Is it just a score? Understanding Training Load Management Practices Beyond Sports Tracking

Armağan Karahanoğlu
University of Twente, Interaction Design Research Group, Enschede, Netherlands
a.karahanoglu@utwente.nl

Aykut Coskun
Koc University, Arcelik Research Center for Creative Industries, Istanbul, Turkey
aykutcoskun@ku.edu.tr

Dees Postma
University of Twente, Human Media Interaction Dept., Enschede, Netherlands
d.b.w.postma@utwente.nl

Bouke Leonard Scheltinga
University of Twente, Biomedical Signals and Systems Dept., Enschede, Netherlands
b.l.scheltinga@utwente.nl

Rúben Gouveia
University of Twente, Interaction Design Research Group, Enschede, Netherlands
r.h.gouveia@utwente.nl

Dennis Reidsma
University of Twente, Human Media Interaction Dept., Enschede, Netherlands
d.reidsma@utwente.nl

Jasper Reenalda
University of Twente, Biomedical Signals and Systems Dept., Enschede, Netherlands
j.reenalda@utwente.nl

ABSTRACT

Training Load Management (TLM) is crucial for achieving optimal athletic performance and preventing chronic sports injuries. Current sports trackers provide runners with data to manage their training load. However, little is known about the extent and the way sports trackers are used for TLM. We conducted a survey (*N*=249) and interviews (*N*=24) with runners to understand sports tracker use in TLM practices. We found that runners possess some understanding of training load and generally trust their trackers to provide accurate training load-related data. Still, they hesitate to strictly follow trackers' suggestions in managing their training load, often relying on their intuitions and body signals to determine and adapt training plans. Our findings contribute to SportsHCI research by shedding light on how sports trackers are incorporated into TLM practices and providing implications for developing trackers that better support runners in managing their training load.

CCS CONCEPTS

• **Human-centered computing** → Empirical studies in HCI.

KEYWORDS

Training load management, SportsHCI, running, human-data interaction, sports tracking, personal informatics

ACM Reference Format:
Armağan Karahanoğlu, Aykut Coskun, Dees Postma, Bouke Leonard Scheltinga, Rúben Gouveia, Dennis Reidsma, and Jasper Reenalda. 2024. Is it just a score? Understanding Training Load Management Practices Beyond Sports Tracking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3613904.3642051

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642051

1 INTRODUCTION

Ongoing developments in mobile apps and smart watches have enabled runners to track various sports performance-related measures [61]. Sports tracking apps like *Run Keeper*[1] and *Map My Run*[2] provide insights into running performance, motivate runners to go for a run, and challenge them to beat their personal best times [74]. Accurate [76] and advanced [54] data becomes critical for many runners in pursuing their running-related goals. Most runners rely on metrics to monitor their performance [46]; some are *measured* metrics, such as distance and time of a workout, while others are *derived* metrics, such as VO_2Max, training frequency, or "suffer score". Tracking and presenting such metrics, sports trackers could help runners with Training Load Management (TLM).

Monitoring and managing the training load requires more knowledge and skill than checking performance metrics [31]. Specific knowledge about the underlying biomechanical and physiological principles of sports training is vital to make sense of the performance measures and maximise the adaptation of an athlete's body to the training load [41]. It might not be evident to runners that measures such as VO_2Max (i.e., the estimate of maximum oxygen intake during exercise) and step frequency depend on one's body characteristics [75]. Such data does not always help improve performance, especially when metrics are not in line with the runner's expectations (e.g., no improvement in VO_2Max despite regular training).

[1] https://runkeeper.com/ (Retrieved on 19 August 2023)
[2] https://www.mapmyrun.com (Retrieved on 19 August 2023)

To effectively use sports trackers in managing training load, runners must transform the data into meaningful narratives and interpretations [23]. Data interpretation is an intricate task involving data collection and organisation, identification of patterns in data, and extraction of meaningful insights [80]. This process can be challenging for runners without sufficient sports training knowledge, potentially leading to inappropriate training load planning. As a general guideline, it is not recommended to increase the training load by more than 10% per week in terms of duration, intensity, or frequency to avoid overloading [70]. Increasing load by over 30% per week, especially for novice runners, raises the risk of running-related injuries [64]. Scenarios such as exceeding the body's capacity with excessively high weekly running distances or having excessive rest between running workouts [34] pose problems, as unsuitable training loads, whether too high or too low, can increase the risk of running-related injuries [1, 83].

Runners have different attitudes towards running performance, which affect how they interpret and interact with running related data. Sensemaking of such data may result in complying with, negotiating with, or ignoring it [17, 18]. These behaviours can vary across data types, with individuals tending to focus more on one favourite type of performance metric while intentionally overlooking unsatisfactory performance measures [17]. Snooks et al., [82] argue that fixating on specific metrics might cause more harm than benefit when people change their behaviour to fulfil sports trackers' expectations. In the long term, this approach may foster an unhealthy obsession with athletic performance and "being emotionally invested" in hitting the numbers [63] rather than a fundamental understanding of the underlying meaning of numbers.

Designing trackers to help runners monitor and understand data for better TLM practices requires sports science and SportsHCI knowledge. Previous work in sports science articulates the underlying dimensions of training load [e.g., 36], how to measure training load by utilising various external (e.g., distance) and internal training load metrics (e.g., heart rate) [e.g., 41], and argues for the role of mental factors [e.g., 22]. In HCI, studies showed the contribution of technology to measure physical metrics by employing wearables that quantify training load measures (e.g., biomechanical and physiological data) [68]. In addition, several studies presented guidelines on how to convey physical activity data to users in a clear [27], engaging [71] and appealing way [30].

Yet, runners have their way of assessing the accuracy of running related data that starts from data interpretation and goes beyond clear and engaging data representations [67]. Furthermore, the primary focus of current sports trackers is more on accurately representing individual metrics (e.g., VO_2Max) than how they relate to each other, which may hinder runners' ability to obtain actionable insights from their data. Finally, in HCI literature, little is known about how runners use running-related data for TLM, how and what they track for managing their training loads and to what extent they comply with the TLM suggestions of trackers.

Recently, the work of Rapp and colleagues [72] provided insightful contributions to our understanding of the use of personal informatics (PI) in sports. They discovered that while amateur athletes (i.e., individuals who are highly trained but not competitive at national /international level) tend to trust the objectivity of the parameters monitored by their devices, elite athletes (i.e., individuals who currently compete at national and international level with national titles [85]) often rely more on their sensations. They also identified a significant gap in the use of PI tools among amateur athletes, arguing that these tools do not adequately guide them in understanding and interpreting their data. This finding is particularly relevant to our research. We argue that misinterpretations or misuses of training data can have significant negative consequences for runners, such as increased risk of injury, and better technology design could alleviate these problems [93]. Hence, TLM is not just a matter of optimizing performance in running but is also crucial for injury prevention in other sports, such as cycling [89], swimming [5], and team sports [59]. Understanding how runners manage their training load more profoundly could help design sports trackers to better guide TLM, and thus help avoid undesirable outcomes. Such an understanding is missing in existing SportsHCI studies.

In this paper, we address this gap through a survey and interview study answering the following questions: (1) How do runners use their sports data for TLM? and (2) how do they perceive and use sports trackers' TLM suggestions? We first explain TLM in sports and illustrate how TLM practices are supported in sports informatics. Then, we report the results of our studies to articulate runners' sports tracker use for TLM purposes. Our findings contribute to SportsHCI research by understanding runners' sports tracking practices in TLM. Through the survey and interviews with runners, we shed light on (1) how runners decide on and adapt their training program with data and guidance provided by sports technology and (2) how this data and guidance help them develop competency in TLM. This understanding allows us to identify implications for future tracking tools that support TLM, contributing towards better designs that address runners' TLM-related needs. We also give novel directions for sports science regarding identifying, presenting, and interpreting data, which will help runners get optimal benefits from their workouts (e.g., performance enhancement and sustained enjoyment).

Furthermore, we inform the development of smarter and personalised sports tracking that support athletes' performance and health, which can go beyond the running context. Such an understanding can lead to developing more integrated sports trackers for TLM and pave the way for more personalised use of sports technology across various sports. Therefore, this paper has the potential to inform not only the field of HCI and sports technology but also the broader domain of sports science and athlete health management.

2 RELATED WORK

2.1 Training Load Management in Sports

Training load is the workload that relates to physical training [10]. It can be either internal or external, where the former refers to the physiological response of the athlete to training, and the latter refers to the physical demands imposed on the athlete [41]. The internal load can be quantified, for example, in terms of the heart rate (HR), blood oxygenation rate, and lactate levels and the external load in terms of, for example, the intensity, duration, and frequency of workouts. Each workout stimulates long-term responses (e.g., increase in VO_2Max) and short-term physiological adaptations (e.g., Excess Post-exercise Oxygen Consumption, EPOC) [50]. Workload is a highly dynamic, individual, and multidimensional construct that a single metric cannot capture.

It is comparatively easier to quantify external load than internal load. One way is to calculate the acute-chronic workload ratio (ACWR) [19, 42]. The acute phase is an athlete's most recent training load, and the chronic phase is the training load the athlete's body is used to and prepared for [33]. ACWR employs several performance measures, such as training intensity and volume (or amount) of training [32]. Some of these are easy to measure and track with current sports trackers. For example, training intensity can be differentiated by looking into a runner's metrics of jogging and sprinting. During running, runners' HR (i.e., internal load) varies as a reaction to factors such as running speed and duration (external load). Hence, HR data [50] across a workout [31, 48], combined with the duration of the exercise, become the input to calculate training intensity [25, 29]. However, athletes' subjective perception of the training intensity is also crucial for accurate internal training load measures, as in the rating of perceived effort (RPE) [5], which is a reliable and common way to quantify and assess internal load [17, 40]. Sports trackers can collect this input by asking how the athlete feels about the intensity of a past workout and calculate ACWR accordingly.

An essential aspect of the ACWR is the definition of a "sweet spot", a ratio between 0.8 and 1.3 (which can be up to 2.5), where the risk for a chronic injury is low and athletic performance can still be enhanced [42]. An athlete can gradually progress to high chronic workloads by consistently implementing training workloads within this sweet spot range. It can be assumed that continuously increasing the training intensity (e.g., running faster), frequency (e.g., running more often) or volume (e.g., running farther) can enhance performance. However, training physiology principles dictate proper resting time before the next training to avoid physiological overload and help recover muscle and tissue damage [69]. Without sufficient rest (i.e., the acute-chronic workload ratio is higher than 1.5), the athlete's body cannot recover from the previous training, increasing the risk of accumulation of tissue damage [7, 32, 69].

Another strategy to manage training load involves integrating the periodisation philosophy of sports training into training plans and providing more flexible micro (i.e. daily), meso (i.e. monthly) and macro (i.e., yearly/life-long) training cycles [8]. A comprehensive macro training cycle is subdivided into meso and micro cycles, where the intensity and the volume of the training sessions are defined based on the athletes' response to micro and meso training cycles. These data-based, personalised cycles can make the training plans fit the athletes' daily routines and effectively mitigate the possibility of overloading to sustain life-long activity [47].

Sports periodisation literature emphasises the significance of adequate recovery in preventing overuse injuries and overtraining syndrome [60, 83]. Monitoring training load emerges as a significant method to avoid such setbacks. Even though preventable with proper TLM [57, 68], overuse-related lower extremity injuries are common among runners [40, 65]. Around 37%–56% of runners have running related overuse injuries yearly [53, 91], and numbers still increase [92, 94]. These facts bring us to the importance of monitoring and managing training loads to avoid the overloading effects of training and mitigate the overuse-related injury risk factors [6, 39, 95]. We think that sports trackers can help in avoiding such overloading effects.

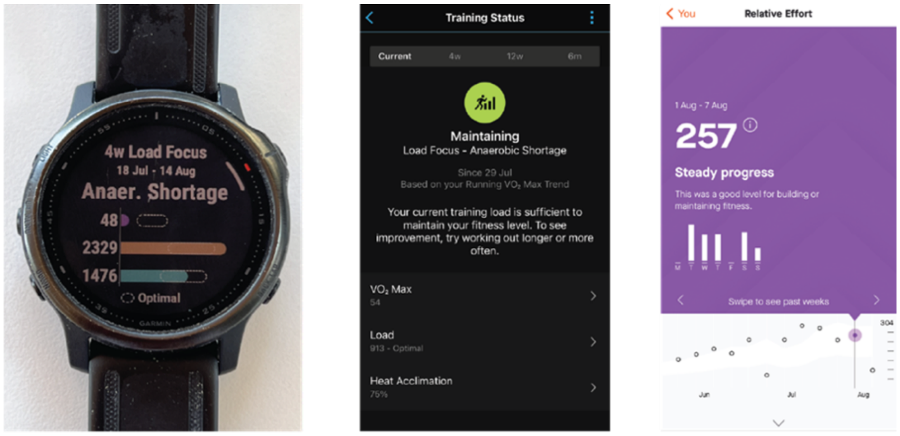
2.2 Training Load Monitoring Practices in Sports Informatics

The primary driving force behind sports technology development is enhancing metric accuracy for predicting human performance [86]. Consequently, current sports trackers offer athletes an immense amount of data [90]. These tools can collect, analyse and synthesise performance-specific data and give individualised feedback and recommendations [35, 56, 74] supporting TLM practices of runners [84], particularly for those who plan the details of their own training. One example is the Garmin sports trackers, which combine HR data with EPOC. By summing the athlete's oxygen consumption measurements over the past seven days and comparing them with the runners' four-week Chronic Training Load[3], these trackers calculate Acute Training Load. Subsequently, they provide the runners with a four-week training load focus (Figure 1a) and illustrate what range the athlete should be training. In the provided example (Figure 1a), it is evident that the runner is underloading the "anaerobic" training, and the tracker shows that the athlete is missing a particular type of training (Figure 1a, purple coloured) and overloading in another (Figure 1a, orange coloured). Within the app interface (Figure 1b), the tracker provides more comprehensive information regarding performance measures and training load and their impact on performance enhancement.

Another sports tracker series, Suunto, uses the Training Stress Score to quantify training load[4]. This score is based on the training impulse, which uses the intensity and duration of the workouts, together with HR data and the runner's pace, to calculate short-term and long-term training loads. The short-term training load is referred to as Acute Training Load and is a 7-day average of the training stress score, while the long-term load is referred to as Chronic Training Load (or fitness), a 42-day weighted average of the training stress score[5]. Running tracking apps, like Strava, Run Keeper, Adidas Running, Map My Run, and Train As One[6], provide limited information about training load to freemium users. In contrast, premium users can access more advanced features, like training plans, based on their training load measures. For example, Strava app (Figure 1c) provides a graphical representation of training load and signals how the relative effort of an athlete changes over time. Training and performance enhancement-focused platforms like Training Peaks[7] also provide advanced features (e.g., peak performance analysis), metrics (e.g., training stress score) and training customisation opportunities.

Although current sports trackers use training knowledge to calculate training load, they do not immediately provide actionable insights about how this information could be implemented for TLM.

[3]https://discover.garmin.com/en-US/performance-data/running/#training-load (Retrieved on 24 July 2023)
[4]https://www.suunto.com/sports/News-Articles-container-page/training-stress-score-in-suunto-app (Retrieved on 24 July 2023)
[5]https://www.suunto.com/sports/News-Articles-container-page/understand-and-manage-your-training-load-with-suunto (Retrieved on 24 July 2023)
[6]https://www.strava.com/features ; https://runkeeper.com ; https://www.runtastic.com ; https://www.mapmyrun.com ; https://www.trainasone.com
[7]https://www.trainingpeaks.com (Retrieved on 23 July 2023)



(a) Garmin sports watch interface (b) Garmin app interface (c) Strava app interface

Figure 1: Example Training Load Information from Garmin and Strava

Furthermore, interpreting training load measurements is complex due to the advanced sports training knowledge [11]. It requires significant effort to reflect on and act upon [46, 86]. Tracker data must be distilled with sports science knowledge to arrive at interpretations about performance and injury proneness [88].

Previous HCI work focused on understanding runners' practices of reflection-on-training through interactive systems. For example, an interactive mirror was employed in a recent study to translate complex training data into qualitative interfaces and understand the negotiation styles runners prefer to achieve life-training balance [77]. Still, HCI researchers should be aware that even though runners know the risks of training overload [55], the majority believe that they already know the limits of their bodies, and they tend not to limit excessive training [81], which can be detrimental due to the reasons explained above. Thus, informing the design of sports trackers, which can successfully guide runners in determining the *sweet spot* for training load, becomes a relevant goal for SportsHCI research.

The importance of TLM stems from its role in balancing athletes' training efforts and recovery time and predicting sports performance and risk of overuse (or overloading) related sports injuries [24, 32, 83]. In addition, it can help prescribe individualised training plans and revise and adjust training schedules that can boost athletes' well-being [48]. Although professional athletes' TLM is done under the supervision of coaches, amateur athletes may not be able to work with a coach, resulting in less supported and informed self-coaching and understanding of sports data [72]. As we stated earlier, TLM requires competency and knowledge. Having sports trackers that support and empower runners is vital to managing the training load effectively. In that regard, we first need to understand runners' current TLM practices, how they track and interpret data from their runs, especially for TLM purposes, and to what extent they follow TLM-related suggestions from trackers. Such an inquiry follows up on the findings around the use of personal informatics in sports [54, 72] and data in running [46, 67]. It opens new research directions in SportsHCI, contributing to the design of sports informatics, specifically for performance enhancement and injury prevention.

3 METHODS

We carried out an online survey followed by in-depth interviews to address (1) how runners use sports trackers when managing their training load (TLM) and (2) how they incorporate TLM-related tracker recommendations and performance predictions into their training. The survey gave us insights into why runners track their runs, particularly trackers' roles in managing training load. However, even after the survey, certain aspects remained unclear, such as specific ways runners integrate trackers into their TLM practices and how they make sense of and act upon TLM-related suggestions provided by trackers. Interviews helped us clarify these aspects and gain in-depth insights into technology-supported TLM practices. We obtained ethical approval from our research institute before participant recruitment.

3.1 Survey Study

Our survey questions drew on literature from HCI [37, 43, 44, 96] and sports science studies [e.g., 31, 36, 41, 49, 72, 74]. All authors, with a background in interaction design, HCI and sports science, reviewed the questions multiple times to arrive at the final form of the survey by checking their appropriateness for addressing the research questions. The survey asked participants their (1) demographics and running history; (2) tracking habits and tracking history; (3) motivations for tracking their runs [adapted from 37]; (4) sources they get help from while setting their running workouts (e.g., coach, technology or experiences runners); (5) reasons for changes in training routines; (6) metrics used to assess training load; and (7) trust in technology within the context of TLM (see Table 1).

Our questions included a combination of open-ended (e.g., age, years of experience in running), single-choice (e.g., running frequency per week) and five-point Likert and rating scale questions (e.g., importance of running metrics for TLM rates on a scale of 1=

Table 1: Questions of the Online Survey

Item	Question	Type	Way of Data Collection
1-2	Age, running and tracking experience, weekly running volume	Open-ended (text box)	Numeric
1	Gender	Closed (single choice)	2 choices (female/male)
1	Running frequency	Closed (single choice)	7 choices (from 1-7 days)
1	Running injury history, training with a trainer	Closed (single choice)	Yes/No
2	Trackers being used	Open-ended (text box)	Text (brand of the sports watch)
2	Self-evaluation of running and TLM experience	Likert Scale (agreement)	2 items (i.e., I am experienced in running/managing my training load)
3	Motivation in tracking	Likert Scale (agreement)	10 items as delineated in [37]
4	Runners' sources of running workouts	Likert Scale (frequency)	6 questions about the way runners plan their running workouts
5	Reasons for changes in training routines	Likert Scale (agreement)	6 questions about the role of trackers in changing routines
6	Metrics to assess training load	Rating Scale (importance)	9 running-related data types that are important for TLM
7	Trust in Technology in TLM	Likert Scale (agreement)	3 questions (training metrics, training load, recovery time)

not at all to 5=very important). We ended the survey by asking respondents to leave their email addresses for follow-up interviews.

3.2 Interview Study

The semi-structured interviews consisted of three sets of questions built upon our survey questions. The first set of questions delved into runners' present motivations for tracking their runs (as was also asked in the survey; see Table 1, Item 3), including their uses and practices with their currently used sports trackers (following up on the survey questions listed Table 1, Item 2). The second set focused on the types of data used in running and TLM (Table 1, Item 6). Lastly, the third set of questions investigated the role of technology (e.g., smartwatches, apps, or other mediums) in runners' TLM and why they trust technology in TLM (Table 1, Item 4). This set also probed participants' trust in trackers' running-related calculations (e.g., VO_2Max calculations) and predictions (e.g., recovery time), commonly used to signal runners' training load (Table 1, Item 7). In short, interview questions were complementary to the survey questions. For example, in the survey, we asked participants why they changed their running routines, and in the interviews, we expanded on why and how these changes occur.

3.3 Participant Recruitment

Our data collection occurred over 4.5 months, between 1 March and 15 July 2023. We aimed to recruit runners with various running, tracking and TLM experience levels. To be included in our study, participants had to run at least twice a week over the last six months and monitor their running workouts with a smartwatch. We shared the online survey link via social platforms of a local marathon event, local running clubs, personal contacts and other social networks like Strava and Instagram. Our survey went online on the JotForm platform and was closed once no responses were received for five consecutive days. In total, we received 263 replies, of which we removed 14 due to respondents' lack of running experience (*N*=2), double entries (*N*=2), and not using a sports watch for tracking runs (*N*=10). The final number of responses was 249 (Table 2).

Survey respondents consisted of 34% (*N*=85) female and 66% (*N*=164) male runners, with an average age of 42.25 (*SD*=11.86). They were running for 10.89 years (*SD*=8.79), on average 3.26 days a week (*SD*=1.11), and 37.44 kilometres per week (*SD*=20.88). Of the respondents, 91% (*N*=227) were not injured by the time of the survey, but most (*N*=181, 73%) had some type of running related injuries before. These distributions are compatible with earlier large-scale running related studies [e.g., 21, 44]. On average, participants tracked their runs for 6.52 years (*SD*=4.70). The majority used a Garmin sports watch (*N*=179, 72%), followed by Apple (*N*=26, 10%) and Polar (*N*=26, 10%) to track their runs. About half (*N*=104, 42%) were following a plan from a running coach, while 145 (58%) did not follow a training plan from a running coach. Most runners thought they were above average experienced in running (*M*=3.69, *SD*=1.01) and in control of their training load (*M*=3.55, *SD*=0.96).

We reached out to runners again for the interview study. We informed them that the interviews would be recorded and conducted in English, and the recordings would be deleted after interview transcription. Of the 112 participants who received an email invitation for the interviews, 28 responded positively. Ultimately, we scheduled an interview with 24 runners who all submitted the survey (Table 3). Two authors interviewed each participant through the online video conferencing platform Microsoft Teams. All sessions were recorded. Prior to the recording, participants gave consent for the interviews to be recorded. The interviews took, on average, 45 minutes.

Interview participants comprised 9 female (38%) and 15 male (62%) runners, a ratio similar to our initial sample in the survey. They had on average 41.25 years of age (*SD*=11.83, Min =20, Max=70), were running for 12.17 years (*SD*=9.96, Min=2.5

Table 2: Demographics of the Survey Participants

	M	SD	Min	Max
Age	42.25	11.86	18	70
Running experience (in years)	10.89	8.79	0.5	50
Tracking experience (in years)	6.52	4.70	0.25	35
Running frequency (in days/per week)	3.26	1.11	1	7
Weekly running volume (in kilometres)	37.44	20.88	5	140
Experience in running (rating scale)	3.69	1.01	1	5
Confidence in controlling training load (rating scale)	3.55	0.96	1	5

Table 3: Characteristics of the Interview Study Participants

	Demographics			Experience (in years)		Running Frequency(in days)	Total weekly distance(in km)
	Age	Gender	Watch brand	Running	Tracking		
P01	41	Male	Garmin	4	4	4	50
P02	38	Male	Garmin	19	3	1	10
P03	46	Male	Garmin	10	10	5	60
P04	48	Male	Garmin	5	4.5	7	140
P05	39	Male	Garmin	7.5	7	5	78
P06	40	Male	Garmin	24	6	3	40
P07	32	Male	Garmin	4	4	3	25
P08	41	Male	Garmin	8	6	2	20
P09	44	Male	Garmin	6	6	4	40
P10	70	Female	Coros	45	8	5	40
P11	48	Female	Garmin	7	7	4	35
P12	58	Male	Polar	15	12	2	20
P13	56	Female	Garmin	25	14	3	25
P14	52	Male	Garmin	8	7	2	12
P15	32	Female	Garmin	12	10	2	20
P16	32	Female	Garmin	2.5	2.5	3	35
P17	48	Male	Polar	7	7	4	55
P18	20	Female	Apple	5	4	5	60
P19	21	Female	Garmin	15	10	2	30
P20	41	Male	Garmin	13	6	3	40
P21	37	Male	Garmin	4	4	6	120
P22	48	Female	Garmin	12	4	5	45
P23	35	Male	Polar	28	20	6	100
P24	22	Female	Garmin	6	3.5	2	10
AVR	42.15			12.17	7.27	3.67	46.25

years, Max=45 years), tracking their runs for 7.27 years (*SD*=4.11, Min=2.5 years, Max=20 years) and running 3.67 days a week recently (*SD*=1.58, Min =1 day, Max=every day). They had been running 47.83 kilometres weekly (*SD*=33.65 km, Min =10, Max=140). Most owned a Garmin watch (*N*=19), while the rest had either a Polar (*N*=3), Coros (*N*=1) or an Apple Watch (*N*=1). None of the respondents were injured by the time of the interviews, while 16 (67%) indicated they had running-related injuries in the past.

3.4 Data Analysis

We employed different analysis methods as the study yielded quantitative and qualitative data. Using IBM SPSS Statistics 26, we calculated means (*M*) and standard deviations (*SD*) for the survey questions with rating and Likert scales and frequencies for the open-ended questions like age, weekly running volume, and type of tracker used. We use these results to report the demographics of the participants.

For the interviews, we conducted qualitative data analysis [2, 62], following the same Reflexive Thematic Analysis (RTA) procedure proposed by [12, 15]. RTA was conducted by the first author, who conducted most of the interviews, and the second author, who was not involved in the interviews. These authors first downloaded the recordings and auto transcriptions provided by Microsoft Teams and reviewed the transcriptions with the recordings to ensure that the interviews were transcribed verbatim. Then, they read the transcripts while taking notes to familiarise themselves with the data (*Familiarization*). Next, the second author labelled the data from the first 12 interviews using inductive and deductive coding approaches (*Coding*). During deductive coding, survey questions were utilised as a reference (e.g., metrics to track while running or motivation to track running-related metrics). The data also induced codes, such as ‘relying on feelings to judge training load”. This coding was followed by reorganising the codes around initial themes, which were later synthesised into a preliminary codebook (*Generating initial themes*). Some example themes from this stage were “compliance with the training related suggestions”, “perceived usefulness of data”, and “feeling-driven training”. The first and second authors come to a common understanding of the themes through a series of discussion sessions (*Developing and reviewing themes*). The first author coded the rest of the transcripts by using the revised codebook and initial themes as references and checked them against the data. All authors communicated frequently during this process to refine and finalise the themes (*Refining, defining and naming themes*). Once we had the final set of themes, we integrated them into a narrative coherent with our research questions (*Writing up*).

Accordingly, we present the results under four headings: motivation to track running, assessing training load, and sources of determining and changing training load to answer the first research question (*How do runners use their sports data for TLM*?), and trust in trackers in TLM to answer the second research question (*How do they perceive and use trackers' TLM suggestions*?).

3.5 Limitations

Our interview sample was predominantly male and Western, with 15 of 24 male participants. A similar gender distribution was observed in survey results (*N*=85 female and *N*=164 male runners). Besides, while we did not collect the country of residence from survey participants, we expect a bias given that the recruitment primarily focused on a local marathon event within Europe, local running clubs and personal contacts from the authors. Future research could explore how TLM practices vary in other cultures and a more equally distributed sample in terms of sex. For example, a study by Niess et al. [66] comparing fitness tracking practices across demographics suggests that Arab users prioritise physiological measurements over goals. This preference may impact the value of metrics and TLM suggestions provided by trackers for this group.

Further, our sample reported that their average experience in running ranges from 6 months to 50 years. The runners' experience levels might influence how they approach TLM and sports data. Literature suggests that more experienced runners are likely to underestimate their skills (and the opposite for less experienced runners: they tend to overestimate their skills, see [51, 52]). However, within the scope of this paper, we used this information for description purposes rather than comparing how runners with different “perceived running experiences” manage their training load. Future research could examine whether novice and experienced runners manage their training load with sports trackers differently.

4 USE OF TRACKERS IN TRAINING LOAD PRACTICES

In this section, we discuss the insights our studies yielded into runners' TLM knowledge and practices, their use of trackers for monitoring and managing their training load (RQ1) and their perception of trackers' TLM-related suggestions (RQ2). In general, we did not observe any instances where interview participants contradicted or disagreed with what was identified in the survey study. Yet, our interview results provided a more nuanced perspective on integrating trackers into runners' TLM practices. We will report the findings from the survey as well as the interviews based on the order of questions provided in Table 1.

4.1 Motivations for Running Tracking

Survey results showed that runners are interested in “*learning how they can improve themselves*” (*M*=4.22, *SD*=0.81). Tracking their runs “*gives pleasure to learn about themselves*” (*M*=4.16, *SD*=0.85) and is “*a good way to improve performance*” (*M*=4.10, *SD*=0.86). They “*feel better when they track their runs*” (*M*=3.96, *SD*=0.93) and find it a good way to “*develop themselves*” (*M*=3.96, *SD*=0.96). These results demonstrated that participants in our survey are intrinsically motivated to track their runs (see Table 4 for all the survey items). Tracking runs for “*preventing running related injuries*” was rarely the focus for tracking (*M*=2.32, *SD*=1.23).

The interview results support these survey findings. A large majority of the interviewees (*N*=19) indicated using trackers to see their progress towards a target running race (e.g., running a half marathon), and assess performance improvements through self-comparisons (e.g., comparing one's performance with six months and one month before the race). The most salient perceived benefit of tracking runs is that various running-related metrics allow runners to assess their performance (*N*=13). For example, P21 illustrated this as:

> *“I think it is good to have an overview. And in running, it quickly became a means to plan. If you plan stuff, you also want to record what you do and see if your realisation is the same as planning. It is* ***very valuable to look back and see patterns, whether they lead to an injury or your performance progresses****. Then, you try to relate what you are doing to whether you progress or not and whether you get injured or not. So, I think those are the most important reasons. Getting a watch that* ***records everything was a thing of convenience****. I wanted to* ***have instant data during the run.*** *A watch was the obvious way to get started.”* (P21, M).

The second prominent motivation for tracking runs is to learn about running behaviour (*N*=14). On the one hand, runners are curious about their performance metrics and what the technology can bring them (*N*=5). On the other hand, they believe that trackers record their running more accurately than their memory, thus supporting better learning (*N*=11). Furthermore, according to our participants, using trackers is not only about assessing performance based on metrics but also about learning how to be a “better runner”. For instance, almost half of the participants mentioned that data provided by the trackers helped them better understand the

Table 4: Motivations for Running Tracking (*N*=249)*

	Motivations for Tracking	M	SD
IM	It is very interesting to learn how I can improve myself.	4.22	0.81
IM	Tracking gives me pleasure to learn more about myself.	4.16	0.85
IM	I find tracking is a good way to improve my performance.	4.10	0.86
IM	I feel better about myself when I track my runs.	3.96	0.93
IM	I have chosen to track my runs as a way to develop myself.	3.96	0.96
IM	Tracking my runs is an integral part of my life.	3.09	1.29
EM	I would feel bad about myself if I did not track my runs.	2.37	1.20
–	I track my runs to prevent running related injuries.	2.32	1.23
EM	Tracking my runs reflects the essence of who I am.	2.27	1.19
EM	Others would disapprove of me if I did not track my runs.	1.33	0.77
EM	The people I care about would be upset if I did not track my runs.	1.18	0.57

*IM refers to the items that measure intrinsic motivation, EM to extrinsic motivation. The item with – was not listed in the original scale.

rationale behind training suggestions given by their coaches or trackers (*N*=10).

Thirdly, runners would use their running data to compare it with their perceived running performance (*N*=13). Participants pointed out that the data's impact on motivation varies based on how well the data aligns with their perception of their running experience. Data occasionally leads to demotivation when there is a mismatch between tracked data and what the runner feels about their performance. Runners continue to track their runs despite this mismatch, as they rely more on their feelings than on data during data-feeling misalignments (see Section 4.4 for more details about this mismatch).

Another motivation for runners to track their runs is having a positive attitude, inherent affection for data and seeing running behaviours in numbers (*N*=9). In contrast with this affection, several participants noted that they typically do not look at their running data while running; nevertheless, they keep tracking their runs for future performance reflections (*N*=9). In simpler words, they are motivated to use trackers for "prospective reflections". Nonetheless, many runners track their runs for future reference. This practice resonates with documenting tracking, as outlined by [79], but differs from it: runners in our study do not only document their runs but also use them for future reflections, which can inform adjustments to their running behaviour. Especially for P20, tracking is to avoid missing values in their dataset: "*The problem is like Duolingo issue. . .* ***You do not want gaps in your data even if you do not use it****. That is, you know, data scientists' worst nightmare: the missing values. So,* ***to avoid those, I just use the watch all the time.***"

Other motivations for tracking runs include the sense of achievement and success runners feel after looking at their data, particularly when they see an improvement in their performance (*N*=7), adapting training load to prevent injuries (*N*=5) and recording running experience holistically by including factors beyond running metrics such as trail choices and favourite running spots (*N*=4). In sum, looking at the survey and the interview results together, it becomes apparent that runners have a desire to regulate their running behaviour to be better in what they do and utilise trackers to objectively capture running data that can be turned into actionable insights once it is combined with their experiences and feelings.

Table 5: The Perceived Importance of Metrics for Training Load Management

Metrics	M	SD
Average heart rate	3.90	1.07
Distance	3.83	0.95
Duration	3.76	1.05
Running Frequency	3.64	1.07
Instant pace	3.60	1.05
Average pace	3.60	1.05
Perceived intensity	3.55	1.07
Total weekly distance	3.52	1.06
Heart rate variability	3.36	1.19

4.2 Assessing Training Load with Tracker Metrics

In the survey, we asked participants about the importance of nine running-related metrics in assessing their training load. Among these metrics (Table 5), average heart rate was reported as the most important metric (*M*=3.90, *SD*=1.07), followed by distance (*M*=3.83, *SD*=0.95), and duration of runs (*M*=3.79, *SD*=1.05). HR variability was ranked as the least important metric for TLM, with a mean rating of 3.36 (*SD*= 1.19), yet above the middle score (i.e., 3.00) of the rating scale.

The interviews helped clarify why certain metrics were more prominent than others and provided a significant distinction between measured and derived metrics. When asked about the metrics checked before, during and after running, all participants reported checking at least one of the following: HR, distance, duration, and pace. In response to the same question, half of the participants mentioned looking at metrics such as VO_2Max (*N*=19) and power (*N*=10) data after running. Furthermore, we found that runners differentiate measured metrics, measured directly by the tracking device (e.g., HR), from derived metrics, which are calculated based on multiple measured metrics (e.g. recovery time). Interestingly, the number of participants who indicated a deliberate use of derived metrics when revising their training program (*N*=9) surpassed those who used measured metrics (*N*=4). We discovered that this tendency is associated with runners' awareness of their performance and how training influences it. The following comment from P4 illustrates how a measured metric (e.g., pace) can become less important in time:

> *"When I started running, **I made all the usual mistakes that one can make, one of which is going too hard and too often**. So, the heart rate data was a bit high. At some point, I started to care about my heart rate. And I wanted to keep it below a certain number. That can be really frustrating... **But now I know what I can do, and on those days that I run without looking at the watch, I can predict within a few seconds what my actual pace is**. So, I do not really need to watch to know what pace I am running."* (P4, M).

With the advancement of technology, the derived metrics used in training load also become more important. For example, for P7, who was tracking his runs for four years, power data became more important than HR data about a year ago:

> *"Also, in the beginning, I used to check only my pace, for long runs and my heart rate for the interval training. And after, I believe, a year ago, I changed to another device and connected my Garmin watch to the Stryd foot pod. And Stryd foot pod calculates your power. I have to say I really love it. **I love it even more than my heart rate values**."* (P7, M).

Furthermore, we found that the runners' emotional and mental state during a specific workout (i.e., whether they are in the mood to run intensively) hold equal importance to both measured and derived metrics in assessing training load (*N*=7). For example, P2 explains how he prioritises bodily sensations over a derived metric:

> *"Because sometimes I run the day after training, **the watch still says "you need to recover"**. But my legs do not feel like they need recovery. Of course, after a very long distance (run), I give two days break, but I also know that some people run a marathon each day, but this is not for."* (P2, M).

Interestingly, a subset of runners incorporates biomechanical metrics into understanding the training load of their running workouts. These metrics encompass vertical ratio, which serves as an indicator of a runner's balance (e.g., P10); training stress score, which quantifies the stress imposed by each workout (e.g., P21); and training readiness, which signals a runner's physical preparedness to another running workout (e.g., P23). While these metrics do not serve as the primary sources for training load assessment for many runners, they provide additional TLM-related insights. For example, P10 emphasised the significance of the *average ground contact time balance*, portraying how the nuances of foot placement and road conditions impact her running balance. Her watch gives her feedback when there is too much imbalance, resulting in changing the side of the road she runs to address any imbalance issues. To sum up, these findings signal the intricate web of factors that runners consider when assessing training load, underlining the blend of objective data, subjective sensations, and evolving personal insights that shape their training journeys.

4.3 Sources for Determining Training Routines and Changes

Survey participants' responses regarding their training program preferences and sources (Table 6) show the highest mean value for the statement "*I schedue my own weekly training program myself"* (*M*=3.43, *SD*=1.40), suggesting a tendency for self-directed training planning. Conversely, the statement "*I follow a training program that my sports watch provides me*" (*M*=1.46, *SD*=0.98) and "*I follow a training program that a running app provides"* received the lowest mean score, indicating a reluctancy to adhere to training suggestions given by technology. The mean value for receiving "*training programs from a running coach"* (*M*=2.41, *SD*=1.57) was moderate (Table 6).

Survey results also showed that runners rely more on their intuition than the trackers suggest in changing their training plans (Table 7). They tend to "*listen to their body before going for a run*" (*M*=3.76, *SD*=0.95), "*decide the pace of their run based on how they feel during running*" (*M*=3.72, *SD*=1.04), "*trust their body signals while planning their running schedule*" (*M*=3.68, *SD*=0.95). Furthermore, runners are less inclined to "*decide the pace of their run based on the data the sports watch provides them during running*" (*M*=2.86, *SD*=1.34). Finally, they are less inclined to "*decide training schedule based on tracker's training suggestions*" (*M*=1.84, *SD*=1.06) and "*based on tracker's recovery hours suggestion*" (*M*=1.84, *SD*=1.11).

Through the interviews, we learned more about why runners tend to schedule their own training program and are not inclined to follow a training program provided by an app or a sports watch. We found that runners do not implement the training suggestions given by the tracker when (1) there is incompatibility between what the runner feels and what data tells; (2) they have a preference for feeling-driven training (*N*=15); (3) they do not want to feel an obligation towards complying with a "machine" (e.g., P5) (*N*=4); the watches or apps do not provide tailored training suggestions (*N*=11) nor realistic and believable predictions regarding performance and recovery time (*N*=7), and actionable training program suggestions (*N*=3). For example, P15 expresses her awareness of physical limits, especially when she feels pain in their knees. Then her judgment becomes "*just about the feeling and not about what I see in the app or something"*.

Through the interviews, we discovered that determining and adapting a training program is a dynamic process for runners, which involves utilising different sources in different situations, feelings, data provided by trackers, coaches' suggestions, and normative training plans. We identified a tendency to prefer certain sources over others (e.g., relying more on one's feelings than tracker data when determining a training plan). The following sections present these sources from the most to the least preferred. However, this does mean that runners always stick with a single source when determining or adapting a training program. For instance, several participants mentioned that (e.g., P1, P5, P7). At the same time, they generally comply with training program suggestions made by their coaches. They also use their feelings to gauge the appropriateness of these suggestions.

Table 6: Source Use for Planning Running Workouts

Question	M	SD
I schedule my own weekly training program myself.	3.43	1.40
I receive a training program from a running coach.	2.41	1.57
I do not follow a structured training program.	2.30	1.51
I follow the training suggestions a runner friend provides me.	1.78	1.19
I follow a training program that a running app provides.	1.75	1.28
I follow a training program that my sports watch provides me.	1.46	0.98

Table 7: Sources for Change in Running Training

Question	M	SD
I listen to my body before going for a run.	3.76	0.95
I decide the pace of my run based on how I feel during running.	3.72	1.04
I trust my body's signals while planning my running schedule.	3.68	0.95
I decide the pace of my run based on the data my sports watch provides me during running.	2.86	1.34
I decide my training schedule based on my tracker's training suggestions.	1.84	1.06
I decide my training schedule based on my tracker's recovery hours suggestions.	1.84	1.11

4.3.1 Learning with Data, Knowing by Feel. We found that the most prevalent approach for runners in determining or adapting their training program was relying on their bodily sensations at a given time, such as being tired, fatigued or stimulated (*N*=17). Interviews demonstrated that most runners have a profound understanding of how their body responds to a running workout and that such self-awareness makes them confident in determining or adapting their training program themselves. The comment below showcases the shared understanding among runners that their body signals are more reliable indicators than what a tracker tells.

> *"I just know what the response is from my body if I run too fast. I listen to my body, and I think I understand my body's reactions quite well now. So, if I feel I am getting tired, I slow down a little bit. I do a little more if I have some headroom and can do a little more.* ***My body tells me much better, in my opinion, than the watch****."* (P12, M).

Runners also acknowledge that their perceived competence in determining training load by feel accumulates in time with the help of trackers. In his later comments, P12 pointed out that "*the technology can help get a relation between* ***how it actually goes and how you feel***", and that the tracker assists him in learning to listen to his body and react to it. In a way, the trackers' primary function becomes providing a ground truth to runners, helping them see their performance more objectively. For example, P17 stated the following:

> *"I did not use a training plan supported by the watch. But* ***I learned to read the information*** *the watch provides about heart rate, distance, speed, and cadence, and I translated them into my (marathon) training plan. So, after the marathon,* ***I look back on the information the watch provides and my feel and (reflect on) what is good and what went well****. Um, do I need to try more? Do I need to train less? Do I need to train in different sessions?"* (P17, M).

This approach of learning through data proved to be instrumental for injury prevention, as described by P21:

> *"When I see (injury is) in onset, I try to be very proactive there. I think I got better and not getting that (injury) at all, and I think the Training Peaks and* ***all the metrics helped a lot in that****. And every once in a while, you feel the niggle, something starts showing up, and if that's something that led to an injury earlier, you know what it is."* (P21, M).

Yet, as P13 highlighted, suggestions provided by trackers might contradict their perceived efforts when a tracker indicates an "acceptable level of training load", but the runner feels that the workout was too difficult to be at an acceptable level. In such situations, runners often prioritise suggestions leaning towards recovery. Such judgements lead to adopting the mantra "*no training is also training*" (P13), recognising the value of recovery in injury prevention.

4.3.2 Incorporating Trackers' Suggestions in TLM Decision-Making. The preference for feeling-driven training does not mean that runners completely ignore their trackers when determining or adapting their training program. For instance, some runners mentioned that they regard tracker suggestions as advice, meaning they incorporate them into their training plans while ultimately relying on their feelings (*N*=14). Furthermore, some tend to follow the tracker recommendations during their initial tracking phases and, in time, become more knowledgeable about the training load and their own bodies. During this process, they consider these suggestions mere advice rather than strict mandates. The comment of P7 encapsulates this transition:

> *"I followed my weight and my morning heart rate just to see if I was overtraining; you might lose weight, and your heart rate in the morning might be too low or too*

> *high compared to the average. That happened a couple of years ago when I was training for my first marathon. Still, I am continuously monitoring my weight and morning heart rate, but* ***even if I see some stability in data, my decision is more based on my feel than my watch****."* (P7, M).

Runners' compliance with the tracker's training load suggestions is related to the perceived usefulness of these suggestions, which is increased when they are presented in a comprehensible manner (*N*=9), tailored to runners' experience and knowledge in training load (*N*=6), and represent different time intervals (*N*=2). Such tailored tracker suggestions, however, cannot reflect outside factors that affect the quality of a running workout. Therefore, runners' decisions during and within the activity are sometimes less dependent on their in-act data. P14 exemplifies how weather conditions can affect the effort they put into a workout:

> *"(My decision to change the workout) mostly depends on the weather. When you go to the beach, it is often very windy, and your speed depends on the wind direction. I start with running in my face and end with the wind in my back. Because you have to run back, and if you are tired and still have the wind in your face, (running) turns very hard sometimes."* (P14, M).

4.3.3 Customising Normative Training Plans. We found that some runners (*N*=10) would customise training plans received from apps, online resources, and athletic clubs. This way of customising normative training plans allows them to change the workout days or the intensities based on their daily routines. Trackers play a role in this adoption when runners use the "recommended workout" function of the trackers. Such functions enable planning their own interval training workout. Some runners perceive these functions as "handy", while others might consider that their trackers do not support customisation that well. Some runners also register for a target race (e.g., a road half marathon or a hilly trail race) (*N*=10) and use this target race as the end of their training planning and adjustments. In making these adjustments, the data can play a role in making the training program better, as described by P23, whose aim was to be in the top 10 of a very competitive race:

> *"I mostly make training programmes for a target race. First, I define a target. Then I make the programme mostly for a year or a season. That programme is only the base. Each month and each week, I look into my data and adjust the weekly training plan. After all,* ***I make the base programme for each target race much better while training.****"* (P23, M).

The interviews revealed that runners considered various factors when customising their training plans, which include the runner's preferred style of training (e.g., how much they want to push themselves) (*N*=4) and the level of exertion needed (e.g., being in an aerobic and anaerobic zone) (*N*=4). For example, P13, who was very much into running long-distance races, stated that after several years of working with a coach, she concluded that slow runs were unsuitable. Therefore, she replaced workouts involving intervals, longer runs with speed variations, or changes in pace because she believed these methods led to better results in her running.

4.3.4 Planning Training Load with a Coach. Sometimes, runners work with a human coach who determines their training program and adjusts the planning together with the runner (*N*=10). In such cases, the tracker data becomes a communication medium between the coaches and the runners. The runners are engaged in discussions with the coach and negotiate with them to decide what is best for their health and goals. In most cases, they use a combination of tracker and input from the coach to determine the right training. The comment from P1 illustrates this:

> *"Well, I do have a running coach. That helps me plan my running. But then again, that's the planning, and I always try, and say: Okay, does this planning feel right?* ***Does it feel right based on how I feel, but also does it feel right based on the metrics that I see on my watch****,* ***on my Strava, on Training Peaks****?"* (P1, M).

Even though the survey results and interviews highlighted that runners are reluctant to follow a training plan from a virtual coach in managing their training load, the openness to such kinds of sources for TLM might depend on the runners' preferences. For example, P16, working with a human coach at the time of the interviews, indicated that she completed a training program provided by Garmin AI coach twice. However, her reason for transitioning from AI coach to human coach was to improve her performance with someone who knew better than her. In her own words, working with a human coach was "*accepting the convenience of having someone who could know better*", implying that a virtual coach cannot know better than a human coach.

4.4 Trust in Trackers in Training Load Management

In the survey, we inquired into participants' trust in the data and suggestions displayed by trackers. We found that runners mostly trust technology in accurately measuring their training metrics (M=3.92, SD=0.95), with their trust level decreasing in training load (M=3.01, SD=1.13) and recovery time calculations (M=2.51, SD=1.19). This finding was also supported by the interview results, which indicated that whether a metric is measured directly by the tracker (HR) or derived from multiple metrics (e.g., energy level) influences runners' trust in this data. Accordingly, rather than its precision, runners like P11 trust "the trend" in the derived metric (e.g., VO_2Max). The main reasons for less trust in such derived metrics are, on the one hand, the lack of transparency on how a derived metric is calculated (*N*=8) and whether this calculation is based on a rigorous method (*N*=9), and on the other hand, the belief that inaccuracy of a measured metric will be multiplied when it is used to calculate a derived metric (*N*=5). P10 illustrates this situation with the following comment:

> *"When Garmin became more capable of measuring and* ***showing performance metrics****, I also became interested in them. I am still a fan of them, but my* ***heart rate data is not always accurate*** *because it is irregular, even during rest. Yet Garmin says (after running), "You ran too hard or too far, too fast". And "take two or three days of rest". I cannot use this suggestion (recovery time) because they look at my heart rate data, which is inaccurate."* (P10, F).

The most prominent factor influencing trust in data is the accuracy in measuring a metric (*N*=20), along with the alignment between multiple measures of the same metric across time (*N*=6) and alignment between metrics measured by different devices (e.g., measuring one's heart rate through a chest sensor and smartwatch sensor) (*N*=4). The interview results indicate that accuracy is not only about measuring the metrics. Some participants mentioned that the validity of tracker predictions (e.g., race predictions) and suggestions (e.g., recovery time) increases their trust in these devices (*N*=15). There were also some instances where runners lost confidence in the accuracy of the trackers. The comment of P16 illustrates this:

> *"In some watches, there's a lot of GPS jumping. For instance,* ***it might show that I ran a kilometre in 2 minutes****. When I see this happening in 2-3 workouts in a row,* ***I lose confidence in the watch's accuracy*** *after a while."* (P16, F).

We also found that many runners use additional sensors to manage their training load precisely. For example, five runners explicitly indicated using a heart rate chest band to measure their training HR accurately. Still, others consider a few beats per minute imprecision in their HR data as not vital, as was stated by P14 that "5% more accuracy" is not very important.

The second most prominent factor influencing trust was the compatibility between what the runner feels and what the data tells. Compatibility then depends on the runner's agreement level with the data (*N*=15), as also elaborated on previously. We discovered that runners tend to compare the tracker data with their in-act experiences (e.g., comparing pace data provided by the tracker with perceived pace). They build trust in time if the gap between what is measured and felt is small (*N*=2). Furthermore, we found that runners do not merely subordinate themselves to the objective data that trackers provide them; instead, they consider the effort they put into diverse running workouts (i.e., easy, long or interval training). They attend to the post-workout somatic signals their body gives them. At times, they leverage the measured (e.g., heart rate) or derived (e.g., recovery time) metrics as informative cues in tailoring and fine-tuning their next workout. In other words, regarding trust in data, they try to see the bigger picture by comparing metrics and data sources. The case of P21 illustrates this nicely:

> *"Let's put it like this:* ***It is never just data I look at.*** *It is always a* ***combination of several metrics and my feelings****. If one is off, I look at the other ones,* ***and there's always a total picture****. And if something is not good during training, it is not just the watch telling me; I will also feel it. And that's fine because it happens every once in a while. If it happens a few trainings in a row, then something's wrong, and you should do something about it. But* ***it's not the watch that throws me out of balance****."* (P21, M).

Finally, survey results show that runners' trust in the accuracy of their data is essential for better TLM. We did a correlation analysis to understand how runners change their running routines based on training load-related suggestions from trackers. The results showed that runners' overall trust in trackers positively affects how they use them in planning and making decisions on their tracking. They tend to follow trackers' recovery time suggestions when they trust in trackers to accurately calculate their training load ($r(249)=0.34$ $p<0.05$, two-tailed) and recovery time ($r(249)=0.49$, $p<0.05$, two-tailed). Correlation analysis also showed that runners who trust their trackers to calculate their training load correctly are more likely to follow suggestions their tracker provides ($r(249)= 0.23$, $p<0.05$, two-tailed).

5 DISCUSSION

Through the deployment of a survey and interviews, we aimed to understand runners' use of sports trackers in their TLM practices. Overall, we discovered that these practices are dynamic, representing diverse relations between runners and their trackers in managing training load. During initial experiences in running, runners rely more on trackers and use them to gain insights into their performance. In time, they become more attuned to the signals of their body and learn, partially through the data, to better interpret these signals. Thus, their reliance on trackers reduces, and they start compensating for potential data flaws and inaccuracies in the data with their bodily sensations. Furthermore, our research, especially in the case of TLM, has revealed that runners engage in two types of TLM: Guided and Self-Directed TLM, differing in terms of the purpose of sports tracker use, sources for determining and adapting training routines and the role of technology in TLM.

In **Guided TLM** (Table 8), a runner aims to develop competency in TLM by learning from various sources. These include (1) running related data provided by trackers in the form of measured metrics (i.e., metrics based on the measurement of a single metric such as HR) and derived metrics (i.e., metrics calculated based on multiple measured metrics., such as VO_2Max); (2) body signals such as feeling exhausted; (3) training-related suggestions of virtual or human coaches, and (4) TLM-related normative information found in online sources or apps. In this type of TLM, runners elaborate on their running performance by comparing their somatic signals with running related data, which in turn helps them understand what these metrics mean for their performance. In some cases, this process of learning about the self is supported by a human or virtual coach, who makes training suggestions based on runners' performance measures, subjective experiences (e.g., perceived effort and fatigue) and training preferences. Thus, in Guided TLM, sports trackers work as data provider or workout advisor.

In **Self-directed TLM**, runners aim to decide the best training plan for them and be autonomous in TLM. In this type of TLM, they rely on their bodily signals to determine or adapt their training plans. In a way, data becomes less critical for TLM as the runner transitions from Guided-TLM to Self-directed TLM. Technology's role as the data provider remains, but runners start to perceive the role of technology more like a supporter, which confirms their decisions and bodily signals rather than an advisor suggesting t what to do.

In their research, Rapp and Tirabeni [72] identified a trend among amateur athletes towards self-coaching. Our findings extend these insights, showing that amateur runners also work with trainers [72, 73]. While there are parallels between our results and the prior work, notable differences emerge in data interpretation, sensory experiences, and coaching methodologies. We propose that TLM

Table 8: Runners' TLM practices and technology's role in these practices

	Type of TLM	
	Guided TLM	Self-directed TLM
Purpose for sports informatics use:	Building competence in TLM Learning from various sources	Being autonomous in TLM
Sources for determining and adapting training routines:	Running related data Somatic signals Coaches Normative training plans	Somatic signals
The role of technology in TLM:	Data provider Advisor	Data provider Supporter

should be viewed as a dynamic, evolving process and found that runners develop their TLM competencies over time, gradually becoming more autonomous. This progression is not strictly tied to runners' status as amateur or elite athletes; instead, it hinges on effectively utilising data. This involves ongoing reflection on personal experiences to enhance TLM proficiency. Hence, our findings indicate that self-coaching, or as we term it, "self-directed TLM", is not exclusively the domain of either amateur or elite athletes. Instead, it emerges as a function of growing competence in TLM. This competence is cultivated over time through continuous interaction with personal data, requiring individuals to interpret it in the context of their physical sensations.

Our work also revealed friction between runners' knowledge about training load, their trust in their trackers' and their compliance with trackers' TLM recommendations (i.e., the importance given to data for managing training load or runners' tendency to lose trust in data when their experience is not aligned with it, as discussed in Section 4.4). Furthermore, although current sports trackers provide essential TLM-related data, runners do not always rely on those (e.g., the case of feeling-driven training, as discussed in Section 4.3.1). Considering these findings, we see several challenges and opportunities for SportsHCI research in improving sports tracking in TLM practices. We discuss these by referring to the relationship between performance metrics and subjective experiences, training load beyond performance data, and runners' trust in trackers in managing their training. Finally, we discuss the potential benefits of an ongoing dialogue between sports science and SportsHCI research.

5.1 Training Load Awareness Beyond Performance Data

Finding the "sweet spot" in training load [42] is highly personal and depends on how the athlete's body reacts to the training load. Therefore, sports science attributes great importance to TLM for improving athletes' performance to keep them physically and mentally healthy (as explained in Section 2.1). In parallel with this need, current sports trackers can measure various TLM-related data (as discussed in Section 2.2).

No doubt that tracking running related data provides runners with insights into their external (e.g., running distance and duration) and internal (e.g., heart rate) training load measures [31, 48, 50]. However, our study revealed that runners do not always show great interest in using this data in determining and adapting training load (Section 4.3). Still, although some runners are interested in using sports trackers' data, they found data-driven TLM limited, as it only allows using the data the technology can track. This is problematic, as TLM requires data beyond those performance metrics, such as runners' physiological (e.g., muscle and tissue damage [69]) and psychological adaptations to training load (e.g., disappointment about performance [20]).

Therefore, our results and the evidence from sports science research signal a discrepancy between what technology affords, what the runners want, and how training should be managed. We do not consider this a limitation but rather an opportunity for SportsHCI. It would be possible to provide a more holistic and personal TLM experience if sports trackers can acknowledge missing data and offer complementary advice to get a fuller picture of TLM. For instance, asking runners their Rating of Perceived Exertion (RPE) after a run and how "strong" they felt is a way to achieve that. This rating should also be combined with measured metrics and presented in a comprehensive manner. As much as identifying missing data, trackers could adapt to the runners' training load management needs, finetuning training load suggestions but not overwhelming them with irrelevant data since they give varying levels of importance to running-related metrics (as explained in Section 4.2.) In the next section, we reflect on how trackers can complement runners' subjective experiences with objective data and help them receive tailored TLM support.

5.2 Complementing Subjective Experience with Performance Metrics

As stated in Section 2, subjective rating of perceived effort (external load) [5] is a reliable and common way to quantify and assess internal load [17, 40]. According to our survey results, the runners did not deem this rating important in determining their training load. Conversely, interviews showed that runners use subjective assessments when managing their training load (e.g., P12, P15). These findings imply that runners tend to see quantified performance values as metrics for assessing training load (e.g., as heart rate), while they do not treat bodily signals as data sources. Yet, they use these signals to determine their exercise load (as in the case of self-directed training). Even though current sports trackers facilitate documenting ratings of perceived effort after running workouts, none of our participants explicitly mentioned using these ratings effectively as part of TLM. Therefore, we think that making the perceived effort easier to be rated at the end of the workouts and eliminating the potential burden on the runners would help them integrate more subjective ratings into their TLM-related assessments.

A recurring theme in the findings was the distinction between measured and derived performance metrics. Runners trust trackers for measured metrics like distance and time, while there is scepticism about the technology's ability to perform the necessary integration in calculating derived metrics and training load. This scepticism could be related to runners' perceptions of trackers' limitations in accurately capturing the subjective aspects of their performance. Therefore, runners often prioritise self-development, self-awareness, and somatic cues in their training routines when they consider metrics derived from the trackers. In doing so, they tend to appreciate performance feedback in relative terms, considering their performance within their circumstances. For instance, runners do not acknowledge a decline in VO_2Max estimates associated with ageing but still perceive their performance as good. This preference for relative measures ties into runners' desire for a personalised understanding of their performance rather than only relying on objective and absolute values.

Training load can be perceived as a score accumulated and quantified from the acute-chronic load ratio (as described in [42]). However, our research reveals a notable emphasis on subjective experiences over quantified metrics, connecting our results to the importance of athlete resilience [see 22, 38]. Additionally, our findings suggest a disparity between scientific recommendations for TLM and how they *are* implemented in runners' TLM practices. Previous HCI studies on self-tracking acknowledge the importance of subjective experiences and suggest complementing quantified data with users' sensations. For example, Rapp and Tirabeni's [72] findings illustrate that amateur athletes trust the objectivity of the monitored parameters, while elite athletes trust their sensations in sports tracking more. However, our findings show that although none of our participants identified themselves as elite athletes, they also relied on their subjective sensations in TLM. Therefore, our studies highlight the complementary role of objective data to subjective experiences, not the other way around, as subjective experiences might be the guiding force in TLM rather than a mere afterthought. This stark contrast was apparent in runners' desire to be autonomous in managing their training workload and relying more on their bodily awareness. Hence, as we observe in the transition from Guided TLM to Self-directed TLM (Table 8), we suggest that future research explore how tracking technology can support the acquisition of competence in TLM while finding a way to remain in the background once such competence is achieved [as in the concept of unremarkable computing 87, 97].

5.3 Building Informed Trust in Trackers

Our results revealed that although many runners trust their trackers in calculating their training load, not all are willing to follow trackers' TLM-related recommendations. We found that trust in TLM recommendations from human coaches or self is higher than in recommendations from trackers. We see several similarities between these findings and ongoing sports technology discussions. Sports science research shows that experienced runners are inclined to trust the recommendations of human coaches more than an information system, and they tend to be more engaged in training when their training plans are developed and remotely supervised by a human [9]. Despite the evidence in their training data indicating a (potential) overload, runners are not always willing to change their training plan [34]. Being overenthusiastic about running might cause a runner to mistrust and ignore the evidence that indicates training overload, and they maintain (or even increase) training frequency. While most runners want to keep a healthy lifestyle, not complying with the evidence of inadequate training loads can cause running-related injuries and quitting running [49].

Combining the results of our study and evidence from sports science, we propose that the TLM support for experienced runners should be designed more like negotiation conversations rather than training suggestions [78]. Such support should not be perceived as coming from an authority but more from a negotiator whose task is to balance runners' beliefs and what the data tells. Examining human-agent collaborations in HCI, Cila [14] suggests that creating a balance in negotiation requires a joint commitment of humans and technological agents to the negotiation activity. Aligned with that, we think the runners should also be aware of the goals of the TLM negotiations. For instance, this negotiation can be done through the sports tracker interfaces and the dashboards that illustrate the benefits of keeping the training load within the sweet spot [42]. At the same time, the sports trackers should not overrule the autonomy of the runners in decision-making. Still, they should support microplanning and negotiations as a way of flexibility [26] in setting running-related goals.

For designing trackers to support runners' TLM, designers should be careful about how the training load is communicated. Since TLM is a crucial practice for runners' health, tracking technology should go beyond tracking and focus on runners' learning [28]. Mere compliance to TLM-related recommendations provided by the trackers would not support runners' learning practices (i.e., learning about the relationship between one's capacity and performance), nor learning how the internal and external factors influence training load. Trackers should communicate how training metrics and data jointly affect training load to support runners' learning. These findings also pave the path for AI-supported TLM applications that support balanced and individualised training load management practices.

Extending Rapp and Tirabeni's [72, 73] findings, our research indicates that non-elite runners also place significant trust in their bodily sensations and subjective feelings. Our findings suggest that SportsHCI should go beyond focusing on quantitative data and incorporate qualitative aspects, such as an athlete's feedback and perceived exertion levels. By doing so, these tools can provide a more comprehensive view of TLM, combining measured metrics with personal, subjective insights. This holistic approach acknowledges the complexity of human performance and the multifaceted nature of training, enabling runners of all levels to make more informed decisions tailored to their unique needs and experiences. In essence, SportsHCI should facilitate a balanced integration of data-driven insights and personal intuition in TLM, ensuring that runners can optimise their training in a scientifically informed and personally resonant way.

Aligned with prior research findings [18], we believe that fostering sensemaking of running-related data would empower runners to manage their training load effectively, especially guide data handling (e.g., seeing the patterns in TLM calculations) and interpretation (e.g., making sense of the patterns in TLM calculations). For example, while describing why the training load should be decreased or increased, the tracker can facilitate hiding or curating the collected data so that the runner observes the impacts of the adjusted training load. This way, any measurement errors in data could be explained, and the runner could be made aware that the collected data is not the absolute truth.

5.4 Towards an Integrated Training Load Management in SportsHCI

Given the results of our study and the challenges and opportunities identified above, we see several distinct future directions in sports informatics and SportsHCI. Better sports informatics systems need to be developed by considering insights from sports science and runners' current TLM practices. These should better fit the runners' training needs by supporting them through data in a manner that they can trust. Still, current HCI practices often focus on precise measurement and data representation, enhanced with human-experience-related factors. We think the SportsHCI can leverage more from sports data for TLM by considering the following directions.

Learning: The challenge of helping runners manage their training load is not solely about measuring, calculating, or estimating data. Instead, SportsHCI should strive for a holistic approach, providing athlete-centred reports that guide data handling and interpretation and help decision-making around training load and learning about TLM [13]. This approach goes beyond employing trackers to learn to regulate athletes' body reactions. It recognizes the importance of integrating subjective sensations and personal experience into TLM practices, even for non-elite athletes [72].

Trust: We found that runners are interested in receiving advice on their TLM but currently distrust the advice of trackers. They do not trust the accuracy of several derived metrics current trackers provide, some of which are essential for TLM (e.g., HR variability). If technology fails to calculate the training load, we need solutions to communicate relevant information to users in a way they can trust. As part of this, TLM support could be designed more like negotiation conversations balancing data with experiences rather than as authoritative suggestions [as suggested in 78].

Recording experiences: Current trackers already allow for subjective self-reporting (e.g., notes) of the workouts. Yet, firstly, not all relevant factors of experience are equally included, and secondly, such functions are not used very much in an integrated manner with the data. Future systems should expand on functions to record experience-related factors, like "muscle soreness", "disappointment in performance", or "feelings like a relaxing run". Furthermore, such data should also be reported more holistically in integration with the numbers rather than keeping them as a separate, unrelated measure. These alternative ways of recording experiences and data can benefit from prior research, such as the use of wearable e-textile displays to support group running [58] or the use of drones in mediating running groups [4] and supporting the well-being of runners [3].

Data vs. experience: Finally, we suggest that future sports trackers could benefit from fundamentally turning around the basic underlying paradigm of data versus experience: Rather than focusing on data and extending it with subjective experiences, they should put the subjective experience first, as the runners themselves do in practice, and then support the runner's experience and perception with data, insights and suggestions from the tracker's quantitative measurements.

It is also essential to recognize the evolving nature of data literacy as a key component in athletic training and decision-making. Data literacy extends beyond mere comprehension of data; it encompasses the ability to critically evaluate and effectively utilise data in everyday practice [45], in college sports [16] and recreational running [67]. This aspect is particularly salient in our study on TLM, where we observe that as runners gain experience, they develop a nuanced form of data literacy. This progression is about accumulating knowledge and cultivating a critical perspective towards the data they encounter. Runners learn to question the relevance, accuracy, and applicability of TLM-related data, which is a crucial step in making informed decisions about their training plans.

TLM as a promising field of research for SportsHCI: We think our results can also be useful for other sports. Our research in running, a sport where athletes extensively self-track and manage training loads to enhance performance, has revealed insights with broad applications across various sports. This approach is particularly relevant for sports with high injury incidence, such as cycling, swimming, triathlon, and multisport disciplines, where athletes can benefit from using personal trackers to monitor and adjust their training loads. Even in team sports like soccer, emerging technologies like smart socks[8] and shin guards[9] are beginning to allow similar data-driven training optimisations. These advancements highlight the importance of incorporating athletes' feedback into technology design, ensuring that devices offer practical, user-friendly advice and interpretations. As more sports adopt these innovative tracking tools, the findings gathered from the running context can guide their development for more personalized and effective training methods across diverse sporting disciplines.

Cross-disciplinary collaboration between sports science, HCI and interaction design could help advance this nascent research area. One potential direction is conducting further integration of the perspective of these fields. For instance, our results showed the mismatches between runners' expectations from sports tracking, tracking capabilities of technology (e.g., tracking heart rate) and how technology utilises/communicates data (e.g., making training suggestions based on heart rate variability). Uncovering the reasons for these mismatches together with sports scientists and addressing them with better interaction design requires explicating the roles users, technology, and science could play in TLM.

6 CONCLUSIONS

In this paper, we showed the importance of training load management for the well-being of runners and illustrated how they

[8]https://www.danusports.com/ (Retrieved on 28 November 2023)
[9]https://humanox.com/en/hx50-shin-guard/ (Retrieved on 28 November 2023)

use sports trackers in TLM practices. Despite its significance in athletes' well-being, and although the current trackers can record much of the necessary data for TLM, HCI literature has largely overlooked TLM till now. Our findings address this gap by providing a novel lens on how the users of sports tracking technology should be supported to help them make decisions related to training load. Accordingly, by closely looking into runners' TLM practices, we unravelled the dynamic nature of TLM, encompassing runners' transition from Guided TLM to Self-Directed TLM by reflecting on their data and running experiences. We have uncovered varying levels of interest in quantifying and utilising running-related metrics for TLM and a desire to personalise training programs according to a multitude of factors beyond performance metrics. We have also highlighted a discrepancy between what sports science literature suggests as the best practice for TLM, what aspects of training load trackers can quantify and what aspects of training load runners are interested in tracking.

Overall, our findings signal a new line of sports trackers prioritising the subjective experience and bodily signals over data provided by the technology, where the latter is only used to complement the former. We conclude that developing such trackers would be only possible with an ongoing dialogue between SportsHCI and Sports Science domains, involving users as experts in their lived experiences. All in all, this paper is an attempt to initiate such a dialogue.

ACKNOWLEDGMENTS

We would like to thank Ir. Amalia Zafeiri for her help in the earlier versions of this research and the participants for their valuable time.

Grant Sponsor

This study was partically supported by Dutch Research Council (NWO) award number: NWO-406.XS.01.112.

REFERENCES

[1] Aicale, R., Tarantino, D. and Maffulli, N. 2018. Overuse injuries in sport: a comprehensive overview. *Journal of orthopaedic surgery and research*, 13, 1, 1-11. https://doi.org/10.1186/s13018-018-1017-5

[2] Auerbach, C. and Silverstein, L. B. 2003 *Qualitative data: An introduction to coding and analysis*. NYU press.

[3] Balasubramaniam, A., Reidsma, D., Obaid, M. and Heylen, D. Exploring Runners' Preferences of Drone Based Feedback to Support their Well-Being. In Proceedings of the Proceedings of the 2nd International Conference of the ACM Greek SIGCHI Chapter).https://doi.org/10.1145/3609987.3609997

[4] Baldursson, B., Björk, T., Johansson, L., Rickardsson, A., Widerstrand, E., Gamboa, M. and Obaid, M. DroRun: Drone Visual Interactions to Mediate a Running Group. In Proceedings of the Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction).https://doi.org/10.1145/3434074.3447148

[5] Barry, L., Lyons, M., McCreesh, K., Powell, C. and Comyns, T. 2021. The relationship between training load and pain, injury and illness in competitive swimming: A systematic review. *Physical Therapy in Sport*, 48, 154-168. https://doi.org/10.1016/j.ptsp.2021.01.002

[6] Bertelsen, M., Hulme, A., Petersen, J., Brund, R. K., Sørensen, H., Finch, C., Parner, E. T. and Nielsen, R. 2017. A framework for the etiology of running-related injuries. *Scandinavian journal of medicine & science in sports*, 27, 11, 1170-1180. https://doi.org/10.1111/sms.12883

[7] Blanch, P. and Gabbett, T. J. 2016. Has the athlete trained enough to return to play safely? The acute: chronic workload ratio permits clinicians to quantify a player's risk of subsequent injury. *British journal of sports medicine*, 50, 8, 471-475. http://dx.doi.org/10.1136/bjsports-2015-095445

[8] Bompa, T. and Buzzichelli, C. 2015 *Periodization Training for Sports, 3E*. Human kinetics.

[9] Boratto, L., Carta, S., Mulas, F. and Pilloni, P. 2017. An e-coaching ecosystem: design and effectiveness analysis of the engagement of remote coaching on athletes. *Personal and Ubiquitous Computing*, 21, 4, 689-704. https://doi.org/10.1007/s00779-017-1026-0

[10] Borresen, J. and Lambert, M. I. 2009. The quantification of training load, the training response and the effect on performance. *Sports medicine*, 39, 9, 779-795. https://doi.org/10.2165/11317780-000000000-00000

[11] Bourdon, P. C., Cardinale, M., Murray, A., Gastin, P., Kellmann, M., Varley, M. C., Gabbett, T. J., Coutts, A. J., Burgess, D. J. and Gregson, W. 2017. Monitoring athlete training loads: consensus statement. *International journal of sports physiology and performance*, 12, s2, S2-161-S162-170. https://doi.org/10.1123/IJSPP.2017-0208

[12] Braun, V. and Clarke, V. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology*, 18, 3, 328-352. https://doi.org/10.1080/14780887.2020.1769238

[13] Cardinale, M. and Varley, M. C. 2017. Wearable training-monitoring technology: Applications, challenges, and opportunities. *International Journal of Sports Physiology & Performance*, 12. http://dx.doi.org/10.1123/ijspp.2016-0423

[14] Cila, N. Designing Human-Agent Collaborations: Commitment, responsiveness, and support. In Proceedings of the CHI Conference on Human Factors in Computing Systems).https://doi.org/10.1145/3491102.3517500

[15] Clarke, V., Braun, V. and Hayfield, N. 2015. Thematic analysis.Qualitative psychology: A practical guide to research methods J. A. Smith, Sage Publications, Great Britain.

[16] Clegg, T., Greene, D. M., Beard, N. and Brunson, J. Data everyday: Data literacy practices in a Division I college sports context. In Proceedings of the Proceedings of the 2020 CHI conference on human factors in computing systems).https://doi.org/10.1145/3313831.3376153

[17] Constantiou, I., Mukkamala, A., Sjöklint, M. and Trier, M. 2022. Engaging with self-tracking applications: how do users respond to their performance data? *European Journal of Information Systems*, 1-21. https://doi.org/10.1080/0960085X.2022.2081096

[18] Coşkun, A. and Karahanoğlu, A. 2022. Data Sensemaking in Self-Tracking: Towards a New Generation of Self-Tracking Tools. *International Journal of Human–Computer Interaction*, 1-22. https://doi.org/10.1080/10447318.2022.2075637

[19] Coyne, J. O., Gregory Haff, G., Coutts, A. J., Newton, R. U. and Nimphius, S. 2018. The current state of subjective training load monitoring—a practical perspective and call to action. *Sports medicine-open*, 4, 1, 1-10. https://doi.org/10.1186/s40798-018-0172-x

[20] de Jonge, J., van Iperen, L., Gevers, J. and Vos, S. 2018. 'Take a Mental Break!'study: Role of mental aspects in running-related injuries using a randomised controlled trial. *BMJ Open Sport & Exercise Medicine*, 4, 1, e000427. http://dx.doi.org/10.1136/bmjsem-2018-000427

[21] Deelen, I., Janssen, M., Vos, S., Kamphuis, C. B. and Ettema, D. 2019. Attractive running environments for all? A cross-sectional study on physical environmental characteristics and runners' motives and attitudes, in relation to the experience of the running environment. *BMC public health*, 19, 1-15. https://doi.org/10.1186/s12889-019-6676-6

[22] Den Hartigh, R. J., Meerhoff, L. R. A., Van Yperen, N. W., Neumann, N. D., Brauers, J. J., Frencken, W. G., Emerencia, A., Hill, Y., Platvoet, S. and Atzmueller, M. 2022. Resilience in sports: a multidisciplinary, dynamic, and personalized perspective. *International Review of Sport and Exercise Psychology*, 1-23. https://doi.org/10.1080/1750984X.2022.2039749

[23] Dourish, P. and Gómez Cruz, E. 2018. Datafication and data fiction: Narrating data and narrating with data. *Big Data & Society*, 5, 2, 2053951718784083. https://doi.org/10.1177/2053951718784083

[24] Drew, M. K. and Finch, C. F. 2016. The relationship between training load and injury, illness and soreness: a systematic and literature review. *Sports medicine*, 46, 6, 861-883. https://doi.org/10.1007/s40279-015-0459-8

[25] Dunbar, C. C., Robertson, R. J., Baun, R., Blandin, M. F., Metz, K., Burdett, R. and Goss, F. L. 1992. The validity of regulating exercise intensity by ratings of perceived exertion. *Medicine & Science in Sports & Exercise*. https://doi.org/10.1249/00005768-199201000-00016

[26] Ekhtiar, T., Karahanoğlu, A., Gouveia, R. and Ludden, G. 2023. Goals for Goal Setting: A Scoping Review on Personal Informatics. In *Proceedings of the Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA). Association for Computing Machinery. https://doi.org/10.1145/3563657.3596087

[27] Epstein, D., Cordeiro, F., Bales, E., Fogarty, J. and Munson, S. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the Proceedings of the 2014 conference on Designing interactive systems* (Vancouver, BC, Canada). Association for Computing Machinery. https://doi.org/10.1145/2598510.2598558

[28] Epstein, D. A., Ping, A., Fogarty, J. and Munson, S. A. 2015. A lived informatics model of personal informatics. In *Proceedings of the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. https://doi.org/10.1145/2750858.2804250

[29] Eston, R. 2012. Use of ratings of perceived exertion in sports. *International journal of sports physiology and performance*, 7, 2, 175-182. https://doi.org/10.1123/ijspp.7.2.175

[30] Fan, C., Forlizzi, J. and Dey, A. K. A spark of activity: exploring informative art as visualization for physical activity. In Proceedings of the Proceedings of

the 2012 ACM Conference on Ubiquitous Computing).https://doi.org/10.1145/2370216.2370229

[31] Foster, C., Rodriguez-Marroyo, J. A. and De Koning, J. J. 2017. Monitoring training loads: the past, the present, and the future. *International journal of sports physiology and performance*, 12, s2, S2-2-S2-8. https://doi.org/10.1123/IJSPP.2016-0388

[32] Gabbett, T. J. 2016. The training—injury prevention paradox: should athletes be training smarter and harder? *British journal of sports medicine*, 50, 5, 273-280. http://dx.doi.org/10.1136/bjsports-2015-095788

[33] Gabbett, T. J., Hulin, B. T., Blanch, P. and Whiteley, R. 2016. High training workloads alone do not cause sports injuries: how you get there is the real issue, 50, 8, 444-445. http://dx.doi.org/10.1136/bjsports-2015-095567

[34] Gabbett, T. J., Kennelly, S., Sheehan, J., Hawkins, R., Milsom, J., King, E., Whiteley, R. and Ekstrand, J. 2016. If overuse injury is a 'training load error', should undertraining be viewed the same way?, 50, 17, 1017-1018. http://dx.doi.org/10.1136/bjsports-2016-096308

[35] Giblin, G., Tor, E. and Parrington, L. 2016. The impact of technology on elite sports performance. *Sensoria: A Journal of Mind, Brain & Culture*, 12, 2. https://doi.org/10.7790/sa.v12i2.436

[36] Halson, S. L. 2014. Monitoring training load to understand fatigue in athletes. *Sports medicine*, 44, 2, 139-147. https://doi.org/10.1007/s40279-014-0253-z

[37] Hancı, E., Lacroix, J., Ruijten, P. A., Haans, A. and IJsselsteijn, W. 2020. Measuring commitment to self-tracking: development of the C2ST scale. *Personal and Ubiquitous Computing*, 24, 735-746. https://doi.org/10.1007/s00779-020-01453-9

[38] Hill, Y., Den Hartigh, R. J., Meijer, R. R., De Jonge, P. and Van Yperen, N. W. 2018. Resilience in sports from a dynamical perspective. *Sport, Exercise, and Performance Psychology*, 7, 4, 333. https://doi.org/10.1037/spy0000118

[39] Hreljac, A. 2005. Etiology, prevention, and early intervention of overuse injuries in runners: a biomechanical perspective. *Physical Medicine and Rehabilitation Clinics*, 16, 3, 651-667. https://doi.org/10.1016/j.pmr.2005.02.002

[40] Hreljac, A., Marshall, R. N. and Hume, P. A. 2000. Evaluation of lower extremity overuse injury potential in runners. *Medicine and science in sports and exercise*, 32, 9, 1635-1641. https://doi.org/10.1097/00005768-200009000-00018

[41] Impellizzeri, F. M., Marcora, S. M. and Coutts, A. J. 2019. Internal and external training load: 15 years on. *International journal of sports physiology and performance*, 14, 2, 270-273. https://doi.org/10.1123/ijspp.2018-0935

[42] Impellizzeri, F. M., Tenan, M. S., Kempton, T., Novak, A. and Coutts, A. J. 2020. Acute: chronic workload ratio: conceptual issues and fundamental pitfalls. *International journal of sports physiology and performance*, 15, 6, 907-913. https://doi.org/10.1123/ijspp.2019-0864

[43] Janssen, M., Goudsmit, J., Lauwerijssen, C., Brombacher, A., Lallemand, C. and Vos, S. 2020. How Do Runners Experience Personalization of Their Training Scheme: The Inspirun E-Coach? *Sensors*, 20, 16, 4590. https://doi.org/10.3390/s20164590

[44] Janssen, M., Walravens, R., Thibaut, E., Scheerder, J., Brombacher, A. and Vos, S. 2020. Understanding different types of recreational runners and how they use running-related technology. *International Journal of Environmental Research and Public Health*, 17, 7, 2276. https://doi.org/10.3390/ijerph17072276

[45] Johnson, B., Rydal Shapiro, B., DiSalvo, B., Rothschild, A. and DiSalvo, C. Exploring Approaches to Data Literacy Through a Critical Race Theory Perspective. In Proceedings of the Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems).https://doi.org/10.1145/3411764.3445141

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

[47] Karahanoglu, A., van Rompay, T. and Ludden, G. D. Designing for Lifelong Sports Experience. In Proceedings of the Design Research Society 2018 Limerick, Ireland). Design Research Society.https://doi.org/10.21606/drs.2018.413

[48] Kellmann, M. 2010. Preventing overtraining in athletes in high-intensity sports and stress/recovery monitoring. *Scandinavian journal of medicine & science in sports*, 20, 95-102. https://doi.org/10.1111/j.1600-0838.2010.01192.x

[49] Kemler, E., Romeijn, K., Vriend, I. and Huisstede, B. 2018. The relationship between the use of running applications and running-related injuries. *The Physician and sportsmedicine*, 46, 1, 73-77. https://doi.org/10.1080/00913847.2018.1412812

[50] Kenney, W. L., Wilmore, J. H. and Costill, D. L. 2021 *Physiology of sport and exercise*. Human kinetics.

[51] Krawczyk, M. and Wilamowski, M. 2019. Task difficulty and overconfidence. Evidence from distance running. *Journal of Economic Psychology*, 75, 102128. https://doi.org/10.1016/j.joep.2018.12.002

[52] Kruger, J. and Dunning, D. 1999. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77, 6, 1121. https://psycnet.apa.org/doi/10.1037/0022-3514.77.6.1121

[53] Kulmala, J.-P., Avela, J., Pasanen, K. and Parkkari, J. 2013. Forefoot strikers exhibit lower running-induced knee loading than rearfoot strikers. *Med Sci Sports Exerc*, 45, 12, 2306-2313. https://doi.org/10.1249/MSS.0b013e31829efcf7

[54] Kuru, A. 2016. Exploring Experience of Runners with Sports Tracking Technology. *International Journal of Human–Computer Interaction*, 32, 11, 847-860. https://doi.org/10.1080/10447318.2016.1202461

[55] Lacey, A., Whyte, E., O'Keeffe, S., O'Connor, S. and Moran, K. 2022. A qualitative examination of the factors affecting the adoption of injury focused wearable technologies in recreational runners. *PloS one*, 17, 7, e0265475. https://doi.org/10.1371/journal.pone.0265475

[56] Lee, E. C., Fragala, M. S., Kavouras, S. A., Queen, R. M., Pryor, J. L. and Casa, D. J. 2017. Biomarkers in sports and exercise: tracking health, performance, and recovery in athletes. *Journal of strength and conditioning research*, 31, 10, 2920. https://www.doi.org.10.1519/JSC.0000000000002122

[57] Maupin, D., Schram, B. and Orr, R. 2019. Tracking training load and its implementation in tactical populations: A narrative review. *Strength & Conditioning Journal*, 41, 6, 1-11. https://doi.org/10.1519/SSC.0000000000000492

[58] Mauriello, M., Gubbels, M. and Froehlich, J. E. Social fabric fitness: the design and evaluation of wearable E-textile displays to support group running. In Proceedings of the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems).https://doi.org/10.1145/2556288.2557299

[59] McLaren, S. J., Macpherson, T. W., Coutts, A. J., Hurst, C., Spears, I. R. and Weston, M. 2018. The relationships between internal and external measures of training load and intensity in team sports: a meta-analysis. *Sports medicine*, 48, 641-658. https://doi.org/10.1007/s40279-017-0830-z

[60] Meeusen, R., Duclos, M., Foster, C., Fry, A., Gleeson, M., Nieman, D., Raglin, J., Rietjens, G., Steinacker, J. and Urhausen, A. 2013. Prevention, diagnosis and treatment of the overtraining syndrome: Joint consensus statement of the European College of Sport Science (ECSS) and the American College of Sports Medicine (ACSM). *European Journal of Sport Science*, 13, 1, 1-24. https://doi.org/10.1080/17461391.2012.730061

[61] Mencarini, E., Rapp, A., Tirabeni, L. and Zancanaro, M. 2019. Designing wearable systems for sports: a review of trends and opportunities in human–computer interaction. *IEEE Transactions on Human-Machine Systems*, 49, 4, 314-325. http://doi.org.10.1109/THMS.2019.2919702

[62] Miles, M. B. and Huberman, A. M. 1994 *Qualitative data analysis: An expanded sourcebook*. Sage.

[63] Mopas, M. S. and Huybregts, E. 2020. Training by feel: wearable fitness-trackers, endurance athletes, and the sensing of data. *The Senses and Society*, 15, 1, 25-40. https://doi.org/10.1080/17458927.2020.1722421

[64] Nielsen, R. Ø., Parner, E. T., Nohr, E. A., Sørensen, H., Lind, M. and Rasmussen, S. 2014. Excessive progression in weekly running distance and risk of running-related injuries: an association which varies according to type of injury. *journal of orthopaedic & sports physical therapy*, 44, 10, 739-747. https://www.jospt.org/doi/10.2519/jospt.2014.5164

[65] Niemuth, P. E., Johnson, R. J., Myers, M. J. and Thieman, T. J. 2005. Hip muscle weakness and overuse injuries in recreational runners. *Clinical Journal of Sport Medicine*, 15, 1, 14-21. https://doi.org/10.1097/00042752-200501000-00004

[66] Niess, J., Woźniak, P. W., Abdelrahman, Y., ElAgroudy, P., Abdrabou, Y., Eckerth, C., Diefenbach, S. and Knaving, K. 'I Don't Need a Goal': Attitudes and Practices in Fitness Tracking beyond WEIRD User Groups. In Proceedings of the Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction).https://doi.org/10.1145/3447526.3472062

[67] Palsa, L. and Mertala, P. 2023. Contextualizing everyday data literacies: The case of recreational runners. *International Journal of Human–Computer Interaction*, 1-12.

[68] Paquette, M. R., Napier, C., Willy, R. W. and Stellingwerff, T. 2020. Moving beyond weekly "distance": optimizing quantification of training load in runners. *journal of orthopaedic & sports physical therapy*, 50, 10, 564-569. https://www.jospt.org/doi/10.2519/jospt.2020.9533

[69] Peake, J. M., Neubauer, O., Della Gatta, P. A. and Nosaka, K. 2017. Muscle damage and inflammation during recovery from exercise. *Journal of applied physiology*, 122, 3, 559-570. https://doi.org/10.1152/japplphysiol.00971.2016

[70] Powers, S. K., Howley, E. T. and Quindry, J. 2007 *Exercise physiology: Theory and application to fitness and performance*. McGraw-Hill New York, NY.

[71] Rapp, A. and Cena, F. 2016. Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. *International Journal of Human-Computer Studies*, 94, 1-17. https://doi.org/10.1016/j.ijhcs.2016.05.006

[72] Rapp, A. and Tirabeni, L. 2018. Personal Informatics for Sport: Meaning, Body, and Social Relations in Amateur and Elite Athletes. *ACM Trans. Comput.-Hum. Interact.*, 25, 3, Article 16. https://doi.org/10.1145/3196829

[73] Rapp, A. and Tirabeni, L. 2018. Personal informatics for sport: meaning, body, and social relations in amateur and elite athletes. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25, 3, 1-30.

[74] Rapp, A. and Tirabeni, L. 2020. Self-tracking while Doing Sport: Comfort, Motivation, Attention and Lifestyle of Athletes Using Personal Informatics Tools. *International Journal of Human-Computer Studies*, 102434. https://doi.org/10.1016/j.ijhcs.2020.102434

[75] Reenalda, J., Maas, M. T. and de Koning, J. J. 2016. The influence of added mass on optimal step length in running. *International journal of sports physiology and performance*, 11, 7, 920-926. https://doi.org/10.1123/ijspp.2015-0182

[76] Restrepo-Villamizar, J., Verhagen, E. and Vos, S. Defining the Individual Injury Profile of Recreational Runners: Integrating Off-Training and Subjective Factors into the Assessment of Non-Professional Athletes. In Proceedings of). Multidisciplinary Digital Publishing Institute Proceedings (MDPI).https://doi.org/10.3390/proceedings2020049087

[77] Restrepo-Villamizar, J., Vos, S., Verhagen, E. and Lallemand, C. 2022. Hyaku: A Qualitative Negotiation-Through-Interaction Interface to Support Runners in Achieving Balanced Training Sessions. In *Proceedings of the Proceedings of DRS* (Bilbao, Spain). https://doi.org/10.21606/drs.2022.569

[78] Rheu, M., Shin, J. Y., Peng, W. and Huh-Yoo, J. 2021. Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human–Computer Interaction*, 37, 1, 81-96. https://doi.org/10.1080/10447318.2020.1807710

[79] Rooksby, J., Rost, M., Morrison, A. and Chalmers, M. 2014. Personal tracking as lived informatics. In *Proceedings of the Proceedings of the SIGCHI conference on human factors in computing systems*. https://doi.org/10.1145/2556288.2557039

[80] Russell, D. M., Furnas, G., Stefik, M., Card, S. K. and Pirolli, P. 2008. Sensemaking. In *Proceedings of the CHI'08 Extended Abstracts on Human Factors in Computing Systems*. https://doi.org/10.1145/1358628.1358972

[81] Saragiotto, B. T., Yamato, T. P., Hespanhol Junior, L. C., Rainbow, M. J., Davis, I. S. and Lopes, A. D. 2014. What are the Main Risk Factors for Running-Related Injuries? *Sports Medicine*, 44, 8, 1153-1163. https://doi.org/10.1007/s40279-014-0194-6

[82] Snooks, K., Whitham, R., Richards, D. and Lindley, J. 2022. Beyond the body: Moving past the metricised bodily goal in self-tracking. In *Proceedings of the DRS2022* (Bilbao, Spain, 25 June - 3 July 2022). https://doi.org/10.21606/drs.2022.501

[83] Soligard, T., Schwellnus, M., Alonso, J.-M., Bahr, R., Clarsen, B., Dijkstra, H. P., Gabbett, T., Gleeson, M., Hägglund, M. and Hutchinson, M. R. 2016. How much is too much?(Part 1) International Olympic Committee consensus statement on load in sport and risk of injury. *British journal of sports medicine*, 50, 17, 1030-1041. http://dx.doi.org/10.1136/bjsports-2016-096581

[84] Sperlich, B., Aminian, K., Düking, P. and Holmberg, H.-C. 2020. Wearable sensor technology for monitoring training load and health in the athletic population. *Frontiers in Physiology*, 10, 1520. https://doi.org/10.3389/fphys.2019.01520

[85] Swann, C., Moran, A. and Piggott, D. 2015. Defining elite athletes: Issues in the study of expert performance in sport psychology. *Psychology of Sport and Exercise*, 16, 3-14. https://doi.org/10.1016/j.psychsport.2014.07.004

[86] Tholander, J. and Nylander, S. 2015. Snot, Sweat, Pain, Mud, and Snow: Performance and Experience in the Use of Sports Watches. In *Proceedings of the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea). ACM. http://10.1145/2702123.2702482

[87] Tolmie, P., Pycock, J., Diggins, T., MacLean, A. and Karsenty, A. Unremarkable computing. In Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems).https://doi.org/10.1145/503376.503448

[88] Torres-Ronda, L., Beanland, E., Whitehead, S., Sweeting, A. and Clubb, J. 2022. Tracking Systems in Team Sports: A Narrative Review of Applications of the Data and Sport Specific Analysis. *Sports Medicine-Open*, 8, 1, 1-22. https://doi.org/10.1186/s40798-022-00408-z

[89] van Erp, T., Foster, C. and de Koning, J. J. 2019. Relationship between various training-load measures in elite cyclists during training, road races, and time trials. *International journal of sports physiology and performance*, 14, 4, 493-500. https://doi.org/10.1123/ijspp.2017-0722

[90] Van Hooren, B., Goudsmit, J., Restrepo, J. and Vos, S. 2020. Real-time feedback by wearables in running: Current approaches, challenges and suggestions for improvements. *Journal of Sports Sciences*, 38, 2, 214-230. https://doi.org/10.1080/02640414.2019.1690960

[91] Van Mechelen, W. 1992. Running injuries. *Sports medicine*, 14, 5, 320-335. https://doi.org/10.2165/00007256-199214050-00004

[92] van Poppel, D., van der Worp, M., Slabbekoorn, A., van den Heuvel, S. S., van Middelkoop, M., Koes, B. W., Verhagen, A. P. and Scholten-Peeters, G. G. 2021. Risk factors for overuse injuries in short-and long-distance running: a systematic review. *Journal of sport and health science*, 10, 1, 14-28. https://doi.org/10.1016/j.jshs.2020.06.006

[93] Videbæk, S., Bueno, A. M., Nielsen, R. O. and Rasmussen, S. 2015. Incidence of running-related injuries per 1000 h of running in different types of runners: a systematic review and meta-analysis. *Sports medicine*, 45, 1017-1026. https://doi.org/10.1007/s40279-015-0333-8

[94] Willwacher, S., Kurz, M., Robbin, J., Thelen, M., Hamill, J., Kelly, L. and Mai, P. 2022. Running-Related Biomechanical Risk Factors for Overuse Injuries in Distance Runners: A Systematic Review Considering Injury Specificity and the Potentials for Future Research. *Sports Medicine*, 1-15. https://doi.org/10.1007/s40279-022-01666-3

[95] Willy, R. W. 2018. Innovations and pitfalls in the use of wearable devices in the prevention and rehabilitation of running related injuries. *Physical Therapy in Sport*, 29, 26-33. https://doi.org/10.1016/j.ptsp.2017.10.003

[96] Woźniak, P., Colley, A. and Häkkilä, J. 2018. Towards Increasing Bodily Awareness During Sports with Wearable Displays. In *Proceedings of the Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore). Association for Computing Machinery. https://doi.org/10.1145/3267305.3267703

[97] Yang, Q., Steinfeld, A. and Zimmerman, J. 2019. Unremarkable AI: Fitting intelligent decision support into critical, clinical decision-making processes. In *Proceedings of the Proceedings of the 2019 CHI conference on human factors in computing systems*. https://doi.org/10.1145/3290605.3300468