Training Quantification in Endurance Sports: Training Load and the Acute Performance Decrement.

This thesis is presented for the Degree of Doctor of Philosophy at the

University of Kent

March 2022

Antonios Panagiotis Kesisoglou

School of Sport and Exercise Sciences

DECLARATION

I confirm that this thesis is my own work and has not been previously submitted for an awarded to this University or any other institution. I certify that to the best of my knowledge, this thesis is not based on anyone's work. Any material from the work of other included in this thesis are fully acknowledged in accordance to common reference practices.

ACKNOWLEDGEMENTS

This thesis is in memory of my principal supervisor Prof. Louis Passfield, who left us very early. Louis, I am grateful for your support, guidance and numerous hours you have spent over the last 6 years to teach me. Your teaching has had an impact both on a personal and a professional level. You were one of the first who believed in me and you provided everything you could for making this thesis happen. Your support in both academic but also lifestyle held me when tough moments came. Both you and I know that during those years a lot of bad moments came, yet you were always there for me and for that, I will be forever grateful. It was many times that you stand by my side as a parental figure and not only as a supervisor. Secondly, I would like to thank Dr. Andrea Nicolo. Andrea, thank you for your patience during all those years working with me. I always felt I was below your expectations in terms of academic commitment. I felt that although a humble coach, I managed to keep up with your standards. I am grateful for all your personal time you gave me all these years. You were there to support me in tough moments and to keep me sane when I needed it. Although we had to work together remotely, I never felt that you were absent, rather you always found time to support me. Thank you for everything.

I would also like to express my gratitude towards the constant support I perceived from my parents. Dad, you must be proud up there looking at everything I have achieved. I kept pushing through all the difficulties because I knew how proud you would be. Mom, thank you for being so unselfish and despite the fact that you were lonely, you never asked me to come back and be around the family who needed me a lot. My girlfriend, Liza Panti, who never thought about herself more than myself and moved in the UK because she wanted to support me at the last, yet very important steps of my PhD. Lastly, I would like to say thank you to some people in my working environment. In the University of Kent colleagues who found some brothers I never had. Sam Smith, Arthur Bossi, Borja Martinez, thank you all the help provided when I struggled. My colleagues at the Canterbury Academy who showed a great understanding about the workload I had to manage. Mark Dayson, you

were a great tutor and you helped me progressed in the profession. You taught me a fine balance between science and coaching. Jesse Sazant, thank you for seeing things on me that I couldn't back in the days. Adam Davies and Billy Beddow, thank you for being such awesome colleagues.

ABSTRACT

The aim of the thesis was to obtain a better understanding into current measurements of Training Load (TL) in endurance performance and whether these TL metrics estimate the training stress imposed by endurance exercise.

In the first study of the thesis (Chapter 3), the concept of Acute Performance Decrement (APD) was introduced. 14 recreational and competitive cyclists were asked to perform a 5 min TT as a baseline, and after various cycling training sessions. The cycling training sessions consisted of a 5 min TT, a 20 min TT, a 20 min and a 40 min sub-maximal cycling session. The resulting APD was calculated as the percentage change in 5 min TT from baseline and was compared with the TL metrics from the corresponding training sessions. The results shown that an APD was found after all training sessions (η_p^2=0.971; P<0.001). A similar APD was observed after the 5 and a 20 min sessions that contradicted the TL metrics. These findings suggest that the basis for TL metrics may be flawed. The TL metrics based on the concept of TWD did not agree with how APD responded after the various cycling training sessions.

In the second study (Chapter 4), the concept of APD was investigated in running. The aim of the study was to investigate whether TL responses were in an agreement with APD. 11 well-trained runners were asked to perform a 1500m TT as a baseline, and after various running sessions. Continuous (CON) and intermittent (INT) running bouts of different duration were examined. The running sessions consisted of either maximal effort for 10 min (10CON, 10INT) or sub-maximal for 25 min (25CON, 25INT). The results showed that an APD was found after all training sessions. TL metrics showed the opposite response compared to APD. APD was found to be similar when compared after the training sessions, with differences found only after 10INT vs 25CON (P=0.02). In contrast, TL metrics provided the opposite pattern with higher scores observed for CON vs INT and lower scores for 10 min vs 25 min training sessions (P<0.001; η_p^2 =0.563).

In the third study (Chapter 5), APD and TL metrics were examined after running at two different intensities and two durations. 12 trained runners were asked to perform a 1500m TT as a baseline and after 4 training sessions. The training sessions consisted for two different intensities (RPE of 5 and 8) and durations (10 min and 40 min). APD was calculated as the percentage change in 1500m TT. The results showed that the training sessions that lasted for 40 min resulted in greater TL metric scores compared to those that lasted for 10 min (P<0.001; η_p^2 >0.781). In contrast, APD was found to be similar after the 10 min (RPE 8) and 40 min (RPE 5) training sessions (P > 0.99).

The findings from this thesis demonstrated that training stress as indicated by APD occurs in response to a combination of exercise intensity and duration that is not reflected by most TL metrics. Across all studies, the magnitude of APD was found to depend mainly on how intense the exercise dose was, however, extended exercise durations were found to increase the training stress. In all experimental studies, an agreement between TL metrics and APD was not found. These findings may provide concern for the validity of the TL metrics but may pave the road for the development of new TL metrics. It is suggested that new TL metrics are required that need to account for the effects of intensity and duration and their interaction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABREVIATIONS	UNITS or TERMS
AU	Arbitrary Units
APD	Acute performance decrement
TRIMPS	Training impulses
bTRIMP	Banister's TRIMP
eTRIMP	Edward's TRIMP
EE	Energy expenditure
LuTRIMP	Lucia's TRIMP
iTRIMP	Individualized TRIMP
TSS	Training stress score
sRPE	Session rating of perceived exertion
rTSS	Running training stress score
FTP	Functional threshold power
NP	Normalized power
HR	Heart rate
HR_{max}	Maximal heart rate
HR_{mean}	Average heart rate
HR_{rest}	Resting Heart Rate
HRV	Heart rate variability
La^-	Blood lactate
La^-_{end}	Post exercise blood lactate
PO	Power Output
PPO	Peak power output
IF	Intensity factor
VO_2	Oxygen consumption
VO_{2max}	Maximal oxygen uptake
VT	Ventilatory thresholds
VT1	First ventilatory threshold
VT2	Second ventilatory threshold
POMS	Profile of mood state
R^2	Coefficient of determination
RPE	Rating of perceived exertion
RPE_{mean}	Average rating of perceived exertion
RPE_{peak}	Peak rating of perceived exertion
CR-10 RPE	0-10 Rating of perceived exertion

PTV	Peak treadmill velocity
SD	Standard deviation
SEM	Standard error
W	Watt
TT	Time trial
TTE	Time to exhaustion
η_p^2	Partial eta squared
CON	Continuous
INT	Intermittent
NASA-TLX	The National Aeronautics and Space Administration Task Load Index
DALDA	Daily Analysis of Like Demands for Athletes
TRQ	Total Recovery Scale
Rest-Q-Sport	Recovery-Stress Questionnaire for athletes
ROF	Rate of Fatigue scale
RCP	Respiratory compensation point
TWD	Total work done
fR	Respiratory frequency

LIST OF PUBLICATIONS

1. Kesisoglou A, Nicolò A, Passfield L (2021). Cycling Performance and Training Load: Effects of Intensity and Duration. Int J Sports Physiol Perform 16 4: 535-543
2. Kesisoglou A, Nicolò A, Howland L, Passfield L (2021). Continuous Versus Intermittent Running:Acute Performance Decrement and Training Load. Int J Sports Physiol Perform 16 12: 1794-1803

Chapter 1: Introduction

When training is applied, the aim is to increase the physical capabilities and to develop a skill, a technique, or performance. This process involves the application of training and recovery with the aim to ultimately improve performance (Bourdon et al. 2017). That concept can be examined both from an acute, and chronic point of view. One of the first attempts to conceptualize training overload and recovery from an acute point of view derives from the work of Yakovlev (1955), which was recently expanded by Issurin (2010). A single training session (i.e acute overload) will result in training-induced fatigue and an acute decrease in the training capacity of an athlete (Issurin 2010). The training stress will consequently trigger adaptations and this is known as the supercompensation cycle (Figure 1.1). Consequently, a stimulus for adaptation occurs if recovery is applied. Endurance athletes train a lot, and applying increasingly high levels of training more frequently is a normal and important part of their training process (Smith 2003). However, there is no clear evidence on the exact amount of training an endurance athlete needs to complete in order to reduce capacity in a purposely way and to ensure performance improvements (Soligard et al. 2016). Although training theories agree that a negative effect (i.e fatigue) will take place (Bompa and Haff 2009; Issurin 2010), no experimental attempts have been made to systematically quantify the reduction in exercise capacity resulted from different acute training sessions. Therefore, Figure 1.1 can only be considered as a theoretical framework and lacks experimental evidence. Specifically, there is a lack of experimental data to demonstrate how exercise capacity changes over time when single or multiple bouts are imposed. If an athlete is not monitored closely, too high levels of overload may result in short-term fatigue which can turn into long-term overtraining (Kuipers and Keizer 1988). Therefore, monitoring this overload in the form of its training load (TL) in a valid, smart, efficient and reliable manner is critical in order to understand how much training has been applied and when its effects triggered.

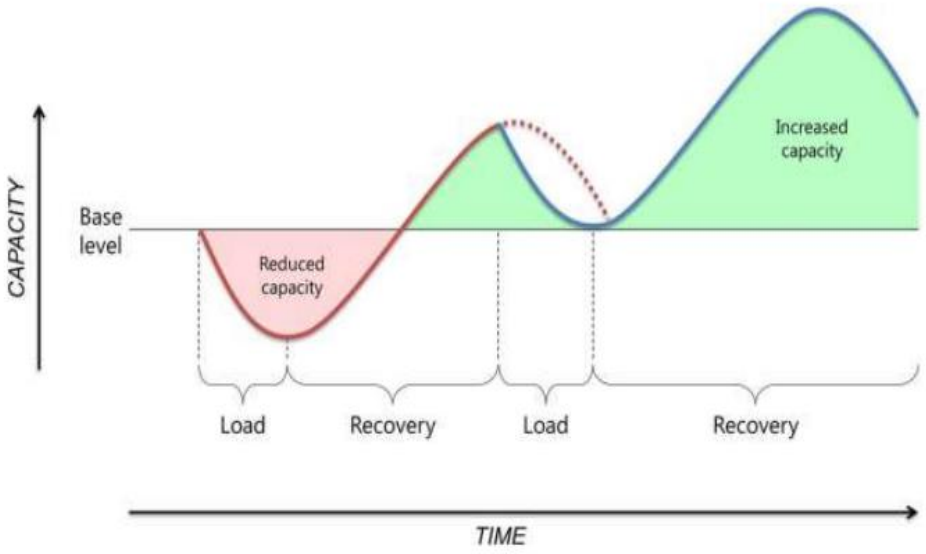

Figure 1-1 Figure presents the relationship between training applied, recovery and improvements in performance(from Meeusen, 2013 and Soligard et al, 2016).

The effects of training are dictated by its exercise intensity, duration and frequency (Borresen and Lambert 2009). To enhance performance, progressive increases of intensity, duration and frequency are necessary. Monitoring training is about the careful evaluation of the parameters mentioned above and their effects. This enables coaches, athletes and scientists to track how much training stress an athlete is experiencing and possibly how much benefit they are obtaining. Based on a recent consensus statement (Bourdon et al. 2017), TL monitoring (Bourdon et al. 2017) is important for several reasons including the understanding of the training responses and their consequent adaptations, the assessment of fatigue, the identification of recovery required after applied overload and understanding or predicting injuries and illnesses. However, different combinations of intensity, duration and frequency may result in similar training effects (Borresen and Lambert 2009). Likewise, similar combinations of these parameters may result in dissimilar training effects because the training response is also affected by other factors. For example, the recovery required following single or multiple periods of training overload play a critical role on the training adaptations and in turn,

performance improvements (Bourdon et al. 2017). Despite the attempts to quantify the training effects via TL monitoring, the relationship between training, its effects and performance is complex.

The use of metrics to evaluate athletes' TL has increased exponentially in recent years but despite this popularity, some concerns have been expressed about its validity, and exploring these concerns forms part of the basis of this thesis. These concerns are various and wide ranging and include the terminology used around the topic (Staunton et al. 2021), the practicality and applicability of the fitness and fatigue model (Taha and Thomas 2003; Hellard et al. 2006) and the use of total work done to compare different sessions (TWD) (Nicolò et al. 2016b; Passfield et al. 2017). Most TL metrics use indices of training duration and intensity by multiplying them in order to estimate the training stress imposed by exercise. By doing that, the effects of training are commonly represented by a single number. This attempts to reduce the complexity of the process and it may result in misleading information about training. The way TL metrics have conceptualized training monitoring may require a new direction. Researchers have attempted to relate TL with performance outcomes in longitudinal studies. However, a missing piece of information is whether the TL metrics currently used can effectively inform the effects from a training session. As Jeffries et al. (2021) suggested, the load imposed by exercise determines acute and chronic training effects that result in sport performance outcomes. When the modulation of load and recovery is not optimal, positive training effects leads to improvements in performance. Conversely, when load or recovery is not well administered, negative training effects might occur and lead to negative changes in performance outcomes. A conceptual issue in what researchers have attempted to investigate so far is the direct link between TL and sport performance outcomes without assessing the link between TL and its acute effects. Prior to longitudinal investigations, the concept of TL needs to be examined in an acute training session context against a sensitive method that does not use indices of intensity and duration to estimate the effects resulting from training. This thesis intents to examine this experimentally and hopefully provide useful insights on TL monitoring.

Chapter 2: Literature Review

2.1 The History of TL

The concept of TL was originally introduced by Banister et al. (1975) although they did not use this term. Their idea was that training could be quantified with training impulses (TRIMPs) which in turn could be used to predict changes in athletic performance (Banister et al. 1975). According to this model which will be further discussed in Chapter 2.3, changes in performance result from the difference between a negative function of training (fatigue) and a positive function of training (fitness) in the short term and by their accumulation in the longer term (Banister et al. 1975). Initially, the term TL was used in a different contexts (e.g. Daniels et al. 1978) and was later linked to the initial work from Bannister and colleagues (Foster et al. 1995). In their first paper, Banister and colleagues modelled the performance of a swimmer resulting from a 105-days training program and the fitness and fatigue it induced. To estimate fitness and fatigue, the authors calculated the participant's swim and weight training and translated these into TRIMPs. The swim distance was recorded in hundreds of meters (i.e. 14,000 meters = 140 units) and then multiplied by an intensity factor for the session to express the session in TRIMPS. The warm up and warm down activities were assigned an intensity factor of 1; the long duration and hard intensities were assigned an intensity factor of 2; the quality training and speed swimming were assigned an intensity factor of 3. The weight training was quantified in units of 500 repetitions. The authors suggested that a unit of weight training (500 repetitions) was equal to 1000 meters of swimming at an intensity 3. Importantly, the findings from this study were only presented next to performance predictions without providing a statistical evaluation. In subsequent studies, the same performance modelling approach was implemented by the same research group (Calvert et al. 1976; Banister and Calvert 1980). Some years later, another scientific paper was published and attempted to statistically evaluate the model and to assess the validity and in particular, the predicted effects on performance (Morton et al. 1990). To do

that, two athletes underwent a 28-day training programme and completed performance trials twice a week for 50 days after training cessation. The authors concluded that the model, as derived from TRIMPs, effectively replicated the participants' actual performances. Based on that work by Morton et al. (1990), Foster et al. (2001) introduced the sRPE method. The sRPE method like TRIMPs multiplies exercise intensity (RPE) and duration in order to create a TL score. Subsequently, several modifications of the way in which TRIMPs are calculated were proposed, with most of the emphasis given to how the exercise intensity should be weighted (Edwards 1993; Lucía et al. 2003; Manzi et al. 2009b). These TL metrics will be discussed in detail below in 2.4.

2.2 TL Terminology and External/Internal TL metrics

In line with the training principles, acute exercise results in training-induced fatigue. It is important to provide some definitions in relation to this relationship. The term fatigue has been widely used in the context of endurance performance (Enoka and Duchateau 2016). Based on a previously proposed taxonomy, fatigue can be differentiated as performance fatiguability and perceived fatiguability (Kluger et al. 2013; Enoka and Duchateau 2016). Performance fatiguability is defined as a drop in a performance assessment over a period of time. On the other hand, perceived fatiguability refers to variations in the sensations related to the decrease in performance. Nonetheless, fatigue can take place with the absence of exercise (Fieo et al. 2013). Due to the complicated nature of fatigue, Enoka and Duchateau (2016) suggested to avoid terms such as central fatigue, peripheral fatigue muscle fatigue, physical fatigue and supraspinal fatigue. Instead, the authors suggested that research should examine fatigue as a single element, with focus on the measurements, as well as the effects of protocols on findings referring to fatigue. In this thesis, the negative effects of training (i.e training-induced fatigue) on performance will be examined (Pageaux and Lepers 2016) and this will be attempted via post-exercise performance assessments (see Chapter 2.8). Nevertheless, fatigue is not the only negative effect resulted from training (Jeffries et al. 2021). Training can result in muscle

damage (Peake et al. 2017) and mood alterations (Morgan et al. 1987). Based on these suggestions (Enoka and Duchateau 2016) and the applied nature of this thesis, assessments of different fatigue mechanisms were avoided. It is also important to acknowledge that this method may not fully describe all elements of fatigue as proposed by Kluger et al. (2013).

Other popular terms in exercise monitoring are "work" and "load". The term "work" implies that mechanical work occurred. The term "work", or mechanical work done is calculated as the outcome between force and the distance in which that force was applied (Winter and Fowler 2009). Notably, evidence suggests that currently used TL metrics are related to the mechanical work done, which was also found to be closely associated with oxygen consumption (VO_2), and has been suggested as the criterion for TL monitoring (Wallace et al. 2014b). However, some researchers have suggested that the way exercise intensity and duration contribute to TWD is more important that the TWD level per se (Renfree et al. 2021). This is an important issue that will be discussed critically discussed further below (2.6.1). The term "load" and its use may have also contributed to the confusion in regards to the TL topic (Staunton et al. 2021). For example, the term "load" in the context of sports is a very broad concept and usually relates to the field of engineering and physics. In the field of engineering, "load" is related to the force exerted on a surface or a body (Rodgers and Cavanagh 1984). Similarly, the force applied on a mass which resist are also external forces that can also be considered as load. In physics, "load" is a term to explain the part of a circuit which consumes electrical power and is measured differently in the SI derived unit. These are two examples that show how the term "load" is used differently in other fields of science and have potentially mislead sports scientists and practitioners. The term load can be also found in non-scientific definitions. For example, Staunton et al. (2021) highlighted how the Oxford English Dictionary defines "load" as "something that is being carried by a person, or a vehicle for example", or as "a burden placed on an individual, structure, machine or system". This may create confusion, because if for example it is

used as a burden upon a person or a system, it may be considered as the amount of work that a person or a machine has to do (Staunton et al. 2021).

Another definition that has been frequently used in the context of exercise monitoring is the term "workload". This term has been traditionally used to describe severity (Knuttgen 1978). Workload consists of terms "work" and "load". The term "load" in the context of workload explains the resistance that an individual experiences when performing the work. The use of the term "workload" has received criticism due to the measurement units that have been used to describe it (Winter 2006). Workload may indicate the load of work that was performed and therefore the accompanied unit measurement should be joules. Alternatively, it may indicate the antagonistic force towards work, and in that case the unit of measurement should be newtons (Winter 2006). It was suggested that neither power output nor speed measurements should be used as a measurement unit for workload (Winter 2006; Winter and Fowler 2009). As said earlier, this term is attempting to describe the severity of the task. For that purpose, Winter (2006) suggested the use of the term intensity instead, as a more universal method that can apply to multiple forms of exercise (Knuttgen 1978; Winter 2006). Nevertheless, the use of that term has been lately used in the context of sports. This term has been proposed as a way to represent the cost for achieving a mission by a human operator (Hart 2012). In sports, it has been proposed to capture that negative effects imposed by mental and physical exercise (Pageaux et al. 2014, 2015; Rozand et al. 2014; Martin et al. 2016; Cutsem et al. 2017). Nonetheless, what the term "workload" exactly represents in the context of sports and how it should be measured is currently not clear. Different authors (Winter 2006; Winter et al. 2016; Staunton et al. 2021) proposed to avoid the use of that term in the context of sports.

The terms "work" ,"load", "workload" and their link with the concept of TL overall may cause confusion. The main problem is what each term represents and how these constructs are connected with TL monitoring. Some examples above used the term ''load'' for something that can be lifted or carried (i.e. mass or weight), while other examples used the term for something that can be quantified

by means of force. The ‘’load’’ in terms of work done seems to better link with the how current TL metrics utilize it in the context of exercise. The terminology used around the topic may also attribute towards the confusion about what the construct of TL represents (Staunton et al. 2021). Some research relates TL to the amount (volume) of exercise or training completed (Impellizzeri et al. 2005; Scott et al. 2013a, b; Halson 2014; Bourdon et al. 2017). Others described TL as a product of volume and intensity of exercise (Impellizzeri et al. 2004). Other authors defined TL as the amount of stress placed on an athlete as a consequence of exercise over a period of time (Jaspers et al. 2017; Eckard et al. 2018) and this is the definition that will be adopted by this thesis as defined by Bourdon et al. (2017). Overall, it seems that there is lack of consistency in terms of how TL is described in the sports science literature and that has led some scientists to suggest it is discarded (Staunton et al. 2021).

Despite the confusion in the terminology around TL, scientists have proposed the external and internal domains for TL quantification (Borresen and Lambert 2009; Bourdon et al. 2017; Mujika 2017) (see figure 2.1). There is no doubt that training monitoring should include mechanical and physiological/perceptual responses of an athlete when training. For example, it is important for a runner to record the distances covered during training sessions. The term external training load is used to define the mechanical work applied by an athlete when exercising (Bourdon et al. 2017). In the review by Bourdon et al. (2017), measurements such as time, distance/milage, movement repetitions, power output, speed and number or jumps were suggested. These measurements are important for understanding if an athlete improves. For example, a runner is aiming to be able to cover greater distances when training. Nevertheless, it is difficult to understand the effects that are caused by a training session by looking at only the external load. This is because different individuals may show different responses to the same external load (Smith 2003; Mann et al. 2014; Barroso et al. 2015; McLaren et al. 2018; Casado et al. 2019; Paquette et al. 2020). These differences are the core information when attempting to quantify training effects via TL monitoring. Collectively,

external TL may be useful to describe the amount of physical work or mechanical work an athlete undergoes (i.e. the amount of training), but not the effects triggered.

Internal TL refers to the physiological and psychological stress as a response to the external TL imposed to an athlete (Bourdon et al. 2017). Common measurements for monitoring internal training load involve the rating of perceived exertion (RPE), Heart Rate (HR) and Blood Lactate (La^-), TRIMPs, questionnaires and measurements of oxygen uptake (Bourdon et al. 2017). Internal TL shows the psychophysiological response to an external TL (i.e. its effects) (Impellizzeri et al. 2019) and seems to play a key role for training adaptations (Viru and Viru 2000). Impellizzeri et al. (2019) suggested the inclusion of both physiological (i.e. HR) and psychological (i.e. RPE) measurements of internal TL. Irrespectively of the method used to describe the internal TL, monitoring training is about understanding the effects of a session and how that will consequently affect performance. With that in mind, the internal TL should not be confused with the external TL. If monitoring the training effect is the goal, then the concept of TL should be linked with the training effects and performance. In other words, the quantification of training via TL metrics should reflect how an athlete internally responds (i.e effects) to a training session (i.e external TL). Nevertheless, whether TL metrics represent the effects of training which in turn will result in changes in performance is unknown. This relationship remains to be established experimentally and will be attempted in later Chapters of this thesis.

2.3 Fitness, Fatigue and Dose-Response Relationship

Based on the model proposed by Bannister et al. (1975), training creates fatigue and fitness responses (figure 2.1). Both responses are then modelled to calculate the cumulative effects of fatigue and fitness on performance. The way athletes respond in terms of their fitness and fatigue decay constants was suggested to be an individualized process (Morton et al. 1990). The fitness responses

represent the physiological adaptations to training which act to improve performance capacity. The fatigue responses represent the deleterious effects of training that result in negative effects on performance (Banister et al. 1975). In this context, the authors suggested that performance at any given point can be explained as the positive training responses minus the negative training responses (i.e. Performance = Fitness – Fatigue). Therefore, the TL imposed on an athlete can result in both negative and positive effects and it is the balance of these will dictate the direction in which performance changes. There are a fair amount of studies that have assessed how periods of applying TL can affect changes in fitness and performance (Manzi et al. 2009b; Akubat et al. 2012; Wallace et al. 2014a; Sanders et al. 2017a). In order to predict the training outcome in an accurate way, an accurate monitoring of TL prescribed and its effects on fitness, fatigue and performance is required. The dose-response relationship between the prescribed training (input) and the training outcome (output) depends upon a reliable and valid method being used by the athlete, coach and scientist. The assumptions underpinning this relationship will be discussed below.

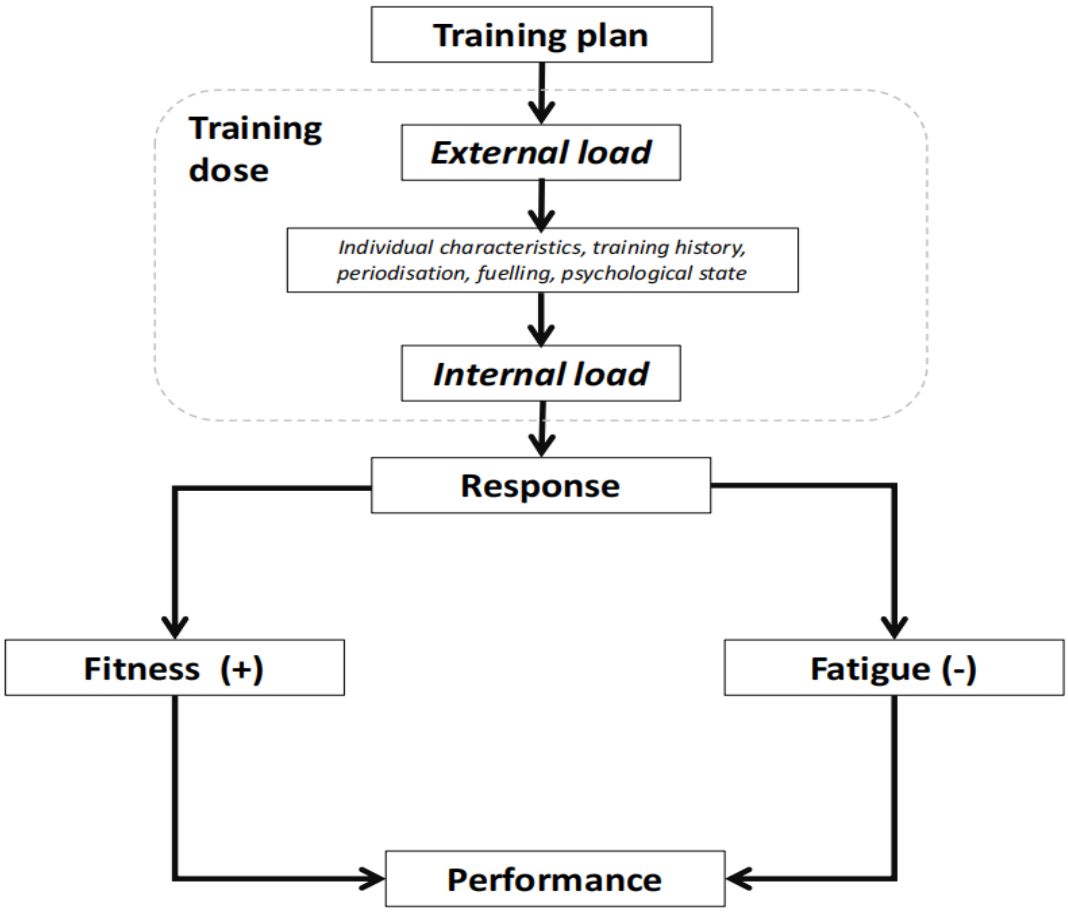

Figure 2-1 The figure above presents the process from the application of the training dose, how this can be quantified in terms of external and internal TL and its effects on performance via changes in fitness and fatigue (Adopted from Impellizzeri et al. 2005; 2019, Coutts, 2016)

Any TL that is implemented for training purposes should relate to the training outcome in a way that reflects the changes in performance capacity. Therefore, a valid way to quantify TL is required to accurately inform the dose-response relationship between a training stimulus and the performance outcome. Thus, it is important to investigate whether a dose-response relationship in endurance training exists. Typically, two different approaches have been used in the literature for validating the dose-response relationship between TL and performance changes in endurance events. The first one uses mathematical modeling in order to assess the modelled performance outcomes compared to the measurements of real performances (Jobson et al. 2009). The second one assesses changes in

performance or fitness after a period where TL was induced (Manzi et al. 2009a; Wallace et al. 2014a). Mathematical modelling was also proposed for predicting the performance following a given TL, as well as conversely to propose a TL required for a given athletic performance (Busso and Thomas 2006). Evidence suggests that the relationship between training and performance can be modelled (Banister et al. 1975; Morton et al. 1990; Mujika, Iñigo et al. 1996; Avalos et al. 2003). Systems modelling attempts describe the performance changes that occur in response to training, by describing changes in performance (output) in relation to the amount of training (input). Several research groups provided insights on how fitness and fatigue correspond to a given TL and investigated the dose-response relationship between TL doses and changes in performance (Banister and Calvert 1980; Busso et al. 1991; Mujika, Iñigo et al. 1996; Hellard et al. 2006; Wallace et al. 2014a). However, if useful insights on how the dose of training affects the performance response are to be obtained, the validity of such approaches must be assured.

In Banister's model, different parameters are used to predict performance. Pt is the performance at time t. P0 is the initial performance capacity. K_a presents the fitness magnitude factor, while K_f is the fitness magnitude factor. The fitness and fatigue decay constants are presented as t_a and t_f, respectively. The known daily or weekly training load is presented as w_t. The time required to restore performance is presented as t_g, while the time required to peak performance after the completion of training is presented as t_g. The equation is as follows:

$$P_t = P_0 + k_a \sum_{s=0}^{t-1} e^{-(t-s)/\tau_a} w_s - k_f \sum_{s=0}^{t-1} e^{-(t-s)/\tau_f} w_s$$

For example, K_a and K_f were used to present the magnitude of the positive adaptations and negative effects, respectively. If $K_f > K_a$ performance will be influenced negatively over the next days after the completion of a single training session (Busso 2003). A variation has been found for K_a and K_f (Hellard et al. 2006) and this seems to be unrelated to the overall TL. For example, K_1 and K_2 was similar in studies were different types and amount of training were performed (Busso et al. 1997;

Hellard et al. 2005). The model suggests it can estimate performance at any time. Time constants are used to describe the decay of the negative and positive effects and have received criticism due to their variability (Hellard et al. 2006). The time constants were also found to be depended on the nature of the training stress imposed. For example, in running, Morton et al. (1990) found a similar negative time constant to that of swimmers (Mujika, Iñigo et al. 1996) but the TL was less in the study of swimmers. However, several authors have suggested that the positive and negative effects may be difficult to interpret from a practical standpoint (Taha and Thomas 2003). For example, a positive correlation was found between testosterone concentrations and the function of fatigue, which contradicts the anticipated results (Busso et al. 1990). Furthermore this modelling approach requires the repeated measuring of not only training doses but the resulting performances and this is impractical for athletes and coaches (Morton et al. 1990; Busso and Thomas 2006).

Several limitations have been reported when it comes to the fitness and fatigue model (Chapter 2.6.3). These limitations may include the parameter interpretability, predictive accuracy, ill-conditioning and the consistency of the predictions when variables change (Busso et al. 1997; Hellard et al. 2006; Ludwig et al. 2019). Time-varying revision models were proposed (Busso et al. 1997; Avalos et al. 2003; Busso 2003). These modifications were proposed in order to better describe the acute negative effects from a single session, as well as to remove artifacts or abnormalities in the modelled changes of performance over a long period of time. Nevertheless, the fit of parameters varies greatly. Some models have shown high correlations between actual and modelled performances after a period of training (Millet et al. 2002; Busso 2003), while others have shown lower correlations between actual and modelled performances (Avalos et al. 2003; Hellard et al. 2005). Others suggested a high within-subject variability between actual and predicted performances that is seen independently of the models used (Morton et al. 1990; Busso et al. 1991). It is possible that the participants' fitness levels across studies may influence the effectiveness of the performance model. For instance, some studies (Busso et al. 1991; Millet et al. 2002; Busso 2003) showed

promising results when modelling performances for an untrained or sedentary population. In contrast, other studies reported lower accuracy when endurance-based athletes with higher fitness levels were participants (Mujika, Iñigo et al. 1996; Hellard et al. 2005). Some studies reported a similar positive and negative time decay constants (Mujika, Iñigo et al. 1996; Morton 1997; Hellard et al. 2006) with only high performance athletes. However, studies with both high performance athletes and less trained athletes reported the same parameters and determination coefficients (Morton et al. 1990; Busso et al. 1991; Mujika, Iñigo et al. 1996; Hellard et al. 2006). In summary, a high variability has been `found in the parameters used for the fitness and fatigue model may and that may affect how changes in endurance performance are computed.

An important detail in relation to the fatigue resulted from training and how the fatigue effects have been used within the mathematical model may be further discussed. For instance, the effects of fatigue have been proposed differently for some models compared to others. Taha and Thomas (2003) questioned the role of fatigue or negative component and whether it should be implemented in the model. Busso et al. (1991) found that fatigue improved the model fit in only 1 out of 8 participants. The authors attributed their results to the low TL prescribed. In a study where a higher TL was applied in triathletes (Millet et al. 2002), the fatigue component was found to assist to the training and performance relationship in running. Nevertheless, for the swimming training and performance fit the model did not seem to benefit from including the fatigue component (Millet et al. 2002). Taha and Thomas (2003) suggested that the exclusion of the fatigue component may be problematic. Their findings suggested that fatigue should be included. Overall, the negative effects of a training session and how this contributes to modelling the performance changes may require further investigation. In contrast to the fitness decay in performance, the fatigue decay takes place immediately after the completion of a training session (Hellard et al. 2006). However, the nature of fatigue has been described as complicated (Enoka and Duchateau 2016). This may suggest the need for a method that will provide information on the acute negative effects resulted from a training session (i.e training-

induced fatigue), before modelling its chronic or cumulative effects on performance. This may be important for how TL is computed. Given that the fitness and fatigue model is sensitive to the model inputs, and evidence has suggested the need for a robust method of monitoring the effects of these for different training sessions (Viru and Viru 2000). However, there is a surprising lack of evidence on the acute negative effects on performance that result from a session, especially given that fatigue tends to be most marked as a training session terminates (Hellard et al. 2006).

In the context of TL, the acute negative effects imposed by exercise can be examined through the use of psychometric questionnaires (Borresen and Lambert 2009; Saw et al. 2016; Jeffries et al. 2020) or neuromuscular assessments such as countermovement jump (Tillin and Bishop 2009). Although the latter approach is promising, no research has examined whether the acute negative effects of training can be found by changes in such tests. A possible explanation for the lack of those findings is due to the fact that many factors contribute to jumping ability (muscle fiber type, tendons elasticity), and also the nature of the countermovement jump which is not common for running and cycling athletes. The use of questionnaires can assess how exercise affects different processes related to fatigue, such as the stress levels perceived from the athletes, the muscle soreness, mood impairments occurring with training, sleep loss and cognitive impairments related to training (Halson 2014). Those factors are often used to examine the relationship between fatigue, wellness and recovery in sport environments (Halson 2014). Examples of these are the Profile of Mood states (POMS) (Morgan et al. 1987), the Daily analysis of like demands for Athletes (DALDA) (Rushall 1990) the Total Recovery scale (TRQ) (Kenttä and Hassmén 1998), the Recovery-Stress Questionnaire for athletes (REST-Q-SPORT) (Kellmann and Kallus 2001) and lately the Rating-of-Fatigue (ROF) scale (Micklewright et al. 2017). In these scales, there is a dimension that assesses fatigue directly or indirectly and therefore they have been proposed for TL quantification purposes (Saw et al. 2016; Coyne et al. 2018). Jeffries et al. (2020) argued that some scales (POMS, REST Q) may require a lot of time to be completed and that this may be at the expense of athlete frustration,

fatigue and boredom. The potential use of these scales may not fully reflect the complicated nature of fatigue as a phenomenon (Enoka and Duchateau 2016). Fatigue has been characterized as universal (Mehta and Parasuraman 2014) multifactorial and a complex phenomenon and therefore holistic approaches have been recommended (Marino et al. 2011). Therefore it may seem unlikely that these scales could describe the negative effects resulted from training. Given how much effort has been spent on modelling the effects of fatigue on performance (Busso et al. 1991; Millet et al. 2002; Busso 2003) it is suprising that no research has directly attempted to measure how fatigue impairs performance capacity, before modelling its effects.

Other subjective measurements of workload have been proposed as a more universal option to capture the negative factors that affect performance (i.e fatigue, stress, illness) (Hart 2012). A classic example here is The National Aeronautics and Space Administration Task Load Index (NASA-TLX) (Hart and Staveland 1988). Despite the criticism on the use of the term "workload" in the sport context (Winter 2006), the NASA-TLX scale has been lately used in endurance performance (Pageaux et al. 2014, 2015; Rozand et al. 2014; Martin et al. 2016; Cutsem et al. 2017). NASA-TLX scale was proposed to assess the subjective workload imposed by a mentally fatiguing task and/or after exercise. For example, NASA-TLX was used after a cognitive task and after running exercise (Pageaux et al. 2014) or cycling exercise (Pageaux et al. 2015; Cutsem et al. 2017). In other cases, NASA-TLX was used only after the completion of cognitive tasks (Rozand et al. 2014; Martin et al. 2016). These are some examples where the NASA-TLX scale was used to assess both physical and/or mental demands. This subjective method is not popular in sports and this may be due to the equivocal use of the term "workload" within the sports literature (Winter 2006), as well as the universal use of the scale due to the multiple subcategories included (Hart and Staveland 1988). Nevertheless, NASA-TLX has not been used systematically to examine the stress imposed by various doses of endurance training.

The NASA-TLX scale consists of six different subcategories and assumes that the amalgamation between these present the workload imposed by various tasks (Hart and Staveland 1988). These subcategories are independent variables, yet users often suggest that they often correlate with each other (Hart 2012). The six subscales include assessments of mental demands, physical demands, temporal demands, frustration, effort and performance. The mental demands (how mentally demanding was the task?) assesses the mental and perceptual demands required to complete a task. The physical demands (how physically demanding was the task?) assesses the physical activity necessary to complete the task. The temporal demands (how harried or rushed was the pace of the task?) assesses the time pressure applied to the performer and whether the pace of the task was perceived as slow or fast. The performance subscale (how successful were you in accomplishing what you were asked to do?) requires from the performer to assess how successful they felt they were in achieving the task set by the experimenter. The effort subscale (how hard did you work to accomplish your level of performance?) assesses the mental and physical resources required to perform a task. Lastly, the frustration subcategory (how unsecured, discouraged, irritated, stressed and annoyed were you?) assesses the negative feelings/emotions perceived by the performer during the task. After the completion of a task, the performer is required to provide a score for each of the different subscales. The overall workload is then presented as the summated score from all six subscales (Hart and Staveland 1988).

The second method for assessing the dose-response relationship between TL and performance changes uses changes in the training outcome (i.e. fitness or performance) during a given training period (Manzi et al. 2009a; Wallace et al. 2014a). These methods assess whether a training intervention resulted in improvements in endurance performance through changes in aerobic fitness or physiological markers related to endurance performance (Beaver et al. 1986; Jones and Carter 2000). After a period of training, scientists often use these tests to assess whether the application of training resulted in positive outcomes (i.e. fitness improvements). In the context of TL, Manzi et al.

(2009b) showed that individualized iTRIMP was significantly related to running performances over 5000m (r = -0.77) and 10000m (r = -0.82) distances. Similarly, Wallace et al. (2014a) found large correlations between different methods for quantifying TL (sRPE, TRIMPS, rTSS) and actual or modelled running performance. Similar results were shown in team-sports, where an individualized method of TL (i.e. iTRIMP) showed stronger dose-response relationships with changes in physiological thresholds associated with aerobic fitness (2 and 4 mmol-1) (Manzi et al. 2013). In professional cycling, Sanders et al (2016) highlighted TL metrics that use individual physiological characteristics (i.e. HR – blood lactate relationship, functional threshold power) iTRIMP (r = 0.81) and the training stress score (TSS) (r = 0.75) (Sanders et al. 2017a). In summary, these studies and others support the existence of a dose-response relationship between TL and changes in fitness (Manzi et al. 2009b; Akubat et al. 2012; Manzi et al. 2013; Malone et al. 2016; Sanders et al. 2017a). The limitations of this method will be critically discussed in 2.6.1, as it is likely that these methods reflect the performance adaptations, rather the training stress imposed by the training session.

Collectively, the evidence above may suggest the need for a new metric that will directly estimate the negative effects resulting from training (i.e training-induced fatigue). The stress imposed by exercise may be reflected by the negative effect on performance and that decrease in performance may be used as an index of fatigue (see Chapter 2.8). Post exercise performance assessments have been found to be sensitive to a variety of environmental, biological and psychological stressors, such as hot environments (González-Alonso et al. 1999), hypoxia (Koglyn L 2013), muscle glycogen depletion (Heigenhauser et al. 1983), muscle fatigue (Marcora et al. 2008) and mental fatigue (Marcora et al. 2009). The concept of assessing performance after exercise has been used by others for TT (Passfield and Doust 2000; Rodríguez-Marroyo et al. 2017; Clark et al. 2018) and TTE tests (Fullerton et al. 2021) but these studies did not also examine the associated TL. Passfield and Doust (2000) showed that a 5 min TT is reduced following a 60 min bout of moderate intensity exercise, compared with a control no-exercise condition. Recently, Fullerton et al (2021) demonstrated that a

TTE performance assessment dropped 10W after a 30 min session. Therefore, there is evidence suggesting that prior exercise compromises performance capacity. Importantly, the extent of that decrement in performance appears to depend on the exercise intensity and duration (Fullerton et al. 2021). Therefore there is a need to be able to assess systematically the negative effects on performance that result from training.

2.4 Training load metrics

2.4.1 Banister's TRIMP (bTRIMP)

The training impulse (TRIMP) (Banister et al. 1975) was developed from Banister and colleagues. However, the predominant formula as a measure of TL based on weighted HR was introduced a couple of years later (Morton et al. 1990). The study from Morton et al. (1990) was the first study where HR data were first used for the TRIMPs' calculation, and a non-linear weighting of exercise intensity was introduced. This weighting scheme was introduced in order to correct bias caused by the low intensity and long duration training sessions (Banister et al. 1975).

The classic bTRIMP formula uses training duration, HR data and a weighting factor, using the formula below:

$$bTRIMP = excercise\ duration\ (minutes) \times \Delta HR \times y$$

Where the multiplying factor (y) is based on the classically described relationship between increases in blood lactate and HR during incremental exercise (Green et al. 1983). This relationship found to be affected by sex and therefore it varies for males and females and is presented via the following equations :

$$y = 0.64e^{1.92x}\ (males)$$

$$y = 0.86e^{1.67x}\ (females)$$

ΔHR is originated from the following equation :

$$\Delta HR = (HR_{ex} - HR_{rest})/(HR_{max} - HR_{rest})$$

bTRIMP has been used in a variety of endurance sports including cycling, running, swimming and triathlon (Morton et al. 1990; Banister et al. 1999; Millet et al. 2002; Padilla et al. 2008; Wallace et al. 2009, 2014b). bTRIMP has been also used in team-sports (Impellizzeri et al. 2004; Borresen and Lambert 2008; Alexiou and Coutts 2008; Akubat et al. 2012). Studies in cycling (Padilla et al. 2000, 2001, 2008) highlighted the complexity of combining intensity and duration in order to quantify TL. For example, Padilla et al. (2000) attempted to quantify different competition time-trials (TT) in 18 professional cyclists. The time-trials consisted of prolonged TTs (race-opening TT), short and long TTs, uphill TTs and team TTs. The results revealed that the team TTs (146 ± 6 AU) resulted in the highest TL scores compared to long TTs (122 ± 27), short TTs (77 ± 23 AU) and prolonged TTs (21 ± 3AU). Although exercise intensity (mean HR) was higher in the short TT, the study highlighted that the higher scores observed in the team TT were due to the longer exercise durations implemented. This is one of the first studies that highlighted the issues with exercise duration within the TL formulae. Padilla et al. (2001) examined the effects of exercise durations on bTRIMP values and suggested that higher TL scores resulted from high-mountain stages (156 ± 38 AU) compared to semi-mountainous (172 ± 31 AU) and flat terrain (156 ± 31 AU). In contrast with findings from the same research group (Padilla et al. 2000) the high-mountain stages showed the highest exercise intensity. Padilla et al. (2008) recruited 16 professional cyclists during a 3-week Grand Tour. The rides were categorized as off category, first and second category, based on the distance and slope grade of the rides. Off category climbs resulted higher scores (115 ± 30 AU), compared to the first

(72 ± 29 AU) and second (41 ± 20 AU) category. The authors suggested that the sequence that those rides take place can also affect the way bTRIMP quantifies TL. Conclusively, the studies suggest that the complex relationship between training intensity and duration in cycling and the use of HR_{mean} might not represent the actual physiological demands of such high training stress, and consequently misinform the resulted TL scores.

In running, a good correlation was found between bTRIMP (r = 0.65 ± 0.13) with modelled and actual (1500m) performances (Wallace et al. 2014a). Despite, authors suggested the use of sub-maximal HR and HRV alongside HRmean in order to appropriately quantify the effects of training-induced fitness and fatigue. Manzi et al(Manzi et al. 2009b) examined the effects of an 8-weeks running training program and compared TRIMPS (bTRIMP, iTRIMP) and actual performances (5000m and 10000m). No relationship was found between bTRIMP and percentage of speed improvements at 2 $mmol^{-1}$ (r = 0.61) and 4$mmol^{-1}$ (r = 0.59). In addition, no relationship was found between bTRIMP and performances on the track for 5000m (r = -0.41) and 10000m (r = -0.54). The authors explained the lack of sensitivity (relationship with fitness variations) and validity (association with performance) due to the gender-dependent coefficient that bTRIMP used was based, compared to individualized methods.(Manzi et al. 2009b)

2.4.2 Edwards TRIMP (eTRIMP)

Edwards (1993) proposed a different model for quantifying TL, based on HR via the use of a different weighting scheme. Instead of using a coefficient based on the kinetics of one physiological parameter, the time spent in five pre-determined HR zones was multiplied by arbitrary coefficients specific to that zone. HR zones derive as percentages of HR_{max} (zone 1: 50-60% HR_{max}, zone 2: 60-70% HR_{max}, zone 3: 70-80% HR_{max}, zone 4: 80-90% HR_{max} and zone 5: 90-100% HR_{max}). A TL score was then provided by summating the total time spent in each of the five zones, and then multiplied with a weighting factor for each zone (weighting factor 1-5). Studies investigated whether eTRIMP

is a valid TL metric and comparisons between other metrics have been proposed in endurance and team-sport athletes (Foster et al. 2001; Impellizzeri et al. 2004; Alexiou and Coutts 2008; Wallace et al. 2009; Manzi et al. 2010). Alexiou and Coutts (2008) demonstrated strong relationship between eTRIMP and sRPE (r = 0.85) in female football players. To the best of the authors knowledge, only two studies have assessed the dose-response relationship of TL quantified by eTRIMP and changes in cycling performance, providing contradictive results. For example, Sanders et al (2016) showed a strong relationship between eTRIMP and endurance performance parameters (power at 2 and 4 mmol/L) and moderate relationship with performance (r= 0.48) for competitive road cyclists during a 10-week training period. On the other hand, Vermaire et al (2019) showed no relationship between eTRIMP and performance for recreational cyclists over a 12-week training period.

2.4.3 Lucia's TRIMP (LuTRIMP)

Another method based on HR was suggested in order to improve the weighting for exercise intensity (Lucía et al. 2003). In this method, the HR zones are based on ventilatory thresholds. Three zones are identified through a laboratory testing exercise. Zone 1 reflects the HR points below the ventilatory threshold (VT). Zone 2 includes HR points between VT and the respiratory compensation point (RCP), while zone 3 represents the HR points above RCP. The time spent in each zone is then multiplied with a coefficient for each zone providing a total LuTRIMP score. Coefficients of 1,2 and 3 are used for zones 1, 2 and 3, respectively. That TL metric received attention mainly from the scope of how endurance athletes distribute their training, mainly in triathletes and runners (Esteve-Lanao et al. 2005; Muñoz et al. 2014b, a). LuTRIMP has also been used in studies for cycling performance (Lucía et al. 2003; Rodríguez-Marroyo et al. 2009). In the study by Lucia et al. (2003) the authors compared the different TL scores from the Tour de France versus Vuelta a Espana. Higher TL scores were observed for the Tour de France (7112 ± 289 AU) compared to Vuelta a Espana (6700 ± 305 AU), possibly due to the higher exercise duration for Tour de France. The effects of duration on

LuTRIMP calculations were showed also in the study from Rodriguez-Marroyo et al. (2009) in a study were cycling races of different durations were compared. Total LuTRIMP scores were higher for the 21-day stage races (7743 ± 166 AU), compared to 5 and 8-day stage races (2005 ± 52 and 3134 ± 123 AU), respectively. The authors claimed that the fitness levels of the athletes, alongside race strategies played a role on the results observed. The effects of the race strategy on TL quantification were also highlighted by others, suggesting that both training duration and intensity can affect the build-up of fatigue (Earnest et al. 2004).

2.4.4 iTRIMP

Although the validity of the original TRIMP model (Banister et al. 1975) has not yet been examined in the literature, Manzi et al.(2009) proposed another modified version (Manzi et al. 2009b). This modified TRIMP version focused on the individualization of the blood lactate-HR response. Morton et al. (1990) suggested an exponential weighting factor (y) based on the exponential rise of blood lactate. However, the fixed values used for males and females derived from mass testing in males and females and they do not reflect the individual responses to exercise (Manzi et al. 2009b). As such, iTRIMP introduced an individualized weighting factor (y_i), in order to weight the exercise intensity based on an individual's blood-lactate response to incremental exercise. Therefore, every HR data point is weighted accordingly. The weighting factors are calculated individually, based on the individual blood lactate-HR response to an incremental test. HR points which are multiplied by the weighting factor corresponding to each point were then summated and present a total iTRIMP score. iTRIMP has been used in runners (Manzi et al. 2009b, a) cycling (Sanders et al. 2017a, b; Vermeire et al. 2021) football (Akubat et al. 2012; Manzi et al. 2013), kayaking (Borges et al. 2014) and hurling (Malone et al. 2016).

Manzi et al (2009) first used iTRIMP in order to identify the dose-response relationship between TL and performance changes during an 8-week training period. The authors showed a very strong

correlations between speeds at 2 and 4 mmol/L and weekly iTRIMP scores. They also showed a large negative correlation between iTRIMP and running-based performance assessments (r= -0.77 for 5000m, r= -0.82 for 10000m). Even though a strong statistical correlation was found, the authors identified some limitations due to the small number of participants (n=8). In another study from the same research group (Manzi et al. 2009a), the dose-response relationship between the autonomic nervous system (ANS) responses and iTRIMP was assessed. Strong correlations between ANS parameters (HR, systolic arterial pressure variability, baroreflex sensitivity) and iTRIMP (r^2 ranged from 0.90 to 0.99; P < 9.001) were found. ANS parameters and marathon performances showed a very strong correlation too (r^2 = from 0.65 to 0.82) when TL peak. Nevertheless, the small sample size (n=8), as well as the number of athletes that successfully finished the marathon (n=6) was suggested as a limitation of this study.

In cycling, Sanders et al. (2016) assessed the dose-response relationship between various TL metrics and aerobic fitness and performance in competitive road cyclists. The results revealed very large correlations between all TL metrics (bTRIMP, eTRIMP, LuTRIMP, iTRIMP, sRPE, TSS) and submaximal fitness variables (power at 2 and 4 mmol/L). The strongest relationship was found between iTRIMP and changes in aerobic fitness variables (r = 0.81 for changes in power at 2mmol/L, r = 0.77 and at 4mmol/L). Moreover, the strongest dose-response relationship with changes in performance (8-min time trial) was also observed with iTRIMP (r = 0.63). Therefore, the authors concluded on the need for individualized physiological characteristics in order to better quantify the relationship between TL and changes in performance.

2.4.5 Session RPE (sRPE)

The session rating of perceived exertion (sRPE) was introduced by Foster et al. (2001). This model uses the CR-10 RPE scale (Borg 1982) as a subjective measure of training intensity. Alongside, the model also considers the exercise duration in minutes. To acquire a single number as

a TL score, the CR-10 RPE value for a given session is multiplied with the duration of exercise. sRPE is an inexpensive method for quantifying TL and has been used in a variety of sports, including endurance sports (Foster et al. 2001; Herman et al. 2006; Seiler and Kjerland 2006; Barroso et al. 2015; Pinot and Grappe 2015; Fusco et al. 2020) and team-sports athletes (Foster et al. 2001; Impellizzeri et al. 2004; Brink et al. 2010; Lovell et al. 2013; Bradley et al. 2015).

Due to the simplicity of that method, research often used sRPE as a method to validate how other TL metrics responded in different forms of exercise. For instance, Borresen and Lamberts (2008) recruited trained males and compared sRPE with other TRIMP methods, highlighting good individual correlations with eTRIMP (r = 0.84) and eTRIMP (r = 0.76). Manzi et al. (2015) also showed a strong correlations between sRPE and iTRIMP in seven recreational long-distance runners (Manzi et al. 2015). Rodriguez-Marroyo et al. (2012) found similar responses between sRPE and LuTRIMP in professional cyclists (Rodríguez-Marroyo et al. 2012). An interesting observation from this study is that, when TRIMPs were normalized (i.e. TL scores divided by distance), a higher correlation was found between RPE and HR-based methods. Despite the findings mentioned above, the validity of the initial model for calculating TL has never been tested or examined and therefore comparisons between TL metrics in that way can misinterpret the relationship between doses of exercises and training adaptations. Only few studies examined the dose-response relation between sRPE and changes in cycling performance, but not in running. Sanders et al. (2016) showed moderate relationships between sRPE and changes in VO2max and Wmax (r= 0.30 – 0.44), however they found strong relationships between sRPE and changes in power output at 2 mmol/L (r= 0.54) and 4 mmol/L (r=0.60), or performance changes (r= 0.51). A six-year monitoring case study assessed the dose-response relationship sRPE and changes in cycling performance (Pinot and Grappe, 2015). The authors showed increments in TL by 83%, attributing this increase to the exercise duration increments (526hours to 943hours) rather than the changes in RPE(Pinot and Grappe 2015).

The time-administration of the CR-10 RPE scale for calculating sRPE has been often challenged in the scientific literature. Foster et al (2001) suggested that athletes should provide how hard/strenuous was the session 30 minutes after the training is terminated. In that regard, authors suggested that the athletes are able to avoid giving higher scores simply because the last part of a training session was intense. Lately though, others suggested the collection of CR-10 RPE responses as long as the session is terminated and considered valid (Christen et al. 2016). Another key questions regarding sRPE is whether it better corresponds to changes in exercise intensity or duration. Green and colleagues (2009) conducted an experiment to examine the effects of exercise intensity and duration distribution on RPE, and consequently sRPE (Green et al. 2009). Authors compared RPE responses in treadmill trials at a fixed workload over different durations (20,30 and 40 minute). The results showed no differences for HR and blood lactate between trials, in an attempt to isolate the effects of exercise duration on RPE. They concluded that exercise intensity plays a more important role in sRPE calculation. On the other hand, Fusco et al (2020) showed that despite swimmers maintained the same external (lap time) and perceived intensity (HR, La^-) in 4 x 10 x 100 yards, sRPE showed increments.(Fusco et al. 2020) The findings from this study, alongside others (Barroso et al. 2015) highlighted that sRPE may provide information beyond exercise intensity.

2.4.6 Training stress score (TSS) and rTSS

The last few years, cycling power meters were proposed a valuable tool for monitoring training and competition . This method however, were derived from training handbooks and not from scientific investigations. Studies suggested that power meters can provide accurate monitoring for coaches, sports scientists and athletes (Paton and Hopkins 2001; Gardner et al. 2004). Power output through power meters can be considered as a direct method for external exercise intensity (Jobson et al. 2009) and can be analyzed in an easy, user friendly, accessible way, with the example of online platforms such as the TrainingPeaks WKO+.

The TSS is a TL metric originally proposed by Allen and Coggan (2003). This TL metric uses the relative intensity alongside the training duration. Relative intensity is calculated based on an individualized power output, the functional threshold power (FTP). Consequently, training quantification via TSS take into account external TL (i.e. power output), as well individualized approaches (i.e. FTP). FTP is the highest average output that a cyclist can maintain over 60 minutes and represents the power output corresponding at lactate threshold, defined as 1 mmol/L increase above resting levels. That estimate derives from previous work, where mean power output during 60-90 minutes of cycling showed a good correlation with lactate thresholds (Coyle et al. 1988; Bishop et al. 1998; Bentley et al. 2001). From a practical standpoint, shorter duration trials (i.e. 20 minutes TT) have also been suggested to provide robust and accurate FTP measurements.(MacInnis et al. 2018).

The equation suggested from Coggan (2003) derives from the original model from Banister (1975)

$$bTRIMP = excercise\ duration\ (minutes) \times \Delta HR \times y$$

was presented in the following formula:

$$TSS = exercise\ duration\ (minutes) \times mean\ power\ output \times\ power\ dependent\ intesity \times weighting\ factor\ (IF)$$

In Banister's model, the weighting factor for exercise intensity was based on the classically described increase of blood lactate for trained males and females during incremental exercise.(Banister et al. 1975) In Coggan's (2003) model, IF was presented as the power output as a percentage of the corresponding blood lactate concentration (1 mmol/L above resting values) from a large number of trained cyclists.(Allen and Coggan 2006)

$$Blood\ lactacte\ (\%\ of\ lactate\ at\ LT) = power(\%\ of\ power\ at\ LT)^{3.90}$$

Showing a strong correlation with an R^2=0.806, n = 76

Mean power output might not represent the physiological demands of training, unless power output remains constant (Jobson et al. 2009), Coggan (2003) suggested smoothing the data via using 30 seconds rolling mean values. That suggestion was based on the fact that physiological responses in relation to exercise intensity changes are not rapid and proposed rolling mean values instead. The 30 seconds rolling means are then raised to the fourth power in order to re-calculate a mean. Eventually, a fourth root of that re-calculated mean is taken and presents the normalized power (NP). Depending how intense and continuous the effort was for a training session, mean power output and normalized power can vary considerably. In order to provide accurate comparisons between individuals, the IF was expressed as a ratio of NP and the TSS was normalized to the amount of work that could be performed during 1 hour of cycling at FTP (100 points). As such, TSS was calculated using the following formula:

$$TSS = \{(\mathrm{t} \times \mathrm{NP} \times \mathrm{IF})/(FTP \times 3600)\} \times 100$$

In running, the concept of TSS was modified and presented as running training stress score (rTSS) (McGregor et al. 2009). Similarly to Cogan (2003), rTSS is used as an intensity factor (IF), the functional threshold pace and exercise duration similarly to Cogan (2003). rTSS uses the gravity order velocity stress score (GOVSS)(Skiba 2006) algorithm. Similarly, rolling means from running speeds are raised to the fourth power for re-calculating the mean and the fourth root was obtained to determine an NP for the entire training session. FTP is calculated based on individualized speed and is defined as the highest average pace that a runner can maintain over 60 minutes.

Only few studies have assessed the dose-response relationship between TSS (or rTSS) and changes in cycling or running performance. In competitive cyclists, Sanders et al. (2016) demonstrated strong relationships between TSS and changes in power output for 2 (r =0.75) and 4 mmol/L (r = 0.79). The relationship between TSS and performance (8mTT cycling test) showed a moderate relationship (r = 0.79)., however when performance improvement were expressed as relative power output ($W \cdot kg^{-1}$), a strong relationship was found (r = 0.61) (Mujika 2017). In running,

Wallace et al (2014) compared different TL metrics and correlated their responses with actual and modelled performances (1500m). All TL metrics showed large correlations between actual and modelled performance, and notably, rTSS was found to correlate better with performance compared to bTRIMP or sRPE. The authors suggested that the results observed in this study were due to the fact that rTSS takes into account the individual performance characteristics compared to other TL metrics used in that study. Again, TSS and rTSS may seem a better choice compared to generalized methods, however the non-scientific background of these methods and the fact that they also derive from TRIMPs may suggest it is used with caution. As stated earlier a limitation for all TL metrics that have been developed based on Banister's work, a key step in the scientific process (validity) has never been tested or examined thoroughly. The capacity of TSS to inform dose-response relationship between TL and performance should be examined in detail and will be thoroughly discussed in Chapter 2.6.3

2.4.7 Conclusion

Many different TL metrics have been proposed in the context of TL. All TL metrics attempt to use and weight the effects of exercise intensity differently. Despite the dose-response relationship between TL and performance (Chapter 2.2.3), there is a debate on the best TL metric to use. So far, TL metrics have found to agree (Wallace et al. 2014a; van Erp et al. 2019a), disagree (Manzi et al. 2015; Sanders et al. 2017b) or agree for some intensities and disagree for some others (Borresen and Lambert 2008; Alexiou and Coutts 2008). Currently, different TL metrics were suggested to inform the dose-response relationship between TL and performance (see Chapter 2.6.4). In summary, no agreement has been found between TL metrics and this may be problematic when training monitoring is of concern.

2.5 The Concept of TWD

The concept of TWD has been used often for both research and applied sport purposes. In the context of science, TWD is related to the prescription of the training dose. The prescription of exercise in endurance performance involves combinations between exercise intensity and duration (Stewart et al. 2016) or equivalent EE (Gibala et al. 2012). In the study from Stewart et al. (2016), two different exercise intensities were used and defined as heavy-intensity and moderate-intensity. Exercise duration was longer (120 min) for the moderate-intensity cycling trial compared to the heavy-intensity cycling trial (90 min). This is an example where the authors attempted to match two endurance cycling trials for the same total mechanical work. The concept of TWD is also used for prescription of continuous or intermittent endurance training sessions (Gibala et al. 2012; MacInnis and Gibala 2017; MacInnis et al. 2017). In the methodology of these studies, exercise intensity and duration was also used to form INT and CON training. Gibala et al. (2012) concluded that when INT and CON exercise is compared on the basis of TWD, INT resulted in similar or superior adaptations in a range of physiological and performance markers. Other researchers criticized the concept of TWD (Chapter 2.6.1) and proposed other methods for prescribing INT and CON training sessions (Seiler et al. 2013; Nicolò et al. 2014). Despite the concerns of the effectiveness of TWD for prescription purposes, this concept still remains a popular method for sports science.

TWD is also used in an applied sport context. The measurement of TWD has been important for coaches and athletes for three reasons. Firstly, the TWD is often related to the training specificity principle. An example here can be training an 800m runner has to undergo. The athlete needs to know ‘’how much’’ training needs to be done and most of the cases it is expressed as the milage that needs to be covered. In that context, TWD may be an important tool. Another example is when a cyclist is looking for training in order to sustain a given work rate (i.e. average PO) for the longest possible time. Secondly, TWD has received attention on the field as a proxy of the rate of energy expenditure. Thirdly, traditional TL measures are computed based on the principle of calculating TWD. That is

due to the multiplication between intensity and duration. That concept was not popular in cycling but mostly in running. Nevertheless, it became popular in cycling through the use of power meters technology. Wallace et al. (2014b) conducted a study and reported that TWD is the most valid measure of TL. Other research groups have also mentioned its significance. Similarly, van Erp et al. (2019) demonstrated a robust correlation between TWD and common internal TL metrics (TRIMPs, sRPE) when training and racing of elite cyclists was analyzed. The TWD concept underpins the TL metrics, however concerns about the TWD concept still exist and will be further explored in Chapter 2.6.1.

2.6 Limitations of the TL Concept

2.6.1 The validity of TL metrics

The TRIMPs method (Banister et al. 1975) was originally introduced in order to predict changes in performance, but was never been validated against a criterion of training stress. In later years, unvalidated TL metrics have often been compared against other unvalidated methods in order to ensure scientific validity. An example is sRPE which was validated against Edward's HR-based TRIMP (Foster et al. 2001). Foster et al. (2001) used eight training sessions and quantified them via the sRPE and eTRIMP method (Edwards 1993). The authors concluded that sRPE can be used as a subjective estimate of TL. Importantly, Foster et al. (2001) did recognize a potential limitation caused by the multiplication of intensity and duration, especially for higher exercise intensities. Another experimental study that examined the validity of the TL concept was performed by Wallace and colleagues (Wallace et al. 2014b). However, the methodology used was inappropriate, as the authors compared internal and external TL metrics against the total VO_2. The study reported that TWD was the most valid measure of TL. There are plenty of reasons why TWD should not be considered the gold standard for TL monitoring and these will be discussed below. Further, this conclusion

contradicted a comment from the same authors in their paper that the internal TL better describes the stress experienced by the athletes.

The first concern regarding the validity of TL load is due to the studies that attempt to estimate the training stress by combining indices of exercise intensity and duration rather than to measure the training stress directly (Manzi et al. 2009b, 2015; Sanders et al. 2017a, b; van Erp et al. 2019b, a). This may be justified because estimating the training stress via TL is more practical than measuring the stress imposed on the human body by exercise, which can be measured at different levels such as cellular, tissue, system and functional levels. Nevertheless, the relationship between TL and changes in performance may be affected by many other factors. For example, the training process cannot ignore aspects such recovery (Bourdon et al. 2017), nutritional strategies (Jeukendrup 2017), percentage of time affected by injuries (Drew et al. 2017), negative life-event stress (Fry et al. 1991), daily hassles (Lehmann et al. 1993) and coaching style (Davis et al. 2018). Collectively, all these aspects have been suggested to affect the relationship between the doses of training and potential adaptations. Therefore, it seems logical to assume that by assessing performance through long-term studies, a greater focus has been given on the adaptations rather than the training stress imposed by training. Nevertheless, the evidence above suggests that there are a lot of factors that affect how an individual adapts to training. As such, the validity of the TL concept may need to be examined in relation to the training stress imposed rather than the long-term adaptations resulted from multiple training sessions. So far, TL metrics are used to estimate the training stress, however , the relationship between TL and training stress remains to be examined experimentally.

Another concern about the validity of the TL concept is due to the link with the TWD. This is due to the fact that both the TWD and the TL concept use the multiplication between intensity and duration. There is plenty of evidence why TWD may not be the foundation for TL metrics and this can be attributed to the following reasons. To begin, TWD may not provide an appropriate representation of the exercise dose, and should not be used for matching sessions when training

intensity and duration are different (Nicolò et al. 2016b; Passfield et al. 2017; Renfree et al. 2021). For instance, Passfield et al. (2017) provided a useful example to show the limitations of the TWD concept. The authors raised that a single, 4 min effort at 400W cannot be differentiated from 4 sets of 1 min efforts at 400W. Importantly, the stress imposed by these two training formats is unlikely to be similar (Passfield et al. 2017). Rather than the TWD, it was suggested that the participants' effort may be a better indicator of the stress experienced from exercise (Nicolò et al. 2016b). Further, the concept of TWD has received criticism in regards to the relationship between TWD and exercise duration (Nicolò et al. 2016b). This criticism (Nicolò et al. 2016b) was based on work from others (Stewart et al. 2016) where increased duration was not accompanied by an increase in TWD and thus, the effort was decreased. As such, when exercise intensity decreases, the TWD also has to be increased rather than fixed, if the aim is to keep an athletes effort constant (Nicolò et al. 2016b). Thirdly, there are differences based on the type of session imposed. These differences can be detected with continuous and intermittent type of endurance sessions are performed (Nicoló et al. 2014). It has been suggested that constant exercise is less stressful compare to interval type sessions, and that a longer TWD is required for the same effort (Nicoló et al. 2014). Furthermore, others (Turner et al. 2006; Davies et al. 2017) proposed that when both the work-to-rest ration in interval exercise and the TWD are fixed, the metabolic strain increases with the duration of the intervals. These studies collectively demonstrated the problems with intensity and duration and how these are reflected in the TWD concept. Lately, Renfree et al. (2021) suggested that the way in which the session is being performed may potentially be more important than its resulted TWD. In summary, the lack of established scientific validity for TWD and related TL metrics as a measure of training dose may not be a small error, rather, it may be a fundamental problem for TL quantification. This will be further supported from the findings of the thesis and will be discussed in Chapters 3,4, 5 and 6.

2.6.2 The Limitations of the Fitness and Fatigue Model

The model proposed from Banister et al. (1975) and used by others later (Millet et al. 2002; Avalos et al. 2003; Busso 2003; Busso and Thomas 2006) has some limitations that need to be noted. The greatest limitation of this model is that the more the athlete trains, the more the performance will increase without any limit. TRIMPs are proportional to the TL and therefore, as TL increases, fatigue and fitness increase to an indefinite point. However, there is plenty of evidence that there is a limit to TL and above that limit no adaptations occur (Fry et al. 1992; Morton 1997). Equally, the model does consider a limit on performance, which from a practical perspective seems unrealistic. Another limitation of that model is that is assumes that performance is explained by the dynamics of TL only. There is plenty of evidence on the fact that athletic performance is known to be a multifactorial process (Banister et al. 1975; Mujika, Iñigo et al. 1996; Avalos et al. 2003). Performance is known to be affected by psychological and nutritional factors (Banister et al. 1975; Morton et al. 1990) and technical factors (Toussaint and Hollander 1994; Wakayoshi et al. 1995). Therefore, changes in performance may not be only related to TL and thus why mathematical models may result in poor predictions (Taha and Thomas 2003).

Another limitation of fitness and fatigue models proposed by Bannister et al. (1975) has been highlighted from Jeffries et al. (2021). It was suggested that training might have more broad effects, and that instead of using fitness and fatigue as terms, a broader terminology should be used such as negative and positive training effects. For example, fitness may be defined by a set of attributes such as muscular strength and endurance, body composition, flexibility, reaction time, power, agility and cardiorespiratory fitness (Wilder et al. 2006). Therefore, Jeffries et al. (2021) argued that these definitions are related to a set of outcome measures (i.e. output) and may not reflect the fitness component (i.e. input) proposed by the fitness and fatigue model (Banister et al. 1975). Similar concerns about the fatigue component in existing model (Banister et al. 1975). Some important complications about the fatigue component were also mentioned earlier in Chapter 2.3. For instance,

the term fatigue proposed by Banister et al. (1975) may not be enough to describe the overall negative effects of training. In particular, the authors proposed that fatigue is not the only negative effect being derived from training. For example, as said in Chapter 2.2, training is known to induce muscle damage (Peake et al. 2017), as well as mood suppression and increases in depression (Morgan et al. 1987). Another example is that mental fatigue has been previously been shown to impair endurance performance (Marcora et al. 2009; Smith et al. 2015), but how mental fatigue specifically contributes to the training-induced fatigue is currently unknown. Lastly, the fitness and fatigue model recognized that contextual factors (i.e. health, nutrition, environmental factors, training status) can have an effect on performance but did not attempt to consider them in the model. Indeed, Banister et al. (1975) acknowledged that emotional factors may affect performance. However, only recent frameworks have attempted to clarify and propose the integration of these factors (Impellizzeri et al. 2019; Jeffries et al. 2021). In summary, these findings suggest that the fitness and fatigue model may has limitations and these may equally apply to the TL concept.

2.6.3 Dose-Response Relationship and TL metrics

There evidence for whether TL metrics can inform the nature of the dose-response relationship is equivocal. So far, there is a fair amount of evidence suggesting a liner dose-response relationship between TL and changes in fitness (Manzi et al. 2009b; Akubat et al. 2012; Manzi et al. 2013; Malone et al. 2016; Sanders et al. 2017a) For example, Sanders et al (2016) highlighted TL metrics that use individual physiological characteristics alike iTRIMP ($r = 0.81$) and the training stress score (TSS) ($r = 0.75$) showed a better dose-response relationship between TL and performance. Similar results were shown in team-sports, where an individualized method of TL (i.e. iTRIMP) showed stronger dose-response relationships with changes in physiological thresholds associated with aerobic fitness (2 and 4 mmol-1) (Manzi et al. 2013). Nevertheless, others have suggested that the relationship

between TL and performance is not straightforward (Foster et al. 1996; Taylor et al. 2017; Vermeire et al. 2021). For example, in a study conducted for cyclists, runners and speed skaters, poor correlations (r = 0.29) between sRPE and performance assessments were found (2 miles TT for runners, 5km or 10km for cycling) (Foster et al. 1996). In cycling (Vermeire et al. 2021) showed that higher weekly scores for all HR-based TL metrics (bTRIMP, eTRIMP, LuTRIMP, iTRIMP) did not correlate with changes in performance. Vermaire et al (2019) concluded that expectations for improvements in performance are highly unlikely to occur due to increases TL. Taylor et al. (2018) suggested iTRIMP and bTRIMP but not LuTRIMP, eTRIMP and sRPE as appropriate methods for monitoring the dose-response between TL and changes in fitness. This may be a surprising finding given that all TL metrics are based on the TWD concept and calculate the TL similarly. Another important detail on the dose-response relationship between TL and performance is that , the human body has a limited capacity to adapt to a given TL (Hellard et al. 2005). Indeed, Hellard et al. (2005) and others (Foster et al. 1996; Taylor et al. 2017; Vermeire et al. 2021) suggested that the relationship between the doses of training and training adaptations are not linear.

Overall, it is unclear which, if any, is the best TL metric for endurance activities. Often, studies have compared how different TL metrics affected performance and proposed some TL metrics compared to others (Padilla et al. 2001; Borresen and Lambert 2008; Alexiou and Coutts 2008; Rodríguez-Marroyo et al. 2012; Wallace et al. 2014b, a; Borges et al. 2014; Manzi et al. 2015; Sanders et al. 2017a, b; van Erp et al. 2019b, a; Van Erp et al. 2020) For example, some studies proposed the use of power output (Sanders et al. 2017a, b; van Erp et al. 2019b; Van Erp et al. 2020) and correspondent metrics (TSS). Others suggested the HR-based metrics better described the TL-performance relationship (Manzi et al. 2009b; Akubat et al. 2012). On the other hand, many other studies found a good relationship between RPE and HR-based metrics (Padilla et al. 2001; Borresen and Lambert 2008; Rodríguez-Marroyo et al. 2012; Wallace et al. 2014a, b; Borges et al. 2014; Manzi et al. 2015; van Erp et al. 2019a). Some other studies showed similar responses between TL metrics

based on RPE and HR for some intensities, yet disagreed in higher exercise intensities (Impellizzeri et al. 2004; Borresen and Lambert 2008; Alexiou and Coutts 2008; Manzi et al. 2015; van Erp et al. 2019a). The characteristics of each variable may explain the contradiction observed between TL metrics. For example, HR as a physiological variable has received criticism on whether it can reflect the demands of exercise (Foster et al. 2001; Borresen and Lambert 2008) at higher intensities and that may have affected how HR-based metrics responded across studies. In contrast, RPE was suggested to reflect exercise intensity better compared HR-based methods when both aerobic and anaerobic metabolic systems are activated (Earnest et al. 2004). Indeed, the use of RPE was also suggested from others (Impellizzeri et al. 2004; Alexiou and Coutts 2008; Manzi et al. 2015; van Erp et al. 2019a). Interestingly, power output (Sanders et al. 2017a; van Erp et al. 2019b), HR (Manzi et al. 2009b; Akubat et al. 2012) or HR and power output-based metrics (Sanders et al. 2017a, b) were suggested as a better choice compared to RPE-based metrics. Several studies have suggested that these differences are due to the nature of the measurements (Borresen and Lambert 2008; Alexiou and Coutts 2008; Rodríguez-Marroyo et al. 2012), while some others suggested that it may be due to the weighting factor used for each metric (Sanders et al. 2017b; van Erp et al. 2019b). Exercise intensity is weighted differently in many TL metrics, with most following a non-linear weighting (bTRIMP, TSS, rTSS, iTRIMP) but some adopting a zone-based weighting (eTRIMP, LuTRIMP). Independently of the different weighting for each TL metric, the variable used to reflect exercise intensity for each method may play an important role. For example, iTRIMP uses lactate as part of determining the non-linear weighting of exercise intensity. Nevertheless, intermittent-type of activities have been shown to alter the blood lactate-HR relationship and consequently affect how the dose of exercise is presented via current TL calculations (Akubat and Abt 2011). Lactate is also found to be affected during competition periods, where anaerobic-type of exercise is often performed more frequently (Seiler 2010). Overall, the contradictive findings in the literature do not provide enough evidence to recommend the use of a particular TL metric.

2.6.4 Training Prescription and Monitoring

The variables of both training intensity and duration are involved for both the prescription and monitoring purposes along with training frequency. The influence of intensity and duration for training prescription and quantification have arguably been biased in the attention they receive. The role of exercise intensity for exercise prescription has been discussed extensively in the literature (Mann et al. 2013; Iannetta et al. 2020). The concepts of power output, speed, VO_2 and HR for training prescription are often used as thresholds or fixed percentages of maximum values such as %PPO, %PTV, % VO_2 and %HR, respectively. The threshold approach has found to be more efficient compared to fixes percentages, although this approach has received criticism due to the many laboratory visits required (Mann et al. 2013). Furthermore, Iannetta et al. (2020) suggested that the responses from any percent-derived or zone-derived value may be unpredictable and will vary among different individuals. Moreover, an inter-individual variability exists when it comes to the training responses and therefore, terms such ‘’responders’’ and ‘’non-responders’’ have become more commonplace (Mann et al. 2014).

There is not a gold standard method of classifying exercise intensity for the purposes of training prescription. The use of RPE is a simple and practical method for exercise prescription in endurance exercise. This method uses the concept of perceived exertion to guide athletes towards pacing the prescription, for example of a “maximal session effort’’ and this is known as isoeffort prescription (Seiler and Sjursen 2004; Seiler et al. 2013; Nicoló et al. 2014). The CR-10 RPE scale (Borg 1982) was also found useful for the prescription of cycling training sessions of different intensities and durations (O’Grady et al. 2021) and resistance-type of training sessions (Helms et al. 2018). Lastly, peak power output (PPO) has also found to be useful for the prescription of interval training (Hawley et al. 1997; Rønnestad et al. 2020) and has shown to be well linked with VO_2max (Hawley and Noakes 1992). In running, the concept of peak treadmill velocity (%PTV) is proposed as a prescription method (Manoel et al. 2017) and it has been suggested to link well with lactate thresholds

(Noakes et al. 1990). Despite the usefulness of these approaches, the determination of VO_2, La^- and %PTV is highly depended on the incremental protocol used (Bentley et al. 2007) and their collection requires a laboratory visit or repeated testing which are not practical options for practitioners and coaches. In summary, there is no clear evidence for determining the best exercise intensity marker to use for training prescription purposes.

Although training prescription is not the main concern of this thesis, the findings discussed above have implications for the concept of TL. The relationship between TL and performance is not straightforward (Foster et al. 1996; Taylor et al. 2017; Vermeire et al. 2021), but exercise intensity seems an important variable for inducing TL (Foster et al. 2001; Impellizzeri et al. 2004; Alexiou and Coutts 2008; Manzi et al. 2009b; van Erp et al. 2019a; Van Erp et al. 2020). There seems to be little clarity in terms of what is the best exercise intensity variable to exercise prescription. The way training is prescribed is also related to monitoring and how training is scheduled, due to the recovery needed between training sessions. Nevertheless, the absence of a gold standard for inducing TL (prescription) may equally pose limitations when it comes to quantify the effects of a training session.

Training duration has also received attention for prescription (Esteve-Lanao et al. 2005; Seiler and Kjerland 2006; Seiler 2010; Hofmann and Tschakert 2017). Different studies have attempted to include intensity and duration together for prescribing exercise (Seiler and Kjerland 2006; Seiler 2010; Hofmann and Tschakert 2017), but exact durations were not proposed for each intensity domain (Esteve-Lanao et al. 2005; Seiler and Kjerland 2006; Seiler 2010). Few attempts have been made to isolate the effects of exercise duration specifically (Tremblay et al. 2005) and its effects in acute and chronic performance adaptations consequently remain currently unclear. It has been suggested that there is a need for an individualized prescription when it comes to the effects of exercise duration (Hofmann and Tschakert 2017). This is reinforced by the findings of a recent study where the independent and individualized effects of exercise duration and distinct duration domains were found (Tschakert et al. 2022). Only few attempts have been suggested to individualize the

prescription of exercise duration. For example, the concept of maximum duration (t_{max}) at a given intensity has been proposed (Lutttikohlt et al. 2006). In this method, training sessions of maximum intensity progressively increase the TWD when plotted against increases in exercise duration. Therefore, every training duration is represented as a percentage of what an athlete can achieve in terms of intensity ($\%t_{max}$) (Hofmann and Tschakert 2017), so that duration is expressed as a percentage rather than being used as an independent multiplier as with most TL metrics. Other authors (Ross et al. 2015; Lin et al. 2021) have suggested that this may be a promising method for addressing heterogeneous acute responses and long-term performance adaptations and that this approach may partially solve the ‘’non-responders’’ problem. Importantly, the prescription of intensity and duration was suggested to influence the recovery time needed between sessions (Tschakert et al. 2022). For example, more intense and short in duration training sessions seem to require longer recovery compared to less intense but longer in duration training sessions (Kenttä and Hassmén 1998; Russell et al. 2013; Hofmann and Tschakert 2017). This is another area where training prescription and monitoring link with an aim to modulate the training dose and consequently affect the design of repeated training sessions. These findings collectively suggest that exercise duration may need to be individualized for prescription purposes and normalized according to overall the training session intensity. This in turn implies that monitoring may also require an individualized approach towards weighting the effects of both intensity and duration.

It seems that the role of exercise duration prescribed matters for the resulted training effects. Durnin et al. (1960). reported that lower exercise intensity but longer session duration resulted in the largest fitness changes in untrained men. Shephard (1968) showed that even with low exercise intensities (<40% of VO2max) longer durations can result improvements in performance of untrained participants. Davies and Knibbs (1971) highlighted the importance of higher intensities for greater training effects compared to duration. Nevertheless, these researchers acknowledge that their training sessions involved shorter exercise durations compared to the protocol used by Shephard (1968). The

effects of duration have been shown to improve VO_2max independently of exercise intensity, frequency and length of the intervention period (Wenger and Bell 1986). In this study, the authors used different training sessions of intensity, duration and frequency and its effects on VO_2max were examined. Their results showed that maximum exercise intensities (90-100% of VO_2max) resulted to improvements in aerobic power. Nevertheless, they also showed that more training sessions of longer durations (>35 min) with lower intensities (50% and 70% of VO_2max) resulted in greater improvements compared to the shorter durations (25-35 min) but higher intensities. Evidence for the relationship between intensity and duration and their effects were previously examined prior to the study from Wegner and Bell. (1986). Despite the findings from these early studies, there seems to be a lack of clarity on the training effects resulted from exercise duration and how that may affect TL monitoring.

Current TL metrics do take into account the effects of exercise duration but the interaction between exercise intensity and duration is not included. Instead, TL metrics use the concept of exercise duration as a multiplying factor in absolute values with training intensity. This is interesting, given how much attention has been given towards exercise intensity and its effects (Sanders et al. 2017a, b; van Erp et al. 2019b), whilst exercise duration has not received much attention in the TL literature. Given the early evidence on the effects of both intensity and duration on performance (Durnin et al. 1960; Shephard 1968; Davies and Knibbs 1971; Wenger and Bell 1986) Banister et al. (1975) could have considered the effects of intensity and duration differently. Years later, the weighting of exercise intensity was adjusted (Morton et al. 1990) but surprisingly without consideration for the effects of exercise duration as well. The role of exercise duration in TL monitoring has been questioned recently (Weaving et al. 2020). In particular, Weaving et al. (2020) found exercise duration to be the main contributor to the variance in TL for professional rugby players. In an opinion piece Renfree et al. (2001) suggested that the inclusion of intensity and duration may not account for the non-linear nature of the training stress. Whilst there are suggestions for the

inclusion of both in a weighted manner when it comes to training prescription (Tremblay et al. 2005; Hofmann and Tschakert 2017; Tschakert et al. 2022), similar suggestions have not yet been proposed for monitoring purposes.

2.7 Statement of the Thesis - Aims and hypothesis

Traditionally, endurance training quantification has been proposed through its resulted TL. The concept derives from the work of Banister et al. (1975) and explained earlier in this thesis (see Chapter 2.1). Despite its long-standing popularity, concerned were expressed earlier in regard to how TL metrics quantify endurance training. For example, some concerns were related to the terminology used and what the TL metrics specifically measure (Chapter 2.2). Moreover, concerns were also expressed on the concepts that underpin the TL monitoring such as the fitness and fatigue model (Chapter 2.6.2), the dose-response relationship between TL metrics and performance (Chapter 2.6.3) and the TWD concept (Chapter 2.6.1). These concerns may suggest the need for an alternative on how TL is conceptualized and quantified by scientists, coaches and athletes. This thesis is aiming to directly assess whether TL metrics are associated with decreases in performance. Specifically, a performance test will be performed after different training sessions in order to quantify their effects. As exercise is expected to lead to an acute performance decrement (APD), this thesis aims to verify if APD is associated with currently used TL metrics when the training characteristics are modulated systematically. Accordingly, 3 studies were designed. The purpose of the first study was to assess the association between APD and TL metrics when comparing training sessions with different intensities and durations. The purpose of the second study was to extend these findings to running and evaluate whether an association is found if continuous and intermittent sessions are compared. Lastly, the third study systematically assessed the effect of exercise duration on TL metrics and APD at different levels of exercise intensities. Collectively, these studies should provide a new direction on how endurance training should be quantified. This thesis hypothesizes that the APD responds to different training sessions will contradict those from TL metrics.

2.8 Methodological considerations for APD

In the view of the concerns mentioned above (Chapter 2.6), there seems a need for alternative approaches for TL quantification. For example, in TL may need to be examined from the scope of the training stress applied by exercise, rather than the scope of training adaptations (Chapter 2.6.1). When exercise, the training stress applied and its effects can be measured at different levels such as cellular, tissue, system and functional levels. Expensive and laborious methods may offer only a snapshot of how the body reacts to exercise. Importantly, the training stress imposed by exercise needs to be examined via a method that is sensitive to a variety of psychobiological stressors that may interact with exercise. From a functional level, the training stress imposed by a single training session and its effects may be conceptualized and quantified as the detrimental effect that this training session had in performance. This conceptualization may be considered in line with how Banister et al. (1975) conceptualized TL. Nevertheless, as said earlier in Chapter 2.3, no attempt has been made to directly assess the negative effects of training on performance.

A useful tool for such purposes may be the APD, or the post exercise performance assessments firstly mentioned in Chapter 2.3. Post exercise performance assessments were shown to be sensitive to a variety of environmental, biological and psychological stressors related to exercise (Heigenhauser et al. 1983; González-Alonso et al. 1999; Marcora et al. 2008, 2009; Koglyn L 2013) and were previously used in endurance exercise (Passfield and Doust 2000; Rodríguez-Marroyo et al. 2017; Clark et al. 2018; Fullerton et al. 2021). These studies in endurance exercise suggested that prior exercise compromises endurance performance. Thus, the extent that performance is affected by prior exercise may be used to quantify the effects that a training session may have on an individual. In that way, the training effects are directly measured, rather than quantifying the amount of training completed, related indices such total oxygen uptake, total work done and similar traditional measures of TL. Conceptually, APD is proposed as an evaluation of the training stress imposed by exercise, rather than a measure of TL. Nevertheless, later in this thesis (Chapters 3,4 and 5), TL metrics and

how they represent the training dose will be examined versus APD. There is a lack of studies where the acute decrement in performance is systematically evaluated after various training sessions consisted of different intensities and durations. This systematic evaluation will be attempted experimentally in Chapters 3,4 and 5. APD will be calculated as the observed change in performance, as a consequence of as training session. Mathematically,

$$Acute\ decrement\ in\ performance = (P2 - P1)/P1 \cdot 100$$

Where P1 is the baseline performance established without prior exercise, and P2 is performance after different training sessions. APD will be calculated as a positive percentage change, to allow comparisons with the corresponding TL metrics. Collectively, APD is a novel methodology and is proposed to address the issues of TL mentioned earlier (Chapter 2.6). It may contribute towards a conceptual shift in terms of how scientists, coaches and athletes think TL and consequently quantify it.

The choice of the performance test to use (i.e TTE or TT) and the experimental design may vary according to the type of sport and competition level. Chapters 3, 4 and 5 used a randomized, crossover, within-subjects design with a careful control of potential confounding factors. In all experimental studies, TT's were preferred over TTE's, due to lower intraindividual variability (Currell and Jeukendrup 2013). In particular, relatively short TT's were implemented for both cycling (Chapter 3) and running (Chapter 4 and 5). The characteristics of a training sessions may also play a role on how performance is reduced. Fullerton et al. (2021) demonstrated that both the intensity and the duration elements of the prior training session seem to affect the extend of that decrement in performance. Later (see Chapters 3,4 and 5) and as part of the systematic evaluation of prior exercise on performance, training sessions with different intensities and durations will be used. This systematic evaluation will provide insights of how APD changes after training sessions with different combinations of intensity and duration. For example, in Chapters 3 and 4, relatively short (≥ 10-min) but very intense training sessions (TT's) were used. Alternatively, longer training sessions (25-

40minutes) with lower intensities were implemented (Chapter 4 and 5). Chapter 5 used a long duration (40-min) and high intensity (RPE 8). The selection of the training sessions aimed to replicate how typically endurance athletes train and were adjusted according to the hypothesis of each experimental chapter.

Measuring performance after a training session may impose several challenges for the participants, as well as the researcher involved. For example, all experimental visits conducted in Chapters 3,4 and 5 required from the participants to perform a maximal effort trial after a training session. That made all visits demanding for participants, and therefore all visits were separated by at least a 48-h window. Another challenge when designing the experimental chapters was the rest time provided between the running or cycling training sessions and the performance assessment. The performance assessment had to take place relatively close to the training session. If the training session and the performance assessment were conducted too far apart, that would risk the possibility to capture the effects from that training session. Therefore, the training session and the performance assessment had to take place immediately after, or with a relative short rest in between. Nevertheless, the exact time that should be provided between the training sessions and the performance assessments was unknown. For example, studies that used the acute decrement in performance used either no rest (Passfield and Doust 2000; Clark et al. 2018; Fullerton et al. 2021) or assessed performance the day after (Rodríguez-Marroyo et al. 2017). The idea of assessing performance immediately after a training session was excluded, due to the purpose of the studies and the thesis collectively. For example, if no rest was provided, the collection of important information for both the training sessions (i.e sRPE scale, NASA-TLX) or the performance assessment trials (i.e motivational scales) would not be possible. Further, it would not allow participants to catch their breath prior to completing a TT. That was important, especially when TT's were implemented as initial training sessions. For those reasons, multiple pilot trials were conducted prior the collecting data for Chapter 3. These pilot

studies suggested that a 5-minute rest is a suffice time frame that allows the collection of data and athletes to exert maximal effort trials.

The aim of all experimental Chapters was to capture the detrimental effects of a training session on APD. Conceptually, this attempt will relate the training dose to its effects, or APD. As said in Chapter 1, there is a lack of evidence on how a single training session precisely develops fatigue and only conceptual attempts have been made (see Figures 1.1). In other words, the magnitude of fatigue resulted after a training session is unknown. The way humans respond to training seems to be a highly individualized process and it is likely to depend on many factors. For example, different training-induced fatigue will be resulted from similar training sessions in individuals with different fitness levels. Similarly, different training sessions of intensity and duration will affect a group of athletes with similar characteristics in a different manner. This add complexity to the conceptual link between the training dose, its resulted fatigue and when APD should be administered.

Ideally, it will be very informative to track the time course of APD after an acute bout of exercise as conceptualized by the supercompensation principle. However, this is not feasible from an experimental perspective and therefore, a within-subjects design was implemented with the 5-minute rest provided between the training sessions and the performance assessments maintained across all visits for all participants.

Chapter 3: Cycling Performance and Training Load: Effects of Intensity and Duration

3.1 Aims and hypothesis

The first aim of Chapter 3 was to experimentally introduce the APD concept (see Chapter 2.8). APD will be used in this Chapter to assess if cycling training sessions with different characteristics will result in a detrimental effect on performance. These cycling training sessions will be consisted of different exercise intensities and durations and their TL will be calculated. As the TL is associated with the TWD, this chapter aimed to challenge that link. The way TL metrics quantify training is via the multiplication between intensity and duration, which is conceptual similar to the TWD concept. To challenge this idea and based on the doubts expressed earlier (Chapter 2.6), cycling training sessions with noticeable differences in their resulted TWD were used. In particular, a 5-min max, a 20-mim max, a 20-min sub-max and a 40-min sub-max. The goal here was to observe how TL metrics quantify a training session with a relative low TWD (i.e 5-min max), compared to training sessions with substantial differences in terms of their TWD (i.e 20-min sub-max and 40-min sub-max). Another aim of this study was to compare how differences in exercise duration affect the way TL metrics quantify cycling sessions when maximal exercise. The 4 times higher duration, despite the same intensity, will provide useful information on how TL metrics assemble compared to how APD develops. Based on the above, it was hypothesized that the traditional TL metrics that depend on the TWD concept may not agree with APD.

3.2 Abstract

Purpose: To examine the effect of cycling exercise intensity and duration on subsequent performance; and to compare the resulting acute performance decrement (APD) with total work done (TWD) and corresponding training load (TL) metrics.

Methods: 14 male cyclists performed a 5-min time-trial (TT) as a baseline, and after 4 initial exercise bouts of varying exercise intensity and duration. The initial exercise bouts were performed in a random order and consisted of a 5- and a 20-min TT, and a 20- and 40-min sub-maximal ride. The resulting APD was calculated as the percentage change in 5-min TT from baseline, and this was compared with the TWD and TL metrics for the corresponding initial exercise bout.

Results: Average power output was different for each of the 4 initial exercise bouts (η_p^2=0.971; P<0.001), and all bouts resulted in an APD. But APD was only different when comparing maximal to sub-maximal bouts (η_p^2=0.862; P<0.001). The APD contradicted TWD and TL metrics, and was not different when comparing 5- and 20-min maximal TT's, or the 20- and 40-min sub-maximal bouts. In contrast, TL metrics were different for all training sessions (η_p^2=0.970; P<0.001).

Conclusion: An APD is found after initial exercise bouts consisting of 5- and 20-min TT's and after 20-, and 40-min of sub-maximal exercise that is not consistent with the corresponding values for TWD or TL. This discrepancy highlights important shortcomings when using TWD and TL to compare exercise bouts of different intensity and duration.

3.3 Introduction

It is well established that an acute performance decrement (APD) is observed as a consequence of prior endurance exercise ranging between 60 and 85% $\dot{V}O_2$peak lasting from 30 to 120-min (Passfield and Doust 2000; Iannetta et al. 2018; Clark et al. 2018). The extent of this APD varies considerably, however comparison between studies is problematic as a 1% change in time-to-exhaustion is not directly proportional to a 1% variation in time-trial (TT) performance (Amann et al. 2008). Nonetheless, it is apparent that the total work done (TWD) during the prior exercise does not correspond to the resulting APD. For example, an increase in work-rate of just 10 W for an exercise bout of 30-min changes APD from -37% to -64% (Iannetta et al. 2018). The discrepancy between these measurements may be important because TWD is widely used as a means of matching exercise bouts of different intensities and durations (Stewart et al. 2016; MacInnis et al. 2017). We and others have previously pointed out the theoretical limitations of this approach (Seiler et al. 2013; Nicoló et al. 2014; Nicolò and Girardi 2016; Nicolò et al. 2016b), but surprisingly a systematic evaluation of the APD that results from exercise of different intensities and durations has never been conducted.

The potential discrepancy between TWD and APD is more than just a theoretical concern. In an applied context coaches and scientists use the concept of training load (TL) to quantify their exercise programs. The increasingly popular TL concept is based on the approach of normalizing the training session for its TWD (Wallace et al. 2014a; van Erp et al. 2019a). The TL concept and its underpinning of TWD originates from the work of Banister and colleagues (Banister et al. 1975), who attempted to quantify the performance benefit (fitness gains) and APD (fatigue) that resulted from individual training sessions. The authors attempted to quantify the short-term "load" imposed by a training session using training impulses (TRIMPs) in order to predict longer-term changes in performance. Thus the long-term performance predictions are derived from TRIMPs, which reflect a short-term assessment of the fitness gains and APD arising from the individual training sessions.

Despite substantial subsequent revisions to this approach, the present-day TL metrics typically still rely on direct measurement or estimations of TWD. As discussed above, if exercise bouts similar in TWD can lead to substantially different APDs and physical effort (Seiler et al. 2013; Nicoló et al. 2014; Iannetta et al. 2018), monitoring training with current metrics of TL may provide limited insight. Considering how TL was originally conceptualized by Bannister and colleagues, the TL for a training session should be broadly reflected by its resulting APD. If this is not the case it may indicate that the basis for calculating TL merits revision. Accordingly, a systematic evaluation of acute bouts of exercise where intensity and duration are manipulated will allow us to determine if the resulting changes in TWD, TL and APD correspond.

The present study was designed to evaluate the short-term effects of manipulating the intensity and duration of an initial exercise bout on subsequent APD. Specifically, we manipulated TL by changing work rate, and duration in different exercise bouts to examine its relation to subsequent APD. We hypothesized that current metrics of TL that are dependent on calculated TWD may not be associated with the subsequent APD.

3.4 Methods

3.4.1 Participants

This study recruited (see Figure 3.1) 14 male recreational and competitive cyclists (mean ± SD: age 30 ± 8 years, height 1.80 ± 0.65 m, weight 75.9 ± 8.3 kg), with at least 3-years of experience and training >8 h/week. They gave informed consent to participate in this study which had local ethics committee approval in compliance with the Declaration of Helsinki. Participants were asked to standardize their pre-testing physical status by, consuming water and a light meal 3-h prior to each test, and refraining from caffeine and alcohol consumption, and vigorous exercise, for 24-h before each test.

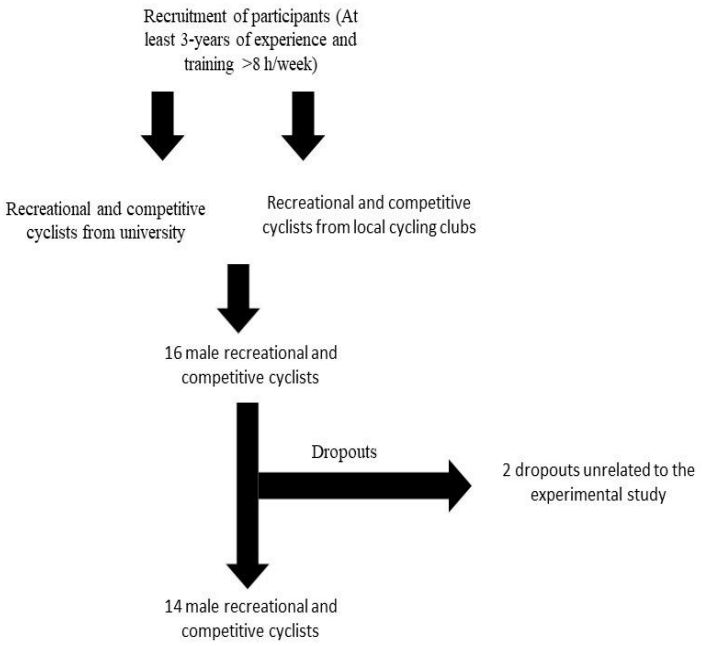

Figure 3-1 Flow chart of recruitment process and dropouts

3.4.2 Experimental Overview

A randomized, crossover within-subjects design was used for this study which required participants to visit the laboratory at the same time of day on 6 separate occasions, separated by at least 48-h over a 4-wk period (see Figure 3.1). Participants completed all exercise protocols using their personal road bicycle, which was mounted on a computer-controlled ergometer, (Cyclus-2, RBM Electronik- Automation GmbH, Leipzig, Germany).

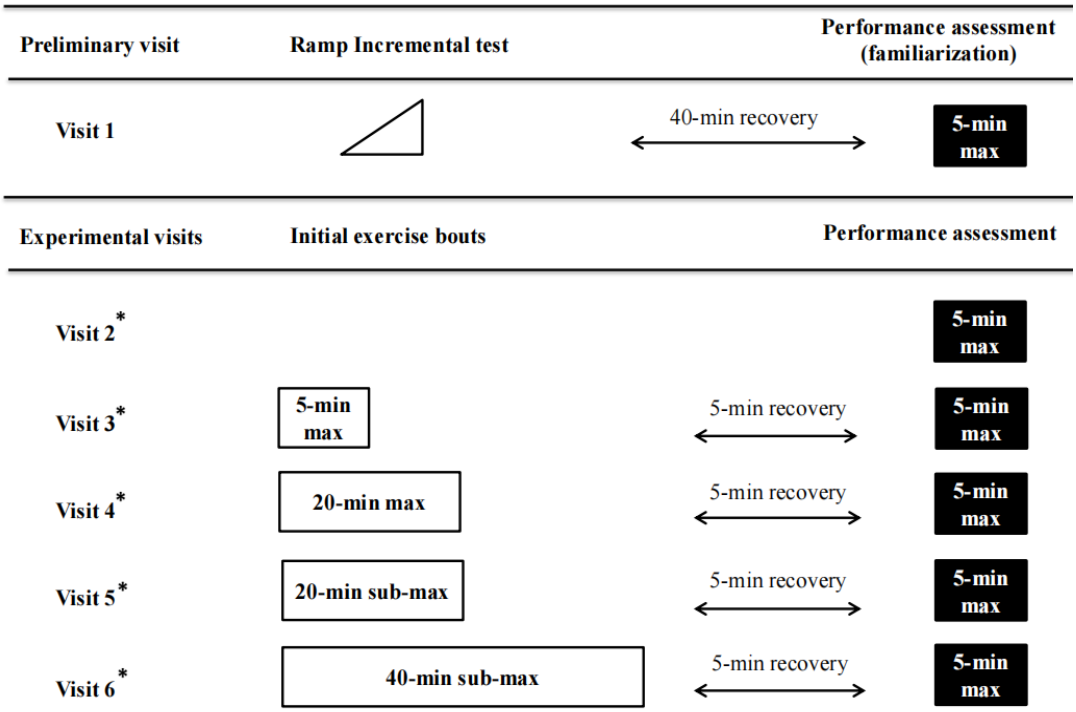

Figure 3-2 Diagram of the experimental design. Visit 1 is the preliminary testing. The baseline, 5-min max, 20-min max, 20-min sub-max and 40-min sub-max visits are the initial exercise bouts (open boxes) performed in randomized order* and the 5-min TT (filled boxes) evaluate APD.

3.4.3 Ramp incremental test

After a 5-min warm up at 100W, participants completed a ramp incremental test to volitional exhaustion. The test started at 80W and was incremented by 30W·min^{-1} after the first 2-min. Participants were instructed to maintain their preferred cadence (87 ± 7 rpm) throughout the test which was terminated when cadence fell below 60 rpm. The participants' $\dot{V}O_{2max}$ was calculated as the highest 30-s average, while peak power output (PPO) and maximum heart rate (HR_{max}) were recorded as the highest absolute values. The first (VT_1) and the second (VT_2) ventilatory thresholds were obtained according to previously described procedures (Beaver et al. 1986). After a 40-min rest participants then performed a familiarization 5-min TT according to the details below.

3.4.4 Time-trials

Participants were instructed to maintain a specific power output for the first 15-s of all TT's. For the familiarization 5-min TT, this power output was computed as a percentage of PPO achieved during the ramp test. For all the other TT's, the targeted power output reflected the average power achieved during the familiarization 5-min TT (105%). We implemented this strategy to reduce the variability in the first part of all TT's (Thomas et al. 2012), in an attempt to ensure consistent pacing strategies across all visits. After 15-s participants were un-paced and only provided with feedback on elapsed time and cadence.

3.4.5 Experimental Protocols

Prior to all laboratory visits, a standardized 15-min warm-up and 5.5-min rest was completed. Participants then completed 5 randomly ordered experimental visits, one to determine their 5-min TT performance (baseline), and another four where a 5-min TT was performed after an initial bout of

exercise. The four initial bouts of exercise were a 5-min, and a 20-min TT (5-min max, and 20-min max, respectively); and a 20-min and 40-min sub-maximal bout, (20-min sub-max, and 40-min sub-max, respectively). The work rates for the two sub-maximal initial bouts were both calculated to accumulate 3 x TWD of the familiarization 5-min TT. This means the work rate for the 40-min sub-maximal trial was always half of that for the 20-min sub-maximal trial. The choice of the four exercise bouts was made to compare: i) the effect of exercise duration on APD when the overall effort is maximal (5-min max vs. 20-min max); and ii) the effect of substantial differences in TWD on APD (5-min max vs. 20-min sub-max and 40-min sub-max). After completing the initial exercise bouts participants were allowed 5-min recovery before completing a 5-min max TT.

3.4.6 Measurements

During the ramp incremental test, $\dot{V}O_2$, respiratory frequency (f_R), and heart rate (HR) were measured breath-by-breath using an online gas analyzer (Cortex Metalyser 3B, Cortex GmbH, Lepzig, Germany) that was calibrated according to the manufacturer's instructions. A thumb-prick blood lactate (La^-) sample (Biosen C-Line analyzer, EKF diagnostics, Wales) was obtained immediately after each test. Participants' rating of perceived exertion (RPE) (Borg 1982) was collected using the 6-20 scale every 1-min. Participants had used the scale previously but were still instructed in its use on their first laboratory visit. The APD was calculated as the percentage change in average power output for the 5-min TT relative to the baseline trial (i.e. without preceding exercise). Detailed information about APD can be found in Chapter 2.8. A thumb-prick blood La^- sample (Biosen C-Line analyzer, EKF diagnostics, Wales) was collected prior any exercise (data not shown), as well as after each exercise bout and TT. Participants' $\dot{V}O_2$, and f_R were recorded breath-by-breath and averaged over 30-s for all TT's and over 1-min for the initial exercise bouts along with HR. The participants' RPE was recorded every minute for 5-min TT's and every 2-min for all other exercise. The National Aeronautics and Space Administration Task Load Index (NASA-TLX) rating scale (Hart and Staveland 1988) was completed immediately after the initial exercise bouts. The

participants' scores for each of the six subcategories (i.e. mental demand, physical demand, temporal demand, performance, effort and frustration) were added together without using a weighting factor to provide an overall score for each initial exercise bout (Hart 2012). Motivation was assessed prior to every exercise bout using a 0 to 9 analogue scale, where 0 corresponds to motivated "not at all" and 9 corresponds to motivated "a lot" (Lurquin et al. 2016). Participants were familiarized with all psychological scales during their first laboratory visit. A bTRIMP (Bannister 1991) score was calculated for all initial exercise bouts as proposed by Bannister: bTRIMP = exercise duration (minutes) · ΔHR · *y*, where $\Delta HR = (HR_{ex} - HR_{rest}) / (HR_{max} - HR_{rest})$, where *y* stands for the classical increase in blood lactate for males ($0.64e^{1.92x}$). Lucia's TRIMP (Lucía et al. 2003) was also calculated, using time spent in three pre-defined HR zones based around ventilatory thresholds (VT), measured during the ramp incremental test. Session RPE (sRPE) was calculated using the 0-10 CR-10 scale (Borg 1982) collected immediately on completion of the initial exercise bout (Christen et al. 2016). Energy expenditure (EE) of initial exercise bouts was calculated from expired gases (Peronnet, F., Massicotte 1991), and a training stress score (TSS^{TM}) (Allen and Coggan 2006) was calculated using the power output as follows: $TSS^{TM} = [(t \times NP^{TM} \times IF^{TM}) / (FTP \times 3600)] \times 100$, where t is the duration of exercise (min), NP^{TM} is the normalized power for the exercise bout and IF^{TM} is the ratio between NP and the individual's FTP (Allen and Coggan 2006). The FTP is the participant's functional threshold power and calculated from 20-min max as previously suggested (MacInnis et al. 2018).

3.4.7 Statistical analysis

An a priori statistical analysis was performed using G*Power (version 3.1.9.2; Kiel University, Kiel, Germany). Expecting a large effect size for the effect of intensity and duration manipulation on APD. A sample size of 11 was required based on $1-\beta = 0.80$ and $a = 0.05$. 16 participants were recruited to account for potential dropping out.

Statistical analysis was conducted using IBM SPSS Statistics 25 (SPSS Inc., Chicago, IL, ISA). A Shapiro-Wilk test was used to check data for normality before analysis. One-way repeated measures ANOVA with Bonferroni *post-hoc* tests were performed to compare values for the initial exercise bouts, their TL metrics, and the subsequent 5-min TTs. Two-way repeated measures ANOVA (time x initial bout) were used to compare the time course of different variables (power output, HR, RPE, f_R and $\dot{V}O_2$) during the 5-min TT's. When the sphericity assumption was violated, the Greenhouse-Geisser correction was used. Effect sizes were calculated as partial eta squared (η_p^2) and small, medium and large effects were taken as $\eta_p^2 \geq 0.01$, $\eta_p^2 \geq 0.059$, and $\eta_p^2 \geq 0.138$ respectively (Cohen 1988). Statistical significance was accepted where $P < 0.05$ was found. The results are presented as means (±SD) in the Tables, and as means (±SEM) in the Figures.

3.5 Results

3.5.1 Incremental Test

Peak $\dot{V}O_2$ and PPO from the ramp incremental test, were 3.9 ± 0.6 L min^{-1} (53 ± 8 ml kg^{-1} min^{-1}) and 396 ± 53W, respectively. Power output and $\dot{V}O_2$ corresponding to VT_1 and VT_2 were 195 ± 54 W and 277 ± 50W, and 2.4 ± 0.6 L min^{-1} and 3.1 ± 0.5 L min^{-1} respectively.

3.5.2 Initial Exercise Bouts

Due to technical problems, HR analysis was performed for 11 participants, and $\dot{V}O_2$ and f_R for 12 participants. For the 4 initial exercise bouts, average power output (P<0.001; η_p^2=0.971) and TWD (P<0.001, η_p^2=0.970), were different, except for 20-min sub-max vs. 40-min sub-max which were designed to result in the same TWD but at different power outputs (Table 3.1 and Figure 3.3). This resulted in 98±2%, 97±3% and 57±43% of the time being spent above the $\dot{V}O_2$ at VT_2 for the 5-min and 20-min max, and 20-min sub-max initial exercise bouts respectively. In contrast for the 40-min sub-max initial exercise bout 89±25% of the time was spent below the $\dot{V}O_2$ at VT_1. An effect of

initial exercise bout was found for $\dot{V}O_2$ (η_p^2=0.903; P<0.001), f_R (η_p^2=0.807; P<0.001), HR (η_p^2=0.956; P<0.001) and RPE (η_p^2=0.939; P<0.001), but not motivation (η_p^2=0.089; P=0.296).

An effect of initial exercise bout was found for the TL metrics, sRPE (η_p^2=0.963; P<0.001), TSS (η_p^2=0.936; P<0.001), TWD (η_p^2=0.97; P<0.001), bTRIMP (η_p^2=0.918; P<0.001) and LuTRIMP (η_p^2=0.921; P<0.001) and all were different for the 5-min max vs. 20-min max initial exercise bouts, except for the NASA-TLX scale (P>0.99).

3.5.3 5-min TT's

An APD was observed after all initial exercise bouts (η_p^2=0.862; P < 0.001) compared with the baseline (Table 3.2). Bonferroni *Post hoc* testing identified differences in APD following all initial bouts, except for between 5-min max vs. 20-min max (P>0.99), and 20-min sub-max vs. 40-min sub-max (P=0.813). 95% Confidence intervals (CI) were calculated following the 5-min max [15.4, 20.5], the 20-min max [16.6, 20.4], the 20-min sub-max [3.8, 9.8] and the 40-min sub-max [2.5, 5.5] initial training sessions. No difference was observed in motivation prior to the 5-min TT's, except after the 5-min max initial bout compared with baseline (P= 0.018). A significant effect of time was observed for all variables during the 5-min TT's (η_p^2=0.861; P<0.001), HR (η_p^2=0.963; P<0.001), power output (η_p^2=0.469; P<0.001), $\dot{V}O_2$ (η_p^2=0.926; P<0.001) and RPE (η_p^2=0.903; P<0.001). A significant effect of initial exercise bout was found for f_R (η_p^2=0.204; P<0.036), HR (η_p^2=0.375; P=0.001), power output (η_p^2= 0.714; P<0.001), $\dot{V}O_2$ (η_p^2=0.315; P=0.002), but not for RPE (P= 0.114). A significant interaction was observed for $\dot{V}O_2$ (η_p^2=0.338; P<0.001), RPE (η_p^2=0.171; P=0.001), f_R (η_p^2=0.32; P<0.001), power output (η_p^2=0.274; P<0.001) and HR (η_p^2=0.135; P=0.024), see Figures3. 4 and 3.5 .

Table 3-1 Physiological, perceptual and mechanical values for the initial exercise bouts.

	Initial Exercise Bouts			
Measurement	**5-min max**	**20-min max**	**20-min submax**	**40-min submax**
Power output (W)	$337 \pm 49^{\dagger\beta}$	$273 \pm 46^{\dagger}$	$236 \pm 34^{\#}$	118 ± 17
$\dot{V}O_2$ ($L \cdot min^{-1}$)	$3.57 \pm 0.5^{\dagger}$	$3.52 \pm 0.56^{\#}$	$3.15 \pm 0.41^{\#}$	1.91 ± 0.22
HR ($beats \cdot min^{-1}$)	$170 \pm 10^{\dagger}$	$171 \pm 9^{\dagger}$	$154 \pm 12^{\#}$	114 ± 10
f_R ($breaths \cdot min^{-1}$)	$50 \pm 8^{\dagger}$	$47 \pm 9^{\dagger}$	$39 \pm 5^{\#}$	28 ± 3
Rating of perceived exertion	$17.9 \pm 0.9^{\dagger\beta}$	$17.1 \pm 1^{\dagger}$	$14.1 \pm 1^{\#}$	10.4 ± 1.2
La^-_{end} ($mmol \cdot L^{-1}$)	$12.7 \pm 2.4^{\dagger}$	$11.4 \pm 2.9^{\dagger}$	$4.3 \pm 1.9^{\#}$	1.7 ± 0.6
Motivation	7.5 ± 0.7	7.5 ± 1.2	8.1 ± 0.7	7.6 ± 1.2

Values are mean ± SD. $\dot{V}O2$= oxygen uptake, HR=heart rate, fR=respiratory frequency, RPEend = Final rating of perceived exertion, La-end = Post exercise blood lactate. Statistical significance is as follows; †$P<0.05$ from 20- and 40-minute sub-max, max, #$P< 0.05$ from 40-min sub-max, β $P<0.05$ from 20-min max.

Table 3-2 Physiological, perceptual, and mechanical values for the baseline 5-minute TT and the 5-minute TT that followed each initial exercise bout

		5-min TT following initial exercise bouts			
Measurement	**Baseline**	**5-min max**	**20-min max**	**20-min sub-max**	**40-min sub-max**
Power Output (W)	351 ± 47*,†	289 ± 46†	287 ± 44†	327 ± 49	337 ± 45
Cadence (rpm)	87 ± 9	86 ± 6	87 ± 6	91 ± 8	89 ± 6
RPE_{end}	19.7 ± 0.3	19.7 ± 0.6	19.7 ± 0.4	19.7 ± 0.4	19.7 ± 0.4
La^{-}_{end} (mmol·L^{-1})	15.2 ± 2.5	14.2 ± 3.0	12.3 ± 3.2	12.7 ± 2.4	12.6 ± 2.6
Motivation	8.5 ± 0.6‡	7.5 ± 0.7	7.5 ± 1.2	7.5 ± 1.1	7.9 ± 0.9

Values are mean ± SD. RPEend=Final Rating of Perceived Exertion, La-end=Post TT blood lactate. Motivation was recorded immediately before the 5-min TT. Statistical significance is indicated as follows; †$P < 0.05$ from 20 and 40-min sub-max, *$P< 0.05$ from 5-min and 20-min max, ‡$P < 0.05$ from 5-min max.

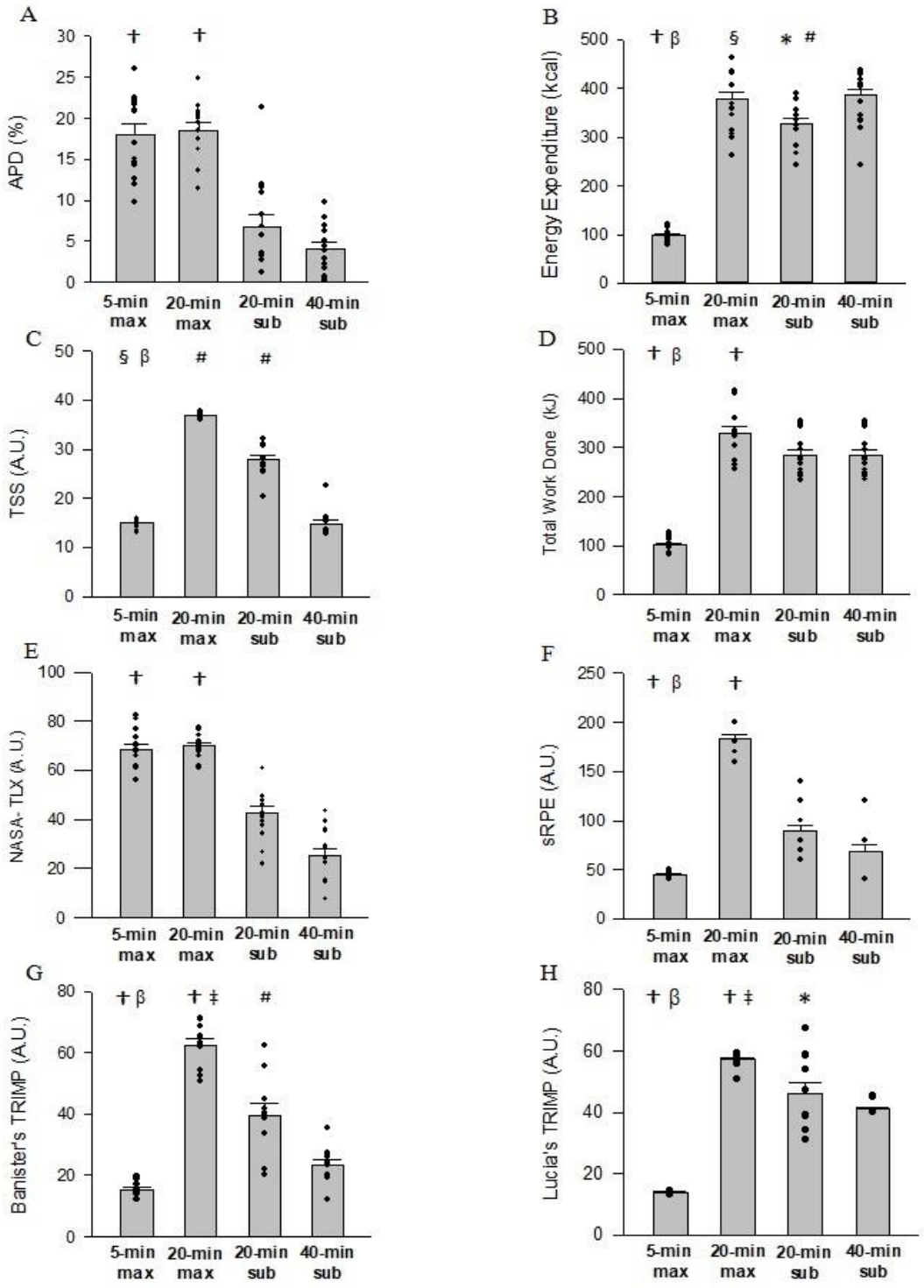

Figure 3-3 Training load metrics for the initial exercise bouts: APD - Figure A, EE - Figure B, TSSTM - Figure C, TWD - Figure D, NASA-TLX - Figure E, sRPE – Figure F, bTRIMP - Figure G, and LuTRIMP – Figure H. See text for abbreviations. Values are mean ± SEM. Significant pairwise comparisons are shown as follows; *P< 0.05 from 5-min and 20-min max, ‡P < 0.05 from 5-min max, §P < 0.05 from 20-min sub-max, †P < 0.05 from 20-min and 40-min sub-max, #P < 0.05 from 40-min sub-max, and βP< 0.05 from 20-min max.

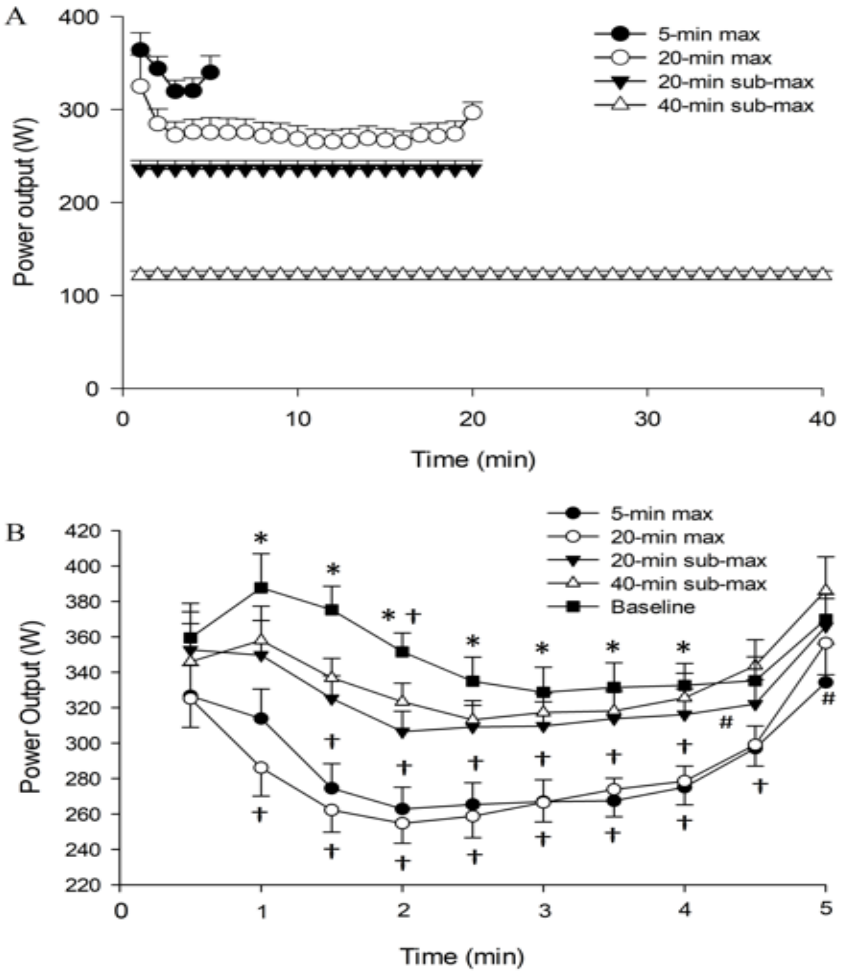

Figure 3-4 Power output for the initial exercise bouts (A) and the subsequent 5-min TT (B). Values are mean ± SEM. Significant condition x time interaction, †P < 0.05 from 20-min and 40-min sub-max, *P< 0.05 from visit 5-min and 20-min max, #P < 0.05 from 40-min sub-max.

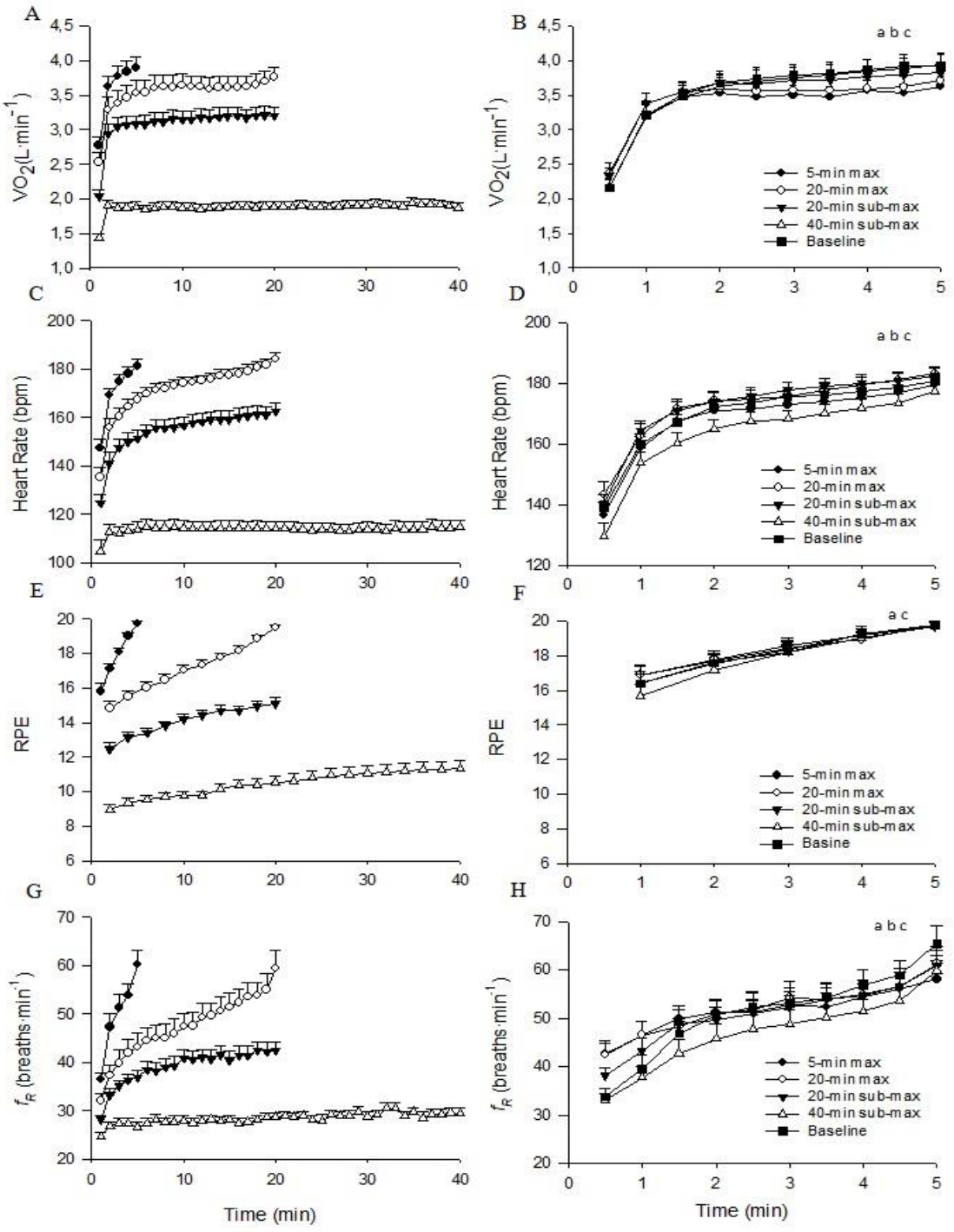

Figure 3-5 The left-hand panel shows responses to the initial exercise bouts and the right-hand panel shows responses to the subsequent 5-min TT's. Individual plots are for V̇O2 (A and B), HR (C and D), RPE (E and F), and fR (G and H). See text for abbreviations. The initial exercise bouts are indicated as follows; baseline (filled squares), 5-min max (filled circles), 20-min max (open circles), 20-min sub-max (filled triangles), 40-min sub-max (open triangles). Values are mean ± SEM.

Statistical significance is indicated as follows; (a) interaction of initial exercise bout by time ($P < 0.05$), (b) main effect of initial exercise bout ($P < 0.05$), (c) main effect of time ($P < 0.05$).

3.6 Discussion

The present study examined the APD that occurs following an initial bout of exercise, and how manipulating TWD and TL by changing exercise intensity and duration affected subsequent performance. The main findings of the present study were that an APD was observed after maximal 5-min, and 20-min TT's, and after sub-maximal 20-min and 40-min initial exercise bouts when compared to a baseline 5-min TT. There were significant differences in the APD that were found following the initial exercise bouts, but some of these were not reflected by corresponding differences in TWD or the calculated TL metrics. Another notable finding was that APD was similar following the 5-min and 20-min maximal initial exercise bouts despite their substantial differences in TWD and TL. These findings show that the TWD and related TL metrics (sRPE, bTRIMP, LuTRIMP, TSS) do not reflect the APD observed. Consistent with the present study a recent case report found that total elevation gain, rather than TWD or TSS, predicted performance on key mountain climbs for a Grand Tour Champion (Van Erp et al. 2020). In the present study, the initial bout with the shortest duration (5-min max) had a lower TWD and TL metric scores, but a greater APD when compared with the longer sub-maximal bouts. These results underline the fact that TWD, and TL metrics do not reflect the effect that dissimilar training sessions have on acute changes in performance. However, where more similar sessions are compared, i.e. the 20-min and 40-min sessions, a more consistent pattern in APD and TL metrics can be seen.

The importance of these findings is difficult to overstate because TWD is widely and increasingly used as the means for matching exercise bouts of different intensity and duration (Seiler et al. 2013; Stewart et al. 2016; MacInnis et al. 2017). The present study findings suggests that this practice of matching TWD when comparing different training sessions or programmes is not always appropriate, and supports several researchers who have critiqued this practice (Seiler et al. 2013;

Nicoló et al. 2014; Nicolò and Girardi 2016). Previous studies have reported that different sessions with the same TWD can result in substantial differences in physical effort (Nicoló et al. 2014; Nicolò and Girardi 2016), and the present study uses the concept of an APD, to demonstrate that an initial exercise bout with a lower TWD and TL can result in a greater APD. Furthermore, the time spent in a given exercise-intensity domain does not appear to be well associated with the resulting APD. The maximal bouts showed large differences in the total time spent above VT_2 (i.e. 4.9 min vs. 19.3 min), but a similar APD; while the sub-maximal bouts were most of the time situated in different exercise-intensity domains (i.e. 11.3 min above VT_2 vs. 35.5 min below VT_1), but showed no significant differences in APD.

In the present study APD was quantified as the percentage change in 5-min TT performance from a baseline (with no preceding exercise). For the 5-min TT we were careful to familiarize participants, and to use the same protocol that has been shown to be sensitive to changes following prolonged moderate-intensity exercise (Passfield and Doust 2000). Self-paced trials performed after different initial exercise bouts can be influenced by uncontrolled variations in pacing and motivation, and so these factors were carefully considered for the present study. We attempted to standardize the influence of pacing by setting the work rate for the first 15-s of every 5-min TT. Figure 3.4 suggests this approach was successful as it shows similar pacing strategies in the power output profiles of each of the 5-min TT's. Motivation was similar before the four initial exercise bouts, but was lower after the 5-min TT initial exercise bout. Despite the difference in reported motivation for this one trial, similar values for RPE and f_R were observed across all 5-min TT's, suggesting this did not prevent participants from producing their maximal effort. Furthermore, a similar physical effort was seen across all 5-min TT's performed to measure APD, with the exception of the one following 40-min sub-max, where lower RPE and f_R were values were observed. Nevertheless, the time course of power output, RPE and f_R suggests that the 5-min TT's were performed to a generally high level of consistency.

As discussed previously, Banister et al. (1975) anticipated an APD, which they called "fatigue", in response to individual training sessions, although they did not try to measure it directly, rather they modelled its long-term effects. In the present study we examine only the short-term implications by systematically manipulating exercise intensity and duration in order to change TL and evaluate its corresponding effects on APD. In agreement with previous studies TWD and TL were strongly associated (Wallace et al. 2014a; van Erp et al. 2019a), but correspondence between these variables and APD was not found. One possible reason for this finding is that the APD is itself misleading. This seems unlikely as a decrement in performance following an initial exercise bout has been variously reported by different laboratories, and was an implicit part of Banister's original model. Moreover, as reported here APD has been shown to be sensitive to variations in exercise intensity and duration (Passfield and Doust 2000; Iannetta et al. 2018; Clark et al. 2018), and to specific experimental interventions (Amann 2011). In contrast to APD however, theoretical concerns with the use of TWD and TL have been highlighted by ourselves and others previously (Seiler et al. 2013; Nicoló et al. 2014; Nicolò and Girardi 2016). It has been suggested that equating exercise bouts of different intensity and duration on the basis of their TWD or related TL metrics is flawed because multiplying exercise intensity and duration does not yield comparable levels of stress or overall effort. This issue is likely compounded if a linear exercise intensity weighting is used as with Lucias's TRIMP, but even when intensity is weighted non-linearly, as with bTRIMP or TSS, good agreement with APD was not found in the present study. Indeed, when the aim is to equate different exercise bouts for overall effort, it appears that a greater TWD is needed for longer, lower intensity exercise bouts, compared with shorter higher intensity ones (Nicolò et al. 2016b, a). Likewise, continuous and intermittent exercise bouts with a similar overall effort result in different values of TWD (Nicoló et al. 2014). Based on the observation that the 5- and 20-min TT's elicited a similar APD, but very different TWD and TL values, we suggest that the confounding variable in these calculations is the exercise duration. Future research is required to determine whether re-weighting the contribution of

exercise duration in the calculation of TL may result in values that are more consistent with the observed APD.

Interestingly, the NASA-TLX was found to be correspond with APD in contrast to the more commonplace TL metrics. Similar values of NASA-TLX were found after 5-min max and 20-min max and were consistent with the observed APD. NASA-TLX is a common measure of overall workload experienced when performing tasks, and takes into account the responses from six different subcategories (Hart 2012). However, its use is not common in the field of sport. Our findings may indicate that NASA-TLX could be a useful instrument to consider alongside the more widely used measures of TL. Further studies are required to evaluate the association between NASA-TLX and APD, and the potential benefits of adopting this instrument for monitoring training.

3.6.1 Limitations and Future Research

We suggest that the findings of the present study should not be seen as argument to replace TWD and related TL metrics with APD. Rather these findings should serve to highlight the limitations of the metrics employed and stimulate the development of new approaches. Banister et al. (1975) introduced the concept of TRIMP that today still underpins the TL concept, despite its acknowledged inherent weaknesses (Busso and Thomas 2006). One of the challenges of Banister's approach to systems modelling was that it required repeated (daily) performance trials. In this regard APD is little different and therefore would be more appropriately used as a way to develop new metrics rather than adopted as a method in itself. We suggest that future research explores a wider range of initial exercise bout durations and intensities to more completely reflect those employed by athletes in their training. It will also be useful to determine whether it is possible to predict the APD that arises from pre-specified exercise bouts and how it is affected by differences in athletes' fitness and exercise modality. A further limitation of this study is that we cannot determine the mechanisms

that underpin the APD. Banister when introducing the TRIMP proposed that training sessions created a fatigue effect - reflected as an APD. It is clear that there is much that is still unknown about how fatigue may have limited performance in the present study and therefore we have avoided use of this term and preferred APD instead (Enoka and Duchateau 2016). Lastly, the relatively small sample size of our study does not exclude the possibility that subtle differences between the APDs between the 20-min sub-max and the 40-min sub-max might be revealed with a larger sample size. The confidence intervals observed appear to support this suggestion.

3.6.2 *Practical applications*

Our findings show that the popular methods of TWD and TL that are used for comparing different exercise bouts do not reflect the APD observed, particularly where short, intense exercise is involved. The large effect sizes observed reinforce this proposition. This finding has important implications for training prescription, monitoring and periodization. As stated above, we are not suggesting APD as an alternative measure of TL, but rather as a theoretical contribution to the development of new TL metrics. However, if future research were to find that it is possible to model the APD arising from specified exercise bouts, this may allow the use of APD to extend from the theoretical to more practical realms. At present, our findings indicate that TWD and related TL metrics, overestimate the stress of longer duration exercise bouts (by using duration as a multiplier). Where colleagues, coaches and athletes use TWD or TL to evaluate training sessions of different exercise intensity and duration they should be aware of the bias caused by using time as multiplier. We also recommend that further consideration be given to the potential benefit of using the NASA-TLX scale as an instrument for evaluating TL.

3.6.3 Conclusion

A significant APD measured as a change in 5-min TT is observed after initial exercise bouts of varying exercise intensity and duration. However, the APD found was not consistent with TWD in the preceding exercise bout or its calculated TL. This finding supports previously expressed theoretical concerns and suggests that the practice of matching training sessions or programs by TWD or TL may be flawed. New methods for calculating TL are consequently required. Finally, this study has identified the NASA-TLX instrument as an instrument worthy of further evaluation for monitoring TL.

Chapter 4: Continuous Versus Intermittent Running: Acute Performance Decrement and Training Load

4.1 Aims and hypothesis

The aim of Chapter 4 is to extend the findings from Chapter 3 with the use of different training sessions in a different exercise modality. Findings from the cycling study showed that an APD occurs after different cycling training sessions, independently of how intense or lengthy they were. Nevertheless, whether an APD will be found as assessed after running training sessions was unknown. Therefore, Chapter 4 was designed to demonstrate that running training sessions of different characteristics will result in an APD. Another aim for Chapter 4 is to investigate whether APD and TL will come in agreement when training is prescribed in either a continuous or an intermittent manner. Training sessions were either continuous (10CON, 25CON) or intermittent (10INT, 25INT) and consisted of different exercise durations (10- and 25-min). Similarly to the cycling study, the running sessions used will be consisted of different intensities and durations to examine their effects on how TL metrics quantify training compared to APD responses. Based on the findings from Chapter 3, it was hypothesized that an APD will be found after the running training sessions. A dissociation between the APD responses and TL metrics was anticipated.

4.2 Abstract

Purpose: To examine the effect of continuous (CON) and intermittent (INT) running training sessions of different durations and intensities on subsequent performance and calculated training load (TL).

Methods: Runners (n=11) performed a 1500m time-trial (1500m TT) as a baseline, and after completing 4 different running training sessions. The training sessions were performed in a randomized order, and were either maximal for 10 minutes (10CON, 10INT) or sub-maximal for 25 minutes (25CON, 25INT). An acute performance decrement (APD) was calculated as the percentage change in 1500m TT speed measured after training compared with baseline. The pattern of APD response was compared to that for several TL metrics (bTRIMP, eTRIMP, iTRIMP, rTSS and sRPE) for the respective training sessions.

Results: Average speed (P<0.001; η_p^2 =0.924) was different for each of the initial training sessions which all resulted in a significant APD. This APD was similar when compared across the sessions, except for a greater APD found after 10INT vs 25CON (P=0.02). In contrast, most TL metrics were different and showed the opposite response to APD, being higher for CON vs. INT, and lower for 10 vs. 25 min sessions (P<0.001; η_p^2 =0.563).

Conclusion: An APD was observed consistently after running training sessions but it was not consistent with most of the calculated TL metrics. The lack of agreement found between APD and TL suggests that current methods for quantifying TL are flawed when used to compare CON and INT running training sessions of different durations and intensities.

4.3 Introduction

Training is frequently quantified by its load (TL), a metric that is a function of exercise intensity and duration, and assumed to reflect the stress of a given session or sessions (Bourdon et al. 2017). In running (Manzi et al. 2009b; Wallace et al. 2014a) and team-sports (Impellizzeri et al. 2004; Akubat et al. 2012; Manzi et al. 2013) research has tended to focus upon the influence of TL on the development of fitness over the longer term involving multiple training sessions. Interest in evaluating TL has grown in recent years, notably driven by interest in injury prevention (Bourdon et al. 2017) and wearable technology development (Emig and Peltonen 2020). For example, data from wearables enabled a recent study where the TLs for 1.6 million training sessions were quantified in order to gain insight into optimal training strategies (Emig and Peltonen 2020). However, some studies have shown that the relationship between TL and performance is not straight forward (Taylor et al. 2017; Van Erp et al. 2020; Vermeire et al. 2021). This may be because relatively little attention has been given to validating how TL is derived from the intensity and duration of individual training sessions (Kesisoglou et al. 2021b), before TL is summed across multiple sessions.

Most TL metrics are based on the concept suggested initially by Banister et al. (1975) that training results in short and long term changes in performance that are dictated by the accumulation of both fatigue and fitness. Banister et al. (1975) derived the training impulse (TRIMP), a function of training intensity and duration, as a means of quantifying training sessions. The TRIMP was used as a TL metric to model how training altered athletes' fatigue and fitness status, and thus changes in their performance. In order to validate the underpinning concepts of TL we measured the acute performance decrement (APD) that results from a single cycling training session (Kesisoglou et al. 2021b). Given how TL is conceptualized, the size of a session's TL should be reflected in a corresponding APD. However, no agreement between the APD and typical TL metrics for that session were found. The largest difference in TL was found when comparing a short intense, maximal 5-min session with a longer 20-min one, yet the APD after both sessions was similar. These findings

indicate that the current basis for calculating TL for cycling sessions is flawed and this could explain why the relationship between TL and performance is not straightforward. However, whether an APD occurs and is consistent with changes in TL following other modes of exercise, such as running, remains to be determined.

The prescription of intermittent (INT) training by perceived exertion rather than using relative exercise intensity is becoming increasingly popular due its simplicity, and its potentially greater benefits compared with continuous exercise (CON) (Buchheit and Laursen 2013a). The process of INT training prescription varies according to the work interval, its exertion rating, duration, and recovery prescribed (Buchheit and Laursen 2013a). Runners and team sport players in particular, often employ INT methods in their training. The CON and INT sessions can be compared by matching for work (Gibala et al. 2012) or effort and duration (iso-effort, iso-time) (Seiler et al. 2013; Nicoló et al. 2014). When two training sessions (CON vs. INT) were compared after matching for effort and duration, substantial differences in total work, oxygen uptake and blood lactate were reported (Seiler et al. 2013; Nicoló et al. 2014). A greater understanding of the comparative effects of iso-effort, iso-time CON and INT sessions may be provided by evaluating their TL and resulting APD but to our knowledge these have not been assessed together.

The present study was designed to evaluate the influence of iso-effort, iso-time CON and INT running training sessions of different durations on several TL metrics and subsequent APD. Specifically, we manipulated TL in different CON and INT sessions to evaluate whether the APD reflected these changes. Based on our previous research, we hypothesized that an APD would be observed after running sessions, but that differences in APD between sessions would contradict those for the respective TL metrics.

4.4 Methods

4.4.1 Participants

The participants (see Figure. 4) in this study were 10 male and 1 female (mean ± SEM; age: 25 ± 2 years, height: 1.77 ± 0.1 m, weight: 67.3 ± 2.5 kg), well-trained runners. Participants had at least 3 years' experience of competing in events ranging from 800m to half-marathon; they trained >8 hours/ week and had personal best times for 1500m of 266 ± 8 s. All participants gave written informed consent to take part in this study which was approved by the ethics committee of the University of Kent in compliance with the Declaration of Helsinki.

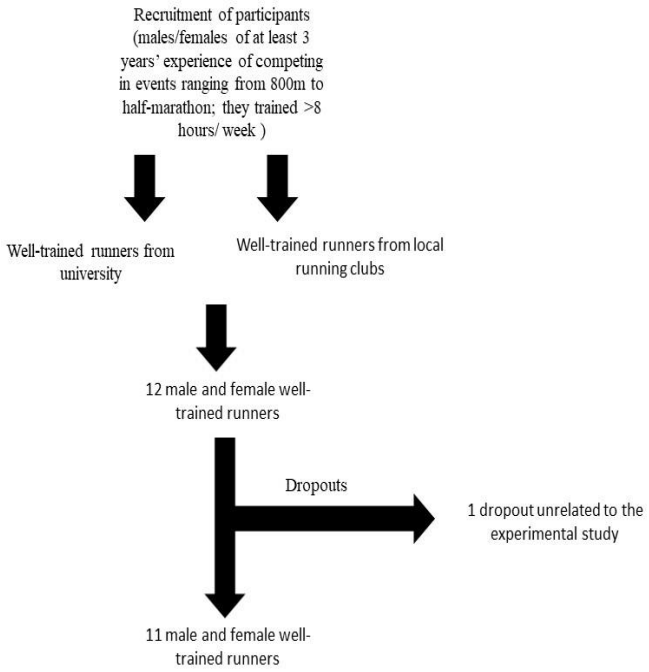

Figure 4-1 Flow chart of recruitment process and dropouts

4.4.2 Experimental Overview

Using a randomized, crossover, within-subjects design, over a 4 week period, participants completed 5 test-sessions following preliminary laboratory testing (see Figure 4.2). All sessions were separated by at least 48 hours. The 5 randomized test-sessions consisted of a baseline 1500m time-trial (1500m TT), and four experimental training sessions followed by a 1500m TT. All training sessions were completed at the same time of the day with participants fully rested and hydrated, having consumed a light meal 3 h before and having refrained from alcohol and caffeine consumption and vigorous exercise for 2 hours previously.

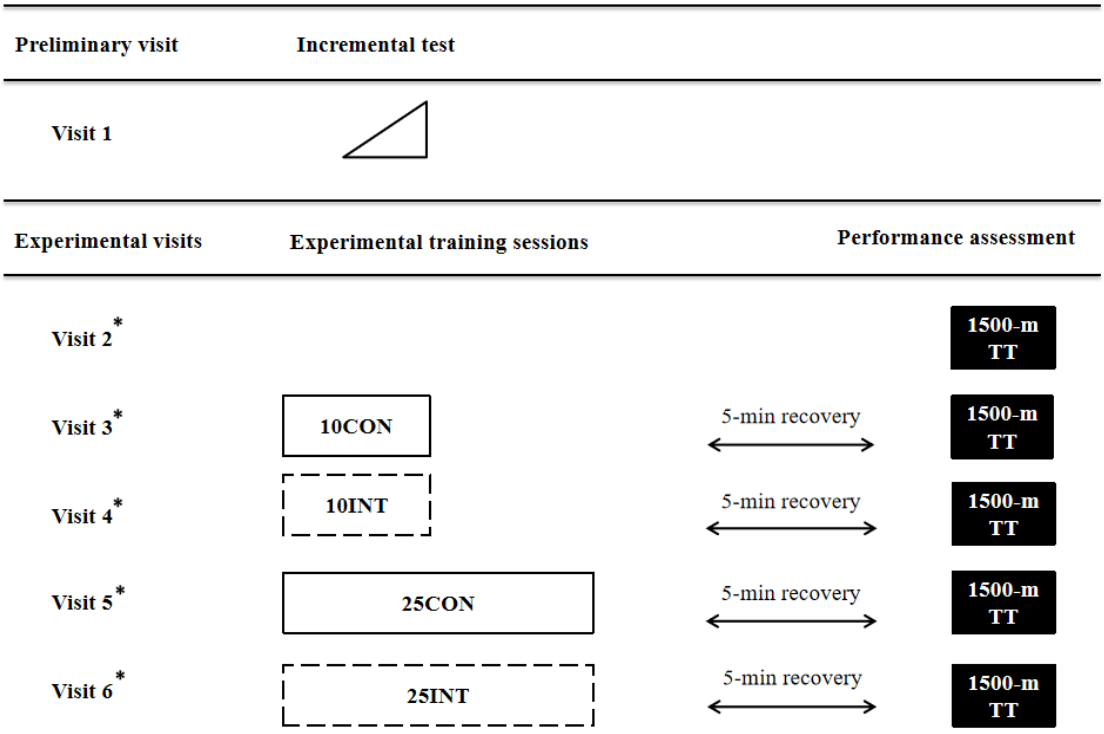

Figure 4-2 Diagram of the experimental design. Visit 1 is for preliminary testing. The baseline, 10CON, 10INT, 25CON and 25INT visits are the experimental training sessions (open boxes) performed in randomized order* and the 1500 TT's (filled boxes) evaluate APD.

4.4.3 Preliminary Laboratory Testing

Participants performed a two-phase incremental test on an indoor treadmill (Woodway ELG, Woodway GmbH, Germany) to determine individual blood lactate concentration (La^-) profiles, and then maximum heart rate (HR_{max}), peak treadmill velocity (PTV), and associated $\dot{V}O_{2max}$ (Manzi et al. 2009b). The first phase consisted of 4 min stages, interspersed with 1 min rest, starting at an initial running speed of 10 $km \cdot h^{-1}$ and increasing by 1 $km \cdot h^{-1}$ each stage. The second test phase began when blood lactate concentrations (La^-) reached 4 mmol/l, whereupon running speed was increased continuously by 0.5 $km \cdot h^{-1}$ every 30-seconds until exhaustion. Thumb-prick La^- samples were collected and analyzed (Biosen C-Line analyzer, EKF diagnostics, Wales) during the 1 min rest periods in phase one, and immediately on completing phase two. Expired gases were measured using a breath-by-breath open-circuit indirect calorimetry system (Oxycon Pro, Erich Jaeger, Germany), calibrated according to the manufacturer's instructions prior to each use. The participants' HR_{max} and PTV were recorded as the highest observed values and $\dot{V}O_{2max}$ was calculated as the highest 30-second average.

4.4.4 Experimental Training Sessions

The 5 experimental sessions were conducted on an outdoor synthetic athletics track. The 4 experimental training sessions consisting of CON or INT running for 10 or 25 min (i.e. 10CON, 10INT, 25CON, 25INT) as shown in Figure 4.2. The CON and INT training sessions were matched for overall effort using the CR-10 scale (Borg 1982), and prescribed as maximal (10/10) for both 10 min sessions, and submaximal (6/10) for both 25 min sessions. The two INT training sessions consisted of 150m efforts performed once per minute, with the passive recovery lasting for the remainder of each minute. Participants self-paced their efforts, and no feedback was provided during or between training sessions. On completing the experimental training sessions participants had 5 min of passive recovery, before performing a 1500m TT.

Timing during all training sessions and 1500m TT's time was performed manually with a stopwatch. Speed and heart rate were measured second-by-second using a GPS wristwatch and a Bluetooth connected foot pod, (Polar Oy, Polar Electro, Kempele, Finland). The wind speed was assessed with a digital anemometer (Protmex, MS6252A) for all experimental sessions which were only performed if it was below 2 $m \cdot sec^{-1}$. Prior to all experimental sessions, HR_{rest} was measured for 3 min, seated in a quiet environment before participants undertook a standardized self-paced 3-lap warm-up and dynamic stretching routine.

4.4.5 Measurements

The APD for each experimental training session was calculated as the percentage change in 1500m TT speed, i.e. APD = (P1-P2)/P1•100, where P1 is the baseline 1500m TT speed and P2 is the 1500m TT speed recorded after the training session. Note that the APD is calculated as a positive percentage change to facilitate comparison with the corresponding TL metrics for the session. More information about APD can be found in Chapter 2.8.

The participants' RPE (scale 6-20) was recorded at 300m, 700m, 1100m, and 1500m during each 1500m TT and every 2 minutes throughout the experimental training sessions (Borg 1982). The National Aeronautics and Space Administration Task Load Index (NASA-TLX) rating scale (Hart and Staveland 1988) was used to assess TL immediately on completion of the experimental training sessions by summating the scores from its 6 sub-categories (Hart 2012). Motivation was assessed prior to experimental training sessions and 1500m TT's (Lurquin et al. 2016). Participants were familiarized with all scales during their laboratory visit.

The TL metrics were calculated using 3 different TRIMP formulae (bTRIMP, iTRIMP and eTRIMP), session RPE (sRPE) (Foster et al. 2001) and a running training stress score (rTSS). The bTRIMP (Bannister 1991) was calculated from HR using a weighting factor according to the formula:

bTRIMP = duration training (minutes) · ΔHR · y, where ΔHR = (HR_{ex} - HR_{rest}) / (HR_{max} - HR_{rest}), where *y* represents a classic increase in blood lactate of $0.64e^{1.92x}$ for males and $0.84e^{1.67x}$ for females. The iTRIMP (Manzi et al. 2009b) was calculated using individualized weighting factors (y_i) for each participant using their phase 1 laboratory test data (Manzi et al. 2009b). The eTRIMP (Edwards 1993) was calculated from the time spent in five pre-defined HR zones (Bannister 1991) multiplied by arbitrary zone-based weighting factors. The sRPE was collected immediately after the training session using the CR-10 scale (Borg 1982) and multiplied by training duration ($sRPE_r$) (Foster et al. 2001) and excluding recovery ($sRPE_{wr}$) (Wallace et al. 2014b). The running training stress score (rTSS) was calculated as previously described (Skiba 2006; McGregor et al. 2009) with functional threshold pace being assumed to be 88% of 10CON (van Djik and van Megen 2017). The TL metrics were also normalized by dividing them by the training session duration in order to remove the influence of duration and evaluate intensity separately.

4.4.6 Statistical Analysis

An a priori statistical analysis was performed using G*Power (version 3.1.9.2; Kiel University, Kiel, Germany). Expecting a large effect size for the effect of exercise intensity manipulation on APD. A sample size of 10 was required based on 1-β = 0.80 and a = 0.05. 12 participants were recruited to account for potential dropping out.

Statistical analysis was conducted using IBM SPSS Statistics 25 (SPSS Inc., Chicago, IL, ISA). Following a Kolmogorov-Smirnov test to check data for normality, where appropriate a one-way repeated measures ANOVA with Bonferroni post-hoc test was performed to compare mean values for the experimental training sessions, including the TL metrics, and the APD. Some TL metrics were not normally distributed (sRPE, eTRIMP, iTRIMP) and these were evaluated with the non-parametric Friedman and Bonferroni post-hoc test instead. Two-way repeated measures ANOVA (session x

time) was used to compare the responses of variables (speed, HR, and RPE) during both the experimental training sessions and the 1500m TT's. Where the sphericity assumption was violated, the Greenhouse-Geisser correction was used. Effect sizes were calculated as partial eta squared (η_p^2) and small, medium, and large effects were taken as $\eta_p 2 \geq 0.01$, $\eta_p^2 \geq 0.059$, and $\eta_p 2 \geq 0.138$ respectively (Cohen 1988). Statistical significance was accepted where $P < 0.05$ was found. All results are presented as mean ± SEM.

4.5 Results

4.5.1 Incremental Test

$\dot{V}O_{2peak}$ and PTV from the incremental test were 4.63 ± 0.22 L min^{-1} (68.8 ± 2.7 ml kg^{-1} min^{-1}) and 19.7 ± 0.5 km·h^{-1}, whilst HR_{max} and La^- were 188 ± 3 bpm and 12.6 ± 1.3 mmol/L respectively.

4.5.2 Experimental Training Sessions

Total distance covered in each session was 2914 ± 77, 1500, 6316 ± 193 and 3750 m for 10CON, 10INT, 25CON and 25INT, respectively. Technical problems meant HR analysis not performed for 1 participant. When comparing 10CON and 10INT a main effect of time was found for HR and RPE (P<0.001; η_p^2 >0.898) but not speed (P=0.306; η_p^2 =0.107). A main effect for session was found for speed and HR (P<0.001; η_p^2 >0.898) but not RPE (P=0.499; η_p^2 =0.047). A significant interaction was found only for HR (P<0.001; η_p^2 =0.463). When comparing 25CON and 25INT, a main effect of time was observed for all variables (P<0.001; η_p^2 >0.829), and a main effect of session was found for speed (P<0.001; η_p^2 =0.87), but not HR or RPE. A significant interaction was found for speed and HR (P<0.042; η_p^2 >0.248), but not for RPE (see Figure 4.3).

Table 4.1 shows average speed was different between all the 4 training sessions (P<0.001; η_p^2 =0.924), %PTV (P<0.001; η_p^2 =0.92), as was HR_{mean} (P<0.001; η_p^2 =0.657), and RPE_{end} (P<0.001;

η_p^2 =0.917), but motivation was similar (P=0.585; η_p^2 =0.062). A difference between sessions was found for all TL metrics, $sRPE_r$ (P<0.001), $sRPE_{wr}$ (P<0.001), rTSS (P<0.001; η_p^2 =0.882), bTRIMP (P<0.001; η_p^2 =0.781), eTRIMP (P<0.001), iTRIMP (P=0.001), and NASA-TLX (P<0.001; η_p^2 =0.74) and are shown in Figure 4.4. Bonferroni post-hoc testing revealed between-session differences for TL metrics as shown in Figures 4.4 and 4.5. Differences in the opposite direction were found when the TL metrics were normalized, sRPE (P<0.001), bTRIMP (P<0.001; η_p^2 =0.662), eTRIMP (P<0.001), iTRIMP (P<0.001) and rTSS (P<0.001; η_p^2 =0.799) and are shown in Figure 4.5. Figure 4.6 presents speed and HR for the 4 training sessions (A and C), and example minute periods (B and D).

4.5.3 The 1500m TT's

An APD was observed after all experimental training sessions (P<0.001; η_p^2 =0.454; β = 0.984). Bonferroni post-hoc testing found no differences in the size of APD except for 10INT vs. 25CON (P = 0.02; see Figure 4.4). For the APDs after the training sessions, the 95% Confidence intervals (CI) were calculated for the 10CON [4.9, 8.9], the 10INT [5.3, 10.1], the 25CON [2.4, 4.8] and 25INT [2.9, 5.7]. Table 4.2 presents the measurements during the 1500m TT's. A main effect of time was observed for all variables measured during the 1500m TT's; speed (P=0.01, η_p^2 =0.429), HR (P<0.001, η_p^2 =0.956), and RPE (P=0.001, η_p^2 =0.869). A main effect of session was found for speed (P<0.001, η_p^2 =0.662), and RPE (P=0.05, η_p^2 =0.227) but not for HR (P=0.2, η_p^2 =0.148). A significant interaction was observed for speed (P=0.03, η_p^2 =0.229), but not for HR (P=0.16) or RPE (P=0.601), see Figure 4.3. No difference was observed in motivation before the 1500m TT's, except for baseline vs. 10CON (P=0.01) and 25CON (P=0.02).

Table 4-1 Physiological and perceptual responses to the training sessions.

	10CON	**10INT**	**25CON**	**25INT**
Speed (km/h)	17.5 ± 1.5$^{\dagger, *}$	22.1 ± 1.8*	15.2 ± 1.6§	18.8 ± 1.7
%PTV	89 ± 3.8$^{\dagger, *}$	112 ± 5.6*	77 ± 5.1§	95 ± 7.1
HR_{mean} (bpm)	173 ± 8$^{\dagger, *}$	162 ± 10§	161 ± 12	154 ± 13
Total Distance (m)	2914 ± 255	1500	6316 ± 639	3750
RPE_{mean}	17 ± 1*	17 ± 1*	13 ± 1	13 ± 1
RPE_{peak}	19 ± 1*	19 ± 1*	15 ± 1	15 ± 1
Motivation	7.1 ± 1.1	7.5 ± 1.2	7.6 ± 0.8	7.5 ± 0.8

Values are mean ± SD. Speed= Average speed, %PTV= Average speed expressed as a percentage of peak treadmill velocity, HR= Average heart rate, RPE_{mean} = Average rating of perceived exertion, PRE_{peak}= Final rating of perceived exertion. Statistical significance: $^{*}P<0.05$ from 25CON and 25INT, $^{\dagger}P<0.05$ from 10INT, $^{\S}P< 0.05$ from 25INT.

Table 4-2 Physiological and perceptual responses to the 1500m TT's performed after each training session and at baseline.

	Baseline	**10CON**	**10INT**	**25CON**	**25INT**
Speed (km/h)	19.4 ± 1.4	18.1 ± 1.7[a]	18 ± 1.8[a, #]	18.7 ± 1.5[a]	18.6 ± 1.6[a]
%PTV	98.2 ± 3.3	91.9 ± 3.9[a]	91.2 ± 3[a, #]	94.7 ± 2.8[a]	94.2 ± 4[a]
HR_{mean} (bpm)	173 ± 7	174 ± 9	172 ± 7	174 ± 8	172 ± 9
RPE_{mean}	18 ± 1	18 ± 1	18 ± 1	17.5 ± 1	17.8 ± 1
RPE_{peak}	19 ± 1	19 ± 1	19 ± 1	19 ± 1	19± 1
Motivation	8.2 ± 0.6[b, #]	7 ± 1.1	7.4 ± 1.2	7.3 ± 1.1	7.7 ± 0.9

Values are mean ± SD. Speed= Average speed, %PTV = Average speed expressed as a percentage of peak treadmill velocity, HR= Average heart rate, RPE_{mean} = Average rating of perceived exertion, RPE_{peak}= Final rating of perceived exertion. Motivation was assessed prior to the 1500m performance trial. Statistical significance: [a]$P<0.05$ from baseline, [b]$P<0.05$ from 10CON, [#]$P<0.05$ from 25CON.

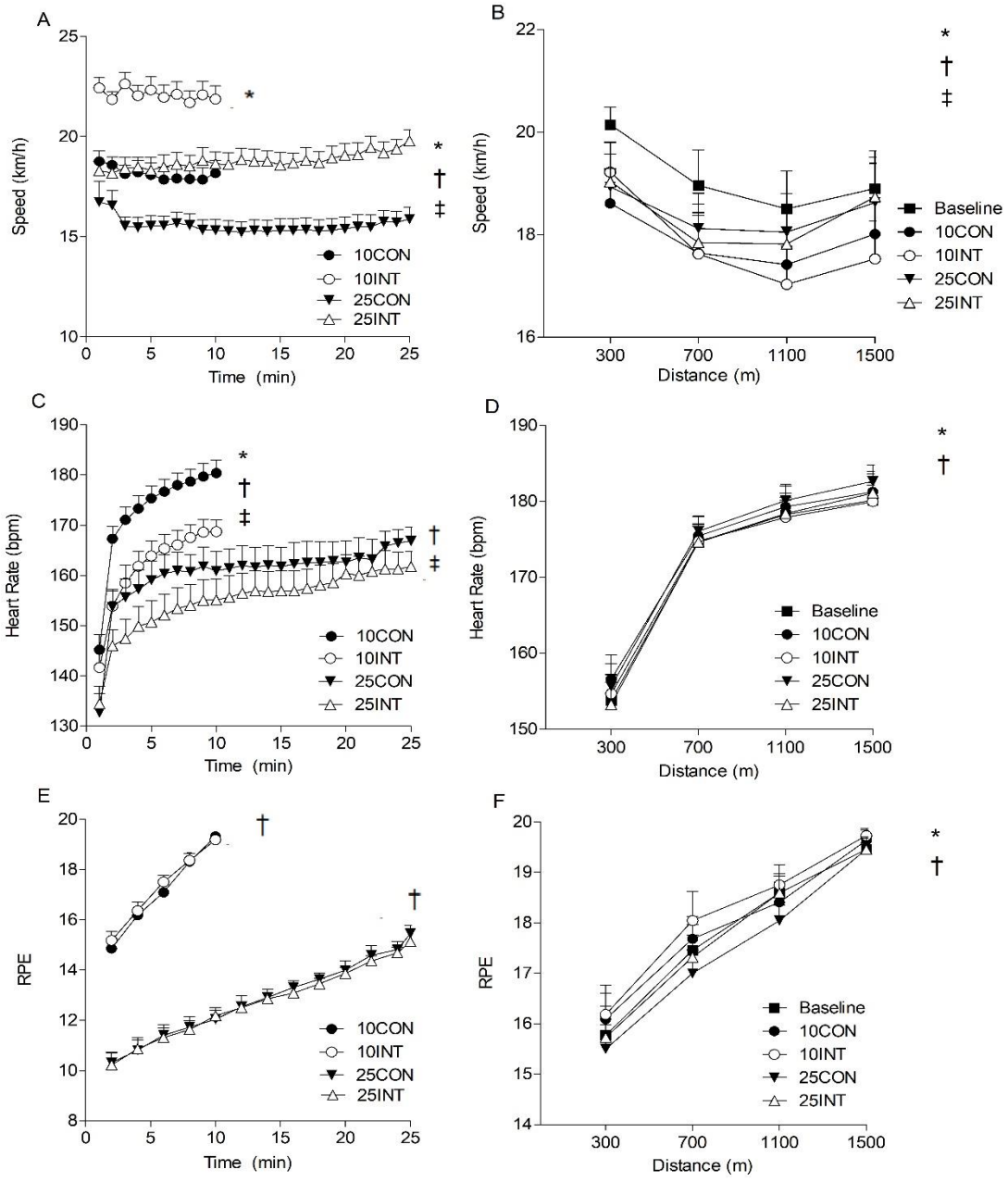

Figure 4-3 The left panels show responses to the experimental training sessions and the right hand panels show responses to the subsequent 1500 TT's. Individual plots are for Speed (A and B), HR (C and D), RPE (E and F). See text for abbreviations. baseline (filled squares), 10CON (filled circles), 10INT (open circles), 25CON (filled triangles), 25INT (open triangles). Values are mean ± SEM. Statistical significance is indicated as follows; (‡) interaction of condition by time ($P < 0.05$), (*) main effect of condition ($P < 0.05$), (†) main effect of time ($P < 0.05$).

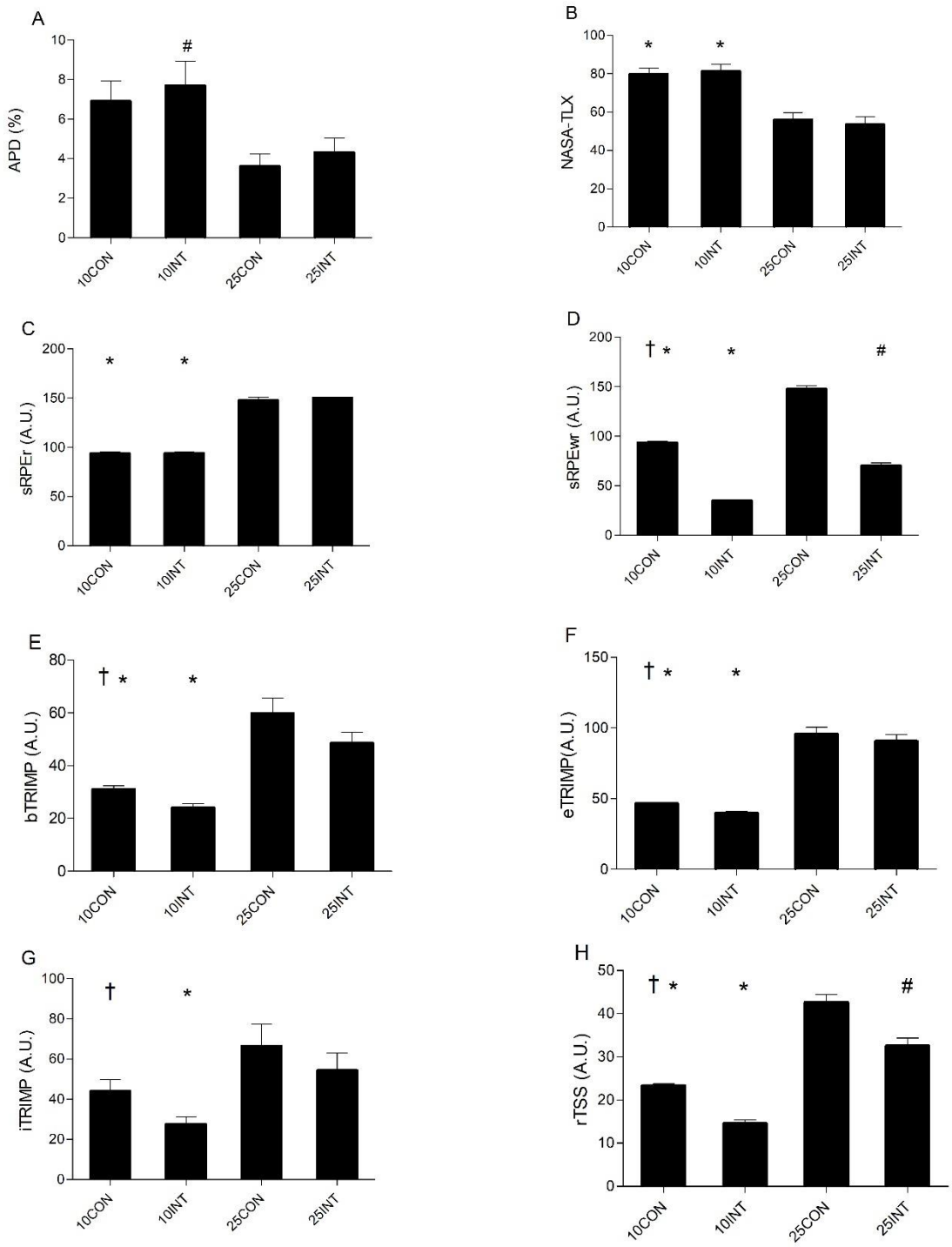

Figure 4-4 Training load metrics for the training sessions: APD - Panel A, NASA-TLX - Panel B, $sRPE_r$ - Panel C, $sRPE_{wr}$ - Panel D, bTRIMP - Panel E, eTRIMP – Panel F, iTRIMP - Panel G, and $rTSS^{TM}$ – Figure H. See text for abbreviations. Values are mean ± SEM. Significant pairwise comparisons are shown as follows; $^{*}P < 0.05$ from 25CON and 25INT, $^{\dagger}P < 0.05$ from 10INT, $^{\#}P < 0.05$ from 25CON.

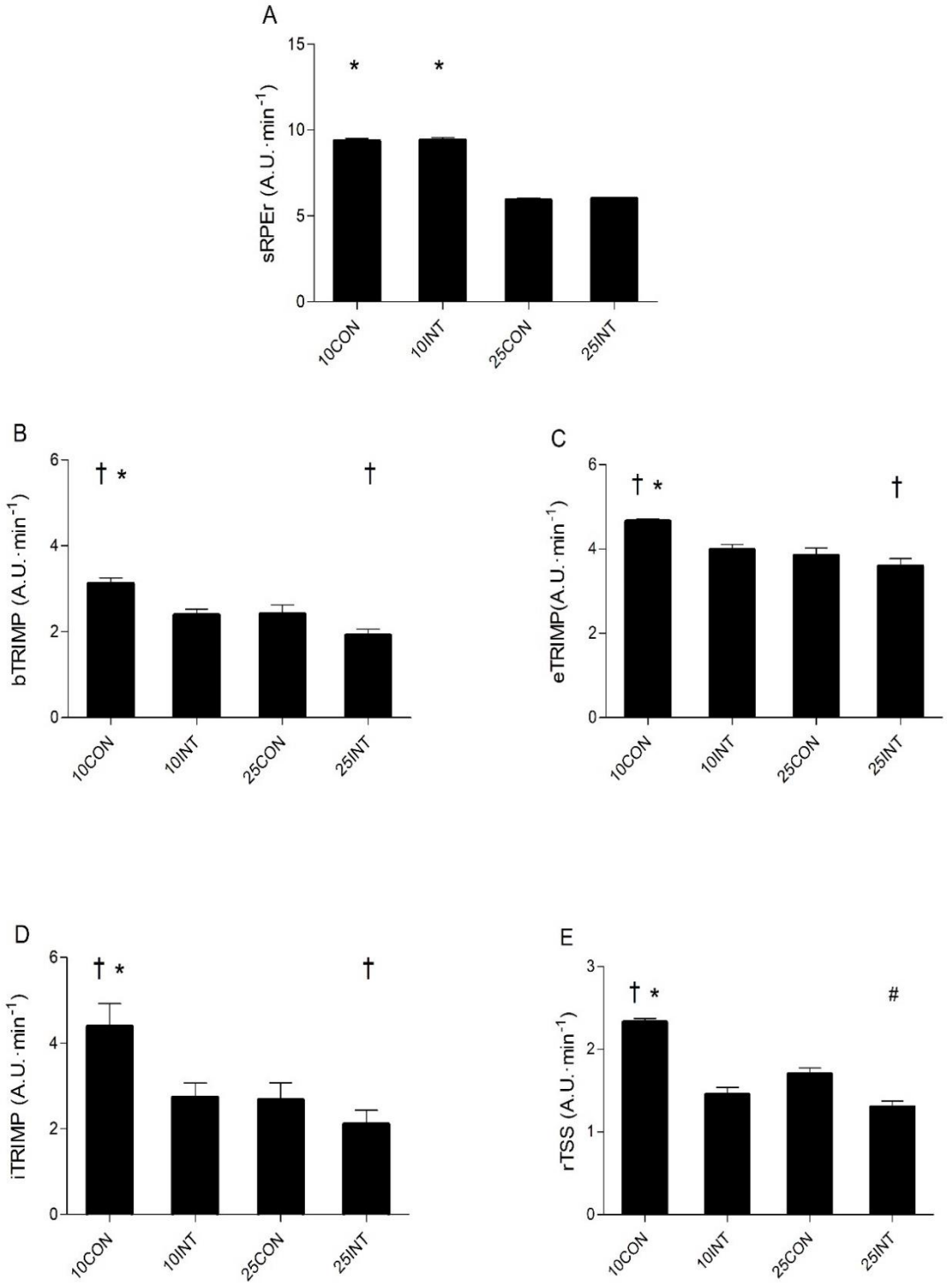

Figure 4-5 Normalized training load metrics for the training sessions: sRPEr- Panel A, bTRIMP - Panel B, eTRIMP - Panel C, iTRIMP - Panel D, rTSS - Panel E. See text for abbreviations. Values are mean ± SEM. Significant pairwise comparisons are shown as follows; $^{*}P< 0.05$ from 25CON and 25INT, $^{\dagger}P < 0.05$ from 10INT, $^{\#}P < 0.05$ from 25CON.

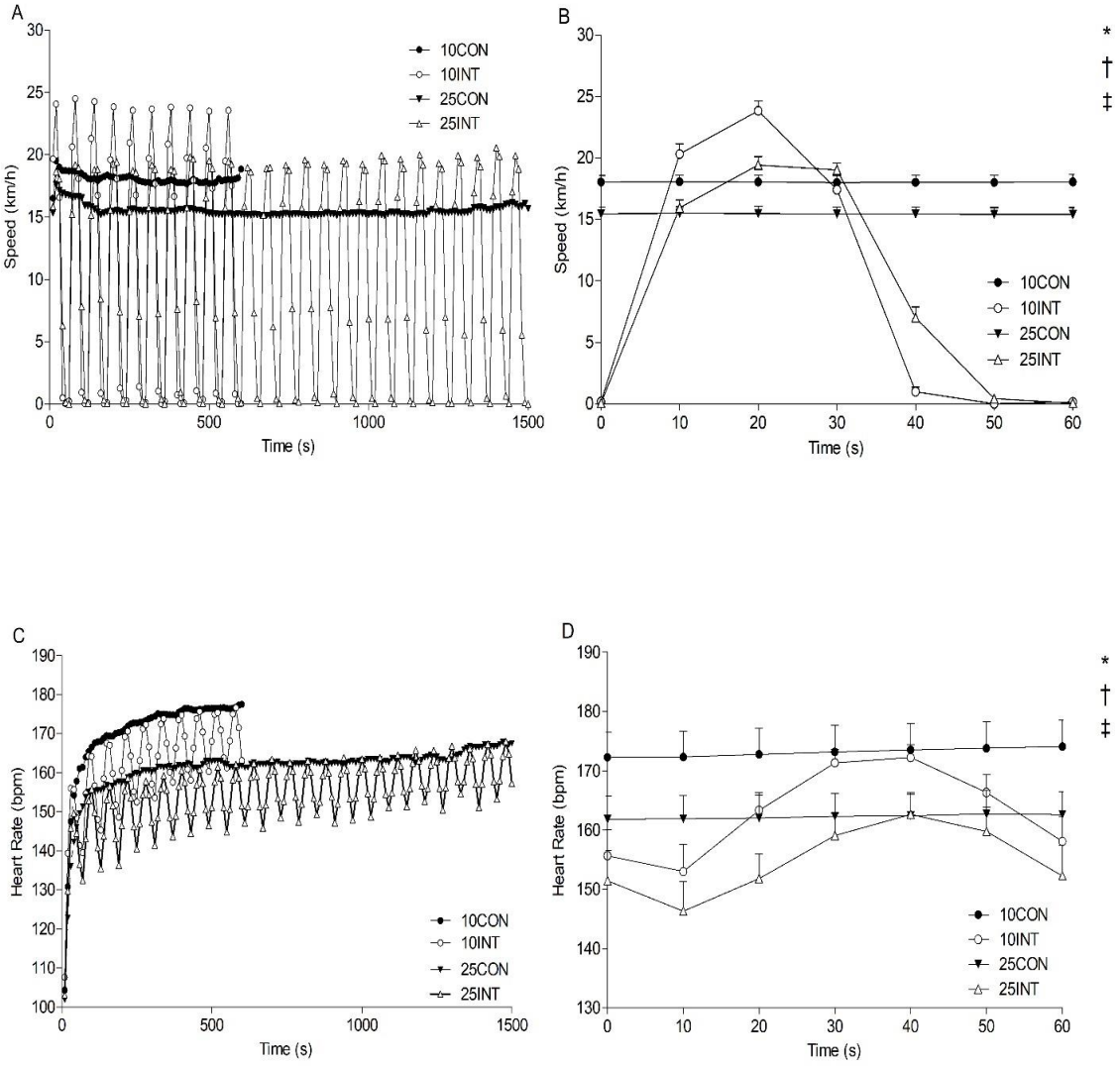

Figure 4-6 The left panels show the group mean time course for Speed (A) and HR (C) from the experimental training sessions. The right panels show the group time course within the work-recovery cycle for Speed (B) and HR (D). Values are mean ± SEM. Statistical significance is indicated as follows; (‡) interaction of condition by time ($P < 0.05$), (*) main effect of condition ($P < 0.05$), (†) main effect of time ($P < 0.05$).

4.6 Discussion

This study examined the APD that results from running training sessions, and the effects of iso-effort CON and INT sessions. The main finding of the study was that an APD measured as a change in 1500m TT speed, was found after all training sessions regardless of their duration or intensity. The magnitude of this APD, was similar for all sessions, except when comparing 10INT with 25CON, where the shorter session resulted in a larger APD. Contradicting the APD response, the session TL metrics were significantly smaller for the shorter sessions when compared with the longer ones (Figure 4.3). Furthermore, whilst APD was not different when comparing CON and INT trials of the same duration, most of the TL metrics were significantly lower for the INT sessions. As can be seen in Figure 4.3, the APD showed the opposite pattern of response to its corresponding TL metrics, with the exception of the NASA-TLX responses, which broadly followed the same pattern as the APD. The large effect sizes observed reinforce the dissociation between APD values and TL metrics. When TL metrics were normalized for session duration however, their pattern of response reversed and was seen to be much more consistent with that of APD.

The disagreement observed between APD and TL metrics questions the usefulness of TL metrics for comparing dissimilar training sessions. These findings are consistent with our previous study of continuous cycling bouts (Kesisoglou et al. 2021b), and extends these observations to situations involving running and CON vs. INT exercise. Previously, we observed a higher APD and lower TL metrics following 5 min maximal cycling compared with 20 and 40 min submaximal bouts. In the present study we found that comparing iso-effort CON and INT exercise highlighted further discrepancies between APD and TL. Typically, TL was lower for INT than CON whilst APD was not different. Previous studies (Nicoló et al. 2014; Nicolò et al. 2017, 2019) have shown that differently structured iso-time and iso-effort bouts require similar levels of effort. These findings are consistent with the APD response observed in the present study, and further support the use of the iso-effort method for prescribing exercise (Figure 4.2:E). The findings from the present study

reinforce our previously expressed concerns on the shortcomings of using TL to compare training sessions of different intensity and duration and highlight further problems with using TL for INT activities.

The HR-based TL metrics (bTRIMP, eTRIMP, iTRIMP) had lower values for 10INT compared to 10CON training sessions (Figure 3). This is probably attributable to the delay in HR response that is most noticeable at the onset and offset of exercise (Figure 4.5). This effect is well-documented for INT activities in football (Nicolò et al. 2019) and cycling (Nicolò et al. 2014, 2017) and here we repeat those findings for INT running sessions. Table 4.1 summarizes the training session data which indicates that the APD's were not proportional to running speed, total distance, session duration or HR but were associated with differences in prescribed exercise intensity. Similar to the findings of our previous study, iso-effort sessions, although different in nature, resulted in similar APD's.

It seems unlikely that issues with the measurement of APD are responsible for its divergence from TL metrics, although the effect of pacing on performance is well known (Jones 2002). Previous studies have used maximum effort TT's effectively to evaluate changes in performance (Passfield and Doust 2000; Currell and Jeukendrup 2013) and over similar distances (Wallace et al. 2014a). Nonetheless, given the field nature of this study, we carefully monitored factors such as pacing and motivation. All participants were experienced middle-distance runners used to running this distance, and they were instructed to follow strictly the same pacing strategy for all their 1500m TT's. Figure 4.2B suggests that this approach was successful, as similar shaped pacing profiles can be seen for all 1500m TT's. Motivation was similar for most 1500m TT's after the training sessions, except when the baseline was compared with 10CON (P=0.012) and 25CON (P=0.016). Moreover, similar values were observed during all 1500m TT's for RPE_{mean}, RPE_{end} and HR_{mean}, which suggests that differences in motivation did not prevent participants from reproducing their maximum effort in all trials. Additionally, the similar time-course for changes in HR, and RPE shown in Figure 4.2:B, also

suggest that the 1500 TT's were performed to a high level of consistency and therefore changes in APD reflect the impact of the prior training session rather than poor 1500m TT execution.

The reason for the contradiction between APD and TL may be related to the way in which exercise duration is incorporated into the TL metrics as a multiplier. Given that the present study used an RPE based prescription that appears to have been applied effectively, it may appear surprising that the sRPE was not consistent with the resulting APD. Most TL metrics, including sRPE, quantify TL based on the model proposed initially by Bannister et al. (1975) multiplying the exercise intensity component (e.g. HR, speed or RPE) by its duration. But it is notable that the APD's observed in the current study were independent of changes in exercise duration. As a RPE scale was used both for prescription (Borg 1982) and to calculate the sRPE (Foster et al. 2001), it seems more likely that it is the use of duration as a multiplier that leads to discrepancies between the resulting APD and TL. This suggestion is reinforced when the TL metrics are evaluated after normalizing them for the training session duration to remove its influence (Figure 4.4). Once normalized the TL metrics show the opposite response and are much more consistent with the changes in APD, especially sRPE. This finding is consistent with that of Weaving et al. (2020) who concluded that duration was the main contributor to the TL variance observed for rugby players over an extended period of time. Notably, in the present study the NASA-TLX was the only metric that tracked APD reasonably, reinforcing the findings of our previous study (Kesisoglou et al. 2021b). The NASA-TLX does not use exercise duration as a multiplier, instead aggregating scores across six different subscales (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, Frustration). Thus, we suggest that the stress imposed by a training session may be better represented by weighting the importance of its duration and intensity separately.

4.6.1 Study Limitations

The field-based nature of this study poses some specific limitations. Due to the nature of the study it is not possible to comment on the potential mechanisms underpinning the APD. Bannister et al. (1975) when first introducing the TRIMP noted training sessions create a fatigue effect, quantified in the present study as an APD. As the etiology of the decrement in 1500m TT performance trials is unknown, we have avoided use of the term fatigue and referred to an APD instead. Indeed, we speculate that the etiology of the APD will likely vary according to the characteristics of the training session. Consequently, the APD measured directly after the session may not be proportional to the recovery time course, nor the overall stimulus for adaptation. Further studies are required to determine how the magnitude of APD varies in response to different sessions, and whether it corresponds to the recovery duration and resulting longer-term adaptations. A further limitation was that because most data collection occurred on an outdoor athletics track, differences in the weather could have influenced the training sessions and subsequent 1500m TT performances. However, environmental conditions and wind speed in particular were monitored for every session which was postponed if wind speed was above 2 $m \cdot sec^{-1}$.

In the present study, the conservative Bonferroni test for post-hoc analysis was adopted to find differences in TL metrics whilst APD was similar. If a less conservative post-hoc test was adopted, the APD's for both 10-min sessions are found to be greater than those for both the 25-min sessions which reinforces the observation as APD and TL metrics change in opposite directions. This is confirmed by the confidence intervals observed. However, given the limited sample size of the present study, the more conservative perspective was preferred. Further research may be needed to examine a more diverse range of training sessions reflecting those typically adopted by athletes and coaches (Buchheit and Laursen 2013a; Nicoló et al. 2014).

4.6.2 Practical applications

Our findings have important implications for comparing training sessions and quantifying INT-based sessions using HR-based TL metrics in particular. Collectively, current TL metrics overestimate the stress imposed by longer running training sessions, independently of whether they are prescribed as CON or INT. Despite different methods implemented for computing TL, such as non-linear (bTRIMP, rTSS), zone-based (eTRIMP) or individualized (iTRIMP), the TL metrics changed in a contradictory manner to APD. In contrast, TL metrics normalized for session duration were reversed in response and much more consistent with APD. When coaches and athletes use TL metrics to evaluate training sessions of varying exercise duration, they should be aware of the potential bias caused by using session duration as a multiplier and consider the influence of duration and intensity separately.

The use of effort-based methods was effective for the purposes of training session prescription for well-trained middle-distance runners as it resulted in an APD following every session. Importantly, APD is not suggested as a replacement for current methods of TL quantification, rather as an objective and theoretical contribution towards the development of new approaches. In this regard, future research could explore the use of NASA-TLX as the APD's found in the present study showed good agreement with these ratings. This finding is consistent with the conclusions of a systematic review on self-report measures (Saw et al. 2016), but the use of NASA-TLX has not been examined thoroughly in a sporting context specifically.

4.6.3 Conclusion

A significant APD was observed as a decrease in 1500m TT performance following training sessions of 10, and 25 min duration comprised of CON and INT running. The changes in APD contradicted the TL metrics for the training sessions both for changes in exercise duration and when performed as CON or INT running. These results suggest that current methods for quantifying TL

may be flawed and new methods need to be developed. Lastly, the present study found NASA-TLX was associated with APD and suggests its potential for evaluating training sessions may be worth exploring further.

Chapter 5: Acute Performance Decrement in Running: Effects of Changes in Exercise Duration.

5.1 Aims and hypothesis

The aim of Chapter 5 is to extend the findings from Chapters 3 and 4. Similarly to Chapter 4, Chapter 5 is based on running performance. Chapters 3 and 4 showed that APD and TL did not agree as they assessed after various intensities and durations. These descrepancies raised question on the effectiveness of TL metrics when it comes to comparing training sessions with dissimilar training characteristics. Importantly, these Chapters also showed that exercise duration and how it is currently incorporated within the TL metrics seem to be the main reason for the discrepancies observed. Chapters 3 and 4 provided evidence that current TL metrics overestimate the importance of exercise duration over exercise intensity. Therefore, Chapter 5 aims to provide a systematic quantification of such an overestimation. Specifically, in Chapter 5, 4 different running training sessions consisted of 2 different intensities and durations will be used (10MIN5, 10MIN8, 40MIN5, 40MIN8). Similarly to Chapter 4, the running training sessions were prescribed via perceived exertion. This design will provide information on the relative contribution of intensity and duration on APD. Based on the above and the findings from Chapters 3 and 4, it is hypothesized that the effect of the changes in intensity and duration on APD will not agree with TL metrics changes.

5.2 Abstract

Purpose: To examine the effects of running training consisting of two different intensities and durations on the resulting acute performance decrement (APD) and corresponding session training load (TL) metrics.

Methods: Twelve trained runners performed a 1500m time-trial (TT) as a baseline performance, and after completing 4 different running training sessions. The training sessions were performed in a random order and consisted of two different exercise intensities (RPE of 5 and 8) and durations (10 min and 40 min). The APD was calculated as the percentage change in 1500m TT performance speed from baseline when measured after the training session. These differences in APD for each session were compared with their TL metrics.

Results: The training sessions lasting 40 min showed greater TL scores compared to those of 10 min (P<0.001; η_p^2 >0.781). These results were in contrast to the APD for the 10 min at RPE 8 and 40 min at RPE 5 which were not different (P > 0.99). However, APD for all other comparisons were different (P<0.05)

Conclusion: The APD response did not agree with TL when compared between sessions**.** The exercise intensity and duration contributed differently to the development of APD and should not be used as independent multipliers for TL quantification purposes.

5.3 Introduction

Training load (TL) is a popular concept and its use in training prescription and for monitoring whether athletes are adapting to their training programme is considered essential (Bourdon et al. 2017). Traditional TL metrics rely on quantifying internal and external markers of exercise intensity such heart rate (HR), blood lactate (La^-), power output (PO), speed, and Rating of Perceived Exertion (RPE), and presenting these as a function of training duration (Borresen and Lambert 2009). The use of TL metrics was inspired by the original work of Banister et al. (1975). These authors proposed that if the training impulse (TRIMPs) for training sessions were quantified, in arbitrary units, these could be used to reflect the accumulation of fitness and fatigue and predict the changes in performance that occur in response. In a later study Morton et al. (1990) demonstrated it was possible to produce a statistically significant prediction of the changes in performance resulting from a 50-day training programme. Notably however, Banister and colleagues never established the validity of TRIMPs as a measure of TL. Subsequent studies that developed new TL metrics also assumed their validity (Edwards 1993; Lucía et al. 2003) or used unvalidated criterion measures (Foster et al. 2001; Wallace et al. 2014b). Two recent studies have compared the acute performance decrement (APD) that occurs following a training session with a range of TL metrics and concluded that most metrics do not seem to reflect the stress of the session (Kesisoglou et al. 2021b, a). The researchers speculated that the reason for the disparity between APD and TL metrics was due in part to the way in which session duration is incorporated into most metrics. Similar concerns regarding the role of duration in the calculation of TL metrics have been expressed by other researchers too (Weaving et al. 2020).

Previous findings in running (Kesisoglou et al. 2021a) and cycling (Kesisoglou et al. 2021b) suggested that exercise intensity and duration contribute differently to the development of APD. This suggestion was also consistent with the findings of Fullerton et al. (2021). For example, Kesisoglou et al. (2021b) found that maximal trials of 5 and 20 min showed a similar APD suggesting that the training stress imposed by the sessions were similar. However, the TL metrics for these sessions

suggested that the training stress was very different. Similarly in running greater APD's were found to be induced by shorter and maximal intensity training sessions, compared to longer in duration but sub-maximal in intensity (Kesisoglou et al. 2021a). A possible explanation for these findings is that the differences in exercise intensity between these sessions were of greater relative importance than the differences in exercise duration. If this speculation is correct, the influence of exercise duration varies according to exercise intensity and therefore should not be used as an independent multiplier in the calculation of TL. While Fullerton et al. (2021) found that when cycling at an intensity corresponding to maximal lactate steady state (MLSS), APD decreased linearly with increased in exercise duration, they only examined the effect of changing exercise duration on APD at one exercise intensity. Further insight into this important issue could be gained if the effects of changing exercise intensity and duration are conducted at two different intensities. Therefore, the present study examined the effect of two training session intensities and durations on the resulting APD and corresponding TL metrics. We hypothesized that the effects of these changes in intensity and duration on APD would not be reflected by their respective TL metrics.

5.4 Methods

Trained runners (9 males, 3 females) (mean ± SD: age: 26 ± 8.7 years old, height: 174 ± 8 cm, weight: 66 ± 8 kg), volunteered to participate (Figure 5.1) in this study. Participants had at least 3 years of experience of competing in events ranging from 800m to a half marathon; they trained >8 hours/week and had personal best times for 1500m of 288 ± 23 s. All participants gave written informed consent to take part in this study which was approved by a local Ethics Committee in compliance with the Declaration of Helsinki.

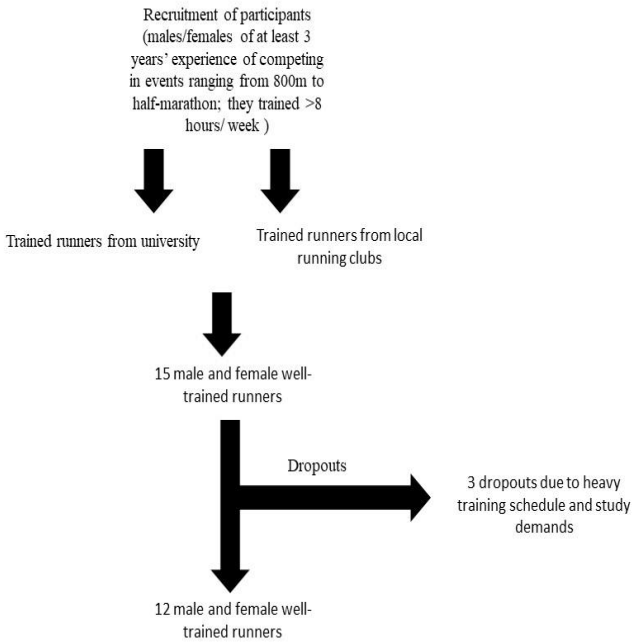

Figure 5-1 Flow chart of recruitment process and dropouts

5.4.1 Experimental Overview

In a randomized, crossover within-subjects design, over a 4-week period, participants completed 5 experimental conditions (see Figure 5.2). All conditions were separated by at least 48 hours. The 5 visits consisted of one baseline 1500m TT (1500m TT), and four experimental training sessions followed by a 1500m TT. All training sessions were completed at the same time of the day (± 2 hours) and participants were asked to attend each session fully rested and hydrated, having consumed a light meal 3 h before, and to have refrained from alcohol, caffeine consumption, and vigorous exercise for 24 hours before each session.

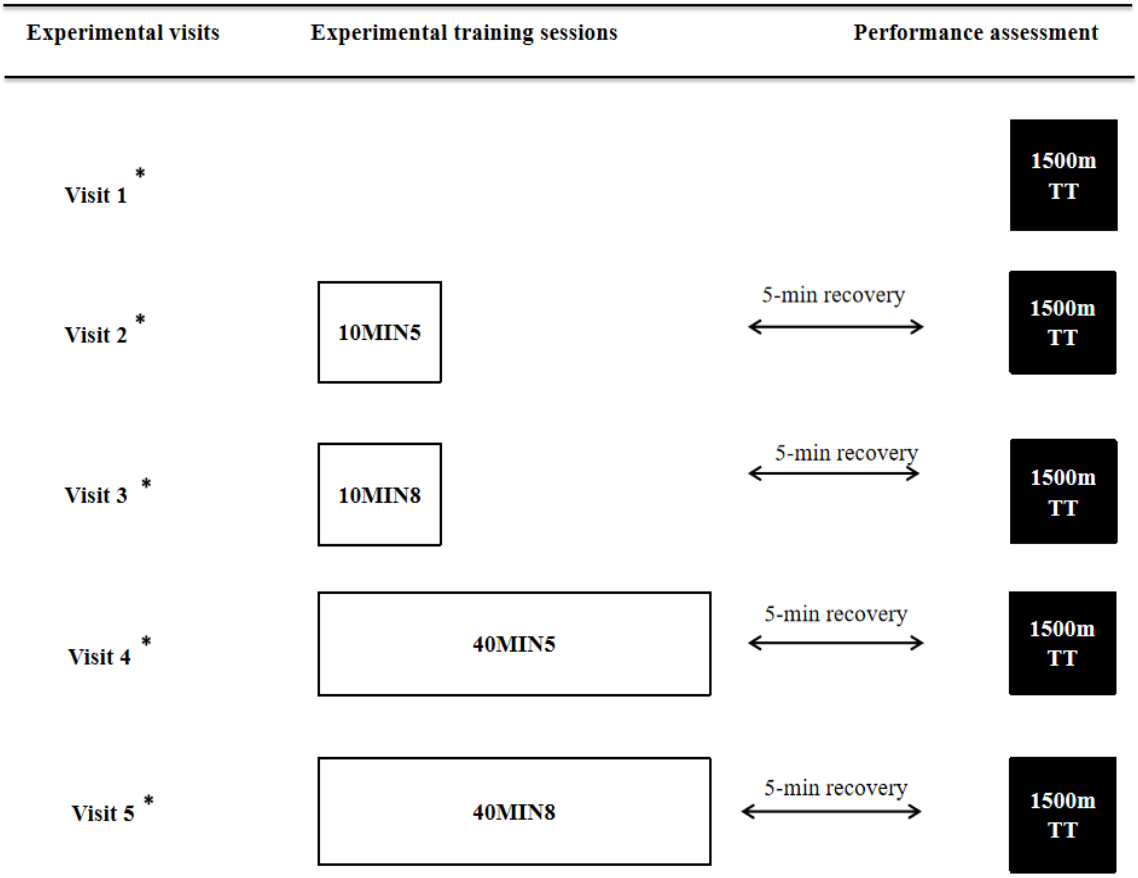

Figure 5-2 The 10MIN5, 10MIN8, 40MIN5 and 40MIN8 are the training sessions (open boxes) and the 1500m TT (filled boxes) used to evaluate the acute performance decrement.

5.4.2 Experimental Training Sessions

Participants completed all 5 visits on an outdoor synthetic athletics track. In addition to the baseline 1500TT. The 4 experimental training sessions consisting of 2, 10 min training sessions, while the other two lasted for 40 min, as shown below in Figure 5.1. Both 10 and 40 min training sessions were performed at a prescribed intensity using the CR-10 scale (Borg 1982), of 5 (10MIN5, 40MIN5 respectively) and 8 (10MIN8, 40MIN8 respectively). As this was an exertion-based exercise prescription, participants were responsible for their own pacing, and no feedback was provided during the training sessions. On completing the experimental training sessions participants were allowed 5 min of passive recovery, before performing a maximum effort 1500m TT. Prior to all experimental visits, participants undertook a standardized self-paced 3-lap warm-up followed by a standardized dynamic stretching routine. Timing during all training sessions and 1500m TT's time was performed manually with a stopwatch. Speed and heart rate were measured second-by-second using a GPS wrist watch and a Bluetooth connected foot pod, (Polar Oy, Polar Electro, Kempele, Finland). The wind speed was assessed with a digital anemometer (Protmex, MS6252A) for all experimental sessions which were only performed if it was below $2 m \cdot sec^{-1}$.

5.4.3 Measurements

The APD for each experimental training session was calculated as the percentage change in 1500m TT speed as calculated in previous chapters (see Chapter 2.8). The APD is presented as a positive percentage change to make comparison with the corresponding TL metrics for that session easier.

RPE (scale 6-20) was measured at 300m, 700m, 1100m, and 1500m during every 1500m TT and every 2 minutes during the experimental training sessions (Borg 1982). The National Aeronautics and Space Administration Task Load Index (NASA-TLX) rating scale (Hart and Staveland 1988)

was used to assess subjective workload immediately after the completion of all experimental training sessions. The NASA-TLX is composed of six subscales that were summed to provide an overall score (Hart 2012). Motivation was assessed prior to all experimental training sessions and 1500m TT's using a 0 to 9 analogue scale (Lurquin et al. 2016). Participants were familiarized with all scales prior visit 1.

TL metrics were calculated via using 2 different TRIMP formulae based on either HR or RPE. The Bannister's TRIMP (bTRIMP) (Bannister 1991) and Edward's TRIMP (eTRIMP) (Edwards 1993) were used as HR-based metrics and session RPE (Foster et al. 2001) (bTRIMP) as a metric based on RPE. The bTRIMP (Bannister 1991) was calculated from HR using a weighting factor according to the formula: bTRIMP = duration training (minutes) · ΔHR · y, where ΔHR = (HR_{ex} - HR_{rest}) / (HR_{max} - HR_{rest}), where *y* represents a classic increase in blood lactate of $0.64e^{1.92x}$ for males and $0.84e^{1.67x}$ for females. The iTRIMP (Manzi et al. 2009b) was calculated using individualized weighting factors (y_i) for each participant using their phase 1 laboratory test data (Manzi et al. 2009b). The eTRIMP (Edwards 1993) was calculated from the time spent in five HR zones that were pre-defined as percentages of HRmax and were multiplied by arbitrary zone-based weighting factors. The sRPE used the participants' scores collected immediately after the training session using the CR-10 scale (Borg 1982) multiplied by training duration. All metrics (bTRIMP, eTRIMP, sRPE) were also normalized by dividing the scores achieved for different metrics by the training duration for each experimental running session.

5.4.4 Statistical Analysis

An a priori statistical analysis was performed using G*Power (version 3.1.9.2; Kiel University, Kiel, Germany). Expecting a large effect size for the effect of exercise intensity and

duration manipulation on APD. A sample size of 11 was required based on 1-β = 0.80 and a = 0.05. 15 participants were recruited to account for potential dropping out.

Statistical analysis was conducted using IBM SPSS Statistics 25 (SPSS Inc., Chicago, IL, ISA). A Shapiro-Wilk test to check data for normality and if data were not normally distributed, the non-parametric Friedman and Wilcoxon signed rank tests were used instead. Following the assessments for normality, a one-way repeated measures ANOVA with Bonferroni post-hoc tests were performed to compare mean values for the experimental training sessions, for the TL metrics, and the subsequent 1500m TT's. Two-way repeated measures ANOVA (session intensity x duration) was used to compare the responses of variables (speed, HR, and RPE) during the experimental training sessions and the 1500 TT's. Where the sphericity assumption was violated, the Greenhouse-Geisser correction was used. Effect sizes were calculated as partial eta squared (η_p^2) and small, medium, and large effects were taken as (η_p^2) ≥ 0.01, (η_p^2) ≥ 0.059, and (η_p^2) ≥ 0.138 respectively (Cohen 1988). Statistical significance was accepted where P < 0.05 was found. All results are presented as mean ± SEM in tables and figures.

5.5 Results

5.5.1 Training Sessions

The total distance for covered for each session was 2149 m ± 136 m, 2639 m ± 142 m, 9176 m ± 601 m, 9693 m ± 714 m for 10MIN5, 10MIN8, 40MIN5 and 40MIN8, respectively.

Table 5.1 shows that average speed was different for all 4 training sessions performed (P<0.001; η_p^2 =0.742), as well as HR (P<0.001; η_p^2 =0.522), RPE_{mean} (P<0.001; η_p^2 =0.845), RPE_{peak} (P<0.001; η_p^2 =0.856) and motivation (P=0.032; η_p^2 =0.232). When comparisons were made for the bouts lasting 10 min (10MIN5 vs 10MIN8), a main effect of time was found for speed (P<0.001; η_p^2

=0.684), HR (P=0.006; η_p^2 =0.514) and RPE (P<0.001; η_p^2 =0.91) (see Figure 5.3, panels A, C and E). Similarly, an effect of session intensity was found for speed (P<0.001; η_p^2 =0.44), HR (P<0.001; η_p^2 =0.789) and RPE (P=0.01; η_p^2 =0.918). No significant interaction was found for speed, HR and RPE (P>0.222; η_p^2 >0.126). When comparisons are performed between the bouts lasting 40 min (40MIN5 vs 40MIN8), an effect of duration was observed for speed (P=0.011; η_p^2 =0.288), HR (P=0.03; η_p^2 =0.296) and RPE (P<0.001; η_p^2 =0.909). A main effect of session intensity was also found for speed (P=0.035; η_p^2 =0.343), HR (P=0.019; η_p^2 =0.405) and RPE (P<0.001; η_p^2 =0.867). No significant interaction was found for speed (P= 0.443; η_p^2 =0.076), HR (P=0.71; η_p^2 =0.296) and RPE (P=0.71; η_p^2 =0.031).

The TL metrics for each session are shown in Figure 5.4. An APD of 3.6%, 6.5%, 6.2% and 11.6% was found after the 10MIN5, 10MIN8, 40MIN5 and 40MIN8, respectively. A difference between 40 min and 10 min training sessions was found for bTRIMP (P<0.001; η_p^2 =0.907), eTRIMP (P<0.001; η_p^2 =0.781), sRPE (P<0.001; η_p^2 =0.991) and NASA (P<0.001; η_p^2 =0.497). When TL metrics were normalized for the duration of each training session, differences were found for bTRIMP (P=0.004; η_p^2 =0.523) but not for eTRIMP and sRPE (P=0.063; η_p^2 =0.237). Bonferroni post-hoc testing revealed between training session differences for bTRIMP, eTRIMP, NASA and sRPE (Figure 5.4). Similar post-hoc testing was performed to identify between session differences for normalized metrics (Figure 5.4).

5.5.2 The 1500-m TT's

The data for the 1500m-TTs are presented in Figure 5.3 (B, D and F Panel) and Table 5.2 An APD was found after all training sessions (P<0.001; η_p^2 =0.62). Bonferroni post-hoc testing found differences in the size of APD for all training sessions, except for 10MIN8 vs. 40MIN5 (P > 0.99;

see Figure 5.4). For the APDs after the training sessions, the 95% Confidence intervals (CI) were calculated following the 10MIN5 [2.3, 4.9], the 10MIN8 [5.3, 7.8], the 40MIN5 [4.3, 8.2] and the 40MIN8 [8.9, 14.3]. A main effect of session duration was found for speed (P<0.001; η_p^2 =0.44), HR (P<0.001; η_p^2 =0.886) and RPE (P<0.001; η_p^2 =0.887) during the 1500-m TTs. A main effect of session intensity was found for speed (P<0.001; η_p^2 =0.78) and RPE (P=0.002; η_p^2 =0.31), but not for HR (P=0.675; η_p^2 =0.051). A significant interaction was found for speed (P=0.001; η_p^2 =0.355), but not for HR (P=0.026; η_p^2 =0.886) and RPE (P=0.071; η_p^2 =0.166).

Table 5-1 Physiological and perceptual responses to the training sessions.

	10MIN5	**10MIN8**	**40MIN5**	**40MIN8**
Speed (km/h)	$14.5 \pm 0.8^{\ddagger,\S}$	$16.1 \pm 0.2^{*}$	$13.7 \pm 0.2^{\#}$	14.5 ± 0.3
HR_{mean} (bpm)	$157 \pm 5^{\ddagger,\#}$	170 ± 4	162 ± 4	170 ± 4
Total Distance (m)	2419 ± 39	2693 ± 41	9176 ± 174	9693 ± 206
RPE_{mean}	$13 \pm 0.3^{\ddagger,\#}$	16 ± 0.3	$12 \pm 0.4^{\#}$	15 ± 0.3
RPE_{peak}	$14 \pm 0.4^{\ddagger,\#}$	18 ± 0.2	$14 \pm 0.4^{\#}$	18 ± 0.2
Motivation	7.1 ± 0.2	7.0 ± 0.2	6.7 ± 0.3	6.3 ± 0.3

Values are mean ± SEM. Speed= Average speed, HR= Average heart rate, RPE_{mean} = Average rating of perceived exertion, PRE_{peak}= Final rating of perceived exertion. Statistical significance: $^{*}P < 0.05$ from 10MIN8, 40MIN5 and 40MIN8, $^{\dagger}P < 0.05$ from 40MIN5 and 40MIN8, $^{\#}P < 0.05$ from 40MIN8, $^{\ddagger}P < 0.05$ from 10MIN8, $^{\S}P < 0.05$ from 40MIN5, $^{\mu}P < 0.05$ from 10MIN5.

Table 5-2 Physiological and perceptual responses to the 1500m TT's performed after each training session and at baseline.

	Baseline	**10MIN5**	**10MIN8**	**40MIN5**	**40MIN8**
Speed (km/h)	17.9 ± 0.3$^{*,\mu}$	17.3 ± 0.3*	16.8 ± 0.4$^{\#}$	16.9 ± 0.4	16.1 ± 0.3
HR_{mean} (bpm)	163 ± 4.7	161 ± 5.4	162 ± 5.1	162 ± 4.3	167 ± 3.7
RPE_{mean}	17 ± 0.4	17 ± 0.3$^{\ddagger, \#}$	17 ± 0.3	17 ± 0.4	18 ± 0.3
RPE_{peak}	20 ± 0.1	19 ± 0.2	20 ± 0.1	19 ± 0.2	20 ± 0.1
Motivation	7.4 ± 0.2‡	7.0 ± 0.2	6.6 ± 0.2	6.8 ± 0.2	6.6 ± 0.1

Values are mean ± SEM. Speed= Average speed, HR= Average heart rate, RPE_{mean} = Average rating of perceived exertion, PRE_{peak}= Final rating of perceived exertion. Statistical significance: $^{*}P< 0.05$ from 10MIN8, 40MIN5 and 40MIN8, $^{\dagger}P < 0.05$ from 40MIN5 and 40MIN8, $^{\#}P < 0.05$ from 40MIN8, $^{\ddagger}P < 0.05$ from 10MIN8, $^{\S}P< 0.05$ from 40MIN5, $^{\mu}P < 0.05$ from 10MIN5

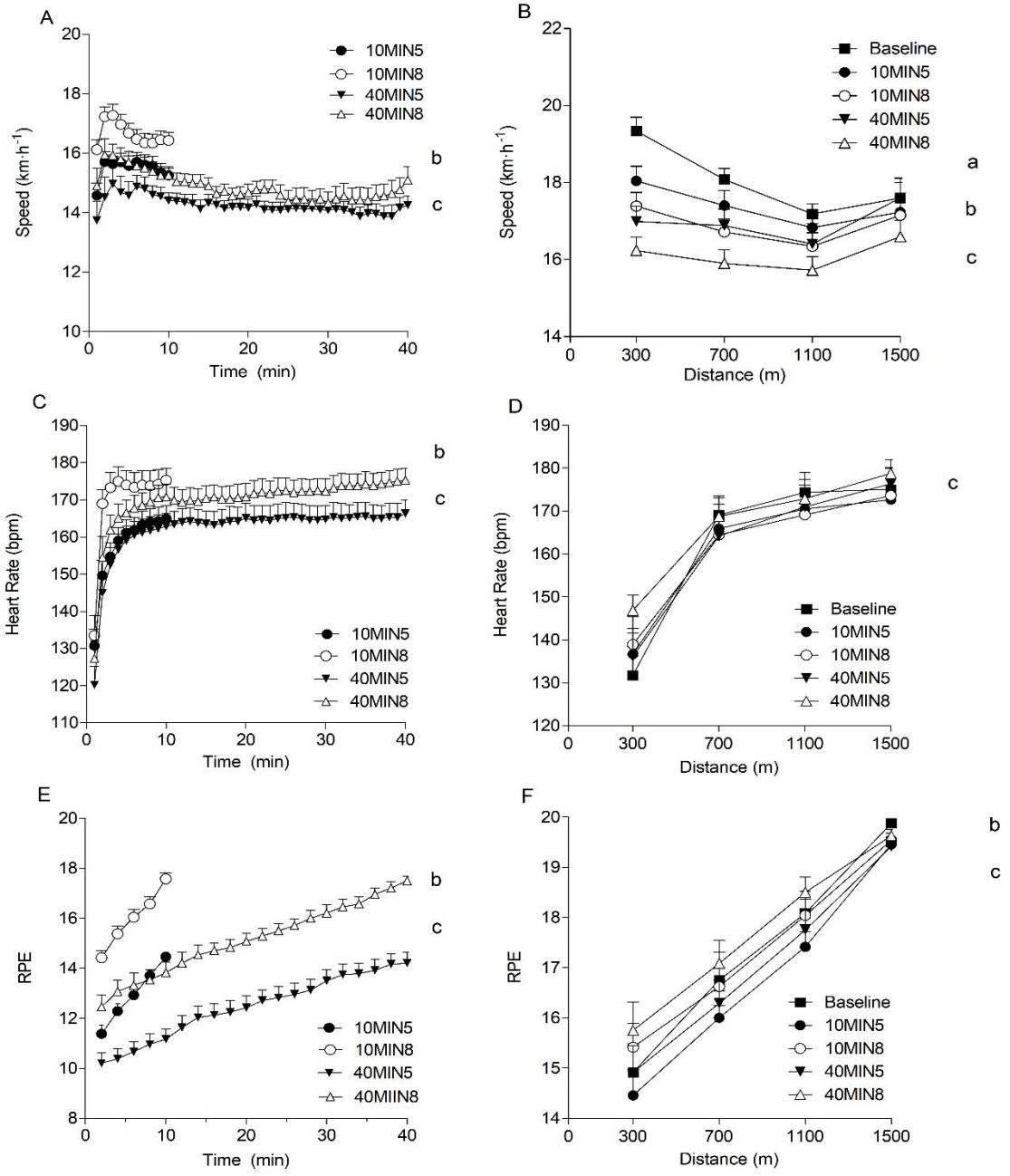

Figure 5-3 The left-hand panels show responses to the training sessions, while the left-hand panel show responses to the subsequent 1500mTT's. Individual plots are for Speed (A and B), HR (C and D) and RPE (E and F). The training sessions are shown as : Baseline (filled squares), 10MIN5 (filled circles), 10MIN8 (open circles), 40MIN5 (filled triangles), and 40MIN8 (filled triangles). [a]interaction of session intensity by time ($P < .05$), [b]main effect of session intensity ($P < .05$), [c]main effect of time ($P < .05$)

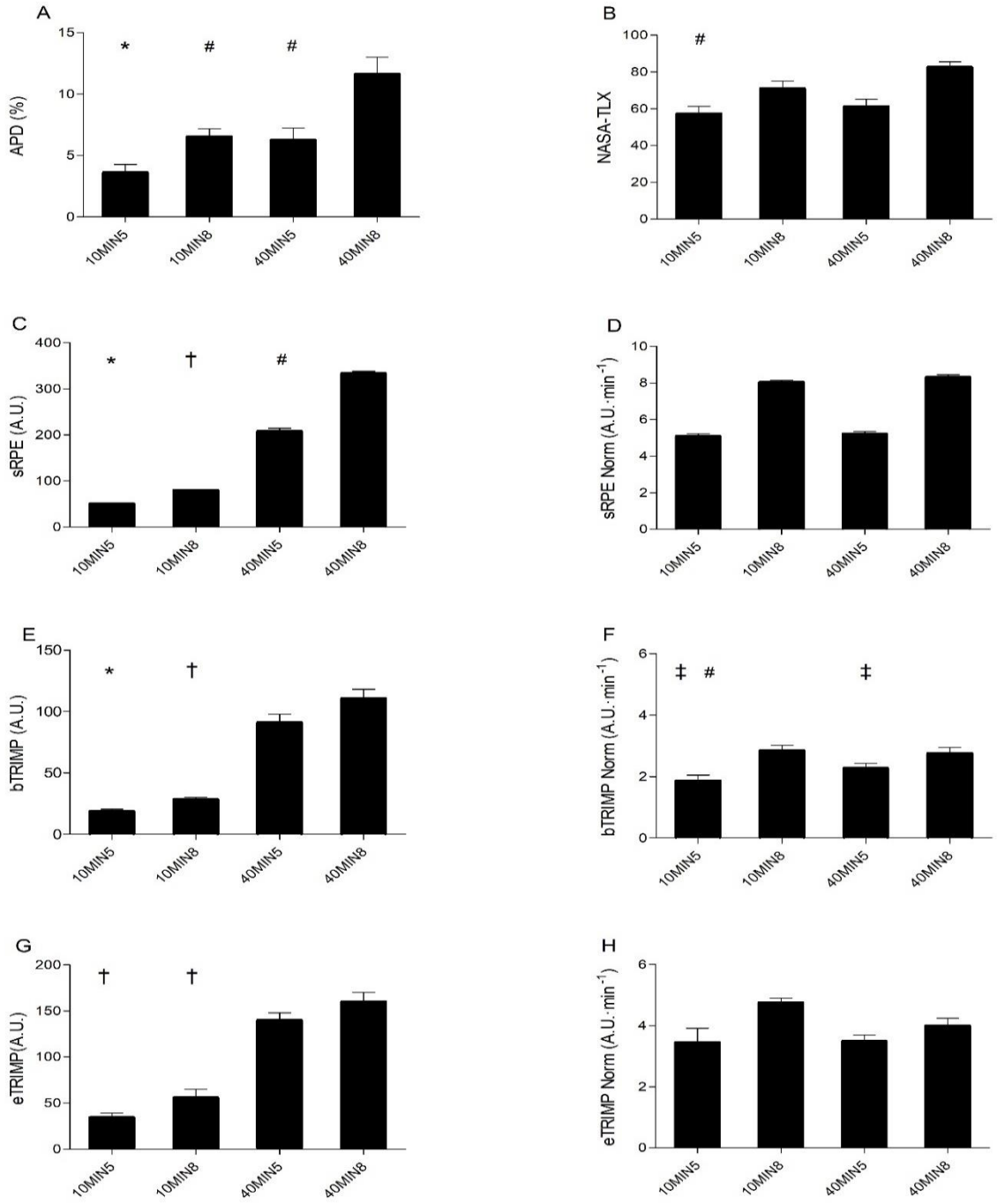

Figure 5-4 Training load metrics for the training sessions (A) APD, (B) NASA-TLX, (C) sRPE, (D) sRPE Normalized, (E) bTRIMP, (F) bTRIMP Normalized, (G) eTRIMP and (H) eTRIMP Normalized. APD indicates acute performance decrement; NASA-TLX, NASA (National Aeronautics and Space Administration Task Load Index; sRPE, session rating of perceived exertion; bTRIMP, training impulse of training session, eTRIMP, Edwards TRIMP. Values are mean ± SEM. Significant pairwise comparisons are shown as follows: *P < .05 from 10MIN8, 40MIN5 and 40MIN8, ‡P < .05 from 10MIN8, †P < .05 from 40MIN5 and 40MIN8, $^{\#}$P < .05 from 40MIN8

5.6 Discussion

The purpose of the study was to examine the effects of running training sessions at two exercise intensities and durations on their APD's and corresponding TL metrics. The main finding of the study was that session APD and TL metrics were not in agreement. This lack of agreement was found regardless of whether duration was included or excluded (i.e. TL metrics were normalized). Greater APD's might be anticipated for the training sessions with the higher exercise intensities (10MIN8, 40MIN8), given previous study findings (Kesisoglou et al. 2021b, a). The findings of the current study confirm that to an extent, as the 40MIN8 training session resulted in the highest APD. However, when the APD's resulting from the 10MIN8 and 40MIN5 training sessions were compared, the longer training session of lower intensity (40MIN5) was found to result in a similar APD to the much shorter higher intensity session (10MIN8) even though both sessions were calculated to have very different TL's. Consequently, it can be seen that relatively small changes in prescribed intensity from 5 to 8, resulted in the same APD despite the large changes in duration from 10 to 40 min. This finding is important as it consistent with our earlier speculation that the use of exercise duration as an independent multiplier in the calculation of TL is inappropriate. These findings are also consistent with a previous study where intensity and duration were found to contribute differently to the development of APD (Fullerton et al. 2021) and indicates that the interaction between cannot be calculated by a simple multiplication. The findings of the present study indicate that the effects of exercise duration vary according to the exercise intensity and merits a systematic evaluation of this interaction between exercise duration and intensity and how this should be modelled by TL metrics. Nonetheless, the fact that the same training stress, as indicated by APD was induced by two very different training sessions in intensity and duration has important implications when considering training prescription and training quantification of running.

The current role of exercise duration in TL quantification has been criticised by researchers (Weaving et al. 2020; Kesisoglou et al. 2021b, a). Current findings suggest that intensity and duration may combine in a disproportionate manner in the development of APD. In a previous study, where exercise intensity was maximal (Kesisoglou et al. 2021b), a fourfold difference in duration did not change the resulting APD. In this example, the two maximal cycling training sessions consisted of 5 min at 337W and 20 min at 273 W whilst APD was 17.9% and 18.4% after both sessions (Kesisoglou et al. 2021b). This finding suggests that the effects of duration are moderated by the exercise intensity. This previous finding suggests that when exercise intensity is maximal APD is similar regardless of the duration. In contrast, in these circumstances TL metrics underestimate the stress imposed by the shorter maximal sessions. In submaximal sessions, the longer and less intense training ones induced a lower APD whilst the TL metrics overestimated their effects presumably because of the bias caused by exercise duration. The current findings may further suggest that when intensity is below maximal levels, the relative contribution of intensity and duration can vary considerably. Training sessions with RPE5 (10MIN5, 40MIN5) were found to induce an APD of 3.6% and 6.2%, respectively. On the other hand, when training intensity was increased (10MIN8, 40MIN8) average APDs of 6.5% and 11.6% were found. These differences in the size of APDs are consistent with the suggestion that the effects of duration depend upon the exercise intensity. Consequently, the findings from the present study are in line with previous findings (Kesisoglou et al. 2021b) and reinforce the suggestion that as intensity decreases, the effect of duration also changes exponentially so that a substantially greater duration is required to achieve a similar APD.

NASA-TLX scale has previously been found to respond in a similar manner to APD (Kesisoglou et al. 2021b, a). These findings were not replicated in the present study as changes in APD did not agree with NASA-TLX. The NASA-TLX responses did follow the same pattern as the normalized TL metrics however (Figure 5.3). This finding reinforces the suggestion from earlier studies that NASA-TLX may be more sensitive to changes in the intensity of the exercise, rather than

the duration. This may explain the agreement between NASA-TLX and APD observed previously for maximum effort training sessions in both cycling and running studies (Kesisoglou et al. 2021b, a). The fact that NASA-TLX followed the same pattern as normalized metrics, but not TL metrics reinforces previous suggestions that exercise intensity and duration may have to be considered separately (Weaving et al. 2020). These findings appear to suggest that NASA-TLX may introduce a bias in the quantification of the training dose which is the opposite of that found for currently used TL metrics. While sRPE, TRIMP and other TL metrics appear to overestimate the importance of exercise duration (Kesisoglou et al. 2021b, a), the NASA-TLX scale seems to underestimate it. However, further research is needed to investigate how NASA-TLX reflects various combinations between intensity and durations of endurance exercise.

5.6.1 Limitations

This study was field-based, and performance was measured on an uncovered outdoor running track. However, it is considered unlikely that the field nature of the study represents a limitation that has meaningfully affected the findings. There are several reasons for this suggestion. For example, a similar experimental design for assessing APD and TL metrics was used in running (Kesisoglou et al. 2021a). A 10-min running training session performed at 10/10 on the CR-10 scale (Borg 1982) resulted in an APD of 7% and 7.5% after continuous and intermittent running respectively. This compares well with the findings of the present study, where an APD of 6.5% was found after 10MIN8 running training session. These results seem consistent with the varying nature of the sessions across both studies and suggest that the standardization of field conditions was effective. A further limitation could be due to changes in the pacing profile between the 1500m TT's. The effects of pacing on performance are known, (Jones 2002) and could have introduced variability in the current results. Participants were instructed to strictly follow the same pacing strategy for the 1500m TT's across all

visits and they all had previous experience of pacing such a distance. The pacing profiles (Figure 5.2B) showed a similar pattern for all TT's and were also similar to those observed previously (Kesisoglou et al. 2021a). Moreover, the HR_{end}, RPE_{mean}, RPE_{peak} were not different between all 1500m TT's. Collectively, these observations suggest that given the field nature of this study the 1500m TT's were performed consistently and maximally and thus did not affect the APD seen.

5.6.2 Practical applications

Our findings provide useful insights in regard to both training prescription and quantification. Current TL metrics should not present a score by multiplying intensity and duration, as the effects of duration seem to vary according to the intensity of the exercise. Coaches and athletes should be aware that intensity and duration interact and that current methods for TL quantification may lead to misjudgements about the stress of a training session. The importance of exercise duration in determining the stress of a session varies with the exercise intensity. The large effect sizes observed reinforce this consideration. This means session duration should not be used as a constant multiplier but varied according to the exercise intensity. Furthermore, this finding may help us reinterpret comparative findings from previously published training studies where session duration and intensity vary.

The fact that the same training effect can be achieved via two training sessions with completely different characteristics can also be of importance for injury prevention and periodization purposes. Traditionally, endurance training heavily relies on how much training can be achieved/sustained. Our current results suggest that coaches can prescribe the same training stress by slightly manipulating effort (i.e. 5 to 8/10), by considerably changing training duration (i.e. 10 to 40 min. Thus, an athlete can be exposed to a less intense training session for greater length of time, and

that often happens with athletes risking with overtraining or returning from injuries, where a gradual/control amount of training needs to occur.

5.6.3 Conclusion

The same APD was induced by two training sessions of different exercise intensity and duration which had markedly different TL metrics. The current findings suggest that the contribution of exercise duration to APD varies according to exercise intensity. In contrast to previous studies NASA-TLX did not agree well with the observed APD. When quantifying the TL imposed by a training session, its intensity and duration should be weighted differently.

Chapter 6: General Discussion

6.1 Is APD meaningful?

In Chapter 3 the concept of APD was introduced, to help quantify the stress of various endurance training sessions and used to evaluate whether TL metrics were in agreement with APD. Previous evidence has suggested that short in duration TTs can provide a reliable measure of performance (Currell and Jeukendrup 2013), which has been found to be reduced by prior exercise but this was in studies that did not also explore the concept of TL (Passfield and Doust 2000; Rodríguez-Marroyo et al. 2017; Clark et al. 2018; Fullerton et al. 2021). All experimental Chapters found that an APD was induced by the training sessions. That is an important finding, as even low intensity training sessions for small or large exercise durations was seen to result in an APD. An example of this is the APD's found after the 40 min cycling training sessions (Chapter 3, see figure 3.2), or after the relatively easy, 10 min running training session (Chapter 5, see figure 5.3). APD was proposed as a means to assess the training-induced stress resulting from an endurance training session and the fact that a drop in APD was found even after relatively easy training sessions suggest that it can detect the stress imposed by exercise for a wide range of training sessions. The magnitude of APD was found to be a fairly consistent when exercise was compared for the same training modality (i.e. running, Chapters 4 and 5). For example, 10 min CON and INT training sessions performed at maximal effort (Chapter 4) resulted in an APD of 7% and 7.5%, respectively. In the later study where running training sessions were used (Chapter 5), an APD of approximately 6.5% was found after the 10MIN8 training session (figure 5.3). The APD's observed from these 3 different sessions may indicate the use of APD as a fairly consistent tool for evaluating the stress imposed by a training session. Comparisons between the cycling (Chapter 3) and running (Chapter 4) APD's, may not be directly comparable as it has been suggested exercise modalities may also affect the TL

(Sandbakk et al. 2021). The findings of this thesis agree with previous indications that APD is sensitive to changes in intensity and duration, and that changes in intensity appear to be the main determinant (Fullerton et al. 2021). The greatest APD values were found when maximal intensity training sessions were completed (Chapters 3 and 4). APD was proposed to quantify the training stress imposed by exercise. TL metrics, which combine indices of intensity and duration rather than measuring the training stress directly, did not agree with APD. The purpose of APD was to further understand how current metrics represent the training stress and to highlight potential limitations with TL metrics.

6.2 Are current TL metrics effective?

The pattern of TL metrics did not agree with the pattern of APD across all the studies presented in this thesis. Throughout the different studies, TL was manipulated in different ways and yet, no agreement was observed. TL was imposed via cycling and/or running, as well as continuous and intermittent type of endurance training. For example, an agreement was not found when TTs of different durations were compared with longer durations but lower intensity training sessions (Chapter 3). Moreover, there were limitations with the TL metrics when comparisons between CON and INT training sessions in running were examined. TL metrics overestimated the stress imposed by CON compared to INT running sessions. When continuous endurance exercise was used (Chapters 3 and 5), TL metrics seemed to favor longer training sessions performed at lower exercise intensities. This was apparent in cycling, where 20 and 40 min submaximal training sessions resulted in TL scores that were higher than the 5 min TT (Chapter 3). This is in line with findings from chapter 5, where the 40 min training sessions resulted in greater TL metric scores compared to the 10 min training sessions (Chapter 4). Throughout all the studies in this thesis, TL metrics seemed to favor the effects of exercise duration over intensity. Three reasons may explain why TL metrics did not

agree with APD and these will be discussed below. Firstly, it may be the weighting of exercise intensity used in current TL metrics. Secondly, it may be the physiological markers used as markers of exercise intensity for TL metrics. Third and perhaps most important, it seems that the role of exercise duration within the calculation of TL metrics is overemphasized when they are compared to APD responses.

6.3 Why TL metrics fail: The role of exercise intensity

Chapters 3, 4 and 5 showed that exercise intensity has been suggested as a key factor towards the development of APD. The role of exercise intensity markers for TL quantification has been extensively discussed in the literature (Foster et al. 2001; Impellizzeri et al. 2004; Alexiou and Coutts 2008; Manzi et al. 2009b; van Erp et al. 2019a; Van Erp et al. 2020). The way exercise intensity is weighted within the current TL metrics is a topic which has received a lot of attention in the TL research. For example, some methods utilize non-linear metrics based on HR (bTRIMP), zone based approaches (eTRIMP, LuTRIMP), individualized methods (iTRIMP), or other nonlinear methods based on power output or Speed (TSS, rTSS). However, none of these methods for TL quantification showed an agreement with APD in Chapters 3, 4 and 5. The relationship between exercise intensity and how much stress is imposed on an athlete is not linear and despite some metrics attempting to take this into account (Morton et al. 1990; Manzi et al. 2009b), an overall agreement between these metrics and APD was not found. Some methods were suggested as more effective compared to others for TL quantification. For example, a metric based on power output (TSS) was suggested to better describe the changes in TL for higher exercise intensities compared to other metrics based on HR or RPE (van Erp et al. 2019b). The authors attributed this finding to the nonlinear (quadratic) weighting that TSS utilizes. TSS was used in Chapter 3 for cycling and Chapter 4 for running (rTSS), and in

neither study agreed with the pattern of APD observed. Other individualized methods based on HR have been suggested to function better compared to other metrics (iTRIMP) for TL quantification (Manzi et al. 2009b). However, the iTRIMP did also fail to agree with APD in chapter 4. Therefore, problems may not only be related to the weighting of exercise intensity, which is non-linear, but also due to the physiological parameters to describe the effects of intensity.

HR is a popular method within TL metrics (bTRIMP, eTRIMP, LuTRIMP, iTRIMP) and has inherent limitations because of its day-to-day variability (Buchheit 2014). Moreover, HR_{mean} has previously received criticism on whether it can accurately reflect the actual physiological demands of exercise. This has been an issue particularly when exercise intensity is very high or very low in cases like professional cycling for example (Sanders et al. 2017a). In Chapters 3 and 4, no HR-based metrics agreed with the APD responses. This finding is in line with other study findings (Alexiou and Coutts 2008; Manzi et al. 2015; van Erp et al. 2019a), where TL metrics where used to quantify low versus high exercise intensities. Similar problems were found when zone-systems were used (Edwards 1993; Lucía et al. 2003), with eTRIMP and LuTRIMP to did not follow the same pattern as APD. This may be due to in part to the different zone-based weighting factors and how they are assigned to each zone. Another important limitation is the fixed blood lactate-to-workload relationship that is used in bTRIMP (Banister et al. 1975) as this may become problematic with changes in an athletes' status over time. Although Morton's weighting factor (Morton et al. 1990) attempted to overcome this issue, the assumption that HR and blood lactate relationship is fixed is a known issue. Although this issue was addressed via individualized methods (Manzi et al. 2009b), others suggested that intermittent exercise can alter the HR- blood lactate relationship and therefore can affect how TRIMPs are calculated (Akubat and Abt 2011). In all our studies, we found that these concepts may be relatively helpful to understand the training stress imposed by the training sessions.

However, popular physiological variables (i.e HR, La^-) and how they are used in different TL metrics did not describe APD in a similar way to perceived exertion.

Collectively in Chapters 3, 4 and 5, perceived exertion better explained APD. It was the RPE and f_R (Nicolò et al. 2016a) responses that were collected in the lab (Chapter 3), and the RPE responses in one of the running studies (Chapter 4). Perceived exertion has been recognized as marker of exercise intensity in the context of TL (Borresen and Lambert 2009). The higher the RPE across training sessions (Chapters 3 and 4), the higher the APD. This was not always the case with the TL metrics. The popular physiological measurements like HR, or VO_2 did not describe APD as well as overall RPE did in all 3 experimental studies. For example, in our cycling study, we found that cycling training at a similar, high perceived exertion showed similar APDs even when total duration was 4 times higher in contrast to other TL metrics (figure 3.2). HR values obtained from the cycling training sessions were also useful for predicting APD (see table 3.1), but not in the running studies (Chapters 4 and 5). This may be due to the different demands placed during running compared to cycling (Millet et al. 2009) and the fact that HR was unable to reflect the higher perceived exertions imposed by the INT running training sessions (Chapter 4). Therefore, it may not only be the way intensity is weighted, but also the physiological variables used to evaluate the effects of intensity. Currently used variables (HR, La^-, Power, Speed) and how these are implemented within the TL calculations do not seem as efficient as using RPE itself in reflecting the overall training stress, especially at the highest intensities. Nevertheless, intensity was the parameter that well described APD. Current findings though suggest the use of RPE rather than popular markers of exercise intensity. Therefore, the use of RPE is also recommended as an inexpensive and relatively easy method for training monitoring purposes. The fact that the training stress applied is mainly dependent on how intense the exercise is may also explain findings from studies where performance was assessed in terms of long-term adaptations and the TWD was fixed between training sessions (Helgerud et al. 2007; Rønnestad et al. 2020). Earlier in this thesis (chapter 2), it was suggested that the studies that investigate the long-

term effects on performance should be interpreted in the context of TL with caution, due to factors such recovery and diet can affect the adaptation process. Although this is a speculation, the findings of these two studies (Helgerud et al. 2007; Rønnestad et al. 2020) may also suggest that greater adaptations were found because of the higher intensities implemented in certain training sessions.

Despite the importance of perceived exertion, a metric which is based on RPE (sRPE) did not reach an agreement with APD in Chapters 3, 4 and 5. When RPE was normalized for time (i.e. just RPE without multiplying by duration) its correspondence with APD improved markedly. This is an interesting finding that questions the use of exercise duration within the metric. Indeed it is speculated that the sRPE score may already reflect the effect of exercise duration before duration is used as a multiplier. For example, RPE has been found to increase at constant work rate (Noakes 2004) and this suggests that end session RPE is inversely proportional to the session duration. This may mean that end RPE already includes the effects of duration. For example, when an athlete is been asked ‘’how was your workout?’’ (Foster et al. 2001), the length of the session may already be included as a factor in providing their response. Evidence suggests that without feedback on session duration, the momentary sensations (i.e. RPE), make maintaining power output and speed more challenging (de Koning et al. 2011). This may be another reason to why sRPE overestimated the effects of cycling and running training sessions. This is only a speculation and future research may want to investigate whether RPE includes duration and how this may affect its use for training quantification purposes.

Several studies suggest that different TL metrics are highly intercorrelated (Foster et al. 2001; Manzi et al. 2015; van Erp et al. 2019a). This is somewhat surprising as many metrics use different approaches. It may not be the weighting factor, or the variable used to describe the intensity of a training session, but instead the way intensity is multiplied by duration that leads to this agreement between TL metrics. This suggestion is supported by the similar findings between Chapters 3, 4 and 5 regarding quantifying the effects of various training sessions on TL metrics and APD. In all 3 studies, HR, RPE, power and speed-related TL metrics showed similar responses, yet none of those

metrics were found to be in an agreement with APD. A lack of agreement was also found when the TL metrics were normalized (Chapters 4 and 5). This was done as an attempt to separately examine the effects of exercise intensity and duration.

6.4 Why TL metrics fail: The role of exercise duration

The problematic role of exercise duration within the TL metrics was found for both cycling and running, independently of the method used for measuring exercise intensity (i.e RPE, Speed, HR). Overall, we found that the role of exercise duration was not straightforward, especially when different intensity and duration training sessions are compared. The findings of chapter 3 showed that 5 and 20 min maximal intensity cycling sessions resulted in the same APD. Similarly in chapter 4, where 25 min training sessions resulted in higher TL metric scores despite lower APD responses. These differences in response suggest that exercise duration cannot be used in absolute values and as an independent multiplier in TL metrics. The problem was more obvious for maximal intensity training sessions (Chapters 3 and 4) and less so for sub-maximal training sessions (Chapter 5). Thus the contribution of exercise duration to the development of APD seems to depend on how intense the exercise is (Fullerton et al. 2021). Current results suggest that session duration needs to be considered differently, outside fixed chronological frames and that this varies with exercise intensity. Previous experimental work on time perception, pacing and exercise intensity may provide an interesting insight (Edwards and McCornick 2017). These authors concluded that the higher the intensity of exercise, the more the perception of time appears to shorten in relation to the chronological time elapsed. This is also in line with previous findings that individuals perceive distance to be further in tasks of greater severity (Pinheiro et al. 2016). This indicates that the higher the intensity of exercise, the more the subjective time shortens making chronological time seem slow (Edwards and McCornick 2017). In a practical example, the 5 min (in cycling) or 10 min (in running) training

sessions may have been perceived as ‘’longer’’ for subjects. Accordingly, in terms of perception, absolute measures of duration may not provide a meaningful assessment of training stress if used a multiplier with HR (TRIMPs), RPE (sRPE) or power/speed related metrics (TSS, rTSS). Therefore, when exercise is performed at high intensities, absolute exercise duration may underestimate the training stress imposed and when intensity is low, exercise duration may over present its effects. Although this is a speculation, future research may further explore whether time should be equated differently based on how intense training is and in turn, better inform the TL metrics.

Although, RPE without the effects of duration responded better compared to traditional TL metrics. Findings from Chapters 3 and 4 both suggest that isolating and monitoring exercise intensity without duration may help athletes and practitioners to better quantify their training. A similar conclusion was reported from Weaving et al. (2020), who analyzed the relative contribution of exercise intensity and duration to daily TL in professional rugby players. While RPE was found to be successful for prescribing sessions in Chapters 4 and 5, it agreed with the observed APD only in Chapter 4. The findings from Chapter 5 and others (Fullerton et al. 2021) suggested that the role of exercise duration matters for APD. Therefore, it is not the intensity alone (Chapters 4 and 5), but both intensity and duration (or distance covered) that explains APD. Thus whilst the effects of exercise duration in current TL metrics are over emphasized, excluding exercise duration completely loses an important piece of the puzzle.

Chapter 5 provided some useful insights of how two perceived exertions and durations interacted and induced TL scores and APD’s. As RPE was reduced (from 8 to 5), 4 times higher duration seemed enough to induce a similar APD. When RPE was maximal (Chapter 3), the two training sessions with 4 times higher durations resulted the same training stress in cycling (Figure 3.2). In running and submaximal intensities (Chapter 5), 4 times higher duration was sufficient to result in the same training stress between two sessions with different characteristics. It is also important to state that the role of exercise duration may be related to the form of exercise, modality, fitness levels and the

sequence of different intensities in a training session. For example, the sequence of “load’’ has found to affect performance in elite tour cycling (Padilla et al. 2008), as well as the structure of a training session (Nicoló et al. 2014; Passfield et al. 2017; Renfree et al. 2021). Future research may want to further examine how different combinations between intensity and duration influence APD and TL metrics and how these are influenced by exercise modalities (i.e. running vs. rowing) and different fitness levels (i.e. reactional runners vs elite runners).

6.5 NASA-TLX and Training Load

Whilst APD was found to contradict TL metrics, this was not the case for the NASA-TLX scale in Chapter 3. NASA-TLX is a subjective measurement of perceived workload, yet it does not include absolute exercise duration for these purposes. It includes the effect of duration for the participants via the Temporal Demand subcategory (i.e. How harried or rushed was the pace of the task). However, it does not use this as an independent multiplier. In the studies where maximal or sub-maximal intensities were implemented (Chapter 3 and 4), NASA-TLX agreed well with APD. NASA-TLX was followed the same pattern as APD when comparisons were performed in cycling and running, as well as between CON and INT training sessions (Chapter 4). Previously, subjective methods have been proposed for training quantification purposes. The last experiment (Chapter 5) provided insights on the limitations and usefulness of the NASA-TLX scale however. Participants were asked to exercise at a 5/10 RPE using the CR-10 scale (Borg 1982), and the exercise duration was extended from 10 to 40 min. Although NASA-TLX responded consistently to APD for the 40 min session in Chapter 3 this was not the case with Chapter 5. Collectively, the findings from the 3 studies may suggest that NASA-TLX better corresponds to changes in session intensity, than duration.

Interestingly, NASA-TLX agreed with normalized TL metrics but not with the traditional TL metrics that use exercise duration. The limitations found in terms of using the NASA-TLX may discourage athletes and practitioners for its use. However, it is possible that NASA-TLX is a worth exploring further as a potential TL metric, but at present some limitations are apparent. Accordingly, the use of NASA-TLX may be considered alongside other simple, subjective methods such as RPE. Consequently, the finding from the present studies may help guide future research towards the further identification of simple, inexpensive methods for quantifying the training stress effectively.

6.6 Practical applications

The repeated finding that current TL metrics do not agree with the pattern of APD changes suggests that the quantification of training stress with current methods is misleading athletes and coaches. In a practical context, athletes and coaches have to decide their appropriate training dose by combining session intensity and duration. Traditional methods of endurance training involve such the high-volume training (HVT), threshold training (THR), high-intensity interval training (HIIT) or a combination of the above known as polarized training (Seiler and Kjerland 2006; Stöggl and Sperlich 2014, 2015). Findings from Chapter 3, 4 and 5 suggest that these training intervention cannot be equated with current TL methods. For example, Chapters 3 and 4 showed how potentially the training stress of low intensity-high volume training sessions may be overestimated, or how threshold training and HIIT training sessions may be underestimated with current TL metrics. The ''easy training easy and hard training hard'' principle when interpreted with current metrics could lead coaches and athletes to misjudge how their sessions should be planned. Additionally, the fact that CON and INT sessions with similar RPE's resulted in the same APD may also have important practical implications. Given the training literature, INT has been suggested to potentially cause superior adaptations over CON training (Buchheit and Laursen 2013b).

Similar problems may arise for the use of the acute : chronic workload ratio (ACWR) e.g. for injury prevention purposes (Hulin et al. 2014). This method summates the sRPE responses over an acute (1-week total) and chronic (4-week average) timescale. Chapters 3 and 4 demonstrated the potential problems with daily sRPE values may misguide the training dose. Notably, the findings from Chapter 5 may also provide useful practical information, as two different training sessions with completely different characteristics resulted in the same training stress. This may be important for periodization and injury prevention. Regarding how training is periodized, the results may suggest that rather than accumulating or sustaining TL, sessions with less duration but higher intensity may be as sufficient as needed for triggering the same training effect. For example, when training at maximum intensity, a shorter duration may suffice based upon the lack of difference seen in APD for a 5 and a 20 min session. Alternatively, exercise therapists and injured athletes may want to use less intense training sessions with longer durations to provide training stress, for example in return-to-play protocols.

A simple approach towards a better quantification of TL may be not use the duration of each training session as a multiplier and to report exercise duration and intensity metrics separately. This was suggested in Chapters 3 and 4 and is in agreement with comments by others (Weaving et al. 2020). Nevertheless, the findings from Chapter 5 suggest that athletes and practitioners may also have to consider them in a combined manner, as they interact to generate the training stress. The findings of the thesis reinforce the notion that TWD and related TL metrics do not provide appropriate scores that agree with APD and therefore do not represent the training stress. Sport practitioners and athletes should use current TL metrics with caution, until their validity has been established.

6.7 Limitations

This thesis used some assumptions and limitations that need to be mentioned. To begin with, an assumption is about the APD methodology (Chapter 2.8) and its conceptual link with the TL

concept as introduced by Banister et al. (1975). Earlier, the concept of Banister et al. (1975) was discussed and this thesis evaluates whether the TL should reflect the APD responses resulted from various endurance training sessions. Nevertheless, the model by Banister et al. (1975) is also based on the long-term (i.e weeks, months of training) effects (training-induced fatigue and performance) of training or a block of training sessions. Therefore, it should be clear that Chapters 3, 4 and 5 examine only the acute effects of training (i.e APD). For that reason, the conceptual link made between Banister et al. (1975) work and APD should be used with more reserve and requires further investigation.

Lately, the concept of APD was used to challenge the validity of the TL concept (Passfield et al. 2022). This attempt has received criticism (McLaren et al. 2022; Slattery et al. 2022) and questioned its use for that purpose. Indeed, it is currently unclear whether APD challenges the validity of the TL concept or the validity of the TL metrics. The way the TL concept was introduced by Banister et al. (1975) suggested a link between the amount of training done by an individual and its long- and short-term effects on performance. As said above, this thesis only attempted to link the short-term effects of a training session with performance. Nevertheless, a lack of agreement was found between APD and TL metrics across all studies. These findings challenge the way TL metrics present the effects resulted from single endurance training sessions. This lack of agreement observed across all studies do not necessarily challenge the validity of the TL concept. However, it does raise questions on whether the future of TL monitoring will involve the use of training characteristics in the way they currently used. This thesis provided a novel methodology to challenge traditional methods of TL metrics and highlight its limitations. This will hopefully provide new directions for conceptualizing the TL concept differently and will stimulate the development of new TL metrics.

The link between TL and the training effects (i.e APD), as the only result of the training dose is another assumption that needs to be mentioned. This thesis attempts to conceptually link the training dose with its effect on performance (i.e APD). For that purpose, APD was suggested, due to

its sensitivity to a variety of interventions (Chapter 2.8). Despite that, the APD may not fully describe all the training effects resulted from training. For example, Jeffries et al. (2021) suggested multiple methods of assessing the training effects and classified them as performance measures, physiological measures, subjective measures and others (cognitive, biomechanical). Therefore, it is important to acknowledge that APD may not fully reflect the training effects resulted from the endurance training sessions used in Chapters 3,4 and 5. Indeed, quantifying endurance training and their effects just by a single number was not expected to be straightforward. In order to conceptually link the effects of training with the TL concept, APD findings may need to be presented in relation to other (short and long-term) training effects. Further research is required to provide an understanding on whether APD can be combined with other methods for better understanding the effects of training.

Until further research is conducted, APD could potentially considered as a method to assess the consequence of fatigue deriving from various endurance sessions. Nevertheless, due to the complicated nature of fatigue (Enoka and Duchateau 2016), the exact mechanisms of fatigue have not been examined in depth in the current thesis (see Chapter 2.2). Therefore, this thesis avoided to refer to the nature of fatigue resulted from the endurance training sessions. The absence of examining the fatigue mechanisms in depth may be considered as a limitation and no further justification about the APDs found can be provided. For example, Chapter 3 showed that a 5- and a 20-min TT had a similar APD. If particular mechanisms were examined, this could have helped towards a justification of the findings.

The demanding nature of the current studies should be considered when designing future studies using the APD methodology. The relatively low sample size used in our studies (11-14 participants) was satisfactory for identifying significant differences when the effect sizes were large. However, when looking for subtle differences between different training sessions, a higher sample size may be required. The fact that a conservative post-hoc analysis was used in all studies does not exclude the possibility that a type II error might have been committed. In some specific instances,

coefficient intervals suggested that the recruitment of a larger number of participants might have allowed us to observe other differences between APDs. For example, in Chapter 3, no difference was found in the APDs as assessed after the 20-min sub-max and the 40-min sub-max training sessions. These sessions could have been found different if more participants were recruited. Another example is Chapter 4. In that study, APDs found to be no different as assessed after the 10CON and 25CON training sessions. It cannot be excluded that with more participants, APDs could have been different. These considerations are useful to provide suggestions for future studies that intend to use the APD methodology. When researchers are interested in the potential differences in APD and the expected effect size is small or moderate, larger sample sizes may be needed.

The time frame in which the test was administered in Chapters 3,4 and 5 could have also affected our findings. In Chapter 2.8, the reasons behind the selection of the 5-min rest between the training sessions and APD were explained. If the test was delivered too late after the competition of the training sessions, capturing the effects of that session could be at risk. For that reason, the rest provided (i.e 5-min) between the training sessions and the performance assessment was based on the thesis aims, pilot testing conducted prior to data collection and feasibility. Nonetheless, the exact time that the performance assessment needed to be administered for capturing the training effects of a training sessions remains under question and more evidence is required. Consequently, whether the 5-min of rest between the training sessions and APD was suffice requires further investigation. Potentially, different APD's would have been observed if the rest period between the training sessions and the performance assessment was 3-, 8- or 10-minutes. Future studies may want to investigate how different rest periods provided between training sessions and APD affect how the training effects are represented.

6.8 Future directions

Earlier in this thesis (Chapters 2, 3,4 and 5), we have highlighted some of the challenges of assessing the validity of TL measures across multiple sessions. Although it is relatively simple to design studies assessing the validity of TL metrics at a single session level, only a few attempts have been made (Morton et al. 1990; Wallace et al. 2014b). At a single session level, a useful measure of TL should provide a reasonable estimate of the training stress for a single exercise session for various combinations of exercise intensity and duration. Our findings showed the problems with the current approaches via using performance assessments to capture the training-induced fatigue caused by various endurance-based training sessions. Understanding how exercise intensity and duration affect training stress in a short-term approach may consequently help researchers address the more complicated issues of how to account for changes in TL over multiple training sessions. Therefore, rather than focusing on how training stress affects performance in the long-term, future research may want to first address fundamental concerns around the validity of TL. Hopefully, the findings presented in this thesis may pave the way for the development of new, validated methods to quantify training. Future research may want to examine multiple combinations of training sessions where exercise intensities and durations vary, and to quantify how these training sessions affect APD. Similarly, the selection of different tests for the determination of APD is something that future work may want to investigate. Importantly, future studies should select a test for the assessment of APD based on the sports characteristics and particular events. For instance, a test lasting for a long time (i.e 5 km TT) should not be a choice if the performance of an 800m runner is of concern. Another issue that future research may want to further examine is the type of test used. For example, Fullerton et al. (2021) used a TTE test, while in this thesis only TT's were assessed. Future studies may want to use TTE tests, in order to remove the need for a careful control of pacing. Lately, studies that used the concept of APD preferred a 3 km TT (Vermeire et al. 2023) or a 20-min TT (Valenzuela et al. 2023) assessment. Collectively, future researchers are encouraged to incorporate various training sessions and performance assessments to better understand the training process.

The recovery time provided between the training sessions and the APD is a factor that may want to be further explored. Lately, a 10-min recovery between endurance training sessions and APD was implemented (Vermeire et al. 2023). Researchers are encouraged to try different recovery periods in a systematic way for future studies. Nevertheless, the longer the recovery implemented, the higher the risk to lose the effects of a training session on APD. Moreover, future studies may want to investigate how APD is impacted by other exercise modalities (e.g rowing) and other formats of interval training. For example, Chapter 4 only assessed intervals based on a certain distance, and therefore, APD may want to be examined in other interval matrix where the work and the rest ratios vary. Lastly, it is also suggested that the effect of a training session on an individual's APD may be dependent on their fitness. Therefore, future research may want to characterize how various session intensities and durations result in APD's for active, trained and well-trained endurance athletes.

Future research may want to address the methodological assumptions behind APD, as mentioned in Chapter 6.7. For instance, future research may want to investigate how APD responds in relation to other methods that can capture the short- and long-term effects resulted from training. Training is a complex phenomenon and will result changes in a variety of assessments such as biomechanical, cognitive and subjective (Jeffries et al. 2021). Reductions in performance may partially capture the effects resulted from training and presented as APD in this thesis. Future research may want to relate APD with other measurements as indicators of the acute training effects. A link between APD and other indicators of the acute training effects need to be established. For example, subjective measures of fatigue after a training sessions such as the subcategory of the POMS (Berger and Motl 2000) scale is a direction that future research may want to examine. When that link is established, then a logical step would be to explore the value of APD for long-term training. For that purpose, well-controlled experiments need to be conducted. APD may want to be assessed over multiple times within a period of training, or at the start and end of a training plan. Again, whether APD is a useful tool for long-term training needs to be confirmed versus measurements of the long-

term training effects. Similarly to example made earlier about the use of the fatigue subcategory of POMS for the short-term training effects, mood may also be used for the long-term effects of training. For example, mood assessments after a period of time may well be linked with APD assessments over a period of time. These are just examples of how APD may be used in the future. Researchers and practitioners are encouraged to search for tools that will link APD with other assessments of the training effects resulted from short- and long-term training.

6.9 Conclusion

This thesis investigated the validity of current methods for TL quantification using APD as an index of training stress. Different TL metrics were used to quantify a cycling and a running-based training dose, however, current methods for quantifying TL showed no agreement with APD in any of the experimental studies. These findings emphasize previously expressed concerns regarding the validity of TL concept and the fact that TL metrics have never been validated against an appropriate criterion of training stress. The important role of exercise duration has been discussed and Chapters 3, 4 and 5 further suggested that the use of training duration in absolute units as a multiplier in current TL metrics may be problematic. Therefore, new methods for TL quantification are required. Current findings suggest that these methods may need to take into account both the effects of exercise intensity and duration and their interaction. Current methods for TL quantification in endurance performance should be used with caution, until their validity has been established. It is acknowledged that APD is impractical for daily interpretation of athletes' training sessions. However, the evaluation of APD may be useful in the development of new TL metrics and for assessing their validity.

Reference List

Akubat I, Abt G (2011) Intermittent exercise alters the heart rate-blood lactate relationship used for calculating the training impulse (TRIMP) in team sport players. J Sci Med Sport 14:249–253. https://doi.org/10.1016/j.jsams.2010.12.003

Akubat I, Patel E, Barrett S, Abt G (2012) Methods of monitoring the training and match load and their relationship to changes in fitness in professional youth soccer players. J Sports Sci 30:1473–1480. https://doi.org/10.1080/02640414.2012.712711

Alexiou H, Coutts AJ (2008) A Comparison of Methods Used for Quantifying Internal Training Load in Women Soccer Players. Int J Sports Physiol Perform 3:320–330. https://doi.org/10.1123/ijspp.3.3.320

Allen H, Coggan A (2006) Level II Coaching Manual. In: Training and racing using a power meter: an introduction. Velo Press, Boulder, CO

Amann M (2011) Central and peripheral fatigue: Interaction during cycling exercise in humans. Med Sci Sports Exerc 43:2039–2045. https://doi.org/10.1249/MSS.0b013e31821f59ab

Amann M, Hopkins WG, Marcora SM (2008) Similar sensitivity of time to exhaustion and time-trial time to changes in endurance. Med Sci Sports Exerc 40:574–578. https://doi.org/10.1249/MSS.0b013e31815e728f

Avalos M, Hellard P, Chatard JC (2003) Modeling the training-performance relationship using a mixed model in elite swimmers. Med Sci Sports Exerc 35:838–846. https://doi.org/10.1249/01.MSS.0000065004.05033.42

Banister EW, Calvert, T W, Savage M V, Bach A (1975) A systems model of training for athletic performance. Aust J Sport Med 7:57–61

Banister EW, Calvert TW (1980) Planning for future performance: implications for long term training. Can J Appl Sport Sci 5:170–6

Banister EW, Carter JB, Zarkadas PC (1999) Training theory and taper: Validation in triathlon athletes. Eur J Appl Physiol Occup Physiol 79:182–191. https://doi.org/10.1007/s004210050493

Bannister EW (1991) Physiological testing of elite athletes. In: Macdouggal, J. D. Wenger, H. A. Green HJ (ed) Physiological testing of Elite Athletes. Human Kinetics, Champaign,IL, pp 403–424

Barroso R, Salgueiro DF, Do Carmo EC, Nakamura FY (2015) The effects of training volume and repetition distance on session rating of perceived exertion and internal load in swimmers. Int J Sports Physiol Perform 10:848–852. https://doi.org/10.1123/ijspp.2014-0410

Beaver WL, Wasserman K, Whipp BJ (1986) A new method for detecting anaerobic threshold by gas exchange. J Appl Physiol 60:2020–2027. https://doi.org/10.1152/jappl.1986.60.6.2020

Bentley DJ, Mcnaughton L, Thompson D, et al (2001) Peak power output, the lactate threshold, and time trial performance in cyclists. Med Sci Sport Exerc 33:2077–2081. https://doi.org/10.1097/00005768-200112000-00016

Bentley DJ, Newell J, Bishop D (2007) Incremental exercise test design and analysis: Implications for performance diagnostics in endurance athletes. Sport Med 37:575–586. https://doi.org/10.2165/00007256-200737070-00002

Berger BG, Motl RW (2000) Exercise and mood: A selective review and synthesis of research employing the profile of mood states. J Appl Sport Psychol 12:69–92. https://doi.org/10.1080/10413200008404214

Bishop DJ, Jenkins DG, Mackinnon L (1998) The relationship between plasma lactate parameters, Wpeak and 1-h cycling performance in women. Med Sci Sport Exerc 30:1270–1275. https://doi.org/10.1097/00005768-199808000-00014

Bompa T, Haff GG (2009) Periodization: Theory and Methodology of Training, 5h edn. Human Kinetics, Champaign,IL

Borg G (1982) Psychophysical bases of perceived exertion. Med Sci Sport Exerc 14:377–381. https://doi.org/10.1249/00005768-198205000-00012

Borges TO, Bullock N, Duff C, Coutts AJ (2014) Methods for Quantifying Training in Sprint Kayak. J Strength Cond Res 28:474–482. https://doi.org/10.1519/JSC.0b013e31829b56c4

Borresen J, Lambert M (2009) The quantification of training load, the training response and the effect on performance. Sport Med 39:779–795. https://doi.org/10.2165/11317780-000000000-00000

Borresen J, Lambert MI (2008) Quantifying Training Load: A Comparison of Subjective and Objective Methods. Int J Sports Physiol Perform 3:16–30. https://doi.org/10.1123/ijspp.3.1.16

Bourdon PC, Cardinale M, Murray A, et al (2017) Monitoring athlete training loads: Consensus statement. Int J Sports Physiol Perform 12:161–170. https://doi.org/10.1123/IJSPP.2017-0208

Bradley WJ, Cavanagh BP, Douglas W, et al (2015) Quantification of Training Load, Energy Intake, and Physiological Adaptations During a Rugby Preseason. J Strength Cond Res 29:534–544. https://doi.org/10.1519/JSC.0000000000000631

Brink MS, Nederhof E, Visscher C, et al (2010) Monitoring Load, Recovery, and Performance in Young Elite Soccer Players. J Strength Cond Res 24:597–603. https://doi.org/10.1519/JSC.0b013e3181c4d38b

Buchheit M (2014) Monitoring training status with HR measures: Do all roads lead to Rome? Front Physiol 5:1–19. https://doi.org/10.3389/fphys.2014.00073

Buchheit M, Laursen PB (2013a) High-intensity interval training, solutions to the programming puzzle: Part II: Anaerobic energy, neuromuscular load and practical applications. Sport Med 43:927–954. https://doi.org/10.1007/s40279-013-0066-5

Buchheit M, Laursen PB (2013b) High-intensity interval training, solutions to the programming puzzle: Part I: Cardiopulmonary emphasis. Sport Med 43:313–338. https://doi.org/10.1007/s40279-013-0029-x

Busso T (2003) Variable dose-response relationship between exercise training and performance. Med Sci Sports Exerc 35:1188–1195. https://doi.org/10.1249/01.MSS.0000074465.13621.37

Busso T, Carasso C, Lacour JR (1991) Adequacy of a systems structure in the modeling of training effects on performance. J Appl Physiol 71:2044–2049. https://doi.org/10.1152/jappl.1991.71.5.2044

Busso T, Denis C, Bonnefoy R, et al (1997) Modeling of adaptations to physical training by using a recursive least squares algorithm. J Appl Physiol 82:1685–1693. https://doi.org/10.1152/jappl.1997.82.5.1685

Busso T, Häkkinen K, Pakarinen A, et al (1990) A systems model of training responses and its

relationship to hormonal responses in elite weight-lifters. Eur J Appl Physiol Occup Physiol 61:48–54. https://doi.org/10.1007/BF00236693

Busso T, Thomas L (2006) Using mathematical modeling in training planning. Int J Sports Physiol Perform 1:400–405. https://doi.org/10.1123/ijspp.1.4.400

Calvert TW, Banister EW, Savage M V., Bach T (1976) A systems model of training for athletic performance. Aust J Sport Med SMC-6:94–102

Casado A, Moreno-Pérez D, Larrosa M, Renfree A (2019) Different psychophysiological responses to a high-intensity repetition session performed alone or in a group by elite middle-distance runners. Eur J Sport Sci 19:1045–1052. https://doi.org/10.1080/17461391.2019.1593510

Christen J, Foster C, Porcari JP, Mikat RP (2016) Temporal robustness of the session rating of perceived exertion. Int J Sports Physiol Perform 11:1088–1093. https://doi.org/10.1123/ijspp.2015-0438

Clark I, Vanhatalo A, Bailey S, et al (2018) Effects of Two Hours of Heavy-Intensity Exercise on the Power–Duration Relationship. Med Sci Sport Exerc 50:1658–1668. https://doi.org/10.1249/MSS.0000000000001601

Cohen J (1988) Statistical power analysis for the behavioural sciences, 2nd edn. Hillsdale, NJ, USA

Coyle EF, Coggan AR, Hopper MK, Walters TJ (1988) Determinants of endurance in well-trained cyclists. J Appl Physiol 64:2622–2630. https://doi.org/10.1152/jappl.1988.64.6.2622

Coyne JOC, Gregory Haff G, Coutts AJ, et al (2018) The current state of subjective training load monitoring—A practical perspective and call to action. Sport Med Int Open 4:. https://doi.org/10.1186/s40798-018-0172-x

Currell K, Jeukendrup A (2013) Validity, Reliability and Sensitivity of Measures of Sporting Performance. Sport Med 43:297–316. https://doi.org/10.2165/00007256-200838040-00003

Cutsem J Van, Pauw K De, Buyse L, et al (2017) Effects of Mental Fatigue on Endurance Performance in the Heat. Med Sci Sports Exerc 49:1677–1687

Daniels JT, Yarbrough RA, Foster C (1978) Changes in VO2 max and Running Performance with Training. Eur J Appl Physiol Occup Physiol 39:249–254. https://doi.org/10.1007/BF00421448

Davies CTM, Knibbs A V. (1971) The training stimulus - The effects of intensity, duration and frequency of effort on maximum aerobic power output. Int Zeitschrift für Angew Physiol Einschließlich Arbeitsphysiologie 29:299–305. https://doi.org/10.1007/BF00697733

Davies MJ, Benson AP, Cannon DT, et al (2017) Dissociating external power from intramuscular exercise intensity during intermittent bilateral knee-extension in humans. J Physiol 595:6673–6686. https://doi.org/10.1113/JP274589

Davis L, Appleby R, Davis P, et al (2018) The role of coach-athlete relationship quality in team sport athletes' psychophysiological exhaustion: implications for physical and cognitive performance. J Sports Sci 36:1985–1992. https://doi.org/10.1080/02640414.2018.1429176

de Koning JJ, Foster C, Bakkum A, et al (2011) Regulation of pacing strategy during athletic competition. PLoS One 6:2–7. https://doi.org/10.1371/journal.pone.0015863

Drew MK, Raysmith BP, Charlton PC (2017) Injuries impair the chance of successful performance by sportspeople: A systematic review. Br J Sports Med 51:1209–1214. https://doi.org/10.1136/bjsports-2016-096731

Durnin J V., Brockway JM, Whitcher W. (1960) Effects of a short period of training of varying severity on some measurements of physical fitness. J Appl Physiol 15:161–165. https://doi.org/10.1152/jappl.1960.15.1.161

Earnest CP, Jurca R, Church TS, et al (2004) Relation between physical exertion and heart rate variability characteristics in professional cyclists during the Tour of Spain. Br J Sports Med 38:568–575. https://doi.org/10.1136/bjsm.2003.005140

Eckard TG, Padua DA, Hearn DW, et al (2018) The Relationship Between Training Load and Injury in Athletes: A Systematic Review. Springer International Publishing

Edwards A, McCornick A (2017) Time perception, pacing and exercise intensity: maximal exercise distorts the perception of time. Physiol Behav 180:98–102. https://doi.org/10.1177/0309133309346882

Edwards S (1993) The Heart Rate Monitor Book. Fleet Feet Press, Sacramaneto

Emig T, Peltonen J (2020) Human running performance from real-world big data. Nat Commun 11:1–9. https://doi.org/10.1038/s41467-020-18737-6

Enoka R, Duchateau J (2016) Translating fatigue to human performance. Med Sci Sports Exerc 48:2228–2238. https://doi.org/10.1249/MSS.0000000000000929

Esteve-Lanao J, San Juan AF, Earnest CP, et al (2005) How do endurance runners actually train? Relationship with competition performance. Med Sci Sports Exerc 37:496–504. https://doi.org/10.1249/01.MSS.0000155393.78744.86

Fieo RA, Mortensen EL, Rantanen T, Avlund K (2013) Improving a measure of mobility-related fatigue (The Mobility-Tiredness Scale) by establishing item intensity. J Am Geriatr Soc 61:429–433. https://doi.org/10.1111/jgs.12122

Foster C, Daines E, Hector L, et al (1996) Athletic performance in relation to training load. Wis Med J 95:370–4

Foster C, Florhaug JA, Franklin J, et al (2001) A new approach to monitoring exercise training. J strength Cond Res 15:109–15

Foster C, Hector LL, Welsh R, et al (1995) Effects of specific versus cross-training on running performance. Eur J Appl Physiol Occup Physiol 70:367–372. https://doi.org/10.1007/BF00865035

Fry R, Morton A, Keast D (1992) Periodisation of training stress-a review. Can J Sport Sci 17:234–240

Fry RW, Morton AR, Keast D (1991) Overtraining in Athletes. Sport Med 12:32–65. https://doi.org/10.2165/00007256-199112010-00004

Fullerton MM, Passfield L, MacInnis MJ, et al (2021) Prior exercise impairs subsequent performance in an intensity- and duration-dependent manner. Appl Physiol Nutr Metab 46:976–985. https://doi.org/10.1139/apnm-2020-0689

Fusco A, Knutson C, King C, et al (2020) Session RPE During Prolonged Exercise Training. Int J Sports Physiol Perform 15:292–294. https://doi.org/10.1123/ijspp.2019-0137

Gardner AS, Stephens S, Martin DT, et al (2004) Accuracy of SRM and Power Tap power monitoring systems for bicycling. Med Sci Sports Exerc 36:1252–1258. https://doi.org/10.1249/01.MSS.0000132380.21785.03

Gibala MJ, Little JP, Macdonald MJ, Hawley JA (2012) Physiological adaptations to low-volume, high-intensity interval training in health and disease. J Physiol 590:1077–1084. https://doi.org/10.1113/jphysiol.2011.224725

González-Alonso J, Teller C, Andersen SL, et al (1999) Influence of body temperature on the development of fatigue during prolonged exercise in the heat. J Appl Physiol 86:1032–9. https://doi.org/10.1152/jappl.1999.86.3.1032

Green HJ, Hughson RL, Orr GW, Ranney DA (1983) Anaerobic threshold, blood lactate, and muscle metabolites in progressive exercise. J Appl Physiol 54:1032–1038. https://doi.org/10.1152/jappl.1983.54.4.1032

Green JM, McIntosh JR, Hornsby J, et al (2009) Effect of exercise duration on session RPE at an individualized constant workload. Eur J Appl Physiol 107:501–507. https://doi.org/10.1007/s00421-009-1153-z

Halson SL (2014) Monitoring Training Load to Understand Fatigue in Athletes. Sport Med 44:139–147. https://doi.org/10.1007/s40279-014-0253-z

Hart SG (2012) Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proc Hum Factors Ergon Soc Annu Meet 50:904–908. https://doi.org/10.1177/154193120605000909

Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index). Adv Psychol 52:139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hawley JA, Myburgh KH, Noakes TD, Dennis SC (1997) Training techniques to improve fatigue resistance and enhance endurance performance. J Sports Sci 15:325–333. https://doi.org/10.1080/026404197367335

Hawley JA, Noakes TD (1992) Peak power output predicts maximal oxygen uptake and performance time in trained cyclists. Eur J Appl Physiol Occup Physiol 65:79–83. https://doi.org/10.1007/BF01466278

Heigenhauser GJF, Sutton JR, Jones NL (1983) Effect of glycogen depletion on the ventilatory response to exercise. J Appl Physiol Respir Environ Exerc Physiol 54:470–474. https://doi.org/10.1152/jappl.1983.54.2.470

Helgerud J, Høydal K, Wang E, et al (2007) Aerobic high-intensity intervals improve V̇O2max more than moderate training. Med Sci Sports Exerc 39:665–671. https://doi.org/10.1249/mss.0b013e3180304570

Hellard P, Avalos M, Lacoste L, et al (2006) Assessing the limitations of the Banister model in monitoring training. J Sports Sci 24:509–520. https://doi.org/10.1080/02640410500244697

Hellard P, Avalos M, Millet G, et al (2005) Modeling the Residual Effects and Threshold Saturation of Training: A Case Study of Olympic Swimmers. J Strength Cond Res 19:67. https://doi.org/10.1519/14853.1

Helms ER, Byrnes RK, Cooke DM, et al (2018) RPE vs. Percentage 1RM loading in periodized programs matched for sets and repetitions. Front Physiol 9:1–10. https://doi.org/10.3389/fphys.2018.00247

Herman L, Foster C, Maher M, et al (2006) Validity and reliability of the session RPE method for monitoring exercise training intensity. South African J Sport Med 18:14. https://doi.org/10.17159/2078-516x/2006/v18i1a247

Hofmann P, Tschakert G (2017) Intensity- and duration-based options to regulate endurance training. Front Physiol 8:1–9. https://doi.org/10.3389/fphys.2017.00337

Hulin BT, Gabbett TJ, Blanch P, et al (2014) Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. Br J Sports Med 48:708–712. https://doi.org/10.1136/bjsports-2013-092524

Iannetta D, Inglis E, Mattu A, et al (2020) A Critical Evaluation of Current Methods for Exercise Prescription in Women and Men. Med Sci Sport Exerc 52:466–473. https://doi.org/10.1249/MSS.0000000000002147

Iannetta D, Inglis EC, Fullerton C, et al (2018) Metabolic and performance-related consequences of exercising at and slightly above MLSS. Scand J Med Sci Sport 28:2481–2493. https://doi.org/10.1111/sms.13280

Impellizzeri F, Marcora SM, Coutts AJ (2019) Internal and External Training Load: 15 Years On. Int J Sports Physiol Perform 14:270–273

Impellizzeri FM, Rampinini E, Coutts AJ, et al (2004) Use of RPE-based training load in soccer. Med Sci Sports Exerc 36:1042–1047. https://doi.org/10.1249/01.MSS.0000128199.23901.2F

Impellizzeri FM, Rampinini E, Marcora SM (2005) Physiological assessment of aerobic training in soccer. J Sports Sci 23:583–592. https://doi.org/10.1080/02640410400021278

Issurin V (2010) New Horizons for the Methodology and Physiology of Training Periodization. Sport Med 40:189–206. https://doi.org/10.2165/11535130-000000000-00000

Jaspers A, Brink MS, Probst SGM, et al (2017) Relationships Between Training Load Indicators and Training Outcomes in Professional Soccer. Sport Med 47:533–544. https://doi.org/10.1007/s40279-016-0591-0

Jeffries AC, Marcora SM, Coutts AJ, et al (2021) Development of a Revised Conceptual Framework of Physical Training for Use in Research and Practice. Sport Med. https://doi.org/10.1007/s40279-021-01551-5

Jeffries AC, Wallace L, Coutts AJ, et al (2020) Athlete-Reported Outcome Measures for Monitoring Training Responses: A Systematic Review of Risk of Bias and Measurement Property Quality According to the COSMIN Guidelines. Int J Sports Physiol Perform 15:1203–1215. https://doi.org/10.1123/ijspp.2020-0386

Jeukendrup AE (2017) Periodized Nutrition for Athletes. Sport Med 47:51–63. https://doi.org/10.1007/s40279-017-0694-2

Jobson SA, Passfield L, Atkinson G, et al (2009) The Analysis and Utilization of Cycling Training Data. Sport Med 39:833–844. https://doi.org/10.2165/11317840-000000000-00000

Jones AM (2002) Bioenergetic constraints on tactical decision making in middle distance running. Br J Sports Med 36:102–104. https://doi.org/10.1136/bjsm.36.2.102

Jones AM, Carter H (2000) The effect of endurance training on parameters of aerobic fitness. Sport Med 29:373–386. https://doi.org/10.2165/00007256-200029060-00001

Kellmann M, Kallus KW (2001) Recovery-stress questionnaire for athletes: User manual.

Kenttä G, Hassmén P (1998) Overtraining and recovery. A conceptual model. Sport Med 26:1–16. https://doi.org/10.2165/00007256-199826010-00001

Kesisoglou A, Nicolò A, Howland L, Passfield L (2021a) Continuous Versus Intermittent Running: Acute Performance Decrement and Training Load. Int J Sports Physiol Perform 1–10. https://doi.org/10.1123/ijspp.2020-0844

Kesisoglou A, Nicolò A, Passfield L (2021b) Cycling Performance and Training Load : Effects of Intensity and Duration. Int J Sports Physiol Perform 16:. https://doi.org/doi.org/10.1123/ijspp.2020-0072

Kluger BM, Krupp LB, Enoka RM (2013) Fatigue and fatigability in neurologic illnesses: Proposal for an unified taxonomy. Neurology 80:409–416. https://doi.org/10.1212/WNL.0b013e31827f07be

Knuttgen HG (1978) Force, work, power, and exercise. Med Sci Sports 10:227–228

Koglyn L KB (2013) Control and sensation of breathing during cycling exercise in hypoxia under naloxone: a randomised controlled crossover trial. Extrem Physiol Med 2:1–10. https://doi.org/10.1016/j.rmed.2005.02.005

Kuipers H, Keizer HA (1988) Overtraining in Elite Athletes: Review and Directions for the Future. Sport Med An Int J Appl Med Sci Sport Exerc 6:79–92. https://doi.org/10.2165/00007256-198806020-00003

Lehmann M, Foster C, Joseph K (1993) Overtraining in endurance athletes: a brief review. Med Sci Sport Exerc 25:854–862. https://doi.org/10.1249/00005768-199307000-00015

Lin D, Potiaumpai M, Schmitz K, Sturgeon K (2021) Increased Duration of Exercise Decreases Rate of Nonresponse to Exercise but May Not Decrease Risk for Cancer Mortality. Med Sci Sports Exerc 53:928–935. https://doi.org/10.1249/MSS.0000000000002539

Lovell TWJ, Sirotic AC, Impellizzeri FM, Coutts AJ (2013) Factors affecting perception of effort (session rating of perceived exertion) during rugby league training. Int J Sports Physiol Perform 8:62–69. https://doi.org/10.1123/ijspp.8.1.62

Lucía A, Hoyos J, Santalla A, et al (2003) Tour de France versus Vuelta a España: Which is harder? Med Sci Sports Exerc 35:872–878. https://doi.org/10.1249/01.MSS.0000064999.82036.B4

Ludwig M, Asteroth A, Rasche C, Pfeiffer M (2019) Including the Past: Performance Modeling Using a Preload Concept by Means of the Fitness-Fatigue Model. Int J Comput Sci Sport 18:115–134. https://doi.org/10.2478/ijcss-2019-0007

Lurquin JH, Michaelson LE, Barker JE, et al (2016) No evidence of the ego-depletion effect across task characteristics and individual differences: A pre-registered study. PLoS One 11:1–20. https://doi.org/10.1371/journal.pone.0147770

Lutttikohlt H, McNaughton LR, Midgley AW, Bentley DJ (2006) A prediction model for peak power output from different incremental exercise tests. Int J Sports Physiol Perform 1:122–136. https://doi.org/10.1123/ijspp.1.2.122

MacInnis MJ, Gibala MJ (2017) Physiological adaptations to interval training and the role of exercise intensity. J Physiol 595:2915–2930. https://doi.org/10.1113/JP273196

MacInnis MJ, Thomas ACQ, Phillips SM (2018) The Reliability of 4-Minute and 20-Minute Time Trials and Their Relationships to Functional Threshold Power in Trained Cyclists. Int J Sports Physiol Perform 14:38–45. https://doi.org/10.1123/ijspp.2018-0100

MacInnis MJ, Zacharewicz E, Martin BJ, et al (2017) Superior mitochondrial adaptations in human

skeletal muscle after interval compared to continuous single-leg cycling matched for total work. J Physiol 595:2955–2968. https://doi.org/10.1113/JP272570

Malone S, Doran D, Akubat I, Collins K (2016) The Integration of Internal and External Training Load Metrics in Hurling. J Hum Kinet 53:211–221. https://doi.org/10.1515/hukin-2016-0024

Mann T, Lamberts RP, Lambert MI (2013) Methods of prescribing relative exercise intensity: Physiological and practical considerations. Sport Med 43:613–625. https://doi.org/10.1007/s40279-013-0045-x

Mann TN, Lamberts RP, Lambert MI (2014) High responders and low responders: Factors associated with individual variation in response to standardized training. Sport Med 44:1113–1124. https://doi.org/10.1007/s40279-014-0197-3

Manoel F, da Silva D, Lima J, Machado F (2017) Peak velocity and its time limit are as good as the velocity associated with VO2max for training prescription in runners. Sport Med Int Open 01:E8–E15. https://doi.org/10.1055/s-0042-119951

Manzi V, Bovenzi A, Castagna C, et al (2015) Training-load distribution in endurance runners: Objective versus subjective assessment. Int J Sports Physiol Perform 10:1023–1028. https://doi.org/10.1123/ijspp.2014-0557

Manzi V, Bovenzi A, Impellizzeri MF, et al (2013) Individual training-load and aerobic-fitness variables in premiership soccer players during the precompetitive season. J Strength Cond Res 27:631–636. https://doi.org/10.1519/JSC.0b013e31825dbd81

Manzi V, Castagna C, Padua E, et al (2009a) Dose-response relationship of autonomic nervous system responses to individualized training impulse in marathon runners. Am J Physiol - Hear Circ Physiol 296:. https://doi.org/10.1152/ajpheart.00054.2009

Manzi V, D'ottavio S, Impellizzeri FM, et al (2010) Profile of weekly training load in elite male professional basketball players. J Strength Cond Res 24:1399–1406. https://doi.org/10.1519/JSC.0b013e3181d7552a

Manzi V, Iellamo F, Impellizzeri F, et al (2009b) Relation between individualized training impulses and performance in distance runners. Med Sci Sports Exerc 41:2090–2096. https://doi.org/10.1249/MSS.0b013e3181a6a959

Marcora SM, Bosio A, de Morree HM (2008) Locomotor muscle fatigue increases cardiorespiratory responses and reduces performance during intense cycling exercise independently from metabolic stress. Am J Physiol Integr Comp Physiol 294:R874–R883. https://doi.org/10.1152/ajpregu.00678.2007

Marcora SM, Staiano W, Manning V (2009) Mental fatigue impairs physical performance in humans. J Appl Physiol 106:857–864. https://doi.org/10.1152/japplphysiol.91324.2008

Marino FE, Gard M, Drinkwater EJ (2011) The limits to exercise performance and the future of fatigue research. Br J Sports Med 45:65–67. https://doi.org/10.1136/bjsm.2009.067611

Martin K, Staiano W, Menaspà P, et al (2016) Superior inhibitory control and resistance to mental fatigue in professional road cyclists. PLoS One 11:1–15. https://doi.org/10.1371/journal.pone.0159907

McGregor SJ, Weese RK, Ratz IK (2009) Performance Modeling in an Olympic 1500-m Finalist: A Practical Approach. J Strength Cond Res 23:2515–2523. https://doi.org/10.1519/JSC.0b013e3181bf88be

McLaren SJ, Macpherson TW, Coutts AJ, et al (2018) The Relationships Between Internal and External Measures of Training Load and Intensity in Team Sports: A Meta-Analysis

McLaren SJ, Tzlil S, Schneider C, Ward P (2022) Comment on Passfield et al: Validity of the Training-Load Concept. Int J Sports Physiol Perform 17:1457

Mehta RK, Parasuraman R (2014) Effects of mental fatigue on the development of physical fatigue: A neuroergonomic approach. Hum Factors 56:645–656. https://doi.org/10.1177/0018720813507279

Micklewright D, St Clair Gibson A, Gladwell V, Al Salman A (2017) Development and Validity of the Rating-of-Fatigue Scale. Sport Med 47:2375–2393. https://doi.org/10.1007/s40279-017-0711-5

Millet GP, Candau RB, Barbler B, et al (2002) Modelling the transfers of training effects on performance in elite triathletes. Int J Sports Med 23:55–63. https://doi.org/10.1055/s-2002-19276

Millet GP, Vleck VE, Bentley DJ (2009) Physiological differences between cycling and running: Lessons from triathletes. Sport Med 39:179–206. https://doi.org/10.2165/00007256-200939030-00002

Morgan WP, Brown DR, Raglin JS, et al (1987) Psychological monitoring of overtraining and staleness. Br J Sports Med 21:107–114. https://doi.org/10.1136/bjsm.21.3.107

Morton RH (1997) Modelling training and overtraining. J Sports Sci 15:335–340. https://doi.org/10.1080/026404197367344

Morton RH, Fitz-Clarke JR, Banister EW (1990) Modeling human performance in running. J Appl Physiol 69:1171–1177. https://doi.org/10.1152/jappl.1990.69.3.1171

Mujika, Iñigo, Busso T, Lacoste L, et al (1996) Modeled responses to training and taper in competitive swimmers. Med Sci Sport Exerc 28:251–258. https://doi.org/10.1097/00005768-199602000-00015

Mujika I (2017) Quantification of training and competition loads in endurance sports: Methods and Applications. Int J Sports Physiol Perform 12:2–17. https://doi.org/10.4324/9781315268149

Muñoz I, Cejuela R, Seiler S, et al (2014a) Training-intensity distribution during an ironman season: Relationship with competition performance. Int J Sports Physiol Perform 9:332–339. https://doi.org/10.1123/IJSPP.2012-0352

Muñoz I, Seiler S, Bautista J, et al (2014b) Does polarized training improve performance in recreational runners? Int J Sports Physiol Perform 9:265–272. https://doi.org/10.1123/IJSPP.2012-0350

Nicoló A, Bazzucchi I, Haxhi J, et al (2014) Comparing continuous and intermittent exercise: An "isoeffort" and "isotime" approach. PLoS One 9:. https://doi.org/10.1371/journal.pone.0094990

Nicolò A, Bazzucchi I, Lenti M, et al (2014) Neuromuscular and metabolic responses to high-intensity intermittent cycling protocols with different work-to-rest ratios. Int J Sports Physiol Perform 9:151–160. https://doi.org/10.1123/IJSPP.2012-0289

Nicolò A, Girardi M (2016) The physiology of interval training: a new target to HIIT. J Physiol 594:7169–7170. https://doi.org/10.1113/JP273466

Nicolò A, Marcora SM, Bazzucchi I, Sacchetti M (2017) Differential control of respiratory frequency and tidal volume during high-intensity interval training. Exp Physiol 102:934–949. https://doi.org/10.1113/EP086352

Nicolò A, Marcora SM, Sacchetti M (2016a) Respiratory frequency is strongly associated with perceived exertion during time trials of different duration. J Sports Sci 34:1199–1206. https://doi.org/10.1080/02640414.2015.1102315

Nicolò A, Montini M, Girardi M, et al (2019) Respiratory Frequency as a Marker of Physical Effort During High-Intensity Interval Training in Soccer Players. Int J Sports Physiol Perform 15:73–80

Nicolò A, Passfield L, Sacchetti M (2016b) Investigating the effect of exercise duration on functional and biochemical perturbations in the human heart: Total work or "isoeffort" matching? J Physiol 594:3157–3158. https://doi.org/10.1113/JP272421

Noakes TD (2004) Letter To The Editor: "linear relationship between the perception of effort and the duration of constant load exercise that remains'''." J Appl Physiol 96:1571–1573. https://doi.org/10.1080/13518040701205365

Noakes TD, Myburgh KH, Schall R (1990) Peak treadmill running velocity during the vo2 max test predicts running performance. J Sports Sci 8:35–45. https://doi.org/10.1080/02640419008732129

O'Grady C, Passfield L, Hopker JG (2021) Variability in Submaximal Self-Paced Exercise Bouts of Different Intensity and Duration. Int J Sports Physiol Perform 16:1824–1833. https://doi.org/10.1123/ijspp.2020-0785

Padilla S, Mujika I, Orbañanos J, et al (2001) Exercise intensity and load during mass-start stage races in professional road cycling. Med Sci Sports Exerc 33:796–802. https://doi.org/10.1097/00005768-200105000-00019

Padilla S, Mujika I, Orbañanos J, Angulo F (2000) Exercise intensity during competition time trials in professional road cycling. Med Sci Sports Exerc 32:850–856. https://doi.org/10.1097/00005768-200004000-00019

Padilla S, Mujika I, Santisteban J, et al (2008) Exercise intensity and load during uphill cycling in professional 3-week races. Eur J Appl Physiol 102:431–438. https://doi.org/10.1007/s00421-007-0602-9

Pageaux B, Lepers R (2016) Fatigue induced by physical and mental exertion increases perception of effort and impairs subsequent endurance performance. Front Physiol 7:. https://doi.org/10.3389/fphys.2016.00587

Pageaux B, Lepers R, Dietz KC, Marcora SM (2014) Response inhibition impairs subsequent self-paced endurance performance. Eur J Appl Physiol 114:1095–1105. https://doi.org/10.1007/s00421-014-2838-5

Pageaux B, Marcora SM, Rozand V, Lepers R (2015) Mental fatigue induced by prolonged self-regulation does not exacerbate central fatigue during subsequent whole-body endurance exercise. Front Hum Neurosci 9:1–12. https://doi.org/10.3389/fnhum.2015.00067

Paquette MR, Napier C, Willy RW, Stellingwerff T (2020) Moving beyond weekly "distance": Optimizing quantification of training load in runners. J Orthop Sports Phys Ther 50:564–569. https://doi.org/10.2519/jospt.2020.9533

Passfield L, Doust JH (2000) Changes in cycling efficiency and performance after endurance exercise. Med Sci Sports Exerc 32:1935–1941. https://doi.org/10.1097/00005768-200011000-00018

Passfield L, Hopker JG, Jobson S, et al (2017) Knowledge is power: Issues of measuring training and performance in cycling. J Sports Sci 35:1426–1434. https://doi.org/10.1080/02640414.2016.1215504

Passfield L, Murias JM, Sacchetti M, Nicolò A (2022) Validity of the Training-Load Concept. Int J Sports Physiol Perform 507–514

Paton CD, Hopkins WG (2001) Tests of cycling performance. Sport Med 31:489–496. https://doi.org/10.2165/00007256-200131070-00004

Peake JM, Neubauer O, Gatta PAD, Nosaka K (2017) Muscle damage and inflammation during recovery from exercise. J Appl Physiol 122:559–570. https://doi.org/10.1152/japplphysiol.00971.2016

Peronnet, F., Massicotte D (1991) Table of nonprotein respiratory quotient : an update. Can J Sport Sci 16:23–29

Pinheiro FA, Santos TM, Pires FO (2016) Conscious distance monitoring and perceived exertion in light-deprived cycling time trial. Physiol Behav 165:211–216. https://doi.org/10.1016/j.physbeh.2016.07.020

Pinot J, Grappe F (2015) A six-year monitoring case study of a top-10 cycling Grand Tour finisher. J Sports Sci 33:907–914. https://doi.org/10.1080/02640414.2014.969296

Renfree A, Casado A, McLaren S (2021) Re-thinking athlete training loads: would you rather have one big rock or lots of little rocks dropped on your foot? Res Sport Med 00:1–4. https://doi.org/10.1080/15438627.2021.1906672

Rodgers MM, Cavanagh PR (1984) Glossary of Biomechanical Terms, Concepts, and Units. Phys Ther 64:1886–1902. https://doi.org/10.1093/ptj/64.12.1886

Rodríguez-Marroyo JA, García-López J, Juneau CÉ, Villa JG (2009) Workload demands in professional multi-stage cycling races of varying duration. Br J Sports Med 43:180–185. https://doi.org/10.1136/bjsm.2007.043125

Rodríguez-Marroyo JA, Villa G, García-Lopez J, Foster C (2012) Comparison of heart rate and session rating of perceived exertion methods of defining exercise load in cyclists. J Strength Cond Res 26:2249–2257. https://doi.org/10.1519/JSC.0b013e31823a4233

Rodríguez-Marroyo JA, Villa JG, Pernía R, Foster C (2017) Decrement in professional cyclists' performance after a grand tour. Int J Sports Physiol Perform 12:1348–1355. https://doi.org/10.1123/ijspp.2016-0294

Rønnestad BR, Hansen J, Nygaard H, Lundby C (2020) Superior performance improvements in elite cyclists following short-interval vs effort-matched long-interval training. Scand J Med Sci Sport 30:849–857. https://doi.org/10.1111/sms.13627

Ross R, De Lannoy L, Stotz PJ (2015) Separate effects of intensity and amount of exercise on interindividual cardiorespiratory fitness response. Mayo Clin Proc 90:1506–1514. https://doi.org/10.1016/j.mayocp.2015.07.024

Rozand V, Pageaux B, Marcora SM, et al (2014) Does mental exertion alter maximal muscle

activation? Front Hum Neurosci 8:1–10. https://doi.org/10.3389/fnhum.2014.00755

Rushall BS (1990) A tool for measuring stress tolerance in elite athletes. J Appl Sport Psychol 2:51–66. https://doi.org/10.1080/10413209008406420

Russell AP, Lamon S, Boon H, et al (2013) Regulation of miRNAs in human skeletal muscle following acute endurance exercise and short-term endurance training. J Physiol 591:4637–4653. https://doi.org/10.1113/jphysiol.2013.255695

Sandbakk Ø, Haugen T, Ettema G (2021) The Influence of Exercise Modality on Training Load Management. Int J Sports Physiol Perform 16:605–608. https://doi.org/10.1123/ijspp.2021-0022

Sanders D, Abt G, Hesselink MKC, et al (2017a) Methods of Monitoring Training Load and Their Relationships to Changes in Fitness and Performance in Competitive Road Cyclists. Int J Sports Physiol Perform 12:668–675. https://doi.org/10.1123/ijspp.2016-0454

Sanders D, Myers T, Akubat I (2017b) Training-Intensity Distribution in Road Cyclists: Objective Versus Subjective Measures. Int J Sports Physiol Perform 12:1232–1237. https://doi.org/10.1123/ijspp.2016-0523

Saw AE, Main LC, Gastin PB (2016) Monitoring the athlete training response: subjective self-reported measures trump commonly used objective measures: a systematic review. Br J Sports Med 50:281–291. https://doi.org/10.1136/bjsports-2015-094758

Scott BR, Lockie RG, Knight TJ, et al (2013a) A comparison of methods to quantify the in-season training load of professional soccer players. Int J Sports Physiol Perform 8:195–202. https://doi.org/10.1123/ijspp.8.2.195

Scott TJ, Black CR, Quinn J, Coutts AJ (2013b) Validity and reliability of the session-rpe method forquantifying training in australian football: A comparison of the cr10 and cr100 scales. J Strength Cond Res 27:270–276. https://doi.org/10.1519/JSC.0b013e3182541d2e

Seiler KS, Kjerland GØ (2006) Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an "optimal" distribution? Scand J Med Sci Sport 16:49–56. https://doi.org/10.1111/j.1600-0838.2004.00418.x

Seiler S (2010) What is best practice for training intensity and duration distribution in endurance athletes? Int J Sports Physiol Perform 5:276–291. https://doi.org/10.1123/ijspp.5.3.276

Seiler S, Jøranson K, Olesen B V., Hetlelid KJ (2013) Adaptations to aerobic interval training: Interactive effects of exercise intensity and total work duration. Scand J Med Sci Sport 23:74–83. https://doi.org/10.1111/j.1600-0838.2011.01351.x

Seiler S, Sjursen JE (2004) Effect of work duration on physiological and rating scale of perceived exertion responses during self-paced interval training. Scand J Med Sci Sport 14:318–325. https://doi.org/10.1046/j.1600-0838.2003.00353.x

Shephard RJ (1968) Intensity, duration and frequency of exercise as determinants of the response to a training regime. Int Zeitschrift für Angew Physiol Einschl Arbeitsphysiologie 26:272–278. https://doi.org/10.1007/BF00695115

Skiba P (2006) Calculation of Power Output and Quantification of Training Stress in Distance Runners: The Development of the GOVSS Algorithm

Slattery KM, Wallace LK, Coutts AJ (2022) Comment on Passfield et al: Defending the Use of

Oxygen Uptake as a Criterion Measure for Training Load. Int J Sports Physiol Perform 17:1458–1459. https://doi.org/10.1123/ijspp.4.2.269

Smith D (2003) A Framework for Understanding the Training Process Leading to Elite Performance. Sport Med 33:1103–1126. https://doi.org/10.4135/9781452283012.n121

Smith M, Marcora S, Coutts AJ (2015) Mental Fatigue Impairs Intermittent Running Performance. Med Sci Sport Exerc 47:1682–1690. https://doi.org/10.1249/MSS.0000000000000592

Soligard T, Schwellnus M, Alonso JM (2016) Infographic. International Olympic Committee consensus statement on load in sport and risk of injury: How much is too much? Br J Sports Med 50:1042. https://doi.org/10.1136/bjsports-2016-096583

Staunton CA, Abt G, Weaving D, Wundersitz DWT (2021) Misuse of the term 'load' in sport and exercise science. J Sci Med Sport. https://doi.org/10.1016/j.jsams.2021.08.013

Stewart GM, Yamada A, Haseler LJ, et al (2016) Influence of exercise intensity and duration on functional and biochemical perturbations in the human heart. J Physiol 594:3031–3044. https://doi.org/10.1113/JP271889

Stöggl T, Sperlich B (2014) Polarized training has greater impact on key endurance variables than threshold, high intensity, or high volume training. Front Physiol 5 FEB:1–9. https://doi.org/10.3389/fphys.2014.00033

Stöggl TL, Sperlich B (2015) The training intensity distribution among well-trained and elite endurance athletes. Front Physiol 6:295. https://doi.org/10.3389/fphys.2015.00295

Taha T, Thomas SG (2003) Systems Modelling of the Relationship between Training and Performance. Sport Med 33:1061–1073. https://doi.org/10.2165/00007256-200333140-00003

Taylor AR, Sanders D, Myers T, et al (2017) The dose-response relationship between training load and aerobic fitness in academy rugby union players. Int J Sports Physiol Perform 13:163–169. https://doi.org/10.1073/pnas.0703993104

Thomas K, Stone MR, Thompson KG, et al (2012) Reproducibility of pacing strategy during simulated 20-km cycling time trials in well-trained cyclists. Eur J Appl Physiol 112:223–229. https://doi.org/10.1007/s00421-011-1974-4

Tillin NA, Bishop D (2009) Factors modulating post-activation potentiation and its effect on performance of subsequent explosive activities. Sport Med 39:147–166. https://doi.org/10.2165/00007256-200939020-00004

Toussaint HM, Hollander AP (1994) Energetics of Competitive Swimming. Sport Med 18:384–405. https://doi.org/10.2165/00007256-199418060-00004

Tremblay MS, Copeland JL, Van Helder W (2005) Influence of exercise duration on post-exercise steroid hormone responses in trained males. Eur J Appl Physiol 94:505–513. https://doi.org/10.1007/s00421-005-1380-x

Tschakert G, Handl T, Weiner L, et al (2022) Exercise duration : Independent effects on acute physiologic responses and the need for an individualized prescription. 1–14. https://doi.org/10.14814/phy2.15168

Turner AP, Cathcart AJ, Parker ME, et al (2006) Oxygen uptake and muscle desaturation kinetics during intermittent cycling. Med Sci Sports Exerc 38:492–503. https://doi.org/10.1249/01.mss.0000188450.82733.f0

Valenzuela PL, Alejo LB, Ozcoidi LM, et al (2023) Durability in Professional Cyclists: A Field Study. Int J Sports Physiol Perform 18:99–103. https://doi.org/10.1123/ijspp.2022-0202

van Djik H, van Megen R (2017) Secret of Running: Maximum Performance Gains Through Effective Power Metering and Training. Meyer & Meyer Sport

van Erp T, Foster C, de Koning JJ (2019a) Relationship Between Various Training-Load Measures in Elite Cyclists During Training, Road Races, and Time Trials. Int J Sports Physiol Perform 493–500. https://doi.org/10.1123/ijspp.2017-0722

van Erp T, Hoozemans M, Foster C, de Koning JJ (2019b) The Influence of Exercise Intensity on Training Load in Professional Cycling. Med Sci Sport Exerc 50:27. https://doi.org/10.1249/01.mss.0000535171.74071.b3

Van Erp T, Hoozemans M, Foster C, De Koning JJ (2020) Case Report: Load, Intensity, and Performance Characteristics in Multiple Grand Tours. Med Sci Sports Exerc 52:868–875. https://doi.org/10.1249/MSS.0000000000002210

Vermeire KM, Caen K, Bourgois JG, Boone J (2023) Training Load and Acute Performance Decrements Following Different Training Sessions. Int J Sports Physiol Perform Ahead of print

Vermeire KM, Vandewiele G, Caen K, et al (2021) Training Progression in Recreational Cyclists: No Linear Dose-Response Relationship With Training Load. J Strength Cond Res 35:3500–3505. https://doi.org/10.1519/JSC.0000000000003340

Viru A, Viru M (2000) Nature of training effects. In: Exercise and Sport Science. Philadelphia, pp 1261–1269

Wakayoshi K, D'Acquisto L, Cappaert J, Troup J (1995) Relationship Between Oxygen Uptake, Stroke Rate and Swimming Velocity in Competitive Swimming. Int J Sports Med 16:19–23. https://doi.org/10.1055/s-2007-972957

Wallace LK, Slattery KM, Coutts AJ (2014a) A comparison of methods for quantifying training load: Relationships between modelled and actual training responses. Eur J Appl Physiol 114:11–20. https://doi.org/10.1007/s00421-013-2745-1

Wallace LK, Slattery KM, Coutts AJ (2009) The Ecological Validity and Application of the Session-RPE Method for Quantifying Training Loads in Swimming. J Strength Cond Res 23:33–38. https://doi.org/10.1519/JSC.0b013e3181874512

Wallace LK, Slattery KM, Impellizzeri FM, Coutts AJ (2014b) Establishing the criterion validity and reliability of common methods for quantifying training load. J Strength Cond Res 28:2330–2337. https://doi.org/10.1519/JSC.0000000000000416

Weaving D, Dalton-Barron N, McLaren S, et al (2020) The relative contribution of training intensity and duration to daily measures of training load in professional rugby league and union. J Sports Sci 38:1674–1681. https://doi.org/10.1080/02640414.2020.1754725

Wenger HA, Bell GJ (1986) The Interactions of Intensity, Frequency and Duration of Exercise Training in Altering Cardiorespiratory Fitness. Sport Med 3:346–356. https://doi.org/10.2165/00007256-198603050-00004

Wilder RP, Greene JA, Winters KL, et al (2006) Physical Fitness Assessment: An Update. J Long Term Eff Med Implants 16:193–204. https://doi.org/10.1615/JLongTermEffMedImplants.v16.i2.90

Winter E (2006) “Workload” - Time to abandon? J Sports Sci 24:1237–1238. https://doi.org/10.1080/02640410601072757

Winter EM, Abt G, Brookes FBC, et al (2016) Misuse of “Power” and Other Mechanical Terms in Sport and Exercise Science Research. J Strength Cond Res 30:292–300. https://doi.org/10.1519/JSC.0000000000001101

Winter EM, Fowler N (2009) Exercise defined and quantified according to the Système International d ’ Unités. J Sports Sci 0414:5. https://doi.org/10.1080/02640410802658461

Yakovlev, N (1955). Survey on sport biochemistry [in Russian]. Moscow: FiS Publisher