals* were met and how to interpret the display of days

Table 3: Summary of participant insights triggered by scaffolding of the goal prompt questions. Col. 1 the participant ID. Col. 3 daily adherence (%-age of days with >1 step). Col. 4 wear-time (median hours of wear per day with >1 step). Col. 5 to Col. 8 the 4 types of insights associated with the goal prompt questions. N=21

1	2	3	4	5	6	7	8
PID	Tutorial	Daily	Hours (med)	Wkend Vs Wkdays	Work Vs Non Work	Holiday	Change Goals
P20	Y	100	18			Y	Y
P21	Y	91	13	Y		Y	Y
P6	Y	62	9				
P18	Y	44	13				
P19	Y	41	12				
P17	Y	38	11				
P7	Y	32	13	Y		Y	
P16	Y	29	N/A	Y			
P5	Y	15	14	Y			Y
P4	-	88	16	Y			
P15	-	100	17				
P10	-	100	18				
P8	-	100	16				
P9	-	100	17			Y	Y
P14	-	100	20	Y	Y		Y
P2	-	96	17				
P1	-	86	12	Y			
P13	-	70	13	Y	Y		
P11	-	60	16				
P3	-	39	14				
P12	-	31	13				
			Total	38% (8)	10% (2)	19% (4)	24% (5)
			Total	38%	10%	19%	24%

with any data. Both groups also performed similarly on wear-time (median hours of wear per day with >1 step), with most people tending to miss this aspect and similar levels of awareness of it between the conditions. Overall, both conditions performed similarly.

4.3 RQ2: Goal Prompts

We now consider the results for the scaffolding for reflection. While the tutorial was done only be the 9 participants in the tutorial group, all participants then used the interface to explore their own data. When they had finishing doing this, the experimenter asked them to open the scaffolding section of the interface to consider the questions intended to help them consider and reflect on key features. Table 3 summarises the new insights that participants gained in this scaffolded reflection phase. Col. 5 to Col. 7of the table show where the goal prompt enabled a participant to gain new insights ***in addition*** to those the experimenter recorded them as making already.

When asked about weekend versus weekdays, 8 participants (38%) identified new insights – shown in Col. 5. For example, P4 commented, *"I'm not doing very well on weekends"*. Interestingly, this was only true more recently in the last year because she was more active during weekends and weekdays in the previous year in her data. P14 was confident during the think-aloud and pre-interview that he was more active on weekend than weekdays. *"More physically active during weekends because I go hiking or something"*. However, his actual data contradicts this belief and during the think-aloud session, he did not seem to identify this. When prompted by the goal prompt panel questions, he considered this more closely and commented, *"When I see on your website, there is a lot of white [referring to weekends]. I wonder why that is"*. During post interview, when asked if he found anything surprising, he commented, *"I thought I was hitting my step goals more than I was on the weekends, it was definitely an eye opener"*. P21 commenting on doing better on weekends, *"I'm actually overshooting my goal, quite consistently"*. P7 commented on the differences between Saturday and Sundays, *"good on Saturday but not on Sundays"*.

Our participants were generally aware of their work versus non-work periods and most made observations about this during the think-aloud session when they reflected on their own data. The goal prompt questions helped 2 participants find something new, shown in Col. 6. P14 commented that he is not doing as well during work times, *"I seem to have a little more difficulty especially when I'm in the car a lot"*. P13 commented, *"Doing well when going to the office, not so well when working from home. More recently working during the week has been less good"*. During the post interview, when asked whether he learned something new, he commented, *"Going through the Goal prompts, it got me to think about things like looking at weekends versus during the week – it definitely got me thinking about different ways to divide up my time"*. This question did not apply to 2 participants (P9,P1) because they did not work or study. For example, P9 commented, *"doesn't matter because weekend and weekdays are pretty much the same to me, I'm retired"*.

During the main think-aloud exploration of *iStuckWithIt*, most participants reflected on the effect of holidays – this was triggered by visible trends in their own goal achievement. The goal prompt question (Col. 7) helped 4 participants (19%) find more. For example, P21 went back to the *iStuckWithIt* display, reviewing holiday periods more closely and commented, *"I'm mixed on holidays, there are some holidays where I do walking, there are holidays where I was in [location] , ... , wasn't wearing it back there, I was wearing it sometimes but not much"*. P7 reflected on their broad wearing patterns, daily and hourly, during holidays, *"generally I forget to wear it, and I don't take it often and I fear losing it, and I don't want to take all the charging cables"*. Others focused on goal achievement consistency, for example, P9 commented, *"I walked on Christmas day. I walked on New Year's Day – so I'm quite consistent"*.

While most participants reflected upon both their activity levels and wearing behaviour in the main session exploring *iStuckWithIt*, the goal prompt questions helped 4 participants (19%) to consider changing their goal targets. For example, P21 commented, *"I guess I should lower it to something more realistic to achieve that the 10K steps per day. Maybe if I change to 8K"*. P14, after realising that he was less active on weekends commented, *"I want to make sure I hit my step goals more on the weekends, walking a little bit more"*. P9 commented, *"I should change my goal because I'm exceeding it too easily."*. Notably, P4 was confused by the wording of the prompt and failed to see the usefulness of this question. She commented, *"I don't understand why this question is here, I compare myself to myself, it doesn't matter to me"*.

Not all participants considered the goal prompts useful. For example, P16 commented, *"I think the ideas of these questions are interesting but these ones are not as useful for me. I prefer it gives me advice or record what I did"*. P16 also suggested that a prompt about how he achieved a goal each day might be useful in helping him remember. *"I think if each day when I achieve the goal, I get a question about how I achieved the goal, I think that could be useful. It might be useful if it can help me remember how I achieved the goal"*. P21 commented, *"I'm not sure how I would use it, I don't have so many goals that I have to write them down"*. He went on to suggest that these types of prompts might be better integrated into the regular weekly email that that Fitbit currently sends – he suggest this should also highlight notable items, like peaks and ask for reflection on those days then. *"I would like to be prompted as part of the weekly email if there was something interesting e.g., 28K steps days, to note things like I went hiking into the note. It would be nice to have it integrated with my weekly email to complement what I currently see"*.

To summarize, the goal prompt helped 10 participants (48%) gain new insights and 7 of these participants gained two or more insights. Broadly, our results indicate that the goal prompt scaffolding for reflection is valuable.

5 DISCUSSION

To understand our long term OLM, and reflect on their long term behaviour and its implications, users needed to *make sense of their long term data*. Our interface transforms that data into an OLM which highlight two key aspects. It shows a person's *apparent goal achievement*, in terms of how they did compared with their target for steps per day, active minutes or distance. However, to related this to the actual activity level, they also need to appreciate the *implications of their wearing behaviour*. This determines the completeness and meaningfulness of their data. The tutorial scaffold was intended to both support this and to enable us to study how readily people could learn about these aspects with all tutorial condition participants doing this in the context of two sets of carefully designed data, for the hypothetical users Alice and Alex. This section first considers what we learnt from the study of the tutorial scaffolding. We then discuss what our findings reveal about goal prompt scaffolding. Finally, we discuss lessons learned and insights for designing future OLMs for this class of lifelong, life-wide learning for an important aspect of health.

No tutorial is needed to understand daily goal achievement trends (RQ1)

Our study results show that all participants, in either the tutorial or control group found it easy to see that blocks of missing data (grey cells) meant there was no data and that other cells indicated whether they met their goal. In the tutorial condition, participants could do this without assistance for the case of both Alice and Alex. They then went on to make a similar interpretation of their own data. The control condition participants also found this aspect of the interface intuitive and a good basis for reflection about the reasons they met their goals, commenting on aspects like holidays, injury and motivation. So, the tutorial was not needed for this case.

The tutorial did not make a difference in helping people understand wear-time (RQ1)

We observed that a minority of participants could *discover this* from the carefully crafted data for Alice and Alex. Only 56% of participants noted the low wear-time for Alice in the later months of tracking. Even fewer speculated or commented on its meaning. Moreover, only 2 participants noticed inconsistent wear-time for Alex over the long term. Our design of the tutorial was based on providing very clear cases that should have made this concept easier to discover. For the participants who did not work this out, the experimenter explained it to them. Our observations of participants studying their own data indicate that both the tutorial and control condition participants performed similarly on this aspect. So, the tutorial was not helpful for this case and further, many participants found it difficult to appreciate the concept of wear-time. This is unfortunate since it is critical to take account of the number of hours with data to judge whether there is enough data to conclude about whether they have met their goal. There appear to be two main ways to tackle this problem. First, we may need to help people appreciate the importance of taking account of wear-time, particularly the number of hours of wear in a day. Second, the interface should make this clearer than it currently does.

Can scaffolding support insights? (RQ2)

We designed *iStuckWithIt* to present and overview model of people's goal achievement for levels of physical activity. The goal prompt scaffolding was designed to help users reflect on questions that they may not have considered when reviewing their own data. Our study showed that even for our population of existing long term physical activity trackers who are highly educated and familiar with technology, the goal prompt scaffolding could help many of them reflect on important aspects that they had not considered. Our results showed that while *iStuckWithIt* design was useful in supporting reflection on its own, the goal prompt scaffold helped many to consider and discover insights they missed.

Adherence scaffolding design: lessons learned

In this section, we discuss lessons learned and opportunities for future user interfaces that aims to support daily adherence and wear-time for reflecting on long term physical activity data.

First, our results suggest that a more adaptive or personalised approach is needed to teach about wear-time. This should take account of the user's actual wear-time. For people who had high wear-time (e.g., 15 hours or more per day) and consistent wear (near 100% of days have >0 steps), there is no need to consider wear-time when interpreting the interface's goal achievement display. In this case, future interfaces could highlight just the low wear-time days. The configuration interface could be enhanced to define suitable thresholds for this. Then the interface could filter the display to show only days that meet the threshold. For those who do *not* have high wear-time, further work is needed. However, since our long term physical activity tracker participants had high wear-time we have limited information for this type of user.

The goal prompt scaffolding findings demonstrate the insights gained by half our participants, by considering salient aspects. This is also an opportunity for personalisation, so that the cases that deserve attention are provided as prompts. This is particularly likely to be valuable for real world use, rather than our laboratory study. It will be important in ensuring users can readily tackle the challenge of reflection, when confronted for the first time with an unfamiliar OLM interface for their long term performance on physical activity. Opportunities for scaffolding include highlighting aspects known to be important, such as weekend versus weekdays, asking users to consider their performance and behaviour across different environments (e.g., work versus non-work). Our findings show that there is a need to personalise such prompts based on the individual's behaviour. Prompts or questions that do not apply to the participants or do not fit the user's circumstances (e.g., work situation) or goals should not be shown.

Finally, it may be useful to send goal prompts or reflective questions via regular email messages to capture contextual and qualitative information about days of interest (e.g., what did they do or how they achieved a peak day or goal-met day). Participant comments suggest they believe capturing such data may help them remember to consider and reflect on goal prompt questions.

5.1 Limitations

Our study was restricted to existing FitBit users. As FitBits had been widely available for many years [5], this allowed us to recruit participants who has already collected long term data. Also, our study is a lab study of the scaffolding designs. This may limit generalisability of our findings for wider populations of activity trackers in authentic settings. Moreover, since the iStuckWithIt user interface and study was designed for long term data, the usefulness of our scaffolding design on short term data was not examined. Our work paves the way for longer term field studies. Further work is still needed on scaffolding designs that help people understand the impact of data incompleteness in short term physical activity tracker data.

6 CONCLUSION

In this study, we explored how to help people reflect on their long term physical activity goal achievement. We extended previous work by exploring two forms of scaffolding (tutorial, goal prompt) and reported on a lab study of 21 existing long term physical activity tracker's experiences. The tutorial scaffolding results reveal that missing or incompleteness in data on a day-to-day basis (daily adherence) is intuitive and well understood. However, for wear-time (hours of wear per day), it may be more appropriate to provide prompts and alerts based on automatic detection of features such as low wear-time on certain days, weeks or highlighting inconsistencies over time. In addition, our prompt scaffolding (reflective questionnaire) proved effective support for reflection, resulting in new insights. We go beyond previous OLM work, particularly in the focus on a lifelong, life-wide learning goal associated with long term physical activity goal achievement. Our findings provide design insights about these two scaffolding approaches apply, with recommendations on future work and potential roles for reflection scaffolding and its personalisation.

[5] http://www.wareable.com/fitbit/fitness-tracker-sales-2015-fitbit-1169

REFERENCES

[1] Norasnita Ahmad and Susan Bull. 2009. Learner Trust in Learner Model Externalisations.. In *AIED*. 617–619.

[2] Lamiya Al-Shanfari, Carrie Demmans Epp, and Susan Bull. 2016. Uncertainty in Open Learner Models: Visualising Inconsistencies in the Underlying Data.. In *LAL@ LAK*. 23–30.

[3] Roger Azevedo and Jennifer G Cromley. 2004. Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of educational psychology* 96, 3 (2004), 523.

[4] Albert Bandura. 2005. The growing centrality of self-regulation in health promotion and disease prevention. *The european health psychologist* 1 (2005), 11–12.

[5] Albert Bandura. 2005. The Primacy of Self-Regulation in Health Promotion. *Applied Psychology* 54, 2 (apr 2005), 245–254.

[6] Debjanee Barua, Judy Kay, Bob Kummerfeld, and Cécile Paris. 2014. Modelling long term goals. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 1–12.

[7] Timothy K. Behrens and Mary K. Dinger. 2005. Ambulatory Physical Activity Patterns of College Students. *American Journal of Health Education* 36, 4 (2005), 221–227.

[8] F Bentley, K Tollmar, and P Stephenson. 2013. Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. *Tochi* 20, 5 (2013), 1–27.

[9] Susan Bull, Paul Brna, and Helen Pain. 1995. Extending the scope of the student model. *User Modeling and User-Adapted Interaction* 5, 1 (1995), 45–65.

[10] Susan Bull, Blandine Ginon, Clelia Boscolo, and Matthew Johnson. 2016. Introduction of Learning Visualisations and Metacognitive Support in a Persuadable Open Learner Model. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK '16)*. ACM, New York, NY, USA, 30–39. https://doi.org/10.1145/2883851.2883853

[11] Susan Bull and Judy Kay. 2010. Open learner models. In *Advances in intelligent tutoring systems*. Springer, 301–322.

[12] Susan Bull and Judy Kay. 2013. Open learner models as drivers for metacognitive processes. In *International handbook of metacognition and learning technologies*. Springer, 349–365.

[13] Susan Bull and Judy Kay. 2016. SMILI: a framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education* 26, 1 (2016), 293–331.

[14] Lisa Cadmus-Bertram, Bess H Marcus, Ruth E Patterson, Barbara A Parker, and Brittany L Morey. 2015. Use of the Fitbit to Measure Adherence to a Physical Activity Intervention Among Overweight or Obese, Postmenopausal Women: Self-Monitoring Trajectory During 16 Weeks. *JMIR mHealth and uHealth* 3, 4 (2015), e96.

[15] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding Self-Reflection: How People Reflect on Personal Data through Visual Data Exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealthâĂŽ17). ACM, New York, NY, USA*, Vol. 10.

[16] Edward L Deci and Richard M Ryan. 2000. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry* 11, 4 (2000), 227–268.

[17] Michel C Desmarais and Ryan S Baker. 2012. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction* 22, 1-2 (2012), 9–38.

[18] Melissa C Duffy and Roger Azevedo. 2015. Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior* 52 (2015), 338–348.

[19] Chris Elsden, David S Kirk, and Abigail C Durrant. 2015. A Quantified Past: Toward Design for Remembering With Personal Informatics. *Human–Computer Interaction* (2015), 1–40.

[20] Endeavour. 2014. Inside wearables part 2, How the Science of Behavior Change Offers the Secret to Long-Term Engagement. *https://blog.endeavour.partners/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3* Jan (2014). https://doi.org/wp-content/uploads/2015/11/2014-Inside-Wearables-Part-2-July-2014.pdf

[21] Carrie Demmans Epp and Susan Bull. 2015. Uncertainty representation in visualizations of learning analytics for learners: current approaches and opportunities. *IEEE Transactions on Learning Technologies* 8, 3 (2015), 242–260.

[22] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the 2014 conference on Designing interactive systems*. ACM, 667–676.

[23] Daniel A Epstein, Jennifer H Kang, Laura R Pina, James Fogarty, and Sean A Munson. 2016. Reconsidering the device in the drawer: lapses as a design opportunity in personal informatics. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 829–840.

[24] Stuart J Fairclough, Lynne M Boddy, Kelly a Mackintosh, Alexandra Valencia-Peris, and Elena Ramirez-Rico. 2014. Weekday and weekend sedentary time and physical activity in differentially active children. *Journal of Science and Medicine in Sport* In Press (2014). https://doi.org/10.1016/j.jsams.2014.06.005

[25] Reza Feyzi-Behnagh, Roger Azevedo, Elizabeth Legowski, Kayse Reitmeyer, Eugene Tseytlin, and Rebecca S Crowley. 2014. Metacognitive scaffolds improve self-judgments of accuracy in a medical intelligent tutoring system. *Instructional science* 42, 2 (2014), 159–181.

[26] Thomas Fritz, Elaine M Huang, Gail C Murphy, and Thomas Zimmermann. 2014. Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 487–496.

[27] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2015. How do we engage with activity trackers?: a longitudinal study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1305–1316.

[28] Julio Guerra. 2016. Open Social Learner Models for Self-Regulated Learning and Learning Motivation. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. ACM, 329–332.

[29] Julio Guerra-Hollstein, Jordan Barria-Pineda, Christian D Schunn, Susan Bull, and Peter Brusilovsky. 2017. Fine-Grained Open Learner Models: Complexity Versus Support. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 41–49.

[30] Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and Jon Bird. 2015. Activity tracking: barriers, workarounds and customisation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 617–621.

[31] Judy Kay. 1994. The um toolkit for cooperative user modelling. *User Modeling and User-Adapted Interaction* 4, 3 (1994), 149–196.

[32] Judy Kay. 2016. Enabling people to harness and control EDM for lifelong, life-wide learning.. In *EDM*. 10–20.

[33] Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H Nguyen. 2015. Why we use and abandon smart devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 635–646.

[34] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 557–566.

[35] Ian Li, Anind K Dey, and Jodi Forlizzi. 2011. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing (UbiComp '11)*. ACM, New York, NY, USA, 405–414.

[36] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2012. Using context to reveal factors that affect physical activity. *ACM Transactions on Computer-Human Interaction* 19, 1 (2012), 1–21.

[37] Yanjin Long and Vincent Aleven. 2011. StudentsâĂŽ understanding of their student model. In *International Conference on Artificial Intelligence in Education*. Springer, 179–186.

[38] Yanjin Long and Vincent Aleven. 2013. Supporting studentsâĂŹ self-regulated learning with an open learner model in a linear equation tutor. In *Artificial intelligence in education*. Springer, 219–228.

[39] Alessandro Marcengo, Amon Rapp, Federica Cena, and Marina Geymonat. 2016. The Falsified Self: Complexities in Personal Data Collection. *UAHCI 2016* 9737 (2016), 351–358. https://doi.org/10.1007/978-3-319-40250-5

[40] Jairo H. Migueles, Cristina Cadenas-Sanchez, Ulf Ekelund, Christine Delisle Nystr??m, Jose Mora-Gonzalez, Marie L??f, Idoia Labayen, Jonatan R. Ruiz, and Francisco B. Ortega. 2017. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Medicine* (2017), 1–25. https://doi.org/10.1007/s40279-017-0716-0

[41] Antonija Mitrovic and Brent Martin. 2002. Evaluating the effects of open student models on learning. In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*. Springer, 296–305.

[42] Antonija Mitrovic and Brent Martin. 2007. Evaluating the effect of open student models on self-assessment. *International Journal of Artificial Intelligence in Education* 17, 2 (2007), 121–144.

[43] Jakob Nielsen. 1994. *Usability engineering*. Elsevier.

[44] Amon Rapp and Federica Cena. 2016. Personal Informatics for Everyday Life: How Users without Prior Self-Tracking Experience Engage with Personal Data. *International Journal of Human-Computer Studies* (2016).

[45] Amon Rapp and Federica Cena. 2016. Personal Informatics for Everyday Life: How Users without Prior Self-Tracking Experience Engage with Personal Data. *International Journal of Human-Computer Studies* 94 (2016), 1–17.

[46] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers Chalmers. 2014. Personal tracking as lived informatics. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 1163–1172.

[47] Brenda Rooney, Kathy Smalley, Jennifer Larson, and Sarah Havens. 2003. Is knowing enough? Increasing physical activity by wearing a pedometer. *WMJ-MADISON-* 102, 4 (2003), 31–36.

[48] Patrick C Shih, Kyungsik Han, Erika Shehan Poole, Mary Beth Rosson, and John M Carroll. 2015. Use and adoption challenges of wearable activity trackers. *iConference 2015 Proceedings* (2015).

[49] Mical Kay Shilts, Marcel Horowitz, and Marilyn S Townsend. 2004. Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. *American journal of health promotion : AJHP* 19, 2 (2004), 81–93. https://doi.org/10.4278/0890-1171-19.2.81

[50] V J Strecher, G H Seijts, G J Kok, G P Latham, R Glasgow, B DeVellis, R M Meertens, and D W Bulger. 1995. Goal setting as a strategy for health behavior change. *Health education quarterly* 22, 2 (1995), 190–200. https://doi.org/10.1177/109019819502200207

[51] Bernardo Tabuenca, Marco Kalz, Hendrik Drachsler, and Marcus Specht. 2015. Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers & Education* 89 (2015), 53–74.

[52] Lie Ming Tang, Margot Day, Lina Engelen, Philip Poronnik, Adrian Bauman, and Judy Kay. 2016. Daily & Hourly Adherence: Towards Understanding Activity Tracker Accuracy. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 3211–3218.

[53] Lie Ming Tang and Judy Kay. 2017. Harnessing Long Term Physical Activity Data: How Long-term Trackers Use Data and How an Adherence-based Interface Supports New Insights. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 2 (jun 2017), 26:1—-26:28. https://doi.org/10.1145/3090091

[54] Rayoung Yang, Eunice Shin, Mark W Newman, and Mark S Ackerman. 2015. When fitness trackers don't'fit': end-user difficulties in the assessment of personal tracking device accuracy. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 623–634.