Providing feedback on cycling metrics in an adaptive training app

A thesis presented for the master's degree of Human-Computer Interaction

In collaboration with JOIN Sports B.V.

Remy van Tussenbroek
Student number: 6480373

Primary supervisor: Judith Masthoff
Secondary supervisor: Marit Bentvelzen
External supervisor: Adrián Méndez

Department of Information and Computing Sciences
Utrecht University
The Netherlands
June 27, 2023

Acknowledgements

I would first like to express my deepest appreciation to my research supervisors Prof. dr. ir. Judith Masthoff and Marit Bentvelzen MSc for their invaluable guidance throughout my master's thesis journey. Their enthusiastic engagement and profound academic knowledge were evident in the regular meetings where they provided constructive feedback, contributing significantly to the quality of my research.

I also could not have undertaken this journey without my external supervisor Adrián Méndez, Product Manager at JOIN Sports B.V. With his extensive experience in the field of product development, coupled with an amusing sense of humor, I not only gained valuable knowledge but also enjoyed a positive and enriching learning experience. I am truly grateful for his mentorship, which has profoundly contributed to my professional growth.

I also would like to express my sincere gratitude to Jim van den Berg, Chief Executive Officer at JOIN Sports B.V., for providing me with the exceptional opportunity to combine my passion for cycling with research at JOIN Sports B.V.

Additionally, this endeavor would not have been possible without the entire team at JOIN Sports B.V. They have welcomed me with open arms and it was a pleasure collaborating with everyone.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Abstract

Cycling technology has developed rapidly. With the advances of Artificial Intelligence and Machine learning, adaptive training applications have appeared. These provide personalised training plans that adapt to users' characteristics such as weekly availability, goal and level. To be able to provide a sports coach-like experience to the user, developers want to imitate the process of sports coaching. One of the tasks of a coach is to provide feedback. Some adaptive training apps try to involve feedback in the form of data analysis and visualisation. However, actual scientific basis for the (design) decisions taken by these commercial applications is missing. In collaboration with the mobile adaptive training app JOIN, this thesis developed a way of providing post-hoc feedback on cycling metrics in a mobile application. We conducted focus groups to find out what kind of post-hoc feedback users of different levels of expertise want on training performed. Both experienced and less experienced users agreed on wanting feedback regarding the quality of specific intervals. Also, an accuracy score seemed to be the most logical way to show how well the planned training was performed. Lastly, they would like to receive more information regarding the way the training plan is adjusted after a performed workout. Using the focus group insights, we iteratively developed low-fidelity and high-fidelity prototypes. An implementation, the Workout Score, was chosen to be evaluated with a survey to assess how much users appreciate the post-hoc feedback on training performed. The survey showed that participants appreciated the Workout Score, finding it a simple and clear way to show the planned versus actual training. Yet, the explanation lacked specificity to better show where users need to improve to increase their score. Furthermore, the inclusion of RPE in the Score was unclear to some. Additionally, the Workout Score can motivate to stick to the suggested workout but may also demotivate when users receive a low score without actionable feedback. To improve the post-hoc feedback feature in the future, separate scores for each component of the Workout Score should be shown. To add to that, we suggest to reconsider the role of RPE or to provide a more comprehensive explanation for its inclusion. Lastly, we propose to make it more evident to users when they should select that they have completed the suggested workout and when it would instead be better to select that they have done something else. In the future, JOIN could explore if the algorithm could potentially take over this decision-making process.

Keywords: cycling, feedback, metrics, adaptive training, mobile application

List of Acronyms

AI Artificial Intelligence

ANN Artificial Neural Networks

BD Back-end Developer

CEO Chief Executive Officer

eFTP estimated Functional Threshold Power

FTP Functional Threshold Power

HCI Human Computer Interaction

IoT Internet of Things

NP Normalized Power

PD Product Designer

PM Product Manager

RPE Rate of Perceived Effort

SVM Support Vector Machines

TSS Training Stress Score

VI Variability Index

Contents

1 Introduction

Cycling is a sport that is open to many different athletes with varying goals. Recently, partly due to the COVID-19 pandemic [4], there has been a strong growth in recreational cycling. The technology that is used in cycling has not stopped either. New sensors, smartwatches [36, 14], cycling computers [5] and training platforms [55, 47] are released every year. With the advances of Artificial Intelligence (AI) and Machine Learning, adaptive training applications have appeared [20, 53, 34, 52, 46, 35, 10]. These provide personalised training plans that adapt to users' characteristics such as weekly availability, goal and level. The advancements are also reflected in research done in the Human Computer Interaction (HCI) field. Wunsch [66] used the Wizard of Oz evaluation method to investigate a virtual safety coach application for supporting (novice) cyclists in real-world scenarios.

1.1 Problem

To be able to provide a sports coach-like experience to the user, adaptive training applications try to imitate the process of sports coaching [68]. One of the tasks of a coach is to provide feedback [50]. Feedback has been researched widely in the HCI context with various feedback modalities used to help cyclists [29, 33, 42]. Research is mainly done on real-time feedback, feedback provided while participating in sports. However, literature on post-hoc feedback is lacking, some adaptive training applications try to involve feedback in the form of data analysis and visualisation [34, 53, 47, 55]. However, actual scientific basis for the (design) decisions taken by these adaptive training applications is missing.

1.2 Goal

In collaboration with the mobile adaptive training app JOIN, this thesis aims to develop a way of providing post-hoc feedback on cycling metrics in a mobile application. The thesis will use a mixed approach of qualitative and quantitative research methods including focus groups and a survey. Using these research methods, we hope to answer the following research question:

RQ How can effective post-hoc feedback on training performed be provided in a cycling application?

We will answer RQ in the context of JOIN, using active users of the app. The objective is to design a post-hoc feedback implementation in JOIN using its users' preferences and expectations regarding receiving feedback on the workouts suggested by the app after they have completed these workouts. Besides RQ, we aim to address the following sub research questions:

SQ1 What kind of post-hoc feedback do users of different levels of expertise want on training performed?

SQ2 How much do users appreciate the post-hoc feedback on training performed?

For SQ1, we will make use of the different Activity profiles available in JOIN, further explained in chapter 3. This allows us to research the possible different post-hoc feedback requirements for experienced and less experienced cyclists. Finally, to address SQ2, we will evaluate users' perspectives on the post-hoc feedback in terms of its understandability, usability, and overall value within the app.

1.3 Structure of this work

This thesis is structured as follows. Chapter two provides an overview of related work. This chapter discusses how cycling is represented in HCI literature and technology, the use of data in cycling, the rise of adaptive training and how feedback is given in cycling. Chapter three discusses the mobile application JOIN, the workings of the app and the philosophy of the company. Chapter four describes the process and results of gathering user insights regarding the post-hoc feedback feature in the JOIN app with the use of focus groups. Chapter five discusses the iterative process of creating prototypes for the design of the post-hoc feedback feature. Chapter six describes the evaluation of the implementation of the post-hoc feedback feature. Chapter seven describes the contributions and limitations of this work, as well as directions for future research. Lastly, a final conclusion is presented in chapter eight.

2 Related work

This chapter provides an overview of the related work. The chapter starts with a section on cycling in HCI, which discusses several cycle-based computing technologies and applications and considers research done on cycling in the HCI context. This section is followed by a section on the use of data in cycling which focuses on the part sensors and metrics play in cycling. Furthermore, we discuss adaptive training, it's definition, examples of commercial applications that provide adaptive training plans and the different algorithms that are used. Lastly, the role of feedback in sports and cycling is described, we discuss real-time and post-hoc feedback and conclude with how feedback can be personalized effectively.

2.1 Cycling in HCI

A range of cycle-based computing technologies and applications have emerged. Whatever the goal, the sport and what's involved is so multi-faceted that it allows for the rise of technologies from different areas such as training, adventure, sleep, recovery and nutrition.

Sensors are an important technology in cycling. These range from simple sensors that measure speed or cadence, to more complex ones that measure power and heart rate. Further, cycling computers [5] are used to display the metrics from these sensors. Most computers also have GPS functionality and the ability to load and view routes with directions. These routes can be created on platforms like RidewithGPS [40] and Komoot [21]. After a workout, cyclist can directly upload their data from the cycling computer to analysis platforms like Strava [47] and Trainingpeaks [55]. Lots of training platforms have arisen to aid cyclists and coaches in planning training. Mobile applications have been developed to provide adaptive and personalised training plans. Additionally, smartwatches from brands like Garmin [14] and Polar [36] allow cyclists to track their lives to the detail. Another tracker that has use cases in cycling is the WHOOP band [62]. This band monitors your recovery, sleep, training, and health, with personalized recommendations and coaching feedback. In addition, to maximize recovery tools like the Normatec 3 from Hyperice [18] have been developed. This pants-like device uses precision pulse technology that helps to increase circulation, revive muscles, and reduce swelling. Also, EatMyRide [9] is a mobile application that provides personalized nutrition advice and insights to balance your real-time energy burn and intake rates. Technology to increase cyclist safety has also seen advancements. Garmin have developed the Garmin Varia [15] lights that has a rear-view radar with taillight that pairs with your bike computer or smartphone and alerts you to vehicles approaching from behind.

Further, often when the weather outside is not safe to ride in, cyclists resort to cycling indoors. This is done on indoor trainers. These can be wheel-on trainers [71], where the back wheel remains on the bike and the bike is put onto the trainer. A level up are the direct-drive trainers [69], where the back wheel is removed and the bike is put onto the trainer that has a cassette mounted on it. Moreover, with the growing popularity of indoor cycling, indoor bikes [59] have been created that replace the road bike entirely. To make cycling indoors more enjoyable, virtual cycling worlds like Zwift [70], BKOOL [2], ROUVY [41] and Wahoo RGT [61] have been developed. These provide engaging environments that replicate riding outdoors as much as possible. They use virtual fantasy worlds or replicate real-world areas in the virtual cycling environment. The social aspect in these applications is quite present as well. Cyclists can partake in virtual races, group rides and workouts. Also, users are able to join a club and organise rides together. Often, users communicate through online platforms like Discord [8]. Here, they discuss race tactics while racing or have a casual chat while doing a group ride.

The advancements in technology in cycling have also reflected in research done in the HCI field. Wunsch [66] used the Wizard of Oz evaluation method to investigate a virtual safety coach application for supporting (novice) cyclists in real-world scenarios. To add to that, Michahelles [27] introduced the idea of an Eternity Bike, a (partly) automated bike with active safety functions that will fully integrate into the automated and cooperative transport systems of the future. Suoheimo et al. [48] looked into what challenges winter conditions set for cyclists. The results show that the most commonly reported challenges relate to fast battery drainage in the cold, and the use of gloves which hinder the use of touch screens and general dexterity with equipment.

2.2 The use of data in cycling

Data and its analysis has become an integral part of our current lives. It has also found its way to sports [38]. Compared to other sports, cycling allows for frequent and accurate athlete monitoring [24]. Therefore, cycling is an easily quantifiable sport [28]. To add to that, to gain the slightest advantage, lots of research is continuously done to improve aspects of the performance like aerodynamics [26], nutrition [57] and training [22].

Sensors play an important part in the data collection process. As mentioned by de Leeuw et al.[24], the fact that the bike and rider are equipped with sensors, enables regular and precise athlete monitoring. Many sensors exist that are well-known for most amateur cyclists like a speed sensor, cadence sensor, heart rate monitor and power meter. Developments in high performance cycling have led to the use of sensors like on-bike aerodynamic sensors [58][31], lactate meters[37], glucose meters [49] and sweat meters[30].

The data that is produced by these sensors is often presented in real-time on a specific cycling computer [5]. These devices have customizable data fields that allow users to personalise the way the data is shown. Recently, mobile applications have been released that replicate such devices [64]. In addition, data is used post-hoc to perform analysis of the completed cycling workout. Importantly, in both cases, the raw sensor data is presented in the form of a metric such as speed, power or heart rate. For most cycling metrics, they can be extended with metrics that provide more insights such as average speed over the total workout, current speed or average speed in an interval.

Cycling metrics help quantify the human performance on a bicycle. Power output measured by a power meter forms the basis of other common modern training metrics. These power-derived training metrics have a basis in physiology, and others are models that have certain benefits and drawbacks [28]. TrainingPeaks [55] is a training platform for cyclists that has laid out a great basis of power-derived metrics that are used widely. Examples from their glossary include metrics such as Functional Threshold Power (FTP), Normalized Power (NP), Training Stress Score (TSS) and Variability Index (VI) [54]. In other platforms, several of these metrics might have a slightly different naming or formula details, but the purpose remains the same.

In addition, there exist metrics that are not derived from sensor data. For example, cyclist often use a rating of perceived effort (RPE) when performing intervals or depicting the intensity of the performed workout.

2.3 Adaptive training

The number of cyclists has grown [4] significantly in the last couple of years. With more novel-level cyclists in the sport, the need grows for guidance to achieve goals, participate and compete in events and generally improve as an athlete. Classically, sports coaches have been around to assist athletes with such aspirations. However, with the recent advancements in technology, the bar has been lowered by apps [53, 34, 52, 46, 35, 10] that provide quality recommendations of workouts. They take into account personal characteristics such as the time you have available or your level of experience in the sport. These applications provide personalized and adaptive training plans at a considerably lower price point than a personal coach. This offers an easy way for cyclists to start structured training. Besides that, these apps have been proven to be effective. Silacci et al.[45] showed that a virtual coach can compete with human experts in making proper personalized training plans.

To be able to provide a sports coach-like experience to the user, developers want to imitate the process of sports coaching. They attempt to do so by looking at what a real sports coach does [68]. The trainer selects the main aspects of the athlete's performance that can be enhanced and plans training sessions to address the problems that they have observed. Coaching focuses on improving the performance level and learning capacity. It involves giving feedback, but also the use of some techniques, such as motivation, efficient question-asking, and the management style conscious adaptation to the athlete's training level, in relation to the objective to be fulfilled [50]. Broadly speaking, coaching actually ensures the athletes' potential realization, for the maximization of their own execution; it consists of supporting the athletes to rather learn themselves than to teach them effectively something [13]. The coach will need to be able to: assist athletes to prepare training programs, communicate effectively with athletes, assist athletes to develop new skills, use evaluation tests to monitor training progress, and predict performance [65]. As depicted, sports coaching is a challenging task that requires recognition of many different facets.

The training plans that most apps provide are personalized in the sense that their algorithms take into account personal characteristics such as their weekly availability, goal, training experience, and current fitness level. The plans are adaptive because they adjust to changes in the users' schedule and performance. This study is performed in combination with the adaptive training app JOIN [20]. There exist other apps that provide adaptive training these include TrainerRoad [53], Pillar [34], AIEndurance [52], Spoked [46], PKRS.AI [35] and Enduco [10]. All provide an adaptive training plan to the user, but some offer features that are not available in others. TrainerRoad is a well-known app within the cycling community and has many users. TrainerRoad is a training platform with a mobile application that has recently launched their Adaptive Training function. With a scientific basis TrainerRoad developed an algorithm optimized by machine learning to provide their Adaptive Training function. Users complete intelligently-recommended workouts. After every workout, Adaptive Training's AI analyses your performance. Progression Levels adjust to reflect your current abilities and training progress. In TrainerRoad, your unique range of cycling abilities is reflected in your Progression Levels, a dynamic metric that quantifies your fitness and training progress across seven power zones (Endurance, Tempo, Sweet Spot, Threshold, VO2 max, Anaerobic and Sprint). As you train and your fitness subtly changes, your Progression Levels adjust to reflect those changes. Progression Levels help Adaptive Training give you the right workout each day, and provide insight into your abilities as they grow and change. Your upcoming workouts adapt to your Progression Levels to make you faster. Also, Pillar offers the ability to contact a real-life coach as an addition to their algorithm recommendation. Aside from cyclists, AIEndurance also targets other types of endurance athletes like runners and swimmers. For example, the app provides runners with predictions for their time at a race like a marathon based on their current training. Furthermore, Spoked shows the users a personalized suggestion regarding how many hours they should have available to ride in a week to achieve their goal. Moreover, PKRS.AI gives the user a daily nutrition plan that syncs with the day's workouts. On endurance days it will give the user nutrition that is balanced to create optimal fat burning. On high intensity days the user will be given a higher carbohydrate meal plan and on strength days the user will be given a higher protein diet. Lastly, Enduco offers a fair way of competing with other app users, using their Enduco Crush feature. The users don't have to be in the same place or ride at the same time. The app makes the performances comparable among each other and creates a ranking.

2.3.1 Algorithms for adaptive training

Machine Learning and other data-driven analytic algorithms are the primary technologies behind adaptive training applications. These algorithms analyze data such as heart rate, power output, and training volume to make recommendations for training intensity, duration, and rest periods. Artificial Neural Networks (ANN) are one of the most used approaches according to [44]. This is due to recent advances in deep learning and ANN-based models, but mainly its generalization capability. Next to ANNs, Super Vector Machines (SVM) are sometimes used, but not as much as Silacci et al.[44] would have expected. SVMs can be quite powerful as they are similar to ANNs but they are able to avoid overfitting. Others have made use of evolutionary computation. Fister [12] used swarm intelligence algorithms to search for the best combination of training sessions within a cycle of duration one week. Connor et al.[6] combined control system theory and evolutionary computing by implementing an intelligent control system. This artificial intelligence-based feedback controller outperformed random and proportional strategies in optimizing future training loads. Also, Kumyaito et al. [23] employed Adaptive Particle Swarm Optimization and incorporated constraint methods to consider physiological constraints like monotony, chronic training load ramp rate and daily training impulse.

2.4 Feedback

People partake in sports for various reasons which can be simply to have fun and enjoy the activity, to get some mental recreation or to enhance the social relations related to the sport context. Importantly, improving performance can also be seen as a motivation to do sports as people also participate in sports to compete and keep fit [43]. One of the most important variables affecting learning and subsequent performance of a skill is feedback [17].

Hawley et al. [17] differentiate two sources of feedback in sports. Firstly, intrinsic or inherent feedback, which is feedback from the athlete's own sensory channels (i.e. sight, hearing, touch). Although some information from intrinsic sources provides clear information (i.e. the ball missed the goal), more detailed information (i.e. coordination of joint activity, amount of force produced) often requires experience from the performer to evaluate. Secondly, extrinsic feedback comes from an external source (usually a coach). It is meant to complement intrinsic feedback, helping the athlete to compare what was done to what was intended. Extrinsic information is regarded to speed up learning for the majority of complex skills and may even be required to help the athlete attain their peak performance levels. The coach should be able to offer useful information about a given moment to help develop that skill along with error detection and correction mechanisms. As a result, extrinsic feedback can be seen as a complement to intrinsic feedback.

This distinction of feedback can be extended. On the one hand, feedback can be given while participating in sports, which we will call real-time feedback. On the other hand, feedback can be given after participating in sports, which we will call post-hoc feedback.

2.4.1 Real-time

Various feedback modalities are used to help cyclists in real-time. Visual feedback is a common method. A domain of cycling where real-time visual feedback can be beneficial is racing. Traditionally, in-race feedback has been done by radio voice-communication or visual signals.

Municio et al. [29] presented and tested prototypes that used Internet of Things (IoT) technologies to provide continuous athlete monitoring in competitive (professional and amateur) cycling environments. The prototypes formed a reliable mesh network with the bikes as nodes, sending real-time sensor information from cyclist to cyclist. The network did not require any internet connection. The real-time sensor included metrics such as heart rate, speed and location. This data was analyzed in real-time to estimate the performance of each rider and derive instantaneous visual feedback. Additionally, Xu et al. [67] presented an integrated inertial sensor and mobile computing system for real-time cycling performance guidance via biomechanical classification. The system determines a user's pedaling profile during cycling and offers real-time visual feedback to help the user increase pedaling efficiency.

Haptic feedback also provides assistance to cyclists in real-time. Peeters et al. [33] developed an indoor training bike that provided real-time vibrotactile feedback on the aerodynamic cycling position. The system calculated the projected frontal area of the cyclist and the bike as an indicator of aerodynamic drag to estimate the wind resistance. Vibrotactile feedback was applied on the neck of the cyclists when they exceeded a certain calibrated projected frontal area value, corresponding to the most aerodynamic position. Compared to two ways of visual feedback, Berentsen et al. [1] showed that feedback using a vibrotactile belt provided the best perceived sense of balance in novice mountain bikers while being aided in adjusting their body position.

Likewise, the auditory sense is used for real-time feedback. Auditory feedback is used in a system that aimed to improve bicycle pedaling [32]. The system makes a short sound every time a pedal crank turns a quarter rotation. The system plays music in addition to the feedback sound, and the users are encouraged to pedal at constant speed by matching the feedback sound with the music beat. To add to that, Schaffert et al. [42] showed that interactive sonification of the forces applied on the pedals improve the ability to identify asymmetries in pedal technique.

2.4.2 Post-hoc

Reijne et al. [39] performed an observation study of the descending technique of members of a World Tour cycling team. The bicycles were equipped with a system that could measure various metrics like position, steer angle and brake force front and rear. Afterwards, feedback about important indicators like brake point and apex position could be used to improve descending.

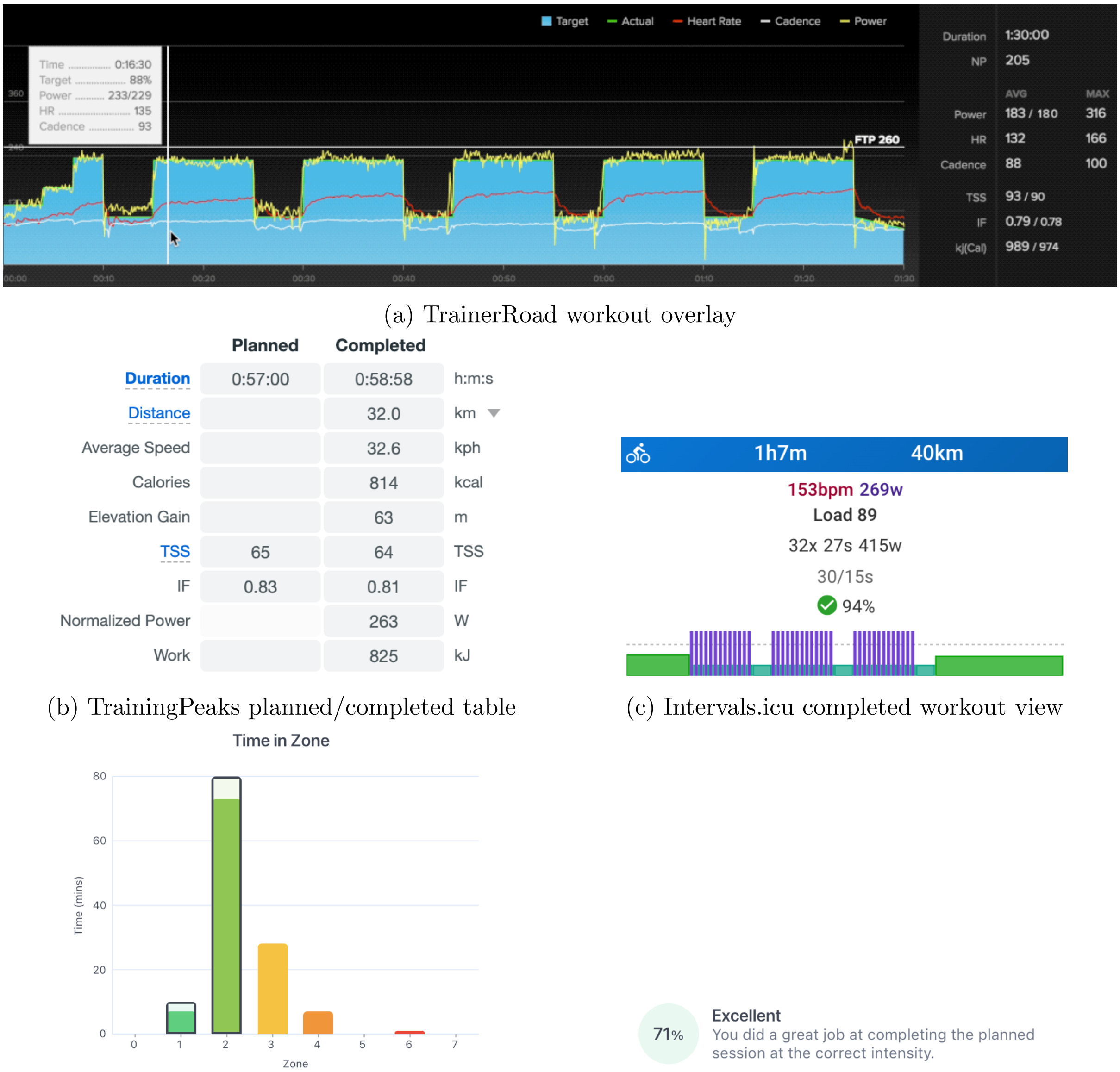
Research regarding post-hoc feedback in cycling is quite limited. However, there exist some commercial applications that provide a way of post-hoc feedback. Mainly, the feedback is given in the form of data visualisation or analysis. Cycling analysis platforms like TrainingPeaks [55], Strava [47] and Today's Plan [51] offer deep analysis of the performed workout. This allows the user to select parts of their workout and analyse them. Various analysis metrics are shown like described in section 2.2. To add to that, colour labeling seems to be a popular way of showing workout compliance. The completed workout card is given a colour label, dependent on how it compares to the planned workout. The analysis platforms use different conditions for the compliance colour labeling. The colour labeling of TrainingPeaks and Today's Plan are displayed in Table 1 and Table 2, respectively. In TrainingPeaks, users can select which value the workout card compliance is based on (duration, distance, or TSS). In Today's Plan the main workout card compliance is based on duration. When users go into the workout analysis, they are presented with a table that shows the planned and completed values for metrics like duration and distance. These have their own compliance colour but follow the same labeling scheme. The only difference is that they are based on each individual metric.

Table 1: TrainingPeaks colour labeling

Colour	**Description**
Green	Completed value is within 15% of planned values.
Yellow	Completed value is between 15%-35% of planned values.
Orange	Completed value is more than 50% above or below planned.
Red	Workout was not completed.
Grey	Unplanned workout.

Table 2: Today's Plan colour labeling

Colour	**Description**
Green	Completed value is within 15% of planned values.
Orange	Completed value is between 15%-35% of planned values.
Red	Completed value is outside 35% of planned values or workout was not completed.
Grey	Rest day or unplanned workout.



	Planned	Completed	
Duration	0:57:00	0:58:58	h:m:s
Distance		32.0	km
Average Speed		32.6	kph
Calories		814	kcal
Elevation Gain		63	m
TSS	65	64	TSS
IF	0.83	0.81	IF
Normalized Power		263	W
Work		825	kJ

(a) TrainerRoad workout overlay

(b) TrainingPeaks planned/completed table

(c) Intervals.icu completed workout view

(d) Pillar time in zones graph

(e) Pillar session score with description

Figure 1: Post-hoc feedback appliances in commercial applications

Additionally, some training platforms like TrainerRoad overlay the planned workout with the completed workout (Figure 1a). Also, a percentage is sometimes shown that refers to how well the completed workout complied to the planned workout. For example, Intervals.icu [19] (Figure 1c) shows a check mark with percentage. The mobile adaptive training application Pillar [34] shows a session score with a description (Figure 1e). To add to that, time in zones is a recurring way of displaying compliance to a plan. Pillar shows seven zones on a bar chart with the scheduled duration per zone marked by a thick lined bar and the actual duration with the corresponding zone colour (Figure 1d). Furthermore, details of specific intervals are often shown in a table. TrainingPeaks shows a table with planned and completed metrics of the workout (Figure 1b). At a glance, users can detect possible differences.

2.4.3 Personalizing

As noted in section 2.2, data is used in abundance in cycling. Sensors collect vast amounts of data during training. For example, power meters, heart rate monitors, cadence and speed sensors. The raw data that is obtained from these sensors are mostly converted into specific metrics that combine data with formulas derived from cycling training theory, such as Normalized Power or TSS. These metrics enable deeper quantitative feedback. However, having access to all these metrics begs the question of what to present to an end user without overwhelming athletes with complex data. As this information overload reduces their capabilities to understand and use the data properly [63]. Moreover, a high amount of data can lead athletes to confusion or misunderstanding and could even make them lose motivation towards using data collection tools [44]. To make the feedback more valuable for each rider, it should be personalized [7]. Thus, it is necessary to consider how and what feedback is presented to a specific user.

Zahran et al. [68] proposed a conceptual framework for the generation of Adaptive Training plans in sports coaching. Their framework introduced the importance of several Behavioral-Change Features including personalized feedback. They noted a crucial point for this kind of feedback: it is most effective when it is specific. Specificity could be integrated by comparing current performance to past accomplishments and previous goals [68].

How and when feedback is presented is also an important factor in real-life coaching. Hawley et al. [17] noted that a coach must consider the precision of feedback. Feedback that is more precise seems to be more beneficial. Interestingly, this does seem to be dependent on the skill level of the athlete. The higher skilled the athlete, the more precise the feedback should be. Additionally, Hawley et al. [17] found that the amount of feedback is in the same way dependent on the skill level of the athlete. Although high amounts of feedback may be beneficial early on in the learning process, too much feedback later in learning may actually impair performance.

Summary Thus, a range of cycle-based computing technologies and applications have emerged. Whatever the goal, the sport and what's involved is so multi-faceted that it allows for the rise of technologies from different areas such as training, adventure, sleep, recovery and nutrition. To add to that, cycling is an easily quantifiable sport [28]. Many sensors exist that allow cyclists to see different metrics in real-time. These metrics help cyclists quantify their performance on the bike. Moreover, adaptive training applications offer an easy way for cyclists to start structured training. They take into account personal characteristics and try to imitate the process of a real-life sports coach. Machine Learning and other data-driven analytic algorithms are the primary technologies behind these applications. Furthermore, feedback is one of the most important variables affecting learning and subsequent performance of a skill [17]. Feedback has two sources according to Hawley et al. [17]: intrinsic and extrinsic. This distinction can be extended by real-time and post-hoc feedback. Research regarding post-hoc feedback in cycling is quite limited. However, there exist some commercial applications that provide a way of post-hoc feedback in the form of data visualisation or analysis. Lastly, to personalize feedback, it should be specific and precise, dependent on the skill level of the athlete.

In this thesis, we aim to develop a way of providing post-hoc feedback on cycling metrics in collaboration with the mobile adaptive training app JOIN [20]. The next chapter describes its workings and brand philosophy.

3 JOIN

This chapter starts with a description of the adaptive training application JOIN [20] and is followed by a short introduction to its brand philosophy. Further, we describe the journey the users take in their training plan. We conclude this section with discussing other functionalities that the app offers.

JOIN is an app available on iOS and Android that provides personalized, adaptive training plans for cyclists. It has a database of over 300 workouts designed by World Tour-level cycling coaches. The training plan will be customized based on the users' input regarding several parameters: training devices, activity profile, type of bike, goal, and weekly availability.

- **Training devices**
  Users can have multiple training devices, they indicate whether they use a power meter and/or a heart rate monitor or solely RPE.
- **Activity profile**
  The activity profile provides a way for the user to pick their level of experience in cycling. Whether that is Recreational (<4 hours/week), Amateur (4-6 hours/week), Intermediate (6-9 hours/week), Advanced (9-11 hours/week), or Pro (>11 hours/week).
- **Type of bike**
  Users can input their preferred type of bike. They can choose from Road, Mountain bike, Gravel, and Trekking.
- **Goal**
  You can select a tailored made training plan or an event as your goal. Additionally, there is an option to create your own custom cycling goal.
- **Weekly availability**
  Users should indicate their weekly availability. You can change your availability daily, and those changes will immediately be reflected in the training plan. The users are prompted at the end of the week to provide their availability for the next week.

Users can connect JOIN to other applications. The app can be connected to Garmin [16], Strava [47], Wahoo [60], Zwift [70], and TrainingPeaks [55]. These applications act as another data source that is used by the JOIN algorithm to provide an even more personalized training plan.

3.1 The JOIN philosophy

JOIN as a brand has its own vision on how to develop an adaptive training app for cyclists. Primarily, their focus is on providing a dynamic and adaptive training plan based on science and pro-level knowledge to improve every cyclist in an easy, accessible and fun way. Their design is clean and minimal, ensuring that the app remains intuitive and straightforward.

3.2 Training plan journey

When a user has completed inserting all the values for the parameters, the training plan can start. The user is presented with a page that provides an insight into the current progress of the plan, how many days of the plan remain, current fitness improvement and the current training phase. The fitness improvement displays whether your fitness is deteriorating, remains the same, improves or that you're doing too much with risk of overtraining. The current training phase can be viewed in more detail showing the upcoming weeks of training with an option to view to accompanied predicted progress in level. Different training phases like build, training and rest are displayed in blue, green and grey respectively. Most importantly, the planned workouts for the week are presented. When the training for the day has been completed, the training data will be automatically uploaded through Strava or the user can add the training manually. After every workout, the user must provide feedback about how the training went. This is done using an RPE scale ranging from 1-10. This feedback is taken into account by the JOIN algorithm for the upcoming workouts. The JOIN algorithm considers when the user does not complete the planned workout or did a little more than planned. Additionally, users can view their future training hours week by week, including the current training phase and the accompanying predicted level increase. Lastly, users can indicate if they are ill or injured. The training plan will automatically provide some days off.

3.3 Other functions

Articles regarding news of the app, recent podcasts and training advice are uploaded to the homepage of JOIN. Podcasts and training advice articles provide theoretical background information for the users to understand their training and cycling better. It is also possible to view the entire training database of JOIN. The 'Workout' page shows different categories of workouts: VO2max, Threshold, Strength, Endurance, DIY tests. There is also a discover function where users can view different cycling events nearby and meet cyclists. People can become friends and follow each other in the app. There is an option to chat with those friends. Lastly, there is also a possibility to create groups with other cyclists.

Now that the workings and the philosophy of JOIN have been established, we continue by describing the process of conducting user research with real users to discover how they want to receive post-hoc feedback from JOIN.

4 User research

This chapter discusses the gathering of user insights regarding the post-hoc feedback feature in the JOIN app. This chapter describes the approach that was used for the recruitment of participant and the analysis of focus group data. This chapter concludes with the results of the thematic analysis of the three focus group sessions.

4.1 Focus groups

The purpose of the focus group sessions was to investigate how to provide more insights for JOIN users into how they performed their workouts. Previously, only the user had an option to provide feedback, but feedback from JOIN on how the user performed was missing. The following aims were used as a guideline for the outcome of each session:

- How would a JOIN user want to receive feedback after a workout from JOIN.
- Find out certain design preferences.
- How much information/graphs should be presented.

The focus group protocol is available in Appendix A.1. The Ethics and Privacy Quick Scan of the Utrecht University Research Institute of Information and Computing Sciences was conducted (see Appendix D). It classified this research as low-risk with no fuller ethics review or privacy assessment required.

4.1.1 Participants

Participants were recruited via a Typeform [56] survey that inquired about their interest in participating in a focus group session at the JOIN office. This survey was sent to users of the JOIN app. A total of 803 invite emails were sent out. The survey got 76 responses, 54 showed interest in participating in a focus group. In total, 16 participants took part in a focus group session, divided over three groups. The first group consisted of 6 participants (6 Advanced). Two participants took part online through Zoom, this sessions was in a hybrid setup. The second group consisted of 5 participants (2 Recreational, 2 Amateur, 1 Advanced), the third group consisted of 5 participants (4 Advanced, 1 Pro). Intentionally, a separation was made between more experienced and less experienced users. The Intermediate category was used as midpoint. Splitting the categories would ensure contribution in focus groups. As it would help avoid less experienced participants to act reserved as they might feel overwhelmed in a group setting with more experienced members. Out of the 16 participants, one was female and the other 15 were male. Their age group ranged from 25 to 65. Video with audio recordings of the sessions were made with consent of the participants. This was done using Microsoft Teams. The main language used in the sessions was Dutch, as this was the native language for all participants. English was occasionally used in the first group as the native language of the session assistant was not Dutch. In total, we collected 5 hours and 2 minutes of video with audio. A focus group recording lasted on average 100 minutes ($M = 100.33, SD = 9.07$).

4.1.2 Analysis of the data

All focus groups were transcribed verbatim, as well as analyzed, in their respective language. The mp4 files were imported into Microsoft Word and MacWhisper[3]. Their automated transcription was used as a base for the transcripts. This base was manually edited by me to fit the recordings accurately. If needed, quotes were translated to English by me. These quotes have been translated as accurately as possible, yet in some quotes grammatical adjustments have been made to ensure readability. The transcripts were loaded into NVivo analysis software[25]. Open coding was used to code all transcripts. Codes were merged into categories after iterative discussions with both supervisors.

4.1.3 Results

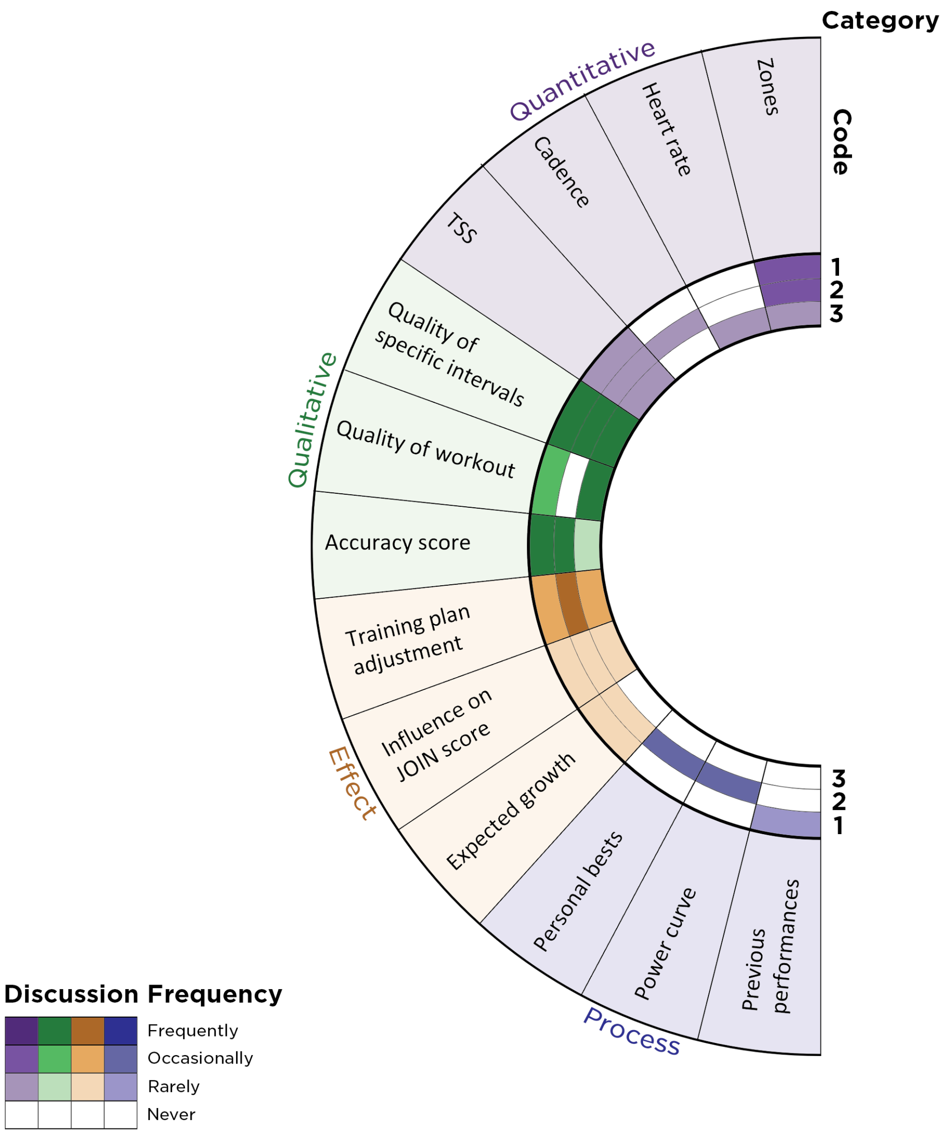


Figure 2: Spectrum display showing the overall thematic analysis of the three focus group sessions.

Quantitative This feedback is based on metric and quantitative data, retrieved from sensors directly or created through formulas. Group 1 and 2 talked frequently about power or heart rate zones as quantitative feedback. The main point made was to have a clear view of the time spent in zones after the performed workout. As P16 noted:

> *"You could show time in zones and then just show what the app expected and what you actually did."*

To add to that, P15 mentioned that a score could be given per zone to show if the user spent the planned duration in the correct zone. Also, P16 preferred to have the zone display in the current colouring used by JOIN. Training Stress Score (TSS) was a metric that all groups discussed occasionally. This number shows the training load of the training. Thus, it can be used as a way to see how hard the training was. P4 said that TrainingPeaks uses TSS as the main metric to show training load. The planned TSS and the TSS you have realized is shown to visualize differences in the completed training.

Qualitative This feedback focuses on the quality of the performance. All sessions discussed qualitative feedback as an important type of feedback that they would want to receive. All groups frequently mentioned wanting to see feedback on the quality of the specific intervals that they had to perform. P11 said:

> *"There would also be a wish for me, to get the feedback on your training itself. Of course, you have a certain training schedule for that ride, so a block of about a minute et cetera. But then I sit and look for myself, like okay, how well have I done those blocks? Did I finish them properly? Are the wattages good on average? But feedback immediately that saves me a few things to lookup."*

The reasoning for feedback on the specific intervals is mostly based on the amount of fluctuations when performing a cycling workout outside compared to indoor. P1 said:

> *"And especially outside, because there is a lot of fluctuation there. If I do a block of 250 watts. And then I do a block of 250 watts, but in those first 5 minutes of those 10 minutes I am sometimes at 220 and sometimes at 300, but the average is 250. Then I will see that I have performed the training properly."*

Outdoor fluctuations occur quite often, P3 mentioned that a traffic light or a road ending can easily force you to finish an interval earlier than planned. There was also a need for the app to recognize the errors and provide actionable feedback. For example, P1 said that when you always miss the last minute of a four minute block, JOIN should see this and prompt you to do something about it. Additionally, P10 said that advice from JOIN would be useful when the first blocks are performed too hard, which leads to worse performance in the later blocks. Providing feedback on the specific intervals requires more than just showing the time in each training zone. Otherwise, the real intent of the intervals would not be correctly displayed. As P13 mentioned:

> *"That [quality of a specific interval] would be different than just time in zones right? Because if you have blocks of 3x10 minutes at a certain wattage, it's different than cycling 30 minutes into a headwind and then 30 minutes easy riding. [...] In the end, the idea of the training theory is that you take the rest after the interval so that you can do the interval again at the right intensity after your rest."*

Overlaying was discussed as a way of visualizing the performance of the intervals. As noted by P11, currently the app shows the suggested workout with coloured blocks, this could be overlaid with the actual training that you have done.

The feedback on quality is also wanted on a broader level, regarding the entirety of the workout. All groups discussed some kind of score to show how well the planned training was actually performed. P11 said that the focus of the score should be on the intended effect of the workout. P11 also mentioned expert involvement. The expert that created the workout should be in charge of deciding the details of the score. As the expert is aware of what is most important for the effect of the training. The score presentation should be simple, in line with the philosophy of the app. P1 recommended using a simple bar or score. Also, P8 mentioned a percentage that would show how you did your training i.e. 80% correct. The display of the score should be dynamic. The overall score can be immediately visible, but details regarding the score and intervals should be visible in a separate screen for users that want to delve deeper. P8 noted:

> *"One day I like to, for example, go deeper into it. Then you know a bit more, you have filled in your RPE and that gives you information. But then you still have a kind of pull-down where you go into the information in more depth or something. But that doesn't mean that you immediately have all that data, because you don't always feel like it, I think. It also has to remain kind of understandable."*

Importantly, the score is preferred to be visible more directly, as P10 said that the score can be a reason to take a look at the details later.

Effect This feedback relates to the effect workout performance has on different facets of cycling. The major point made here was that users want to see the effect of the completed workout on the future workouts. When a workout is loaded into the app, the training plan is adjusted according to that workout. But as P7 said, this change is often not that clear:

> *"Well, if you did a workout for example. It was not as heavy as intended or much heavier. Then you will see that the training schedule will adjust during the week. You can see that in the app, but only if you notice it yourself. Then why is that? Then why does that happen? Is that because I trained too heavily there? I would like to get a little more information and feedback on that. Doesn't have to be very deep but just short."*

Additionally, users would like to see the influence of a workout on their progression. For example, regarding the development of your estimated FTP (eFTP), P13 liked that in Intervals.icu you can see per workout how much your eFTP has increased due to that specific workout. Also, P6 said that it would be interesting to see the impact on the JOIN score that is used in the app.

Process This feedback is based on the performances that have been achieved in the process of training. Personal bests were mentioned in Group 2, these are records that users achieved in training (i.e. a 5 min power average of 400w). P11 said that it might not be the first thing to look at, but it might be nice to see after training:

> *"You really don't care much about it. But if you look once and you see this is your best value, then that's nice. That does matter."*

In line with that, a power curve was discussed as a nice-to-have. This curve plots all personal bests into a line, which allows for an all-in-one way of showing achievements. P9 said that showing a curve for the workout and your best past six months would be a great option to have.

Summary Thus, to effectively design post-hoc feedback on how JOIN users performed their suggested workouts, three findings can be highlighted. Overall, experienced and less experienced users agreed on these findings. Firstly, users require feedback regarding the quality of specific intervals. The app should recognize where errors occur and provide actionable feedback. This type of feedback requires more than simply showing time in zone and is mainly interesting for outdoor rides. Secondly, an accuracy score is the most logical way to show how well the planned training was actually performed. The focus of the score should be on the intended effect of the workout. A simple, dynamic score is preferred. Lastly, users would like to receive more information regarding training plan adjustment after a performed workout.

These findings will form the base for the design of the post-hoc feedback feature in the JOIN app discussed in the next chapter. Initial sketches and designs will take inspiration from the focus group results. Making sure the insights of the users are involved in the design process as soon as possible.

5 Design

This chapter describes the iterative process of creating prototypes for the design of the post-hoc feedback feature in the JOIN app. It discusses how the results of the user research are used to develop low-fidelity prototypes that are iteratively adapted to eventually be translated into high-fidelity prototypes. It concludes with making a selection of parts of the high-fidelity prototype that will be implemented in the JOIN app.

5.1 Prototypes

The results of the focus groups offered a base for the design of the prototypes. In this phase, the typical procedure of the company was followed. This process can be described as an iterative presentation of the progress to an increasing number of stakeholders, as depicted in Figure 3.

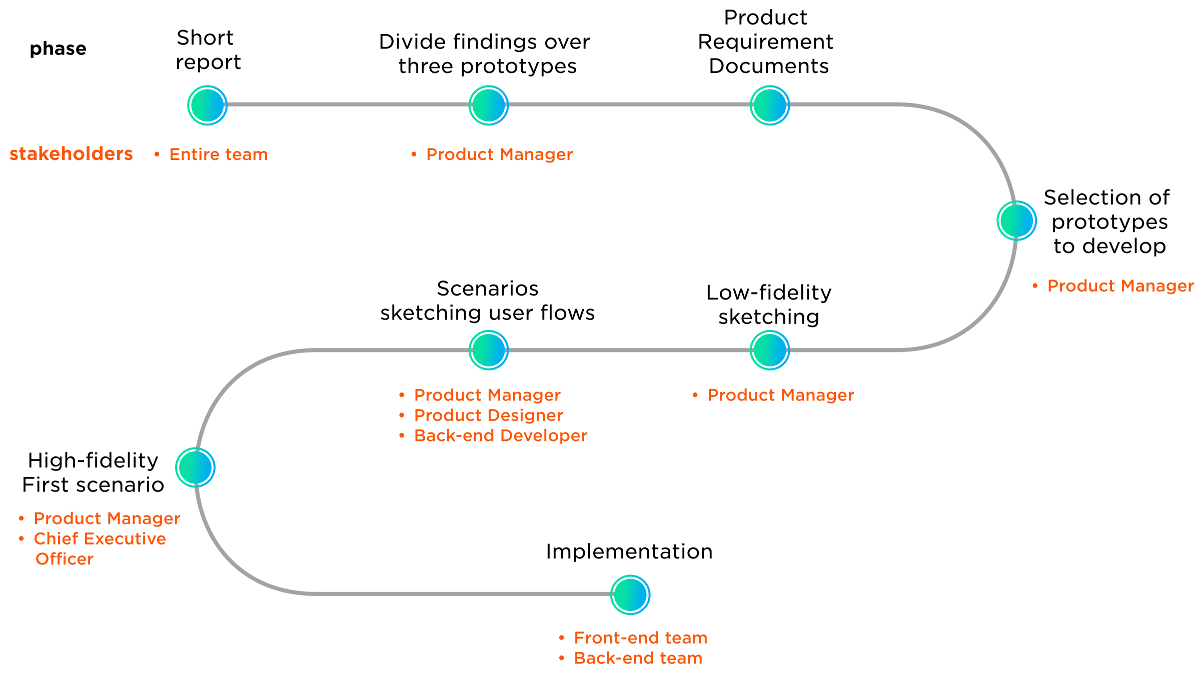


Figure 3: The phases of Iterative Product Development at JOIN with stakeholders per phase. The phases are displayed above every bullet in black. The stakeholders are listed under every phase in orange.

Initially, a short report with key findings of the focus groups was shared with the company. In collaboration with the Product Manager (PM), we found that the findings could best be divided over three separate prototypes:

1. **Quality of intervals**
   This prototype should make users aware of how they have performed their intervals compared to the planned workout. Additionally, users should be able to take insights from the reported quality metrics to improve in the future. Lastly, the app should recognize deviations in performed workouts and present the user with actionable feedback regarding the deviations.

2. **Workout Score**
   This prototype should allow users to see, at a glance, if the intended effect of the planned

workout was met. To add to that, users should be able to interact dynamically with the Workout Score. The overall Workout Score is visible immediately, but the details regarding this score and the workout intervals are displayed in a separate, connected screen.

3. **Influence of the workout**
   This prototype should allow users to see what kind of influence deviations from the planned workout mean for the adjustment of the training plan. Users should also be able to view what the influence of a workout is on their progression.

The next step was to fill in a Product Requirement Document for each prototype (see Appendix B.1). These documents acted as artifacts that displayed the progress on the prototype. After discussion with the PM, we concluded that the Quality of intervals would be put on hold. This prototype would require automatic detection of intervals, which was too complex and time-consuming to implement at this stage.

5.1.1 Low-fidelity

To start visualising the first ideas for the Workout Score and the Influence of the workout, some sketches were made that are shown in Figure 4 and Figure 5. I chose to make use of some of the designs that were already present in JOIN. This might help with understandability, due to users being accustomed to that way of presenting information. In these sketches, I altered them to fit the purpose of both prototypes.

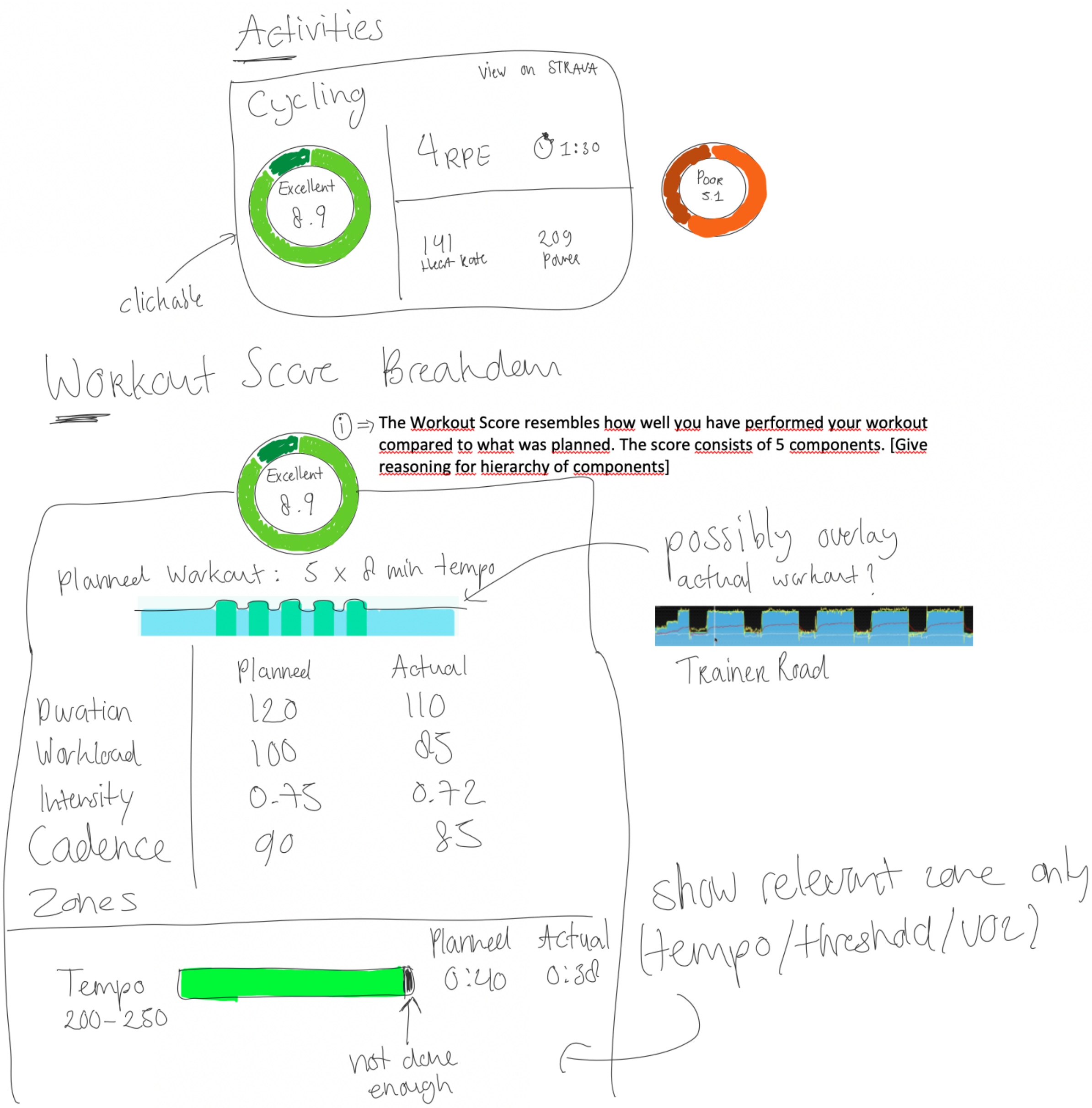


Figure 4: First sketches made for the Workout Score. The top sketch shows an adapted version of the Activity screen that users see when they have completed their workout. Tapping on the circle would direct the user to a separate screen for the Workout Score Breakdown (bottom sketch). This screen shows the planned workout and a planned/actual comparison table for the components of the Workout Score.

Workout Score In Figure 4, the top sketch shows an adapted version of the Activity screen that users see when they have completed their workout. The score is displayed in a circle design, which is also used for other features in JOIN. Tapping on the circle would direct the user to a separate screen for the Workout Score Breakdown (bottom sketch). The Workout Score is shown at the top, followed by the planned workout of that day. This planned workout could possibly be overlaid with the actual workout, similarly to how TrainerRoad [53] has done this. A table is shown comparing the planned and actual values of the components of the Workout Score. The time spent in zones is displayed in a bar, where the differentiation in colour displays that the user has not spent enough time in the zone. Only the relevant zone for the planned workout is shown. The information symbol next to the Workout Score circle would direct the user to a separate screen where the Workout Score is explained.

Figure 5: First sketches made for the Influence of the workout. The top sketch shows a way of visualizing the influence of a workout on the level that is used in JOIN and the estimated FTP (eFTP). The bottom sketch shows a version where only the level increased.

Influence of the workout In Figure 5, the top sketch shows a way of visualizing the influence of a workout on the level that is used in JOIN and the estimated FTP (eFTP). The bottom sketch shows a version where only the level increased. The green arrows highlighting the increase in level or eFTP are also used in other parts of JOIN that indicate similar increases. These sketches are closely related to the Workout Score Breakdown that is shown in Figure 4 as they would replace the way the circle is currently displayed in the Workout Score Breakdown if there is an influence on the level, eFTP or both.

The sketches were shared with the PM. To further develop the design, we decided to create scenarios for the prototypes. This ensured that we accounted for other user flows, instead of focusing solely on one particular user flow. The scenarios integrated both the Workout Score and the Influence of the workout. Fundamentally, there are three different ways a user can complete a workout in JOIN.

1. **Suggested workout**
   An user can complete a workout that has been suggested by JOIN. The user can complete the suggested workout as planned or not as planned. When the user has not completed the suggested workout as planned, they have either exceeded or subceeded the prescribed workout intensity of the suggested workout.

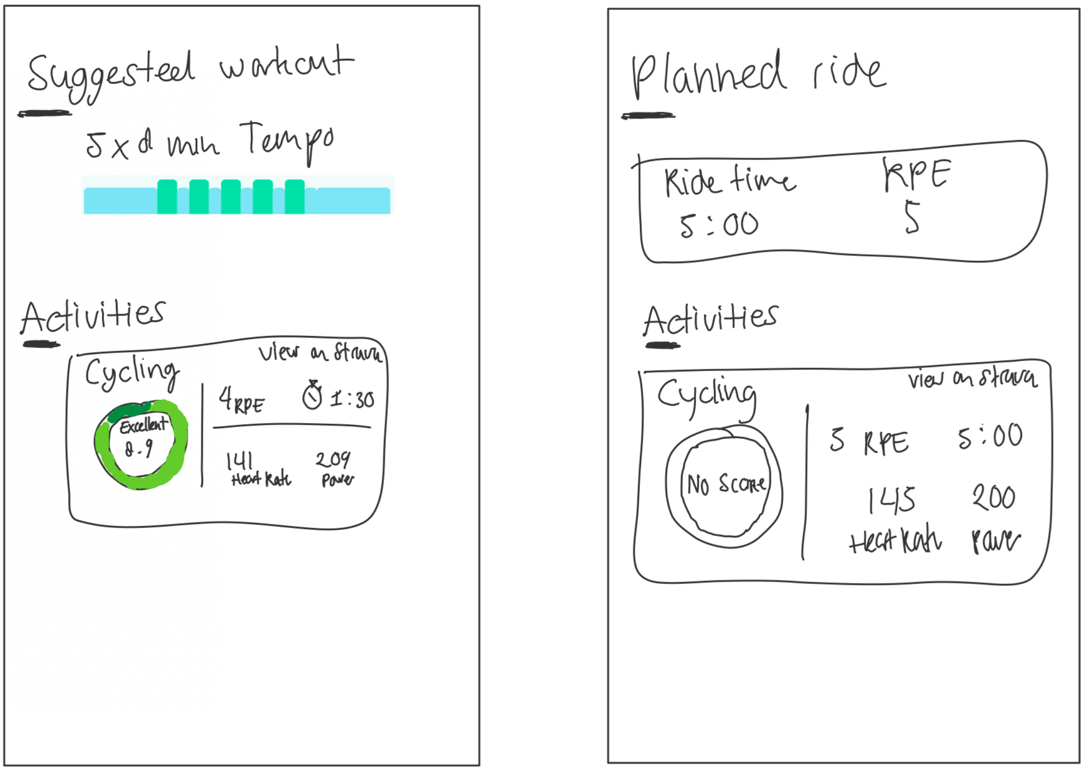
2. **Something else**
   An user can do something completely else than was suggested by JOIN. When the user has done something else, they have either exceeded or subceeded the prescribed workout intensity of the initially suggested workout.

3. **Planned ride**
   An user can plan a ride on their own. In the upcoming seven days, a user can plan a ride that they want to do. The user can set the estimated duration and RPE, alongside an importance factor ranging from C-A that determines the way JOIN will make it fit in your training plan. The user can complete the planned ride as planned or not as planned.

When the user has not completed the planned ride as planned, they have either exceeded or subceeded the estimated duration and/or RPE of the planned ride.

To clarify, the difference between completing a suggested workout not as planned and doing something else is that after loading the workout in the app, the user has to select if they, in their eyes, have completed the suggested workout, whereas doing something else is determined by selecting the 'Did something else' option after loading the workout.

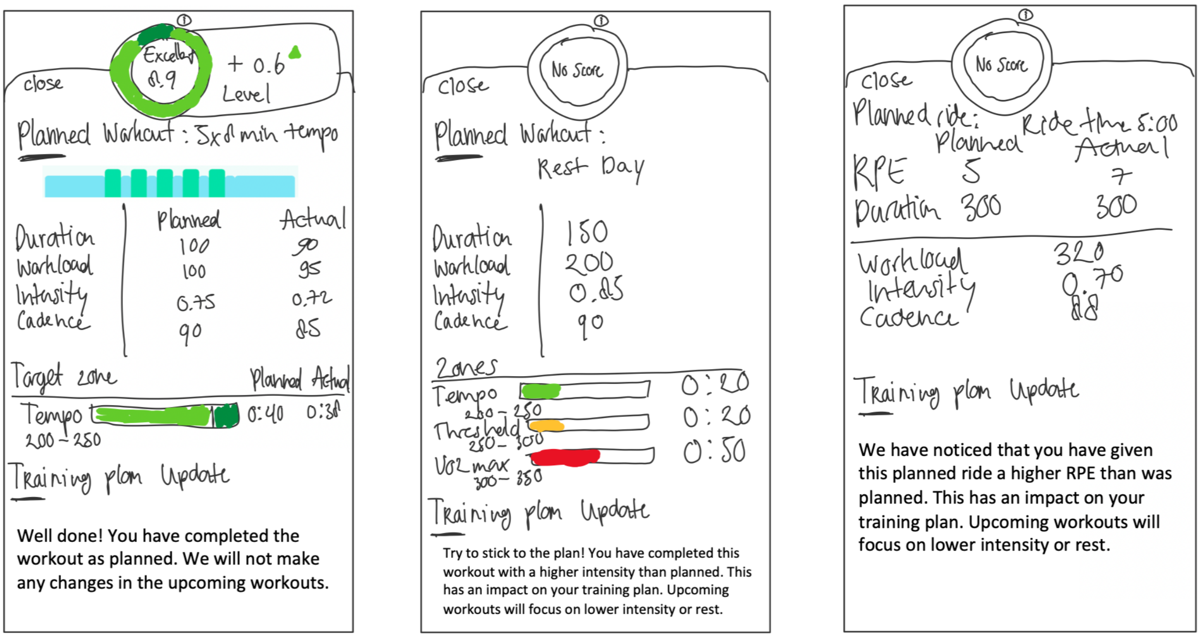
These three ways and their variations give us eight scenarios. A sketch was made for every scenario. Differences in the sketches will be discussed below (see the Product Requirement Documents in Appendix B.1 for an overview of all sketches). The sketch of each scenario contains three app screens: the Activity screen that users see when they have completed their workout, the Workout Score Breakdown and an information screen with the Workout Score Explanation.



(a) (b)

Figure 6: Examples of the Activity screen taken from the complete user flow sketches for each scenario.

Activity screen The Activity screen shows the Workout Score of the loaded workout, depending on the scenario a different score is given. It was decided that a Workout Score would only be given when the user has completed a workout that was suggested by the app. Therefore, the scenarios where the user has done something else or has planned a ride, a 'No score' circle is shown (bottom of Figure 6b). To add to that, when the user has completed a workout that has been suggested by JOIN or when the user has done something else, the suggested workout is shown (top of Figure 6a), whereas when the user has planned a ride (top of Figure 6b) the estimated duration and RPE of that planned ride is shown.



(a) (b) (c)

Figure 7: Examples of the Workout Score Breakdown taken from the complete user flow sketches for each scenario.

Workout Score Breakdown The Workout Score Breakdown is similar to the first sketch as displayed in Figure 4 but with a few modifications. The Workout Score is shown at the top of the screen, with the addition of the influence on level and/or eFTP (Figure 7a) if applicable for the scenario. Additionally, when the user has done something else, the planned/actual table only shows the actual values for every component, as no comparison is made with planned values to create a Workout Score (Figure 7b). When the user has completed a planned ride, RPE is added to the planned/actual table (Figure 7c). Again, only the relevant zones are shown for suggested workout. However, when an user does something else that exceeds the prescribed workout intensity of the initially suggested workout, the higher zone is also shown (Figure 7b). This is done to show the user how much they exceeded the intensity in terms of time in zone. In all other scenarios, the zone information is left out. Lastly, a new section is added to incorporate the Influence of the workout. Under Training plan Update (bottom of examples in Figure 7), a text is shown that is unique for every scenario describing this influence (i.e. higher intensity than planned will have a text describing that upcoming workouts will focus on lower intensity or rest, see Figure 7b).

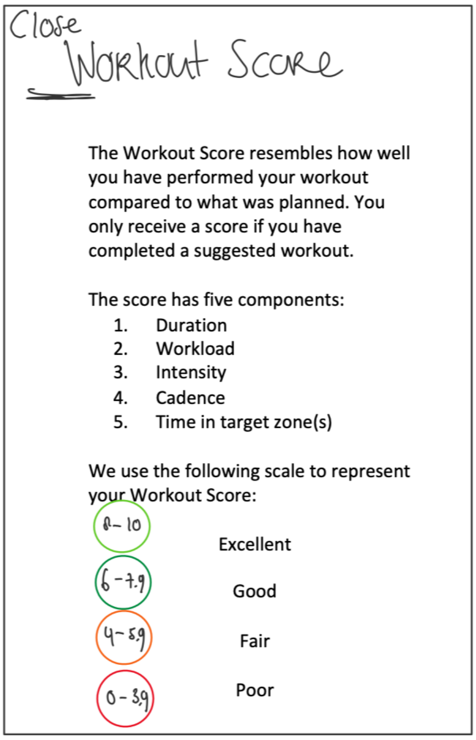


Figure 8: The Workout Score Explanation, similar for every scenario.

Workout Score Explanation The Workout Score Explanation is the same for every scenario (Figure 8). The workings of the score are shortly described in a small summary. Further, it is made clear what the components of the score are. Lastly, a scale is displayed. For every range (Poor, Fair, Good, Excellent), a small explanation should be added to show the user what the scale range means according to JOIN.

When the sketches for every scenarios were done, the Product Designer (PD) was included into the discussion. After discussions with the PM and PD, we agreed on the current status of the user flows and scenarios. Later, a Back-end Developer (BD) joined the procedure to give an opinion and insights into what's possible on the back-end. It became clear that Workload and Intensity should have a different display, instead of numbers. Under the hood, these components are formed in a complex way that a plain number would not resemble well enough. Importantly, the influence of the workout needed to be reconsidered. Instead of showing influence on the training plan, the focus should be on what it does with fatigue and fitness long term. Implementing the proposal above was not realistic as the algorithm takes into account many facets in adapting the training plan. It is not as simple as saying that the app will recommend a higher intensity training after a low intensity day.

5.1.2 High-fidelity

The next step was to create a high-fidelity prototype using Figma [11] for the first scenario, when a user has completed a suggested workout as planned. Creating this prototype comprised of translating the low-fidelity sketches into a design that used the current style of JOIN. Here, we will highlight components of the first version of the high-fidelity prototype and discuss them individually (see Appendix B.2 for overview).

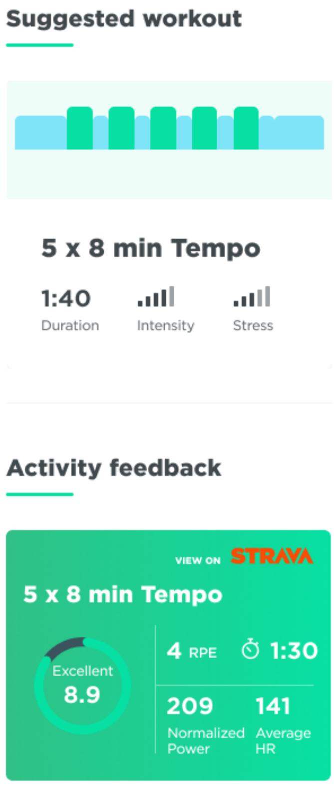
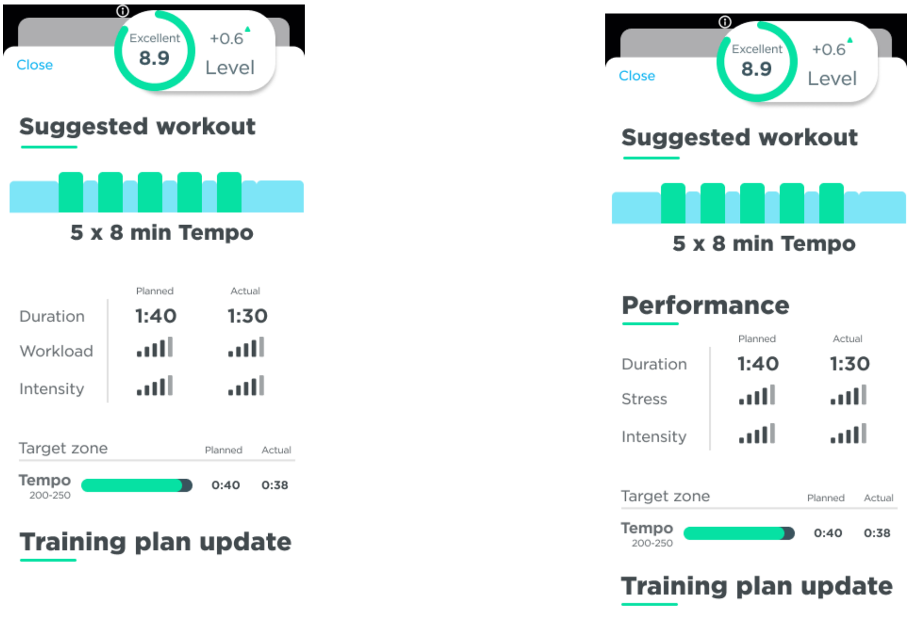


Figure 9: The high-fidelity Activity screen.

Activity screen The Activity screen (Figure 9) was upgraded to higher fidelity where components of the current Activity screen in JOIN were used. No changes were made compared to the lower-fidelity version of the Activity screen (Figure 6).



(a) (b)

Figure 10: The two versions of the high-fidelity Workout Score Breakdown. The left figure (a) shows the version with Workload component. The right figure (b) shows the version with Performance header and Stress component.

Workout Score Breakdown Next to upgrading to a higher fidelity, the Workout Score Breakdown (Figure 10) received some changes compared to the lower-fidelity version. As previously discussed, it became clear that Workload and Intensity should have a different display, instead of numbers. Hence, we introduced a bar display for both (Figure 10). This type of display is also used throughout the app to show Intensity and Stress (i.e. top of Figure 9). At this point, it was not clear if the Workload component was equivalent to the Stress component that is already used throughout the app. Therefore, two versions were created to show either Workload (Figure 10a) or Stress (Figure 10b). Additionally, one of the versions added a small variation that included a Performance header, to potentially increase understandability (Figure 10b). To add to that, the Cadence component was removed from the table of components. At this stage, Cadence would not be a part of the Workout Score. Lastly, the Training plan Update section remained blank (bottom of Figure 10), as at that time we were still looking into how to incorporate a text that explains what the workout does with fatigue and fitness long term. It was later decided to solely focus on the Workout Score and to postpone the Influence of the workout. This prototype would be developed outside of this thesis.

Figure 11: High-fidelity Workout Score circle variations.

In Figure 11, different variants of the Workout Score circle are displayed. We experimented with different colourways.

Figure 12: The high-fidelity Workout Score Explanation

Workout Score Explanation The Workout Score Explanation (Figure 12) was upgraded to higher fidelity where the Workout Score circle was reused to show every scale. Also, some temporary text was added to make room for future explanations the scale. Lastly, the Cadence component was removed from the list of components.

Figure 13: Adaption of high fidelity prototype for Completed as planned scenario: Zone information now displaying all zones with a split on the target zone(s) and the other zones.

The full prototype was discussed with PM and the Chief Executive Officer (CEO). A remark was made regarding the zone information. It would possibly be better to show all zones instead of only the target zone of the specific workout. Using this feedback, the prototype was adapted to show all zones as displayed in Figure 13. Additionally, cadence would be added as a component of the score in a later stage. Lastly, the scale should be changed to a five-point scale as this type of scale is used throughout JOIN.

In a similar fashion, high-fidelity prototypes were made for scenario 2 and 3 (see Appendix B.2). The prototype for scenario 1 was then divided into elements which allowed the team to discuss each element in more detail (see Appendix B.2).

5.2 Implementation

After meeting with the front-end development team, a first iteration was chosen to implement and evaluate. This includes the Activity screen with the Workout Score and the Workout Score Explanation. Tapping on the Workout Score shows the explanation window. To fit the time frame of this thesis, the Workout Score Breakdown would be implemented in a second iteration and evaluated outside of this thesis.

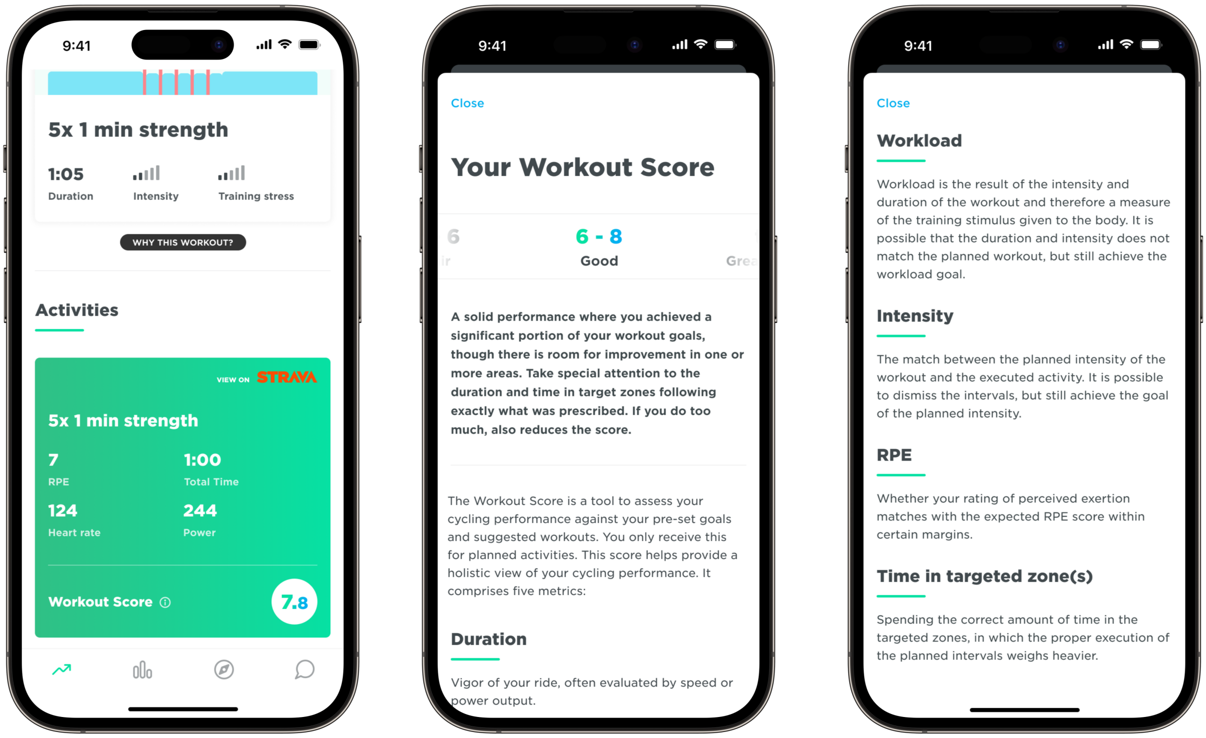
Two changes were made for the implementation. Workload was used as a component of the Workout Score instead of Stress. To add to that, the scale was changed from a four-point scale to a five-point scale. Explanations for each range were added by the PM. The implementation uses the following scale:

- **Perfect Execution (9-10)**: You've done the workout exactly like planned. Great job!
- **Great Job (8-9)**: You've done very well on most criteria of the workout with minor details for improvement.
- **Good (6-8)**: A solid performance in which you've done a big part of the goal of the workout. Take some attention to the intensity and time in zones and try to follow even more precise what was prescribed. Doing more is not always better and can also reduce the score.

- **Fair (3-6)**: You've managed to more or less train in the spirit of the workout, but there are several areas that need significant improvement for a better score. You managed to meet some of the workout goals, but there are several areas that need significant improvement for a better score.

- **Poor (0-3)**: The training done did not meet the goals of the planned training.

To incorporate these changes and to finalize the design, the PD made some final adjustments to the high-fidelity prototype to form the implementation (see Figure 14).



(a) (b) (c)

Figure 14: The implementation including the Activity screen and Workout Score Explanation.

Activity screen In the Activity screen (Figure 14a), the display of the Workout Score was altered to ensure visibility throughout the colour scheme of the app. This was done by using a white circle as a background for the score. Also, the Workout Score was moved to the bottom of the Activity card and the colour of the score itself was changed to a gradient of green to blue.

Workout Score Explanation In the Workout Score Explanation (Figures 14b,14c), instead of showing all score ranges at once, the personal score of the user was shown first. This score was accompanied with an explanation of that score range. The user could swipe to see the other explanations of each score range. The page showed a short text explaining what the Workout Score entailed and when the user swiped up, the five components of the score are shown with for each component a concise explanation.

After iterative development, an implementation was chosen for release to the public. To evaluate the reception of the feature, we used a survey. The process of evaluation is described in the next chapter.

6 Evaluation

To assess the implementation of the Workout Score, we conducted a survey. This chapter describes the approach that was used for the recruitment of participants, the content of the survey and the results of the survey.

6.1 Participants

Participants were recruited via an email blog post that included a link to a Typeform [56] survey. This blog post was sent after the Workout Score was released to all users in the JOIN app. A follow-up email was sent four days after the initial blog post, to recruit more participants. Two survey versions were distributed to accommodate the language preferences of JOIN app users. Depending on their language settings, users received either an English or Dutch version of the survey. In total 562 people completed the survey. 494 filled in the Dutch version, 68 completed the English version. 537 were male, 25 were female. The participants were aged 18-69 ($M = 41.61, SD = 10.06$). Most participants had been using the app for 3-11 months (39.3%) or 1-2 years (36.3%). The experience level in cycling of the participants was a little more varied with users having set their Activity profile to Amateur (19.8%), Intermediate (38.3%) or Advanced ($27,2\%$). See Figure 15 for the complete distributions.

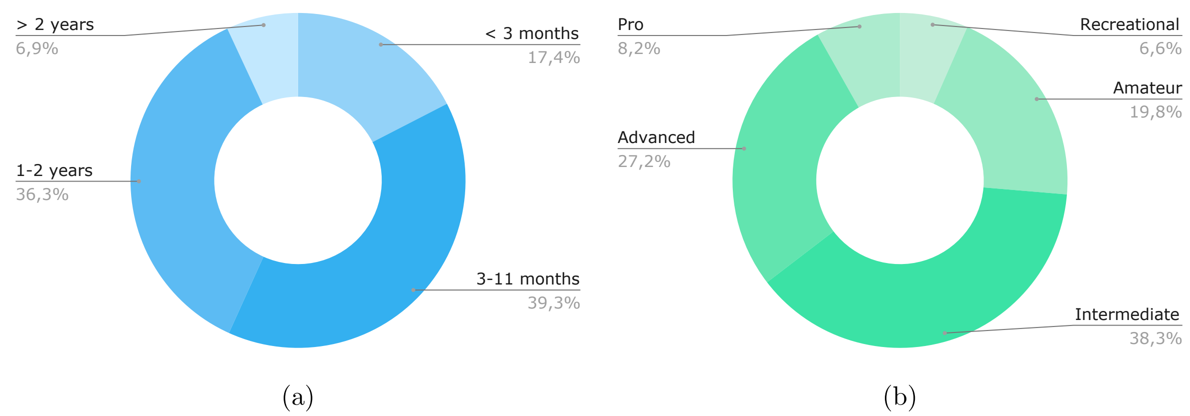


Figure 15: Pie charts displaying the distribution of participants regarding how long they have been using JOIN (a) and their Activity profile in JOIN (b).

6.2 Survey

The purpose of the survey was to evaluate the implementation of the Workout Score. To do so, four key questions (Table 3) were used to form a foundation for the content of the survey.

Table 3: The key questions and the related five-point Likert scale statements per question.

	Key question	**Statement**
1.	Does the Workout Score reflect planned vs. actual well enough?	1
2.	Is the Workout Score easy to understand?	2, 12, 13
3.	Is the Workout Score easy to use?	6, 11
4.	Does the Workout Score add value to the current app?	3-5, 7-10, 14-16

The survey introduced the participants to the research and a consent form was required to be filled in order to continue. The Ethics and Privacy Quick Scan of the Utrecht University

Research Institute of Information and Computing Sciences was conducted (see Appendix D). It classified this research as low-risk with no fuller ethics review or privacy assessment required. Participants were asked for their e-mail, age and gender. Also, participants had to indicate how long they had been using JOIN and select what their Activity profile was. Participants were then asked to indicate to what extent they agreed with 16 five-point Likert scale statements (Table 4). Lastly, they were asked to answer five open questions (Table 5).

Table 4: The 16 five-point Likert scale statements.

	Statement
1.	The Workout Score accurately reflects how well I have completed my suggested workout.
2.	I find the Workout Score easy to understand.
3.	The Workout Score helps me be more effective with my training.
4.	I would look at the Workout Score every time I have completed a suggested workout.
5.	I think the Workout Score is a valuable addition to the JOIN app.
6.	I feel that the way the Workout Score is visually presented in the app makes sense.
7.	The Workout Score provides sufficient information to understand my workout performance.
8.	I will be using the Workout Score in the future.
9.	The Workout Score helps me identify areas of improvement in my workout performance.
10.	The Workout Score helps me stay on track with my training plan.
11.	I think the layout of the explanation page is well-organised.
12.	In my opinion, the Workout Score could benefit from additional information or context about how it's calculated.
13.	I think it is clear how the Workout Score works.
14.	The Workout Score provides me with a sense of accomplishment and fulfilment after completing a workout.
15.	I feel motivated to improve my Workout Score and achieve higher accuracy in my future workouts.
16.	The Workout Score has increased my accountability and commitment to completing my suggested workouts.

Table 5: The five open questions.

	Open question
1.	What did you like the most about the Workout Score?
2.	What do you think should be improved about the Workout Score?
3.	Could you shortly describe how you think the Workout Score works?
4.	Please describe your overall experience with the Workout Score.
5.	If you have anything else to add, please share it here.

6.3 Results

To discuss the results of the evaluation, we will utilize the four key questions outlined in section 6.2. We will describe the findings for each key question by discussing the insights of the group of Likert-scale statements that is linked to the question. These insights are displayed in a figure per group. See Appendix C for a figure with all statements and a table with the mean and standard deviation of each Likert-scale statement. Additionally, the open question answers are used to give context to some of these insights. We used quotes from the open question answers. If needed, quotes were translated to English by me. These quotes have been translated as accurately as possible, yet in some quotes grammatical adjustments have been made to ensure readability.

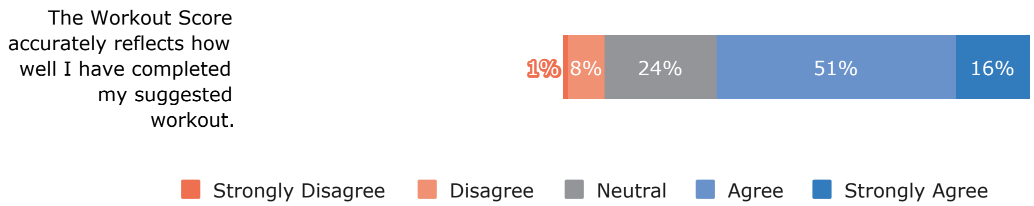


Figure 16: A stacked bar representation of the distribution of the five-point Likert-scale statement agreement related to the first key question.

Does the Workout Score reflect planned vs. actual well enough? As Figure 16 shows, most participants agree (51% Agree, 16% Strongly Agree) that the Workout Score accurately reflects how well they have completed the suggested workout. The answers to the open questions generally support this observation. A participant mentioned the following when asked what they liked the most about the Workout Score:

> *"That JOIN as a digital coach succeeds in providing feedback on the performed training."*

Furthermore, participants liked the fact that the score allows them to see at a glance how the training went. Lastly, the score gives an useful insight in the execution of the training in a simple and clear manner.

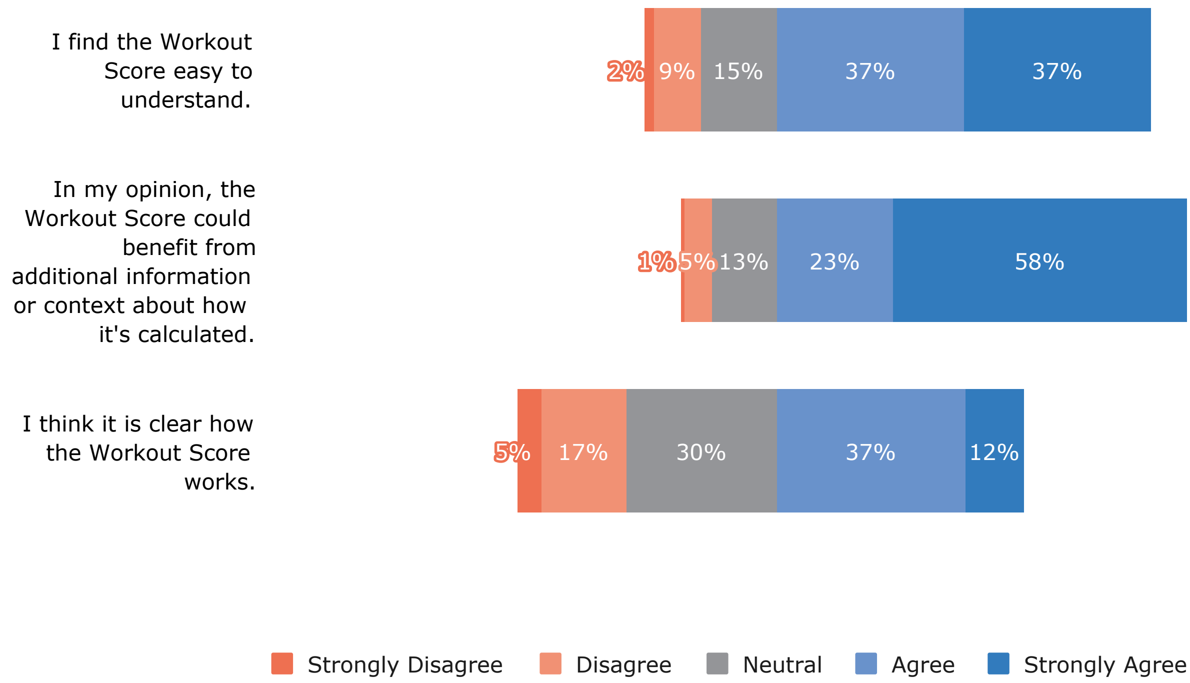


Figure 17: A stacked bar representation of the distribution of the five-point Likert-scale statement agreements related to the second key question.

Is the Workout Score easy to understand? As seen in Figure 17, when asked if they found the Workout Score easy to understand the majority of participants agreed (37% Agree, 37% Strongly Agree). To add to that, the workings of the Workout Score seemed clear to people (37% Agree, 12% Strongly Agree). However, a great part remained neutral regarding this statement and some disagreed. Importantly, the vast majority agreed (23% Agree, 58% Strongly Agree) that the Workout Score could benefit from additional information or context about how it is calculated. This is also reflected in the answers to some of the open questions. When asked what should be improved about the Workout Score, the main insight that can be highlighted is that participants wanted more explanation about what went well or did not go as planned. The explanation should be made more specific in order to clearly indicate where users need to improve to increase their score. It would help if separate scores are shown for the components of the Workout Score. A participant said the following:

> *"Break down the score per criteria. For example, if I score a 9, I would like to know if I lost points (from a 10) in the intensity, duration, RPE, etc. Knowing which aspect of the score needs improvement would help improve the training as a whole."*

Additionally, when asked directly how they thought the Workout Score worked, participants generally described the workings of the score quite well. They included the five components of the feature in their answers and often they referred to the Workout Score Explanation page:

> *"That is actually explained in the description in the app. It looks at duration, training load, intensity, RPE and training in set zones. A kind of score and weighting will probably be attached to this, so that a total score emerges."*

However, there were some that were not able to give a description of how the score worked. To add to that, participants assumed that the app detected the intervals automatically. Though, as mentioned in section 5.1, this is currently not in the implementation.

To add to that, we noticed that a few participants did not understand why the RPE was part of the Workout Score. According to them, RPE is something that relates more to the influence of additional factors like a stressful day at work or how well they have slept or eaten. It should not have an effect on the execution of a training if they experience a training as more or less hard.

> *"I can perform a training session almost to the desired standard, and on a bad day, it feels more challenging. However, I believe that doesn't necessarily mean I did it less well."*

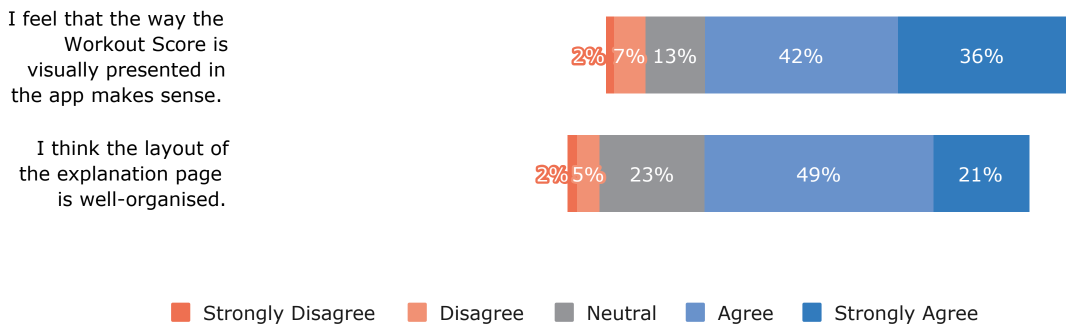
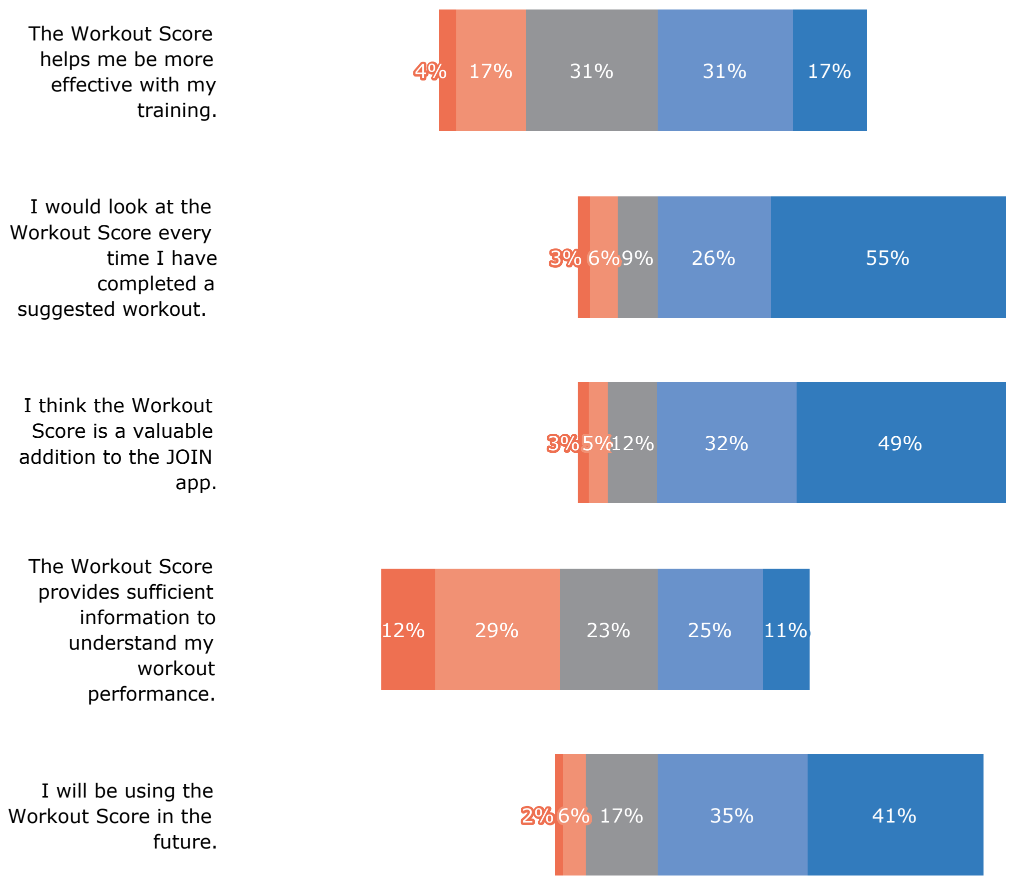


Figure 18: A stacked bar representation of the distribution of the five-point Likert-scale statement agreements related to the third key question.

Is the Workout Score easy to use? The distribution in Figure 18 shows that the way the Workout Score was visually presented in the app made sense for most participants (42% Agree, 36% Strongly Agree). To add to that, the layout of the explanation page was well-organised according to the majority of participants (49% Agree, 21% Strongly Agree).



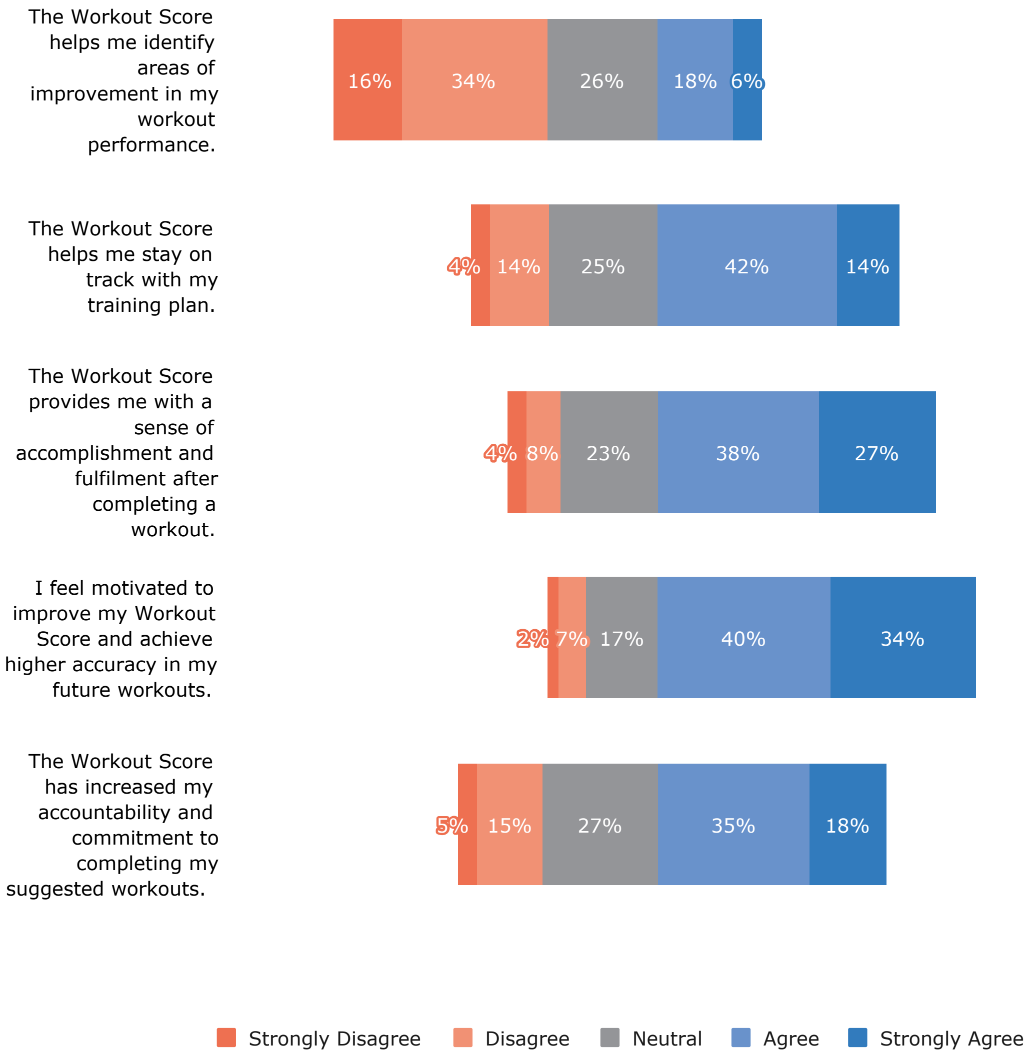


Figure 19: A stacked bar representation of the distribution of the five-point Likert-scale statement agreements related to the fourth key question.

Does the Workout Score add value to the current app? Figure 19 shows that most participants (49% Strongly Agree) saw the Workout Score as a valuable addition to the JOIN app. To add to that, the majority of the participants (55% Strongly Agree) indicated that they would look at the Workout Score every time they have completed a suggested workout. However, the Workout Score did not help enough to identify areas of improvement (34% Disagree, 16% Strongly Disagree). Also, it could provide more information to help users understand their performance (29% Disagree, 12% Strongly Disagree). This is supported by the open question answers. As mentioned earlier, the explanation should be further elaborated to clearly indicate the areas where users need to improve in order to increase their score.

As seen in Figure 19, most participants (40% Agree, 34% Strongly Agree) feel motivated to improve their Workout Score and achieve higher accuracy in future workouts. The same sentiment is evident in some answers to the open questions. Many participants felt that the Workout Score motivated them to follow the prescribed training. For example, a participant said:

> *"It is motivating to execute intervals down to the smallest detail and to fine-tune routes even more precisely to fit the training."*

Also, they mentioned that getting a high score motivated them. However, some participants also noted that the Workout Score can be quite demotivating. Especially when participants got a low score, while it was unclear what they had done wrong. Additionally, participants felt demotivated when they had just completed a fun, enjoyable and challenging ride where they might have ridden a little too hard or long and then got back with a low Workout Score. However, others used the Workout Score to stick to the prescribed plan instead of going too hard:

> *"When I feel good, I often deviate from the prescribed power targets. Knowing that deviating affects my score, I make a greater effort to adhere to the specified power levels."*

7 Discussion

In this chapter we describe the scientific contributions of this work, as well as its limitations, concluding with recommendations for future work.

7.1 Contributions

This thesis has contributed in two ways. First of all, we provide an overview of related work to present a foundation that is required to answer our research question. The overview showed that cycling as a sport and what's involved is so multi-faceted that it allows for the rise of technologies from different areas such as training, adventure, sleep, recovery and nutrition. To add to that, cycling is an easily quantifiable sport [28]. Many sensors exist that allow cyclists to see different metrics in real-time. Moreover, adaptive training applications offer an easy way for cyclists to start structured training. They take into account personal characteristics and try to imitate the process of a real-life sports coach. Furthermore, feedback is a crucial factor influencing skill learning and performance [17] and it can originate from intrinsic and extrinsic sources, supplemented by real-time and post-hoc feedback. Research regarding post-hoc feedback in cycling is quite limited. However, there exist some commercial applications that provide a way of post-hoc feedback in the form of data visualisation or analysis. Lastly, personalized feedback should be specific, precise, and consider the athlete's skill level.

Second, this thesis has contributed by using a scientific based approach to explore a way of providing effective post-hoc feedback on training performed in the cycling application JOIN. Other adaptive training apps have tried to involve feedback in the form of data analysis and visualisation. However, actual scientific basis for the (design) decisions taken by these commercial applications was missing. We have conducted focus groups with JOIN users to gather insights that were used to find out what kind of post-hoc feedback users of different levels of expertise want on training performed. Both experienced and less experienced users agreed on wanting feedback regarding the quality of specific intervals. To add to that, an accuracy score seemed to be the most logical way to show how well the planned training was performed. Lastly, they would like to receive more information regarding the way the training plan is adjust after a performed workout. Using the focus group insights, we iteratively developed low-fidelity and high-fidelity prototypes. An implementation of the Workout Score was chosen to be evaluated with a survey to assess how much users appreciate the post-hoc feedback on training performed. The Workout Score seemed to reflect the planned versus actual training well enough. Participants indicated that they could see how they have performed the training thanks to the useful, simple and clear score. Furthermore, the majority of participants thought that the Workout Score was easy to understand. However, the explanation should be made more specific in order to clearly indicate where users need to improve to increase their score. This could be done by showing separate scores for each component of the Workout Score. Additionally, the inclusion of RPE in the Score was unclear to some. To add to that, the Likert-scale statement results showed that the Workout Score was easy to use for participants. Lastly, most participants saw the Workout Score as a valuable addition to the JOIN app. However, more information and help could be provided to understand the workout performance and to identify areas of improvement. The Workout Score is an addition that can motivate participants to improve their Workout Score, achieve higher accuracy in future workouts and stick to the prescribed plan. Though, it can also demotivate users. Especially when it was unclear what they had done wrong if they received a low Workout Score. But also when they had an enjoyable ride where they might have exceeded some factors of the suggested workout and then got back with a low

Workout Score.

7.2 Limitations

There are several limitations to this research, which we address in this section. Firstly, a limitation of this thesis is that the Quality of specific intervals prototype was put on hold early on in the iterative design process. After the division of the user research findings over three prototypes, it became clear that this prototype would require automatic detection of intervals, which was too complex and time-consuming to implement. As a consequence, this thesis does not fully meet the users' requests for specific feedback on the individual interval performance.

Additionally, eventually postponing the Influence of the workout prototype forms a limitation. In section 5.1.1, it became evident that the prototype needed to be reconsidered. Instead of showing influence on the training plan, the focus should be on what it does with fatigue and fitness long term. Implementing the low-fidelity proposal was not realistic as the algorithm takes into account many facets in adapting the training plan. It was not as simple as saying that the app would recommend a higher intensity training after a low intensity day. As discussed in section 5.1.2, we were still unsure how to exactly incorporate a text that explains what the workout does with fatigue and fitness long term. At that moment, we decided to postpone the prototype entirely. As a result, this thesis does not fully address the users' requests for feedback on the influence of the workout.

Furthermore, the fact that the email blog post was sent out quickly after the release of the implementation of the Workout Score can be seen as a limitation. As a consequence, several participants indicated that they did not have enough experience with the feature to describe their experience with the Workout Score. Ideally, we would have allowed the users to get acquainted with the feature for a longer period. However, to ensure that the implementation could be evaluated within the time frame of this thesis, we decided to sent out the survey quickly after release.

Lastly, a limitation of this thesis is the fact that the implementation did not include the Workout Score Breakdown. As said in section 5.2, to implement the first iteration within the time frame of this thesis, the Workout Score Breakdown was left out. As a consequence, the implementation did not include a more detailed analysis of the completed workout, which could have changed the users' reception of the feature.

7.3 Future work

Several future work possibilities have emerged from this thesis. First of all, there are some areas of improvement for future updates to the Workout Score according to the results of the survey. The findings indicated that the explanation should be made more specific in order to clearly indicate where users need to improve to increase their score. This could be done by showing separate scores for each component of the Workout Score. This would also help reduce the chance of demotivating users with lower Workout Score. Furthermore, the inclusion of RPE in the Score was unclear to some. We suggest that it might help to reconsider the role of RPE or to provide a more complete explanation of the reasoning for its inclusion. Additionally, participants felt demotivated when they had just completed a fun, enjoyable and challenging ride where they might have ridden a little too hard or long and then got back with a low Workout Score. We propose to make it more evident to users when they should select that they have completed the suggested workout and when it would instead be better to select that they have done something else. In the future, JOIN could explore if the algorithm could potentially take over this decision-making process.

Second, the user research conducted in this thesis can be used to further design the two prototypes that were not part of the implementation. The results that are presented in section 4.1.3 form a base for potential future designs based on real users.

Lastly, the Workout Score Breakdown was not implemented. This part will be implemented by JOIN in a second iteration of the Workout Score feature and evaluated separately from this thesis.

8 Conclusion

The aim of this work was to explore how effective post-hoc feedback on training performed can be provided in a cycling application. After setting a foundation with related literature, we conducted focus groups in collaboration with the mobile adaptive training app JOIN to find out what kind of post-hoc feedback users of different levels of expertise want on training performed. Both experienced and less experienced users agreed on wanting feedback regarding the quality of specific intervals. An accuracy score seemed to be the most logical way to show how well the planned training was performed. Users would also like to receive more information regarding the way the training plan is adjusted after a performed workout. Using the focus group insights, we iteratively developed low-fidelity and high-fidelity prototypes. An implementation was evaluated with a survey to assess how much users appreciate the post-hoc feedback on training performed. The survey showed that participants appreciated the Workout Score, finding it a simple and clear way to show the planned versus actual training. Yet, the explanation lacked specificity to better show where users need to improve to increase their score. Furthermore, the inclusion of RPE in the Score was unclear to some. Additionally, the Workout Score was easy to use and a valuable addition to the app according to the participants. The Workout Score can motivate to stick to the suggested workout but may also demotivate when users receive a low score without actionable feedback. To improve the post-hoc feedback feature in the future, separate scores for each component of the Workout Score should be shown. To add to that, we suggest to reconsider the role of RPE or to provide a more comprehensive explanation for its inclusion. Lastly, we propose to make it more evident to users when they should select that they have completed the suggested workout and when it would instead be better to select that they have done something else. In the future, JOIN could explore if the algorithm could potentially take over this decision-making process.

References

[1] Mark J. Berentsen, Marit Bentvelzen, and Paweł W. Woźniak. “MTBalance: Assisting Novice Mountain Bikers with Real-Time Proprioceptive Feedback”. In: *Proceedings of the ACM on Human-Computer Interaction* 5.ISS (Nov. 2021), p. 506. ISSN: 25730142. DOI: 10.1145/3488551. URL: https://dl.acm.org/doi/10.1145/3488551.

[2] BKOOL. *BKOOL, The most realistic Indoor Cycling Simulator*. URL: https://www.bkool.com/en.

[3] Jordi Bruin. *MacWhisper*. URL: https://goodsnooze.gumroad.com/l/macwhisper.

[4] Ralph Buehler and John Pucher. “COVID-19 Impacts on Cycling, 2019–2020”. In: *https://doi-org.proxy.library.uu.nl/10.1080/01441647.2021.1914900* 41.4 (2021), pp. 393–400. ISSN: 14645327. DOI: 10.1080/01441647.2021.1914900. URL: https://www-tandfonline-com.proxy.library.uu.nl/doi/abs/10.1080/01441647.2021.1914900.

[5] Hannah Bussey. *Best cycling computers — Cycling Weekly*. URL: https://www.cyclingweekly.com/group-tests/cycling-gps-units-buyers-guide-181254.

[6] Mark Connor, Marco Beato, and Michael O’Neill. “Adaptive Athlete Training Plan Generation: An intelligent control systems approach”. In: *Journal of Science and Medicine in Sport* 25.4 (Apr. 2022), pp. 351–355. ISSN: 1440-2440. DOI: 10.1016/J.JSAMS.2021.10.011.

[7] Mathias De Brouwer et al. “Personalized Real-Time Monitoring of Amateur Cyclists on Low-End Devices: Proof-of-Concept & Performance Evaluation”. In: (2018). DOI: 10.1145/3184558.3191648. URL: https://doi.org/10.1145/3184558.3191648.

[8] *Discord — Your Place to Talk and Hang Out*. URL: https://discord.com/.

[9] EatMyRide. *Cycling Food Plan? Get Personalized Plans — EatMyRide*. URL: https://eatmyride.com/.

[10] *enduco*. URL: https://www.enduco.app/.

[11] Figma. *Figma: the collaborative interface design tool.* URL: https://www.figma.com/.

[12] Iztok Fister. “Generating the training plans based on existing sports activities using swarm intelligence”. In: (2017).

[13] W. T. Gallwey. *The Inner Game of Tennis*. 1996. DOI: 10.1097/00005053-197603000-00015.

[14] Garmin. *Fitness Watches — Sport Watches — Smartwatches — Garmin*. URL: https://www.garmin.com/en-GB/c/wearables-smartwatches/.

[15] Garmin. *Garmin Varia™ RCT715 — Rearview Bike Radar*. URL: https://www.garmin.com/en-GB/p/721258.

[16] *Garmin Connect — Free Online Fitness Community*. URL: https://connect.garmin.com/.

[17] Willis D Hawley and Unda Valli. *The Essentials of Performance Analysis*. 1999. 2012. ISBN: 9780140136661.

[18] Hyperice. *Normatec 3 Legs — Hyperice*. URL: https://hyperice.com/products/normatec-3-legs/.

[19] Intervals.icu. *Intervals.icu*. URL: https://intervals.icu/.

[20] JOIN.cc. *JOIN Cycling App for iOS & Android - Improve Your Ride.* URL: https://join.cc/.

[21] Komoot. *Komoot — Find, plan and share your adventures.* URL: https://www.komoot.com/.

[22] Magnus Kristian, Mathias Vedsø, and Ernst Albin. "Maximal accelerations for twelve weeks elicit improvement in a single out of a collection of cycling performance indicators in trained cyclists". In: (2023). DOI: 10.3389/fspor.2022.1027787. URL: https://doi.org/10.3389/fspor.2022.1027787.

[23] Nattapon Kumyaito, Preecha Yupapin, and Kreangsak Tamee. "Planning a sports training program using Adaptive Particle Swarm Optimization with emphasis on physiological constraints". In: *BMC Research Notes* 11.1 (Jan. 2018), pp. 1–6. ISSN: 17560500. DOI: 10.1186/S13104-017-3120-9/TABLES/2. URL: https://bmcresnotes.biomedcentral.com/articles/10.1186/s13104-017-3120-9.

[24] Arie Willem de Leeuw et al. "Exploiting sensor data in professional road cycling: personalized data-driven approach for frequent fitness monitoring". In: *Data Mining and Knowledge Discovery* (Dec. 2022), pp. 1–29. ISSN: 1573756X. DOI: 10.1007/S10618-022-00905-5/TABLES/3. URL: https://link-springer-com.proxy.library.uu.nl/article/10.1007/s10618-022-00905-5.

[25] Lumivero. *NVivo.* URL: https://lumivero.com/products/nvivo/.

[26] Fabio Malizia and Bert Blocken. "Impact of Disc Wheel Geometry on Aerodynamic Performance: A Computational Fluid Dynamics Investigation". In: *Journal of Fluids Engineering, Transactions of the ASME* 145.1 (Jan. 2023). ISSN: 1528901X. DOI: 10.1115/1.4055698/1146451. URL: https://asmedigitalcollection.asme.org/fluidsengineering/article/145/1/011209/1146451/Impact-of-Disc-Wheel-Geometry-on-Aerodynamic.

[27] Florian Michahelles. "The Eternity Bike - Bringing Active Safety to the Bicycle; The Eternity Bike - Bringing Active Safety to the Bicycle". In: (2021).

[28] Kolie Moore and B S Biology. "A Coaching Perspective on Moder n Training Metrics and Retur n from Injury and Illness". In: (). DOI: 10.1016/j.pmr.2021.08.012. URL: https://doi.org/10.1016/j.pmr.2021.08.012.

[29] Esteban Municio et al. "Continuous Athlete Monitoring in Challenging Cycling Environments Using IoT Technologies". In: *IEEE Internet of Things Journal* 6.6 (Dec. 2019), pp. 10875–10887. ISSN: 23274662. DOI: 10.1109/JIOT.2019.2942761.

[30] Nix. *Hydration Monitor — Wearable Real-time Sensor — Nix Biosensors.* URL: https://nixbiosensors.com/collections/all-products/products/hydration-biosensor.

[31] Notio. *Notio Aerometer.* URL: https://notio.ai/products/notio-ride-tracker.

[32] Ryo Okugawa et al. "Bicycle pedaling training system using auditory feedback". In: *ACM International Conference Proceeding Series* 16-19-Nove (Nov. 2015). DOI: 10.1145/2832932.2832972. URL: https://dl.acm.org/doi/10.1145/2832932.2832972.

[33] Thomas Peeters, Raman Garimella, and Stijn Verwulgen. "An Indoor Training Bike to Provide Real-Time Feedback on the Aerodynamic Cycling Position Using Frontal Area Calculations". In: (2020). DOI: 10.15221/20.24. URL: https://doi.org/10.15221/20.24.

[34] Pillar. *Best Cycling Training App — Bike Apps — Workout Plan.* URL: https://www.pillar-app.com/.

[35] *PKRS AI Coaching for Endurance Athletes, Triathletes, Cyclists and Runners.* URL: https://www.pkrs.ai/.

[36] Polar. *All Polar Watches — Polar Global.* URL: https://www.polar.com/en/all-watches.

[37] Praxisdienst. *Lactaatmeter Lactate Pro 2 — Praxisdienst.* URL: https://www.praxisdienst.nl/nl/Laboratoriumbenodigdheden/Laboratorium+apparatuur/Lactaatmeters/Lactaatmeter+Lactate+Pro+2.html?gclid=Cj0KCQiAxbefBhDfARIsAL4XLRqYrsAX9WNRStW4inXw8sgsC VfR2Iv6I1kT4aCkEM8hYGkgaAj0YEALw_wcB.

[38] Alen Rajšp and Iztok Fister. "A Systematic Literature Review of Intelligent Data Analysis Methods for Smart Sport Training". In: *Applied Sciences 2020, Vol. 10, Page 3013* 10.9 (Apr. 2020), p. 3013. ISSN: 2076-3417. DOI: 10.3390/APP10093013. URL: https://www.mdpi.com/2076-3417/10/9/3013/htm%20https://www.mdpi.com/2076-3417/10/9/3013.

[39] M. M. Reijne, D. J. J. Bregman, and A. L. Schwab. "Measuring and Comparing Descend in Elite Race Cycling with a Perspective on Real-Time Feedback for Improving Individual Performance". In: *Proceedings 2018, Vol. 2, Page 262.* Vol. 2. 6. Multidisciplinary Digital Publishing Institute, Feb. 2018, p. 262. DOI: 10.3390/proceedings2060262. URL: https://www.mdpi.com/2504-3900/2/6/262/htm%20https://www.mdpi.com/2504-3900/2/6/262.

[40] RidewithGPS. *Ride with GPS — Bike Route Planner and Cycling Navigation App.* URL: https://ridewithgps.com/.

[41] ROUVY. *The Indoor Cycling Reality App for Training & Racing — ROUVY.* URL: https://rouvy.com/.

[42] Nina Schaffert et al. "Towards an application of interactive sonification for the forces applied on the pedals during cycling on the Wattbike ergometer". In: *Displays* 50 (Dec. 2017), pp. 41–48. ISSN: 01419382. DOI: 10.1016/j.displa.2017.09.004.

[43] Ørnulf Seippel. *The meanings of sport: Fun, health, beauty or community?* Jan. 2006. DOI: 10.1080/17430430500355790.

[44] Alessandro Silacci, Redha Taiar, and Maurizio Caon. "Towards an AI-based tailored training planification for road cyclists: A case study". In: *Applied Sciences (Switzerland)* 11.1 (2021), pp. 1–18. ISSN: 20763417. DOI: 10.3390/app11010313.

[45] Alessandro Silacci et al. "Designing an e-Coach to Tailor Training Plans for Road Cyclists". In: *Advances in Intelligent Systems and Computing* 1026 (2020), pp. 671–677. ISSN: 21945365. DOI: 10.1007/978-3-030-27928-8{_}102/FIGURES/6. URL: https://link-springer-com.proxy.library.uu.nl/chapter/10.1007/978-3-030-27928-8_102.

[46] *Spoked — AI Cycling Coach.* URL: https://www.spoked.ai/.

[47] Strava. *Strava — Running, Cycling & Hiking App - Train, Track & Share.* URL: https://www.strava.com/.

[48] Mari Suoheimo and Jonna Häkkilä. "Technology Use with Winter Cycling". In: (2021).

[49] Supersapiens. *Perform Stronger and Recover Faster with CGM.* URL: https://www.supersapiens.com/en-EN/.

[50] Silvia Teodorescu and Constanţa Urzeală. "Management Tools in Sports Performance". In: *Procedia - Social and Behavioral Sciences* 81 (2013), pp. 84–88. ISSN: 18770428. DOI: 10.1016/j.sbspro.2013.06.392.

[51] Todaysplan. *Today's Plan: Smarter training for every athlete.* URL: https://www.todaysplan.com.au/.

[52] *Train Smarter — AI Endurance.* URL: https://aiendurance.com/.

[53] TrainerRoad. *Cycling Planning, Training, & Analytics - TrainerRoad.* URL: https://www.trainerroad.com/.

[54] TrainingPeaks. *Glossary of TrainingPeaks Metrics — TrainingPeaks.* URL: https://www.trainingpeaks.com/learn/articles/glossary-of-trainingpeaks-metrics/.

[55] TrainingPeaks. *TrainingPeaks — Train With Confidence.* URL: https://www.trainingpeaks.com/.

[56] Typeform. *Typeform: People-Friendly Forms and Surveys.* URL: https://www.typeform.com/.

[57] Pedro L. Valenzuela et al. "Pre-sleep protein supplementation in professional cyclists during a training camp: a three-arm randomized controlled trial". In: *https://doi-org.proxy.library.uu.nl/10.1* 20.1 (2023), p. 2166366. ISSN: 15502783. DOI: 10.1080/15502783.2023.2166366. URL: https://www-tandfonline-com.proxy.library.uu.nl/doi/abs/10.1080/15502783.2023.2166366.

[58] Velocomp. *AeroPod® V5 – Velocomp LLC.* URL: https://velocomp.com/products/aeropod.

[59] Wahoo. *KICKR Smart Bike — Indoor Cycle Bike — Wahoo Fitness EU.* URL: https://eu.wahoofitness.com/devices/indoor-cycling/smart-bikes/kickr-bike-buy.

[60] *Wahoo ELEMNT Companion on the App Store.* URL: https://apps.apple.com/us/app/wahoo-elemnt-companion/id1049469683.

[61] Wahoo RGT. *Train. Race. Connect - Wahoo RGT is the next generation of virtual cycling.* URL: https://www.rgtcycling.com/.

[62] WHOOP. *WHOOP — Your Personal Digital Fitness and Health Coach.* URL: https://www.whoop.com/.

[63] Martijn C. Willemsen, Mark P. Graus, and Bart P. Knijnenburg. "Understanding the role of latent feature diversification on choice difficulty and satisfaction". In: *User Modelling and User-Adapted Interaction* 26.4 (Oct. 2016), pp. 347–389. ISSN: 15731391. DOI: 10.1007/S11257-016-9178-6/TABLES/10. URL: https://link-springer-com.proxy.library.uu.nl/article/10.1007/s11257-016-9178-6.

[64] Miguel A. Wister, Pablo Pancardo, and Pablo Payro Campos. "A Comparative Study of Cycling Mobile Applications". In: *2018 3rd International Conference on Smart and Sustainable Technologies, SpliTech 2018.* 2018, pp. 1343–1348. ISBN: 9789532900835. URL: https://ieeexplore-ieee-org.proxy.library.uu.nl/stamp/stamp.jsp?tp=&arnumber=8448360%20https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7781053%0Ahttps://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5485276%0Ahttps://ieeexplore.ieee.org/stamp/st.

[65] Robert Wood. *Role of the Coach.* URL: https://www.topendsports.com/fitness/coach.htm.

[66] Matthias Wunsch. "Evaluation of Interactive Systems for Cyclists Using Wizard of Oz Prototypes In The Wild; Evaluation of Interactive Systems for Cyclists Using Wizard of Oz Prototypes In The Wild". In: (2021). URL: https://doi.org/tba..

[67] James Y. Xu et al. "Integrated inertial sensors and mobile computing for real-time cycling performance guidance via pedaling profile classification". In: *IEEE Journal of Biomedical and Health Informatics* 19.2 (Mar. 2015), pp. 440–445. ISSN: 21682208. DOI: 10.1109/JBHI.2014.2322871.

[68] Laila Zahran, Mohammed El-Beltagy, and Mohamed Saleh. "A Conceptual Framework for the Generation of Adaptive Training Plans in Sports Coaching". In: *Advances in Intelligent Systems and Computing* 1058 (2020), pp. 673–684. ISSN: 21945365. DOI: 10.1007/978-3-030-31129-2{_}62/TABLES/2. URL: https://link-springer-com.proxy.library.uu.nl/chapter/10.1007/978-3-030-31129-2_62.

[69] Zwift. *Direct Drive Trainers*. URL: https://support.zwift.com/en_us/direct-drive-trainers-B1oH2meS.

[70] Zwift. *Indoor Cycling & Running Virtual Training App - Zwift*. URL: https://eu.zwift.com/.

[71] Zwift. *Wheel-on Trainers*. URL: https://support.zwift.com/en_us/wheel-on-trainers-r1B5nQWxS.

A User research

A.1 Focus group protocol

Focus Group Protocol- EN

Ethics and Privacy

Participants have been given an Information sheet and have signed the provided Consent form.

Research Introduction

- Introduction of myself (and Adrian?)
    - Hello everyone, first of all thank you for being here and participating in this focus group. I am Remy van Tussenbroek, 22 years old and I am currently doing my Master in Human Computer Interaction at the University of Utrecht. I am doing my thesis in combination with JOIN. Adrian is the Product Manager here at JOIN. Adrian will help me today, his language preference is English so please address him in English.
- Short overview of research
    - In this research, we are investigating how we can provide more insights for JOIN users into how they performed their workouts. At this moment, there is only an option to provide feedback by the user, but feedback from JOIN on how you did is missing.
- Aims of today
    - How you would want to receive feedback from JOIN
    - Find out certain design preferences
    - How much information/graphs should be presented
- We are recording this focus group and we will use this recording for our analysis. When we write about our findings we might use direct quotes from this focus group but we will not identify you individually and any quotes that we do use will be anonymized.
- I want to remind you that we are here to hear your opinion and ideas. Our intention really is to keep the discussion very open so please let us know everything you think of!
- Questions?

Participant Introduction (ice-breaker)

- Write your name on a folded sheet of paper
- Short introduction of participants
    - Name
    - When they started using JOIN
    - Cycling background

Session

- How do you currently track your progress and measure your performance?
- What metrics are important to you when tracking your cycling workouts? (post it notes)
- What kind of feedback would you like to receive after completing a workout?
    - What kind of metrics/graphs? (draw on a4 or post it)
- How would you like to be presented with this feedback?
    - What is important to you: objective data or is it more than just that? (reward/motivation)
- Do you think it is helpful to provide information on how a wrongly/perfectly executed workout affects your future workouts?

B Design

B.1 Product Requirement Documents

Quality of intervals

Table of contents

Problem statement

- Currently, the app is not providing feedback on the quality of the intervals that users have performed.
- There is no possibility to recognise when intervals have been performed differently compared to the planned intervals.

Goals

- Users are aware of how they have performed their intervals compared to the planned workout.
- Users can take insights from the reported quality metrics to improve in the future.
- The app recognises deviations in performed workouts and presents the user with actionable feedback regarding the deviations.

Non-goals

- Auto-detection of intervals / details per specific block
    - Too complex/time-consuming for now

- Providing full analysis like TrainingPeaks, Intervals.icu

Focus group findings

- "*How well have I done those blocks? Did I finish them properly? Are the wattages good on average? But feedback immediately, that saves me looking for a few things in Strava.*"
- JOIN should recognise where errors occur and provide actionable feedback.
- Requires more than simply showing time in zone."*If you have intervals of 3x10 minutes at a certain wattage, it's different than riding hard into a headwind for 30 minutes and then cycling for 30 minutes easy.*"
- Mainly interesting for outdoors rides, as indoor has less variation.

Proposed solution

For now see Workout Score and Influence of the workout.

What are the high level architectural changes?

What are the high level data model changes?

What are the main changes to the UI?

Risks

What risks might be introduced by this set of changes? Consider running a pre-mortem to raise risks. Be sure to capture mitigating these risks in the Implementation and Rollout Plans.

Are there any backwards-incompatible changes?

TBD

Does this project have special implications for security and data privacy?

TBD

Could this change significantly increase load on any of our backend systems?

TBD

Does this project have any dependencies?

Workout Score

Alternative solutions

What alternatives did you consider? Describe the evaluation criteria for how you chose the proposed solution.

Implementation and rollout plan

Fill this section out based on what is relevant for the size and scope of this project. This section can also be TBD as the project is started, but you should gradually fill this in as the project progresses towards launch.

Does this project require a migration?

If an extensive migration is necessary, write a separate tech spec for it, and link it here. Describe how to rollback in the event of an unsuccessful migration.

Is this project in an experiment or feature flagged?

Describe how to support an incremental release if needed.

Success Criteria

How will you validate the solution is working correctly?

Describe what automated and/or manual testing you will do. Does this project need load or stress testing? This can also be a separate Testing Plan doc that is shared with QA, and linked here.

What monitoring and alerting will you do to ensure this project doesn't decrease performance and reliability?

E.g. Increased requests, latency, and error rates.

Meeting 19 Apr

- Potential to do it only with workout player workout, but not worth the effort
- Export .tcx for in-between

Workout Score

Table of contents

Problem statement

- There is no intuitive, simple metric that is used to show the accuracy of a performed workout vs the planned workout.

Goals

- At a glance, users can see if the intended effect of the planned workout was met.
- Users can interact dynamically with the Workout score. The overall Workout score is visible immediately, but the details regarding this score and the workout intervals are displayed in separate, connected view.

Non-goals

Focus group findings

- Focus of score should be on the intended effect of the workout (i.e. rest periods at VO2max, cadence in strength intervals)

- A simple score is preferred over having to dive into the data/graphs. However, zone information might be helpful, displayed in current JOIN colouring.
- Display of the score should be dynamic. The overall score can be immediately visible, details regarding the score and intervals can be in a separate screen for users that want to delve deeper.

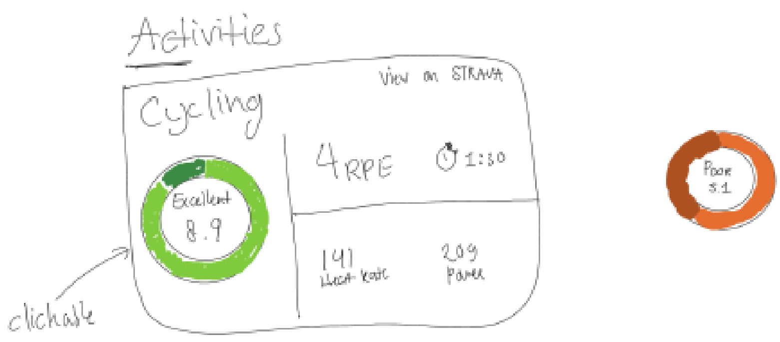
Proposed solution

We propose to add a Workout score that resembles how well an user has performed their workout compared to what was planned. This score consists of the following parameters, listed by importance:

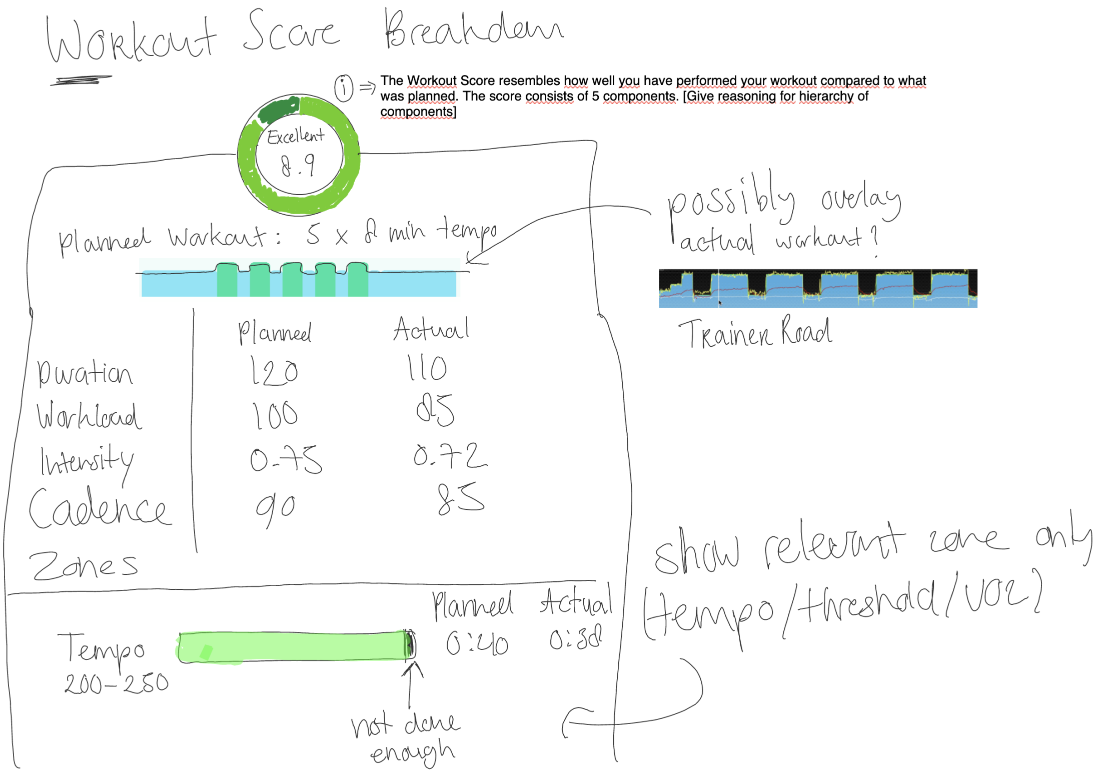
	Goal	Execution	[redacted]	Stars				[redacted]
Duration (min)	120	110		1	1	1	1	
Workload (Stress)	60	55		1	1	1	1	
Intensity (IF)	0,6	0,5		1	1	1	1	
Time spend in target zones (min)				1	1	1	1	
Tempo	25	20						
Threshold	10	12						
Vo2max	9	8						
Average cadence (rpm)	90	85		1	1	1	1	
Total				1	1	1	1	

Users did not have a clear preference for the type of scale of the accuracy score. So could be scaled to 0-100%, 1-10, 1-5 etc.

1. The score is visible in the Activities section after the user loads in their workout and has given RPE.



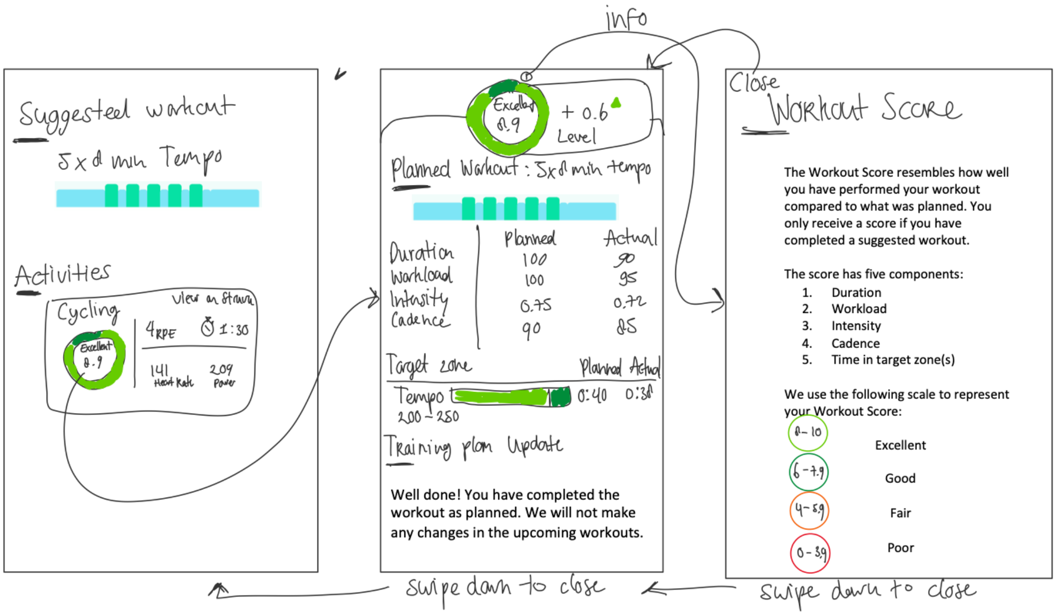
2. Details of the score can be viewed in a separate screen, when the score is clicked. Here, the Time spent in target zones (min) should preferably be displayed in the current colour scheme.



Workout Score Scenarios

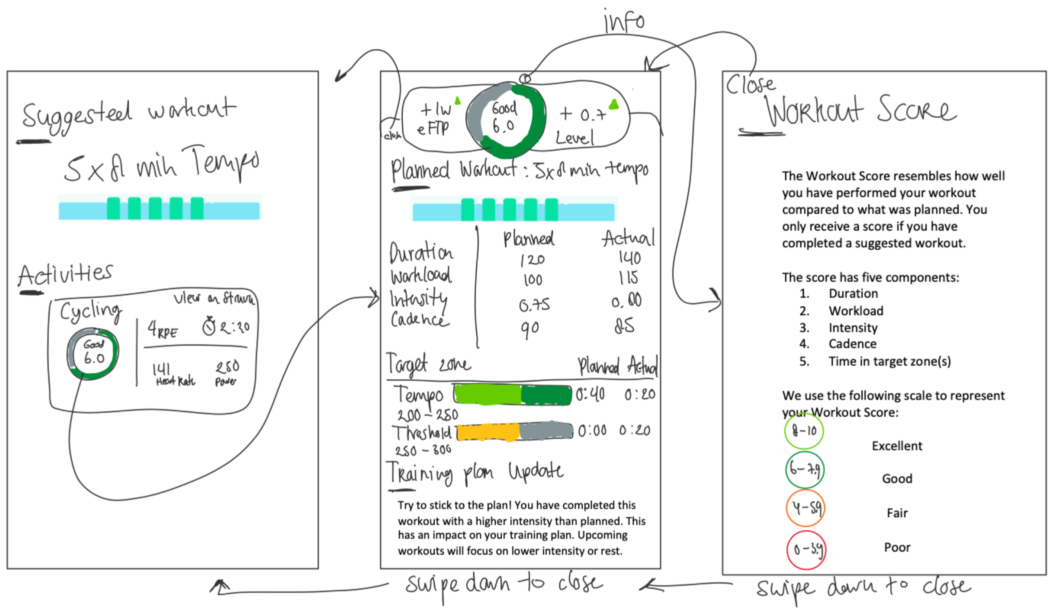
1. **Completed as planned**

    a. User has completed the workout as planned and has a generally good Workout Score. The eFTP remained the same but the JOIN level increased with 0,6. There is no impact on the training plan.



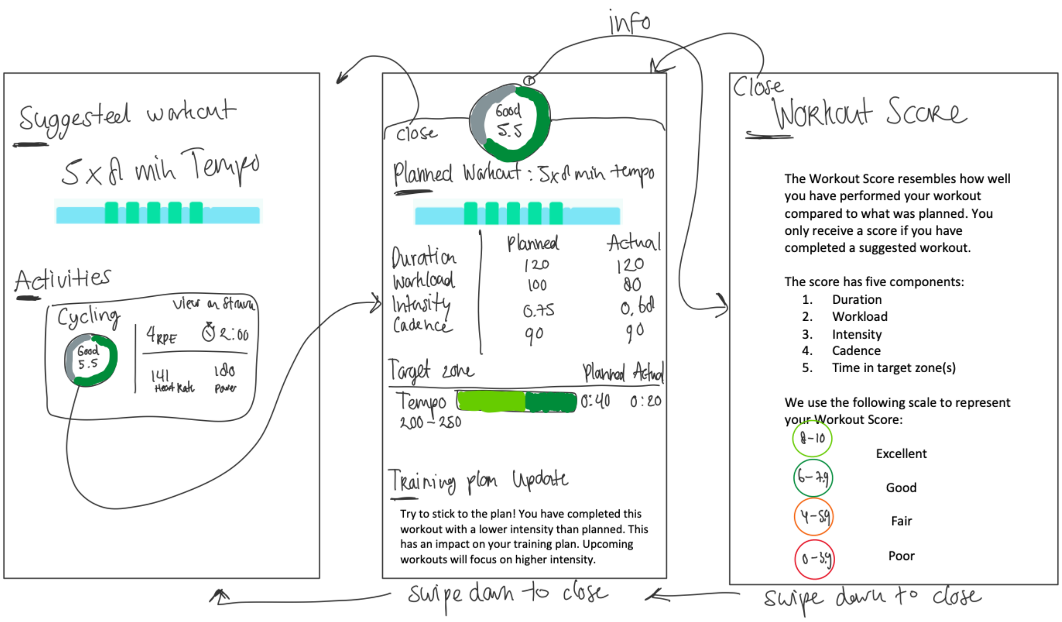
2. **Completed not as planned (higher intensity)**

   a. User has completed the workout but has not completed it as planned. The user did the intervals above the power targets and gets a worse Workout Score. The eFTP increased with 1w and the JOIN level increased with 0,7. This behaviour has an impact on the training plan, the upcoming days JOIN recommends a rest day or some workouts at lower intensity.



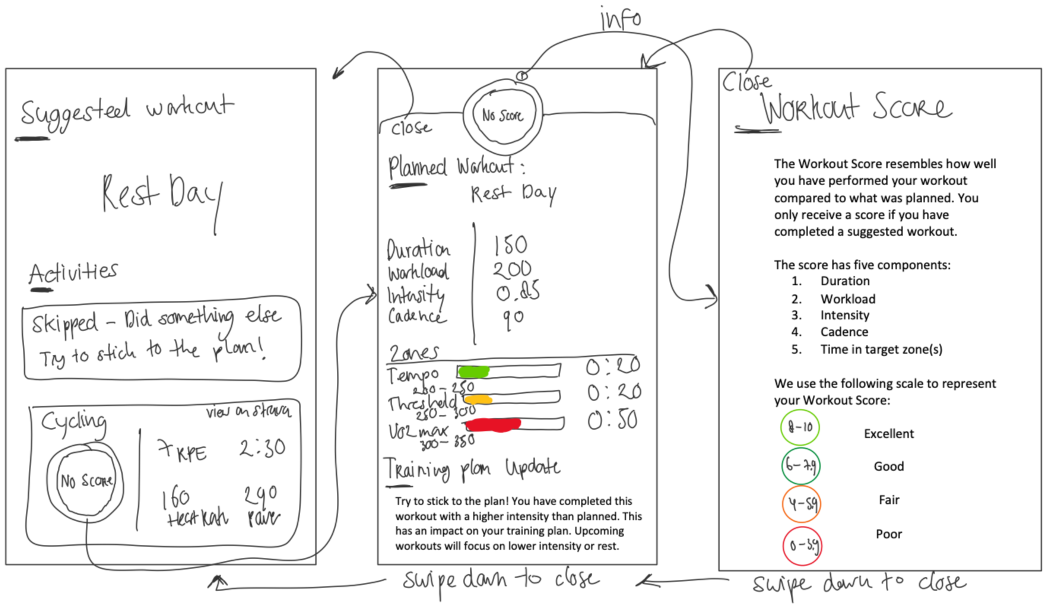
3. **Completed not as planned (lower intensity/rest)**

    a. User has completed the workout but has not completed it as planned. The user did the intervals under the power targets and gets a worse Workout Score. This behaviour has an impact on the training plan, the upcoming days JOIN recommends some higher intensity work.



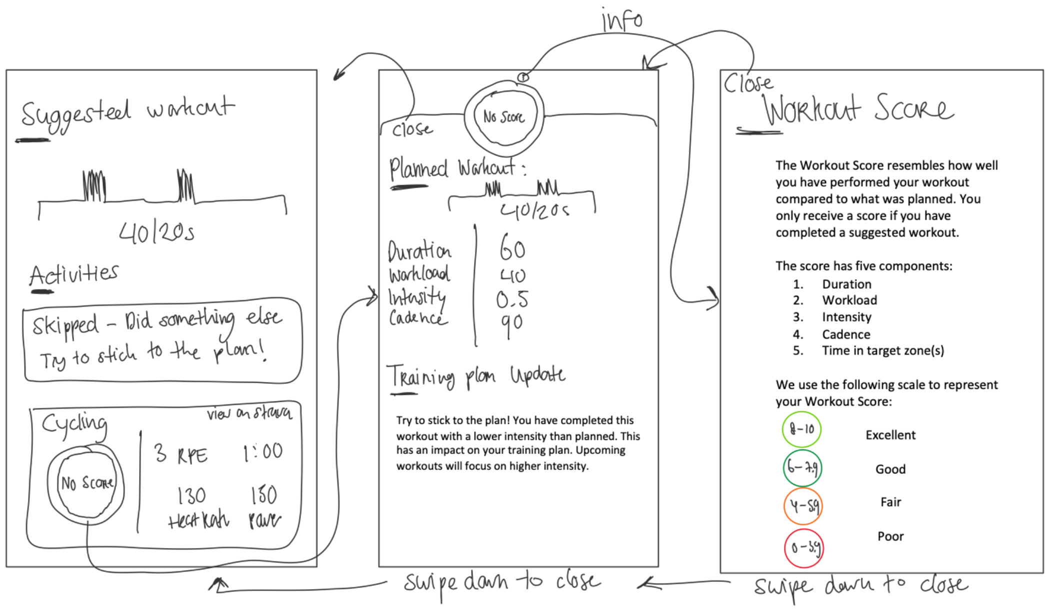
4. **Something else (higher intensity)**

    a. User has done something completely else than was planned. The user did lots of VO2max work instead of the scheduled rest day. The user does not get a Workout Score. This behaviour has an impact on the training plan, the upcoming days JOIN recommends a rest day or some workouts at lower intensity.



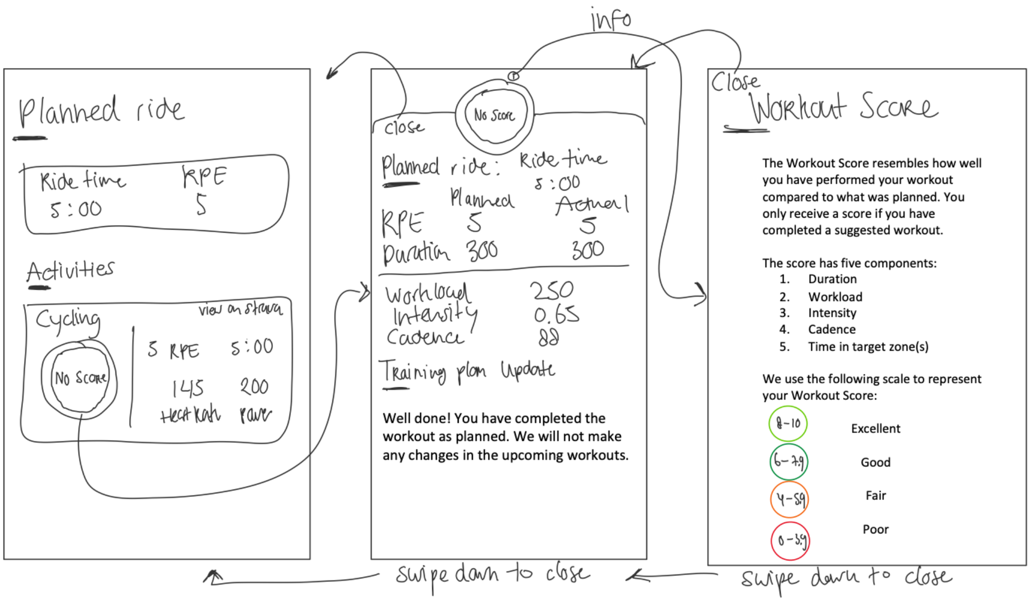
5. **Something else (lower intensity/rest)**

   a. User has done something completely else than was planned. The user did a low endurance ride instead of the scheduled 40/20s. The user does not get a Workout Score. This behaviour has an impact on the training plan, the upcoming days JOIN recommends some higher intensity work.



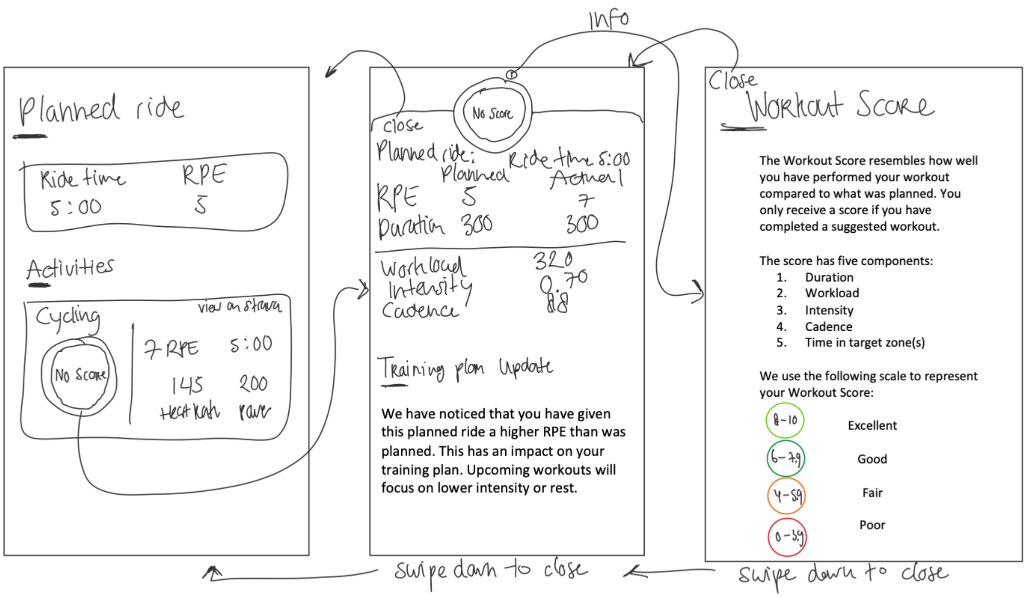
6. **Planned ride (as planned)**

   a. User had planned a ride on the day. The user completed the ride according to the same RPE as given up front. The user does not get a Workout Score. There is no impact on the training plan.



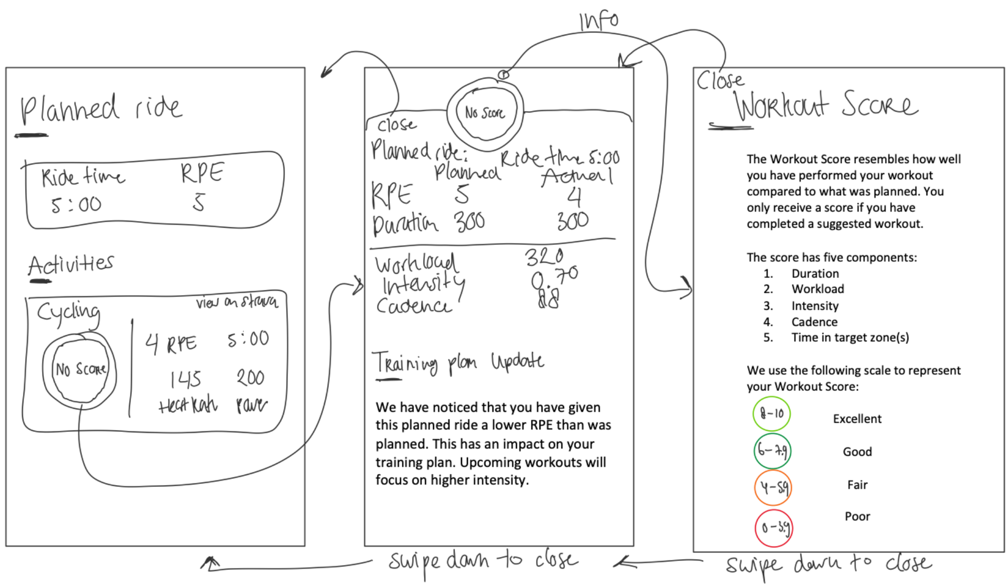
7. **Planned ride (higher intensity)**

   a. User had planned a ride on the day. The user completed the ride with a higher RPE as given up front. The user does not get a Workout Score. This behaviour has an impact on the training plan, the upcoming days JOIN recommends a rest day or some workouts at lower intensity.



8. **Planned ride (lower intensity/rest)**

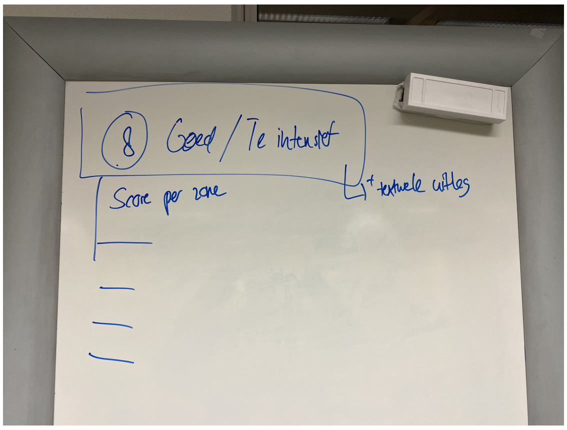
   a. User had planned a ride on the day. The user completed the ride with a lower RPE as given up front. The user does not get a Workout Score. This behaviour has an impact on the training plan, the upcoming days JOIN recommends some higher intensity work.

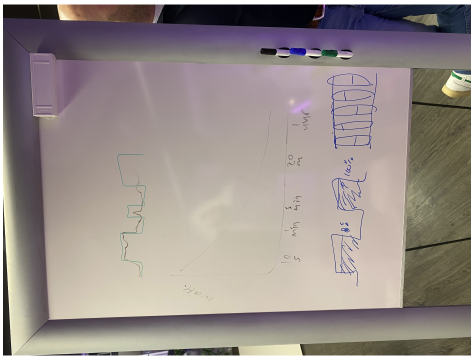


Scenario notes

- In general, only show tempo/threshold/vo2max as these are used in score? No endurance/rest?
- Scenario 2 should we show zones that show the higher intensity or solely the target zone
- In scenario 4 show all zones for did something else
- In scenario 5 did not show it because it was lower intensity so not one of three zones
- In planned ride scenarios show zones?

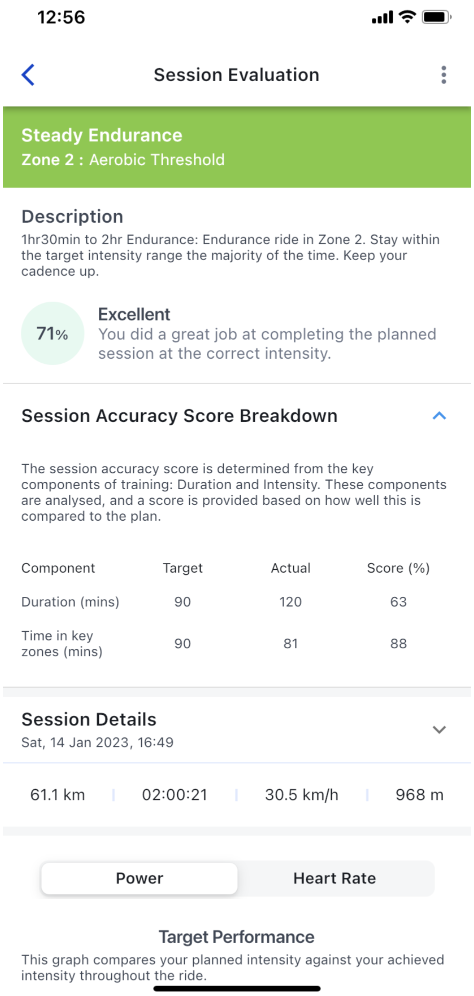
Focus group Whiteboard sketches

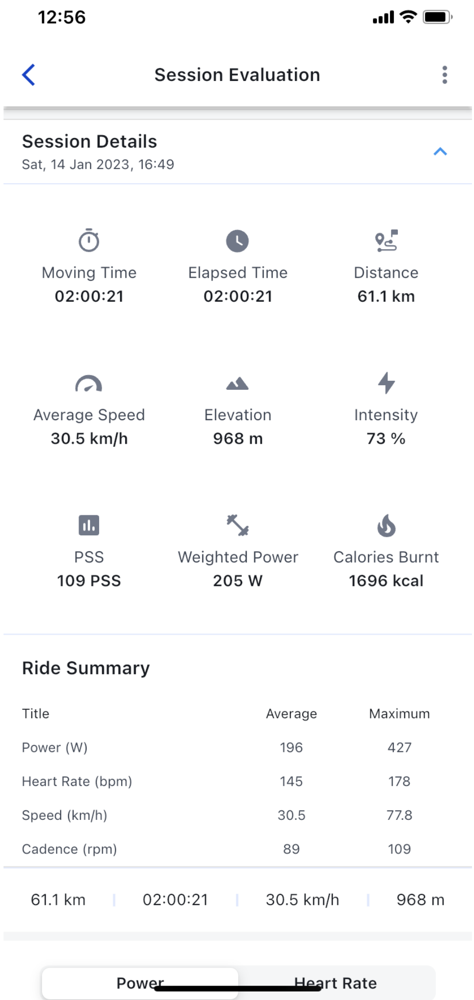


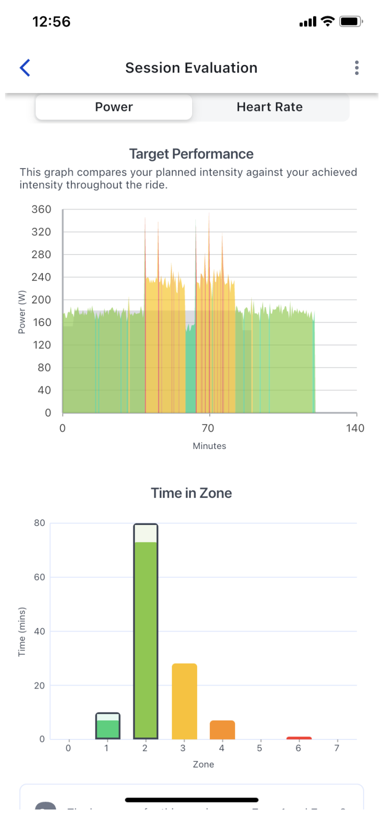


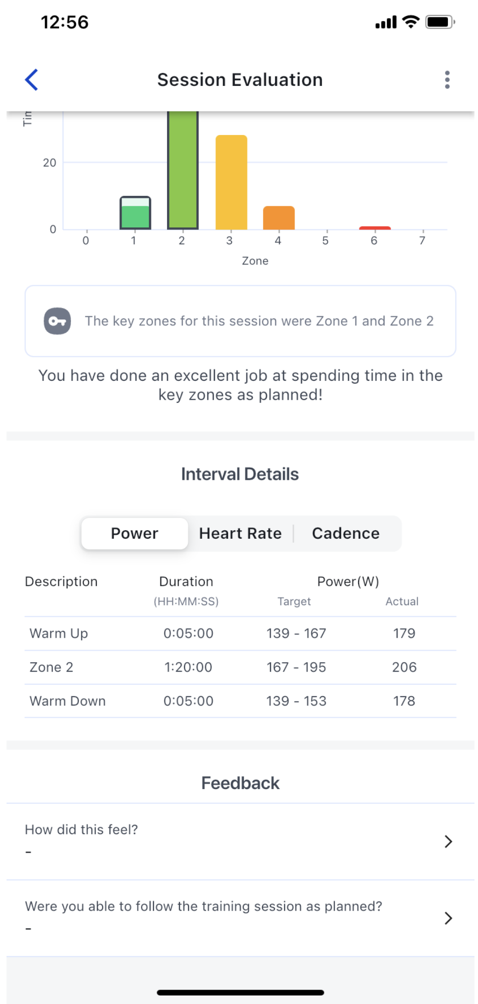
Focus group Post-it categorisation

Session evaluation page from Pillar app


12:56
Session Evaluation
Steady Endurance
Zone 2 : Aerobic Threshold
Description
1hr30min to 2hr Endurance: Endurance ride in Zone 2. Stay within the target intensity range the majority of the time. Keep your cadence up.
71%
Excellent
You did a great job at completing the planned session at the correct intensity.
Session Accuracy Score Breakdown
The session accuracy score is determined from the key components of training: Duration and Intensity. These components are analysed, and a score is provided based on how well this is compared to the plan.
Component
Target
Actual
Score (%)
Duration (mins)
90
120
63
Time in key zones (mins)
90
81
88
Session Details
Sat, 14 Jan 2023, 16:49
61.1 km
02:00:21
30.5 km/h
968 m
Power
Heart Rate
Target Performance
This graph compares your planned intensity against your achieved intensity throughout the ride.


12:56
Session Evaluation
Session Details
Sat, 14 Jan 2023, 16:49
Moving Time
02:00:21
Elapsed Time
02:00:21
Distance
61.1 km
Average Speed
30.5 km/h
Elevation
968 m
Intensity
73 %
PSS
109 PSS
Weighted Power
205 W
Calories Burnt
1696 kcal
Ride Summary
Title
Average
Maximum
Power (W)
196
427
Heart Rate (bpm)
145
178
Speed (km/h)
30.5
77.8
Cadence (rpm)
89
109
61.1 km
02:00:21
30.5 km/h
968 m
Power
Heart Rate





What are the high level architectural changes?

TBD

What are the high level data model changes?

TBD

What are the main changes to the UI?

- Activities screen updated to show Workout Score
- Add Workout Score Breakdown screen

Risks

What risks might be introduced by this set of changes? Consider running a pre-mortem to raise risks. Be sure to capture mitigating these risks in the Implementation and Rollout Plans.

Are there any backwards-incompatible changes?

TBD

Does this project have special implications for security and data privacy?

TBD

Could this change significantly increase load on any of our backend systems?

TBD

Does this project have any dependencies?

TBD

Alternative solutions

What alternatives did you consider? Describe the evaluation criteria for how you chose the proposed solution.

Implementation and rollout plan

Fill this section out based on what is relevant for the size and scope of this project. This section can also be TBD as the project is started, but you should gradually fill this in as the project progresses towards launch.

Does this project require a migration?

If an extensive migration is necessary, write a separate tech spec for it, and link it here. Describe how to rollback in the event of an unsuccessful migration.

Is this project in an experiment or feature flagged?

Describe how to support an incremental release if needed.

Success Criteria

How will you validate the solution is working correctly?

After the implementation phase, we will use an A/B testing setup to test implementation.

- 2 user groups: with/without implementation
- 2 questionnaires: one for both groups, one solely for with implementation to go into detailed quality of function
- Questionnaires with Likert scale questions
- Quantitative analysis

What monitoring and alerting will you do to ensure this project doesn't decrease performance and reliability?

E.g. Increased requests, latency, and error rates.

Influence of the workout

Table of contents

Problem statement

- If a workout is performed differently than planned, the workouts for the upcoming days might be changed. For users, it is not clear what kind of influence possible deviations from the planned workout mean for the adjustment of the training plan.
- Users require more information regarding the influence of a workout on their progression (JOIN level/towards goal).

Goals

- Users should be able to see what kind of influence deviations from the planned workout mean for the adjustment of the training plan.
- Users should be able to view what the influence of a workout is on their progression (JOIN level/towards goal)

Non-goals

- Showing impact of performed week on level and eFTP.

Focus group findings

- All groups showed interest in getting information regarding training plan adjustment after a performed workout
- What is the influence on the plan when a workout is performed more/less intensive?
- But also, what is the influence of a workout on my progression (JOIN level/towards goal)?

Proposed solution

We propose to show the difference in eFTP and JOIN level impact after the workout was loaded into the app. This could go hand-in-hand with the Workout Score Breakdown display after the workout.

Additionally, if the performed workout affects the following workouts in the schedule, users are presented with this after their workout. See the scenarios in Workout Score.

Notes

- Showing CTL/ATL delta could also help show the impact of the individual workout

What are the high level architectural changes?

TBD

What are the high level data model changes?

TBD

What are the main changes to the UI?

- Add a visualisation for eFTP and JOIN level change/impact in Workout Score screen.
- Add a display for influence of workout on trainingplan adjustment

Risks

What risks might be introduced by this set of changes? Consider running a pre-mortem to raise risks. Be sure to capture mitigating these risks in the Implementation and Rollout Plans.

Are there any backwards-incompatible changes?

TBD

Does this project have special implications for security and data privacy?

TBD

Could this change significantly increase load on any of our backend systems?

TBD

Does this project have any dependencies?

- Workout Score

Alternative solutions

What alternatives did you consider? Describe the evaluation criteria for how you chose the proposed solution.

Implementation and rollout plan

Fill this section out based on what is relevant for the size and scope of this project. This section can also be TBD as the project is started, but you should gradually fill this in as the project progresses towards launch.

Does this project require a migration?

If an extensive migration is necessary, write a separate tech spec for it, and link it here. Describe how to rollback in the event of an unsuccessful migration.

Is this project in an experiment or feature flagged?

Describe how to support an incremental release if needed.

Success Criteria

How will you validate the solution is working correctly?

Describe what automated and/or manual testing you will do. Does this project need load or stress testing? This can also be a separate Testing Plan doc that is shared with QA, and linked here.

What monitoring and alerting will you do to ensure this project doesn't decrease performance and reliability?

E.g. Increased requests, latency, and error rates.

Meeting 19 Apr

Use this approach for influences in fitness like fatigue and fitness, freshness. and the other way around. Longer term approach

B.2 High-fidelity prototypes

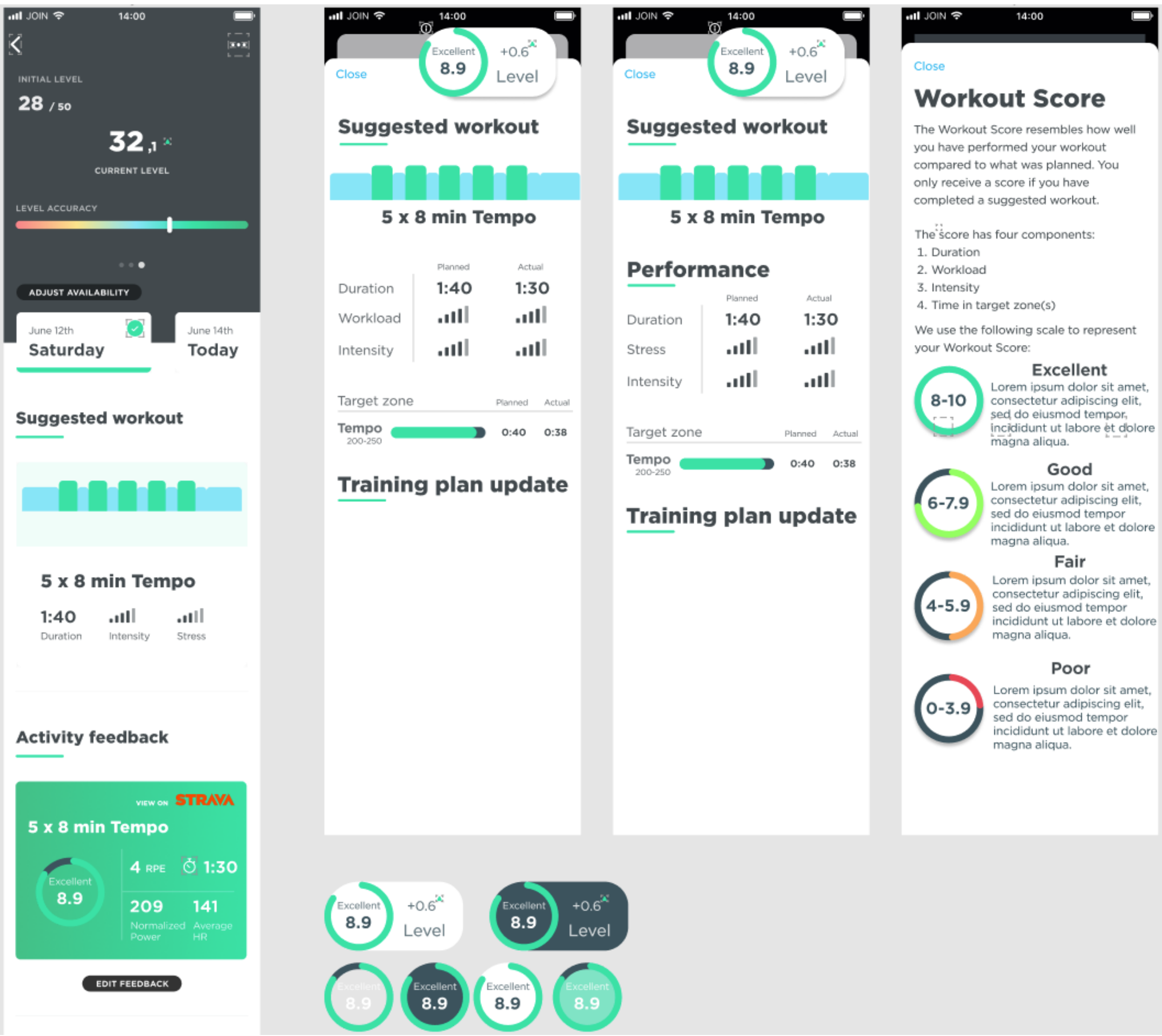


Figure 20: First version of high-fidelity prototype for Completed as planned scenario

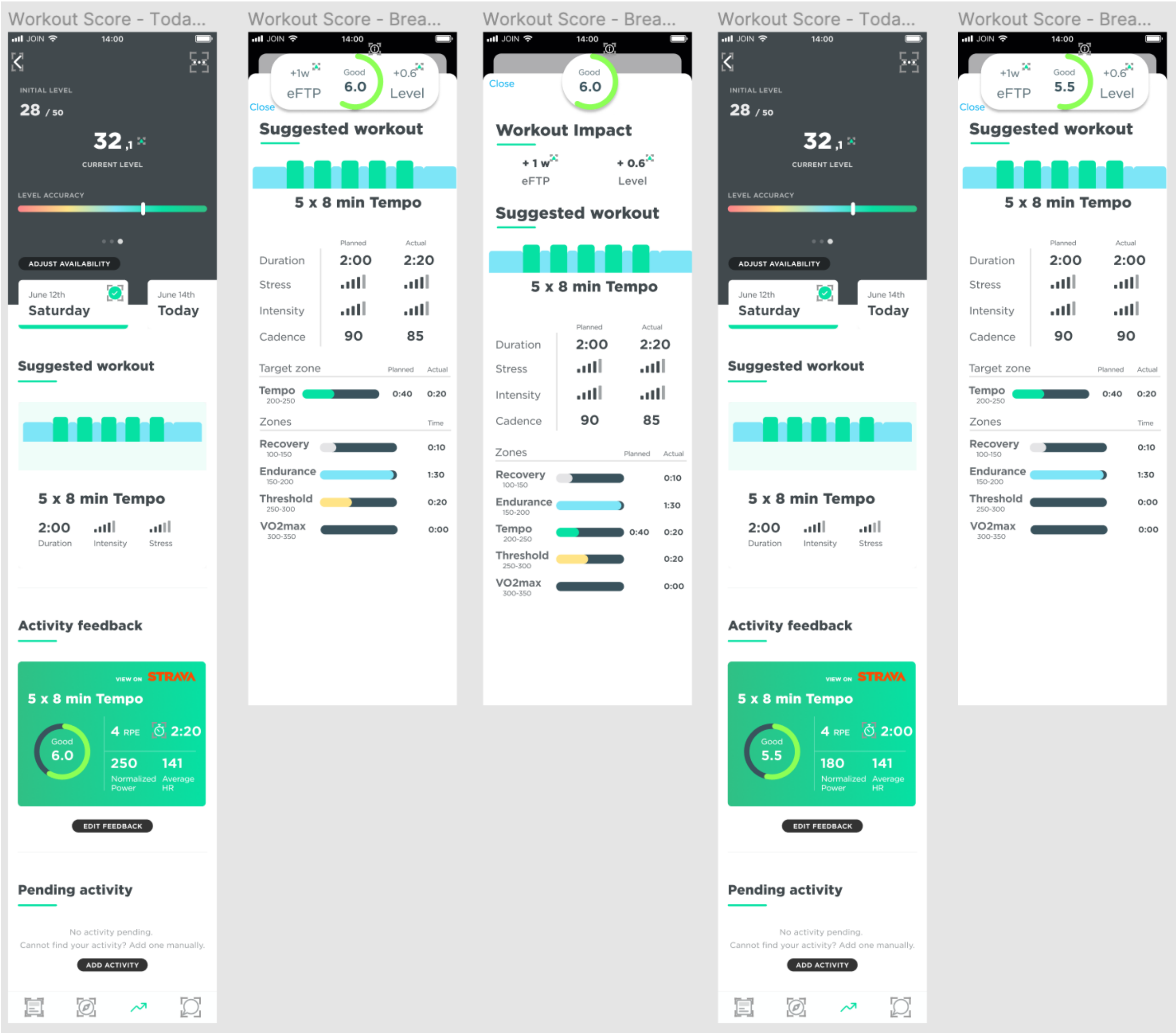
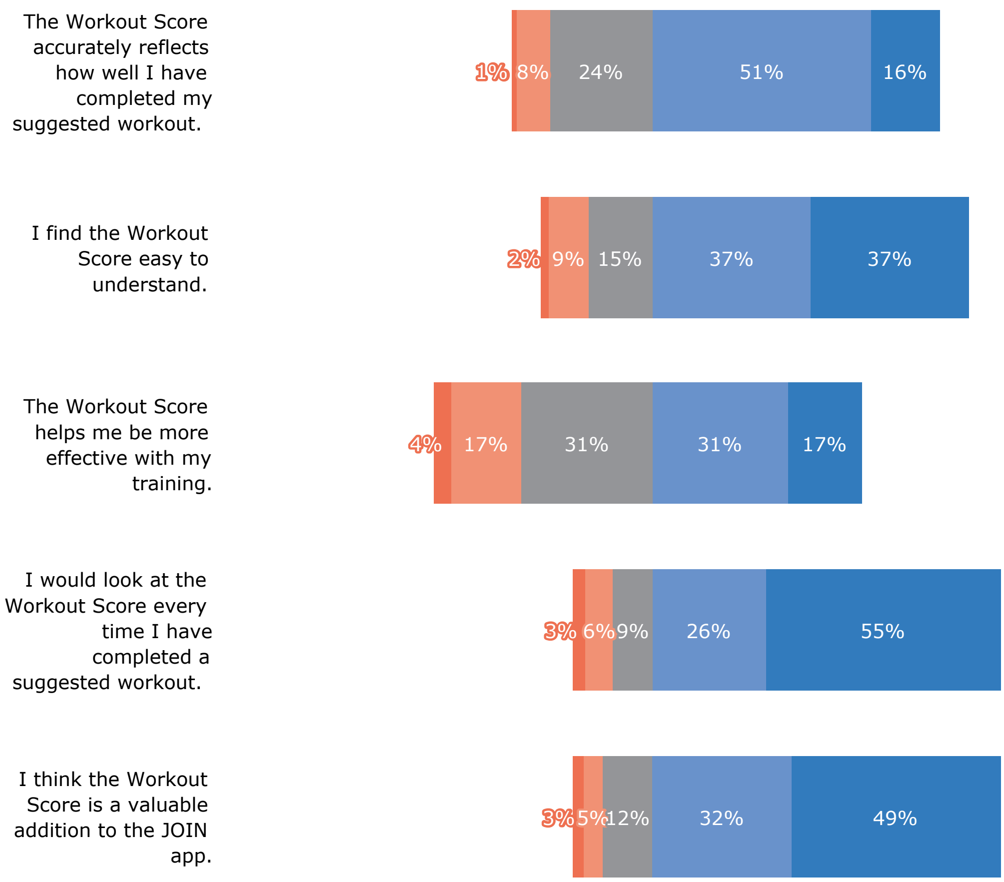


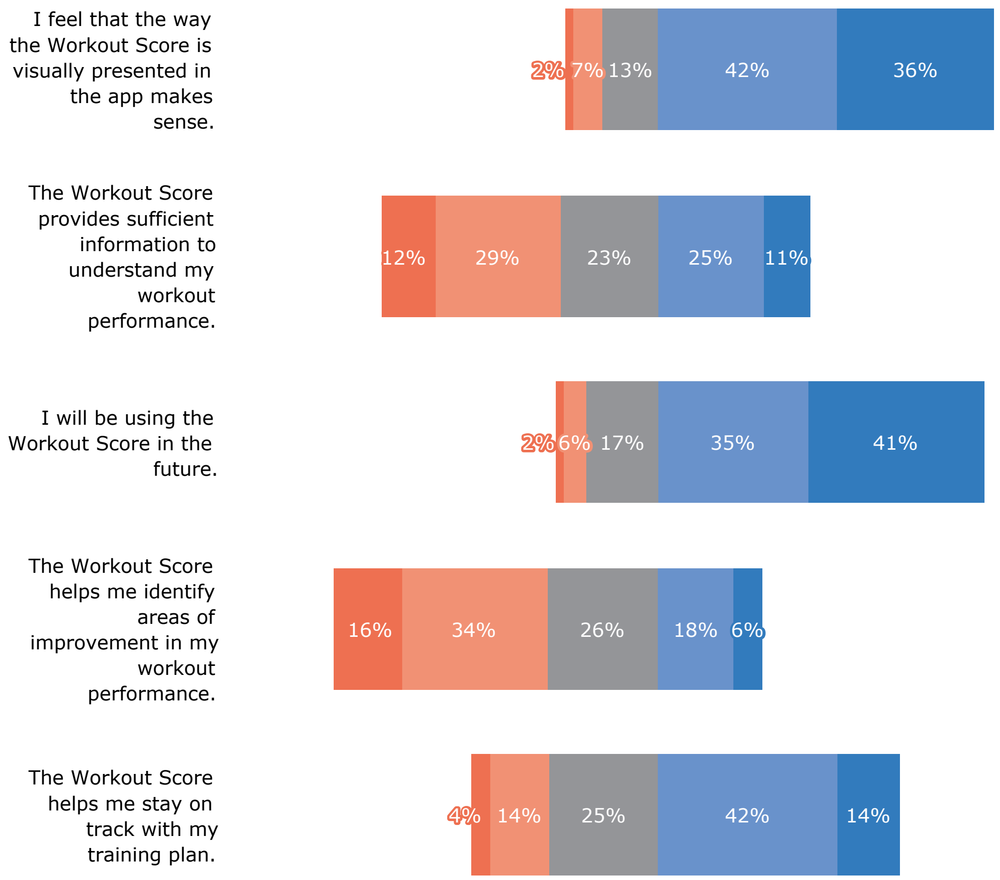
Figure 21: High-fidelity prototypes for Completed not as planned (higher intensity) and Completed not as planned (lower intensity/rest) scenario

Figure 22: Division of elements in high-fidelity prototype of Completed as planned scenario

C Evaluation

Stacked bar graph with all Likert-scale statements.




I feel that the way the Workout Score is visually presented in the app makes sense.
2% 7% 13% 42% 36%
The Workout Score provides sufficient information to understand my workout performance.
12% 29% 23% 25% 11%
I will be using the Workout Score in the future.
2% 6% 17% 35% 41%
The Workout Score helps me identify areas of improvement in my workout performance.
16% 34% 26% 18% 6%
The Workout Score helps me stay on track with my training plan.
4% 14% 25% 42% 14%

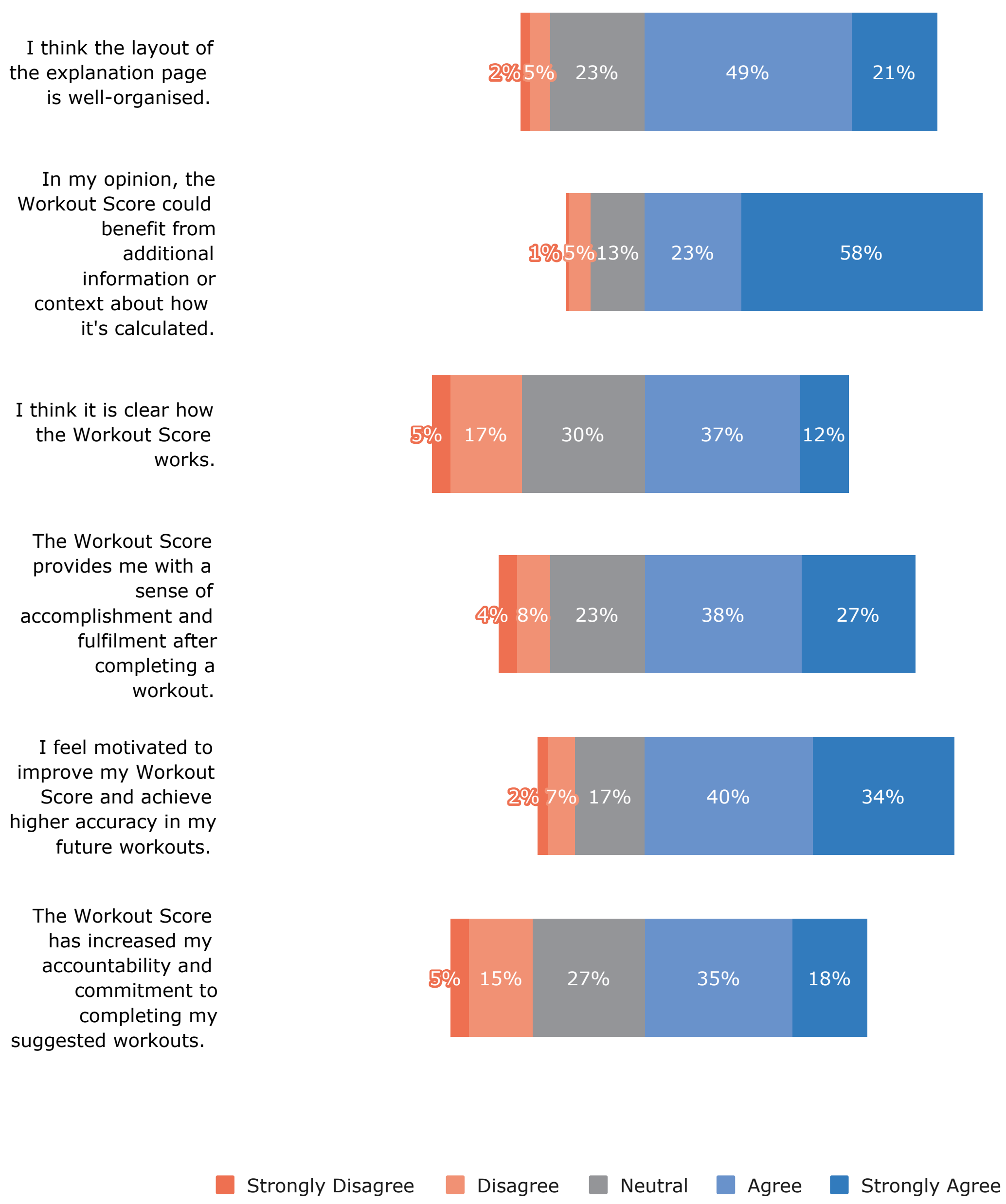

I think the layout of the explanation page is well-organised.
2% 5% 23% 49% 21%
In my opinion, the Workout Score could benefit from additional information or context about how it's calculated.
1% 5% 13% 23% 58%
I think it is clear how the Workout Score works.
5% 17% 30% 37% 12%
The Workout Score provides me with a sense of accomplishment and fulfilment after completing a workout.
4% 8% 23% 38% 27%
I feel motivated to improve my Workout Score and achieve higher accuracy in my future workouts.
2% 7% 17% 40% 34%
The Workout Score has increased my accountability and commitment to completing my suggested workouts.
5% 15% 27% 35% 18%
Strongly Disagree
Disagree
Neutral
Agree
Strongly Agree

Table 6: Mean and Standard Deviation agreement for every Likert-scale statement.

	Statement	***M***	***SD***
1.	The Workout Score accurately reflects how well I have completed my suggested workout.	3.72	0.87
2.	I find the Workout Score easy to understand.	3.98	1.03
3.	The Workout Score helps me be more effective with my training.	3.42	1.08
4.	I would look at the Workout Score every time I have completed a suggested workout.	4.24	1.06
5.	I think the Workout Score is a valuable addition to the JOIN app.	4.20	0.98
6.	I feel that the way the Workout Score is visually presented in the app makes sense.	4.04	1.14
7.	The Workout Score provides sufficient information to understand my workout performance.	2.92	1.04
8.	I will be using the Workout Score in the future.	4.08	0.89
9.	The Workout Score helps me identify areas of improvement in my workout performance.	2.64	0.97
10.	The Workout Score helps me stay on track with my training plan.	3.48	1.21
11.	I think the layout of the explanation page is well-organised.	3.82	0.99
12.	In my opinion, the Workout Score could benefit from additional information or context about how it's calculated.	4.32	0.94
13.	I think it is clear how the Workout Score works.	3.34	1.04
14.	The Workout Score provides me with a sense of accomplishment and fulfilment after completing a workout.	3.75	1.08
15.	I feel motivated to improve my Workout Score and achieve higher accuracy in my future workouts.	3.97	1.00
16.	The Workout Score has increased my accountability and commitment to completing my suggested workouts.	3.46	1.09

D Ethics and Privacy Quick Scan

Response Summary:

Section 1. Research projects involving human participants

P1. Does your project involve human participants? This includes for example use of observation, (online) surveys, interviews, tests, focus groups, and workshops where human participants provide information or data to inform the research. If you are only using existing data sets or publicly available data (e.g. from Twitter, Reddit) without directly recruiting participants, please answer no.

- Yes

Recruitment

P2. Does your project involve participants younger than 18 years of age?

- No

P3. Does your project involve participants with learning or communication difficulties of a severity that may impact their ability to provide informed consent?

- No

P4. Is your project likely to involve participants engaging in illegal activities?

- No

P5. Does your project involve patients?

- No

P6. Does your project involve participants belonging to a vulnerable group, other than those listed above?

- No

P8. Does your project involve participants with whom you have, or are likely to have, a working or professional relationship: for instance, staff or students of the university, professional colleagues, or clients?

- No

Informed consent

PC1. Do you have set procedures that you will use for obtaining informed consent from all participants, including (where appropriate) parental consent for children or consent from legally authorized representatives? (See suggestions for information sheets and consent forms on the website.)

- Yes

PC2. Will you tell participants that their participation is voluntary?

- Yes

PC3. Will you obtain explicit consent for participation?

- Yes

PC4. Will you obtain explicit consent for any sensor readings, eye tracking, photos, audio, and/or video recordings?

- Yes

PC5. Will you tell participants that they may withdraw from the research at any time and for any reason?

- Yes

PC6. Will you give potential participants time to consider participation?
- Yes

PC7. Will you provide participants with an opportunity to ask questions about the research before consenting to take part (e.g. by providing your contact details)?
- Yes

PC8. Does your project involve concealment or deliberate misleading of participants?
- No

Section 2. Data protection, handling, and storage

The General Data Protection Regulation imposes several obligations for the use of **personal data** (defined as any information relating to an identified or identifiable living person) or including the use of personal data in research.

D1. Are you gathering or using personal data (defined as any information relating to an identified or identifiable living person)?
- Yes

High-risk data

DR1. Will you process personal data that would jeopardize the physical health or safety of individuals in the event of a personal data breach?
- No

DR2. Will you combine, compare, or match personal data obtained from multiple sources, in a way that exceeds the reasonable expectations of the people whose data it is?
- No

DR3. Will you use any personal data of children or vulnerable individuals for marketing, profiling, automated decision-making, or to offer online services to them?
- No

DR4. Will you profile individuals on a large scale?
- No

DR5. Will you systematically monitor individuals in a publicly accessible area on a large scale (or use the data of such monitoring)?
- No

DR6. Will you use special category personal data, criminal offense personal data, or other sensitive personal data on a large scale?
- No

DR7. Will you determine an individual’s access to a product, service, opportunity, or benefit based on an automated decision or special category personal data?
- No

DR8. Will you systematically and extensively monitor or profile individuals, with significant effects on them?
- No

DR9. Will you use innovative technology to process sensitive personal data?
- No

Data minimization

DM1. Will you collect only personal data that is strictly necessary for the research?

- No

DM2. Will you only collect not strictly necessary personal data because it is (1) technically unfeasible not to collect it when collecting necessary data, or (2) needed as a source of necessary data?

- Yes

DM3. Will you (1) extract any necessary data as soon as possible from the collected not strictly necessary data and (2) delete the not strictly necessary data immediately after any required extraction?

- Yes

DM4. Will you anonymize the data wherever possible?

- Yes

DM5. Will you pseudonymize the data if you are not able to anonymize it, replacing personal details with an identifier, and keeping the key separate from the data set?

- Yes

Using collaborators or contractors that process personal data securely

DC1. Will any organization external to Utrecht University be involved in processing personal data (e.g. for transcription, data analysis, data storage)?

- Yes

DC2. Will this involve data that is not anonymized?

- Yes

DC3. Are they capable of securely handling data?

- Yes

DC4. Has been drawn up in a structured and generally agreed manner who is responsible for what concerning data in the collaboration?

- Yes

DC5. Is a written contract covering this data processing in place for any organization which is not another university in a joint research project?

- Yes

International personal data transfers

DI1. Will any personal data be transferred to another country (including to research collaborators in a joint project)?

- No

Fair use of personal data to recruit participants

DF1. Is personal data used to recruit participants?

- No

Participants' data rights and privacy information

DP1. Will participants be provided with privacy information? (Recommended is to use as part of the information sheet: For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at www.uu.nl/en/organisation/privacy.)

- Yes

DP2. Will participants be aware of what their data is being used for?

- Yes

DP3. Can participants request that their personal data be deleted?

- Yes

DP4. Can participants request that their personal data be rectified (in case it is incorrect)?

- Yes

DP5. Can participants request access to their personal data?

- Yes

DP6. Can participants request that personal data processing is restricted?

- Yes

DP7. Will participants be subjected to automated decision-making based on their personal data with an impact on them beyond the research study to which they consented?

- No

DP8. Will participants be aware of how long their data is being kept for, who it is being shared with, and any safeguards that apply in case of international sharing?

- Yes

DP9. If data is provided by a third party, are people whose data is in the data set provided with (1) the privacy information and (2) what categories of data you will use?

- Yes

Using data that you have not gathered directly from participants

DE1. Will you use any personal data that you have not gathered directly from participants (such as data from an existing data set, data gathered for you by a third party, data scraped from the internet)?

- No

Secure data storage

DS1. Will any data be stored (temporarily or permanently) anywhere other than on password-protected University authorized computers or servers?

- No

DS4. Excluding (1) any international data transfers mentioned above and (2) any sharing of data with collaborators and contractors, will any personal data be stored, collected, or accessed from outside the EU?

- No

Section 3. Research that may cause harm

Research may cause harm to participants, researchers, the university, or society. This includes when technology has dual-use, and you investigate an innocent use, but your results could be used by others in a harmful way. If you are unsure regarding possible harm to the university or society, please discuss your concerns with the Research Support Office.

H1. Does your project give rise to a realistic risk to the national security of any country?

- No

H2. Does your project give rise to a realistic risk of aiding human rights abuses in any country?

- No

H3. Does your project (and its data) give rise to a realistic risk of damaging the University's reputation? (E.g., bad press coverage, public protest.)

- No

H4. Does your project (and in particular its data) give rise to an increased risk of attack (cyber- or otherwise) against the University? (E.g., from pressure groups.)

- No

H5. Is the data likely to contain material that is indecent, offensive, defamatory, threatening, discriminatory, or extremist?

- No

H6. Does your project give rise to a realistic risk of harm to the researchers?

- No

H7. Is there a realistic risk of any participant experiencing physical or psychological harm or discomfort?

- No

H8. Is there a realistic risk of any participant experiencing a detriment to their interests as a result of participation?

- No

H9. Is there a realistic risk of other types of negative externalities?

- No

Section 4. Conflicts of interest

C1. Is there any potential conflict of interest (e.g. between research funder and researchers or participants and researchers) that may potentially affect the research outcome or the dissemination of research findings?

- No

C2. Is there a direct hierarchical relationship between researchers and participants?

- No

Section 5. Your information.

This last section collects data about you and your project so that we can register that you completed the Ethics and Privacy Quick Scan, sent you (and your supervisor/course coordinator) a summary of what you filled out, and follow up where a fuller ethics review and/or privacy assessment is needed. For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information. Please see the guidance on the ICS Ethics and Privacy website on what happens on submission.

Z0. Which is your main department?

- Information and Computing Science

Z1. Your full name:

Remy van Tussenbroek

Z2. Your email address:

r.vantussenbroek@students.uu.nl

Z3. In what context will you conduct this research?

- As a student for my master thesis, supervised by::
  Judith Masthoff

Z5. Master programme for which you are doing the thesis

- Human-Computer Interaction

Z6. Email of the course coordinator or supervisor (so that we can inform them that you filled this out and provide them with a summary):

j.f.m.masthoff@uu.nl

Z7. Email of the moderator (as provided by the coordinator of your thesis project):

j.f.m.masthoff@uu.nl

Z8. Title of the research project/study for which you filled out this Quick Scan:

Providing feedback on cycling metrics in an adaptive training app

Z9. Summary of what you intend to investigate and how you will investigate this (200 words max):

In collaboration with the mobile adaptive training app JOIN, this study aims to develop a way of providing post-hoc feedback on cycling metrics in a mobile application. The study will use a mixed approach of qualitative and quantitative research methods including focus groups and a survey. This Quick Scan is for the focus group part of my research. For later user testing, another quick scan will be filled in. Using the research methods, the main purpose is to discover how effective post-hoc feedback on training performed can be provided in a cycling application. Additionally, we intend to find out what kind of post-hoc feedback users of different levels of expertise want on training performed. Lastly, this study looks into how this kind of feedback can be effectively provided and how much users appreciate this feedback.

Z10. In case you encountered warnings in the survey, does supervisor already have ethical approval for a research line that fully covers your project?

- Not applicable

Scoring

- Privacy: 0
- Ethics: 0

Response Summary:

Section 1. Research projects involving human participants

P1. Does your project involve human participants? This includes for example use of observation, (online) surveys, interviews, tests, focus groups, and workshops where human participants provide information or data to inform the research. If you are only using existing data sets or publicly available data (e.g. from Twitter, Reddit) without directly recruiting participants, please answer no.

- Yes

Recruitment

P2. Does your project involve participants younger than 18 years of age?

- No

P3. Does your project involve participants with learning or communication difficulties of a severity that may impact their ability to provide informed consent?

- No

P4. Is your project likely to involve participants engaging in illegal activities?

- No

P5. Does your project involve patients?

- No

P6. Does your project involve participants belonging to a vulnerable group, other than those listed above?

- No

P8. Does your project involve participants with whom you have, or are likely to have, a working or professional relationship: for instance, staff or students of the university, professional colleagues, or clients?

- No

Informed consent

PC1. Do you have set procedures that you will use for obtaining informed consent from all participants, including (where appropriate) parental consent for children or consent from legally authorized representatives? (See suggestions for information sheets and consent forms on the website.)

- Yes

PC2. Will you tell participants that their participation is voluntary?

- Yes

PC3. Will you obtain explicit consent for participation?

- Yes

PC4. Will you obtain explicit consent for any sensor readings, eye tracking, photos, audio, and/or video recordings?

- Not applicable

PC5. Will you tell participants that they may withdraw from the research at any time and for any reason?

- Yes

PC6. Will you give potential participants time to consider participation?
- Yes

PC7. Will you provide participants with an opportunity to ask questions about the research before consenting to take part (e.g. by providing your contact details)?
- Yes

PC8. Does your project involve concealment or deliberate misleading of participants?
- No

Section 2. Data protection, handling, and storage

The General Data Protection Regulation imposes several obligations for the use of **personal data** (defined as any information relating to an identified or identifiable living person) or including the use of personal data in research.

D1. Are you gathering or using personal data (defined as any information relating to an identified or identifiable living person)?
- Yes

High-risk data

DR1. Will you process personal data that would jeopardize the physical health or safety of individuals in the event of a personal data breach?
- No

DR2. Will you combine, compare, or match personal data obtained from multiple sources, in a way that exceeds the reasonable expectations of the people whose data it is?
- No

DR3. Will you use any personal data of children or vulnerable individuals for marketing, profiling, automated decision-making, or to offer online services to them?
- No

DR4. Will you profile individuals on a large scale?
- No

DR5. Will you systematically monitor individuals in a publicly accessible area on a large scale (or use the data of such monitoring)?
- No

DR6. Will you use special category personal data, criminal offense personal data, or other sensitive personal data on a large scale?
- No

DR7. Will you determine an individual’s access to a product, service, opportunity, or benefit based on an automated decision or special category personal data?
- No

DR8. Will you systematically and extensively monitor or profile individuals, with significant effects on them?
- No

DR9. Will you use innovative technology to process sensitive personal data?
- No

Data minimization

DM1. Will you collect only personal data that is strictly necessary for the research?
- No

DM2. Will you only collect not strictly necessary personal data because it is (1) technically unfeasible not to collect it when collecting necessary data, or (2) needed as a source of necessary data?
- Yes

DM3. Will you (1) extract any necessary data as soon as possible from the collected not strictly necessary data and (2) delete the not strictly necessary data immediately after any required extraction?
- Yes

DM4. Will you anonymize the data wherever possible?
- Yes

DM5. Will you pseudonymize the data if you are not able to anonymize it, replacing personal details with an identifier, and keeping the key separate from the data set?
- Yes

Using collaborators or contractors that process personal data securely

DC1. Will any organization external to Utrecht University be involved in processing personal data (e.g. for transcription, data analysis, data storage)?
- Yes

DC2. Will this involve data that is not anonymized?
- Yes

DC3. Are they capable of securely handling data?
- Yes

DC4. Has been drawn up in a structured and generally agreed manner who is responsible for what concerning data in the collaboration?
- Yes

DC5. Is a written contract covering this data processing in place for any organization which is not another university in a joint research project?
- Yes

International personal data transfers

DI1. Will any personal data be transferred to another country (including to research collaborators in a joint project)?
- No

Fair use of personal data to recruit participants

DF1. Is personal data used to recruit participants?
- No

Participants' data rights and privacy information

DP1. Will participants be provided with privacy information? (Recommended is to use as part of the information sheet: For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at www.uu.nl/en/organisation/privacy.)
- Yes

DP2. Will participants be aware of what their data is being used for?
- Yes

DP3. Can participants request that their personal data be deleted?
- Yes

DP4. Can participants request that their personal data be rectified (in case it is incorrect)?
- Yes

DP5. Can participants request access to their personal data?
- Yes

DP6. Can participants request that personal data processing is restricted?
- Yes

DP7. Will participants be subjected to automated decision-making based on their personal data with an impact on them beyond the research study to which they consented?
- No

DP8. Will participants be aware of how long their data is being kept for, who it is being shared with, and any safeguards that apply in case of international sharing?
- Yes

DP9. If data is provided by a third party, are people whose data is in the data set provided with (1) the privacy information and (2) what categories of data you will use?
- Yes

Using data that you have not gathered directly from participants

DE1. Will you use any personal data that you have not gathered directly from participants (such as data from an existing data set, data gathered for you by a third party, data scraped from the internet)?
- No

Secure data storage

DS1. Will any data be stored (temporarily or permanently) anywhere other than on password-protected University authorized computers or servers?
- Yes

DS2. Does this only involve data stored temporarily during a session with participants (e.g. data stored on a video/audio recorder/sensing device), which is immediately transferred (directly or with the use of an encrypted and password-protected data-carrier (such as a USB stick)) to a password-protected University authorized computer or server, and deleted from the data capture and data-carrier device immediately after transfer?
- No

DS3. Does this only involve data stored with a collaborator or contractor?
- Yes

DS4. Excluding (1) any international data transfers mentioned above and (2) any sharing of data with collaborators and contractors, will any personal data be stored, collected, or accessed from outside the EU?
- No

Section 3. Research that may cause harm

Research may cause harm to participants, researchers, the university, or society. This includes when technology has dual-use, and you investigate an innocent use, but your results could be used by others in a harmful way. If you are unsure regarding possible harm to the university or society, please discuss your concerns with the Research Support Office.

H1. Does your project give rise to a realistic risk to the national security of any country?
- No

H2. Does your project give rise to a realistic risk of aiding human rights abuses in any country?
- No

H3. Does your project (and its data) give rise to a realistic risk of damaging the University's reputation? (E.g., bad press coverage, public protest.)
- No

H4. Does your project (and in particular its data) give rise to an increased risk of attack (cyber- or otherwise) against the University? (E.g., from pressure groups.)
- No

H5. Is the data likely to contain material that is indecent, offensive, defamatory, threatening, discriminatory, or extremist?
- No

H6. Does your project give rise to a realistic risk of harm to the researchers?
- No

H7. Is there a realistic risk of any participant experiencing physical or psychological harm or discomfort?
- No

H8. Is there a realistic risk of any participant experiencing a detriment to their interests as a result of participation?
- No

H9. Is there a realistic risk of other types of negative externalities?
- No

Section 4. Conflicts of interest

C1. Is there any potential conflict of interest (e.g. between research funder and researchers or participants and researchers) that may potentially affect the research outcome or the dissemination of research findings?
- No

C2. Is there a direct hierarchical relationship between researchers and participants?
- No

Section 5. Your information.

This last section collects data about you and your project so that we can register that you completed the Ethics and Privacy Quick Scan, sent you (and your supervisor/course coordinator) a summary of what you filled out, and follow up where a fuller ethics review and/or privacy assessment is needed. For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information. Please see the guidance on the ICS Ethics and Privacy website on what happens on submission.

Z0. Which is your main department?

- Information and Computing Science

Z1. Your full name:

Remy van Tussenbroek

Z2. Your email address:

r.vantussenbroek@students.uu.nl

Z3. In what context will you conduct this research?

- As a student for my master thesis, supervised by::
  Judith Masthoff

Z5. Master programme for which you are doing the thesis

- Human-Computer Interaction

Z6. Email of the course coordinator or supervisor (so that we can inform them that you filled this out and provide them with a summary):

j.f.m.masthoff@uu.nl

Z7. Email of the moderator (as provided by the coordinator of your thesis project):

j.f.m.masthoff@uu.nl

Z8. Title of the research project/study for which you filled out this Quick Scan:

Providing feedback on cycling metrics in an adaptive training app

Z9. Summary of what you intend to investigate and how you will investigate this (200 words max):

In collaboration with the mobile adaptive training app JOIN, this study aims to develop a way of providing post-hoc feedback on cycling metrics in a mobile application. The study will use a mixed approach of qualitative and quantitative research methods including focus groups and a survey. Earlier, a Quick Scan has been filled for the focus groups. This Quick Scan is for the survey part of my research. Using the research methods, the main purpose is to discover how effective post-hoc feedback on training performed can be provided in a cycling application. Additionally, we intend to find out what kind of post-hoc feedback users of different levels of expertise want on training performed. Lastly, this study looks into how this kind of feedback can be effectively provided and how much users appreciate this feedback.

Z10. In case you encountered warnings in the survey, does supervisor already have ethical approval for a research line that fully covers your project?

- Not applicable

Scoring

- Privacy: 0
- Ethics: 0
