universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„An Analysis of Dashboard Usability and Design for Rowing Coaches"

verfasst von / submitted by

Elisabeth Katharina Wimmer, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2024 / Vienna, 2024

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 645
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Data Science
Betreut von / Supervisor:	Univ.-Prof. Torsten Möller, PhD

Acknowledgements

I would like to thank Torsten for his exceptional mentorship and Nikolaus for his input and for making the crucial contact with the coaches of the Austrian national rowing team. Special thanks go to Anna and Robert and the other rowing coaches who participated in my research for their expertise, without which this would not have been possible. I would also like to thank Christian and Jin for donating their time and visualization expertise to test my prototypes and answer my questions.

Finally, I would like to thank my friends and family for their support throughout this process. Thank you Lena, Marlene and Pavla for all the study sessions and support. I will always be grateful that we spent our study years together. Thank you Laurids for the encouragement to keep going, your patience and for always being there for me. I would also like to thank my cats Azymondias, Rumo and Appa who have cheered me up more times than I can count when I was frustrated during my studies. Of course the list goes on and on and I am grateful to everyone who has supported me along the way. Thank you Mama, Papa, Lini, Nina, Dani, Nini, Fini, just to name a few.

This thesis is dedicated to Azymondias, the best red fluffy cat you could ever wish for, who left us way too soon.

Abstract

In high-performance rowing, the ability of coaches to effectively analyze and interpret athlete data is crucial for optimizing training and overall athlete performance. Dashboard solutions for visualizing sports or health data are already part of many commercial tools. However, despite the availability of general fitness visualization tools, no known solution specifically targets rowing coaches and their needs.

This thesis aims to bridge the gap by introducing a custom dashboard exclusively designed to cater to the requirements of rowing coaches. Domain knowledge was gained through close collaboration with coaches from the Austrian rowing national team and local rowing coaches. Coaches can use the dashboard to explore and analyze all training and well-being data their athletes provide. The coaching dashboard design is created using rapid prototyping and evaluated through usability tests with domain experts. While the implementation needs to be revised for bug and security reasons, the design and functionalities of the dashboard are already appealing to the users.

The core contributions of this thesis are represented by (i) an in-depth exploration of data visualization challenges when it comes to training and well-being data from rowers, (ii) a user-centered approach to dashboard design that integrates domain knowledge from professional rowing coaches, and (iii) the development and iterative refinement of dashboard prototypes.

Kurzfassung

Im Hochleistungsrudern ist die Fähigkeit der Trainer:innen, die Daten der Athlet:innen effektiv zu analysieren und zu interpretieren, entscheidend für die Optimierung des Trainings und der Gesamtleistung der Athlet:innen. Dashboard-Lösungen zur Visualisierung von Sport- oder Gesundheitsdaten sind bereits Bestandteil vieler kommerzieller Tools. Trotz der Verfügbarkeit von allgemeinen Fitness-Visualisierungstools, ist jedoch keine Lösung speziell auf Rudertrainer:innen und deren Bedürfnisse zugeschnitten.

Das Ziel dieser Arbeit ist es, diese Lücke zu schließen, indem ein maßgeschneidertes Dashboard vorgestellt wird, das speziell auf die Bedürfnisse von Rudertrainer:innen zugeschnitten ist. Durch die enge Zusammenarbeit mit Trainer:innen des österreichischen Rudernationalteams und lokalen Rudertrainer:innen wurde Fachwissen gewonnen. Trainer:innen können das Dashboard dazu verwenden, um alle Trainings- und Wohlbefindensdaten, die ihre Athlet:innen bereitstellen, zu untersuchen und analysieren. Das Design des Coaching Dashboards wurde mittels Rapid Prototyping erstellt und durch Usability Tests mit Domänenexpert:innen evaluiert. Während die Implementierung aus Fehler- und Sicherheitsgründen überarbeitet werden muss, sind das Design und die Funktionalität des Dashboards für die Benutzer:innen bereits ansprechend.

Die zentralen Beiträge dieser Arbeit sind (i) eine umfassende Untersuchung der Herausforderungen der Datenvisualisierung in Bezug auf Trainings- und Wohlbefindensdaten von Ruder:innen, (ii) ein nutzerzentrierter Ansatz für Dashboard-Design, der das Fachwissen von professionellen Rudertrainer:innen integriert, und (iii) die Entwicklung und iterative Verfeinerung von Dashboard-Prototypen.

Contents

List of Tables

List of Figures

1 Motivation

In high-performance sports, coaches must keep track of their athletes' training performance and overall well-being. Although many visualization tools are available that enable tracking and analysis of sports or health data, those are usually designed for a broad audience and may not cater to the specific requirements of individual sports disciplines.

Data has gained in importance over the past years in competitive sports. Nowadays, it is frequently utilized to predict or increase the performance of professional athletes. Although there have been many advances in analyzing the behavior of athletes during sports, there are no guidelines on how to design dashboards for coaches. Dashboards in sports, however, can play a distinct role in coaching. They aggregate and visualize vast amounts of data in an easily digestible format, giving an overview of important athlete data and guiding coaches' decisions. Goudsmit et al. highlighted this through their research [22]. They proposed a coaching dashboard design specifically supporting decision-making in athlete development within cyclic sports. Their target user group comprised coaches working with speed skating and running talents, so coaches in individual sports [22]. Rowing is also considered a cyclic sport and can constitute an individual or a team sport. It is competitive; thus, innovative ideas, such as monitoring and utilizing athlete data, are crucial to set oneself apart from the competition. Just like in other cyclic sports, physiological aspects play a vital role in success [34]. Those and other aspects of rowing training can easily be measured and analyzed using visualizations, facilitating the application of a dashboard for coaching. Thus far, there have been no initiatives similar to the work of Goudsmit et al. with a focus on rowing, leaving the specific needs for professional rowing coaches in terms of data visualization and dashboard usability unaddressed.

Daily tasks of professional rowing coaches include training management and providing individually tailored athlete feedback. Those tasks are subject to their domain knowledge and the analysis of various features corresponding to training or well-being, investigating patterns and correlations. Given the exploratory nature of the analysis, a visualization tool can significantly assist coaches.

Providing rowing coaches with a well-designed dashboard can help them make sense of the numerous parameters that need to be analyzed. The multivariate time series data includes ordinal (well-being data such as sleep quality or stress) and quantitative (values obtained by measuring such as training time or resting heart rate) variables. Thus, different visualization approaches are needed. Choosing the most useful techniques for exploring the various features and their dependencies constitutes one of the main challenges.

By including coaches in the design process from beginning to end, the relevant data can be carefully selected and illustrated with their requirements in mind. Successful completion of a helpful coaching dashboard requires a deeper understanding of the needs and preferences of the prospective users, as well as their day-to-day tasks and responsibilities. This is only achievable by working closely with the users and including them in the design process. However, this also leads to challenges. Firstly, conducting interviews and user tests is very time-intensive. Furthermore, diverging opinions or incompatible requirements stated by different coaches call for deliberate and sometimes tough choices.

Implementations for athlete monitoring do exist, but none cater to rowing coaches' unique needs. My thesis aims to fill this gap by creating a tailored dashboard that meets the requirements of rowing coaches and considers their tasks.

Hence, the key components of my approach are defined through several steps. First, a requirement analysis (including task and data analysis) is carried out. These are performed through interviews with Austrian national team rowing coaches to identify their tasks and demands. Based on the discussions, low-fidelity prototypes are created. Subsequently, user tests with rowing coaches are conducted using these prototypes. Feedback from these user tests leads to their refinement and the development of a high-fidelity prototype. This prototype is again tested and further improved, leading to the final implementation of the coaching dashboard.

1.1 Contributions

The following represent the main contributions of the thesis:

- An in-depth exploration of the data visualization challenges unique to the training and well-being data collected from rowers
- A user-centered approach to dashboard design, integrating insights from interviews with professional rowing coaches, ensuring that the resulting prototype is based on real needs.
- The development and iterative refinement of coaching dashboard prototypes, emphasizing usability and tailored data visualization techniques, tested with prospective users.

2 Related Work

While substantial work has been performed in data analysis and visualization in sports in previous years [38, 14], very few relevant contributions were found with a focus on rowing. In this chapter, an overview of previous work in the fields of sports data visualization and dashboard design is given. First, relevant research regarding visualization in sports is reviewed, followed by an evaluation of the literature available on dashboard design. Finally, state-of-the-art methods for visualizing sports data in commercial tools are analyzed.

2.1 Sports Data Visualization

Perin et al. [38] and Du and Yuan [14] respectively surveyed the state of the art of sports data visualization, reviewing research conducted in the domain. The publication of Du and Yuan focuses on competitive sports data and therefore emphasizes that in the research domain, there often is a focus on data gathered during games [14]. This is shown by several examples such as SnapShot [39], SoccerStories [37], TenniVis [40], StatCast [31], Director's Cut [46] or iTTVis [48]. While the papers feature different types of sports, namely ice hockey, soccer, tennis, table tennis, or baseball, all focus on data collected during games. Moreover, all approaches besides StatCast represent a case study involving sports collaborators such as analysts, coaches, or other domain experts. Therefore, they constitute design studies [45], which, according to Perin et al. [38], depict a significant share of research in sports visualization and frequently require a tasks and requirements analysis. Consequently, those studies provide new insights into the respective domains and can serve as a starting point for future visualization research. Our work also represents a design study and helps gain a deeper understanding of the tasks and needs of rowing coaches.

The objective in competitive sports is almost always winning. Therefore, it is not surprising that performance usually plays a role in some way in the research field of sports data visualization. While TenniVis [40] provides a visualization system for the analysis of player performance during individual tennis matches, Pileggi et al. [39] developed a system visualizing National Hockey League shot data with which experts can analyze shot performance over specific periods. In our coaching dashboard, the coaches can also investigate the athletes' performances during competitions. However, the focus lies on daily health and training data and the dependencies between well-being, amount of training, and performance.

2.2 Dashboard Design

Dashboards have grown in popularity for several years and are applied to show detailed information at a glance [16]. However, even though many use them, often enough, too little thought is given to effective visual design, leading to solutions that are not necessarily useful [16, 30].

Nevertheless, before focusing on the correct design, the definition of a dashboard, in general, is of interest. Sarikaya et al. [43] identified two different dashboard types: the visual genre and the functional genre. The former refers to the visual representation of data using simple charts depicted in a tiled layout. The latter genre includes interactive displays, which make it possible to keep track of dynamically updating data in real-time. According to them, tools created for mobile devices often exhibit a visual design quite different from the standard dashboard design while still supporting the same functions. The coaching dashboard developed in the scope of this thesis is likely an excellent example of the functional genre. In his earlier work, Few characterizes dashboards in a manner similar to the visual genre, describing them as visual representations of the most significant information needed to achieve certain objectives [17]. Specifically, this information is arranged on a single screen for easy monitoring. In later writings, however, he differentiates between this concept and dashboards designed for analytical tasks [17]. Furthermore, Sarikaya et al. [43] allocated each dashboard of their collection to one of the following groups: dashboards for decision-making, static dashboards for awareness, and dashboards for motivation and learning. The purpose, audience, visual features, and data semantics of the dashboards differ depending on the cluster.

Concerning good designs, research agrees that it is essential to display precisely the right amount of data and nothing unnecessary [49, 30]. Moreover, further high-level guidelines exist, but generic and practical design guidance is still missing [4]. Bach et al. [4] introduce dashboard design patterns, providing common solutions to designers and supporting them in their choices. During their research, they came up with eight design pattern categories, where the first three correspond to content and the latter five to composition: Data Information, Meta Information, Visual Representation, Page Layout, Screenspace, Structure, Interactions, and Color. Furthermore, Bach et al. [4] detected six different dashboard types that display similar characteristics, have common design pattern combinations, and share contexts or goals. The genres are as follows: static dashboards, analytic dashboards, magazine dashboards, infographic dashboards, repository dashboards, and embedded mini dashboards. According to their description, our dashboard for rowing coaches would correspond to analytic dashboards since they use complete visualizations and tables to display large and detailed data sets. Moreover, they mention interactivity and the possibility of having multiple pages in one dashboard in the context of analytic dashboards [4], which also applies to our prototype.

Visualization research often faces challenges like functional flexibility and visual and analytic literacy. These are even more pronounced for dashboards because of their diversified usage scenarios [43]. According to Sarikaya et al., there is a significant demand for more flexible dashboard features, where the automatic adjustment to different users or display environments is an unresolved issue. During our research, we also encountered those two challenges. Furthermore, Sarikaya et al. state that in the visualization research literature, supporting non-expert users in navigating dashboards still poses a challenge. Issues discussed by Bach et al. [4] included showing the correct information, targeting a specific audience, and putting data displayed into context.

2.3 Commercial Tools

In recent years, personal fitness and health tracking have become popular [33, 42]. Several different apps such as Strava [29], Garmin Connect [32], WHOOP [47] or Fitbit [28] display health and/or training data of users in the form of a dashboard, amongst other functionalities. All collected data is neatly arranged, and on-demand training events or health time series details are displayed. However, as mentioned before, choosing the amount of information depicted constitutes a balancing act. The apps support various sports for a diverse target group, differing in performance level and time spent training. Therefore, they have to cater to the needs of a broad spectrum of potential users, leading to unnecessary information being shown. Moreover, all of the mentioned apps include game features, e.g., badges, levels, leaderboards, progress bars, or quests that encourage users to become better [20, 8]. Results can also be shared amongst friends or strangers, facilitating a comparison with others [20, 8, 42].

Unlike our work, those apps need to create dashboards targeted at a very diverse group of users, while our target group only consists of rowing coaches. This makes it possible for us to focus on a more targeted selection of data and functionalities, whereas apps such as Strava [29], Garmin Connect [32], WHOOP [47] or Fitbit [28] need to include much more. The strength of apps such as Garmin Connect [32] lies in the massive amount of data displayed, as long as one collects it through, e.g., a Garmin watch. There is a multitude of features and possibilities in how to view one's data. However, this simultaneously constitutes a weakness since the user has to invest a lot of time to figure out what they want to view and need for their analysis. Moreover, the objective of our dashboard for rowing coaches is to show the data of all athletes to the respective coach, while the commercial tools mainly aim at showing people their own data. None of the apps is sufficient for helping the rowing coaches with their tasks, illustrating the need for a new, more specialized tool.

2.4 Existing Solutions for Rowing Coaches

While there has been extensive research and commercial development in the realm of fitness and health tracking for various sports, few focus specifically on rowing. However, apart from the conventional dashboard designs and health tracking metrics available in the mentioned apps, there has been a recent innovation aiming to provide real-time feedback to rowing coaches with the help of augmented reality (AR).

ARrow [26] is a real-time AR application explicitly developed for rowing coaches and athletes. Unlike conventional video analysis methods, which are retrospective and time-consuming, ARrow utilizes computer vision techniques to estimate the rower's 3D pose and detect their stroke cycle in real-time. This allows coaches and athletes to receive instantaneous feedback on the rower's body position and stroke, facilitating a more efficient and effective coaching process. The application provides three levels of feedback:

1. Tracking of basic performance metrics over time, similar to those typically observed in dashboards.
2. Visual guidance on the rower's stroke sequence to ensure biomechanically sound practices.
3. A rowing ghost view, a novel feature that aids in synchronizing two rowers' body movements, is essential to team rowing events.

Developed in close collaboration with international rowers and rowing coaches, ARrow represents a significant change in how coaches interact with and train their athletes. A user study with athletes and coaches further validates its effectiveness and usefulness in real-world scenarios [26].

Integrating AR technologies like ARrow in the rowing domain highlights the potential of utilizing advanced technologies to cater to specific coaching needs. Contrary to our work, ARrow analyzes the rowing motion using computer vision. At the same time, no visual data is collected and incorporated into our coaching dashboard.

3 Methodological Approach

The development of the coaching dashboard was approached in accordance with the nine-stage design study methodology framework proposed by Sedlmair et al. [45]. This approach was characterized as a stakeholder-first ordering by Oppermann et al. [36], emphasizing that the stakeholders, in our case, the rowing coaches, dictate the relevant data.

Figure 3.1 depicts a timeline illustrating the development process, highlighting the most important steps along the way.

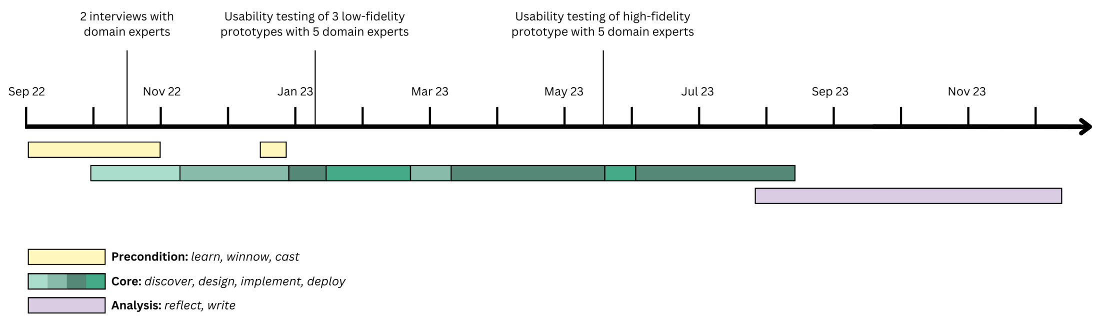


Figure 3.1: *Timeline depicting the development process of the coaching dashboard, indicating the different phases and stages.*

A deeper understanding of the domain context was gained through close collaboration with rowing coaches from the Austrian national team and local rowing coaches. The coaches participated in interviews and usability tests of the different prototypes, providing their expert insights.

During the preconditioning phase, the available athlete data was examined, and existing visualization solutions provided in apps such as Garmin Connect [32] or WHOOP [47] were analyzed. This prepared the interviews with two national team rowing coaches and one local rowing coach, respectively. During the interviews, a requirement analysis was performed, where potential tasks and data demands were discussed with the domain experts. The first two interviews were conducted with the two national team coaches. One local rowing coach was interviewed for comparison since requirements vary on different coaching levels.

As advised by Sedlmair et al. [45], a rapid prototyping approach was pursued. To begin with, simple dashboard sketches were created and continually reviewed through regular talks with the thesis supervisors. During these talks, expert opinions were provided, given that one of them is a visualization and data analysis specialist, and the other is a former professional rowing athlete and coach. After several rounds of feedback, the best and most unique ideas were refined into three different low-fidelity prototypes.

Five domain experts were included in the usability testing of the low-fidelity prototypes. Taking their feedback and comments into account, a high-fidelity prototype was created as a combination of the best components of the previous ones. Five potential users also tested this coaching dashboard prototype. Both sets of user tests included short introductions, followed by the users performing predefined tasks and being able to explore the dashboards. They were asked to speak about their actions throughout the test and provide feedback and comments immediately. The users were asked to rate the prototypes at the end of the respective usability tests. While the low-fidelity prototypes were ranked and rated using three Likert items, the high-fidelity prototype was evaluated using the System Usability Scale (SUS) [10]. The usability tests were conducted remotely using Zoom [12], and help was offered if necessary. To ensure a smooth procedure of the usability tests, two pilot user tests of the low-fidelity prototype were conducted with visualization experts beforehand.

4 Terminology

The following chapter provides definitions of several terms related to sports medicine that are used in the context of this thesis.

Term	**Definition**
RPE	The session rating of perceived extortion (RPE) scale was developed by Foster [19]. Athletes are asked to rate the intensity of their entire training session on a scale from 1 to 10.
Load	For the assessment of internal training load, the athlete's RPE is multiplied by the duration of the training session in minutes [23].
ACWR	Hulin et al. calculated the acute:chronic workload ratio (ACWR) by dividing the acute workload by the chronic workload [25]. While acute workload refers to a week's load, chronic workload is computed as the 4-week rolling average of the acute workload. They concluded that ACWR can be used as a reliable predictor of injury. Subsequently, Gabbett defined the ACWR zone of 0.8-1.3 as the 'sweet spot' and considers values exceeding 1.5 within the 'danger zone' [21].
RHR	RHR is the common abbreviation used for resting heart rate. It is measured in beats per minute.

Table 4.1: *Terminology used within the scope of this work.*

5 Data and Tasks

In this chapter, the *what* and the *why* are characterized as proposed by Tamara Munzner [35]. The characterization is accomplished through an abstraction of the data used for the coaching dashboard and the tasks performed by target users. As previously outlined in chapter 3, two separate interviews were conducted to acquire sufficient information concerning the required data and potential tasks of rowing coaches. The first interview involved two head coaches of the Austrian national rowing team, while the second one was organized with a local rowing coach. By including coaches from different levels of the organization, it was possible to gain a holistic view of the coaching domain in rowing. So, while the total number of interviews was limited to two due to time and availability, the three experienced coaches provided ample insights into their tasks and consequent needs.

5.1 Data Abstraction

The data was provided by Austrian rowing athletes participating in the "AIROW - Artificial Intelligence in Rowing" project [1]. As part of the project, the athletes enter their daily well-being and training data into an app, which is then available to us in real time. In total, the data consists of three datasets of type *table*:

- *activity* comprises all training activities which the athletes enter
- *wellbeing* comprises all well-being entries provided by the athletes
- *auth* holds information on the athletes

Figure 5.1 lists the attributes and their corresponding types from all three datasets utilized for the prototype.

	Attribute Type	Attributes
activity	categorical	activity ID, user ID, activity class, type of activity, boatclass, type of event, comment
	ordinal	intensity, RPE, joy, place in competition
	quantitative	date, time, training distance, duration of training, training load, intervals, race distance, test distance, watt
wellbeing	categorical	ID, user ID, injuries, sickness status, symptoms, training scheduled, holiday status
	ordinal	sleep quantity, sleep quality, state of rest, state of muscles, state of wellbeing, state of stress
	quantitative	date, time, weight, resting heart rate
auth	categorical	user ID, firstname, surname, responsible coach
	ordinal	-
	quantitative	-

Figure 5.1: *A summary overview of the three datasets and their corresponding attributes. The attributes are allocated to the respective attribute type as introduced by Munzner [35].*

Activity Data. This data frame holds all information corresponding to the rowers' training activities. A new activity is created whenever they enter their training into the system. Depending on the type of activity, only certain features are retrieved.

Well-being Data. The athletes must record their physical and mental well-being daily, saved as entries in the *wellbeing* data frame. While *injuries* and *symptoms* are only filled in in the event of either, the other variables are recorded daily.

Authentication Data. The primary purpose of this table is connected to the user authentication of the app, which collects the athletes' data. Still, the unique identifiers and names of the athletes and their coach's names are essential for filtering the data.

5.2 Task Abstraction

Essentially, rowing coaches must supervise their athletes' practice and well-being to guide their training as effectively as possible. Through interviews with the domain experts, we identified the following tasks.

T1. Select Athlete. Coaches must be able to select a specific athlete to evaluate each athlete individually.

T2. Evaluate Training Content. The coaches need to know exactly what was trained. Moreover, they must learn how the athlete felt during the training activity.

T3. Evaluate Training Extent. Sometimes, an athlete is training too much or too little. Identifying either is another task of the coaches. This happens by examining the training duration and comparing it with past data.

T4. Evaluate State of Health. Gaining knowledge about the overall health condition and various factors of influence connected to well-being is essential for the coaches.

T5. Recognize Correlations. The coaches want to identify correlations between, e.g., race or test performance and training or the training schedule and well-being.

T6. Analyze Progress. The possibility of analyzing the athlete's improvement is important for the coaches.

5.3 Design Goals

The following objectives should be achieved through the dashboard.

G1. Clean Design. The dashboard must be structured clearly so users can quickly find what they want.

G2. Focus on the most important features. The dashboard should focus on the most important variables, pre-defined by the national team coaches as ACWR, training load, training duration, weight, sleep quantity, sleep quality, sickness, and injury. Moreover, concerning weight, comparing the preceding day's value and the sleep quantity of the last week's mean is relevant.

G3. More content on demand. There should be a possibility to add further variables of interest.

G4. Detailed Analysis. The users must be able to examine the data in every detail within the dashboard.

G5. Put emphasis on individuals. The athletes' performance shall only be compared to their past performance. There must not be any comparison between athletes.

G6. Views of different time periods. The dashboard should give overviews of the last week, month, and three months. Moreover, the ability to select any other time frame should be available.

5.4 Linking Tasks and Data

Figure 5.2 brings the abstract tasks from section 5.2 into relation with the datasets from section 5.1. Additionally, potential domain user questions, whose answers are needed to solve the corresponding tasks, are included. The colored boxes indicate which dataset is required to answer the respective question.

	Tasks		Questions	activity	wellbeing	auth
T1	Select Athlete	Q1	Which athlete's data needs to be reviewed?			
T2	Evaluate Training Content	Q2	What did the athlete train recently?			
		Q3	How did the athlete feel during the training activity?			
T3	Evaluate Training Extent	Q4	How much time did the athlete spend training?			
		Q5	What training load did the athlete achieve?			
		Q6	In which zones does the athlete's ACWR fall into?			
T4	Evaluate State of Health	Q7	How does the athlete feel overall at the moment?			
		Q8	Is the athlete sick or injured?			
		Q9	Is the athlete sleeping poorly or short on sleep?			
		Q10	Has the athlete lost more weight than is normal?			
		Q11	How do the values compare to the athlete's typical well-being?			
T5	Recognize Correlations	Q12	Is there a relation between the athlete's well-being and their trainings?			
		Q13	Is there a relation between the athlete's performance and their training or well-being?			
		Q14	What could be the cause of the athlete feeling unwell?			
T6	Analyze Progress	Q15	Has there been any improvement on the tests?			
		Q16	Does the athlete enter their training and well-being data regularly?			

Figure 5.2: *A table linking high-level tasks and corresponding domain user questions with the available data determined during the data abstraction. Colored boxes indicate the data required to answer the respective questions.*

The questions were deduced from the information gathered during the two interview sessions with the domain experts. The color mapping in the table is relatively straightforward. Information entailed in *auth* is needed to select the correct athletes for each coach and display the respective athlete's name in the dashboard. The *activity* and the *wellbeing* datasets are of equal value. As indicated by their names, one provides information on what was trained and details on the activity, while the other holds information on the athlete's health and well-being.

T1. Select Athlete. The coaches must ultimately answer Question Q1, but data can support them. Their athlete lists are created with the *auth* dataset. In addition, data from the *activity* and *wellbeing* datasets are used to calculate a compliance score for each athlete. This score is then displayed next to the athlete's name in the selection list and provides information about which athletes regularly submit their data on time. Coaches can also take this into account.

T2. Evaluate Training Content. The assignment of the first two questions is unambiguous. Questions Q2-Q3 essentially capture what task 1 is about, the content of the training session, and how the athlete perceived the training. This information can be found in the *activity* dataset.

T3. Evaluate Training Extent. This task is also performed by analyzing the *activity* data. The training extent is derived from the duration of the training session, which is the response to question Q4 and the athlete's perceived exertion. Question Q5 is answered by combining exactly those two points. Moreover, whether the athlete is under or over-training is the focus of question Q6.

T4. Evaluate State of Health. As opposed to the first five questions, the next five, Q7-Q11, exclusively require insights from the *wellbeing* dataset. Current and past data concerning the athlete's physical and mental health can be found there.

Coaches especially want to keep track of the athlete's sleeping behavior and weight since significant changes in those can indicate upcoming illness.

T5. Recognize Correlations. Both datasets are important for detecting correlations between the athlete's training schedule, well-being, and performance. The coaches try to discover patterns and relationships within and across the two datasets to prevent injuries and optimize the athlete's workout routine regarding performance and well-being.

T6. Analyze Progress. Ultimately, coaches want to help their athlete improve their performance. Therefore, one of their tasks includes the analysis of results from the regular athlete tests, for which the *activity* data is necessary. Furthermore, athlete compliance is tracked through continuous checks of the *activity* and *wellbeing* datasets so that the coaches can verify whether the athlete data is complete and accurate.

6 Coaching Dashboard

This section presents the visualization framework developed after eliciting the requirements (see chapter 5). The specific needs of the domain experts motivated our design decisions.

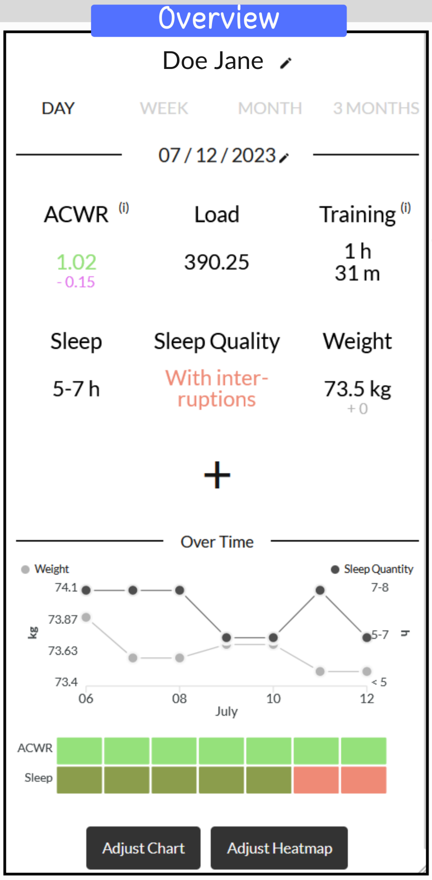
6.1 Implementation

The front end of the coaching dashboard was implemented using `React.js` [3], mainly making use of the libraries `ApexCharts.js` [2] and `D3.js` [9]. Real-time data is accessed from a server via the back end using HTTP requests. The back end was created in `R` [41] with the `plumber.R` [44] package. Three data sets are used: *auth*, *activity*, and *wellbeing*. Furthermore, three functions are implemented in the back end to calculate three more features from the available data, namely: *ACWR*, *compliance*, and *training duration*.

6.2 Overview

The following section provides an overview of the coaching dashboard using a data comic. The idea of creating a data comic to present the final dashboard prototype was inspired by a data comic by Feierl et al. for their visual framework *SunScreen* [15]. They decided to use a data comic instead of just describing the framework's components in words, influenced by the work of Bach et al. [5]. Data comics are engaging and straightforward. They help explain dashboard features in a way that's easy to understand while still capturing all the essential details.

The following data comic was created with Canva [11].



Hello! I will give you a short introduction to the Dashboard for Rowing Coaches!

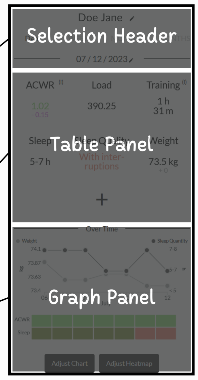
To the left you can see a screenshot of the dashboard.

It is divided into three sections:

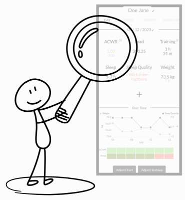
The **Selection Header** lets the user choose what data should be displayed and shows which data is currently selected.

The **Table Panel** shows data in tabular form.

The **Graph Panel** shows a Chart and a Heatmap of selected variables.



Each section will be explored in the upcoming pages.



But before we go into details, let's have a quick look at this summary:

Selection Header

H1 A header showing the user which athlete is currently selected.

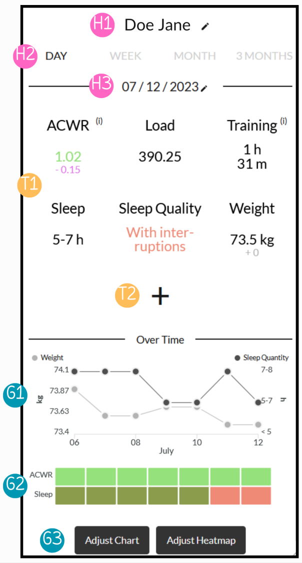
H2 The buttons show which time span is displayed in the table and graph below.
Disclaimer: for **day** the graph panel shows a week of data.

H3 The date button indicates which date is selected.

Table Panel

T1 Table displaying information on selected features.

T2 Button to change the selected features.



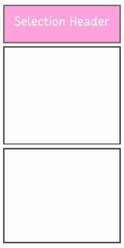
Graph Panel

G1 A plot displaying the values of two selected features over time.

G2 A heatmap that shows the values of selected categorical variables represented by colors over time.

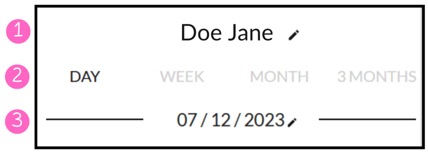
G3 Buttons to change the selected features.

Selection Header

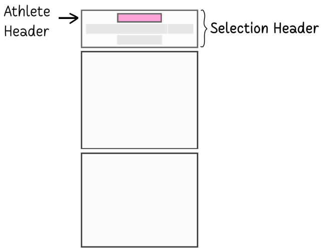


The selection header provides information about the currently displayed data.

It is composed of three different subheaders:



The Athlete Header



Each rowing coach supervises several athletes.

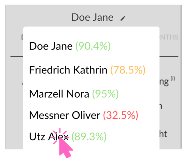
The athlete header shows which athlete's data is currently displayed in the dashboard.

Through clicking on the edit pencil, ...

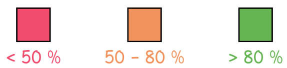
... an athlete list pops up.

It shows all the athletes managed by the coach in alphabetical order and the current compliance score is displayed next to their name.

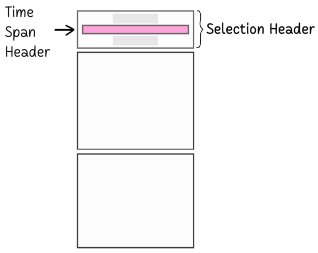
By clicking on a name, the coach can select another athlete.



The color of the compliance score makes it easy for coaches to see who is submitting their data on time.



The Time Span Header

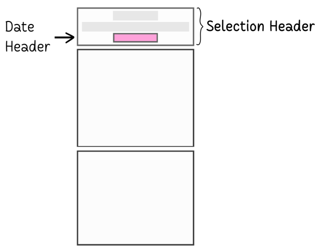


There are four different time spans available:

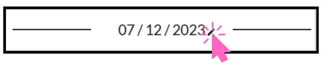
Day is selected by default, and the time spans are arranged by ascending order.

By clicking on another time span, the displayed data on the dashboard is adjusted accordingly.

The Date Header



The coaches are usually interested in the more recent data. Therefore, the current date is selected by default.



Through clicking on the edit pencil, ...

... a calendar pops up, where the coach can pick a different date by clicking on it.

For **week**, **month**, and **3 months**, the selected date represents the last day from the period of data displayed.

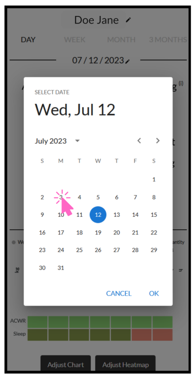
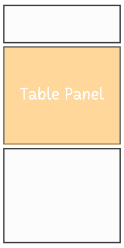
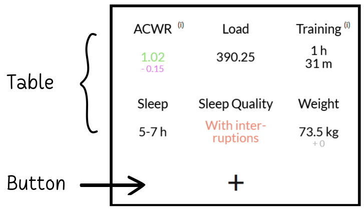


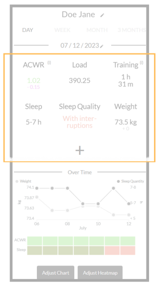
Table Panel



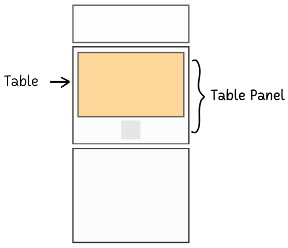
The table panel features the athlete data in a tabular format.

It consists of a table and a button with which the selection of displayed variables can be adjusted.





The Table



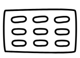
The following variables are preselected in the table:

- ACWR
- Load
- Training (duration)
- Sleep (quantity)
- Sleep Quality
- Weight

Training (ACWR, Load, Training (duration))
Well-being (Sleep (quantity), Sleep Quality, Weight)

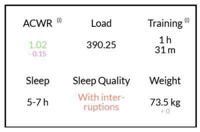
The granularity of the table values depends on the selected time span.

While there are daily values displayed for **day** and **week**, ...



... weekly aggregated values are shown for **month** and **3 months**.

By default, the **day** view is displayed. It shows the exact variable values for the selected date.

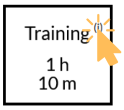


Colors, in particular a mix of the channels hue and saturation, are used for variables that are ordinal, such as sleep quality, or can be categorized, such as ACWR. Changes in ACWR and weight are also displayed in color.

Color scales

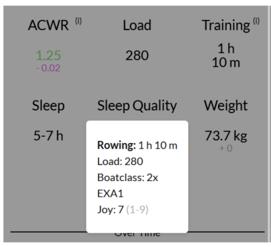
The intermediate color for well-being is always beige, except for sleep quality.

For some of the cells, pop-up boxes have been implemented.
If additional information is available, it is indicated by an (i) icon in the upper right corner of the cell.



Through clicking on it, ...

... a box with the respective details pops up.



For the Training variable, all training sessions for the selected date(s) are displayed in detail. Same for Race and Test, with personal bests added for the Test variable.
For the ACWR variable a color legend is displayed and for the RHR variable the full name Resting Heart Rate is shown.

Clicking anywhere outside the box will hide it again.

The tables corresonding to the other three time spans, **week**, **month**, and **3 months**, look as follows:

week:

	Tue 04.07.	Wed 05.07.	Thu 06.07.	Fri 07.07.	Sat 08.07.	Sun 09.07.	Mon 10.07.
ACWR	0.95	1.11	1.18	1.24	1.25	1.27	1.25
Load	1041	1098	650	703	820	475	280
Training (i)	3 h 30 m	3 h 50 m	2 h 10 m	2 h 55 m	2 h 55 m	1 h 35 m	1 h 10 m
Sleep	5-7 h	7-8 h	7-8 h	7-8 h	7-8 h	5-7 h	5-7 h
Sleep Quality	Good	With inter-ruptions	Good	Good	Good	Good	Good
Weight	73.8 kg +0	73.9 kg +0.1	73.9 kg +0	73.6 kg -0.3	73.6 kg +0	73.7 kg +0.1	73.7 kg +0

month:

	13. - 19. Jun	20. - 26. Jun	27. - 03. Jun/Jul	04. - 10. Jul	11. - 17. Jul
ACWR	1.19	1.34	0.97	1.18	0.94
Load	4160	3934	3024	5067	3409
Training (i)	13 h 35 m	12 h 35 m	11 h 59 m	18 h 5 m	13 h 21 m
Sleep	7-8 h	7-8 h	8 h	7-8 h	7-8 h
Sleep Quality	Good	Good	Good	Good	With inter-ruptions
Weight (kg)	73.77 +0.11	73.53 -0.24	73.59 +0.06	73.74 +0.15	73.66 -0.08

3 months:

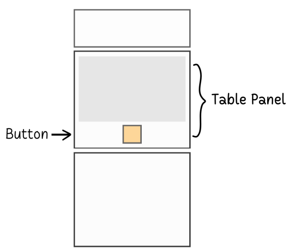
	25. - 01. Apr/May	02. - 08. May	09. - 15. May	16. - 22. May	23. - 29. May	30. - 05. May/Jun	06. - 12. Jun	13. - 19. Jun	20. - 26. Jun	27. - 03. Jun/Jul	04. - 10. Jul	11. - 17. Jul
ACWR	1.02	0.99	1.19	0.93	0.79	1.09	0.93	1.19	1.34	0.97	1.18	0.94
Load	2510	2337	3808	2269	2465	2634	2278	4160	3934	3024	5067	3409
Training (i)	10 h 41 m	8 h 55 m	12 h 45 m	8 h 53 m	8 h 25 m	9 h 39 m	9 h 14 m	13 h 35 m	12 h 35 m	11 h 59 m	18 h 5 m	13 h 21 m
Sleep	7-8 h	7-8 h	7-8 h	7-8 h	7-8 h	5-7 h	7-8 h	7-8 h	7-8 h	8 h	7-8 h	7-8 h
Sleep Quality	Good	Good	Good	Good	Good	Good	With inter-ruptions	Good	Good	Good	Good	With inter-ruptions
Weight (kg)	72.93 -0.53	73.59 +0.66	73.57 -0.02	73.47 -0.1	73.07 -0.4	73.07 +0	73.66 +0.59	73.77 +0.11	73.53 -0.24	73.59 +0.06	73.74 +0.15	73.66 -0.08

The **3 months** table is easier to view when the mobile screen is rotated because it is very wide.

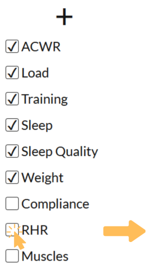
In the **month** and the **3 months** views where data values have been aggregated, the user can click on the row name to check which summary statistic was used. Clicking on the name will display a pop-up box stating e.g., 'Weekly Mean'. Like the pop-up boxes mentioned above, it disappears when the user clicks elsewhere on the screen.

With the arrow buttons < > it is possible to go back or forward by exactly one week without having to open the calendar.

The Button



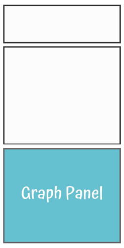
By clicking on the button at the bottom of the table, ...



... the drop-down list for the selection of the variables will appear.

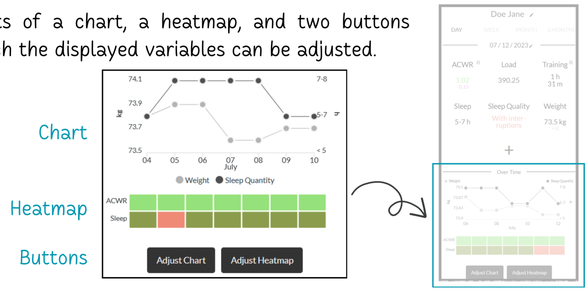
Clicking on an empty box will add the variable to the table, and vice versa if the box is already checked.

Graph Panel

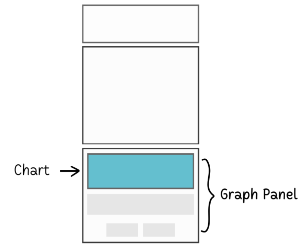


The graph panel features visual encodings of the selected variables.

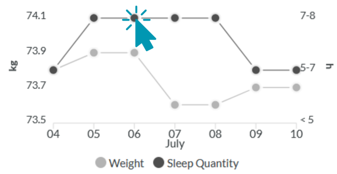
It consists of a chart, a heatmap, and two buttons with which the displayed variables can be adjusted.



The Chart

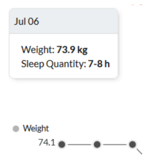


The chart can display up to two quantitative variables, with Weight and Sleep Quantity selected by default.

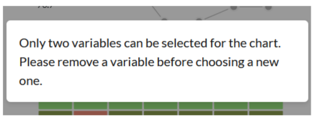


If you click on a data point in the plot, ...

... the exact values for that date are displayed above the plot.



When someone tries to add a third variable, ...

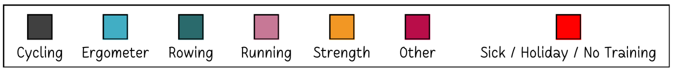


... the message above pops up.

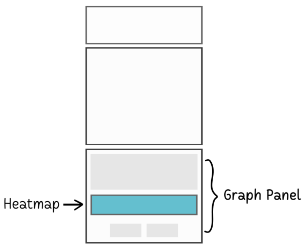
And if one of the selected variables is Training, this data is displayed as a stacked bar chart ...

... showing the distribution of training time across the different categories.

The following color palette was chosen for the stacked bar chart:



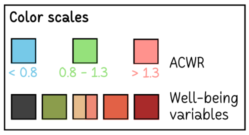
The Heatmap



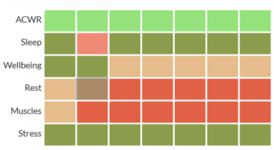
The heatmap can show all ordinal well-being data, and the ACWR values grouped into three categories.

ACWR and Sleep Quality are selected by default, ...

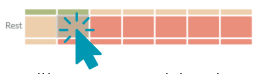
... and it uses the same color scales as the table.



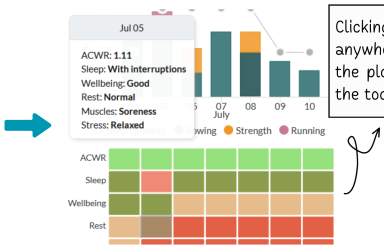
There is no limit to the number of variables that can be displayed simultaneously.



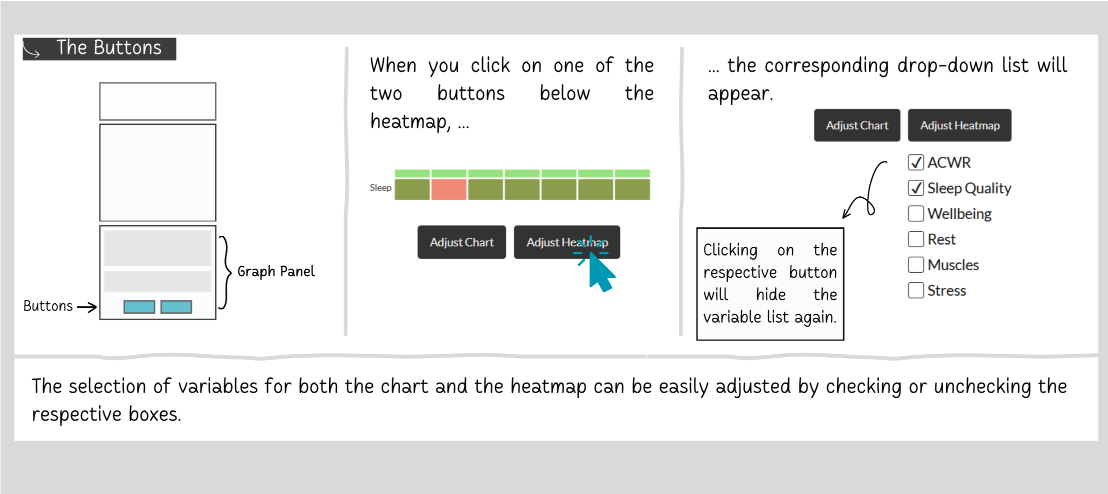
Clicking on one of the cells ...



... will open a tooltip above the heatmap, as in the case of the chart. It shows all the values for that date.



Clicking anywhere next to the plot will hide the tooltip again.


The Buttons
Graph Panel
Buttons
When you click on one of the two buttons below the heatmap, ...
Sleep
Adjust Chart
Adjust Heatmap
... the corresponding drop-down list will appear.
Adjust Chart
Adjust Heatmap
ACWR
Sleep Quality
Wellbeing
Rest
Muscles
Stress
Clicking on the respective button will hide the variable list again.
The selection of variables for both the chart and the heatmap can be easily adjusted by checking or unchecking the respective boxes.

6.3 Design Decisions

This section presents the findings from feedback sessions and user testing of our rowing coach dashboard, designed using rapid prototyping. Feedback from thesis advisors and domain experts was instrumental in refining the dashboard's design. The reasoning for choosing visual encodings for different dashboard elements is explained below.

6.3.1 Why are the specific periods chosen?

During the initial interviews with domain experts, they indicated that the most crucial periods for analyzing their athletes' data were day, week, month, and three months. Therefore, we stuck to these periods when creating the visual encodings for the low-fidelity prototypes. Since there was no objection or negative feedback regarding the chosen periods, they remained the same in the final prototype.

6.3.2 Why use a table to display data?

The table was a central piece of one of the three low-fidelity prototypes. Including a table in one of these prototypes was a choice to provide diversity since the goal was to create three prototypes that were as different as possible. While a visualization expert thought the table was a terrible choice during pilot user testing, it was extremely popular with domain experts. According to them, it is beneficial to see the exact training and well-being values simultaneously, in addition to the plots.

6.3.3 Why use a heatmap to encode ordinal data?

The heatmap shows all of the ordinal well-being data and the ACWR values, which are grouped into three different levels. Using a heatmap, the color channel can express sentiment, while time is encoded on the x-axis and the variable is encoded on the y-axis. This provides an intuitive, compact view of the data from which relationships can be easily derived.

6.3.4 How were the colors chosen for the heatmap?

Each well-being variable has five categories, from very negative to very positive, with a neutral value in between. A diverging colormap with the hues red and green at the endpoints and beige as the midpoint was chosen to represent the categories. The only exception is the sleep quality variable, which has "with interruptions" as an intermediate category. Here, the feedback from a national team coach during one of the user tests suggested choosing a color that shows a slightly negative sentiment since it is not a neutral category. Therefore, we replaced the neutral beige with a low-saturated red shade. Because our well-being data is categorical, specifically ordinal, our color map is segmented into discrete color bins.

The ACWR values were categorized based on the relationship established by Gabbett et al. between ACWR values and the risk of injury in athletes [21]. Here, we used the hue channel to distinguish between the three categories and only used low-saturation colors. Values within the sweet spot range of 0.8 - 1.3 were colored green. Anything less than 0.8 was colored blue, indicating undertraining, and anything greater than 1.3 was colored red, indicating an increased risk of injury due to overtraining. The low-saturation colors were explicitly chosen so that the green and red would differ from those used for the well-being variables. It was considered to use completely different hues, but red and green best convey positive and negative sentiments. Since none of the domain experts had problems distinguishing between the colors and no concerns were raised, we stuck with the original color choice.

Changes in ACWR and weight are shown in the table and are also color-coded. We have chosen to use the hue of magenta to indicate a decrease in value and the hue of orange to indicate an increase in value. Gray is used as a neutral color when there is no change in value. The colors were chosen because they are distinguishable from each other and were not present in the table part of the prototype.

6.3.5 Why add graphs showing a week of data to the day view?

While focusing on daily data, it can be helpful for coaches to have a quick overview of past days when reviewing a specific date. The alternatives included no graphs in the day view or one graph with only daily data. Not having any graphs would result in a waste of space, and a graph with only one observation per variable does not provide any additional information to the numbers displayed in the table. We chose to show a week in the graphs because it was the next larger time period the coaches wanted to see. Compared to the week view, the only difference is that the training duration is shown as a line, just like all the other variables, and not as a stacked bar chart. This decision was made because the focus was on a quick overview, not details.

6.3.6 Why use a stacked bar chart for the training duration?

The workouts performed by the athletes can be assigned to different categories. Stacked bar charts can display information on two-dimensional tables with two keys [35]. They are, therefore, suitable for displaying the duration of the training with the date and the category of the training session as two key values.

6.3.7 How were the colors chosen for the chart?

If Training is not one of the displayed variables, a gray hue with a different luminance is used for each. Since there are already a lot of colors used within the dashboard, the two shades of gray were chosen. They allow for an easy distinction between the variables without confusing the user with more color hues.

However, a larger color palette is required if Training is one of the variables displayed. Six different training categories need to be differentiable in the plot. A seventh color indicates days when the athlete was either sick, on vacation, or had a "no training" day. The training categories are distinguished using a segmented color map with appealing but distinguishable colors. As the national team coaches requested, red was chosen as an additional color to indicate special days. We decided to use the same color for all three to keep the palette small since the exact event - sick, vacation, or no training - can be read from the table or the tooltip. For the second variable, displayed aside from Training, we kept the gray with the higher luminance.

6.3.8 Why use a dual-axis chart?

One of the coaches' tasks is to identify correlations between certain variables. A dual-axis graph is a good choice because it allows them to see how two variables tend to change together. While Few discusses the drawbacks of dual-axis charts and generally discourages their use [18], we have not developed a more appropriate way to display the data over time without taking up too much space. If space were unlimited, juxtaposing views would have been a valid alternative. However, we managed to comply with most of his conclusions. He strongly recommends using a dual-scaled axis only when you have two data sets with different units of measurement, which is valid for all the features that can be selected for the plot, except for training duration and sleep quantity. However, the latter is currently measured in time categories, not exact time measurements. He also states that a dual-scaled graph should never encode values exclusively as bars but concludes that only lines are appropriate. We mostly used lines and only used the stacked bar chart for the duration of the training.

6.3.9 Why limit the amount of features displayed in the chart?

Since all numeric variables have different units of measurement, we could not find a way to display more than two of them in one plot while also using the x-axis to display time.

6.3.10 Why use pop-up boxes and tooltips for displaying details on demand?

In the dashboard, tooltips and pop-up boxes are deliberately used to enhance user interaction while ensuring that not too much is displayed simultaneously. Clickable tooltips in the plot and heatmap are designed to reveal the data associated with each visualization. They are a common way to show details on demand while allowing for an overview and providing an interactive means for coaches to explore the visualized data.

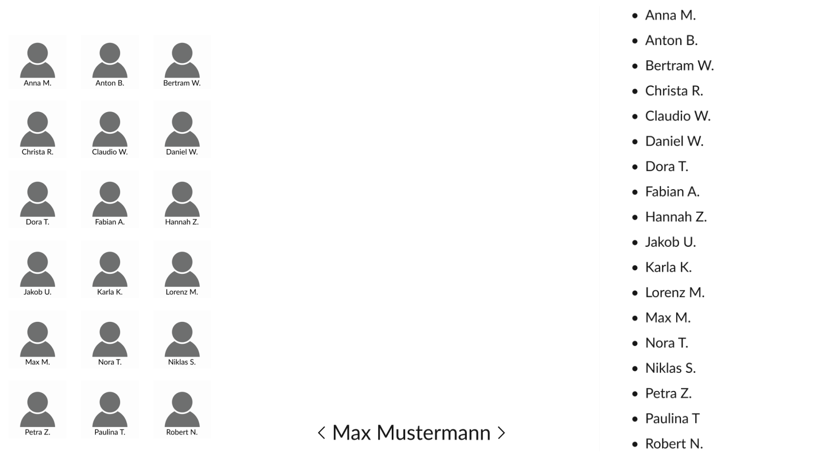
Pop-up boxes in the table context serve two purposes. They provide additional information, such as color legends or definitions of key terms (e.g., RHR for resting heart rate) and, in the case of training data, detailed information about all training sessions for the selected period. Information icons (i) indicate the pop-up boxes in the table. We also added an informative pop-up box in the plot area that activates when a user attempts to select a third variable. This pop-up informs the user of the limitation to select up to two variables, so they know why selecting another variable is not working.

6.3.11 Why use a drop-down list with checkboxes to select variables?

Visualization experts recommended the drop-down list during pilot user testing. Showing the selection only when necessary saves a lot of space. The checkboxes are a simple way to filter the variables.

6.3.12 Why use a list of names for the athlete selection?

We tested three different versions of the athlete selection in the three low-fidelity prototypes. The three variants are shown below: 6.1(a) consists of a grid of profile pictures including and sorted by athlete names, 6.1(b) only provides the ability to switch between an alphabetical list of athletes by clicking left or right, and 6.1(c) is a simple list of names sorted alphabetically.



(a) Grid using pictures. (b) Clicking left and right. (c) List of names.

Figure 6.1: *Three versions of athlete selection.*

The domain experts felt that the list 6.1(c) was the best way to select athletes. In addition, the coaches gave feedback that the surnames of the athletes are more important than the first names of the athletes. So, we changed the alphabetical order to surnames. During user testing of the high-fidelity prototype, one of the coaches requested that the compliance scores be displayed next to the names. This gives the coaches a first impression of who is providing the data and, therefore, suitable for analysis. We incorporated this change in the revision of the high-fidelity prototype.

7 Evaluation

This section presents the evaluation results of our rowing coach dashboard. The main focus is on the usability tests carried out with the low-fidelity prototypes and high-fidelity prototypes, respectively.

7.1 Feedback Rounds

Through feedback rounds with the thesis supervisors, we converted several ideas of how to visualize the training and well-being data of the rowers into three distinct prototypes. We ensured that the dashboard versions varied significantly in visually encoding the data.

7.2 Usability Testing

As indicated in chapter 3, usability tests were carried out to assess low-fidelity and high-fidelity prototypes. All user tests were done remotely through Zoom [12] since this simplified the scheduling of the tests.

7.2.1 Pilot User Tests

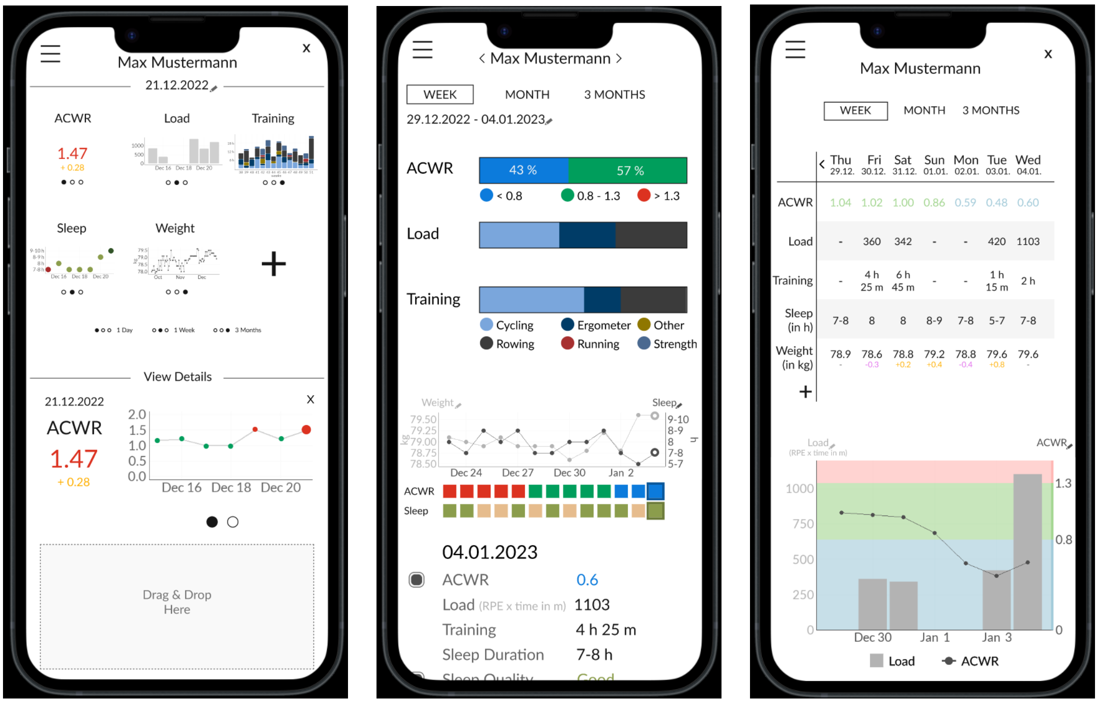
We conducted two pilot user tests with VIS experts. The purpose of those tests was to ensure the comprehensibility of the tasks and to get an estimate of the required time for one round of testing.

Throughout the pilot user tests, it became clear that interactive versions of the dashboard prototypes would considerably improve the testing experience. Therefore, we converted the static sketches with different views into basic interactive prototypes using Figma [27]. This way, users can perform their desired actions independently without instructing the person conducting the test to continue to the respective view. Other adjustments that were made between the pilot and the actual user tests included:

- Using genuine data for creating the dashboards
- Choosing colors that are easily distinguishable from one another
- Adding obvious hints to indicate interactivity
- Keeping all text in English instead of a mixture of English and German

In both trial runs, prototype A was rated best. While one VIS expert preferred prototype C over B, the other valued B more than C. To be exact, the second expert considered prototype C, which relied heavily on tables as a visualization technique, insufficient.

Screenshots of the low-fidelity prototypes are shown in Figure 7.1.



(a) Prototype A. (b) Prototype B. (c) Prototype C.

Figure 7.1: *Screenshots of the initial views of the three low-fidelity prototypes.*

7.2.2 Usability Testing of Low-Fidelity Prototypes

Five domain experts, two national team coaches, and three local rowing coaches participated in user tests of the first round of prototypes. The order in which the three prototypes were shown to the users was shuffled to reduce bias. The original plan was to schedule six user tests, testing three different sequences twice each. This way, we would still compare the same sequences while eliminating bias. However, we ended up with only five user tests. Thus, two users tested them in the sequence "A - B - C", and "C - B - A", respectively, while one user tested the prototypes in the sequence "B - C - A".

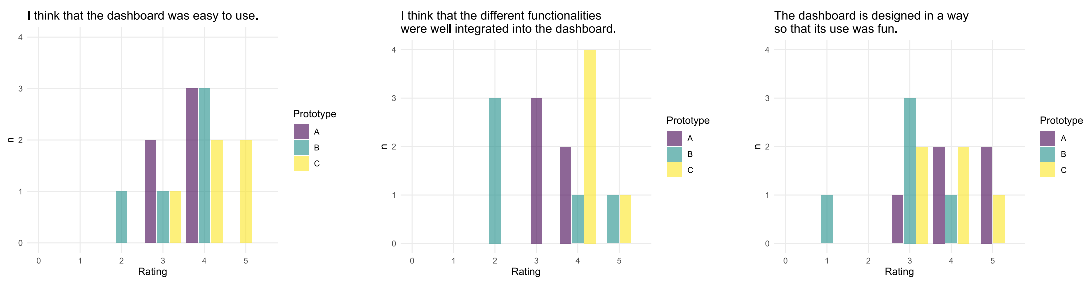
At the beginning of the test, they received a short introduction to the dashboard, which took at most five minutes. The usability was observed during the tests through a task-based approach [13], where the user had to perform predefined tasks on all three tested prototypes. In addition, they could explore the prototypes and were asked to talk about what they were trying to do. Moreover, the users were asked to provide feedback right away if something came to their mind to prevent them from forgetting anything. This took roughly 40 minutes.

After a user tested the prototypes, they at first had to rate the following statements concerning each of the three prototypes on a six-level Likert-type scale, ranging from "strongly disagree" to "strongly agree":

- I think that the dashboard was easy to use.
- I think that the different functionalities were well integrated into the dashboard.
- The dashboard is designed so that its use is fun.

We excluded a neutral rating to avoid the ambiguity of interpreting the midpoint response in Likert scales [24]. In addition, we wanted to prevent the participants from taking the easy way out by choosing a neutral option, which could mask their true opinion on the matter.

Figure 7.2 shows the results of the Likert-type evaluation. Overall, prototype C scored best, followed by prototype A. The users mainly favored those two instead of prototype B because of their cleanliness. Prototype B was described as too packed and unclear. This is directly reflected in the bar charts below. It affects usability in general, the ability of users to easily recognize and use all the functionalities, and their enjoyment.



(a) Ease of use. (b) Functionalities. (c) Enjoyment.

Figure 7.2: *Evaluation results of the three different low-fidelity prototypes regarding the Likert-type statements. 0 corresponds to "strongly disagree" and 5 to "strongly agree".*

They also had to put the prototypes in order of their preference, and on top of that, they assigned each prototype a rating on a scale from 1 to 10, where 10 constitutes the best rating. As expected, prototype C ranked first three out of five times overall, prototype A was placed second, and prototype B was placed last in all but one user test each. Figure 7.3 below indicates that more than half of the domain experts were pretty satisfied with prototype C, assigning it a rating of 8 out of 10. Prototype A received mediocre to decent ratings due to its design being too minimalist for comparing longer time spans.

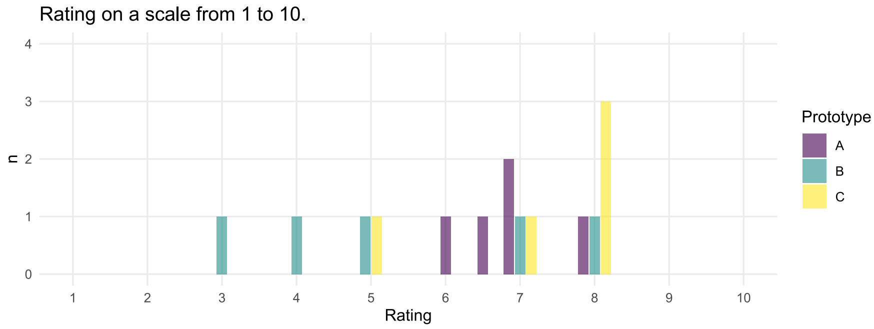


Figure 7.3: *Ratings on a scale from 1 to 10 assigned to the prototypes by the domain experts.*

The users' comments throughout the test and their ratings in the end were used to pinpoint issues and detect which features of the prototypes were functioning well. The daily view of prototype A was combined with the tables and plots of prototype C corresponding to weekly, monthly, and three-monthly data. Furthermore, the idea of vertically juxtaposing a heatmap-like graph depicting categorical variables and a line chart showing two numeric features was incorporated from prototype B.

7.2.3 Usability Testing of High-Fidelity Prototype

Five domain experts again performed the usability test of the high-fidelity prototype. However, while four out of the five coaches already participated in the usability testing of the low-fidelity prototypes, one of the local coaches was substituted with another local coach due to unavailability.

The beginning of the test was used for a short introduction, informing the users about the procedure. Then, the same tasks as during the first round of user tests were assigned, and again, exploring the prototype was possible. The testing of the dashboard took approximately 30 minutes.

The prototype was evaluated using the System Usability Scale (SUS) score [10]. The SUS measures user evaluation with ten Likert-based statements, which address usability aspects such as support, training, and complexity. Through alternating between positive and negative sentiment, the user is forced to consider the statements more thoughtfully when rating them on a scale ranging from "strongly agree" to "strongly disagree".

Figure 7.4 shows the results from the System Usability Scale evaluation, put into perspective through the comparison of acceptability ranges, a school grading scale, and adjective ratings, as proposed by Bangor et al. [7, 6].

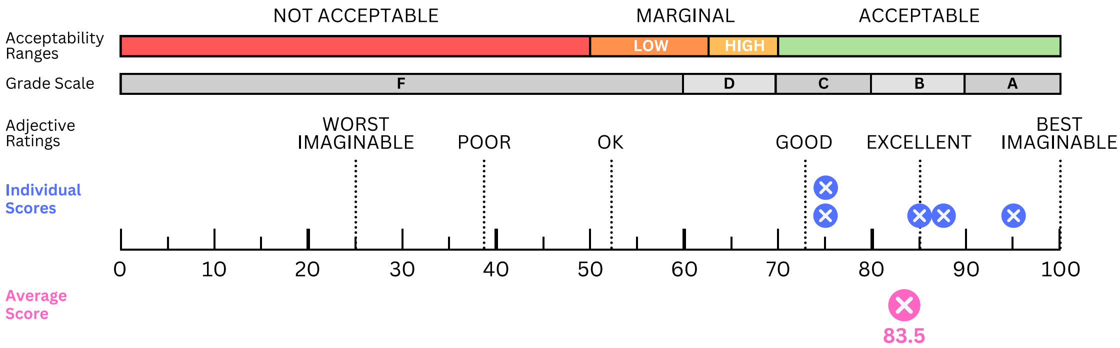


Figure 7.4: *SUS scores for the high-fidelity prototype of the coaching dashboard. The blue cross markers correspond to the individual scores of the users, while the slightly bigger pink cross marker represents the average score.*

The resulting SUS scores span from 75 to 95, with an average score of 83.5. According to a study by Bangor et al. [6], this constitutes acceptable usability ranging from good to excellent and even better according to the adjective rating. The statement that received the poorest rating was "I found the system very cumbersome/awkward to use". According to the testers' explanation, their main issues with the dashboard were technical. Firstly, the data took a very long time (several seconds to half a minute) to load. Secondly, one error message kept popping up on their screens occasionally. Moreover, two minor issues, which were only mentioned by one individual each, concerned a lack of knowledge. One domain expert stated they might need a short introduction to understand all functionalities fully. At the same time, the second explained that they needed some more time to get used to the data displayed since it was still unfamiliar.

8 Discussion

The following chapter presents a reflection of the project in its entirety, with a focus on the feedback provided by the domain experts. First, we reflect on the lessons learned with respect to Visual Design. Then, potential improvements to the rowing coach dashboard are discussed, followed by an analysis of the pitfalls encountered within the project's scope.

8.1 Lessons Learned

Perhaps one of the most difficult aspects was the limited screen space on which many different visual encodings had to fit. In order to fit the table, chart, and heatmap on an average cell phone screen, we decided to use a two-axis chart. While most literature discourages this, we tested it by including it in one of our low-fidelity prototypes. The domain experts liked the chart, which became integrated into the high-fidelity prototype. The users could all explore the data in the graph and interpret the values correctly. Therefore, we conclude that for visual encodings where space is limited, dual-axis graphs are sometimes a viable alternative to juxtaposed views. Additional space was saved by including drop-down lists for selecting variables for the respective visual encoding, which can be shown or hidden by clicking the appropriate buttons.

Another valuable lesson was learned regarding the table. To the eyes of a visualization expert, it looks crowded, especially the month and 3-month views display many tiny numbers on the small screens of cell phones. However, the domain experts appreciated being able to see the exact values of the data all at once and were able to gain valuable insights from it. This showed us that when using rapid prototyping, it is worth including a visual encoding in a low-fidelity prototype, even if it seems inadequate at first. Because low-fidelity prototypes can be easily adjusted or discarded, the extra time spent on them is well worth the effort if it results in a high-fidelity prototype that meets the needs of domain users.

What is evident to one person may be utterly unintuitive to another. Therefore, including icons such as the pencil or the information "i" to guide users through the dashboard's functionalities was essential. Both icons are widely used in digital applications and are, therefore, easily understood by everyone. While the pencil indicates that something can be edited, the information icon clearly suggests that there is more to discover and can be used in many cases.

8.2 Future Work

As described in section 7, some issues persist that still cause inconveniences when using the coaching dashboard. However, the evaluation highlighted that the most disruptive problems are of a technical nature and are not directly connected to the design or basic functionalities of the final prototype. Overall, the feedback from the domain experts was very positive. The current goal is to revise the implementation of the prototype in an upcoming scientific work and eliminate all bugs and other issues.

The suggested modifications below could help further improve the dashboard for rowing coaches and prepare it for deployment.

Linking chart and heatmap.
According to the low-fidelity prototypes, the goal would have been to link the chart and the heatmap. In particular, linked highlighting was planned, but a common tooltip between the two visual encodings was also planned to show details of all selected variables on demand. However, due to programming difficulties, this has not yet been implemented.

Storing user preferences.
Currently, some default variables are selected at the beginning of every session. We determined those variables during the precondition phase of the project based on the importance conveyed in the interviews. However, these features do not necessarily provide the most valuable insights for all coaches. Some coaches might want to explore more features or simply different ones. Those coaches then have to change the default selection every time after logging into the app. Saving the selection at the end of the session and using this preference for the new one will likely enhance the user experience.

More accurate data.
Early in the project, some athlete data, e.g., sleep duration or resting heart rate, was collected based on the athlete's perception. However, several wearables that record this data are currently in use. The data is then retrieved from the tracking devices and enables the usage of more detailed data.

Data Security.
The athlete data is sensitive. In particular, personal health information must be safeguarded against unauthorized access or disclosure. Therefore, specific security standards must be complied with by the app before it can be deployed and used on the personal devices of the rowing coaches.

Fixing bugs.
One bug in particular keeps popping up. However, closing the error message without any other noticeable impacts is possible. Therefore, the bug was not a high priority, and we could not eliminate it due to time issues. Interestingly, the bug was encountered a lot more frequently on iOS devices. Overall, all the problems

due to bugs in the code need to be fixed.

Reworking the database connection.
The current implementation setup is not ideal. Retrieving data from the database takes a long time, especially for the "3 months" view. It takes several seconds for the table and graphs to be generated. The long time lag is cumbersome and drastically reduces the user experience. Finding faster ways to retrieve the data is essential to improve usability.

8.3 Encountered Pitfalls

Several pitfalls were encountered during the study. The following section describes the challenges faced and compares them to common pitfalls identified by Sedlmair et al. [45].

Evolving data.
The data for our design study was provided by the "AIROW - Artificial Intelligence in Rowing" project [1]. Since the data collection had just started at the beginning of our design study, very little data was available initially. As the study progressed, more and more data was collected. However, the data format changed in some cases, requiring the prototypes to be adapted. Also, for some variables, the data has already changed to more granular data, so the final prototype is not current and must be adapted before deployment. This is comparable to the pitfall *PF-4* identified by Sedlmair et al. [45]. In our case, there was enough data to start the study. However, the changes to the data format made from time to time were because the "real" data was not yet available. It might have been better to start our work after the data format was finalized, but the stakeholders wanted our dashboard created as soon as possible due to time constraints.

Not so rapid prototyping.
Another pitfall discussed in the paper by Sedlmair et al. [45] is non-rapid prototyping (*PF-22*). The goal was to use rapid prototyping techniques, starting with hand-drawn prototypes. However, the initial prototyping phase took longer than necessary. Too much time was spent on details of initial ideas that were eventually discarded. It would have been helpful to set a time limit to keep this phase to a minimum.

9 Conclusion

This thesis presents a custom dashboard tailored for rowing coaches to optimize athlete performance analysis and subsequent training. Developed in collaboration with professional rowing coaches, it allows for detailed tracking and analysis of rowing-specific training and well-being data.

Through interviews with national team coaches, we identified the features that were most important to them in terms of training and well-being data. They were most interested in ACWR, training load, training duration, weight, sleep quantity, sleep quality, illness, and injury. During the interviews, we also found that they cared about daily, weekly, monthly, and 3-month views. In addition, we identified six tasks during the requirement analysis: selecting an athlete, evaluating training content, training extent and state of health, recognizing correlations, and analyzing progress. No existing solution allows us to perform all these tasks based on the data collected from the rowing athletes. With our dashboard, rowing coaches can accomplish all of these tasks. It helps them make informed decisions by allowing them to explore and analyze all the training and well-being data provided by their athletes.

The results of our evaluation emphasized the importance of involving domain users in the design process since the opinions of visualization experts and domain users can be very different, as in the case of the table. However, during the final usability tests, it became clear that some things in the dashboard still need improvement. While most of the problems are related to technical issues, it will only be possible to deploy the dashboard once these issues have been resolved. Incorporating data security measures and fixing persistent bugs are the least that should be done. Other improvements include the ability to save user preferences regarding the pre-selected variables displayed, and the link between the chart and the heatmap.

Since the overall feedback was very good in terms of usability, our conclusion is that the dashboard can definitely be an asset for professional rowing coaches. A particularly strong foundation is provided by the involvement of national team coaches. As the project focuses on rowing trainers only, the results are limited to this field. However, the experience gained throughout the process could prove valuable for designing similar coaching dashboards in other sports.

Bibliography

[1] Airow - artificial intelligence in rowing. https://airow.univie.ac.at/. Accessed: 2023-08-28.

[2] Apexcharts. https://apexcharts.com/. Accessed: 2023-08-28.

[3] React. https://react.dev/. Accessed: 2023-08-28.

[4] Benjamin Bach, Euan Freeman, Alfie Abdul-Rahman, Cagatay Turkay, Saiful Khan, Yulei Fan, and Min Chen. Dashboard design patterns. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):342–352, 2022.

[5] Benjamin Bach, Nathalie Henry Riche, Sheelagh Carpendale, and Hanspeter Pfister. The emerging genre of data comics. *IEEE computer graphics and applications*, 37(3):6–13, 2017.

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

[7] Aaron Bangor, Philip T Kortum, and James T Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

[8] Paul Barratt. Healthy competition: A qualitative study investigating persuasive technologies and the gamification of cycling. *Health & place*, 46:328–336, 2017.

[9] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. D^3 data-driven documents. *IEEE transactions on visualization and computer graphics*, 17(12):2301–2309, 2011.

[10] John Brooke. Sus: a "quick and dirty'usability. *Usability evaluation in industry*, 189(3):189–194, 1996.

[11] Canva. Canva. https://www.canva.com/, 2023. Accessed: 2023-12-06.

[12] Zoom Video Communications. Zoom. https://zoom.us/, 2023. Accessed: 2023-05-20.

[13] Michael D Corry, Theodore W Frick, and Lisa Hansen. User-centered design and usability testing of a web site: An illustrative case study. *Educational technology research and development*, 45(4):65–76, 1997.

[14] Meng Du and Xiaoru Yuan. A survey of competitive sports data visualization and visual analysis. *Journal of Visualization*, 24:47–67, 2021.

[15] Lukas Feierl, Torsten Möller, and Peter Luidolt. Sunscreen: Visual fault detection for solar-thermal systems. *IEEE Computer Graphics and Applications*, 2023.

[16] Stephen Few. *Information dashboard design: The effective visual communication of data*. O'Reilly Media, Inc., 2006.

[17] Stephen Few and Perceptual Edge. Dashboard confusion revisited. *Perceptual Edge*, pages 1–6, 2007.

[18] Stephen Few and Perceptual Edge. Dual-scaled axes in graphs. *Retrieved April*, 25:2010, 2008.

[19] CARL Foster. Monitoring training in athletes with reference to overtraining syndrome. *Medicine and science in sports and exercise*, 30(7):1164–1168, 1998.

[20] Aristea Fotopoulou and Kate O'Riordan. Training to self-care: Fitness tracking, biopedagogy and the healthy consumer. *Health Sociology Review*, 26(1):54–68, 2017.

[21] Tim J Gabbett. The training—injury prevention paradox: should athletes be training smarter and harder? *British journal of sports medicine*, 50(5):273–280, 2016.

[22] Jos Goudsmit, Ruby T. A. Otter, Inge Stoter, Berry van Holland, Stephan van der Zwaard, Johan de Jong, and Steven Vos. Co-operative design of a coach dashboard for training monitoring and feedback. *Sensors*, 22(23), 2022.

[23] Shona L Halson. Monitoring training load to understand fatigue in athletes. *Sports medicine*, 44(Suppl 2):139–147, 2014.

[24] David R Hodge and David Gillespie. Phrase completions: An alternative to likert scales. *Social Work Research*, 27(1):45–55, 2003.

[25] Billy T Hulin, Tim J Gabbett, Daniel W Lawson, Peter Caputi, and John A Sampson. The acute: chronic workload ratio predicts injury: high chronic workload may decrease injury risk in elite rugby league players. *British journal of sports medicine*, 50(4):231–236, 2016.

[26] Elena Iannucci, Zhutian Chen, Iro Armeni, Marc Pollefeys, Hanspeter Pfister, and Johanna Beyer. Arrow: A real-time ar rowing coach. 2023.

[27] Figma Inc. Figma. `https://www.figma.com/`, 2023. Accessed: 2023-05-20.

[28] Fitbit Inc. Fitbit. `https://www.fitbit.com/global/at/home`, 2023. Accessed: 2023-06-18.

[29] Strava Inc. Strava. `https://www.strava.com/`, 2023. Accessed: 2023-06-18.

[30] Andrea Janes, Alberto Sillitti, and Giancarlo Succi. Effective dashboard design. *Cutter IT Journal*, 26(1):17–24, 2013.

[31] Marcos Lage, Jorge Piazentin Ono, Daniel Cervone, Justin Chiang, Carlos Dietrich, and Claudio T Silva. Statcast dashboard: Exploration of spatiotemporal baseball data. *IEEE computer graphics and applications*, 36(5):28–37, 2016.

[32] Garmin Ltd. Garmin connect. `https://connect.garmin.com/`, 2023. Accessed: 2023-06-18.

[33] Francois Modave, Jiang Bian, Trevor Leavitt, Jennifer Bromwell, Charles Iii, and Heather Vincent. Low quality of free coaching apps with respect to the american college of sports medicine guidelines: A review of current mobile apps. *JMIR mHealth and uHealth*, 3:e77, 07 2015.

[34] Iñigo Mujika. Quantification of training and competition loads in endurance sports: methods and applications. *International journal of sports physiology and performance*, 12(s2):S2–9, 2017.

[35] Tamara Munzner. *Visualization analysis and design*. CRC press, 2014.

[36] Michael Oppermann and Tamara Munzner. Data-first visualization design studies. In *2020 IEEE Workshop on Evaluation and Beyond-Methodological Approaches to Visualization (BELIV)*, pages 74–80. IEEE, 2020.

[37] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. Soccerstories: A kick-off for visual soccer analysis. *IEEE transactions on visualization and computer graphics*, 19(12):2506–2515, 2013.

[38] Charles Perin, Romain Vuillemot, Charles D Stolper, John T Stasko, Jo Wood, and Sheelagh Carpendale. State of the art of sports data visualization. In *Computer Graphics Forum*, volume 37, pages 663–686. Wiley Online Library, 2018.

[39] Hannah Pileggi, Charles D Stolper, J Michael Boyle, and John T Stasko. Snapshot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2819–2828, 2012.

[40] Tom Polk, Jing Yang, Yueqi Hu, and Ye Zhao. Tennivis: Visualization for tennis match analysis. *IEEE transactions on visualization and computer graphics*, 20(12):2339–2348, 2014.

[41] Posit team. *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA, 2023.

[42] Damian J Rivers. Strava as a discursive field of practice: Technological affordances and mediated cycling motivations. *Discourse, Context & Media*, 34:100345, 2020.

[43] Alper Sarikaya, Michael Correll, Lyn Bartram, Melanie Tory, and Danyel Fisher. What do we talk about when we talk about dashboards? *IEEE transactions on visualization and computer graphics*, 25(1):682–692, 2018.

[44] Barret Schloerke and Jeff Allen. *plumber: An API Generator for R*, 2023. https://www.rplumber.io, https://github.com/rstudio/plumber.

[45] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics*, 18(12):2431–2440, 2012.

[46] Manuel Stein, Halldór Janetzko, Thorsten Breitkreutz, Daniel Seebacher, Tobias Schreck, Michael Grossniklaus, Iain D Couzin, and Daniel A Keim. Director's cut: Analysis and annotation of soccer matches. *IEEE computer graphics and applications*, 36(5):50–60, 2016.

[47] WHOOP. Whoop. `https://www.whoop.com/`, 2023. Accessed: 2023-08-24.

[48] Yingcai Wu, Ji Lan, Xinhuan Shu, Chenyang Ji, Kejian Zhao, Jiachen Wang, and Hui Zhang. ittvis: Interactive visualization of table tennis data. *IEEE transactions on visualization and computer graphics*, 24(1):709–718, 2017.

[49] Ogan M Yigitbasioglu and Oana Velcu. A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, 13(1):41–59, 2012.