Towards an AI-Based Tailored Training Planification for Road Cyclists: A Case Study

Alessandro Silacci [1,2,*] **, Redha Taiar** [3] **and Maurizio Caon** [1]

1   School of Management Fribourg, HES-SO, 1700 Fribourg, Switzerland; maurizio.caon@hes-so.ch
2   Department of Information Systems, University of Lausanne, 1015 Lausanne, Switzerland
3   Department of Sport Science, University of Reims, 51100 Reims, France; redha.taiar@univ-reims.fr
*   Correspondence: alessandro.silacci@hes-so.ch; Tel.: +41-76-574-45-83

Abstract: In a world where the data is a central piece, we provide a novel technique to design training plans for road cyclists. This study exposes an in-depth review of a virtual coach based on state-of-the-art artificial intelligence techniques to schedule road cycling training sessions. Together with a dozen of road cycling participants' training data, we were able to create and verify an e-coach dedicated to any level of road cyclists. The system can provide near-human coaching advice on the training of cycling athletes based on their past capabilities. In this case study, we extend the tests of our empirical research project and analyze the results provided by experts. Results of the conducted experiments show that the computational intelligence of our system can compete with human coaches at training planification. In this case study, we evaluate the system we previously developed and provide new insights and paths of amelioration for systems based on artificial intelligence for athletes. We observe that our system performs equal or better than the control training plans in 14 and 24 week training periods where it was evaluated as better in 4 of our 5 test components. We also report a higher statistical difference in the results of the experts' evaluations between the control and virtual coach training plan (24 weeks; training load: $X^2 = 4.751$; resting time quantity: $X^2 = 3.040$; resting time distance: $X^2 = 2.550$; efficiency: $X^2 = 2.142$).

Keywords: virtual coaching; e-coach; artificial intelligence; training plan; road cycling

1. Introduction

Performance is tracked and optimized everywhere at any time. The sports world is an area where one can observe this ever growing need to perform better assessed by the growing market of mobile performance applications (see Deloitte 2017's report). In the business field, companies try to maximize their benefits by increasing their performance in multiple levels [1]. The common aspect between these two fields is that one can track the performance and get an idea of it through quantitative information. Data is nowadays ubiquitous, and devices are always collecting data. This makes it easy to access micro-quantitative and visual feedback of our performance but does not necessarily explain the meaning in a macro perspective. Indeed, amateurs as well as professionals in the sports field are relying on wearable devices as a motivational object, but more importantly to perform better [2].

The growing market of wearables devices enables the tracking of new body-related information to be more precise and more detailed [3]. These ameliorations enable one to get a deeper quantitative feedback, and might also overwhelm the athletes with numerical and complex data, reducing their capabilities to understand and use it properly [4]. This high amount of data can lead the athletes to confusion or misunderstanding and could even make them lose motivation towards using such tools [5].

Making sense of one's data may become a complicated task. The athlete's level will also determine how one interacts and uses the collected data. However, one can fall into overtraining syndrome, which is experienced when a person trains disproportionally

and ends up training in an inappropriate way compared to their capabilities and what their body can manage [6]. Additionally, such syndrome can occur when the equilibrium between the training load and the resting time is not respected and can lead to critical situations, both related to physical and mental health [7]. Amateur athletes are also more prone to develop such syndrome, as they do not benefit from coach feedback on the way they train [6].

Personalized training feedback may be seen as a way to mitigate the issues induced by performance tracking in sports. Professional athletes rely on human coaches with expertise in managing training, efforts, and resting time. We suppose that amateur and semi-professional athletes are not interested in the price of a human expert to manage their training as this would be too much for their needs. This level of athletes use training software that lacks a proper personalization component and may experience a decrease in their motivation past a certain time [8].

In our study, we present a novel approach to training management. We leverage on new artificial intelligence-based work in order to create a virtual coach with training personalization capabilities. The technique we use is able to tailor and take into consideration any level of athletes, as it bases the feedback instructions on the capabilities of its users. Using state-of-the-art technologies and specific measurements, we ensure a proper and convenient training plan designed and managed around the athletes. Thus, we provide amateurs and semi-professionals with a convenient training coaching system, enabling them to get expert-like instructions and feedback. Additionally, this solution can provide human coaches with a different point of view on their athletes' training and introduce more diversity in training plans.

2. Related Work

Through our research work, we reviewed an interesting collection of research papers that demonstrated the trends in computer intelligence applied to sports. We particularly looked at the presence of machine learning and algorithms applied to cycling. The results of the research still yielded additional sports that we comprised in this review. We coded the results of our state-of-the-art review according to Lv. L. and Ye. C.'s categories [9]. We augmented the category set by adding new groups of systems, as reported in Table 1.

Table 1. Applications of machine learning in sports.

Applications	Number of Papers
Training optimization	10
Quantifying athlete's performance	6
Predicting results of athlete	6
Training motivation	4
Injury prevention	2
Rule control	1
Gear settings optimization	1

As a general statement, we report a higher involvement of researchers into systems dedicated to endurance sports, like running [10–13] and cycling [13–18,18–20]. One of the main reasons for this is that these physical activities are quite easy to practice and affordable. Thanks to this ease of access and higher popularity, they both benefit a wider range of sensors that can track the athletes [21]. These sports are also ones that may require less data manipulation in order to work with machine learning. Indeed, rankings in running and cycling competitions are based on incremental scores [14], and one can observe changes in the athletes' performance through the collected timeseries data coming from many different sensors [21].

2.1. Usable Features for Sports

A primary data source for performance measurement is sensors. They can be integrated in specific devices or by using the ones embedded in smartphones. Raw data

from these sensors may be used, but they are usually processed in order to remove noise. Sensors provide data with a high frequency and thus may not be used directly. For example, information such as the 3-axis acceleration enables the understanding of a movement relative to its past state but can highly fluctuate across short periods of time since the sensors are particularly sensitive. Most of the time, we observed a higher usage of feature combinations, where features get merged in order to create a new one, a process commonly known as feature extraction. Such measurements are time-related and are highly impacted by past values. In cycling, for instance, one can use Training Stress Score (TSS) to explain the effort required by a training session [22]. Using sensors allows measuring signals without intruding the user's activity and, therefore, without impact on their performance. In contrast, measurements based on self-reporting such as Rating of Perceived Exertion (RPE) require that the athlete provide the required information in an explicit manner. For the RPE, the user should have a certain knowledge of its current capabilities as they have to report the perceived exertion on a 20-point scale. Despite being more accurate than other sensors' measurements, the RPE is hard to use [19].

Systems are no longer dependent on wearable sensors, unlike the ones used for running or cycling. Researchers also calculated statistics based on players or teams' actions and scores, since sports like football where access to a video of matches in high-ranking teams is quite difficult for external users to access. Thus, statistical data from the matches are used in order to predict future scores [23]. Similarly, in the case of fitness, the data gathered on the athlete cannot accurately describe their effort; rather, one needs to use additional devices to track the movements and state on the machine [24,25]. Features explaining the performance of a team or an individual may not be directly related to some sensors data. In fact, one can understand the performance of a football team by checking their scores and enrich the statistics with additional information on whether it was a home win or an away win [26]. External information feeds can also help to understand one's performance, since people share a lot of information through social media platforms. Thus, mining the information shared on social media and using Natural Language Processing (NLP) can provide a rich source of feedback data about an athlete's performance [11].

Sports where the body position has a high impact on performance is also being helped by computer vision capabilities as well as recent research in deep learning. Indeed, it is possible to extract a great number of features and information from a camera feed. The biomechanical data can be treated in order to enhance the movements or correct the postures of the athletes and thus help them perform better. Sports such as golf, tennis or javelin are benefitting from these techniques to track the athlete's position and treat it through image processing [27–29]. One's body shape may partially define their ability to perform at a certain level, as well as the extent to which one can perform a gesture. We observed the usage of data coming from specific sensors or measurements made to explain the current organs and physiological status of an athlete. We obviously find a high usage of the heartrate, despite being criticized for its high variance and dependence on the athlete's form [19]. While heartrate explains the evolution of one's heart beats per minutes, other features are not sensor-based and may require experts to perform measurements. For example, kinanthropometric data is a set of information composed of the body size, its shape or even its composition and may be used to evaluate one's potential performance [30].

In Table 2, we summarize the observed features and classify using our own taxonomy. Compared to outer body-related data, the inner body information may vary faster. The two categories are linked but still separated by a thin line, which is the latter's fluctuation across a short period of time. Athlete-related data make a clear categorization of the athlete using their age, sex and anthropometric information. Additionally, computer vision (CV) based features can be extracted from video frames processing. These are measurements based on the skeleton's joint position and enable the understanding of one's body structure and position through time.

Table 2. The categorization of used data in sports dedicated systems.

Sensor-Based	Computer Vision	Installation Specific	Inner Body	Ambient	Athlete Related	Subjective Measurements	Performance Data	Injury Related	External Sources	Computed Values
Acceleration	Video	Displacement of reference point	HR	Humidity	Sex	Feelings	Phase of play	Injury type	Social Media	TSS
Gyroscope	X-Factor movement	Cable force	Strength	Location	Age	RPE	Stage of the season	Injured side		TRIMP
Velocity	Skeleton joint positions		Flexibility	Temperature	Anthropometric measures		Home win	Injured body part		VO2max
Power			Technical efficiency				Away win	Reoccurrence		Exercise economy
Distance			Force development				Results history	Days unavailable		Anaerobic capability
Speed			Kinanthropometric measures				Average scores pro/con	# Injured players		Lactate threshold
			Body mass				Overall performance	Player availability		Fatigue
			Capillarization				Rankings			Max strength
			Oxidative enzyme activity							Max speed
										Endurance performance
										Diffusion distance

A high number of sports rely on the athlete's position in time, thus the information may be key to an enhancement in performance [27]. Processing video feeds can provide a high amount of information in professional-level sports, since the data is undisclosed due to the competition aspect.

2.2. Machine Learning Usage in a Sports Context

Not all machine learning techniques can be used to treat sports problems. Through the papers we reviewed, we found some trends in the selected techniques. The type of data, the number of features and the distribution of the data are some of the most determinant aspects to consider while choosing a machine learning model. Thus, depending on the features and the quantity of available data, researchers privileged some models among others. In Table 3, we present the count of machine learning techniques used in the reviewed papers.

Table 3. Comparison of machine learning models usage in sports.

Techniques	Number of Papers
ANN	7
SVM	3
CV	3
Clustering	3
Fuzzy logic	3
Naive bayes	3
Feature selection/extraction	2
Ontology	2
K-NN	2
Random forest	1
LogiBoost	1

Artificial Neural Network (ANN) are the most used approaches of our review. This is due to recent advances in deep learning and ANN-based models, but mainly its generalization capability. ANNs tend to be, when well used, models that can treat any type of data. The architecture is flexible, as one can parametrize the number of neurons, the layer types (which depends on the data type), and the starting weights. The only issue is that ANN are known to be overused and, in some cases, may also overfit the dataset quite easily. The Support Vector Machine (SVM) approach is very similar to ANN in terms of generalization capabilities. As for ANNs, these models can also support high amounts of data. We are surprised that there are not more papers using this approach. Indeed, the use of SVM ensures avoiding overfitting on the available data. On the contrary to supervised machine learning techniques, unsupervised machine learning tries to solve one of the main issues of creating a dataset, which is data labelling. The goal of adopting such techniques is generating clusters of data that have similar characteristics or values. The finality of such an approach is mostly an explanation of the dataset's content where one can find and extract trends or patterns. They seem to be quite efficient at determining winning strategies in multi-staged competitions [14,15].

We also accounted papers using some older techniques to interpret sports' data, such as fuzzy logic and ontologies. Both techniques provide results that can be easily understood and may also be used with unlabeled data. Fuzzy logic coupled with fuzzy inference can provide an idea of the different states, or classes, of the given data [25]. Thus, one could use it to extract classes from a given dataset that may not be initially understandable. In counterparts, ontologies are used to have a defined number of states and transitions and one uses semantic reasoning to apply the data to it and extract results. The latter technique will adapt the states and their transition to a dataset, thus it can be used to provide recommendations of another one's performance with a certain degree of flexibility [11,17].

Computer vision (CV) is also not directly linked to machine learning techniques and can beused when one needs to extract data from image processing in order to construct a

statistical analysis or a usable dataset for machine learning techniques [31]. We found many applications (mainly in sports) where it was not possible to use on-bodied sensors to track the user's movements. Thus, the tracking of data such as the skeleton joints can provide information on gestures and postures performed by the athletes and react to them [27]. CV can also be used to track objects' movements and explain specific behaviors [28].

2.3. Sports Coaching Based on Machine Learning

Virtual coaches can be defined as, "computer systems capable of sensing relevant context, determining user intent and providing useful feedback with the aim of improving some aspect of the user's life" [32]. We observed that virtual coaching has evolved as fast as machine learning research, enabling management of larger quantities of information, more data types and offering new models to rely on. However, we hereafter point out aspects linked to coaching and a common gap of all reviewed solutions. E-coaching consists of virtual support for human real-life activities and it can be deployed in a plethora of different contexts, including sports.

A coaching support can be provided in many ways, but it is important that the medium chosen for the interaction with the user still needs to be adequate to its application, especially in sports since athletes may not get the information in all contexts (i.e., before, during or after their effort). We relied on the modalities proposed by [33] to construct our synthesis:

- audio communication;
- video communication;
- synchronous text-based communication;
- asynchronous text-based communication.

Audio communication is particularly interesting in sports, since it allows conveying information of the current performance to an athlete without engaging them in a high workload activity. We define high workload as any activity where the person needs to think and focus in order to get information. On the contrary, a video communication system represents a high workload during the effort, since the athlete needs to focus on the screen and not on the effort they are making. However, head-mounted displays may reduce the induced workload as the athlete's focus on their main activity is only overlaid by a small piece of information, such as direction information [16]. Video cues and indications may be more efficiently used in sports or periods of effort where one is not stressed by a time constraint or can easily switch their focus. Sports such as weight lifting or trail track running benefit from video-based indications, since the athlete has time to focus on a screen before performing an exercise [34]. Additionally, indoor and individual sports do not provide the same triggers for athletes' performances. Thus, involving the user in a totally virtual environment can benefit their performance at particular tasks in sports like home cycling [35].

Text-based communication, either synchronous or asynchronous, is presenting the same issues as video communication. Additionally, they may not enable one to provide as rich information as in the audio or video communication media. However, people rely on text for off-activity information gathering; thus, an athlete could easily look at coaching information while not practicing sports. Systems like training planners are only available through text-based communication since they are used pre or post the sports activity [11,20,36].

Athletes use coaches to push themselves to their limit and enhance their performance. Virtual coaching can provide motivational cues in an implicit or more explicit manner depending on the communication medium and the coaching type used. Team sports may benefit from the group effect as a motivation source but can also be enhanced by adding virtual information to the group [12,18,37]. Although video communication is not recommended in certain sports, it can provide additional and richer information than audio or text-based feedback. People feel more comfortable and motivated when using a video-based coach since the emotional engagement may be helped by facial micro-

expressions [38]. Additionally, it is easier to get a correct execution of exercises using a video or virtual character showing the movements, rather than explaining them via text or using audio information [39].

We observe a trend in the use and availability of computational intelligence-based systems to support athletes in sports. In our review, systems were directed at assisting the athletes in order to correct or provide feedback on their performance. This growth is also explained by the increasing capabilities of machine learning and artificial intelligence, enabling new analysis and modeling of the athletes' performance. However, we identified a gap in these systems, since no research was properly leveraging on these techniques to properly personalize its feedback to the athlete. Thus, we introduce a new reinforcement learning-based virtual coach. The main aim of this study is to test our virtual coach training plans on the following five different components: the training sessions' distribution, training load, resting time quantity, resting time distribution, and efficiency. We involve professional coaches into the process of evaluation and compare our approach for training planification to an established training coaching platform and gather quantitative and qualitative results that we further analyze.

3. Materials and Methods

We involved six road cycling experts to provide a qualitative and quantitative feedback on three series of training plans. We generated the training plans using our virtual coaching system that bases the planification on a balance of TSS and TSB values. The system functioning is further explained in a previous research paper [36]. The experts are provided with three different training plans each designed for a specific training period. Thus, coaches have to evaluate plans for the following training periods:

- 42 days (6 weeks)
- 98 days (approximately 14 weeks)
- 168 days (approximately 24 weeks)

The selected time periods were dependent on the availability of training plans provided by TrainingPeaks' coaches as we use it as a control training plan platform. Using the selected time periods, we generated new virtual coach-based training plans and anonymized both training sources. We have built the training plan and provided the anonymized data of a semi-professional road cyclist after consent agreement.

Both control and virtual coach training systems were provided with the athlete's training data in order to build up the training. The training plans were composed of a date, a session type and the TSS score. The TSB score was recorded at some point in the time period in order to inform the participants about the timeline of the sessions, their type, the load induced by the session and its implication on the athletes' overall form. In the case of our virtual coach system, we selected the most rewarded attempt at planning for each time period [36].

The proposed sessions are standardized in a set of categories: endurance, anaerobic intervals, recovery and rest. This implies that the sessions proposed in the control training plans are converted into these categories. We further explain the categorization process in the following section.

The selected control training plans were described as not intended for professional athletes and we qualitatively evaluated them as adequate in terms of the proposed sessions for our athlete's profile. Additionally, these training plans were also selected because they were not merging too many sessions of different sports in order to build up performance. Thus, at the end of our selection process we have chosen the following training plan as the control:

- 6 weeks: "Peak Criterium Racing Form in 6 Weeks! Cat 2, 3 & 4, Masters (Workout Builder), 8 h/week," from Simon Kessler;
- 14 weeks: "Your Best 75 Mile Sportive/GF Ever-40 Week S&C Plan, Support Groups & $25 off next plan!," from Pav Bryan; and

- 24 weeks: "Intermediate Peak form and Base for faster Criterium + email access to Coach (WKO4i Levels)," from Richard Rollinson.

3.1. The Training Evaluation Questionnaire

The experts were provided an online questionnaire that mixed quantitative questions with a 5-point Likert scale and more qualitative ones to gather comments and insights on cycling coaches. The set of questions was the same for each time period to allow a comparison across them. The questionnaire was composed of the following questions:

1.　Which of the two planning seems the most appropriate for this period of time?
2.　Explain your choice in the previous question.
3.　Please evaluate the training sessions' distribution (1 = very bad, 5 = very good).
4.　Please evaluate the training load for the two proposed planning (1 = very low, 5 = very high).
5.　Please evaluate the resting time quantity (1 = not enough, 5 = too much).
6.　Please evaluate the resting time distribution (1 = not enough, 5 = too much).
7.　Please evaluate the efficiency of the proposed planning (1 = very bad, 5 = very good).
8.　Additional comments on Planning 1?
9.　Additional comments on Planning 2?

The identification of the two proposed planning was made using "Planning 1" and "Planning 2" as names for each of the set time period. As the set of provided planning for each time period did not mention if the planning represented either the control or the virtual coach training plans, both could be identified as either Planning 1 or Planning 2 in each time period. The training plans in each time period were randomly attributed to the two particular identifiers.

3.2. TrainingPeaks' Converted Sessions

For comparison purposes between the two proposed training plans, we use a defined set of training types. Thus, some of the control plan session names had to be adapted. Since each planning period had specific workout names, we transformed their names depending on their initial name (if it mentioned the training type), the effort distribution along the session and the session's TSS load.

Hereafter are presented the original sessions' names and the converted ones with an additional explanation as to why they were categorized like this. The six week planning sessions converted as anaerobic intervals were threshold intervals, sub threshold + VO2/AP, threshold + VO2 intervals, sprint (15 and 30 sec sprints), sprint training, VO2 intervals, threshold + 30/30 intervals, sub threshold plus breakaway, combination big gear + sprint/finish, endurance with on/off sprints, tempo progression, and sprint + VO2 (race simulation). We motivate this categorization by observing the training intensity fluctuation during the training. All these sessions were composed of scheduled periods of different intensities denoting intervals.

In the same control training plan, we further grouped the following sessions under endurance: combination big gear + sprint/finish, endurance with on/off sprints, tempo progression and sprint + VO2 (race simulation). Even if these sessions were segmented, we considered the fact that their duration was longer than what was presented in intervals. The intensity in these sessions was mainly set to be the same during a long period of time.

Recovery sessions were considered by their low TSS load and are meant to allow the athlete to recover. We considered the following sessions as being part of the Recovery category: recovery + sprints, recovery + high cadence and accelerations, sprints on/off (10 s), and pre-race (one day before).

Adaptations to the 14-week control training plan were conducted. We categorized the later sessions as being anaerobic intervals: spin ups, 3′ aerobic test, strengthreps (set 1; 8 × 2.5′ FTP(HR); 2 × 10′ FTP(HR) into 10′ zone 2), FTP Test, STME intervals (one minute; FTP(HR) over/unders), hill reps–set, short spin, activation session and hilly ride-low RPM. The main reason for this categorization is the presence of multiple segments with

alternated intensities, which happened either in the whole training session or in most of it. We combined two training sessions (core strength and flexibility workout and leg speed set 1–short), as Recovery, since the combined TSS value was low.

In the same training period, we categorized a small number of sessions as Endurance due to the long segments in the same TSS value. Thus, the following sessions were considered Endurance: activation session and hilly ride-low RPM.

The last training period (24-weeks) was composed of a higher number of sessions considered anaerobic intervals compared to the rest of the planning we used previously: anaerobic capacity FR × 4; sprints × 4; Vo2 max 10 m × 3; tempo intervals 2 × 15 m; FTP 2 × 20 m; PD curve test: short; PD curve test: medium; force strength; optimized max (optimized intervals) × 5; extensive anaerobic (optimized intervals) × 4; sweet spot intervals 2 × 45 m; max aerobic (optimized intervals) × 4; Vo2 max micro 6 × 8 m; anaerobic capacity FR 2 × 5/30 s 1:2; sprints FR 2 × 10; velocity speed sprints × 6; sweet spot matches 2 × 30 m; intensive aerobic (optimized intervals) × 2; intensive anaerobic (optimized intervals) × 4; and opener 1 h. Again, the fact that these sessions were considered intervals was mainly due to their repeated intensity segment alternations during the session.

For this last training period, we identified the following training sessions as pertaining to the Endurance session type: hills >400 m 1 h 30 m; velocity skill 120 rpm; PD curve test: long; endurance 1 h 30 m; tempo intervals 1 × 45 m; goup (or Dartlek Solo) 2 h hard; rolling tempo 1 × 55 m; and race: circuit 1 h. These sessions were all composed of a single or two long segments of the same intensity. Surprisingly, we could not find any session to define as Recovery in the 24-week training plan.

3.3. Measurements

Results are reported in a contingency table for each question having a Likert-scale answer for each condition. The results are grouped into a positive part, a neutral part and a negative part where we summed values that were in the positive middle part (between 3 and 5), the neutral part (3) and the negative middle part (between 3 and 1). Additionally, we provide an χ^2 value for each table in order to provide information on the dependency of the results we gathered. We also gather qualitative data on the reasons people have preferred one training plan over another for a certain period, as well as their thoughts on the proposed planning.

4. Results

The experiment was conducted with six sports coaching experts in the cycling domain. As we extend our previous research, the study was done across the span of one year, but the data and the questionnaire remained untouched.

In this section, we first present the results in terms of suitability of the proposed training plans for each training period. Then, we analyze the results through contingency tables for each question and each training period.

4.1. Suitability of the Training Plans

The binary answer provided by each participant informed which planning was better in a given period of training. We asked the coaches for a detailed explanation to understand their choice for each training period. The difference was higher in the six week training period, where 60% of the participants selected control planning. From the qualitative data collected on the reasons that they chose the latter, three of the experts mentioned too high TSS and TSB values in the virtual coach's proposed planning. Most importantly, one mentioned that the TSS values of 457 must not be present in a training plan. The smoother approach to the TSB increase presented by the control training plan has been described by two of these three coaches as a more adequate approach to training. The last one that voted for the control planning mentioned the diversity of the trainings, while the two participants who voted for the virtual coach plan mentioned the session distributions as "slightly better" and "more polarized" compared to the control plan.

Results for the same questions across a longer time period demonstrated less discrepancy between the two training plans. Indeed, three coaches voted for the control planning as the best solution for the 14 weeks planning period, while three others voted for the virtual coach training plan. Similar to the previous virtual coach-based plan, the presence of training sessions with a 457 TSS value has been criticized as being too high. The control planning was mentioned as having a more gradual approach with easier workout sessions in terms of TSS and avoids TSB peaks in the high-risk zone (Blog post by Joel Friel: https://joefrielsblog.com/managing-training-using-tsb/). The last control planning vote provided the motive that sessions were more diverse and proposed a higher number of anaerobic training. However, it seems that coaches all have their preferred style and approach, as one of them mentioned that the virtual coach planning better suited the training period since it proposed a higher amount of endurance sessions relative to intervals. Another vote was given to the virtual coach plan as the participant noted the gradual increase over the period as preferable. The same expert also mentioned that the planning was better if one planned to build a strong foundation for long-term training. The last coach who voted for the virtual coach mentioned that the control training plan was too poor in terms of the provided training load.

As in the previous situation, the 24 week training plan proposed as the control received three votes, compared to three votes for the virtual coach-based one. The same expert voted for the control as the virtual coach was presenting 457 TSS sessions. Another participant for the control planning mentioned that the planning was, "more efficient without tiring the rider so much in each session," while the last one that also voted for the latter plan mentioned the main presence of anaerobic intervals. One coach mentioned that the virtual coach planning would be more interesting for long-term performance building and argued that this training plan could also be perceived as "more encouraging from a moral perspective" with the help of small peaks. Another participant mentioned the fact that the interval quantity was too important in the control training plan and opted for the virtual coach proposition despite a mentioned high amount of endurance sessions. Lastly, the virtual coach planning proposed a fewer number of interval sessions than in the control plan, which resulted in the second vote for the virtual coach-based training plan.

4.2. Training Sessions' Distribution

The second question of our questionnaire targeted the distribution of the sessions through the week or the whole training period. The collected data provide insights on the way that the virtual coach schedules and the number of a certain type of training sessions, compared to the distribution of the sessions in the control training plan. In Table 4, we present the results of the 6 week training plans in a contingency table.

Table 4. The contingency table of the 6 week sessions' distribution.

$\chi^2 = 13.897$	Good	Neutral	Bad	Total
VC	0	3	8	11
Control	12	3	4	19
Total	12	6	12	30

We observe that the virtual coach's sessions' distribution across this time period appears worse than the control. However, according to the χ^2-test, the results are not significantly different ($\chi^2 = 13.897$). We observed no statistical difference in the two other time periods of planning (14 weeks: $\chi^2 = 20.925$ and 24 weeks: $\chi^2 = 6.303$). Comparing the contingency tables, the distribution of the virtual coach sessions revealed to get better with longer spans as demonstrated by Table 5, but still could not match the distribution quality of the control training plans.

Table 5. The contingency tables of the 14 and 24 week sessions' distribution.

	14 Weeks					24 Weeks			
χ^2 = 20.925	Good	Neutral	Bad	Total	χ^2 = 6.303	Good	Neutral	Bad	Total
VC	0	9	4	13	VC	0	12	2	14
Control	12	0	4	16	Control	4	6	2	12
Total	12	9	8	29	Total	4	18	4	26

Thus, we conclude that the system has to be enhanced on the distribution of the sessions, but we see an enhancement as the training period is augmented.

4.3. Training Load

Experts were asked to rate the training load of the two proposed training plans for the three given periods. Results demonstrate no significant difference between both conditions in the 6 and 14 week training periods, while the results from the 24 week planning were significantly different (χ^2 = 6.303). In the latter time period, the virtual coach was outperforming the control plan by having a lower training load according to the participants. Table 6 shows the resulting contingency table for the 24 week planning period, where the control plan has been defined by experts as providing a training load that is too high. Comparing the training plans, we observed that the control plan has maintained peaks at a higher TSS value, while the virtual coach-generated planning provided peaks with 457 TSS, but still maintained some training sessions with lower TSS in between.

Table 6. The contingency table of the 24 week training load.

χ^2 = 4.751	High	Neutral	Low	Total
VC	4	9	2	15
Control	12	3	2	17
Total	16	12	4	32

The same results appeared in the two other training periods, where the virtual coach generated a training plan that has been evaluated as either more balanced or in the lower range of values than the control training plan. We also observed that most coaches evaluated the control training plan as providing a training load that was too high compared to the virtual coach, in the 14 week training period (three times given a four for the control).

In summary, one could say that the training load was more balanced in the training plans provided by the virtual coach in the longest training period of 24 weeks.

4.4. Resting Time Quantity

The results obtained on the evaluation of the resting time quantity demonstrated a statistically significant difference for all the training periods. In the 14 and 24 week time period, the virtual coach training plans were more balanced in terms of resting time quantity and were evaluated as providing a more decent amount of it compared to the control plans, as reported in the contingency tables of Table 7.

Table 7. The contingency tables of the 14 and 24 week resting time quantity evaluation.

	14 Weeks					24 Weeks			
χ^2 = 1.903	Too Much	Neutral	Not Enough	Total	χ^2 = 3.040	Too Much	Neutral	Not Enough	Total
VC	8	9	2	19	VC	12	9	0	21
Control	12	6	2	20	Control	13	6	2	21
Total	20	15	4	39	Total	25	15	2	42

In the shortest time period, the control training plan provided a more satisfying resting time quantity compared to the virtual coach. We observed that the coaches mainly evaluated it as "neutral" in the control with four of them providing a three, as demonstrated in Table 8.

Table 8. The contingency table of the 6 week resting time quantity evaluation.

$\chi^2 = 4.399$	Too Much	Neutral	Not Enough	Total
VC	4	12	2	18
Control	0	15	2	17
Total	4	27	4	35

The resting time quantity reveals to be more adequate and balanced for the training proposed by the virtual coach. We observed better results in the 14 week training plans, while both propositions were judged as excessively providing resting time (compared to their respective planning). In the 6 week time period, the control training plan outperformed the virtual coach training plan.

4.5. Resting Time Distribution

Through this component of the questionnaire, we observed if the resting time was correctly distributed across the training period, according to the coaches. We observed that only the 14 and 24 week training plans received statistically significant results. The 6 week training plans had no significant difference according to the χ^2-test ($\chi^2 = 7.733$). Additionally, this time period is the only one where the virtual coach was worse than the control training plan, as shown in Table 9. In this 6 week time period, the virtual coach was either putting too much resting days in the same week or not enough, which means that the distribution of the between sessions' resting days was not balanced to the training sessions' intensity.

Table 9. The contingency table of the resting time distribution in the 6 week plans.

$\chi^2 = 7.733$	Too Much	Neutral	Not Enough	Total
VC	8	0	8	16
Control	4	6	6	16
Total	12	6	14	32

We observed that the training plans were also getting equally evaluated by the experts on the 24 week training period as shown in Table 10. However, the main discrepancy of Table 10 was observed in the 14 week' training plans, where the control planning was evaluated as providing too much resting time in a consecutive manner (two more coaches gave a four than for the virtual coach).

Table 10. The contingency tables of the 14 week and 24 week resting time distribution evaluation.

	14 Weeks					24 Weeks			
$\chi^2 = 4.751$	Too Much	Neutral	Not Enough	Total	$\chi^2 = 2.550$	Too Much	Neutral	Not Enough	Total
VC	8	6	3	17	VC	8	9	2	19
Control	16	3	2	21	Control	9	12	0	21
Total	24	9	5	38	Total	17	21	2	40

Results in the 14 week and 24 week training evaluations were statistically independent according to the χ^2-tests (14 weeks: $\chi^2 = 4.751$ and 24 weeks: $\chi^2 = 2.550$). The virtual

coach was always evaluated as better in the distribution of the resting days in longer periods of time. Thus, we observed that the virtual coach needs more constraints on the short period of time and still has to be more balanced between training and resting in periods exceeding 14 weeks.

4.6. Efficiency of the Training Plans

Results from the evaluation of the planning efficiency were statistically dependent on the 6 week and the 14 week training periods, while the 24 week results were statistically independent. In the two former time periods, the virtual coach was never evaluated as efficient and was considered more neutral, while the coaches mostly agreed on the control training plan, as demonstrated in Table 11.

Table 11. The contingency tables of the 6 and 14 week efficiency evaluation.

	6 Weeks					**14 Weeks**			
$\chi^2 = 14.437$	**Good**	**Neutral**	**Bad**	**Total**	$\chi^2 = 14.777$	**Good**	**Neutral**	**Bad**	**Total**
VC	0	9	5	14	VC	0	9	5	14
Control	13	6	2	21	Control	12	3	4	19
Total	13	15	7	35	Total	12	12	9	33

As for the other evaluations, the virtual coach was better noted for the longer periods of time but received its worst results for 6 weeks. Table 12 shows the 24 week training plan scores, where we observe that the virtual coach received even less disagreement on the efficiency than the control training plan.

Table 12. The contingency table of the 24 week training plans efficiency evaluation.

$\chi^2 = 2.142$	**Good**	**Neutral**	**Bad**	**Total**
VC	4	9	3	16
Control	8	6	4	18
Total	12	15	7	34

Efficiency seems to be one of the areas where our virtual coach was not performing better than the control training plans. We further looked at the qualitative evaluation of the whole training plans from the coaches and made the link to this lacking component. However, coaches have not found the training plans out of place but had a more neutral evaluation of them, indicating that the basis is there and only needs some changes to get full agreement.

5. Discussion

Our virtual coach demonstrated good results on the longer time periods. We observed that the results are better managed by our virtual coach generated training plan, compared to the training plans provided by real human coaches. Experts composed of professional cycling experts, evaluated our virtual coach planning to be just as appropriate as the human coaches for the 14 and 24 week training period. We observed better results on the components relating to resting time and resting distribution. These results help us identify paths of amelioration where adjustments to the virtual coach optimization conditions could lead to better management of the athletes' training plan.

Based on these expert evaluations, we identified flaws in the design of the generated training plans. Indeed, most of the coaches mentioned the control plan as the best suited for the provided training period. The reason for this is the fact that looking at the TSB curves, we observe that the control is less peaking in each extreme and TSB values are smaller too. Thus, the virtual coach appeared more chaotic in the management of TSB compared to the

smoother curves of the control training plan. Additionally, the peaks of TSS sessions were higher in the virtual coach than in the control training plans, and the experts interpreted them as very high intensity training plans in the case of the virtual coach. The results for the appreciation of the training plan are something that is mainly induced by the lack of information provided to the experts. Providing them with a capabilities summary of the user we built the training plan for might have changed the results. We still think that the training plans are excessive in the performance requirement and we leverage on these first comments to adapt the conditions we apply to the training.

Through the results, we observed that the main issues faced by the virtual coach training plans happened in the 6 week training duration where the system was proposing high intensity trainings and provided unbalanced and intensive training plans for a short period. In fact, in such a reduced time frame, one would not want to see such high intensity training sessions and may need a more relaxed training plan. We need to reduce the TSS threshold for the sessions selected according to the length of the training requested. Additionally, the time of the season plays a role in training, since one would not start their training season with such high-intensity trainings. Adapting the range of TSS sessions proportionally to the current season's trainings might be a better strategy than providing the athletes with a training plan based on their whole sports life. The distribution of the sessions received bad evaluations for the virtual coach because it was either proposing too much or not enough of the appropriate training sessions. This is clearly seen in some training plans generated where the virtual coach is filling the planning with resting days in order to reduce the TSB values and maintain the athlete in the proper zone. As previously mentioned, the intensity regulation should also have the proper effect on this component. Proposing lower intensity trainings will reduce the necessity of recovery and resting days and will thus reduce the probability of a chain of these sessions.

In the longer time periods, we observe that our virtual coach system can perform as well as the human coaches' plans. In terms of resting quantity and resting time distribution, the virtual coach was outperforming the control training plans. In the intensity of the training, and despite the peaks of high TSS values in some sessions, the experts revealed that the virtual coach provided the athletes with a coherent training plan. In some cases, the training plans were in both judged as excessive, such as in terms of resting times distribution (see Table 10).

There are indeed points where the virtual coach's plans were not the better than the control ones and other components where this tendency was inverted. We believe that working on a shorter time period and using less data for the initial training plan could result in a more balanced training plan for each time period. We also need to put more focus in the management of the athlete's TSB evolution in order to smoothen the progression in the proposed trainings. We, however, keep in mind that there were still some components where the experts were less likely to notice a difference between both presented training plans. Thus, there is positive as well as negative management of components in both approaches. One should consider the proposed planning from the virtual coach and learn from it and use it, while the virtual coach should be inspired by the way a human coach interprets athletes.

6. Limitations

From a technical point of view, the developed virtual coach is currently basing the performance evaluation on the whole dataset of provided trainings. We observed that this could mislead the creation of our training plans, as one's performances may not be the same over the span of years. Thus, the over-evaluation of an athlete's capabilities could also lead them towards a negative effect of training as they might fall into overtraining after the first week of the proposed planning. However, such limitations would require more testing in real time, with participants trying to follow the planning and using the virtual coach. We think that in such a situation, the model will get adjusted as well as the planning and this would mitigate the overtraining risks. The virtual coach only takes the

athlete into consideration and not his ambient. The latter is also causing issues as a sports year is dependent on components like seasonality.

Taking the experimental point of view, we observed that the information shown in the generated training plan lacked context. Experts were not able to understand and were sometimes frustrated by the fact that there were peaks in the proposed sessions' TSS. Providing them with an athlete profile would certainly have changed their point of view on the proposed planning. Our case study does not experiment the usage of such technology-based coaching with a real-life athlete, but this will be explored in future work.

7. Conclusions

The field of sports science is being highly impacted by the availability of wearable data. Such devices enable anyone to get information on their performance, mainly in endurance sports. We observed the growing trend to apply computational intelligence as well as artificial intelligence-based systems to understand and enhance athletes in many kinds of sports. The provided tools are not always as simple as smart bracelets or smart watches; rather, they involve a certain quantity of installations such as tracking cameras. Endurance sports, as mentioned earlier, do have a better solution as wearables and specific measurement tools are vastly developed.

Using such information, one intends to understand her training or get some feedback on her performance. It may foster a positive effect on motivation to practice sports that may also tend to go through overtraining. Thus, researchers are looking for virtual coaching solutions to accompany the new and current athletes in their training, by providing algorithm-based training advices. Such tools have a limit, as mentioned, and are mainly based on heuristic-based rules to determine training planning.

Through our empirical study, we propose a novel approach on training planning, by tailoring the training sessions proposed by the past performance of the athlete to their current performance and evolution through time. We proposed our reinforcement learning-based approach in past research and test it in this case study with cycling coaches as experts. The quantitative results demonstrated that our virtual coach has been equally, or even better, evaluated compared to training plans from real coaches. Qualitative data and the former quantitative results also pointed out lack of context of the developed virtual coach. Thus, we will further go into the development and research of our virtual coach by adding more contextualization in the training plans creation process.

References

1. Neely, A.D. (Ed.) *Business Performance Measurement: Theory and Practice*; Cambridge University Press: Cambridge, UK, 2002; ISBN 978-0-521-80342-7.
2. Seshadri, D.R.; Drummond, C.; Craker, J.; Rowbottom, J.R.; Voos, J.E. Wearable Devices for Sports: New Integrated Technologies Allow Coaches, Physicians, and Trainers to Better Understand the Physical Demands of Athletes in Real Time. *IEEE Pulse* **2017**, *8*, 38–43. [CrossRef] [PubMed]
3. Aroganam, G.; Manivannan, N.; Harrison, D. Review on Wearable Technology Sensors Used in Consumer Sport Applications. *Sensors* **2019**, *19*, 1983. [CrossRef] [PubMed]
4. Willemsen, M.C.; Graus, M.P.; Knijnenburg, B.P. Understanding the Role of Latent Feature Diversification on Choice Difficulty and Satisfaction. *User Model User Adap. Inter.* **2016**, *26*, 347–389. [CrossRef]
5. Hassan, L.; Dias, A.; Hamari, J. How Motivational Feedback Increases User's Benefits and Continued Use: A Study on Gamification, Quantified-Self and Social Networking. *Int. J. Inf. Manag.* **2019**, *46*, 151–162. [CrossRef]
6. Montesano, P.; Di Silvestro, M.; Cipriani, G.; Mazzeo, F. Overtraining Syndrome, Stress and Nutrition in Football Amateur Athletes. *J. Hum. Sport Exerc.* **2019**, *14*, S957–S969. [CrossRef]
7. Susta, D. The Overtraining Syndrome. In *Triathlon Medicine*; Migliorini, S., Ed.; Springer International Publishing: Cham, Switzerland, 2020; pp. 345–354, ISBN 978-3-030-22357-1.
8. Rabbi, M.; Pfammatter, A.; Zhang, M.; Spring, B.; Choudhury, T. Automated Personalized Feedback for Physical Activity and Dietary Behavior Change with Mobile Phones: A Randomized Controlled Trial on Adults. *JMIR mHealth uHealth* **2015**, *3*, e42. [CrossRef]
9. Lv, L.; Ye, C. Comprehensive Analysis Method Model of Aerobics Performance Based on Artificial Intelligence. In Proceedings of the 2016 Eighth International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Macau, China, 11–12 March 2016; pp. 151–154.
10. Vales-Alonso, J.; López-Matencio, P.; Gonzalez-Castaño, F.J.; Navarro-Hellín, H.; Baños-Guirao, P.J.; Pérez-Martínez, F.J.; Martínez-Álvarez, R.P.; González-Jiménez, D.; Gil-Castiñeira, F.; Duro-Fernández, R. Ambient Intelligence Systems for Personalized Sport Training. *Sensors* **2010**, *10*, 2359–2385. [CrossRef]
11. Mata, F.; Torres-Ruiz, M.; Zagal, R.; Guzman, G.; Moreno-Ibarra, M.; Quintero, R. A Cross-Domain Framework for Designing Healthcare Mobile Applications Mining Social Networks to Generate Recommendations of Training and Nutrition Planning. *Telemat. Inform.* **2018**, *35*, 837–853. [CrossRef]
12. Mulas, F.; Carta, S.; Pilloni, P.; Manca, M. Everywhere Run: A Virtual Personal Trainer for Supporting People in Their Running Activity. In Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, Lisbon, Portugal, 8–11 November 2011; pp. 70:1–70:2.
13. Rønnestad, B.R.; Mujika, I. Optimizing Strength Training for Running and Cycling Endurance Performance: A Review. *Scand. J. Med. Sci. Sports* **2014**, *24*, 603–612. [CrossRef]
14. Ofoghi, B.; Zeleznikow, J.; MacMahon, C.; Dwyer, D. A Machine Learning Approach to Predicting Winning Patterns in Track Cycling Omnium. In *Artificial Intelligence in Theory and Practice III*; Bramer, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 67–76.
15. Ofoghi, B.; Zeleznikow, J.; Dwyer, D.; Macmahon, C. Modelling and Analysing Track Cycling Omnium Performances Using Statistical and Machine Learning Techniques. *J. Sports Sci.* **2013**, *31*, 954–962. [CrossRef]
16. Sörös, G.; Daiber, F.; Weller, T. Cyclo: A Personal Bike Coach Through the Glass. In Proceedings of the SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications, Hong Kong, 19–22 November 2013; pp. 99:1–99:4.
17. De Brouwer, M.; Ongenae, F.; Daneels, G.; Municio, E.; Famaey, J.; Latré, S.; De Turck, F. Personalized Real-Time Monitoring of Amateur Cyclists on Low-End Devices: Proof-of-Concept & Performance Evaluation. In Companion Proceedings of the The Web Conference 2018, Lyon, France, 23 April 2018; pp. 1833–1840.
18. Walmink, W.; Wilde, D.; Mueller, F. "Floyd" Displaying Heart Rate Data on a Bicycle Helmet to Support Social Exertion Experiences. In Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction, Munich, Germany, 16 February 2014; pp. 97–104.
19. Rodríguez-Marroyo, J.A.; Villa, G.; García-López, J.; Foster, C. Comparison of Heart Rate and Session Rating of Perceived Exertion Methods of Defining Exercise Load in Cyclists. *J. Strength Cond. Res.* **2012**, *26*, 2249–2257. [CrossRef] [PubMed]
20. Rauter, S. New Approach for Planning the Mountain Bike Training with Virtual Coach. *TRENDS Sport Sci.* **2018**, 69–74. [CrossRef]
21. Camomilla, V.; Bergamini, E.; Fantozzi, S.; Vannozzi, G. Trends Supporting the In-Field Use of Wearable Inertial Sensors for Sport Performance Evaluation: A Systematic Review. *Sensors* **2018**, *18*, 873. [CrossRef] [PubMed]
22. Allen, H.; Coggan, A. *Training and Racing with a Power Meter*, 2nd ed.; VeloPress: Boulder, CO, USA, 2010; ISBN 978-1-934030-55-4.
23. McCabe, A.; Trevathan, J. Artificial Intelligence in Sports Prediction. In Proceedings of the Fifth International Conference on Information Technology: New Generations (ITNG 2008), Las Vegas, NV, USA, 7–9 April 2008; pp. 1194–1197.
24. Novatchkov, H.; Baca, A. Artificial Intelligence in Sports on the Example of Weight Training. *J. Sports Sci. Med.* **2013**, *12*, 27–37. [PubMed]
25. Novatchkov, H.; Baca, A. Fuzzy Logic in Sports: A Review and an Illustrative Case Study in the Field of Strength Training. *Int. J. Comput. Appl.* **2013**, *71*, 8–14. [CrossRef]

26. Hucaljuk, J.; Rakipović, A. Predicting Football Scores Using Machine Learning Techniques. In Proceedings of the 34th International Convention MIPRO, Opatija, Croatia, 23–27 May 2011; pp. 1623–1627.
27. Kelly, P.; Healy, A.; Moran, K.; O'Connor, N.E. A Virtual Coaching Environment for Improving Golf Swing Technique. In Proceedings of the 2010 ACM Workshop on Surreal Media and Virtual Cloning, Firenze, Italy, 25–29 October 2010; pp. 51–56.
28. Kovalchik, S.; Reid, M. A Shot Taxonomy in the Era of Tracking Data in Professional Tennis. *J. Sports Sci.* **2018**, *36*, 2096–2104. [CrossRef] [PubMed]
29. Bartlett, R. Artificial Intelligence in Sports Biomechanics: New Dawn or False Hope? *J. Sports Sci. Med.* **2006**, *5*, 474–479.
30. Silva, A.J.; Costa, A.M.; Oliveira, P.M.; Reis, V.M.; Saavedra, J.; Perl, J.; Rouboa, A.; Marinho, D.A. The Use of Neural Network Technology to Model Swimming Performance. *J. Sports Sci. Med.* **2007**, *6*, 117–125.
31. Wong, P.K.C. Developing an Intelligent Assistant for Table Tennis Umpires. In Proceedings of the First Asia International Conference on Modelling Simulation (AMS'07), Phuket, Thailand, 27–30 March 2007; pp. 340–345.
32. French, B.; Tyamagundlu, D.; Siewiorek, D.P.; Smailagic, A.; Ding, D. Towards a Virtual Coach for Manual Wheelchair Users. In Proceedings of the 2008 12th IEEE International Symposium on Wearable Computers, Pittsburgh, PA, USA, 28 September–1 October 2008; pp. 77–80.
33. Geissler, H.; Hasenbein, M.; Kanatouri, S.; Wegener, R. E-Coaching: Conceptual and Empirical Findings of a Virtual Coaching Programme. *Int. J. Evid. Based Coach. Mentor.* **2014**, *12*, 165.
34. Buttussi, F.; Chittaro, L.; Nadalutti, D. Bringing Mobile Guides and Fitness Activities Together: A Solution Based on an Embodied Virtual Trainer. In Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services, Helsinki, Finland, 12 September 2006; pp. 29–36.
35. Sañudo, B.; Abdi, E.; Bernardo-Filho, M.; Taiar, R. Aerobic Exercise with Superimposed Virtual Reality Improves Cognitive Flexibility and Selective Attention in Young Males. *Appl. Sci.* **2020**, *10*, 8029. [CrossRef]
36. Silacci, A.; Khaled, O.A.; Mugellini, E.; Caon, M. Designing an e-Coach to Tailor Training Plans for Road Cyclists. In *Human Systems Engineering and Design II*; Ahram, T., Karwowski, W., Pickl, S., Taiar, R., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 1026, pp. 671–677, ISBN 978-3-030-27927-1.
37. Mulas, F.; Pilloni, P.; Carta, S. *Everywhere Race!: A Social Mobile Platform for Sport Engagement and Motivation*; ThinkMind: Venice, Italy, 2012; pp. 63–69.
38. Kang, S.; Baek, H.; Jung, E.; Hwang, H.; Yoo, S. Survey on the Demand for Adoption of Internet of Things (IoT)-Based Services in Hospitals: Investigation of Nurses' Perception in a Tertiary University Hospital. *Appl. Nurs. Res.* **2019**, *47*, 18–23. [CrossRef] [PubMed]
39. Ruttkay, Z.; van Welbergen, H. Elbows Higher! Performing, Observing and Correcting Exercises by a Virtual Trainer. In *Intelligent Virtual Agents*; Prendinger, H., Lester, J., Ishizuka, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 409–416.