UNIVERSITY OF CHICHESTER

An Accredited Institution of the UNIVERSITY OF SOUTHAMPTON

Faculty of Sport, Education & Social Sciences

PHYSICAL TRAINING FOR LOADED MARCHING PERFORMANCE AMONG BRITISH ARMY RECRUITS

By

Pieter E H Brown

Thesis for the degree of Doctor of Philosophy

June 2009

ABSTRACT

UNIVERSITY OF CHICHESTER
ABSTRACT
FACULTY of SPORT, EDUCATION & SOCIAL SCIENCES
Doctor of Philosophy
PHYSICAL TRAINING FOR LOADED MARCHING PERFORMANCE AMONG BRITISH ARMY RECRUITS
By Pieter Edward Hudson Brown

Study 1 quantified the validity and repeatability of an automated on-line (ON) gas analysis system during sub-maximal loaded marching (LM) against that of the Douglas Bag (DB) approach. The 95% ratio Limits of Agreement (LoA) revealed the ON system systematically overestimated $\dot{V}O_2$ by ~16% (1.16 (×/÷1.19). The Bland and Altman plots revealed DB repeatability was almost two-fold better than ON (~9% vs. ~15%), thus the DB approach should be used subsequently to measure human expired gases.
Study 2 investigated the difference between a LM maximal oxygen uptake protocol (LM_p) versus a standard running protocol (R_p). The LM_p $\dot{V}O_2max$ was lower than R_p (48.6 ± 4.3 $ml \cdot kg^{-1} \cdot min^{-1}$ vs. 51.3 ± 4.0 $ml \cdot kg^{-1} \cdot min^{-1}$, P=0.001). Thus, the quantification of sub-maximal LM exercise intensity will be underestimated by ~5% if derived from a running $\dot{V}O_2max$ protocol.
Study 3 investigated the repeatability of accepted and potential determinants of Loaded Marching Performance (LMP). The LoA revealed the repeatability of Loaded Marching Economy (LME) (0.98 (×/÷1.09)), $\dot{V}O_2max$ (1.01 (×/÷1.07)), upper body dynamic strength (1.01 (×/÷ 1.11)), and anthropometric measures (1.00 (×/÷ 1.02)) to (1.00 (×/÷ 1.07)) was reasonable, but dynamic leg strength (1.06 (×/÷1.14)) and isometric strength (1.00 (×/÷1.12)) to (0.99 (×/÷1.16)) were large.
Study 4 established the determinants of 2.4 km LMP from a test battery performed at the beginning of British Army infantry training. The best mathematical model of LMP included the independent variables of, LME (r=0.65), 2.4 km run time (r=0.42), and peak static lift strength (r=0.48). This explained 65% of the variation in LMP, and had a prediction error of ± 51 s. Mathematically, LME and 2.4 km run time exerted the greatest influence on LMP, whereas the influence of static lift strength on LMP was small.
Study 5 investigated the physical and physiological responses of the established determinants of LMP during 24 weeks of British Army infantry training. Loaded marching performance improved 7.0% (900 s to 837 s, P=0.001), LME 9.6% (2.28 $l \cdot min^{-1}$ to 2.06 $l \cdot min^{-1}$, P<0.01), 2.4 km running performance 3.6% (617 s to 595 s, P=0.002), and $\dot{V}O_2max$ 2.6% (3.74 $l \cdot min^{-1}$ to 3.84 $l \cdot min^{-1}$, P=0.007), however peak static lift strength did not change (126.0 kg to 122.0 kg, P=0.249). Thus, infantry physical training should be modified to further improve $\dot{V}O_2max$, 2.4 km run time, and muscular strength, in order to further improve LMP.
Study 6 investigated the efficacy of a modified (MOD) physical training programme designed to improve $\dot{V}O_2max$, 2.4 km run time, and muscular strength, for the purposes of further improving LMP. MOD physical training consisted of high intensity interval training, and field based resistance training. Between-groups no differences were observed in 2.4 km LMP (832 s vs. 826 s, P=0.187), $\dot{V}O_2max$ (4.01 $l \cdot min^{-1}$ vs. 4.06 $l \cdot min^{-1}$, P=0.828), 2.4 km running time (571 s vs 570, P=0.208), and static lift strength (126.5 kg vs. 119.0 kg, P=0.218) at the end of training. Unexpectedly, the control group performed better in, 6.4 km LMP (49.7 mins vs. 51.5 mins, P=0.005) at the end of training, and 2.4 km running performance in the middle of training (588 vs. 566 s, P=0.001). Thus, the MOD intervention was no better than existing infantry physical training at improving LMP, as well as the determinants of LMP.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF PUBLISHED ABSTRACTS

Brown P.E.H., Fallowfield J.L., Wilkinson D.M., Bilzon J.L.J (2007). The reproducibility of loaded marching economy using two different gas analysis systems. Journal of Sport Sciences. 25(3), 254.

Brown P.E.H., Fallowfield J.L., Blacker S.D., Izard R.M., Wilkinson D.M., Bilzon J.L.J (2007). Does loaded marching economy explain some of the variance in loaded marching performance? Medicine and Science in Sports and Exercise. 39, S154.

Brown P.E.H., Fallowfield J.L., Blacker S.D., Izard R.M., Wilkinson D.M., Bilzon J.L.J (2008). The training response to 24-weeks of British Army line infantry recruit training. Medicine and Science in Sports and Exercise. 40, S159.

DECLARATION OF AUTHORSHIP

I, PIETER BROWN, declare that the thesis entitled "Physical Training for Loaded Marching Performance Among British Army Recruits" and the work presented in the thesis are both my own work, and have been generated by me as the result of my own original research. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Parts of this work has been published before submission (see published abstracts).

Signed:..

Date:..

ACKNOWLEDGMENTS

First, a huge thanks to the Headquarters Army Recruiting and Training Division for supporting this work. Working with the British Army has been a privilege that I will cherish. Also, many thanks to the University of Chichester for putting a roof over my head for the past three years or so. The completion of this thesis has been long, but rewarding, I would liken the process to running a marathon, with my shoe-laces tied together. Several people have been influential in the development of this thesis, who I would like to thank.

I would first like to thank Dr James Bilzon for commissioning this thesis, and providing valuable support and supervision throughout.

Thanks to Dr Charles Minter and Phil Stevens who gave advice on building the gas analysis system used in this thesis.

Thanks to Major Martin Colclough, Captain Ken Carter, WO2 Steve Spiers, Corporal Jimmy Corkhill, Corporal Richie Peet, and the staff and recruits of the Rifles Company, and the Prince of Wales Company at the Infantry Training Centre (Catterick).

Thanks to Sam Blacker, Katie Griggs, Mark Robinson and Rachel Izard for assisting in the data collection process.

Thanks to Dr Dave Wilkinson for his supervision, excel spreadsheets and encyclopaedic knowledge.

Finally, huge thanks go to my Director of Studies Dr Jo Fallowfield, who stayed committed and enthusiastic towards this programme of research even after leaving the University of Chichester for the Institute of Naval Medicine. I would like to thank Jo for tolerating my coping mechanism (whinging), particularly in the final stages of the PhD. This an era of my life I will be sorry to leave behind, particularly as I have only recently learnt to translate Jo's hieroglyphic like handwriting.

MILITARY DEFINITIONS AND ABBREVIATIONS

ACFT 1	**Advanced Combat Fitness Test 1** A 2.4 km British Army loaded marching performance test
APTCI	**Army Physical Training Corps Instructor** A trained physical training instructor in the British Army
CEG	**Career Employment Group** Refers to the Trade, Arm or Service a soldier specialises in the British Army
CIC	**Combat Infantryman's Course** 24 to 26-week infantry training course conducted at the ITC(C)
CMS(R)	**Common Military Syllabus (Recruits)** Phase 1 British Army training
CS 95	**Combat Soldier 95** British Army issue clothing
HQ ARTD	**Headquarters Army Recruiting and Training Division** Division of the British Army responsible for all recruiting and training
ITC(C)	**Infantry Training Centre (Catterick)** The training unit responsible for training all infantry recruits in the British Army
ITG	**Initial Training Group** Part of HQ ARTD that is responsible for all Phase 1 recruit training
PFT	**Personal Fitness Test** A standard British Army fitness test that consists of, 2-minutes of press-ups, 2-minutes of sit-ups, and 2.4 km run test
Phase 1	**Phase 1 training** 14 weeks of initial British training, also called CMS(R)
Phase 2	**Phase 2 training** Follows Phase 1, trains British Army recruits for specialist roles
Platoon	A group of four sections that equate to a total of 40 to 50 recruits
PLCE	**Personal Load Carriage Equipment** The load-carriage ensemble of the British Army
PSS(R)	**Physical Selection Standards (Recruits)** A series of tests conducted before recruit training to predict the likelihood of a potential recruit to pass the RMT physical output standard at the end of military training
PT	**Physical Training**
PTI	**Physical Training Instructor** See APTCI
RMT	**Representative Military Task** Job specific tasks that include (loaded marching, single lift, carry, and repetitive lift) which recruits must complete to a minimum standard to pass recruit training
Streaming	Placing recruits into ability based groups during PT lessons, in an attempt to optimise performance, and reduce injuries

CHAPTER 1

INTRODUCTION

The Headquarters Army Recruiting and Training Division (HQ ARTD) is responsible for initial (Phase 1) and specialist (Phase 2 and 3) training of all Officer Cadets and Soldiers joining the British Army, and has commissioned this thesis. HQ ARTD's mission statement is:

"To deliver the required number of appropriately trained and motivated officers and soldiers to meet the operational requirements of the Army and Defence".

Operating under the direction of HQ ARTD, Initial Training Group (ITG) is responsible for delivering initial (Phase 1) training to all non-infantry recruits across 5 different training centres. During this period recruits are taught generic skills fundamental to soldiering. Subsequently, recruits transfer to Phase 2 training establishments where they begin specific skills relevant to their chosen Trade, Arm or Service, which are generically referred to as Career Employment Groups (CEG). Infantry recruits are the exception to this rule, who undergo a combined (Phases 1 and 2) 24 to 26-week Combat Infantryman's Course (CIC) at the Infantry Training Centre (Catterick) (ITC(C)).

The Representative Military Task (RMT) of loaded marching is a fundamental military occupational task (Rayson *et al.*, 2000), particularly for the infantry soldier. This is reflected in the infantry CEG physical output standards that must be achieved towards the end of recruit training. Infantry initial training physical output standards are among the hardest in the British Army, and require a recruit to complete a 12.8 km march, with a load of 25 kg, in a duration of no longer than 2 hours (British Army, 2002). A good example demonstrating this requirement is the Falkland's War. Here troops were forced to march long distances across the island, with loads of up to 68 kg prior to engaging the enemy (McCaig and Gooderson, 1986). In such a scenario loaded marching ability could have made the difference between life and death.

Presently, the HQ ARTD mission is not being met, hence strategies to maximise recruit output to the Field Army have already been implemented. These have included the extension of Phase 1 training from 12 to 14 weeks, infantry training from 24 to 26 weeks, and a 3-week Soldier Pre-Conditioning Course aimed at the least aerobically fit individuals (i.e. the 5th quintile of 2.4 km run times at the Army Selection and Development Centre) (Bilzon *et al.,* 2007). The present thesis will focus on the content of physical training (PT) during British Army infantry training as a means of improving physical fitness, specifically for loaded marching performance (LMP).

Exercise performance measures alone (i.e. LMP time) do not elucidate the mechanisms that determine exercise performance. To determine the mechanisms that explain exercise performance individual components of fitness should be considered (Wilkinson, 1999). Therefore an appraisal process that measures individual components of fitness can uncover the specific morphology (i.e. capillary density, cardiac output etc.) and functional abilities (i.e. $\dot{V}O_2max$, exercise economy etc.) that determine exercise performance (Coyle, 1995). The adequacy (or not) with which these components of fitness respond (or not) to a PT programme will highlight which morphological, functional, or performance measures should be targeted subsequently to further improve exercise performance (Wilkinson, 1999). This has been demonstrated by quantifying the changes that functional running abilities had on the subsequent improvements to long distance running performance (Wood, 1999). Here, a mathematical simulation showed that improvements in $\dot{V}O_2max$, maximal sustainable $\%\dot{V}O_2max$, and running economy exerted a marked influence on long distance running performance, but improvements in $\dot{V}O_2$ kinetics and anaerobic capacity did not. From such information physical training programmes can prioritise the development of the most influential components of fitness, in order to maximise improvements in exercise performance. A similar model of exercise performance evaluation, physical training evaluation, and physical training prioritisation to maximise the improvements in LMP will be applied in the present thesis.

Whilst the loaded marching literature has attempted to quantify the determinants of LMP (Rayson *et al.,* 2000; 2002a, Pandorf *et al.,* 2002), and physical training interventions have been implemented to improve components of fitness that are thought to influence LMP (Kraemer *et al.,* 1987, 2001, 2004), only one programme of research has investigated these questions in relation to LMP in British Army recruits (Williams *et al.,* 1999, 2002, 2004; Williams and Rayson, 2006). However, a general criticism of the loaded marching research is the over-reliance of $\dot{V}O_2max$ (in the form of direct assessment and indirect estimation) as the sole parameter of aerobic fitness in the establishment of the determinants of LMP, and the monitoring of training responses. Thus, an incomplete understanding of the mechanisms that determine LMP, and how best to train to improve LMP exists as other parameters of aerobic fitness, such as exercise economy, lactate/ventilatory threshold, $\dot{V}O_2$ kinetics (Whipp *et al.,* 1982), the maximal sustainable percentage of $\dot{V}O_2max$ (Wood, 1999), and the velocity attained at $\dot{V}O_2max$ (Noakes *et al.,* 1990) have been ignored.

The primary purpose of this thesis was to improve LMP in British Army infantry recruits undergoing the Combat Infantryman's Course (CIC) at the Infantry Training Centre (Catterick) (ITC(C)). To investigate this, a valid and reliable test battery of accepted and potential determinants of LMP was developed (Chapter 4, 5 and 6), which was then applied to the field. In the field (i.e. ITC(C)) a mathematical model established the determinants of LMP (Chapter 7), and the training responses of these determinants were tracked over 24 weeks of British Army infantry training (Chapter 8). The determinants of LMP and the training response of these determinants highlighted the strengths and weaknesses of infantry training, in relation to LMP progression. Finally, a modified physical training programme was developed to target the existing weaknesses of infantry physical training, with the aim of further improving LMP. The efficacy of this modified physical training programme was compared against a matched group of controls (Chapter 9).

CHAPTER 2

THE REVIEW OF LITERATURE

2.1 Contextualisation

'Loaded marching' and 'load-carriage' are terms used interchangeably in this thesis. Unless otherwise stated these terms refer to the carriage of a load, positioned on the torso during human bipedal locomotion (e.g. rucksack / military bergen). Load-carriage is fundamental to military strategy, and is used for tactical and administrative purposes around a military combat zone. The success of a loaded march (LM) is characterised by a punctual rendezvous with sufficient energy in reserve to complete a mission (Headquarters Department of the Army, 1990). The importance of load-carriage to the British Army is reflected in the loaded marching physical output standards that must be attained during initial training, and subsequently maintained throughout a career in the Field Army. In the case of the line infantry, and infantry foot guards these physical output standards refer to the completion of a 12.8 km loaded march, carrying 25 kg, in a maximum time of 2 hours (British Army, 2002).

Load-carriage is multifaceted. It has physiological (the focus of this thesis), biomechanical, and injury related factors (Knapik *et al.*, 2004), all of which have the potential to influence mobility. Five strategies have been proposed to improve load-carriage mobility and include: the development of lighter components; the use of computer scenarios to plan load configurations; the development of specialised load-carriage equipment (e.g. handcarts etc.); a change in doctrine (e.g. reduce amount of ammunition in loads by the improvement of marksmanship); and improved physical training specifically for load-carriage (Knapik *et al.*, 2004).

In leading towards the question of improving physical training (PT) for loaded marching performance (LMP), this review of literature will first discuss the existing determinants of running performance, and how this information may be useful in establishing the determinants of LMP. It will then discuss the accepted and potential determinants of LMP, and how these can be improved through PT to further develop LMP.

2.2 Determinants of running performance

Whilst running is different to LM (Rayson *et al.*, 1995), it represents the closest exercise modality to LM (i.e. continuous, rhythmical, and cyclical), and has a plethora of empirical findings that could transfer, and help further develop the understanding of LMP. Running performance is the product of individual components of fitness that combine to collectively influence performance (Wilkinson, 1999). This same notion will ultimately be applied to LMP.

There is a hierarchical structure in the factors that determine endurance performance (Coyle, 1995). Here *morphological components* (e.g. muscle fibre type composition, maximal stroke volume etc.) influence *functional abilities* (e.g. $\dot{V}O_2$max, exercise economy, lactate threshold etc.), which then influence *performance abilities* (e.g. performance velocity etc.). This is a classic case of the structure of an organism dictating its function (Figure 2.1).

Reviews of the running literature revealed that the influence of aerobic factors (i.e. $\dot{V}O_2$max, anaerobic threshold, and running economy), and anaerobic factors (i.e. anaerobic capacity and power) combine to influence running velocity, which influences middle distance running performance (Brandon *et al.*, 1995). In a more comprehensive model of middle and long distance running performance, $\dot{V}O_2$max, $\dot{V}O_2$kinetics, maximal sustainable percentage of $\dot{V}O_2$max, anaerobic capacity, and running economy were identified as determinants of performance (Wood, 1999). This model demonstrated that a training related improvement in $\dot{V}O_2$max and running economy would exert the largest improvement in 5 km to 42 km running performance. Furthermore, as running distance increased from 5 km to 42 km the maximal sustainable percentage of $\dot{V}O_2$max became a very influential determinant of running performance, but the influence of $\dot{V}O_2$kinetics and anaerobic capacity were negligible on running performance. The influence of physiological measures alone (i.e. functional abilities), which included $\dot{V}O_2$max, the %$\dot{V}O_2$ at lactate threshold, and running economy also emerged as determinants of marathon running speed (Joyner, 1991). More recently, a curvilinear relationship between the $\dot{V}O_2$max / running economy ratio explained 95.9% of the variation in middle distance running performance (Ingham *et al.*, 2008).

Most of the established determinants of running performance are physiological, hence relate to the *functional abilities* of a runner (Coyle, 1995; Figure 2.1). As these functional abilities hierarchically precede *performance abilities* it's logical that performance measures will be the combined influence of *morphological components* and the *functional ability* of a runner. Due to the fact that *performance abilities* are composite measures of fitness (i.e. include more than one component of fitness), these may be better determinants of running performance. In support of this notion, Noakes *et al.,* (1990) reported running performance over race distances of 10 km to 90 km (a *performance ability*) as the strongest correlates of other endurance race performance, and not measures of *functional ability.* Although these *performance abilities* appear to be the most relevant and specific determinants of running performance, the fact that these are the product of several individual components of fitness (Wilkinson, 1999) means that the establishment of the determinants of running performance solely from *performance abilities* does little to further the understanding of the mechanisms that determine exercise performance (Figure 2.1).

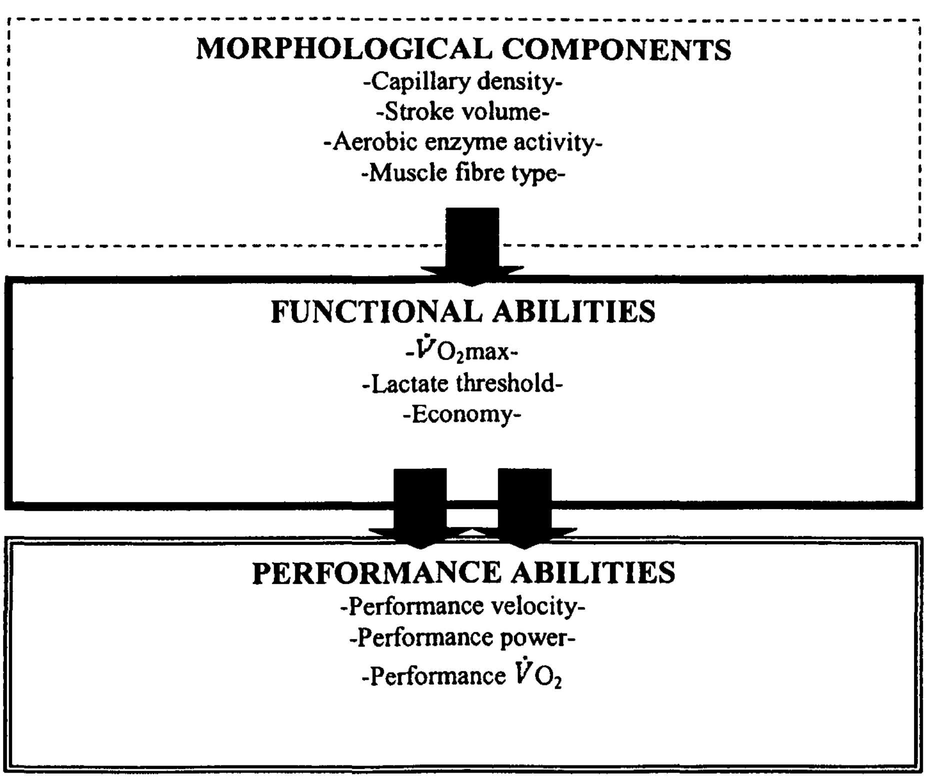


Figure 2.1. A theoretical model illustrating the hierarchical cascade of structural, and functional abilities, and how these determine endurance performance. The examples provided are not definitive. Adapted from (Coyle, 1995).

2.3 Determinants of loaded marching performance

The theoretical concept proposed by Coyle (1995), which suggested a hierarchical order to the determinants of endurance performance (Figure 2.1) can be applied to LMP. Understanding the determinants of LMP is pertinent to the British Army, as this would prioritise the development of the individual components of fitness that are most influential to LMP.

Several studies have investigated the determinants of LMP (Table 2.1). These studies established the determinants by quantifying the strength of the relationship between test battery measures and a criterion LMP test. The considerable variation in the distances (2 km to 20 km) and loads (14 kg to 61 kg) of these criterion LMP tests may have altered the importance of certain components fitness between models (Mello *et al.*, 1988). This should be taken into consideration when interpreting the relevance and applicability of these findings. Nevertheless, a common theme to emerge from these data is that LMP is determined by three general aspects of fitness that include parameters of aerobic fitness, strength, and anthropometry.

Table 2.1. A review of the studies that investigated the determinants of loaded marching performance.

Study	Sample	Criterion Load Carriage Test	Load-Carriage Performance Predictors
Dziados *et al.*, (1987)	49 male soldiers	16 km 18 kg	Absolute $\dot{V}O_2$max (r = -0.37) Leg torque (r = -0.34 to -0.45; R^2 = 0.21)
Mello *et al.*, (1988)	28 male soldiers	2, 4, 8 & 12 km 46 kg	**2 km** – No predictors **4 km** – No predictors **8 km** - % fat (r = 0.49) Leg torque (r =-0.48 to r =-0.64) **12 km** – Leg torque (r = -0.48 to r =-0.59)
Knapik *et al.*, (1990)	96 male soldiers	20 km ~46 kg	Age (r = -0.27); Body mass (r = -0.22); Fat Free Mass (r = -0.26) Absolute $\dot{V}O_2$max (r = -0.31) Peak leg power (r = -0.23); Abdominal strength (r = -0.45) Hand grip strength (r = -0.30); Upper torso strength (r = -0.32) Leg strength (r = -0.22 to r =-0.27); Plantar flexor strength (r = -0.24 to r =-0.29)
Rayson *et al.*, (1993)	18 female soldiers	6.4 km·h^{-1} (treadmill) Mean load 37kg	Absolute $\dot{V}O_2$max (r = 0.57); Relative $\dot{V}O_2$max (r = 0.52) Plantar flexion torque (r = 0.50); Stature (r = 0.49) Maximum load = absolute $\dot{V}O_{2max}$, Plantar flexion, Age, % fat (R^2 = 0.71)
Frykman and Harman (1995)	13 males	3.2 km 34 and 61 kg	**34 and 61 kg** Body mass (r =-0.60, r =-0.63); Height (r = -0.63, r =-0.87) Leg length (r = -0.61, r =-0.75); Trunk length (r = -0.58, r =-0.80) Shoulder diameter (r = -0.74, r =0.65); Bitroch. diameter (r = -0.54, r =-0.76) Neck circumference (r = -0.72) 61 kg only

Harman and Frykman (1995)	13 males	3.2 km 34 and 61 kg	**34 and 61 kg** Absolute $\dot{V}O_2max$ (r = -0.84, r =-0.74) Leg strength/endurance (r = -0.68, r =-0.72)
Rayson *et al.*, (2000)	82 – 100 male & female soldiers	12.8 km 15, 20 & 25 kg	**15 kg** - Absolute $\dot{V}O_2max$, Body mass, Static arm flexion endurance (R^2 = 0.40) **20 kg** – Multi stage fitness test, gender (R^2 = 0.55) **25 kg** – Multi stage fitness test, Ln static arm flexion endurance, % fat (R^2 = 0.75)
Rayson *et al.*, (2002a)	24 – 161 male & female British Army recruits	9.7 to 12.8 km 15, 20 & 25 kg	**9.7 km (15 kg)** – 2.4-km run time, 2.4-km run time^2, Static lift strength (R^2 = 0.50) **9.7 km (20 kg)** – Relative $\dot{V}O_2max^2$, Absolute $\dot{V}O_2max$ (R^2 = 0.39) **12.8 km (25kg)** – Absolute $\dot{V}O_2max$, Back extension strength (R^2 = 0.52)
Rayson *et al.*, (2002b)	Male & female Junior Entry British Army recruits	9.7 km 15, 20 & 25 kg	**15 kg** – 2.4-km run time, 2.4-km run time^2, Static lift strength **20 kg** – Relative $\dot{V}O_2max^2$, Absolute $\dot{V}O_2max$ **25 kg** – Absolute $\dot{V}O_2max$, Back extension strength
Pandorf *et al.*, (2002)	12 female soldiers	3.2 km 14, 27, 41 kg	**14 kg** – 3.2-km run time, Shoulder width, Hip width (R^2 = 0.82) **27 kg** – 3.2-km run time, Absolute $\dot{V}O_2max$, Hip width (R^2 = 0.73) **41 kg** – 3.2-km run time, Number of press-ups, Absolute $\dot{V}O_2max$ (R^2 = 0.69)
Williams & Rayson (2006)	56 – 92 male and female British Army recruits	3.2 km 15 & 25 kg	**15 kg** - (gender pooled) Gender, Stature, % fat, Shuttle run time (Adjusted R^2 = 0.91) - (male) Shuttle run time, % fat, FFM, Body mass (Adjusted R^2 = 0.79) - (female) Stature, Static lift strength (Adjusted R^2 = 0.75) **25 kg** – (male) Age, Stature, Shuttle run time (Adjusted R^2 = 0.40)

Simpson *et al.*, (2006)	18 elite soldiers	3.2 & 29 km 20 kg	**3.2 km** – load-carriage maximal treadmill test (r =-0.57), mental toughness (r =-0.53) (R^2 0.51) **29 km** - load-carriage maximal treadmill test (r=-0.66), 3.2 km load-carriage (r =0.77). 3.2-km load-carriage, relative $\dot{V}O_2max$ (R^2 0.70)
Graham *et al.*, (2007)	13 elite soldiers	12.8 km 20 kg	**12.8 km** – Velocity at lactate break-point (r =-0.86), Δ 1mM, 2.5 mM, 3.5 mM, 4.0mM [BLa] (r =-0.71 to r =-0.82), $\dot{V}O_2peak$ (r =-0.40)
Griggs *et al.*, (2008)	62 British Army Officer cadets	12.8 km 20 kg (female) 25 kg (male)	**12.8 km** – 2.4 km LMP (R^2 0.74, SEE 124 s) **12.8 km** – 2.4 km run time and body mass (R^2 0.51, SEE 224 s)

2.3.1 Parameters of aerobic fitness

Parameters of aerobic fitness are the most frequently cited determinants of LMP (Table 2.1), which makes intuitive sense considering the repetitive, continuous and sub-maximal nature of load-carriage. The parameters of aerobic fitness refer to the maximal oxygen uptake ($\dot{V}O_2max$), exercise economy, lactate/ventilatory threshold, $\dot{V}O_2$ kinetics (Whipp *et al.*, 1982), the velocity attained at $\dot{V}O_2max$ (Noakes *et al.*, 1990), and the maximal sustainable percentage of $\dot{V}O_2max$ (Wood, 1999). A criticism of the parameters of aerobic fitness that have emerged as determinants of LMP (Table 2.1) is that these do not reflect the well established parameters of aerobic fitness in the running literature, hence an incomplete picture on the mechanisms that determine LMP exists. The measures investigated in the LMP literature have typically included indirect and direct assessment of $\dot{V}O_2max$ ($l \cdot min^{-1}$, $ml \cdot kg^{-1} \cdot min^{-1}$), and maximal effort running performance tests (2.4 km and 3.2 km). Other parameters of aerobic fitness that are known determinants of endurance performance have been ignored (e.g. exercise economy, velocity attained at $\dot{V}O_2max$, maximal sustainable percentage of $\dot{V}O_2max$, and lactate/ventilatory threshold).

Early attempts to describe the determinants of LMP produced at best moderate associations. In the first of a series of trials, trained male soldiers carried 18 kg for 16 km, which produced weak inverse correlations ($r = -0.37$) with absolute $\dot{V}O_2max$ (Dziados *et al.*,1987). This implied that individuals with a larger absolute $\dot{V}O_2max$ performed better in a loaded marching task. With a heavier load (46 kg) Knapik *et al.*, (1990) also reported weak inverse relationships ($r = -0.31$) over the longer march distance of 20 km, which was attributed to a decrease in subject motivation over extended march distances. Thus, psychological factors may have impacted upon the physiological / performance relationship with LMP in this study. Strong ($r = -0.84$) correlations with absolute $\dot{V}O_2max$ have been reported over 3.2 km, with loads of up to 61 kg (Harman and Frykman, 1995), but the small sample size used ($n = 13$) means these results should be viewed with caution (Fallowfield *et al.*, 2005).

The weak to moderate correlations between $\dot{V}O_2max$ and LMP are far from conclusive evidence that this parameter of aerobic fitness is a major determinant of LMP. However, simple bivariate comparisons might be an inadequate statistical tool to explain all of the variance in LMP, as factor analysis revealed that 79% of LMP is explained by a multitude of factors that include, size, overweightness, muscular strength/endurance and aerobic fitness (Rayson *et al.,* 2000). Others have suggested that load-carriage relies on two very different physical factors (i.e. aerobic fitness and strength) (Kraemer *et al.,* 1987, 2004; Williams *et al.,* 2004). Hence, the use of just one measure to explain variation in LMP should be viewed as poor practice. Instead, a multivariate approach may be more appropriate at explaining the multi-factorial nature of LMP.

A multivariate approach was used to develop and validate criterion models of LMP to determine if a potential British Army recruit would be capable of successfully completing the LM physical output standard towards the end of British Army recruit training. This mathematical model to predict LMP is commonly referred to as Physical Selection Standards for (Recruits) (PSS(R)) (Rayson *et al.,* 2000). To develop this model, personnel representing each Arm and Service of the British Army were assessed (over 300 trained military personnel). The test battery consisted of 34 separate measures, which were analysed against a 15, 20 and 25 kg, 12.8-km 'best effort' LMP test. Stepwise multiple regression explained 75, 55, and 40% of the variance in LMP for the 15, 20 and 25 kg load-carriage tasks, respectively. These data show that as load magnitude increased, the ability to explain the variation in LMP decreased. It is possible that this observation is a statistical artefact, as the 15 kg and 20 kg loaded marching models were gender pooled (male and female subjects), whereas the 25 kg model had male only data. Thus, as R^2 is sensitive to data heterogeneity (between-subject variation) (Altman and Bland, 1983), the larger range of gender pooled data (15 kg and 20 kg) compared to male only (25 kg load-carriage) probably explains the better fit of the 15 kg and 20 kg LMP models.

Parameters of aerobic fitness that were included in these LM models were indirect estimates of $\dot{V}O_2max$ (Rayson *et al.,* 2000). Interestingly, indirect absolute $\dot{V}O_2max$ emerged as a determinant in the heaviest LMP test (i.e 25 kg), whereas the Multi Stage Fitness Test score best explained 15 kg and 20 kg LMP. Absolute $\dot{V}O_2max$ is a physiological measure that reflects both the maximal rate the body can transport and utilise oxygen and body mass (or more likely) fat-free mass (Haisman, 1988). This indicated that as the magnitude of an absolute load increases, the importance of a high $\dot{V}O_2max$ may begin to decrease, but body size increase.

Changes to the British Army selection and training necessitated that the original PSS(R) LMP models (Rayson *et al.,* 2000) were re-validated (Rayson *et al.,* 2002a). The need for the re-validation was driven by the increased duration of training (from 11 weeks to 24-26 weeks) for CEGs required to carry 25 kg over 12.8 km, as well as the reduced distance (9.7 km) for CEGs required to carry 15 kg or 20 kg at the end of the 12-week Phase 1 training. As the process from Army selection to the completion of British Army recruit training had systematically altered a failure to re-validate the PSS(R) LMP model would have caused *shrinkage* (i.e. a decrease in explained variance), and an increased prediction error of the PSS(R) LMP model (Thomas and Nelson, 1996). This meant the selection standards for the British Army were no longer valid. The re-validation of PSS(R) meant that the number of misclassifications would be reduced (i.e. selecting out individuals that would have actually reached the physical output standard at the end of training and *vice* versa).

The PSS(R) re-validation revealed parameters of aerobic fitness to be determinants of LMP in each of the 3 models (15, 20, 25 kg) (Rayson *et al.,* 2002a). Again the combined importance of a high aerobic fitness and a large body size emerged, as indirect absolute $\dot{V}O_2max$ was a predictor of 25 kg LMP. For the lighter LM models of 15 kg and 20 kg, 2.4-km run time and indirect relative $\dot{V}O_2max$ emerged as predictors. It should be noted that a small sample size developed the 25 km model, hence these data should be interpreted tentatively. Similar trends were observed in 20-42 weeks of British Army Junior Entry recruit training (Rayson *et al.,* 2002b). Pandorf *et al.,* (2002) reported 3.2 km running performance as a determinant of LMP over the same distance, with three different loads (14, 27, and 41 kg). Furthermore, absolute $\dot{V}O_2max$ was also more influential as load size increased from 27 kg to 41 kg.

Multivariate approaches to modelling LMP explained moderate to high amounts of variation in LMP, which was greater than the low to moderate associations reported in early bivariate analyses. This suggests that aerobic fitness underpins LMP, and that as the size of the external load increases the importance of $\dot{V}O_2max$ shifts from relative / performance measures, to absolute measures of $\dot{V}O_2max$. These absolute measures of aerobic fitness imply that body size may also be an important determinant of LMP. However, a glaring gap is apparent in these studies (Table 2.1). Section 2.2 highlighted the importance of sub-maximal physiological measures in determining running performance (i.e. economy, lactate/ventilatroy threshold, and maximum sustainable percentage of $\dot{V}O_2$max), which has barely been acknowledged in the LM literature (Graham *et al.*, 2007).

2.3.1.1 Sub-maximal parameters of aerobic fitness

As previously discussed sub-maximal parameters of aerobic fitness are known determinants of endurance running performance (Joyner, 1991; Brandon, 1995; Wood, 1999; Ingham *et al.*, 2008), but have all but been ignored in the LMP literature (Graham *et al.*, 2007) (Table 2.1). Running economy is one such measure, and is defined as:

"the oxygen uptake required at a given absolute exercise intensity."

(Jones and Carter, 2000)

Running economy is influential to running performance, as a theoretical 10% improvement in this measure will elicit a concomitant improvement of 6-8% in middle and long distance running performance (Wood, 1999). Sixty-five percent of the variation in 10 km running performance, in individuals homogenous in $\dot{V}O_2$max is explained by differences in running economy (Conley and Krahenbuhl, 1980), and a low $\dot{V}O_2$max can be compensated for with good running economy in individuals homogenous for 10 km running performance (Weston *et al.*, 2000). Based on observations such as these there is potential for an equivalent LM parameter to have a similar influence on LMP. Bilzon *et al.*, (2001) observed that the strength of the relationship between sub-maximal $\dot{V}O_2$ (i.e. an equivalent measure to running economy) and body mass almost doubled from (r=-0.47) during an unloaded running at 9.5 km·h^{-1}, to (r=-0.87) in an 18 kg loaded marching task at 9.5 km·h^{-1}. Furthermore, heavier individuals performed better in a LM test to exhaustion. This suggested that a better *Loaded Marching Economy* (LME) (note – this is a previously unused term, that is analogous to running economy) predisposed an individual to perform better in a LM performance test to exhaustion. Furthermore, Graham *et al.*, (2007) reported sub-maximal indices of aerobic running fitness (e.g. velocity at lactate breakpoint, lactate changes of 1.0, 2.5, 3.0, 4.0 mM, and break-points in ventilation) to strongly correlate with 12.8 km, 20 kg LMP. Thus, the rationale from the running literature and the indications from the LM literature, would suggest that LME is a potential determinant of LMP, and as such requires further investigation.

2.3.2 Parameters of anthropometry

The importance of absolute $\dot{V}O_2max$ would imply that body size, hence anthropometric measures, could have an important role to play in determining LMP. Anecdotally, it has been suggested that muscular / stocky individuals, typified by Gurkha Regiment in the British Army are ideal for infantry tasks such as load-carriage (Kilpatrick, 1989). This is a physical requirement specific to load-carriage, as in unloaded running tasks heavier individuals are disadvantaged (Vanderburgh and Mahar, 1995). In this study positive correlations were observed between 3.2 km run time and body mass (r=0.68), in a population heterogeneous in body mass (82.6 ± 12.1 kg). This evidenced that an increase in body mass (and to a lesser extent fat free mass r=0.32) was associated with slower run times. This contradicts that observed in LM, where individuals with a high $\dot{V}O_2max$ or fast run times are not necessarily the best performers in a load-carriage task. These tend to be heavier individuals (Bilzon *et al.*, 2001). In this study the time to exhaustion in an 18 kg, 9.5 $km{\cdot}h^{-1}$ load-carriage task was 11.8 minutes longer in volunteers with a body mass greater than 80 kg compared to those with a body mass of less than 80 kg. The better performance of the heavier individuals was evident even though relative $\dot{V}O_2max$ was 5.9 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ lower. This indicated that factors other than relative $\dot{V}O_2max$ are important determinants of LMP; one such factor may be body mass.

One technique to eliminate the body mass bias in running performance, is to allometrically scale run time against body mass, and fat free mass with an exponent of 0.31 (Vanderburgh and Mahar, 1995). A simpler way to eliminate the body mass bias is to perform the running task with an external load (Vanderburgh and Flanagan, 2000). In this study a mathematical simulation to estimate loaded 3.2 km run time with loads ranging from 0 kg to 50 kg was conducted. Correlations revealed no relationship between loaded run time and body mass when external load was increased from 20 kg to 50 kg. This suggested that heavier individuals are not disadvantaged when loads of at least 20 kg are used in 'best effort' running tests. Hence, an increase in absolute external load will favour larger individuals.

Body size and body dimensions are likely to vary with body mass. Hence, larger body size and body dimensions are associated with better LMP times. Frykman and Harman (1995) found moderate to high inverse correlations (r=-0.54 to r=-0.74) between 3.2 km, 34 kg LMP and the 7 anthropometric measures of, body mass, body stature, limb lengths, diameters and circumferences. When the external load was increased to 61 kg the strength of this association improved in 5 out of 6 measures, ranging from r=-0.63 to r=-0.87. To a lesser extent Pandorf *et al.,* (2002) reported an increase in the significance and strength of the relationship of five anthropometric measures as the magnitude of an external load increased from 14 kg to 41 kg. These studies highlight that increased size of an individual is beneficial to LMP, and that the heavier an absolute load the more important body size becomes. However, both studies discussed had small sample sizes (n = 13 and 12, respectively), hence caution is required in the interpretation of these data. Larger sample sizes have shown more modest findings, with weak inverse correlations between body mass (r=-0.22), fat free mass (r=-0.26) and 20 km LMP (Knapik *et al.,* 1990).

Multivariate analyses have also revealed anthropometric measures to be determinants of LMP. Williams and Rayson (2006) reported three out of four independent variables in a 15 kg loaded marching task were accounted for with the anthropometric measures of body mass, fat free mass and percent body fat. Body fat percentage (fat mass) is one anthropometric measure that exerts an influence on LMP in a counter intuitive manner. Studies have observed that as fat mass increases, LMP also increases (Rayson *et al.,* 1993; Rayson *et al.,* 2000; Pandorf *et al.,* 2002). Hence, the 'fatter' an individual the better the LMP. Intuitively this is surprising as subcutaneous fat has no mechanical purpose, and acts as additional 'dead weight', increasing the burden on the load-carrier. Therefore the most plausible explanation for the direction of this relationship is that increases in body mass and/or fat free mass are also accompanied by increases in fat mass.

2.3.3 Parameters of strength

Strength is the third general physical fitness component to influence LMP, which may simply be a substitute measure of body size. As Figure 2.1 demonstrated, morphology influences physiology, which influences performance (Coyle, 1995). Hence, the structure of a biological organism will influence its ability to function, and human skeletal muscle is no different (Marieb, 1992). Assuming fibre type expression, series-elastic elements and motor unit recruitment are constant, the arrangement of sarcomeres either in parallel (i.e. muscle cross-sectional area), or in series (i.e. muscle length) are key morphological characteristics that modulate muscular force and velocity generation, respectively (Jones and Round, 1990). Furthermore, cross-sectional area was moderately correlated (r=0.56) with muscle function (Maughan *et al.,* 1984), and the increase in cross-sectional area with training has accounted for 40% of the improvement in muscle function (Narici *et al.,* 1989). Thus, a bigger muscle can partly explain why it is stronger. Previous sections on *aerobic fitness* and *anthropometry* constructed an argument to suggest that the structure of an individual might influence load-carriage ability. Therefore, as structure dictates function, larger individuals should have greater strength, and assuming a reasonable aerobic fitness this would facilitate LMP. Consequently, the involvement of strength as a determinant of LMP may simply reflect the biological structure (hence anthropometry) of an individual.

An association between strength and LMP appears to have a reasonable basis. Several of the studies reviewed (Table 2.1) identify strength as a determinant of LMP. These confirm that stronger individuals exhibit better load-carriage ability. Moderate to low inverse correlations were observed in isokinetic hamstring, quadriceps and plantar flexor torque; squat strength/endurance; and isometric hand grip, upper torso and mid torso strength (Dziados *et al.,* 1988; Mello *et al.,* 1988; Knapik *et al.,* 1990; Rayson *et al.,* 1993; Harman and Frykman, 1995). The direction of this relationship was also confirmed in isometric measures of leg / back strength, and arm flexion endurance during multivariate analyses (Rayson *et al.,* 2000; Rayson *et al.,* 2002a).

One seminal training study highlighted the potential importance of strength for LMP. Kraemer *et al.,* (1987; 2004) observed that individuals involved in concurrent aerobic and resistance training programmes significantly improved 3.2 km LMP, however individuals who solely conducted resistance or endurance training did not improve. This suggested that strength alone was not an appropriate training stimulus for improving LMP, but when prescribed in combination, endurance and resistance training maybe complementary to the development of LMP. The combined influence of strength and endurance confirms the multi-factorial nature of load-carriage.

2.3.4 Task specificity

Specificity is a strongly advocated principle (Tanaka, 1994), both for training, and the evaluation of sporting / occupational performances. A dearth of the LM literature has considered this concept in the determination of LMP (Table 2.1). Instead, the LM research has included generic field measures (e.g. 2.4 km run, hand grip strength and skinfold measurements), and more labour intensive laboratory measures (e.g. $\dot{V}O_2max$, isokinetic dynamometry and underwater weighing) in the development of models for LMP. Most of these studies overlooked specific measures of LM which is surprising considering that the established running literature strongly focuses on task specific measures (Joyner, 1991; Brandon, 1995; Wood, 1999; Ingham *et al.*, 2008).

Loaded marching is a highly specific exercise modality (Vanderburgh and Flanagan, 2000), and it cannot be substituted with other criterion tests (e.g. running) as these impose a body mass bias against heavier individuals (Vanderburgh and Mahar, 1995). Just two studies have investigated the influence of specific LM measures on subsequent LMP (Simpson *et al.*, 2006; Griggs *et al.*, 2008). Simpson *et al.*, (2006) observed that maximal load-carriage treadmill test duration (r=-0.66), and 3.2 km LMP (r=0.77) were the only two measures to correlate with 29 km (20 kg) LMP. Furthermore, the regression model revealed the maximal load-carriage treadmill test duration best explained 3.2 km LMP (R^2=0.51), and both 3.2 km LMP and LM $\dot{V}O_2max$ best explained 29 km LMP (R^2=0.70). However, whilst these task specific measures emerged as better determinants of LMP compared to general components of fitness, the strength of these associations were similar to the general fitness measures reported elsewhere (Table 2.1). Others confirmed the validity of a LM specific test to predict 12.8 km LMP (Griggs *et al.*, 2008). Here, 2.4 km LMP explained 23 % more variance (74% vs. 51%), and reduced prediction error by almost 50% (124 vs. 224 s) compared to 2.4 km running performance and body mass measures.

Summary: Parameters of aerobic fitness, anthropometry and strength influence LMP. Aerobic fitness is widely applied and appears to be the most prevalent determinant of LMP. As the magnitude of a load increases absolute $\dot{V}O_2max$ increases in importance, suggesting that both fat free mass and maximal oxygen delivery capacity are important to factors when carrying absolute loads (e.g. 25 kg). The role of other parameters of aerobic performance has been ignored in the LM literature. Exercise economy (or loaded marching economy) is one such measure, which has the potential to be a determinant of LMP.

Anthropometric measures revealed heavier, larger and less lean individuals perform better in load carriage tasks. When absolute loads are involved, the likely explanation for this is that these individuals will carry a smaller load relative as a percentage of body mass, thereby reducing the relative amount of physical stress. As the structure of a biological organism dictates its function, strength's influence on LMP may be a reflection of anthropometry characteristics. Thus stronger, hence larger individuals perform better in load-carriage tasks. Finally, the role of task specificity in determining LMP needs further investigation.

2.4 The effect of physical training on loaded marching performance

The determinants of LMP (i.e. parameters of aerobic fitness, anthropometry, strength) outline the ideal physical profile required for LMP, and as such these components of fitness should be the focus of a physical training intervention to improve LMP. The paucity of data investigating the influence of specific sub-maximal (i.e. LME), and maximal measures of LMP (section 2.3) indicates the current knowledge of the determinants of LMP is incomplete. This may have implications on training prescription advice for improving LMP.

It has been suggested that existing military training programmes do not develop occupational fitness adequately (Williams *et al.*, 1999). In this study, changes in LMP were assessed over 11 weeks of British Army phase 1 training. Seventy-one periods of PT were undertaken consisting of; sports (19 periods), circuit training (18), agility (10), swimming (9), running (7), material handling (2), and prolonged bouts of load-carriage and drill practice. The effectiveness of the training programme was judged against the improvement in a 3.2 km, 15 and 25 kg LMP. Loaded marching performance significantly improved by 15.7 % in male recruits (25 kg), but not in a mixed gender platoon that carried the 15 kg loads. Furthermore, two out of seven measures of material handling performance (i.e. maximal box lift, repetitive lift and carry and loaded marching) improved over the 11 weeks, which coincided with an absence in strength development. These findings suggested that British Army recruit training was inadequate at developing physical performance in occupationally relevant tasks (i.e. LM), which was of direct concern to the British Army.

2.4.1 Training interventions

Kraemer *et al.,* (1987; 2004) made the first attempt to improve PT for LMP. Thirty-five trained male soldiers were divided into 4 separate training groups, and participated in 12 weeks of physical training, 4 days per week. The four groups were: (Group 1) high intensity endurance training with upper and lower body strength; (Group 2) high intensity endurance training with upper body strength only; (Group 3) upper and lower body strength only; and (Group 4) high intensity endurance training only. A combination of steady state / interval running, and 5 – 15 repetition maximum (RM) constituted the endurance training and strength training programmes, respectively. Loaded marching performance was measured over 3.2 km with a load of ~45 kg, and significantly improved by 14% and 11% in the endurance training groups that were supplemented with either a combination of upper and lower (Group 1) or just upper body strength (Group 2), respectively. This suggested that concurrent endurance and strength training programmes best suited the physical requirements of a high intensity, short duration load-carriage task. Furthermore, no significant differences existed between these groups 1 and 2, hence the benefit of additional lower body strength training for LMP may be negligible. Although the findings of this seminal study made intuitive sense (i.e. aerobic fitness is important during ambulatory tasks, and strength is important when carrying absolute loads), design weaknesses were present. These included small group sample sizes (n = 8/9), differences in baseline aerobic fitness levels (3.2-km run time of 13.4 to 15.4 minutes), two-fold differences in total training volume in the concurrent training groups (8 vs. 4 sessions·week^{-1}), and a high frequency of PT, which maybe unrealistic for busy military training programmes. Subsequently, several PT programmes have been supplemented with strength training, to enhance the potential gains in occupational tasks in female (Knapik and Gerber, 1996; Harman *et al.,* 1996; 1998; Kraemer *et al.,* 2001; Reynolds *et al.,* 2001), and male populations (Williams *et al.,* 2002; Visser *et al.,* 2005), and have reported mixed outcomes.

The findings of the seminal load-carriage training studies (Kraemer *et al.,* 1987;2004) were subsequently confirmed. A 24-week intervention, on untrained women matched for body size and strength (but not aerobic fitness) improved upon the original study design (Kraemer *et al.,* 2001). With a larger group sample size, four different strength interventions of 3 to 12 RM, and one field-based strength intervention, 3.2 km (34.1 kg)

LMP significantly improved. The potential importance of resistance training for LMP was emphasised as the group exposed to aerobic training alone (5 days·week^{-1}, 25 to 40 minutes·day^{-1}) was the only training group not to improve LMP.

A concurrent endurance and strength training programme in women yielded the largest training response reported in LMP. Here, 32 civilian women undertook a PT programme initially over 14 weeks (Harman *et al.*, 1996), which subsequently extended to 24 weeks (Harman *et al.*, 1998; Reynolds *et al.*, 2001). Training was conducted 5-days·week^{-1} (4 strength , 2 run, 1 hiking/marching, and 2 varied drills), and each session lasted 1 to 1.5 hours. Loaded marching performance with 34 kg over 3.2 km improved by 8.6 minutes, from 36.2 to 27.6 minutes (23.8%). This is the largest reported training response in LMP in the literature. This improvement was underpinned by gains of 14 % in $\dot{V}O_2max$, and 18 % to 47 % in manual material handling, which are higher than those previously reported (Knapik and Gerber, 1996). However, the absence of an exclusion criteria for low baseline fitness levels in these civilian women might explain the size of this training response. This is supported by the inverse relationship observed between baseline test performance and the magnitude of the subsequent training response in this study. Based upon the population of volunteers and the relatively high volume of training, it might be postulated that these results represent the upper limits of a LMP training response. Consequently, it might be unrealistic for trained new entry Army recruits that are relatively homogenous for fitness, entering an arduous period of initial military training to improve LMP by anything approaching this magnitude.

Other studies have reported very modest improvements in LMP with a concurrent training intervention (Knapik and Gerber, 1996). Here 13 female soldiers trained for 5 days·week^{-1} (1 hour·day^{-1}) over 14 weeks. The three resistance (whole body, 13 RM) and 2 endurance training (steady state and interval) sessions conducted each week improved LMP over 5 km, with 19 kg by 3.6%. This modest improvement occurred despite substantial improvements in 3.2 km run time (9.4%), manual material handling (14% to 24%), and 1 RM (15% to 39%). This small response in LMP was markedly less than the 11% to 14% (Kraemer *et al.*, 1987; 2004), and 23.8% (Harman *et al.*, 1998; Reynolds *et al.*, 2001) previously reported.

It is good scientific practice to incorporate a control group in the experimental design of PT intervention, as this will allow the true effect of the intervention to be quantified, and separated from sources of noise (e.g. a learning effect) in the data collection process. The use of controls was employed by Kraemer *et al.,* (1987; 2001; 2004). As previously discussed, 11 weeks of standard British Army military training elicited both marked (15.7%) and modest (3.6%) improvements in LMP (Williams *et al.,* 1999). Subsequently, a modified military training programme, with a strength training component was conducted with the aim of improving material handling and general physical fitness (Williams *et al.,* 2002). The total number of PT periods remained the same (71), however 28 periods of strength (whole body, 6 RM) were supplemented within this quota, which was made possible by truncating the volume of circuit training and general sport periods. Offsetting this reduction in aerobic based activities, the number of periods dedicated to endurance training was doubled from 7 to 15, and within these sessions loaded marching with individually prescribed loads was conducted. Additionally, prolonged bouts of load-carriage and drill practice were also part of the training programme. Over a 3.2 km course, male (25 kg) and mixed gender (15 kg) platoons significantly improved LMP by (16.7%), and (8.9%), respectively. This was also accompanied by improvements in material handling (8.0% to 18.5%), estimated $\dot{V}O_2$max (9.6%) and dynamic strength (14.2%). Using previous data (Williams *et al.,* 1999) as a control, the strength intervention produced a significantly greater training response in LMP with 15 kg, but not 25 kg LMP. Furthermore, body mass, fat free mass, absolute $\dot{V}O_2$max, dynamic strength, and one out of four material handling tasks also improved more in the modified training group. This suggested that initial military training programmes can develop body size, strength, $\dot{V}O_2$max, and to a degree LMP with a reduction in aerobic training volume, and an emphasis on strength development. One unexplained finding from these studies is the lack of difference in LMP during the 25 kg load-carriage task. Based on the evidence from the determinants of LMP (section 2.3) it might be anticipated that an increase in strength would be of more benefit to heavier (25 kg), not lighter (15 kg) load-carriage tasks. One possible explanation for this counter-intuitive observation is that both males and females conducted the 15 kg, and only males the 25 kg load-carriage tasks. Thus, females (with lower initial strength) may have benefited more from the resistance intervention than the males, causing LMP to improve to a greater extent in the lighter load-carriage task.

Several (Harman *et al.*, 1996; 1998; Williams *et al.*, 1999; Reynolds *et al.*, 2001; Williams *et al.*, 2002), but not all (Kraemer *et al.*,1987; 2001: 2004; Knapik and Gerber, 1996) LMP training studies have used the exercise modality of LM within a training programme. Just two studies have applied LM training alone as the sole independent variable during a training intervention. In addition to a standardised strength and endurance PT programme, Knapik *et al.*, (1990) conducted 9 weeks of loaded march training on 137 soldiers. Four different training groups were formed that included, no march training, or march training once, twice or four times per month. The LMP training response was measured over a distance of 20 km, with a load of 46 kg. It was concluded that LMP could be improved more effectively when training was conducted ideally 4, or at least 2 times per month. However, marked changes in the ambient temperature between the pre and post test conditions (5-21 °C), meant the heat stress of the post test was responsible for poorer performances. Thus, conclusions could not based on within group (pre to post) comparisons, instead they were based on between group comparisons of LMP in the post test alone. This was not ideal as it only reflected differences in LMP between the various training groups, not the magnitude of the training response, hence the efficacy of the intervention. Furthermore, at baseline there was a trend for the groups that trained 2 and 4 times per month to be faster load-carriers ($P>0.05$), which may have biased these findings.

Visser *et al.*, (2005) also used a standardised endurance and strength training programme, manipulating only the load-carriage aspect of training. Four different loaded marching training programmes were conducted on 76 male and female officers from the Netherlands Royal Military Academy, over 8 weeks. Training either focused on the duration (march distances of 8.3 km to 16.5 km), or the intensity of march training (load sizes of 45% to 67.5% of body mass for males and 35% to 55% of body mass for females). In addition the frequency of training was varied between two or four times per month. Pooled data showed that LMP improved by 5% over a 3.2-km march distance. Furthermore, an incremental loaded march test reported intensity-based loaded march training produced a two-fold greater training response than duration-based training (13.5% vs. 6.5%). Similarly, weekly bouts of loaded march training were significantly more effective then bi-monthly training bouts (12.6% vs. 7.4%). Thus regular, weekly bouts of loaded march training of short duration and high intensity were the most effective at improving LMP. However, whilst the short duration, high intensity model of LM training maximised performance it was confounded by the fact this was the most injurious loaded marching training intervention.

As such it was suggested that LM training should be conducted every 10 days, which might be sufficient to find the correct health-performance balance.

2.4.2 Individualising the training emphasis

The principle of 'one size fits all' is common to military PT programmes. This is understandable considering the high number of training sessions Army Physical Training Corps Instructors (APTCI) are required to deliver, to large numbers of recruits. Although this is a pragmatic solution to training prescription on a mass scale, it is not a *smart* approach to training. Consequently, the training needs of recruits entering the British Army may not be optimally met. Recognising this need, it has been proposed that recruits entering the British Army should be *streamed* into ability groups according to aerobic fitness, to reduce the injury risk, and optimise the physical stress (Bilzon, 2003). Deriving from this concept, one study has suggested that the emphasis of PT (i.e. endurance or strength) should be varied on an individual basis to improve the LMP response to training (Williams *et al.*, 2004).

Based on the proposition that concurrent endurance and strength training programmes may optimise the LMP training response (Kraemer *et al.*, 1987; 2001; 2004, Williams *et al.*,1999; 2002), Williams *et al.*, (2004) suggested that training prescription should be based around the diagnosis of a baseline ratio of endurance to strength. Such a ratio would highlight an individual's area of physical deficiency (i.e. endurance or strength), which would form the focus of a subsequent training intervention. The efficacy of the PT programme was appraised by the magnitude of the training response in LMP. In testing this hypothesis two groups were recruited over 10 weeks of training. One group was exposed to standard British Army training that was based around endurance/circuit training (Williams *et al.*, 1999), and the other group underwent a modified version of British Army training that had a strength training emphasis (Williams *et al.*, 2002). In the standard training programme 'good' (20.4 ± 3.3%) and 'poor' (10.3 ± 5.1%) responders to the training could be differentiated by the ratio of 20 m shuttle run to back extension strength or 38 cm upright pull strength. 'Good' responders had a lower endurance to strength ratio, whereas 'poor' responders had a significantly higher ratio score. (e.g. an endurance / strength ratio of 577 s / 1059 N = 0.55, would represent a 'good' responder, whereas a ratio of 615 s / 963 N = 0.64, would represent a 'poor' responder). This finding made intuitive sense, as the 'good' responders to standard military training (endurance/circuits based) were those who had the lower initial endurance scores. Hence, the endurance/circuit based training programme focused on the area of physical deficiency for the 'good', but not the 'poor' responders. This is the first study to suggest that simple field measures of endurance and strength, could help prescribe the most appropriate training emphasis on an individual basis (i.e. endurance or strength) to improve the LMP response. This simple and straightforward approach seems both feasible and realistic.

However, based on this ratio of endurance to strength, data from the modified training group (strength based) was unable to differentiate between 'good' (22.8 ± 3.7%) and 'poor' (11.0 ± 6.4%) responders. Here, baseline, 3.2 km LMP (25 kg) was the only measure able to differentiate between training responses. Interestingly, this LMP training response was inversely related to baseline LMP times in both training groups, which suggested that initial load-carriage ability was the most robust method of differentiating 'good' and 'poor' LMP responders. However, applying this measure to gauge training response would not work for two reasons. Firstly, Army training establishments do not expose recruits to load carriage of this intensity so early in training, based on injury concerns. Second, although the LMP test might project the magnitude of a LMP training response, it would not identify an individual's area of physical deficiency, hence a specific training emphasis could not be advised. This study proposed a novel method of prescribing training to improve the gains in LMP based upon quantifiable physical deficiencies, which seems worthy of future investigation.

Summary: Seminal work on load-carriage suggested that concurrent endurance and strength is the most effective training intervention to improve LMP. However, confirmatory work mostly conducted on females has shown mixed results, and has failed to include control groups. An improvement in LMP of ~24 % appears to represent the upper ceiling of LMP improvement in response to a training programme. Just two studies have investigated LMP response in actual army recruits, undergoing initial training. The benefit of resistance training on subsequent LMP was mixed. The identification of the physical deficiencies an individual may have, and prescribing a training programme to focus on those deficiencies may be an effective strategy to improve LMP, however this concept is in its infancy. Little research that has attempted to optimise PT for LMP. Consequently, well founded conclusions and recommendations as yet cannot be made.

2.5 Conclusion

Concepts from the running literature can be applied to LM to identify the determinants of LMP. This process revealed gaps in the LMP literature which was due to the limited fitness components included in the models of LMP. Thus, the influence of potential determinants of LMP such as maximal measures of LM, and LME seem both valid, but underrepresented in the literature. As such, these should be investigated in addition to accepted determinants of LMP (i.e. $\dot{V}O_2max$, running performance, anthropometry and strength), to further develop the understanding of mechanisms that determine LMP. This will allow PT to be designed to target the most important components of fitness that influence LMP.

CHAPTER 3

GENERAL METHODS

3.1 Subject recruitment

Chapters 4, 5, and 6 interrogated the validity and repeatability of the experimental measures applied in this thesis. This early work was undertaken in the British Association of Sport and Exercise Sciences (BASES) Accredited laboratories, at the University of Chichester. These preliminary investigations involved volunteers drawn from the population of staff, post-graduate and under-graduate students of the University. In Chapters 7, 8 and 9 this validated test battery was applied to a field environment where the physical training progression during initial British Army infantry training was evaluated over 24 to 26 weeks. The military recruits in Chapters 7 to 9 were younger (19.2 ± 2.0 vs. 28.8 ± 7.1 years), and lighter (69.0 ± 8.5 vs. 81.3 ± 6.6 kg), but had a similar relative $\dot{V}O_2max$ (53.1 ± 4.2 vs. 51.2 ± 3.7 $ml \cdot kg^{-1} \cdot min^{-1}$) in comparison to the civilian volunteers in Chapters 4 to 6.

3.2 Anthropometric testing

3.2.1 Body mass

Wearing shorts, T-shirt and socks body mass was measured to the nearest 0.1 kg (Seca, Hamburg, Germany).

3.2.2 Body stature

Wearing shorts, T-shirt and socks body stature was measured to the nearest mm (Avery Berkel, Smethwick, UK).

3.2.3 Identification of skinfold and circumferential girth sites

Skinfold and circumferential girth sites were identified from bony protrusions, or anatomical features that provided fixed points of reference. These were usually identified by palpation of the skin or by measuring a specified distance from an anatomical feature. The identification of anatomical landmarks, and subsequent derivation of each measurement site ensured that measurements were reliably taken from the same location for repeated observations. Failure to standardise measurement sites can increase skinfold variability by as much as 2-3 mm (Ruiz *et. al.*, 1971).

With a non-permanent pen, a single dot on the right hand side of the body identified each anatomical landmark / measurement site. The dot was equivalent to the centre of a horizontal line, or the centre of a cross, which are typical landmarking conventions. A single dot was used to optimise time, without compromising accuracy.

3.2.4 The skinfold measurement technique

A pair of skinfold callipers (Harpenden, Body Care, UK), accurate to within 0.2 mm were zeroed prior to use. Each measurement site was grasped between the fore-finger and thumb to create a fold of cutaneous and subcutaneous tissue. The callipers were placed perpendicular to the fold, 1 cm below the finger and thumb, at a mid-fingernail depth. In all cases the calliper jaws were released for a count of 2 seconds, at which point the skinfold was recorded. The average of 2 measurements were taken at each site, and had to agree ±10 %. Failure to agree required a third measurement to be taken, and an average of the 2 closest scores was recorded (International Society for the Advancement of Kinanthropometry ISAK, 2001).

3.2.5 The eight skinfold sites

Skinfolds were taken from eight skinfold sites (biceps, triceps, subscapular, iliac crest, supraspinale, abdomen, front thigh and medial calf) (Figure 3.1), and provided a restricted profile as suggested (ISAK, 2001). The additional measures of estimated percentage body fat, estimated fat-free mass, and estimated fat mass were derived from the four skinfold sites of the biceps, triceps, subscapular, and supraspinale (Durnin and Wormsley, 1974).

3.2.5.1 Upper arm (biceps and triceps)

A characteristic of data collection in the field in this thesis was that a high number of subjects were analysed over a limited time period. As such an adapted technique to find the measurement site for the biceps and triceps was employed that was efficient, but did not compromise measurement objectivity and reproducibility. Current convention (ISAK, 2001) advocates that both the *acromiale* and *radiale* bony landmarks are located in order to find the mid-point (*mid-acromiale-radiale*) of the upper arm. From this landmark both biceps and triceps measurements sites can easily be found.

This ISAK method was adapted to produce a more time efficient method. Here, the mid-point of the upper arm was measured from 8 subjects that ranged in stature (1.75 – 1.89 m), and hence humerus length (0.32 – 0.36 m). The range of limb lengths (0.34 m) was halved (0.17 m) and a wooden measurement rod cut to size. This standardised distance was measured from the *acromiale* to a point on the lateral part of the arm. From this point, the most anterior and posterior mid-line identified the sites for the vertical biceps and triceps folds, respectively. Although the distance between the *acromiale* and *radiale* would not have been equidistant for all subjects, it would be identical within-subjects. This was deemed essential for tracking temporal responses.

3.2.5.2 Subscapular

The *subscapulare* bony landmark was identified as the inferior angle of the scapula. A diagonal skinfold was taken at a 45° angle, 2 cm inferior and lateral.

3.2.5.3 Iliac crest

The *iliocristale* was identified as the most lateral aspect of the *iliac tubercle*. With the thumb placed on this landmark, a horizontal fold was taken immediately superior, and a mark was placed in the centre of this fold to identify the measurement site.

3.2.5.4 Supraspinale

A diagonal skinfold was taken from the point where two imaginary lines intersected (i.e. a horizontal line from the *iliocristale* towards the midline of the body, and a vertical imaginary line from the *iliospinale* to the natural crease where chest and arm meet).

3.2.5.5 Abdominal

A vertical fold was taken 5 cm horizontal from the midpoint of the naval.

3.2.5.6 Front thigh

The *front thigh* skinfold site was also adapted for reasons of measurement efficiency as stated for the *biceps* and *triceps* skinfolds. ISAK (2001) advocate that the *front thigh* skinfold be equidistant, between the *inguinal fold,* and the *anterior patella.* Initially, this technique was adhered to for the development of the adapted technique. Sitting in the upright position, with a 90° angle at the knee, the *front thigh* mid-point was identified using the ISAK method on a sample of 8 subjects. Subjects ranged in stature (1.75 – 1.89 m), and hence femur length (0.42 – 0.45 m). The mid-range limb length (0.44 m), was halved (0.22 m), and a wooden measured rod was prepared. The length of wood was placed in the *inguinal fold* along the midline of the quadriceps, which identified the measurement site.

3.2.5.7 Medial Calf

At the point of maximum circumference, a vertical fold was taken at the most medial part of the calf.

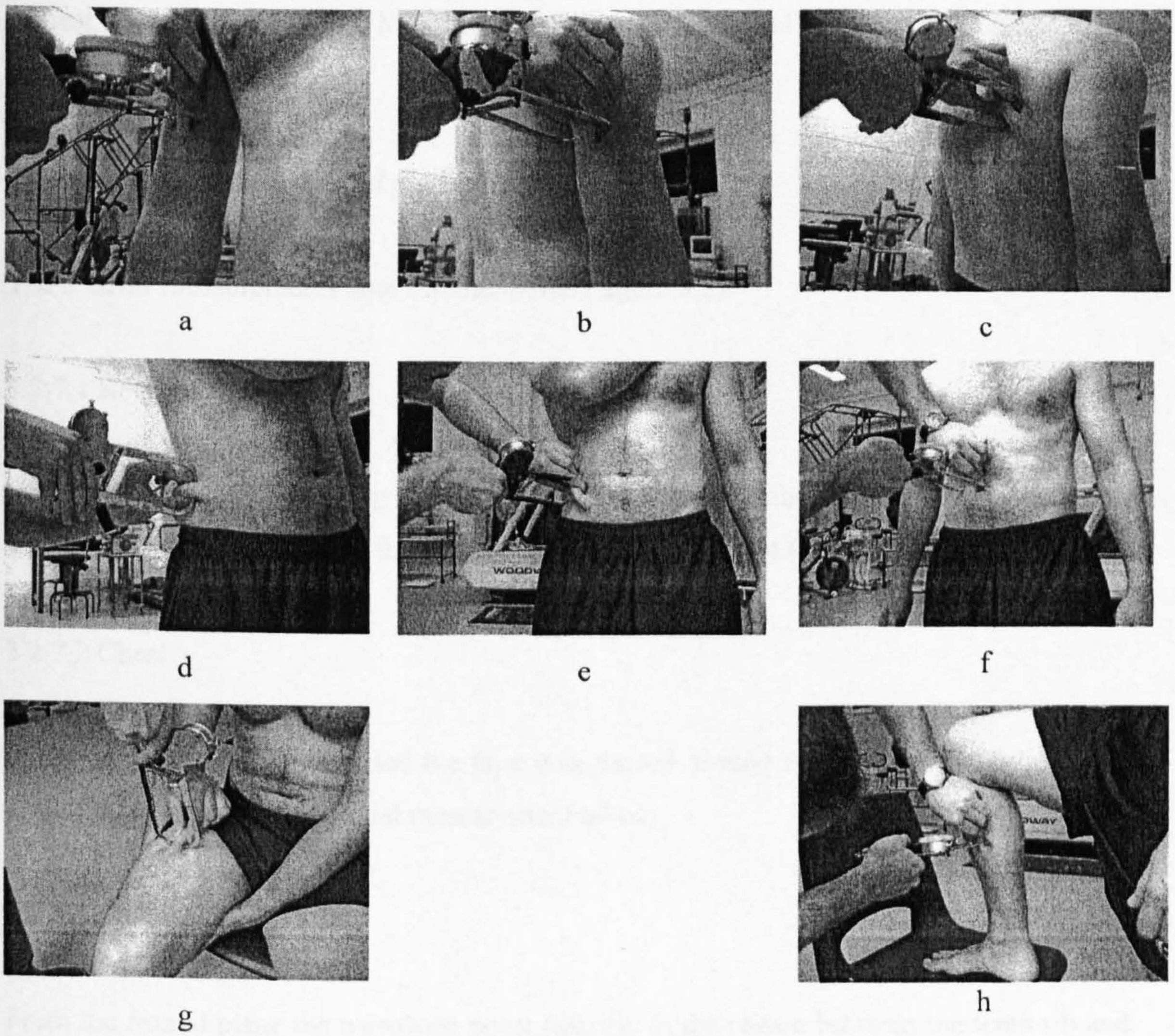


Figure 3.1. The skinfold measurement sites for the: Biceps (a); Triceps (b); Subscapular (c); Iliac Crest (d); Supraspinale (e); Abdomen (f); Front Thigh (g); and the Medial Calf (h).

3.2.6 The circumferential girth measurement technique

Girth measurements were taken from six sites (relaxed arm, chest, waist, gluteals, thigh and calf), and measured to the nearest millimetre using a metal tape (Silverflex, Rabone Chesterman, England). The *cross-hand* technique was used to measure all girths. The tape was maintained perpendicular to the longitudinal axis of the body segment measured. Visual cues such as, an indentation on the skin (tape too tight), or gaps appearing between the tape and skin (tape too loose) were used to apply consistent levels of tension during each measurement. The mean of 2 girth measurements was taken at each site. Failure to agree within ±2% required a third measurement to be taken, and the mean of the two nearest measures calculated. Measurements were always read with the eyes at the same level as the tape (figure 3.2).

3.2.7 The six circumferential girth sites

The 6 girths measurements sites are illustrated (figure 3.2).

3.2.7.1 Relaxed Arm

With the arm relaxed, hanging beside the body, and slightly abducted the tape was passed around the arm at the level of the biceps / triceps landmark site (section 3.2.5.1)

3.2.7.2 Chest

With the arms slightly abducted the tape was passed around the torso at the level of the *areola* (nipple), and an end tidal measurement taken.

3.2.7.3 Waist

From the frontal plane the narrowest point (usually in the region between the tenth rib and the *iliac crest*) was estimated and an end tidal measurement taken.

3.2.7.4 Gluteals

Viewed from the sagittal plane, the largest posterior protuberance of the buttocks indicated the measurement site.

3.2.7.5 Thigh

In a standing position, the front thigh landmark site (section 3.2.5.6) marked the level at which the measurement was taken.

3.2.7.6 Calf

In a standing position the largest circumferential girth signified the point of measurement.

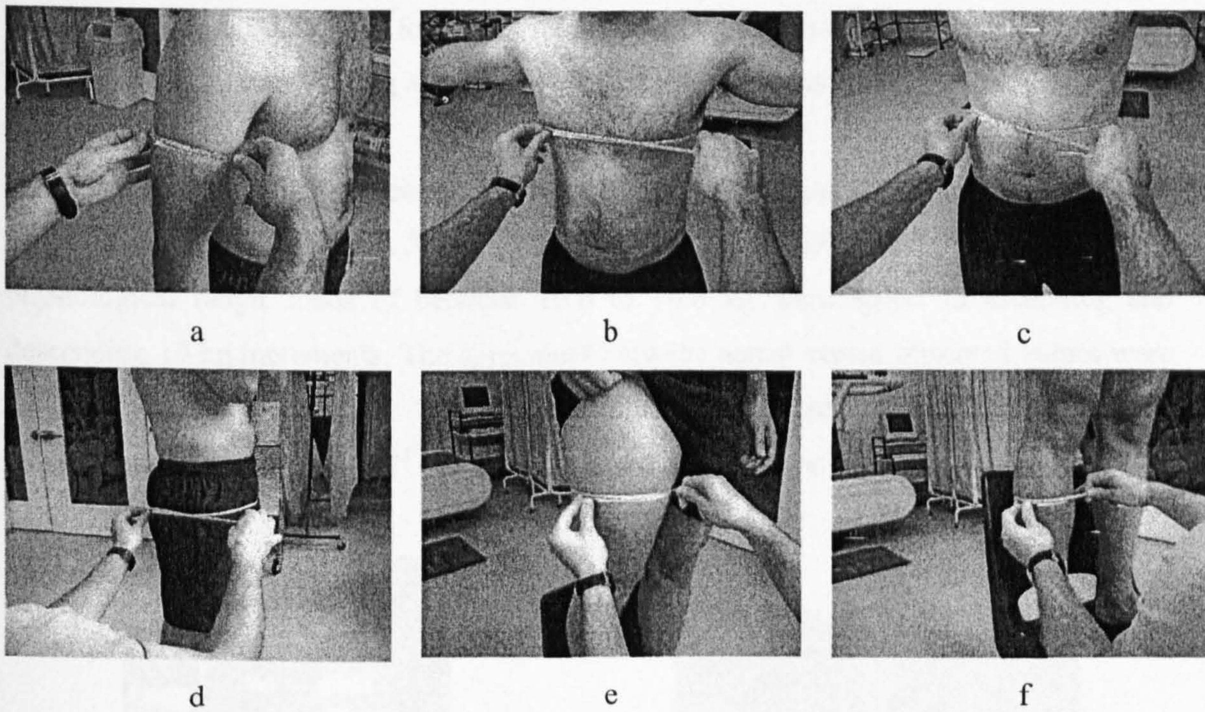


Figure 3.2. The girth measurement sites for the: Relaxed Arm (a); Chest (b); Waist (c); Gluteals (d); Thigh (e); and Calf (f).

3.3 Isometric strength measures

Verbal instruction, practical demonstration, and one sub-maximal warm up repetition at ~50% of maximum preceded all maximum isometric efforts. For each effort subjects were instructed to gradually increase the force over 2 to 3 s, then apply a maximal force for a period of 2 to 3 s. A rest period of 30 s separated each repetition (British Army, 2002).

3.3.1 Hand grip strength

A slightly abducted, straight arm was held at the side of the body for all hand grip measurements (Figure 3.3). The hand grip dynamometer (Model 5001, Takei, Japan) was set to a single standardised grip span of 6 cm. Using the dominant arm subjects were instructed to exert a maximal voluntary force.

Preliminary laboratory observations were conducted to establish the number of repetitions required to elicit a peak force. A total of 5 repetitions, separated by 30 s rest (British Army, 2002) were undertaken. Peak force was attained during the first repetition, thus to ensure peak force was attained during all subsequent trials 3 repetitions were performed.

Checks for measurement linearity were made by hanging calibrated weights from the dynamometer handle using a T-bar, chain and sling hoist (Figure 3.4). To reflect the physiological range, loads of between 10.6 to 70.6 kg were added in ascending and descending 10 kg increments. The agreement between actual versus observed values were fitted with a linear function. The slope (1.0071), y-axis intercept (-0.6588), and the coefficient of determination (R^2 = 0.9997) demonstrated that the response was linear.

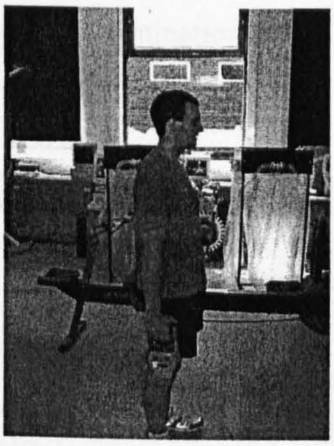


Figure 3.3. The hand grip dynamometer technique.

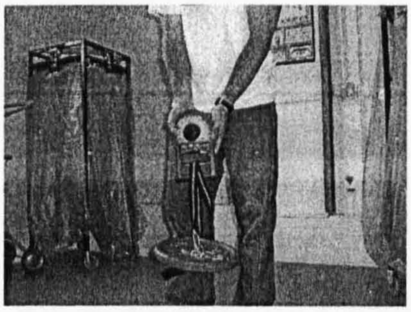


Figure 3.4. The method used to evaluate the linearity of the hand grip dynamometer.

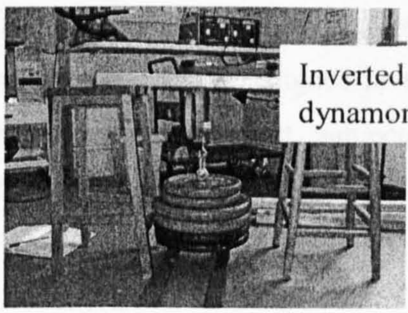
3.3.2 *Static Lift Strength (SLS)*

Potential British Army recruits are required to perform the SLS as part of the Physical Selection Standards Recruits (PSS(R)). With the T-bar positioned 38 cm above the dynamometer platform (Knapik *et al.*, 1981), leg and back strength were measured (5402, Takei, Japan). The validity and reliability of the dynamometer has been confirmed previously (Coldwells *et al.*, 1994). Only if a safe 'squat' position was assumed were subjects asked to exert maximal force (Figure 3.5). Previously, this technique has correlated with isometric leg and back strength, as well as dynamic measures of muscular performance (Birch *et al.*, 1994).

Preliminary laboratory observations were conducted to establish the number of repetitions required to elicit a peak force. A total of 5 repetitions, separated by 30 s rest were undertaken (British Army, 2002). Peak force was normally attained during the third repetition, thus during all subsequent trials 3 repetitions were performed.

Checks for measurement linearity were made by inverting the dynamometer. Calibrated weights (20 to 100 kg) were hung from the T-bar in ascending and descending 10 kg increments (Figure 3.6). The agreement between actual versus observed measurements were fitted with a linear function. The slope (0.9905), y-axis intercept (0.5), and the coefficient of determination ($R^2 = 0.9997$) demonstrated that the response was linear.

Figure 3.5. The Static Lift Strength technique.



3.6. The method used to evaluate the linearity of the Static Lift Strength dynamometer.

3.4 Dynamic strength measures

The sensitivity of isometric measures to track training adaptations of muscular function has been questioned (Wilson and Murphy, 1996), thus dynamic strength was also assessed. A dynamometer that measured whole body function was employed to assess dynamic strength (DYNO, Concept 2, Nottingham, UK). A rotating flywheel mechanism based on the same technology as Concept 2 indoor rowing ergometers quantified force (kg). An applied force rotated the flywheel, thereby creating wind resistance. The magnitude of the applied force was proportional to the speed of flywheel rotation, and hence the overall air resistance. Force was calculated by an electronic Force Monitor that measured the flywheel acceleration, speed of rotation and the moment of inertia. The derivation of force was a reflection of the force applied to the flywheel.

3.4.1 Strength measures

Three compound movements constituted the dynamic strength measures (Figure 3.7). For each individual the machine settings were adjusted, recorded, and replicated in subsequent trials. Prior to each movement, 3 sub-maximal warm up repetitions were performed. During the trials subjects were instructed to exert a force "as hard and fast as possible". The strength measures included, the seated chest press, seated back pull, and the leg press:

3.4.1.1 Seated chest press

This was deemed to be a measure of pectoral, anterior deltoid and triceps strength (Concept 2). Subjects sat upright with feet flat on the floor and the bar aligned level with the sternum. From the starting position with the bar touching the chest, subjects fully straighten their arms. To standardise range of motion subjects were instructed not to roll the shoulders forwards at the end of the movement (Figure 3.8a).

3.4.1.2 Seated back pull

This was deemed to be a measure of latissimus dorsi, trapezius, posterior deltoid, and biceps strength (Concept 2). Subjects sat upright with feet flat on the floor and the bar aligned level with the sternum. Taking a neutral grip the arms were fully extended and subjects pulled the bar towards the body until the hands touched the chest. To standardise the range of motion subjects were instructed to keep the chest in contact with the supportive pad (Figure 3.7b).

3.4.1.3 Leg press

This was deemed to be a measure of quadriceps, hamstrings, and gluteus maximus strength (Concept 2). Subjects sat upright with a 90° angle at the knee. A wooden block was fixed to the monorail to standardise the range of motion (Figure 3.8). Subjects were instructed to push through the heels, fully straighten the legs, and allow the toes to come away from the platform (Figure 3.7c).

3.4.2 Strength protocol design

To ensure reproducible results, Concept 2 DYNO recommend the standardisation of four variables. These are: load; range of movement; number of repetitions; and duration between repetitions.

3.4.2.1 Load

Eight dampers regulate air flow to the flywheel, which in turn alters the load (or drag). As more dampers are opened drag increases, thus a greater force is required to rotate the flywheel. Previously miscalculations have been reported when six or more dampers were open (i.e. the heaviest, slowest setting) (Concept 2, 2005). Thus just four consecutive dampers were opened for all test procedures, which equated to a mid-range load selection.

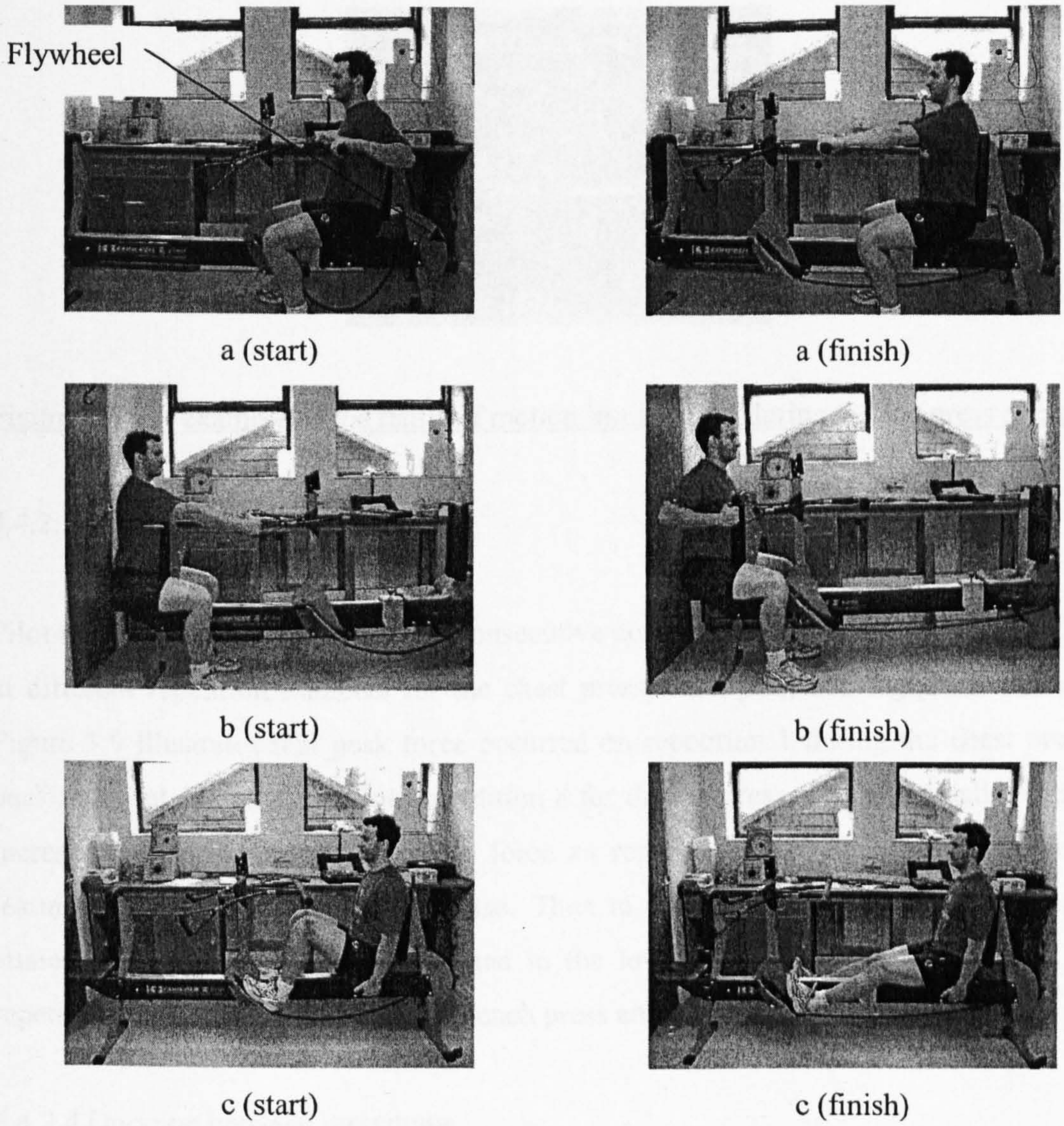


Figure 3.7. The three dynamic strength measures of the Concept 2 DYNO: Chest press (a); Back Pull (b); and the Leg Press (c).

3.4.2.2 Range of motion

The range of motion was standardised for each individual, as previously described (section 3.4.1). A change in range of motion can alter the acceleration / speed of rotation of the flywheel, resulting in measurement variability. Range of motion was standardised by adherence to a strict technique, replicating the machine setup within-subjects, and in the case of the leg press using a range of motion limiter (Figure 3.8).

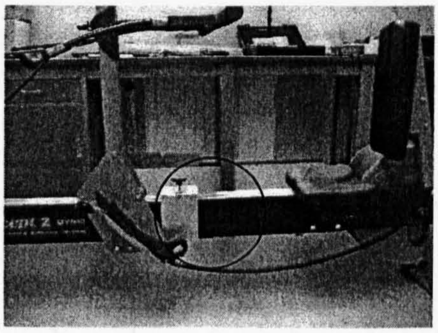


Figure 3.8. An example of the range of motion limiter used during the leg press exercise.

3.4.2.3 Number of repetitions

Pilot testing revealed that across 10 consecutive contractions, the peak force was achieved at different repetition numbers for the chest press, back pull, and leg press movements. Figure 3.9 illustrates that peak force occurred on repetition 1 during the chest press and back pull, but did not occur until repetition 8 for the leg press. The statistically significant increase ($P<0.05$) in peak leg press force as repetition number increased suggested a learning effect existed in this exercise. Thus to ensure a true peak measurement was attained, 10 repetitions were conducted in the lower body measures (leg press), and 5 repetitions on upper body measures (bench press and back pull).

3.4.2.4 Duration between repetitions

Exactly 5 seconds separated each repetition. This ensured the flywheel had a standardised time to slow down between repetitions.

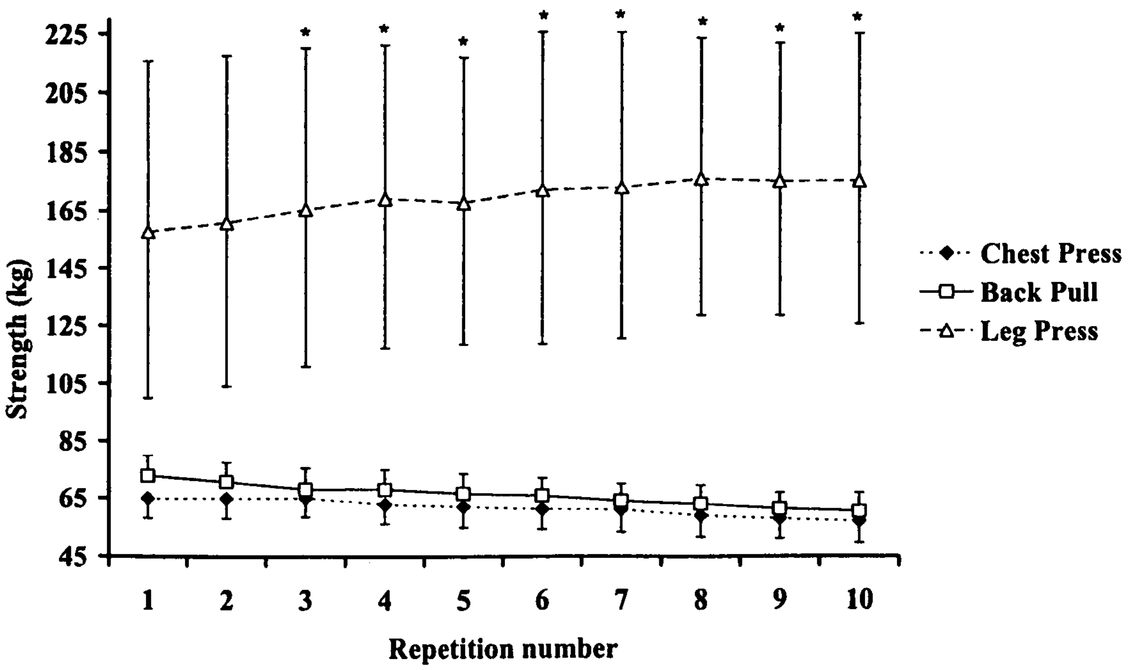


Figure 3.9. Measure-dependent occurrence of peak strength for the chest press, back pull and leg press, taken from 20 observations. * = different from repetition 1 (P<0.05).

3.5 Treadmill-based protocols

3.5.1 The treadmill

All laboratory tests were undertaken on a motorised, slat-belt treadmill (Ergo ELG 70, and PPS 70med-I, Woodway, Germany). A Woodway User System (WUS) controlled speed and gradient to the nearest 0.1 $km \cdot h^{-1}$, and 0.1%, respectively. The belt width was 0.7 m, which had a circumferential length of 5.22 m (Ergo ELG 70) and 3.6 m (PPS 70med-I). Both speed and gradient calibrations were conducted prior to each major data collection phase.

3.5.1.1 Treadmill speed calibration

Speed calibrations were performed at speeds of 4, 8, 12, 16 and 20 $km \cdot h^{-1}$. An identifiable mark placed on one slat of the treadmill allowed whole revolutions of the belt to be counted. For each 4 $km \cdot h^{-1}$ increment, 15 full revolutions were timed to the nearest one hundredth of a second. Thus, 15, 30, 45, 60 and 75 revolutions were timed at 4, 8, 12, 16 and 20 km·h^{-1}, respectively. The actual belt speed was calculated in km·h^{-1} (Equation 3.1) and data reported to the nearest two decimal places.

Actual belt speed =

$$(N^o \text{ revolutions} \times \text{belt length (m)} / \text{time of revolutions (s)} \times 60 \times 60 / 1000.$$

(Equation 3.1)

Differences between actual vs. displayed belt speed were no greater than ~ 1%. A typical data set is presented (Table 3.1).

Table. 3.1. The actual *versus* displayed treadmill belt speed

Display Speed (km·h^{-1})	N° Revolutions	Time Taken (s)	Actual Speed (km·h^{-1})
4.00	15	48.30	**4.02**
8.00	30	48.26	**8.05**
12.00	45	48.26	**12.08**
16.00	60	48.09	**16.16**
20.00	75	48.11	**20.20**

3.5.1.2 Treadmill gradient calibration

In the 0% gradient position, a spirit level was used to ensure that the treadmill was initially level. Agreement between observed and displayed % gradient was verified at 1, 2, 3, 4, 5, 10, 15, and 20%, with a gradient measuring device. The measuring device consisted of a 'plum line' that moved against a graduated scale that represented 1% increments in gradient.

3.5.2 Treadmill protocols

Treadmill protocols in this thesis consisted of sub-maximal load-carriage tests, as well as maximal running and maximal load-carriage tests. All load-carriage tests were conducted at a constant speed of 6.5 km·h^{-1}, which approximates the minimum average speed British Army recruits must maintain during the Basic Combat Fitness Test (BCFT) physical output test, undertaken towards the end of recruit training (British Army, 2002).

3.5.2.1 Sub-maximal load-carriage test 1

The load-carriage test consisted of one 5-minute unloaded walking warm-up at 6.5 km·h^{-1}, followed by 48 minutes of exercise, and was performed in Chapter 4. Wearing training shoes, and British Army issue lightweight Combat 95 (CS95) jacket and trousers, subjects conducted the 48-minute exercise bout on a motorised treadmill, at 6.5 km·h^{-1}. Four, 9-minute stages were performed with 4 different external loads (0, 15 20 and 25 kg), in a randomised order. The loads reflected the various CEGs in the British Army (British Army, 2002). Four minutes of unloaded walking at 6.5 km·h^{-1} separated each stage. Within each 9-minute stage Douglas Bag and On-Line expired gas collections were taken between minutes 4 – 6 and 7 – 9 of each stage (Figure 3.10).

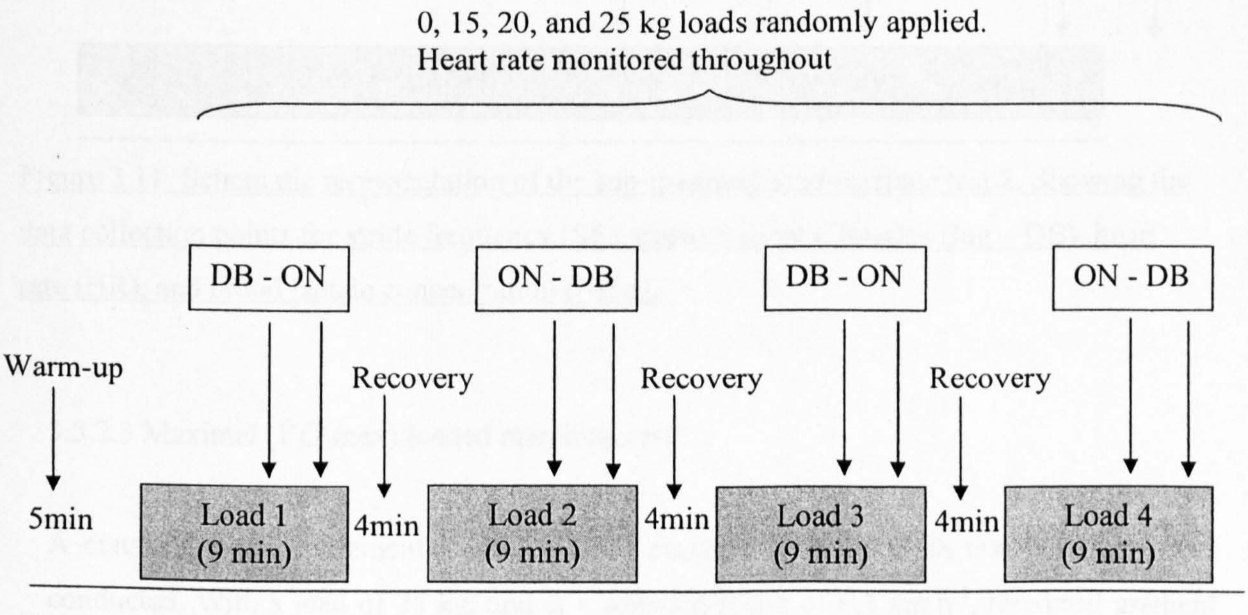


Figure 3.10. Schematic representation of the sub-maximal load-carriage test 1. DB and ON denote the use of Douglas Bag and On-line gas analysis systems.

3.5.2.2 Sub-maximal load-carriage test 2

A truncated sub-maximal load-carriage exercise model that reflected the CEG requirements of infantry recruits (i.e. 25 kg load-carriage task) was performed in Chapters 6, 7, 8, and 9. The exercise model consisted of one, 8-minute bout of load-carriage with a 25 kg load. The civilian subjects in Chapter 6 wore training shoes, whereas the infantry recruits in Chapters 7, 8, and 9 were asked to wear British Army issue boots. Heart rate was monitored continuously and recorded in the final 15 seconds of the stage. Stride frequency was measured over a one minute period between minutes 4 and 5. Expired gases were collected with the Douglas Bag approach between minutes 6 – 8, and 30 seconds post test duplicate finger tip capillary blood samples were drawn and analysed for whole blood lactate concentration (Figure 3.11).

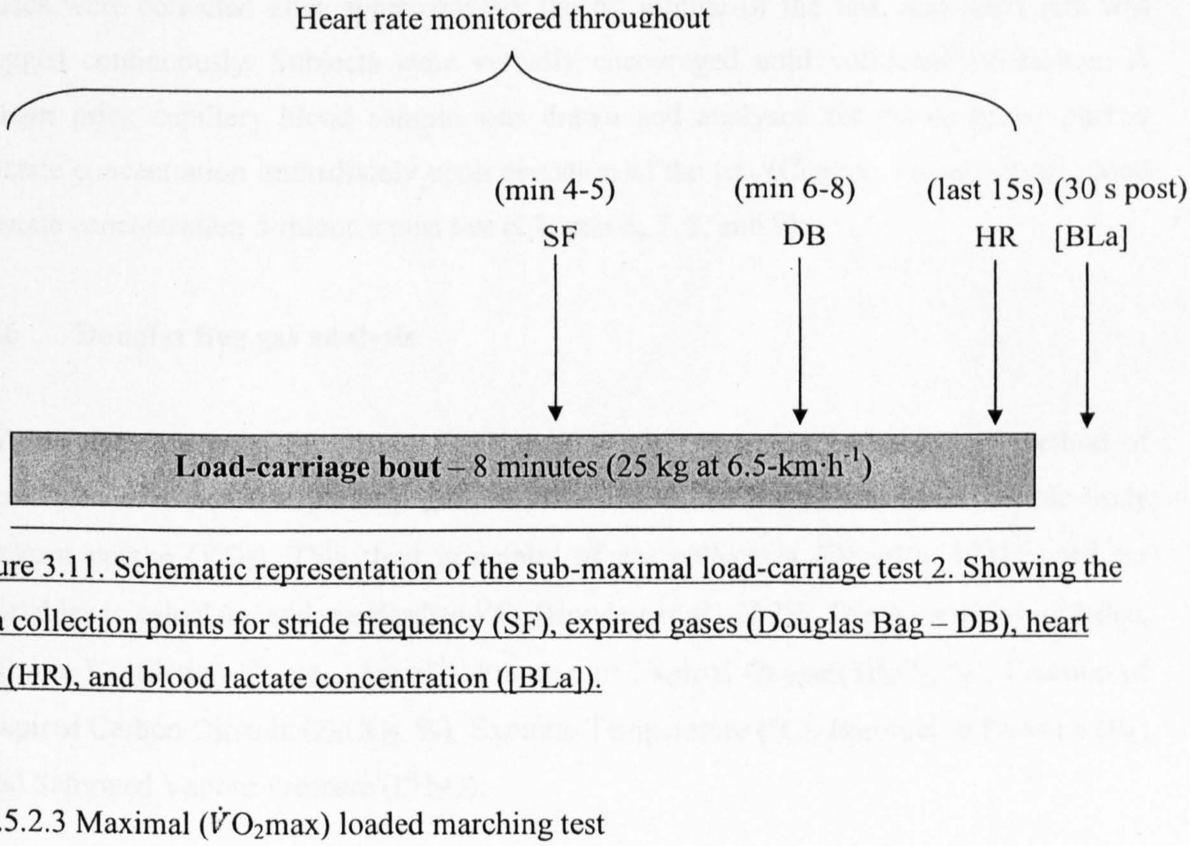


Figure 3.11. Schematic representation of the sub-maximal load-carriage test 2. Showing the data collection points for stride frequency (SF), expired gases (Douglas Bag – DB), heart rate (HR), and blood lactate concentration ([BLa]).

3.5.2.3 Maximal ($\dot{V}O_2max$) loaded marching test

A constant speed, incremental gradient (IG) maximal loaded march test ($\dot{V}O_2max$) was conducted. With a load of 25 kg, and at a constant speed of 6.5 km·h^{-1}, treadmill gradient was ramped from 0% at a rate of 1%·minute^{-1} (0.5% every 30 s). This protocol was adapted from that of Balke and Ware (1959). Expired gases were collected after approximately the

6^{th} minute of the test, and heart rate was logged continuously. Subjects were verbally encouraged until volitional exhaustion. A finger prick capillary blood sample was drawn and analysed for whole blood plasma lactate concentration immediately upon cessation of the test.

3.5.2.4 Maximal ($\dot{V}O_2max$) running test

An incremental speed (IS) maximal running test ($\dot{V}O_2max$) was conducted. An initial test speed of 7.0 $km{\cdot}h^{-1}$ (Chapter 5) was subsequently increased to 9.0 $km{\cdot}h^{-1}$ (Chapters 6, 7, 8, and 9), as this was a more comfortable initial running speed. This speed was maintained for the first 5 minutes of the test, during which time treadmill gradient increased from 1% to 5%. After 5 minutes the incline remained constant (5%), and treadmill speed was increased at a rate of 1.2 $km{\cdot}h^{-1}{\cdot}min^{-1}$ (0.2 $km{\cdot}h^{-1}$ every 10 s) (Draper *et al.*, 2003). Expired gases were collected after approximately the 6^{th} minute of the test, and heart rate was logged continuously. Subjects were verbally encouraged until volitional exhaustion. A finger prick capillary blood sample was drawn and analysed for whole blood plasma lactate concentration immediately upon cessation of the test (Chapter 5), and whole blood lactate concentration 5-minutes post test (Chapter 6, 7, 8, and 9).

3.6 Douglas Bag gas analysis

Within the context of this thesis the 'Douglas bag' technique refers to the method of collection of expired respiratory gases for the determination of the rate of whole body oxygen uptake ($\dot{V}O_2$). This 'first principle' of gas collection (Douglas, 1911) used six variables to calculate, and standardise $\dot{V}O_2$ (Howley *et al.*, 1995). These variables included, Minute Ventilation ($\dot{V}_{E\ ATPS}$, $l{\cdot}min^{-1}$), Fraction of Expired Oxygen (F_EO_2, %), Fraction of Expired Carbon Dioxide (F_ECO_2, %), Expirate Temperature (°C), Barometric Pressure (P_B), and Saturated Vapour Pressure (PH_2O).

In its simplest form $\dot{V}O_2$ can be calculated by subtracting expired $\dot{V}O_2$ from inspired $\dot{V}O_2$. A more detailed expression of this equation is presented (Equation 3.2):

$$\dot{V}O_2 = (\dot{V}_I \times F_IO_2) - (\dot{V}_E \times F_EO_2)$$

(Equation 3.2)

The Fraction of Inspired Oxygen (F_IO_2) was assumed to be constant (20.93%). The only unknown variable in Equation 3.2 required for the calculation $\dot{V}O_2$ was the Inspired Minute Ventilation ($\dot{V}_I$). This was calculated using the *Haldane Transformation*, which assumes that Nitrogen (N_2) is inert (neither used nor produced by the body), such that the volume of inhaled N_2 (F_IN_2) is assumed equivalent to the volume of exhaled N_2 (F_EN_2). Making these assumptions, and knowing three variables ($\dot{V}_E$, F_IO_2, and F_EO_2), $\dot{V}_I$ was calculated using (Equation 3.3).

$$\dot{V}_I = \frac{(\dot{V}_E \times F_EN_2)}{F_IN_2}$$

(Equation 3.3)

All $\dot{V}O_2$ measurements were standardised for temperature and pressure. This was essential as *Charles' Law* states that the volume of a gas is proportional to its temperature, and *Boyle's Law* states that the volume of a gas is inversely related to pressure. Thus, all $\dot{V}_E$, $\dot{V}O_2$, and $\dot{V}CO_2$ values in this thesis have been corrected to a Standardised Temperature (0°C or 273°K), Barometric Pressure (760 mmHg or sea level), and Saturated Vapour Pressure Dry (STPD) (Powers and Howley, 1997).

3.6.1 The collection of expired gas ($\dot{V}_{E\ ATPS}$).

Wearing a nose clip, subjects expired through a 2-way non-rebreathing mouthpiece (2700, Hans Rudolph, Kansas, USA) that was attached to a 1.8 m EVA breathing tube, on to a 200 l plastic Douglas bag (Cranlea & Co., Birmingham, UK). To allow continuous, serial gas collections, 4 Douglas bags were mounted on a rack. This allowed expirate to be directed into a specified bag, or vent to atmosphere (Figure 3.12). At the point of collection expirate reflected ambient conditions, hence was not standardised for temperature, pressure and saturation ($\dot{V}_{E\ ATPS}$).

The approximate dead space of this closed loop circuit was calculated by multiplying the length of breathing tube (1.8 m) by its internal cross sectional area (0.000908 m^2). Cross sectional area was calculated from the internal diameter of the tubing (0.034 m) with the formula ($\pi \cdot r^2$). For example, expirate collected in bag 1 had the furthest, and bag 4 the shortest distances to travel. These respective dead spaces ranged from (4.55 m × 0.000908 m^2), ~4.13 l (bag 1), to (1.82 m × 0.000908 m^2) ~1.65 l (bag 4). Thus, the minimum duration subjects were required to breathe through this system to flush the dead space varied dependent upon the bag number.

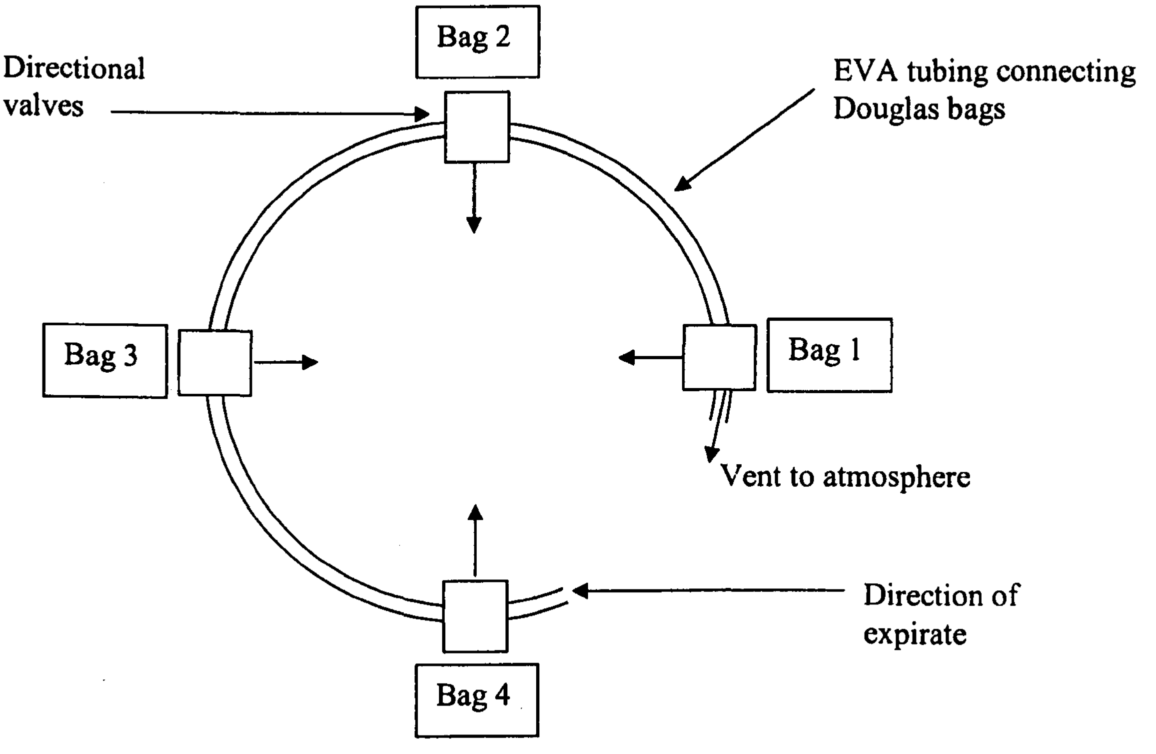


Figure 3.12. Schematic representation of the Douglas bags rack system that allowed continuous, serial collections of expired human gas.

For example loaded marching (the least intense exercise model used in this thesis), imposed an oxygen demand of 1.5 $l{\cdot}min^{-1}$ on the fitter subjects, which equated to a $\dot{V}_{E\ ATPS}$ of 30 $l{\cdot}min^{-1}$. Thus, in the worst case scenario (i.e. $\dot{V}_{E\ ATPS}$ 30 $l{\cdot}min^{-1}$, dead space 4.13 l (bag 1)) subjects were required to breathe through the mouthpiece for a minimum of 7 s to flush the dead space prior to collection.

With a lap stopwatch (100 memory, Fastime, Cranlea & Co., Birmingham, UK), whole breaths of expirate were collected and the collection time was recorded to the nearest one hundredth of a second. Collection periods varied from ~60 s during $\dot{V}O_2max$ assessments (sections 3.5.2.3 & 3.5.2.4) to ~120 s during the sub-maximal loaded marching assessments (sections 3.5.2.1 & 3.5.2.2). As previously mentioned, sub-maximal loaded marching tasks produced low ventilation volumes (e.g. $\dot{V}_{E\ ATPS}$ 30 l·min^{-1}, thus a $\dot{V}O_2$ of 1.5 l·min^{-1}). Such low ventilation volumes have been shown to increase the variability of $\dot{V}_E$ and calculated $\dot{V}O_2$, two to three-fold, respectively (Wood, 1999). This provided a rationale for all sub-maximal loaded marching gas collection periods to be increased to ~120 s.

Occasionally, Douglas bag racks were changed during a maximal $\dot{V}O_2$ tests. In such an instance a collection period of ~50 s was made in the final bag (bag 4) (Figure 3.12) prior to changing to a new set of Douglas bags. Approximately 5 s were required to disconnect the rack containing expirate, and re-connect an evacuated Douglas Bag rack. Throughout the transition the nose clip and mouthpiece were worn, and when re-connected a period of ~5 s was allowed to clear the dead space.

At the point of volitional exhaustion gas collection periods were often < 60 s during the final stage of the maximal running and loaded marching tests. Data from this final stage was only considered valid if a subject exercised for 30 s or longer (Wood, 1999), as this would ensure adequate ventilation volumes ($\dot{V}_{E\ ATPS} \geq 50$ l·min^{-1}). If the gas collection period was < 30 s the 60 s collection from the penultimate stage was recorded as $\dot{V}O_2max$.

3.6.2 The calibration and measurement of expired gas volumes ($\dot{V}_{E\ ATPS}$)

A dry gas meter measured expired gas volume (6162, Harvard Apparatus, Edenbridge, UK). Prior to each major data collection period the dry gas meter was calibrated by comparing the display reading to a known gas volume. To achieve this a 7 l syringe (4900, Hans Rudolph Inc., Kansas City, USA) injected 5 known volumes (35, 70, 105, 140 and 175 l) that spanned the physiological range. All four Douglas bags (Figure 3.12) were filled, and subsequently evacuated with the dry gas meter. The mean of the four Douglas bag collections were calculated for each of the five gas volumes. The display volume was then plotted against the known volume, and fitted with a linear trendline. The regression equation constant was removed by driving the intercept through zero, which left a single value that was the calibration correction factor. This calibration correction factor was used to improve the accuracy of $\dot{V}_{E\ ATPS}$, $\dot{V}_{E\ STPD}$ measurement, and $\dot{V}O_2$, $\dot{V}CO_2$ calculation.

3.6.3 Fraction of expired oxygen (F_EO_2) and carbon dioxide (F_ECO_2)

3.6.3.1 Avoiding contamination from prior analyses

To reduce the effect of expirate contamination, evacuated Douglas bags were flushed with ambient air prior to gas collections (Douglas, 1911). This was deemed to be important as the gas concentrations within the residual volume of an evacuated Douglas bag will vary depending upon the exercise intensity of previous gas collections (i.e. oxygen 15 to 18%, carbon dioxide 3 to 5%). This may have contaminated subsequent gas collections, hence introduce a source of error. Thus, all Douglas bags were flushed and evacuated with ~70-80 l of ambient air prior to gas collections, hence standardising the concentration of residual oxygen and carbon dioxide.

3.6.3.2 Standardising flow rate

A gas analyser (1440, Servomex, Crowborough, UK) that consisted of paramagnetic and infrared transducers measured the fraction of expired oxygen and carbon dioxide, respectively. An external vacuum pump (VP0140, Nitto, Watford, UK) drew aliquots of expirate from each Douglas bag, which was pumped through a pressurised system, at a standardised rate of 1 $l \cdot min^{-1}$ (Figure 3.13). As this partial pressure gas analyser was sensitive to flow rate, this was standardised using an ultrafine control flowmeter (Uniflux, Cache Instruments Ltd., Wakefiled, UK). Knowledge of flow rate, and the duration over which a sample was drawn allowed the volume of each aliquot to be calculated. The volume of this aliquot was subsequently added to the $\dot{V}_{E\ ATPS}$ measurement (section 3.6.2).

3.6.3.3 Standardising the water vapour of gases

During the calibration of the gas analyser and the determination of F_EO_2 and F_ECO_2, three gases of varying water vapour content were passed through the analyser. These included human expirate, ambient air, and calibration gas. Typically, human expirate is fully saturated, ambient air has a degree of saturation (dependent upon ambient temperature and humidity), and calibration gases are completely dry. If gases are not standardised for water vapour content, the determination of oxygen will be underestimated by as much as 25%

(Beaver, 1973). One technique to standardise the water vapour content of gases entering the analyser is to use drying agents. However, the degree dehumidification (hence standardisation) is dependent upon the type, batch, quantity and level of saturation of the drying agent. Furthermore, gases can be absorbed into the drying agent, which may contaminate subsequent analyses (Elia *et al.*, 1986).

Due to the limitations of drying agents, the water vapour content of gases were standardised by the humidification of gases. All gases were passed through a length of semi permeable Nafion® tubing (Cortex, Leipzig, Germany) that was submerged in distilled, room temperature water. The tubing equilibrated the gas to the ambient conditions, which in this case was complete 100% humidification (saturation). To prevent the humidified expirate from condensing in the inside of the analyser, it was cooled to a dew-point of 5 °C (manufacturer's recommendations) by a thermoelectric gas cooler (PKE4, Bühler, Patterson Instruments Ltd., Billericay, UK). Consequently, all three gases (i.e. expirate, ambient air, and calibration) were standardised for water vapour content (level of humidification) prior to entering the gas analyser (Figure 3.13).

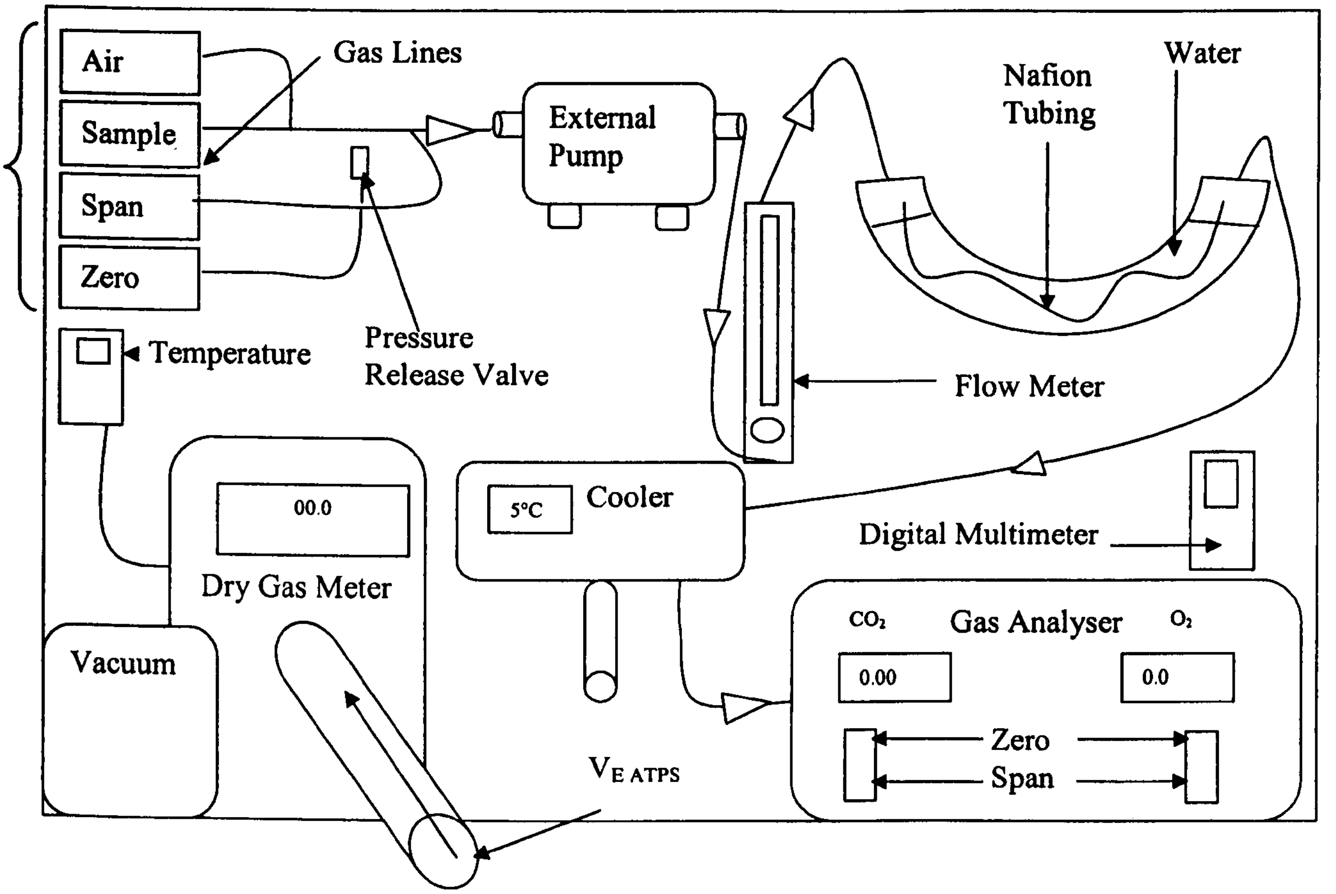


Figure 3.13. Schematic representation of the Douglas bag analysis system for the determination of $\dot{V}O_2$, from $\dot{V}_{E\ ATPS}$, F_EO_2 and F_ECO_2.

3.6.3.4 Analyser response time

At a standardised flow rate of 1.0 $l \cdot min^{-1}$ the response kinetics of oxygen and carbon dioxide were analysed at 10 s time intervals, over a duration of 110 s. Carbon dioxide readings were displayed to 2 decimal places, however the oxygen transducer displayed data to one decimal place. To improve the resolution of F_EO_2 measurements (i.e. measurement to 2 decimal places) a digital multimeter (328, Rapid Electronics Ltd., Colchester, UK) was connected to the RS232 port on the gas analyser (Figure 3.13). Across the sub-maximal to maximal exercise continuum the improved resolution would theoretically increase the sensitivity of F_EO_2 by between 2.4 to 3.5%, or 0.5 to 1.9 $mlO_2 \cdot kg^{-1} \cdot min^{-1}$, respectively. Thirty observations showed that oxygen and carbon dioxide readings reached a steady state after 80 s. Thus all F_EO_2 and F_ECO_2 readings were recorded at this fixed time point.

3.6.3.5 Calibration of the gas analyser for the determination of F_EO_2 and F_ECO_2

A single calibration was completed for each subject during the preliminary laboratory studies (Chapters 4 , 5, and 6). However, in Chapters 7, 8, and 9 the analyser was re-calibrated after two subjects had completed the sub-maximal loaded marching 2 (section 3.5.2.2), and the maximal running (section 3.5.2.4) tests. Each time an identical, three-point calibration involved setting fixed "zero" and "span" points.

The role of zero and span adjustments were to establish the intercept and slope, creating linearity over the physiological range. The calibration process involved drawing through nitrogen, with a minimum purity of 99.998% (BOC, Guildford, UK) through the gas analyser to establish the zero setting. Subsequently, ambient air set the span for oxygen at 20.93%. A mixed gas balanced with nitrogen (BOC, Guildford, UK) set the span for carbon dioxide at 5.18%, and the linearity of the oxygen calibration was checked with a 16.05% oxygen gas. To verify the calibration, ambient air was measured, and a reading of 20.93% for oxygen and 0.03% for carbon dioxide was required (±0.01). If these readings were not achieved, the process was repeated until the gas analyser settled on normal ambient values. All gases were drawn through the system at a standardised flow rate (1.0 $l \cdot min^{-1}$), for a timed duration of 90 s.

3.6.4 Barometric pressure

Barometric pressure was measured on a digital weather forecast station (BAA913HG, Oregon Scientific Inc., Tualatin, USA) during Chapters 6, 7, 8, and 9. The forecast station expressed pressure in millibars (mb), at a resolution of 1 mb, over the range of 795 to 1050 mb. To allow the derivation of $\dot{V}O_2$, barometric pressure was converted into millimetres of mercury (mmHg), using a meteorological conversion factor (mmHg = mb × 0.75218) (www.csgnetwork.com). This converted pressure reading was regularly checked for agreement against a mercury barometer (Stanley and Co., London, UK), which had a resolution of 0.05 mmHg, and was used in Chapters 4 and 5.

3.6.5 Temperature

The temperature of expirate was measured coincident with the measurement of gas volume ($\dot{V}_{E\ ATPS}$ – section 3.6.2). A gas monitoring thermometer (Cranlea & Co., Birmingham, UK) was inserted into the inlet port of the dry gas meter (Figure 3.13) and measured expirate temperature immediately prior to entering the dry gas meter. The thermometer had a resolution of 0.1 °C.

3.7 Online gas analysis

An automated on-line gas analysis system (VMax29, Sensormedics, Yorba Linda, USA) was compared to the Douglas Bag system in Chapter 4, to establish the most valid and reliable gas analysis system for future investigations. Prior to using the analyser at least 30 minutes were allowed for the machine to warm-up.

3.7.1 The determination of flow rate (gas volume)

The On-line system measured flow rate with a hot wire anemometer (mass flow sensor), that was inserted into a 2.6 l mixing chamber in the rear of the device. Within the mass flow sensor the rate at which heat was lost from a pair of heated stainless steel wires was electronically integrated, and flow rate subsequently calculated. Relative humidity was assumed to be constant, whereas both ambient temperature and barometric pressure (P_B) were continuously monitored for standardisation.

3.7.2 Flow rate calibration

A flow rate calibration was conducted prior to each test. First, room air was used to purge the mass flow sensor with two strokes of a 3 l syringe (Sensormedics, Yorba Linda, USA). Following a 10 s pause, the mass flow sensor automatically calibrated flow rate to "zero". Five inspiratory and expiratory calibration strokes were performed with the syringe, at a flow rate of between 3.0 to 6.0 $l \cdot s^{-1}$. A correction factor was automatically applied to ensure that the mass flow sensor read within ± 3% of the known syringe volume. Subsequently, five additional inspiratory and expiratory verification strokes were performed. To meet the American Thoracic Society's flow rate calibration criteria (Crapo *et al.*, 1994), at least one of the expiratory flow rates were required to be <0.5 and >3.0 $l \cdot s^{-1}$. If one or more strokes were not accurate to within ± 3% of the target value, the flow rate accuracy standard was deemed not to have been met. Thus, the calibration procedure was repeated.

3.7.3 The determination of F_EO_2 *and* F_ECO_2

A paramagnetic analyser (a diagmagnetic glass dumbbell suspended in a magnetic field, rotating in proportion to the surrounding partial pressure of oxygen), and a non-dispersive infrared analyser (a beam of infrared energy through a gas sample, measured the amount of infrared energy absorbed) measured expired F_EO_2 and F_ECO_2, respectively. These analysers had a resolution of ± 0.015% and ± 0.0023%. Samples of O_2 and CO_2 were drawn through a length of Perma Pure® Nafion® tubing that equilibrated the sample to ambient conditions. Thus, expirate was dehumidified, and calibration (dry) gas was humidified to ambient conditions. The system operated on the 'mixing chamber mode for exercise' setting, and $\dot{V}O_2$, and $\dot{V}CO_2$ was calculated at 20 s intervals. To replicate the sampling period used during *load-carriage test 1* (section 3.5.2.1) the 20 s intervals were averaged over 2-minute periods.

3.7.4 The calibration of gas concentrations (F_EO_2 and F_ECO_2)

A two step gas calibration was conducted prior to each test. A gas fraction sample line was detached from the mass flow sensor and placed in the front of the analyser module, and O_2 and CO_2 analysers underwent calibration simultaneously. Step 1 involved the calculation of correction factors on three different gas concentrations (16% O_2, 4% CO_2; 26% O_2, 0% CO_2, and 20.93%, O_2, 0.03% CO_2 (room air)). Step 2 involved the verification of these gas concentrations to within ± 5% of absolute operating range. Only when flow rate and gas concentration calibrations were complete did data collection begin.

3.8 Blood lactate measurement

3.8.1 Laboratory measure of blood lactate concentration

Initially, (Chapter 5) blood lactate concentration was measured using a bench top analyser (YSI 2300 Stat Plus, Yellow Springs Instruments, Ohio, USA). To prepare the measurement site an isopropyl alcohol swab (Medlock Medical, Oldham, UK) wiped the area. The first blood droplet was wiped away to avoid contamination from perspiration. The sample was drawn into a 44.7 μl heparinised capillary tube (1006637, Vitrex, Herlev, Denmark), and decanted into a 0.0004 l vessel (Greiner Bio-One, Germany). A 25 μL sample of blood was analysed for whole blood plasma lactate concentration. Prior to measurements the instrument was calibrated (0.45gl (~5 $mmol\cdot l^{-1}$)) (YSI, 2747, standard), and stabilised (YSI, 2357, buffer) with known solutions. On a weekly basis linearity checks were performed with a range of standards (2.5, 5, 7.5, 10, 15, 25, and 30 $mmol\cdot l^{-1}$).

3.8.2 Portable measure of blood lactate concentration

Blood lactate measurements were lysed and measured for whole blood lactate concentration in the field (Chapters 7, 8, and 9). The repeatability of which was interrogated in the laboratory (Appendix A.3). To perform this task a portable analyser, measuring whole blood lactate concentration was used (Lactate Pro ™, Arkray, Shiga, Japan) (Pyne *et al.*, 2000). This has been demonstrated to have good repeatability, and agreement against laboratory based systems (McNaughton *et al.*, 2002). The measurement site was prepared (section 3.8.1), and a 5 μl sample was automatically drawn into an enzymatic reagent test strip that was inserted into the analyser. The sample was analysed over a 60 s period using an amperometric technique. Here the lactate sample reacted with potassium ferrocyanide and lactate oxidase, to form potassium ferrocyanide and pyruvate. The ferrocyanide was oxidised when exposed to a voltage, which released electrons that created a current. The current underwent amperometric measurement that was proportional to the concentration of the blood lactate sample (Pyne *et al.*, 2000). To avoid the possibility of continued glycolysis from the erythrocytes (thus a time dependent increase in lactate concentration), all drawn samples were assayed immediately. Each time the analyser was used a 'check strip', with a reference blood lactate concentration confirmed the analyser was operating correctly. In addition a calibration strip ensured that each batch of test strips was matched to the analyser.

3.9 Personal Load Carriage Equipment (PLCE)

Current British Army PLCE consists of a bergen (backpack), hip webbing, and chest webbing / close operations vest (Leamon and Scott, 2005). In Chapters 4 and 5 British Army issue bergen, and hip webbing comprised the PLCE ensemble. The webbing load totalled 5.7 kg, and consisted of one yoke, one hip belt, and four pouches (Figure 3.14). The pouches contained: one full water bottle, one set of mess tins, and four small containers filled with sand (simulating ammunition). A short back bergen (with side pouches) sat on top of the webbing. Within-subjects, all straps (two × shoulder for webbing and bergen, and one × waist for webbing) were measured and replicated in subsequent visits. The bergen contents consisted of one helmet and body armour, and the remaining space was tightly packed with clothing to prevent movement of the load's centre of mass. The bergen weighed a total of 9.3 kg, thus the combined weight of the webbing and bergen was 15.0 kg (equivalent to the lowest CEG load carried in the British Army) (Figure 3.15). Disc weights (each weighing 2.5 kg) were added to the most medially-anterior part of the bergen's side pouches to increase the load to 20 kg and 25 kg (the other British Army CEG loads).

Changes were made to the PLCE in Chapters 6, 7, 8 and, 9. The webbing was removed, such that the load ensemble was configured of the bergen alone. This meant fewer straps required adjustment, saving time during intense periods of data collection in the field. The load was configured of clothing, tightly packed in the bottom half of the bergen, with a bag of gravel placed in the upper part as close to the body as possible and held in position with additional clothing. It was believed the simplicity of this load configuration would minimise the movement of the load's centre of mass over the 2-year data collection (Figure 3.16).

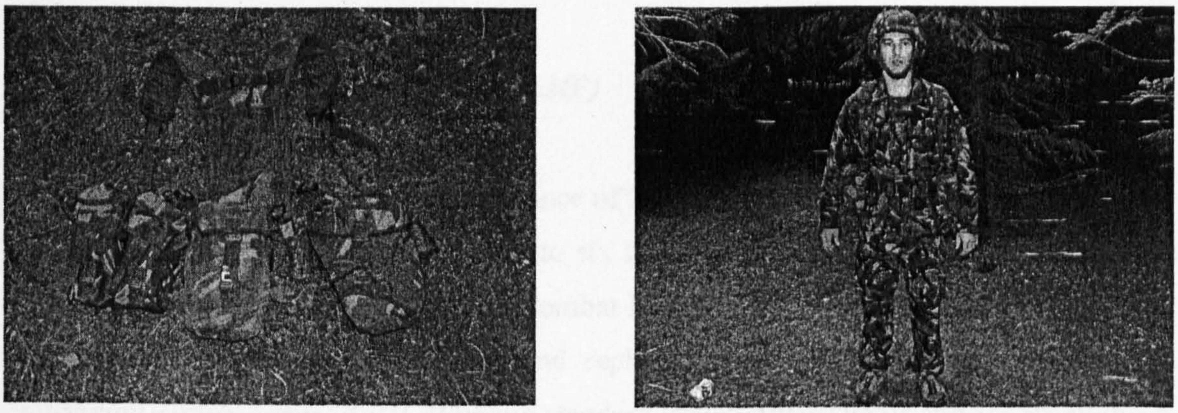


Figure 3.14. The configuration of webbing, the first layer of the PLCE ensemble.

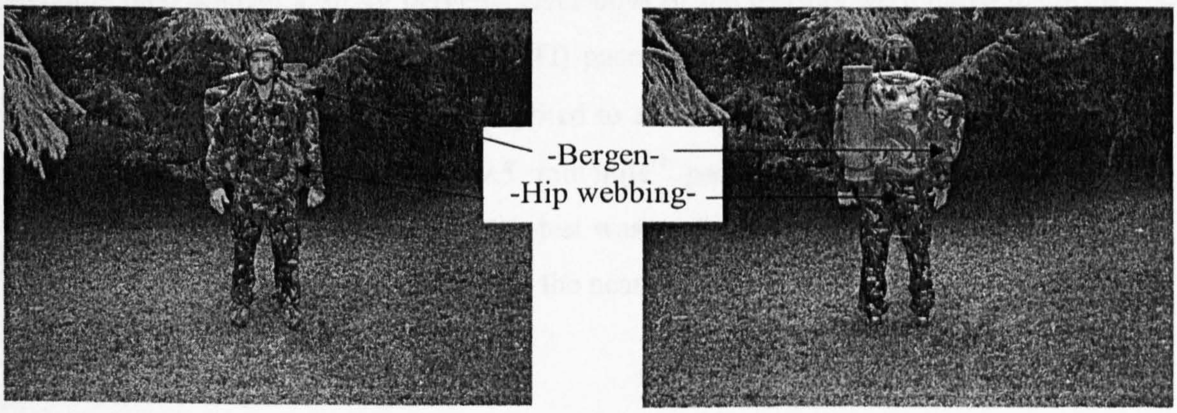


Figure 3.15. The webbing and bergen PLCE ensemble.

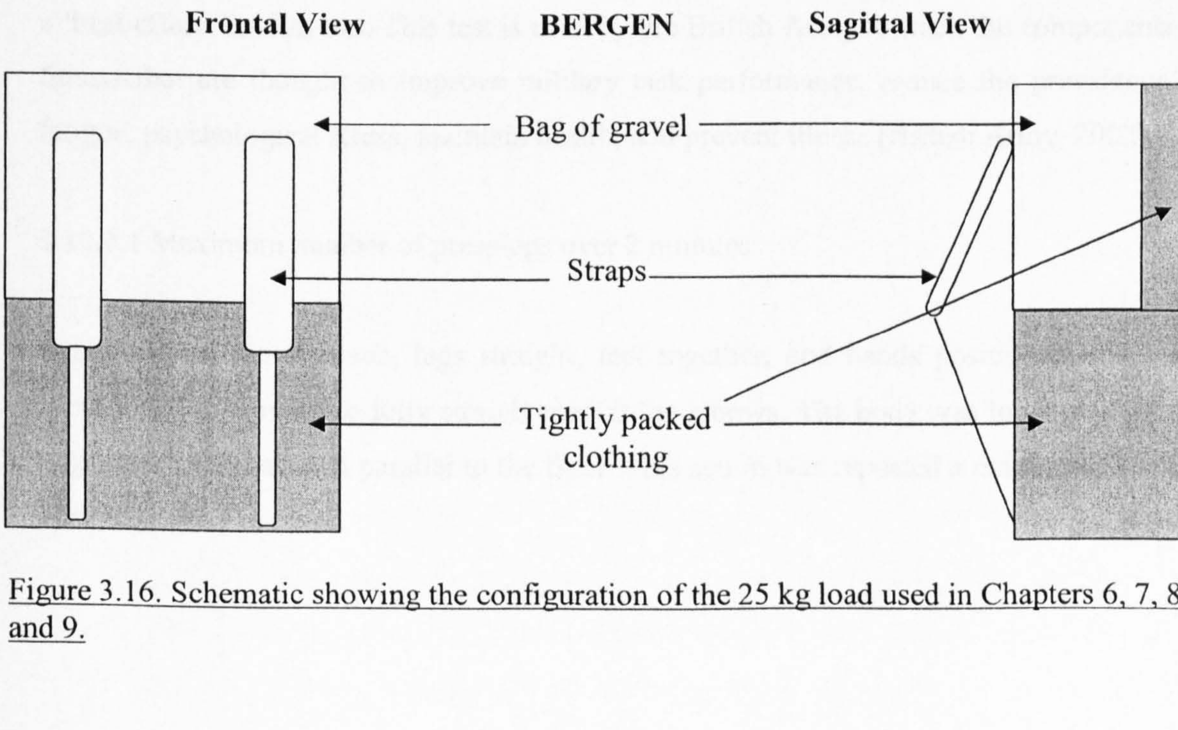


Figure 3.16. Schematic showing the configuration of the 25 kg load used in Chapters 6, 7, 8 and 9.

3.10 Military performance tests

3.10.1 Loaded Marching Performance (LMP)

The LMP test was performed over a distance of 2.4 km at the ITC(C), on a standard 400 m all weather athletics track. This equated to six full laps of the track. This speed march is commonly referred to as the Advanced Combat Fitness Test 1 (ACFT1) and is used by the British Army to test aerobic power, and replicate the operational demands of urban patrolling (British Army, 2002). Wearing standard issue military boots and CS 95 clothing, recruits completed an 800 m running warm up as a squad, at a speed of ~10.0 $km \cdot h^{-1}$. Recruits then donned a 20 kg bergen. Over 60% of the recruits were naïve load-carriers, hence a Physical Training Instructor (PTI) paced the first half (1.2 km) of the test at a speed of 9.6 $km \cdot h^{-1}$ (this is the pace required to successfully pass the ACFT1). The actual pace achieved was 10.2 $km \cdot h^{-1}$ (9.5 $min \cdot mile^{-1}$ pace), which was replicated in all subsequent tests. The second half of the test was performed as an individual 'best effort'. The time for each recruit was recorded to the nearest second.

3.10.2 Personal Fitness Test (PFT)

The PFT comprised of three discrete tests that included: the maximum number of press-ups performed over 2 minutes; the maximum number of sit-ups performed over 2 minutes, and a 'best effort' 2.4 km run. This test is used by the British Army to track the components of fitness that are thought to improve military task performance, reduce the prevalence of fatigue, psychological stress, maintain health, and prevent illness (British Army, 2002).

3.10.2.1 Maximum number of press-ups over 2 minutes

Lying flat on the stomach, legs straight, feet together, and hands positioned under the shoulders, the arms were fully straightened at the elbows. The body was lowered until the upper arms were at least parallel to the floor. This action was repeated a maximum number of times over 2 minutes.

3.10.2.2Maximum number of sit-ups over 2 minutes

With the arms folded across the chest, fingers placed in the clavicle depression, knees bent (70 to 110°), and a partner securing the feet to the floor, the torso was curled-up to the vertical position. The torso was then lowered until the shoulder blades touched the floor, and the action was repeated a maximum number of times over 2 minutes.

3.10.2.3 The 2.4 km run

On the ITC(C) all weather athletics track, and wearing PT kit (shorts, T-shirt, and trainers) an 800 m squadded warm-up was conducted, at an average speed of 9.6 $km \cdot h^{-1}$. This was followed by a 'best effort' 2.4 km run, which equated to six full laps of the track. Time was recorded to the nearest second.

3.10.3 Basic Combat Fitness Test (BCFT)

The BCFT is a 12.8 km physical output Representative Military Task. This loaded marching test is conducted towards the end of infantry training. Its aim is to test loaded marching ability to ensure recruits satisfy a minimum level of loaded marching fitness prior to embarking on a career in the field Army. Wearing standard issue boots, CS 95 jacket and trousers, and carrying a 25 kg load (inclusive of bergen, webbing and weapon), the first 6.4 km of the march were completed as a squad, at an average speed of 6.4 $km \cdot h^{-1}$. To test loaded march performance the second 6.4 km of the test was completed as an individual 'best effort'. This is a test model that was used to develop PSS(R) (Rayson *et al.*, 2002a).

3.11 Statistical analysis

3.11.1 Choosing an appropriate statistical technique to quantify the validity and repeatability of measurement instrumentation.

Several different methods have been proposed to evaluate repeatability (Atkinson and Nevill, 1998). It is important that these tests consider systematic bias (mean differences), random error (within-subject variation), and re-test correlation (the maintenance of a rank order on a re-test). Of these parameters, the minimisation of random error is the most important, due to its role in sample size estimation (Hopkins *et al.,* 1999).

When used in isolation, certain statistical techniques are considered inappropriate for method comparison (validity), or repeatability (reliability) analyses. Paired t-tests and Analysis of Variation (ANOVA) are such examples that only detect systematic bias, and do not quantify random error (Atkinson and Nevill, 1998). Furthermore, the absence of systematic bias is often erroneously interpreted as verification that repeatability has been satisfied, which is likely an artefact of a large random error (Altman and Bland, 1983). Correlation, intra class correlation, and regression are also criticised for being sensitive to the heterogeneity of a data set, and in the case of correlation for being insensitive to systematic bias (Atkinson and Nevill, 1998).

Absolute measures of repeatability are more appropriate statistical tools for method comparison and repeatability (Atkinson and Nevill, 1998). Two statistical approaches that measure absolute reliability are 95% Limits of Agreement (LoA) (Altman and Bland, 1983; Bland and Altman, 1986), and Typical Error (TE) (Hopkins, 2000a). Whilst these both quantify random error, only LoA quantifies systematic bias. Previously, the validity of each method was debated (Atkinson and Nevill, 2000; Hopkins, 2000b). Typical Error, and its derived coefficient of variation (CV) only accounts for 52% of test-retest differences. Consequently, it lacks the prudence that is provided with 95% LoA (Woolford and Gore, 2004). For example, from an identical data set systematically smaller TE values (by a factor of 2.77) are observed compared to LoA. Thus, levels of apparent repeatability differ almost 3-fold between these tests. However, LoA are in fact more appropriate, as these represent the majority (95%) of the observed population, not just 52%.

3.11.2 The application of LoA for method comparisons and repeatability

The repeatability (or reliability) of the same measure, and the method comparison (or validity) of two different methods are concepts fundamental to the early part of this thesis (Chapters 4 and 6). Limits of Agreement (Bland and Altman, 1983; 1986, 1995; 1999; 2003) determined the repeatability of all measures (Chapter 6; Appendix A), and the agreement between Douglas bag and On-line gas analysis systems (Chapter 4). Bland and

Altman plots graphically expressed between-test differences on the y-axis, and the mean of the two tests on the x-axis, for each individual subject (e.g. Figure 4.2).

The assumptions of Normality and Homoscedasticity were checked. Normality of distribution was checked on the between test differences (i.e. observation 2 – observation 1), and was deemed present if the Skewness ÷ SE Skewness, or Kurtosis ÷ SE Kurtosis was ≥ than 2 (Fallowfield *et al.,* 2005). Heteroscedasticity (commonly referred to as measurement error proportional to the mean) was also checked using between test differences (i.e. observation 2 – observation 1). Here the *absolute* between test differences (i.e. all differences expressed in a positive direction) were plotted (y-axis) against the pooled mean of observation 1 and 2 (x-axis) (e.g. Figure 6.1). A positive Pearson's correlation coefficient indicated the presence of heteroscedasticity (Nevill and Atkinson, 1997). Violations to Normality and homoscedasticity were criteria for data to be logarithmic transformed, and expressed as Ratio LoA (Nevill and Atkinson, 1997a; 1997b).

Two descriptive statistics were generated by the LoA analysis that quantified the level of repeatability or agreement. These were, *systematic bias* (i.e. test 2 mean score – test 1 mean score), and *random error* (i.e. standard deviation of the between test differences (SD_{diff})). Systematic bias highlighted the direction of the mean differences around the zero line (Figure 4.2). A *t*-test objectively quantified the presence of a systematic bias ($P<0.05$). Random error was quantified by multiplying the SD_{diff} by the z-score of 1.96, to provide a 95% LoA. *Absolute LoA* were expressed in the units of measurement used in the test, as **systematic bias (± 95% LoA)** (Atkinson and Nevill, 1998).

To indicate the precision of the systematic bias (Equation 3.4 and 3.5) and 95% LoA (Equation 3.6 and 3.7), 95% confidence intervals (CI) were derived from the calculation of standard error (SE). This indicated the application of the LoA to the wider population (Bland and Altman, 1986; 2003). A *t* distribution table (Thomas and Nelson, 1996) selected the appropriate *t* statistic for calculation of CI. Here, a two-tailed level of significance was selected at 0.05, with n – 1 degrees of freedom.

$$\textit{Systematic Bias SE} = \textit{squrt}\ (SD^2/n)$$

(Equation 3.4)

$$\textit{Systematic Bias CI} = \textit{systematic bias} \pm (\textit{t statistic}\cdot (\textit{squrt}\ (SD^2/n)))$$

(Equation 3.5)

$$\textit{Random Error SE} = \textit{squrt}\ (3 \cdot SD^2/n)$$

(Equation 3.6)

$$\textit{Random Error CI} = (\textit{systematic bias} \pm 95\%\ LOA)\ \pm (\textit{t statistic}\cdot (\textit{squrt}\ (3 \cdot SD^2/n)))$$

(Equation 3.7)

Due to the high level of non-Normally distributed data and heteroscedasticity, data were transformed into the logarithmic scale to produce *Ratio LoA* (Bland and Altman, 1996b). Here data were transformed with a natural logarithm (ln) that was subsequently antilogged, to return the value back to the natural scale (Bland and Altman, 1996a). As subtraction and addition in the natural scale (i.e. absolute LoA) are equivalent to multiplication and division logarithmic scale (i.e. Ratio LoA) the 95% ratio LoA were calculated as **systematic bias (x/÷ 95% ratio LoA)**. Ratio LoA data can also be expressed as a percentage by subtracting 1 from the 95% ratio LoA (e.g. 1.14 -1= 0.14 = 14%) (Bland and Altman, 1996e).

It has been suggested that a sample size of >40 is required to ensure 95% LoA can be applied to the wider population (Atkinson and Nevill, 1998). However, access to a subject data pool of this size can be problematic (Atkinson *et al.*, 2005), and is often not achieved (Nevill and Atkinson, 1997a).

3.11.3 Inferential statistical analyses in military physical training studies

Normality of distribution was satisfied in just 13 out of 59 measures (22%) across Chapters 8 and 9. Thus to ensure continuity of analysis, non-parametric tests (the merits of which are discussed below) were universally applied in these chapters.

Inferential statistical tests offer an objective tool by which data sets can be appraised, however the existence of a single correct way to analyse these data is a common misconception (Maxwell and Delaney, 1990). Typically, parametric tests are the correct statistical choice if Interval or Ratio scale data satisfy the assumptions of; normality, homogeneity of variance, and independence of observations (Thomas and Nelson, 1996). However, when the parametric assumptions are violated the 'correct' statistical decision is less clear. In such an instance the options include (1) running a parametric test despite the violation (Maxwell and Delaney, 1990; Tabachnick and Fidell, 1996), (2) transforming the data to produce normality and then performing a parametric test, (3) performing a non-parametric test (Roberts and Russo, 1999; Field, 2001; Pallant, 2007).

Performing a parametric test despite violations is possible as parametric tests are 'robust' to violations in the assumptions when, relatively equal sample sizes between-groups, no outliers, and over 20 degrees of freedom for error are present (Tabachnick and Fidell, 1996). However, these may not provide the most powerful test when the assumptions have been violated (Maxwell and Delaney, 1990). Whilst data transformations are a suggested solution to violations they are not universally recommended (Tabachnik and Fidell, 1996). Transformations alter the scale of measurement into arbitrary units. Although this may be a valid decision statistically it threatens data interpretation, which has scientific implications (Maxwell and Delaney, 1990; Grayson, 2004). The final option is to perform assumption free, non-parametric tests. These tests are generally considered less powerful than parametric tests, thus are more conservative hence more likely to produce a Type II error (false-negative) (Thomas and Nelson, 1996). Philosophically, the use of non-parametric statistics within this thesis may be prudent as these findings will influence physical training policy in recruit training across the whole British Army. Thus it may be wise that data analysis and interpretation be conservative, as were significant differences to be found, one can be more confident that this is a true effect and not a Type I error (false-positive). Undoubtedly, this may inflate the risk of a Type II error (false-negative) however; this approach should ensure inappropriate recommendations are not made to the end-user. In any case the notion that parametric tests must be used due to the greater statistical power is an oversimplification, as non-parametric tests can be more powerful than parametric alternatives, particularly when normality is violated (Maxwell and Delaney, 1990).

Effect size (r) represents the strength of an effect in relation to the variability in the data (Winter *et al.*, 2001) and should be reported with all significant tests (Thomas and Nelson, 1996). As such effect size (r) (Equation 3.8) is reported for all inferential statistical tests in Chapters 8 and 9 (Field, 2005). The meaningfulness of an effect can be classified as: a small effect (r = 0.10), medium effect (r = 0.30), and large effect (r = 0.50) (Cohen, 1992).

Effect size (r) = (z score/(squrt of total number of observations))

(Equation 3.8)

Due to the prevalence of non-Normally distributed data, central tendency in Chapters 8 and 9 was expressed as median, not mean (Tabachnik and Fidell, 1996). As such the range, and not the standard deviation reported the distribution of data around the median (Field, 2005).

CHAPTER 4

A VALIDATION OF THE ON-LINE SENSORMEDICS GAS ANALYSIS SYSTEM.

4.1 Introduction

The measurement of human expired gases (gas analysis) is the primary physiological measure in this thesis, and will monitor sub-maximal loaded marching economy (LME) (i.e. the rate of oxygen uptake at an absolute exercise intensity (Jones and Carter, 2000)) and running $\dot{V}O_2max$ responses of British Army recruits during infantry training (Chapter 7, 8, and 9). A reliable, precise and accurate gas analysis system is likely to be important for the assessment of LME, as exercise economy and $\dot{V}O_2max$ are known to respond modestly to training (Sjödin *et al.*, 1982; Conley *et al.*, 1984; Svedenhag and Sjödin, 1985; Berg *et al.*, 1995; Franch *et al.*, 1998; Billat *et al.* 1999; Jones *et al.*, 1999; Spurrs *et al.*, 2003). The only study to have tracked the training-related adaptation in LME reported a small improvement (~1 $ml \cdot kg^{-1} \cdot min^{-1}$) over 8 weeks of military style training (Gutekunst *et al.*, 2006). This was smaller than that considered acceptable day-to-day variation in human expired gases measurement (± 2 or 3 $ml \cdot kg^{-1} \cdot min^{-1}$) (Thoden, 1991), thereby highlighting the importance of accurate and reliable gas analysis systems, especially for small sample sizes.

The Douglas Bag (DB) approach to gas analysis refers to the collection of expired gases (Douglas, 1911) and subsequent analysis. This technique is based upon first principles of gas analysis, and as such is considered the *gold standard* approach. However, the established validity and repeatability of this approach is offset by the time-consuming nature of data processing, conversley data processing speed is not an issue for modern automated on-line systems (ON). Automated systems lack hardware and software transparency, meaning that the genesis of calculated $\dot{V}O_2$, and error propagation are hard to identify (Macfarlane, 2001; Hodges *et al.*, 2005).

It is prudent to objectively assess a measurement approach prior to its widespread use. This process involves a systematic validation that evaluates validity (e.g. method comparison – DB vs. ON system), and quantifies repeatability (e.g. test-retest differences of DB and ON systems) (Bland and Altman, 1986). Several gas analysis studies have conducted incomplete validations that completed method comparisons, but not repeatability assessments (La Mere *et al.*, 1993; Miles *et al.*, 1994; Unnithan *et al.*, 1994; Yule *et al.*, 1996; Engebretson, 1998; Hiilloskorpi *et al.*, 2000; Midownik *et al.*, 2000; Bassett *et al.*, 2001; Rietjens *et al.*, 2001; Gore *et al.*, 2003). Failing to quantify repeatability is a fundamental flaw, as a test measure cannot be valid if it is not repeatable (Thomas and Nelson, 1996).

A lack of statistical standardisation in the literature makes comparisons between validation studies problematic (Hodges *et al.*, 2005). Bland and Altman's 95% Limits of Agreement (LoA) (Altman and Bland, 1983; Bland and Altman, 1986) were specifically developed for the process of validation. Unlike other statistical approaches, one advantage of this method is that it quantifies both systematic and random errors (Atkinson and Nevill, 1998). Refer to section 3.11 for a more detailed discussion on LoA.

The purpose of this study was to conduct a validation of an ON gas analysis system against a *gold standard* DB system. The processes of method comparison (validity), and repeatability (reliability) will determine the most suitable gas analysis system for the description and detection of potentially small physiological responses to British Army infantry training.

4.2 Method

4.2.1 Subjects

Twelve male University students and staff volunteered to participate in the study. All were actively involved in sport and exercise, and had previous experience of physiological laboratory testing. The research was approved by the University of Chichester Ethics Committee.

Table 4.1. Physiological characteristics of the subjects (mean ± SD).

Age (years)	Stature (m)	Body Mass (kg)	$\dot{V}O_2max$ ($ml \cdot kg^{-1} \cdot min^{-1}$)
28.8 ± 7.1	1.81 ± 0.05	81.3 ± 6.6	51.2 ± 3.7

4.2.2 Experimental design

Subjects reported to the laboratory on four separate occasions, each visit was separated by at least 48 hours. Visits 1 and 2 consisted of loaded marching (LM) familiarisation. Visits 3 and 4 (observation 1 and 2 – Figure 4.1) were identical loaded march bouts designed to assess the validity and repeatability of the ON gas analysis system against the DB system. During each test session the order of DB and ON expired gas collections were alternated, this order remained identical within-subjects, but was counterbalanced between-subjects (Figure 4.1).

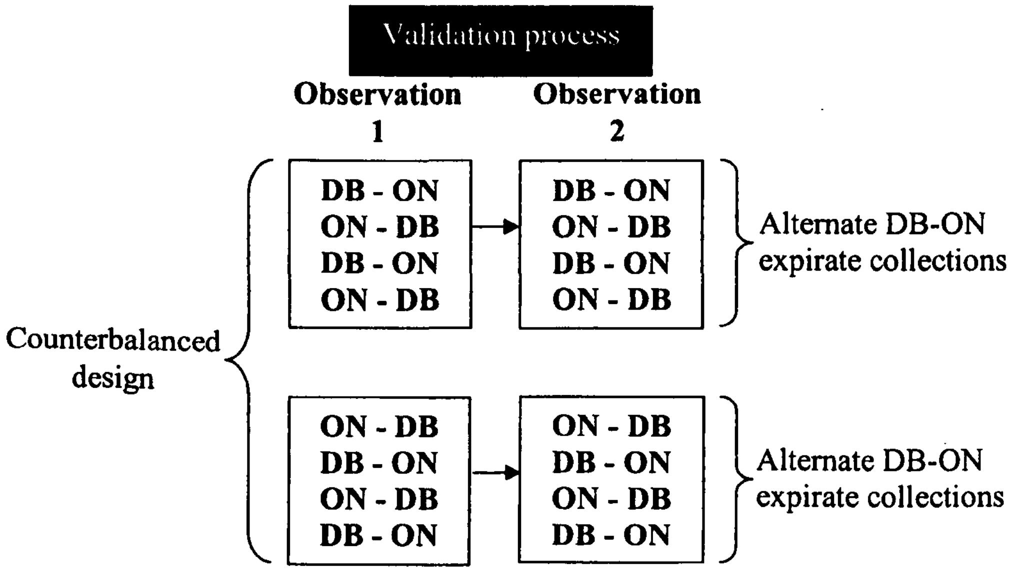


Figure 4.1. Schematic representation of the counterbalanced experimental design. DB = Douglas Bag, ON = On-line gas analyser.

4.2.3 Protocols

4.2.3.1 Loaded marching familiarisation

Subjects wore training shoes, British Army CS 95 lightweight jacket and trousers, and a PLCE ensemble that consisted of a bergen (rucksack) and hip webbing (section 3.9). Twenty nine minutes of marching at 6.5 $km \cdot h^{-1}$ with varying loads was completed. Each load (0, 15, 20 and 25 kg) was carried in the order of increasing mass for 5 minutes, with each bout of load-carriage separated by 3 minutes of unloaded walking.

4.2.3.2 The validation process (observation 1 and 2)

The *sub-maximal loaded march test 1* (section 3.5.2.1) was performed for the validation. A 5-minute unloaded walk at 6.5 $km \cdot h^{-1}$ constituted the warm-up, and preceded the loaded march test. The identical clothing and PLCE ensemble as *per* the familiarisation was worn, and subjects undertook a total of 48 minutes of ambulation on a motorised treadmill at 6.5 $km \cdot h^{-1}$. Four, 9-minute stages were performed with 4 different external loads (0, 15, 20 and 25 kg), in a randomised order. Four minutes of unloaded walking at 6.5 $km \cdot h^{-1}$ separated each stage. Within each 9-minute stage two-minute DB and ON collections were taken between minutes 4 – 6 and 7 – 9 (Figure 3.10), the order of which was alternated between stages (Figure 4.1).

4.2.4 Statistical analysis

Method comparison and repeatability were quantified with Bland and Altman's 95% LoA (Altman and Bland, 1983; Bland and Altman, 1986) (section 3.11.2). Ratio LoA were converted to a percentage by subtracting 1 from the 95% Ratio LoA (i.e. 1.14 – 1 = 0.14 = 14%) (Bland and Altman, 1996a). As previously suggested, multiple observations were made on each subject (one per LM bout i.e. 0, 15, 20, 25 kg) (Crouter *et al.*, 2006), as this increased the number of data points recommended for LoA analysis (Atkinson *et al.*, 2005). Due to the repeated observations a correction was applied to prevent the LoA from becoming falsely narrow (Bland and Altman, 2007). Outliers in the data were identified from DB and ON test differences. Cases outside of the normal range of boxplot scores were considered outliers (Field, 2001), but were still included in the LoA analysis (Bland and Altman, 1999).

4.3 Results

With the DB analyser LME was 16.8 ± 1.1 $ml \cdot kg^{-1} \cdot min^{-1}$ (range 14.6 $ml \cdot kg^{-1} \cdot min^{-1}$ to 17.7 $ml \cdot kg^{-1} \cdot min^{-1}$); 20.5 ± 1.6 $ml \cdot kg^{-1} \cdot min^{-1}$ (range 17.6 $ml \cdot kg^{-1} \cdot min^{-1}$ to 22.5 $ml \cdot kg^{-1} \cdot min^{-1}$); 21.6 ± 1.9 $ml \cdot kg^{-1} \cdot min^{-1}$ (range 18.1 $ml \cdot kg^{-1} \cdot min^{-1}$ to 24.5 $ml \cdot kg^{-1} \cdot min^{-1}$); and 23.2 ± 2.2 $ml \cdot kg^{-1} \cdot min^{-1}$ (range 19.7 $ml \cdot kg^{-1} \cdot min^{-1}$ to 26.5 $ml \cdot kg^{-1} \cdot min^{-1}$) during the 0, 15, 20 and 25 kg LM stages, respectively. The variability of each of the four load-carriage stages (0, 15, 20 and 25 kg) was similar (~15%) so data were pooled as *per* Crouter *et al.,* (2006). The pooled mean LME was 20.5 ± 3.0 $ml \cdot kg^{-1} \cdot min^{-1}$ for DB, and 23.8 ± 3.4 $ml \cdot kg^{-1} \cdot min^{-1}$ for ON gas collections.

4.3.1 Method comparison (validity)

A systematic bias existed for the ON system in 5 out of 6 gas analysis measures. The ON relative $\dot{V}O_2$, absolute $\dot{V}O_2$, $\dot{V}_E$ and $\dot{V}CO_2$ were overestimated by 12% to 16%, and F_EO_2 underestimated by 1% in comparison to the DB system ($P<0.05$). F_ECO_2 was the only parameter not to exhibit a systematic bias ($P>0.05$) (Table 4.2).

Two separate method comparisons were made (observations 1 and 2) (Figure 4.1), which yielded inconsistent findings. For example $\dot{V}O_2$ ratio LoA (i.e. systematic bias (×/÷ 95% LoA)) of ~38 % (or 1.16 (×/÷ 1.19)) in observation 1 (Figure 4.2), were markedly different to the ~24% (or 1.13 (×/÷1.10)) reported in observation 2 (Figure 4.3). When translated into absolute units of measurement, 95% of the ratios for ON vs. DB $\dot{V}O_2$ (in observation 1) should lie between the systematic bias (1.16), multiplied or divided by the random error ratio (×/÷ 1.19). Thus the LoA are (1.16 × 1.19 = **1.38**, to 1.16 ÷ 1.19 = **0.97**). Consequently, ON $\dot{V}O_2$ varied by +0.6 to +7.8 $ml \cdot kg^{-1} \cdot min^{-1}$ around the mean DB $\dot{V}O_2$ of 20.5 $ml \cdot kg^{-1} \cdot min^{-1}$ (for an example see Figure 4.2). Ratio LoA were up to ~32% (or 1.16 (×/÷ 1.14)) for minute ventilation ($\dot{V}_E$), ~38% (or 1.15 (×/÷ 1.20)) for $\dot{V}CO_2$, ~13 % (or 1.01 (×/÷ 1.12)) for F_ECO_2, and ~5 % (or 0.99 (×/÷ 1.04)) for F_EO_2 (Table 4.2).

Table 4.2. The method comparison (validity) between Douglas Bag (DB) and On-line (ON) gas analysis systems during simulated loaded marching, over two separate observations. Data are expressed as (Systematic bias (95% LoA)) in ratio and absolute terms.

Measure	**Mean ± SD**		**Ratio LoA**	**Absolute LoA**
	DB	**ON**	**Bias (×/÷ 95% LoA)**	**Bias (± 95% LoA)**
Observation 1				
$\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	20.5 ± 3.0	23.8 ± 3.4	1.16** (×/÷ 1.19)	3.3** (± 3.9)
$\dot{V}O_2$ (l·min^{-1})	1.67 ± 0.27	1.93 ± 0.26	1.16** (×/÷ 1.19)	0.26** (± 0.28)
$\dot{V}_E$ (l·min^{-1})	36.4 ± 6.4	42.4 ± 7.7	1.16** (×/÷ 1.14)	6.0** (± 5.8)
$\dot{V}CO_2$ (l·min^{-1})	1.52 ± 0.26	1.74 ± 0.27	1.15** (×/÷ 1.20)	0.22** (± 0.28)
F_EO_2 (%)	16.40 ± 0.55	16.20 ± 0.52	0.99** (×/÷ 1.04)	-0.20** (± 0.69)
F_ECO_2 (%)	4.22 ± 0.43	4.27 ± 41	1.01 (×/÷ 1.12)	0.05 (± 0.48)
Observation 2				
$\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	20.5 ± 2.9	23.2 ± 3.7	1.13** (×/÷ 1.09)	2.8** (± 2.4)
$\dot{V}O_2$ (l·min^{-1})	1.66 ± 0.23	1.89 ± 0.28	1.13** (×/÷ 1.10)	0.22** (± 0.19)
$\dot{V}_E$ (l·min^{-1})	35.8 ± 6.3	41.3 ± 6.9	1.15** (×/÷ 1.14)	5.5** (± 5.4)
$\dot{V}CO_2$ (l·min^{-1})	1.51 ± 0.24	1.69 ± 0.28	1.12** (×/÷ 1.10)	0.18** (± 0.17)
F_EO_2 (%)	16.34 ± 0.48	16.22 ± 0.45	0.99* (×/÷ 1.04)	-0.12** (± 0.59)
F_ECO_2 (%)	4.26 ± 0.43	4.24 ± 0.42	0.99 (×/÷ 1.08)	-0.03 (± 0.34)

* statistically significant systematic bias (P<0.05), ** (P<0.01).

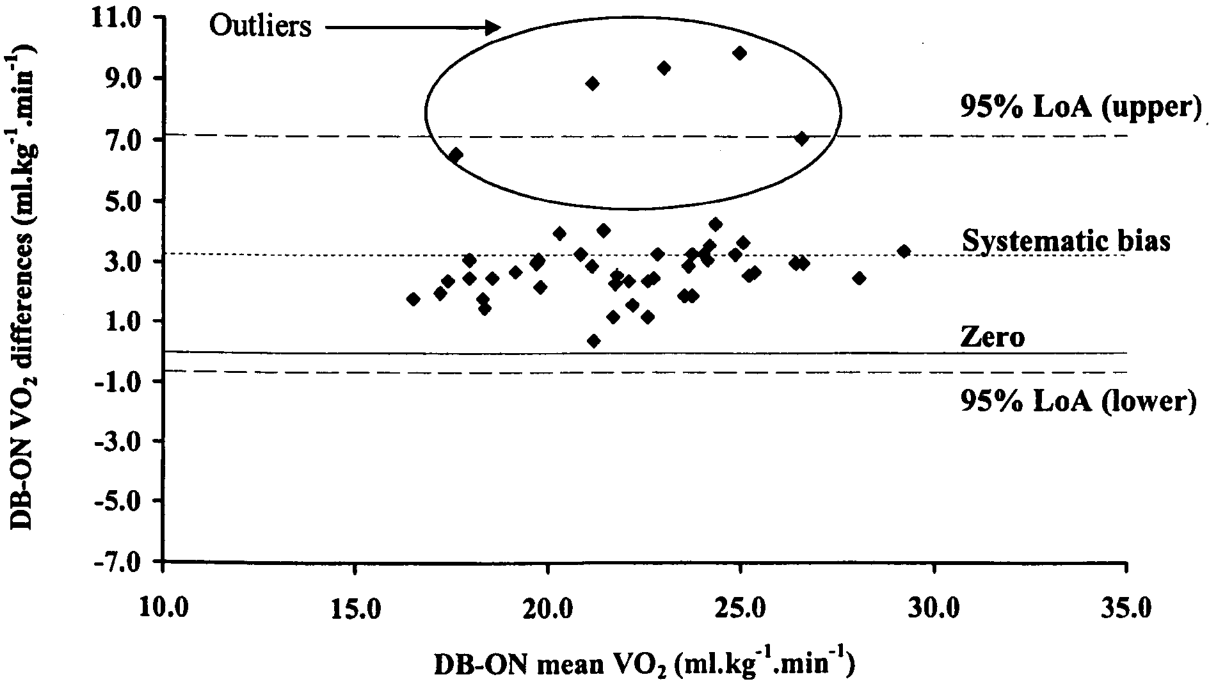


Figure 4.2. Bland and Altman method comparison plot showing systematic bias, absolute 95% LoA, and outliers for DB vs. ON $\dot{V}O_2$ during observation 1.

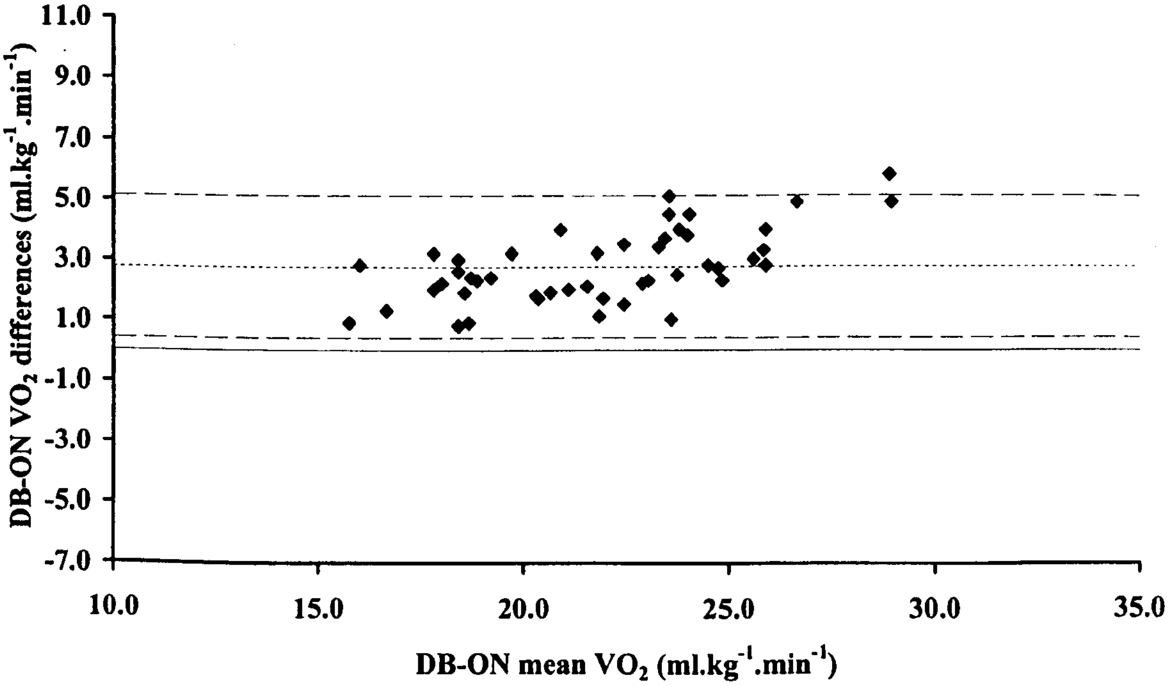


Figure 4.3. Bland and Altman method comparison plot showing systematic bias, and absolute 95% LoA for DB vs. ON $\dot{V}O_2$ during observation 2.

4.3.2 Repeatability

None of the DB test re-test measures were systematically biased ($P>0.05$), however four ON measures (relative $\dot{V}O_2$, absolute $\dot{V}O_2$, $\dot{V}_E$, and $\dot{V}CO_2$) had systematic biases of ~3% ($P<0.05$) (Table 4.3).

Ratio LoA were ~16% (or 0.97 (×/÷ 1.15)) for ON $\dot{V}O_2$, which was comparable to the DB $\dot{V}O_2$ of ~15% (or 1.00 (×/÷ 1.15)) (Figures 4.4 and 4.5). However, removal of four DB outliers (Figure 4.4, section 4.2.4) almost halved the DB ratio LoA (from ~15% to ~9%). Removal of the one outlying ON data point (Figure 4.5) did not change ON ratio LoA (Table 4.3). Ratio LoA were ~20% (or 0.97 (×/÷ 1.21)) for ON $\dot{V}_E$, which was greater than the DB $\dot{V}_E$ of ~16% (or 0.98 (×/÷ 1.14)). LoA were ~18% (or 0.97 (×/÷ 1.19)) for ON $\dot{V}CO_2$, which was greater than the DB $\dot{V}CO_2$ of ~12% LoA (or 0.99 (×/÷ 1.12)). F_EO_2 was comparable for both ON and DB systems (~4% to 5%). The ON F_ECO_2 LoA of ~9% (or 0.99 (×/÷ 1.08), was half that of the DB F_ECO_2 (1.01 (×/÷ 1.14)) (Table 4.3).

Table 4.3. The repeatability (reliability) of Douglas Bag (DB) and On-line (ON) gas analysis systems during simulated loaded marching. Data are expressed as (Systematic bias (95% LoA)) in ratio and absolute terms.

Measure	**Mean ± SD**		**Ratio LoA**	**Absolute LoA**
	Test 1	**Test 2**	**Bias (×/÷ 95% LoA)**	**Bias (± 95% LoA)**
DB				
$\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	20.5 ± 3.0	20.5 ± 2.9	1.00 (×/÷ 1.15)	-0.1 (± 2.7)
$\dot{V}O_2$ (l·min^{-1})	1.67 ± 0.27	1.66 ± 0.23	1.00 (×/÷ 1.15)	-0.01 (± 0.21)
$\dot{V}_E$ (l·min^{-1})	36.4 ± 6.4	35.8 ± 6.3	0.98 (×/÷ 1.14)	-0.6 (± 4.7)
$\dot{V}CO_2$ (l·min^{-1})	1.52 ± 0.26	1.51 ± 0.24	0.99 (×/÷ 1.13)	-0.01 (± 0.18)
F_EO_2 (%)	16.40 ± 0.55	16.34 ± 0.48	1.00 (×/÷1.05)	-0.06 (± 0.82)
F_ECO_2 (%)	4.22 ± 0.43	4.26 ± 0.43	1.01 (×/÷ 1.14)	0.04 (± 0.56)
ON				
$\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	23.9 ± 3.5	23.3 ± 3.6	0.97* (×/÷ 1.15)	-0.6* (± 3.5)
$\dot{V}O_2$ (l·min^{-1})	1.94 ± 0.27	1.89 ± 0.28	0.97* (×/÷ 1.15)	-0.05* (± 0.29)
$\dot{V}_E$ (l·min^{-1})	42.7 ± 8.0	41.5 ± 7.0	0.97 (×/÷ 1.21)	-1.3* (± 7.7)
$\dot{V}CO_2$ (l·min^{-1})	1.75 ± 0.28	1.70 ± 0.28	0.97* (×/÷ 1.19)	-0.05* (± 0.31)
F_EO_2 (%)	16.21 ± 0.52	16.23 ± 0.44	1.00 (×/÷ 1.04)	0.02 (± 0.65)
F_ECO_2 (%)	4.27 ± 0.41	4.23 ± 0.41	0.99 (×/÷ 1.08)	-0.03 (± 0.34)

* statistically significant systematic bias ($P<0.05$).

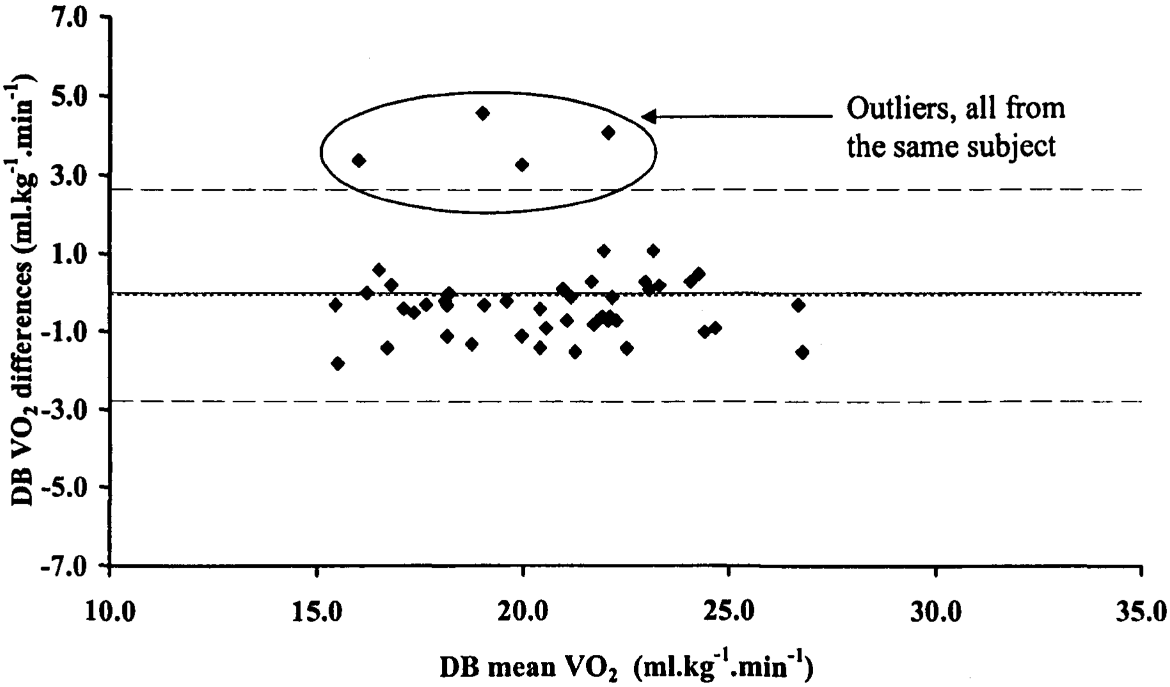


Figure 4.4. Bland and Altman repeatability plot showing systematic bias, and absolute 95% LoA of relative DB $\dot{V}O_2$. Four outlying data points (all from the same subject) circled.

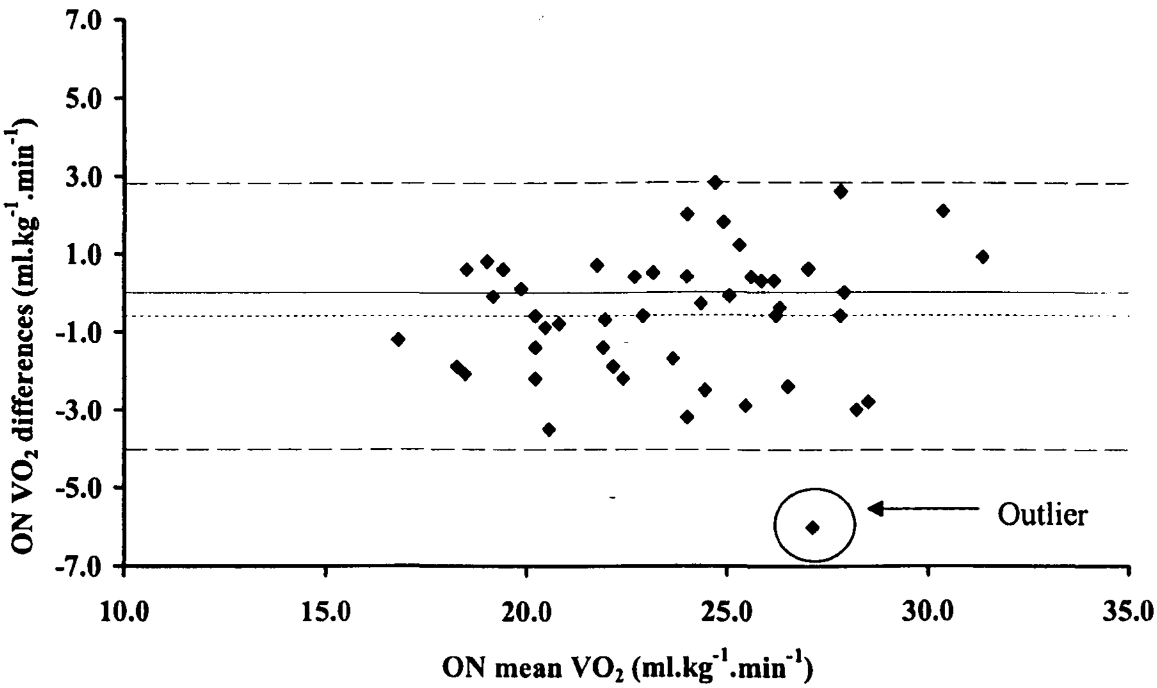


Figure 4.5. Bland and Altman repeatability plot showing systematic bias, and absolute 95% LoA of relative ON $\dot{V}O_2$. One outlying data point circled.

4.4 Discussion

Two main findings emerged from the present study. First, the ON gas analysis system systematically overestimated $\dot{V}O_2$ compared to the *gold standard* DB system, by 13 to 16%. Second, whilst the 95% LoA statistic found the repeatability of DB and ON analysers to be comparable, inspection of the Bland and Altman plots indicated that the ON had poorer repeatability. This was confirmed when outlying DB and ON data points were removed, showing ON repeatability to be almost two-fold poorer than the DB system.

4.4.1 Method comparison

The ON system systematically overestimated $\dot{V}O_2$ by 13 to 16% compared to the DB. This equated to a systematic bias of around 3.3 $ml \cdot kg^{-1} \cdot min^{-1}$, or 0.26 $l \cdot min^{-1}$. Others have reported the ON SensorMedics to systematically overestimate $\dot{V}O_2$ by almost double that of the present study (5.4 $ml \cdot kg^{-1} \cdot min^{-1}$) (Babineau *et al.*, 1999). As measurement error is typically proportional to the magnitude of the mean value (Bland and Altman, 1996c), the smaller systematic bias in the present study may be explained by the sub-maximal exercise model adopted (i.e. loaded marching), compared to the maximal model undertaken by Babineau and colleagues. Other automated systems have demonstrated a systematic bias of similar magnitude (0.27 $l \cdot min^{-1}$) and in the same direction (La Mere *et al.*, 1993), and a similar magnitude (11%) but in the opposite direction (Gore *et al.*, 2003) to the present study. In the latter study it was concluded that systematic bias of 11% questions the accuracy (thus validity) of an analyser.

In the present study the $\dot{V}O_2$ random error in the 95% LoA (i.e. the 95% LoA statistic without the systematic bias) was worse (Bassett *et al.*, 2001), similar (Rietjens *et al.*, 2001; Crouter *et al.*, 2006) and better (Crouter *et al.*, 2006) than previously reported. However, when the marked systematic bias in $\dot{V}O_2$ (Table 4.2) was combined with the random error component to produce the 95% LoA statistic (i.e. systematic bias (×/÷ random error)) the LoA become excessive, such that these became markedly worse (Bassett *et al.*, 2001), worse (Rietjens *et al.*, 2001; Crouter *et al.*, 2006), and comparable (Crouter *et al.*, 2006) to other studies. The method comparison in the present study demonstrated that the ON system lacked validity, which is a consequence of the large systematic bias in the calculation of $\dot{V}O_2$.

It is inaccurate to refer to the measurement error of $\dot{V}O_2$ (and $\dot{V}CO_2$), as $\dot{V}O_2$ is not measured but calculated from the three primary variables ($\dot{V}_E$, F_EO_2 and F_ECO_2), and three secondary variables (barometric pressure, gas temperature, and water vapour pressure of gas) (Howley *et al.*, 1995). Therefore the error propagation in the calculation of $\dot{V}O_2$ primarily resides in the measurement of $\dot{V}_E$, F_EO_2 and F_ECO_2. The genesis of the systematic bias in ON $\dot{V}O_2$ seems to be a consequence of a systematic overestimation in minute ventilation ($\dot{V}_E$), and not F_EO_2 or F_ECO_2 (Table 4.2). This was verified by the observation that the 16% (or 1.16) systematic bias in $\dot{V}O_2$ was proportional to that of $\dot{V}_E$, which has been corroborated elsewhere (Withers *et al.*, 2000).

The potential source of discrepant $\dot{V}O_2$ values during method comparison studies has received some attention. In agreement with the present study, one study cited $\dot{V}_E$ as the source for the overestimation in $\dot{V}O_2$, observing an overestimation in $\dot{V}O_2$ during maximal exercise of 0.27 $l \cdot min^{-1}$, which coincided with a proportional overestimation in $\dot{V}_E$ of 3.1 $l \cdot min^{-1}$ (La Mere *et al.*, 1993). Other studies did not observe systematic differences in $\dot{V}O_2$ between automated and *gold standard* systems, but did observe differences in the primary variables of $\dot{V}_E$, F_EO_2 or F_ECO_2 (La Mere *et al.*, 1993; Unnithan *et al.*, 1994; Engebretson *et al.*, 1998; Jensen *et al.*, 2002; Carter and Jeukendrup, 2002). The mechanism for the larger error in the calculation of $\dot{V}O_2$ in ON systems has been suggested to be due to the inaccuracy of automated flow measuring devices that measure $\dot{V}_E$ (Midowink *et al.*, 2000). Reviews of the literature support this notion, stating that automated ON systems typically over or underestimate $\dot{V}_E$ by up to 10.5% compared to reference systems (Hodges *et al.*, 2005).

Minute ventilation ($\dot{V}_E$) is the likely source of the overestimation of ON $\dot{V}O_2$ in the present study. The cause of this problem may originate from two potential sources. These being, (1) the different approaches used to quantify minute ventilation, and (2) the potentially different equations used to calculate $\dot{V}O_2$ between DB and ON systems. First, DB minute ventilation is measured by the collection of expirate into a slightly permeable PVC bag, which is then drawn through a dry gas meter. An optically coupled detector in the dry gas meter drives an LED display (Harvard Apparatus Ltd) and measures the volume of the gas in relation to its ambient temperature, pressure and saturation level ($\dot{V}_{E\ ATPS}$). Concurrently, the temperature of the expirate is measured with a thermometer inserted into the dry gas meter inlet port. In contrast, automated ON systems does not measure gas volume, but instead flow-rate from which gas volume is calculated (Atkinson *et al.*, 2005). A criticism of this method of gas volume determination is that it is a derivative calculation, not a direct measurement (Macfarlane, 2001). The Sensor Medics Vmax uses a hot-wire anemometer (mass-flow sensor) to measure flow rate. This consists of two heated, stainless steel (gold plated) wires. The rate at which these wires cool as expirate passes over them is proportional to the number of gas molecules, or flow. These flow signals are integrated into a computer programme and produce a volume measurement that is linearised, unaffected by water vapour, and sensitive to ambient and expirate temperature changes (Sensor Medics reference manual). These fundamental differences in the determination of minute ventilation could be the cause of the systematic overestimation in ON minute ventilation, hence calculated $\dot{V}O_2$.

The second explanation for the discrepant $\dot{V}O_2$ method comparison is the (potentially) different equations used to calculate $\dot{V}O_2$. DB $\dot{V}O_2$ in the present study was calculated with the *Haldane Transformation* (section 3.6), which applied the three primary variables (Howley *et al.*, 1995) of, $\dot{V}_E$ Standardised for Temperature, Pressure and Dry (STPD), F_EO_2 and F_ECO_2. However, the variables used to calculate ON $\dot{V}O_2$ were elusive. Neither the ON system's users handbook, nor the manufacturers could confirm with certainty if $\dot{V}O_2$ was calculated from $\dot{V}_{E\ STPD}$ or $\dot{V}_{E\ BTPS}$. Assurances were made that $\dot{V}O_2$ was calculated from $\dot{V}_{E\ STPD}$, however this was an intuitive assumption that could not be verified. The lack of transparency in the calculation of $\dot{V}O_2$ was a good example of a black-box scenario, which is a common complaint of automated systems (Macfarlane, 2001; Hodges *et al.*, 2005). This ambiguity was problematic as $\dot{V}O_2$ can be calculated from two different $\dot{V}_E$ measures ($\dot{V}_{E\ STPD}$ or $\dot{V}_{E\ BTPS}$). When $\dot{V}O_2$ is calculated from $\dot{V}_{E\ STPD}$ the gas temperature is standardised to 0°C (or more accurately 273° Kelvin), but if calculated from $\dot{V}_{E\ BTPS}$ (Body Temperature, Pressure, Saturated) the gas temperature is higher due to the addition of body temperature (310°Kelvin = 273°Kelvin + 37°C (body temperature)). *Charles' Law* states that the volume of a gas is proportional to its temperature, hence if $\dot{V}O_2$ is calculated from $\dot{V}_{E\ BTPS}$ it will systematically inflate $\dot{V}O_2$ calculations by the same proportion (Powers and

Howley, 1997). In the present study DB $\dot{V}O_2$ was ~16% greater when $\dot{V}_{E\ BTPS}$ (instead of $\dot{V}_{E\ STPD}$) was used in the calculation of $\dot{V}O_2$. This brought DB $\dot{V}O_2$ inline with the observed ON $\dot{V}O_2$. Furthermore, $\dot{V}_{E\ BTPS}$ was ~16% larger than $\dot{V}_{E\ STPD}$ for the ON system. Either the systematic overestimation of ON $\dot{V}O_2$ was a coincidence, or the calculated DB and ON $\dot{V}O_2$ were derived from different variables (i.e. from $\dot{V}_{E\ BTPS}$ in the case of the ON, and from $\dot{V}_{E\ STPD}$ in the case of the DB system). Whilst this conclusion is both plausible and logical, a lack of transparency in the ON system makes this speculative.

Attributing the 16% overestimation in ON $\dot{V}O_2$ solely to systematic differences $\dot{V}_E$ may be an oversimplified interpretation, as; gas saturation levels entering the analysers, the measurement of the fraction of inspired air (F_IO_2 and F_ICO_2) used in the *Haldane Transformation*, and calibration procedures between the DB and ON systems were different, and may have also contributed to the systematic differences observed.

4.4.2 Repeatability

Repeatability provides a baseline of instrument variability against which method comparisons can be assessed (Bland and Altman, 1999), hence adequate levels of repeatability are essential if measurement systems are to be considered valid (Thomas and Nelson, 1996). Therefore, validation studies should conduct both repeatability and method comparisons (Bland and Altman, 1986), however this basic principle is often ignored (La Mere *et al.*, 1993; Miles *et al.*, 1994; Unnithan *et al.*, 1994; Yule *et al.*, 1996; Engebretson, 1998; Hiilloskorpi *et al.*, 2000; Midownik *et al.*, 2000; Bassett *et al.*, 2001; Rietjens *et al.*, 2001; Gore *et al.*, 2003).

Repeatability analysis revealed systematic biases ($P<0.05$) in 4 out of 6 ON measures (absolute and relative $\dot{V}O_2$, $\dot{V}CO_2$, and $\dot{V}_E$). A systematic bias in LoA analysis means a measure cannot be truly replicated (Bland and Altman 1986; 1999; 2003). Whilst this systematic bias was smaller than the systematic bias observed in the method comparison (-3% vs. 16%) it nevertheless indicated that the ON measurements were systematically altered during the process of test re-test. The direction of the ON bias (-3%) indicated lower values were attained during observation 2, which would be indicative of a learning effect. Although a systematic bias was absent from all DB analyses (Table 4.3), the removal of four outlying DB data points (Figure 4.4) also revealed a systematic bias in 3 out 6 DB measures (absolute and relative $\dot{V}O_2$, and $\dot{V}CO_2$) (data not tabulated). This DB systematic bias of -2% may also have been symptomatic of a learning effect. Systematic biases in $\dot{V}_E$ and $\dot{V}O_2$ have been reported in the same direction to the present study with a computerised system (gas analyser = Beckman LB-2 and Applied Electrochemistry; gas meter = Parkinson-Cowan CD-4). Here the ventilatory and O_2 demands during sub-maximal cycling and running decreased over 4 laboratory visits suggesting an order effect may have caused the systematic biases (Armstrong and Costill, 1985), which was supported elsewhere (Morgan *et al.*, 1991; Crouter *et al.*, 2006). Thus, in the present study two familiarisation sessions conducted prior to data collection may have been insufficient to negate the learning effect in the novel exercise modality of loaded marching. Hence, the possible existence of a learning effect in LME should be considered in future investigations that track changes in LME over the duration of a training programme in order to separate physiological adaptation from biomechanical learning (i.e. improved motor co-ordination).

The repeatability of ON and DB $\dot{V}O_2$ was comparable (~15% to 18%) (Table 4.3), which has also been reported with the coefficient of variation statistic (Carter and Jeukendrup, 2002; Jensen *et al.*, 2002). However, visual inspection of the Bland and Altman plots (Figure 4.4 and 4.5) revealed that the DB data points were more tightly grouped around zero line than the ON system. This suggested the DB system had better repeatability (even though the 95% LoA statistic did not). In the one gas analysis validation similar in design to the present study (i.e. applied Bland and Altman's 95% LoA, a sub-maximal exercise model, and conducted 5 repeated observations on the same 10 subjects at different exercise intensities) (Crouter *et al.*, 2006), the smallest ratio LoA were reported with the automated system (0.98 (×/÷ 1.19)), then DB (0.98 (×/÷ 1.24)), and finally the portable system (0.94 (×/÷ 3.30)). In all instances the LoA were wider than both the DB (1.00 (×/÷ 1.15)) and ON systems (0.97 (×/÷ 1.15)) in the present study, by at least 4 %. The authors claimed that the LoA of ~20% in the automated systems represented reasonable repeatability. This seems an erroneous interpretation as an individual with a sub-maximal $\dot{V}O_2$ of 35 $ml \cdot kg^{-1} \cdot min^{-1}$ could expect $\dot{V}O_2$ to randomly vary by ± 7 $ml \cdot kg^{-1} \cdot min^{-1}$, which is unacceptable considering this would have approached the between-subject variation in LME for 25 kg loads in the present study of 6.8 $ml \cdot kg^{-1} \cdot min^{-1}$ (from 19.7-26.5 $ml \cdot kg^{-1} \cdot min^{-1}$). Furthermore, LME also improves by a much smaller amount (~1 $ml \cdot kg^{-1} \cdot min^{-1}$) in response to a physical training programme (Gutekunst *et al.*, 2006).

It has been suggested that an acceptable daily and weekly variation in $\dot{V}O_2$ is ± 2 $ml \cdot kg^{-1} \cdot min^{-1}$, and ± 3 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively (Thoden, 1991). Whilst these values approximate DB and ON repeatability (Table 4.3), they refer to maximal, and not sub-maximal exercise (i.e. LME). As such, acceptable absolute LoA (i.e. repeatability expressed in actual units of measurement) for maximal exercise are likely to be wider than for sub-maximal exercise due to the inherent presence of heteroscedasticity in ratio scale data (Nevill and Atkinson, 1997). Consequently, in the present study the absolute LoA of ~2.7 $ml \cdot kg^{-1} \cdot min^{-1}$ (at best) may be larger than those deemed acceptable (Thoden, 1991). Therefore, the potentially small training effect for LME of ~1 $ml \cdot kg^{-1} \cdot min^{-1}$ (Gutekunst *et al.*, 2006) may be too small to detect adaptations in LME, if inadequate sample sizes are selected. Therefore, to avoid a lack of statistical power in subsequent studies prospective sample size estimations should be conducted.

In the case of the DB system (but not the ON) outlying data points were a likely explanation for inflated $\dot{V}O_2$ LoA. In total five outlying data points (4 DB and 1 ON) were identified (Figures 4.2, 4.4 and 4.5). Removal of DB outliers almost halved the LoA from ~15% to ~9% (absolute units = ~2.7 $ml \cdot kg^{-1} \cdot min^{-1}$ to ~1.4 $ml \cdot kg^{-1} \cdot min^{-1}$), which approached a level of repeatability that is deemed acceptable (Thoden, 1991). No such improvement was observed when the single ON $\dot{V}O_2$ outlier was removed (Figure 4.5). Cases were considered potential outliers if test re-test differences exceeded at least 1.5 box lengths outside the range of normal boxplot values (Field, 2001). The four outlying DB data points (all from the same subject) seem to be a consequence of inflated F_EO_2 values during observation 1. Here, F_EO_2 ranged between 17.1% to 17.4%, which are values typically observed during intense/maximal exercise, not light/moderate sub-maximal exercise. This finding was pertinent as errors in F_EO_2 exert the greatest influence on the calculation of $\dot{V}O_2$ (Withers *et al.*, 2000; Atkinson *et al.*, 2005). As the outlying values were all from the same subject / observation it is possible the higher F_EO_2 values were a result of a systematic error during the calibration of the gas analyser. Applying this logic, it is suggested that DB repeatability is almost two-fold better than the ON system, and approaches that deemed acceptable (Thoden, 1991).

As the repeatability of $\dot{V}O_2$ was similar between DB and ON analysers (when outliers are included) (Table 4.3) $\dot{V}_E$, F_EO_2 and F_ECO_2 can be compared to identify the cause of the variability within each system. The main sources of DB error (compared to ON data) appear to be differences in F_EO_2 (5 % vs. 4%) and F_ECO_2 (~14 % vs. ~8%), and not $\dot{V}_E$. Although the 1% difference between DB and ON F_EO_2 seems negligible this could have elicited as much as a 6.5% error in $\dot{V}O_2$ calculation (Withers *et al.*, 2000). The main source of ON error (compared to the DB data) was $\dot{V}_E$, which was approximately two-thirds larger than DB error (~21% vs. ~14%). This was consistent with method comparison findings, and previous reports (Hodges *et al.*, 2005). Unfortunately, inconsistent reporting of the primary $\dot{V}O_2$ variables (i.e. $\dot{V}_E$, F_EO_2 and F_ECO_2) in the literature makes empirical sources of error in $\dot{V}O_2$ elusive (Cullum *et al.*, 1999; Skinner *et al.*, 1999; Carter and Jeukendrup, 2002; Crouter *et al.*, 2006).

In conclusion the novel physiological measure of loaded marching economy has the potential to be an important determinant of loaded marching performance, and has a relative $\dot{V}O_2$ range of 19.7 $ml \cdot kg^{-1} \cdot min^{-1}$ to 26.5 $ml \cdot kg^{-1} \cdot min^{-1}$, in recreationally active civilian males (in a 25 kg simulated loaded march). The ON gas analysis system systematically overestimated $\dot{V}O_2$ by ~16% compared to the DB, which may have been a function of differences in the calculation of $\dot{V}O_2$. Following two familiarisation sessions an apparent learning effect of 2-3% was evident in loaded marching economy, which should be considered in the interpretation of results in subsequent studies. The removal of outlying data points as a consequence of test re-test differences scores, visual inspection of the Bland and Altman plots, and suspect F_EO_2 values revealed that ON repeatability was ~15% compared to ~9% for the DB. The primary cause of the variation in $\dot{V}O_2$ calculation appeared to be variability in the calculation of flow rate in the case of the ON, and measurement expired gas fractions (i.e. F_EO_2 and F_ECO_2) in the case of the DB system. As the ON system systematically overestimated $\dot{V}O_2$, and had poorer repeatability compared to the DB system it is suggested that the DB system be the sole instrument used to measure the rate of oxygen uptake subsequently in this thesis.

CHAPTER 5

THE DIFFERENCE BETWEEN RUNNING AND LOADED MARCHING $\dot{V}O_2max$ PROTOCOLS

5.1 Introduction

The identification of an appropriate measure of maximal oxygen uptake in the appraisal of a physical training programme designed to improve loaded marching performance (LMP) is warranted, as $\dot{V}O_2max$ underpins LMP (Rayson *et al.*, 1993, 2000, 2002a, 2002b; Harman and Frykman, 1995; Pandorf *et al.*, 2002). Maximal oxygen uptake (or $\dot{V}O_2max$) is the highest rate oxygen can be utilised by the body during severe exercise, at sea level (Bassett and Howley, 2000). The values attained during these tests are exercise modality dependent, as specific protocols such as cycling (Hermansen and Saltin, 1969), skiing and swimming (Åstrand and Saltin, 1961) and walking (Froelicher *et al.*,1974; Sheehan *et al.*, 1987) have all elicited lower maximal oxygen uptakes than standard running protocols. Nonetheless, these specific measures of maximal oxygen uptake are valid as they stress the relevant central and peripheral aspects of a physiological system (Shephard, 1984).

Whilst specificity is a well endorsed training principle (Tanaka, 1994), the application of a non-specific running $\dot{V}O_2max$ protocol (using both direct and indirect assessments) has been utilised in the load-carriage research to describe training progression (Harman *et al.*, 1998; Williams *et al.*, 1999, 2002, 2004; Reynolds *et al.*, 2001), mathematically model LMP (Dziados *et al.*, 1987; Knapik *et al.*, 1990; Rayson *et al.*, 1993, 2000, 2002a, 2002b, Frykman and Harman, 1995; Harman and Frykman, 1995; Pandorf *et al.*, 2002, Williams and Rayson, 2006), and express LM relative exercise intensity (Rayson *et al.*, 1995). This is surprising considering other exercise modalities, such as running consistently use exercise specific protocols to quantify $\dot{V}O_2max$ for running performance (Noakes *et al.*, 1990; Joyner, 1991). Just one study has used a loaded march specific $\dot{V}O_2max$ protocol to describe the physiological ceiling of load-carriers, and subsequently model LMP (Simpson *et al.*, 2006). This study reported *moderate* correlations between 3.2 and 29 km LMP, and indices of a constant speed-incremental gradient maximal loaded march protocol.

Implementing a specific LM $\dot{V}O_2max$ protocol may be valid, as the relative exercise intensity of LM (i.e. %$\dot{V}O_2max$) is calculated from the $\dot{V}O_2max$ (Rayson *et al.*, 1995). In this study LM $\dot{V}O_2max$ was lower compared to a running $\dot{V}O_2max$ protocol. Thus, the expression of sub-maximal LM exercise intensity as a percentage of running $\dot{V}O_2max$ and running heart rate maximum (HR_{max}) (not loaded marching, $\dot{V}O_2max$ and HR_{max}), underestimated relative exercise intensity by 20.4% and 9.5%, respectively. Consequently, the prescription and/or description of LM exercise intensity from a non-specific $\dot{V}O_2max$ protocol (i.e. running) should be interpreted in the knowledge that the actual exercise intensity of LM may be greater than that calculated from a running $\dot{V}O_2max$ protocol.

The purpose of this study was to quantify differences in peak physiological measures between running and loaded marching $\dot{V}O_2max$ protocols. This would establish the extent LM exercise intensity was under or overestimated when derived from a running $\dot{V}O_2max$ protocol, as this will be the method used to determine $\dot{V}O_2max$ subsequently in this thesis.

5.2 Method

5.2.1 Subjects

Ten male University students and staff volunteered to participate in the study (Table 5.1), which had previously been approved by the University of Chichester Ethics Committee. All were recreationally active and had prior laboratory and treadmill experience. None were engaged in habitual load-carriage, but all had engaged in some form of load-carriage previously (i.e. hiking with backpacks). None reported a previous history of back problems. Measurements took place in an air conditioned laboratory that was controlled for ambient temperature (17.9 ± 0.8 °C) and relative humidity (54 ± 7%).

Table 5.1. Subject characteristics (mean ± SD)

Age (years)	Stature (m)	Body mass (kg)	Estimated body fat (%)
29.2 ± 7.4	1.80 ± 0.05	80.3 ± 5.1	17.1 ± 3.9

5.2.2 Experimental design

A within-subjects cross-over design was performed over two visits to the laboratory, separated by at least 48 hours. Each visit consisted of a LM familiarisation period that preceded each of the maximal incremental tests. The maximal tests included a standard run protocol (R_P) and a specific loaded march protocol (LM_P). To prevent an order effect the tests were counterbalanced between-subjects (Figure 5.1).

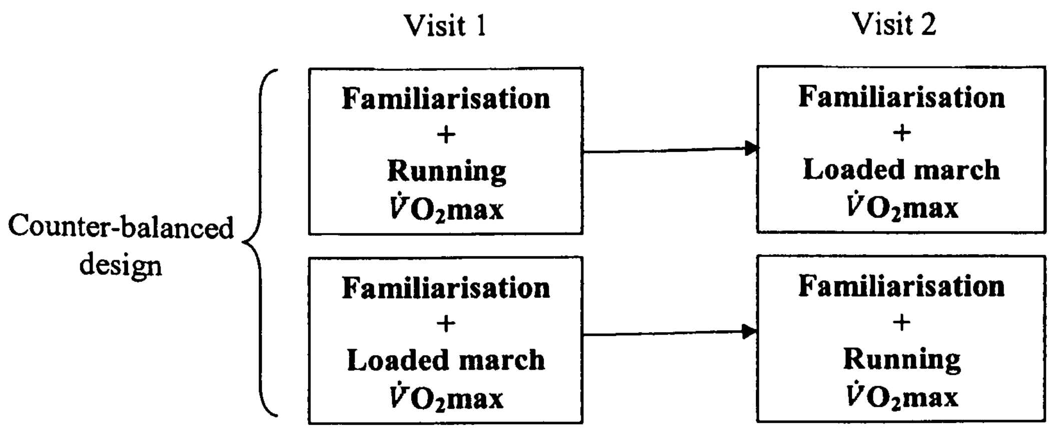


Figure 5.1. A schematic representation of the 2-visit, counterbalanced experimental design.

5.2.3 Testing protocols.

5.2.3.1 LM familiarisation

LM familiarisation was conducted as previously described (section 4.2.3.1).

5.2.3.2 Maximal run protocol (R_P)

In training shoes, shorts and T-shirt an incremental speed and gradient maximal run protocol was performed (Draper *et al.*, 2003). A speed of 7 $km \cdot h^{-1}$ was maintained for the initial 5 minutes, during which time gradient increased from 1 to 5%, at a rate of $1\% \cdot minute^{-1}$. From the 6th-minute the gradient remained constant, and treadmill speed increased at a rate of 1.2 $km \cdot h^{-1} \cdot minute^{-1}$ (0.2 $km \cdot h^{-1} \cdot 10\ s^{-1}$) until the point of volitional exhaustion. Serial Douglas Bag collections were made (section 3.6.1) from the 6th-minute of the test. To minimise the variability observed in $\dot{V}O_2max$, expired gas collections were only accepted if the final bag time was ≥30 s (Wood, 1999). Peak heart rate (HR) was the retrospective average of the final minute of the test. Blood was drawn immediately upon cessation of the test from a finger tip capillary sample, and analysed for whole blood plasma lactate concentration ([BLa]) (YSI 2300, Stat Plus, Yellow Springs Instruments, Ohio, USA). RER was calculated from the respiratory quotient ($\dot{V}CO_2/\dot{V}O_2$).

5.2.3.3 Maximal loaded march protocol (LM_P)

In training shoes, CS 95 clothing, and PLCE weighing 25 kg (section 3.9) a constant speed, incremental gradient maximal loaded march protocol was performed, similar to that of Balke and Ware (1959). A constant speed of 6.5 $km \cdot h^{-1}$ was maintained throughout the test. Gradient was initially 0% and increased at a rate of 1 $\% \cdot minute^{-1}$ (0.5 $\% \cdot 30\ s^{-1}$) until the point of volitional exhaustion. The measures of $\dot{V}O_2$, HR, [BLa], and RER were determined in the same manner as the R_P.

5.2.4 $\dot{V}O_2max$ criteria

The term $\dot{V}O_2max$ and $\dot{V}O_2peak$ have been used inconsistently in the literature. It has been suggested that $\dot{V}O_2peak$ should only be used in clinical settings when exercise tolerance is limited by a pathophysiological condition (Midgley *et al.*, 2007). The term $\dot{V}O_2max$ will solely be used to describe the maximal rate of oxygen uptake in this thesis (regardless of the exercise modality). To verify the attainment of $\dot{V}O_2max$, and the appropriateness of both protocols a $\dot{V}O_2$-plateau criterion was applied. To maximise the incidence of the $\dot{V}O_2$-plateau, separate criteria were devised for each protocol (Freedson *et al.*, 1986). The $\dot{V}O_2$-

plateau was classified as a rise in $\dot{V}O_2$ during the final stage of the test of less than 1.96 SDs of the mean change in $\dot{V}O_2$, during the preceding three, 1-minute expired gas collection periods (Rowland and Cunningham, 1992). This was considered valid as it ensured a separate $\dot{V}O_2$-plateau criterion was devised for each subject (Midgley *et al.*, 2007). Three secondary citeria (Howley *et al.*, 1995) of, [BLa] $\geq$ 8 mmol·l^{-1} (Åstrand, 1952), RER $\geq$ 1.15 (Issekutz *et al.*, 1962), and HR $\pm$ 10 b·min^{-1} of age predicted maximum (Åstrand, 1960) also verified the attainment (or not) of maximum effort during the test.

5.2.5 *Statistical analysis*

Data are expressed as mean ± SD. Data were all normally distributed (i.e. Skewness ÷ SE Skewness, or Kurtosis ÷ SE Kurtosis ≥2) (Fallowfield *et al.*, 2005). A paired samples *t*-test identified differences in peak variables between the two test protocols. Significance was accepted as $P<0.05$.

5.3 Results

5.3.1 *Peak physiological responses.*

All 10 subjects had a lower $\dot{V}O_2max$ in the LM_P compared to the R_P (Figure 5.2).

Time to volitional exhaustion was 66 s longer in the LM_P, compared to the R_P (t=-3.168, P=0.011). Five out of six peak LM_P measures were lower than the R_P measures ($P<0.05$) (Table 5.2). Of these lower values, relative $\dot{V}O_2max$ was 2.7 ml·kg^{-1}·min^{-1} (t=5.046, P=0.001); absolute $\dot{V}O_2max$ 0.21 l·min^{-1} (t= 4.662, P=0.001); RER 0.04 (t=2.436, P=0.038); HR 2 beats·min^{-1} (t=3.233, P=0.012); and $\dot{V}_E$ 8.9 l·min^{-1} (t=2.729, P=0.023) lower than the R_P. Circulating blood plasma lactate concentration ([Bla]) was the only variable not statistically different between protocols (t = -1.513, P=0.165).

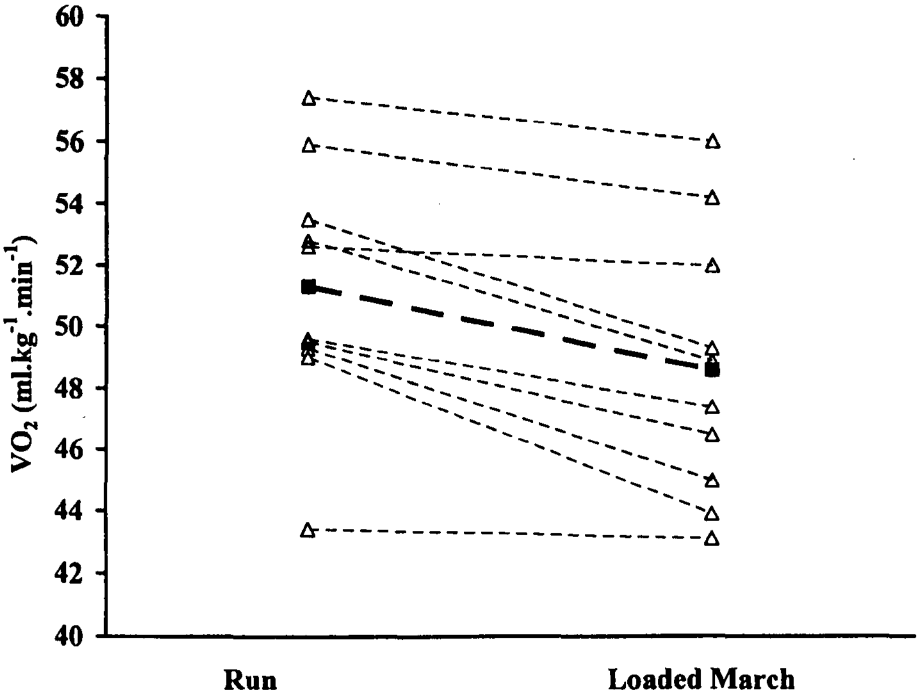


Figure 5.2. Data from two separate laboratory visits, as indicated by dashed line. Individual (triangles) and mean (square) $\dot{V}O_2max$ values for the run (R_P) and loaded march (LM_P) protocols.

Table 5.2. Differences (mean ± SD) in peak physiological variables during run and loaded marching protocols (n = 10).

	Run	Loaded March
$\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	51.3 ± 4.0	48.6 ± 4.3 *
$\dot{V}O_2$ (l·min^{-1})	4.11± 0.39	3.90 ± 0.36 *
[BLa] (mmol·l^{-1})	8.0 ± 1.5	8.6 ± 2.0
RER	1.23 ± 0.07	1.19 ± 0.06 *
HR (beats·min^{-1})	187 ± 11	185 ± 11 *
$\dot{V}_E$ (l·min^{-1})	132.2 ± 9.8	123.3 ± 13.1 *

* statistical significance ($P < 0.05$)

5.3.2 $\dot{V}O_2max$ criteria

None of the four $\dot{V}O_2max$ criteria achieved an incidence rate of 100%. The $\dot{V}O_2$-plateau incidence was 70% (7 out of 10 subjects) for both protocols. The [BLa] criterion was 60% (6/10) for LM_P and 50% (5/10) for R_P. The RER criterion was 70% (7/10) for LM_P and 90% (9/10) for R_P. Finally, the HR criterion was 66% (6/9) for LM_P, and 70% (7/10) for R_P (Figure 5.3).

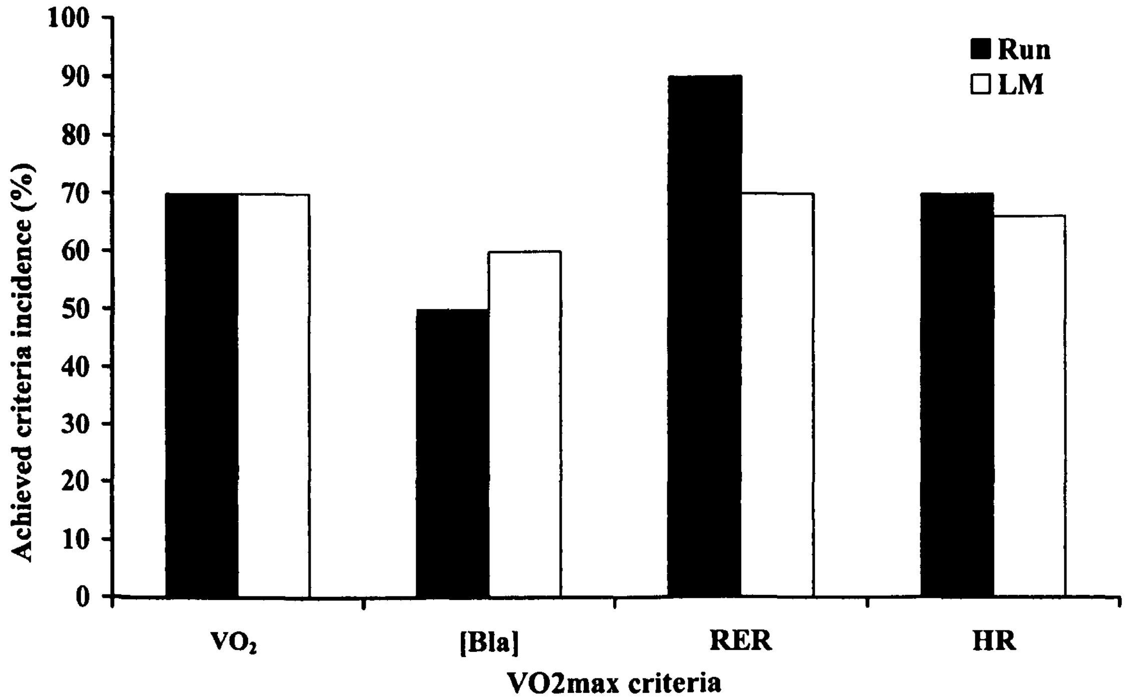


Figure 5.3. The incidence of $\dot{V}O_2max$ criteria for R_P and LM_P (n = 10).

5.4 Discussion

5.4.1 Main findings

The main finding of this study was that an incremental loaded march protocol elicits a maximal $\dot{V}O_2$ 5.3% lower than a standard running protocol.

5.4.2 Peak physiological responses

Whilst it took longer to reach volitional exhaustion in the LM_P (12:36 ± 1:34 min:s vs. 11:30 ± 0:47 min:s, $P<0.05$) the mean duration of both protocols were consistent with that recommended (Buchfuhrer *et al.*,1983). The $\dot{V}O_2max$ attained in the LM_P and R_P were different in magnitude. In the LM_p, $\dot{V}O_2$ (5.3%), RER (3.3%), HR (1.1%), and $\dot{V}_E$ (6.7%) were lower than the R_p. Consequently, sub-maximal LM exercise intensities would be underestimated by 5.3% if expressed relative to running $\dot{V}O_2max$. The expression of LM exercise intensity relative to running $\dot{V}O_2max$ could result in a miscalculation of the sustainable exercise intensity during prolonged marches (Christie and Scott, 2005).

Different exercise modalities elicit different $\dot{V}O_2max$ scores. The $\dot{V}O_2max$ during cycling (Hermansen and Saltin, 1969), as well as skiing and swimming (Åstrand and Saltin, 1961) were 5-7% lower than treadmill running protocols. More closely related to the present study, walking protocols elicit a $\dot{V}O_2max$ 6.5% to 9.7% lower than running protocols (Froelicher *et al.*,1974; Kang *et al.*, 1999). One possible explanation for smaller differences observed between the $\dot{V}O_2max$ protocols in the present study (5.3%) could be due to walking protocols being constrained by the walk-run transition. In the present study, subjects were allowed to alter gait from a walking pattern to a slow jog (or shuffling) movement towards the end of the test. This replicated the reality of military loaded marching that is not constrained to a walking gait. This allowed subjects to continue for longer than would have been possible if a walking gait alone were adhered to. Hence the duration of the LM_p was extended, which may have further augmented $\dot{V}O_2max$, hence explain the smaller difference between protocols in the present study compared to previous walking *versus* running comparisons.

Just one study has previously investigated differences between maximal running and loaded marching protocols (Rayson *et al.*, 1995). Unlike the constant speed (6.5 $km \cdot h^{-1}$) incremental gradient ($1\% \cdot min^{-1}$) protocol in the present study, a constant gradient (0%) and speed (6.4 $km \cdot h^{-1}$), but incremental load (5.0 kg to 2.5 kg every 4 minutes) protocol was compared to a running protocol. The LM $\dot{V}O_2max$ of 2.12 $l \cdot min^{-1}$ was 0.54 $l \cdot min^{-1}$ (20.4%), and HR_{max} of 172 $beats \cdot min^{-1}$ was 18 $beats \cdot min^{-1}$ (9.5%) smaller than the running protocol. This was a markedly larger difference than the 5.3%, and 1.1% difference in $\dot{V}O_2max$ and

HR_{max}, respectively in the present study. The steeper gradient at the point of volitional exhaustion during the LM_p in the present study compared to Rayson *et al.*, (1995) (12.0 ± 1.6% vs. 0%) may explain the smaller difference observed between R_p and LM_p in the present study. An increase in treadmill gradient produces higher $\dot{V}O_2max$ values that are thought to be due to the recruitment of 'accessory' muscles (i.e. synergist and fixators) (Taylor *et al.*, 1955). Furthermore, a 10% difference in treadmill gradient can recruit an additional 0.5 kg of active muscle mass in the lower limb that will demand additional oxygen, thereby increasing $\dot{V}O_2max$ by 1.3 $ml \cdot kg^{-1} \cdot min^{-1}$ (3%) (Sloniger *et al.*, 1997). Hence, the steeper gradient reached during the LM_p in the present study may partly explain the smaller difference between LM and running $\dot{V}O_2max$ values in the present study, compared to Rayson *et al.*, (1995).

However, the most likely explanation for such variation in the differences between LM and running protocols in these studies might be the relative size of the external loads used in the loaded march protocols (i.e. % body mass). A constant load of 25 kg was applied in the present study, however Rayson *et al.*, (1995) increased the load with increments of 5.0 kg and 2.5 kg, but did not report the size of the load at volitional exhaustion. Compared to increases in speed ($km \cdot h^{-1}$) or gradient (%), an increase in load size (kg) is the least effective method to increase the oxygen demand of LM (Santee *et al.*, 2003a; Christie and Scott, 2005). For example marching at 6.4 $km \cdot h^{-1}$ with a load of 70 kg (~93% of body mass) elicited a $\dot{V}O_2$ of 92.5% of $\dot{V}O_2max$ (Soule *et al.*, 1978). It is therefore probable that Rayson and colleagues loaded subjects to a very high percentage of body mass to reach volitional exhaustion, which may be the reason why these subjects did not appear to be centrally limited in the loaded march protocol. Thus, it would appear that constant load – incremental gradient LM protocols like that used in the present study are a more valid assessment of loaded marching $\dot{V}O_2max$, than that of a constant gradient – incremental load protocol.

The LM_P was a valid measure of maximal LM $\dot{V}O_2max$, however the higher $\dot{V}O_2max$ attained in the R_P suggest that the LM_P was not limited centrally by the same magnitude. One possible explanation for the smaller central response in the LM_p may have been impaired pulmonary function, which is often reported during load-carriage tasks. Pulmonary function decreases inversely with the magnitude of a load carried (Muza *et al.*, 1989), and occurs with loads as light as 6 kg (Legg and Cruz, 2004). This impairment is attributable to a combination of load size, and strap tightness that restricts the mechanical expansion of the thoracic cage (Bygrave *et al.*, 2004). As pulmonary function is a potential physiological factor to limit $\dot{V}O_2$max (Bassett and Howley, 1997) it is of little surprise that the maximal rate of oxygen transport decreased during the LM_p.

Minute ventilation ($\dot{V}_E$) during the LM_P was 6.7%, and $\dot{V}O_2$max 5.3% lower than the R_P. As a change in minute ventilation results in a change in calculated $\dot{V}O_2$ of approximately the same proportion (Withers *et al.*, 2000), it is likely the lower minute ventilation in the LM_P explains the proportional decrease in calculated $\dot{V}O_2$, which may reflect the impaired mechanical movement of the thoracic cage (Bygrave *et al.*, 2004). Reduced thoracic cage movement causes a reduction in the pressure gradient between the atmosphere and the lungs, thus the process of *Bulk flow* is affected as the movement of air from high pressure (i.e. atmosphere) to low pressure (i.e. lungs) is impaired due to the smaller pressure gradient. As a result less air enters the lungs, and minute ventilation will decrease (Powers and Howley, 1997). Tests of pulmonary function were not conducted in the present study, thus establishing the lower minute ventilation as the cause of the lower $\dot{V}O_2$max during the LM_P is speculative. However, loads of 10 kg and 30 kg have caused decreases in Maximal Voluntary Ventilation over 15 s (9.5%), Forced Vital Capacity (6%), and Forced Expiratory Volume in 1 second (6.7%) (Muza *et al.*, 1989), which was similar in magnitude to the 6.7% minute ventilation decrease in the present study. Thus, the lower $\dot{V}O_2$max may reflect an impairment to thoracic cage movement during a maximal load-carriage task. This may limit pulmonary function hence oxygen delivery, thereby highlighting the specific nature of LM that may differentiate it (physiologically) from other exercise modalities such as running.

The RER was 3.3% lower during the LM_P. RER is a secondary criteria that verifies the validity of maximum effort (Howley *et al.*, 1995). Whilst the mean, peak RER for LM_p exceeded the RER criterion of 1.15 (Issekutz *et al.*, 1962), it was lower than the R_p (1.19 vs. 1.23). This indicated that the LM_p acid-base balance was not challenged as severely compared to the R_P, and is a reflection of a smaller production of non-metabolic CO_2 (Jeukendrup and Wallis, 2005). In addition, the absence of a difference in [BLa] was an unexpected finding. This observation seemed counterintuitive, as the lower RER in the LM_p would suggest less non-metabolic CO_2 was produced (thus lower $\dot{V}CO_2/\dot{V}O_2$ quotient) as less bicarbonate would be required to buffer hydrogen production (Beaver *et al.*, 1986).

This anomaly may be due to the manner with which finger tip capillary blood samples were taken. Capillary blood samples were drawn immediately upon cessation of both protocols. In the case of the LM_P blood was drawn whilst subjects were wearing the bergen, thus the straps may have occluded the local blood flow to the arms, thereby creating a partially ischemic environment.

5.4.2 $\dot{V}O_2max$ criteria.

The $\dot{V}O_2max$ criteria verify the appropriateness of a $\dot{V}O_2max$ protocol, and the honesty of effort in a maximal test (Midgley *et al.*, 2007). Both test protocols failed to satisfy all four $\dot{V}O_2max$ criteria, in all 10 subjects (Figure 5.3). Between-subject differences in training status or training emphasis (i.e. sprint vs. endurance training) explain why certain individuals may, or may not satisfy the $\dot{V}O_2max$ criteria (Midgley *et al.*, 2007). Additionally, three of the four $\dot{V}O_2max$ criteria (i.e. RER, [BLa], age predicted HRmax) were developed on different subjects, and protocols, thus it should be no surprise that some criteria were not satisfied (Duncan *et al.*, 1997).

The $\dot{V}O_2$-plateau is the primary $\dot{V}O_2max$ criterion (Howley *et al.*, 1995). Its relevance derives from the early work of Hill and Lupton, (1923) that reported an absent increase in $\dot{V}O_2$ with an increase in work-rate as oxygen consumption reached a maximum rate. Although the existence of the $\dot{V}O_2$-plateau in maximal oxygen uptake tests has been challenged (Noakes, 1988; 1997; 1998) the widely held belief remains that it does exist, and represents a central limitation in the maximal rate of oxygen supply (Bassett and Howley, 1997; 2000). In the present study the incidence of the $\dot{V}O_2$-plateau was identical for both protocols (70%), despite a significantly higher $\dot{V}O_2max$ in the R_P. This may indicate that both protocols were capable of eliciting a true $\dot{V}O_2max$ relative to each exercise modality, despite the observation that $\dot{V}O_2max$ was not limited centrally by the same extent in the LM_P. A possible reason for this may be due to specific $\dot{V}O_2$-plateau criteria devised for each test protocol, and each individual subject as recommended. (Freedson *et al.*, 1986; Midgley *et al.*, 2007). This is further supported by the fact that the same 7 subjects exhibited the $\dot{V}O_2$-plateau during the R_p and LM_p.

The incidence of the $\dot{V}O_2$-plateau in the present study was higher, (< 50%; Froelicher *et al.*, 1974; Niemelä *et al.*, 1980; Sheehan *et al.*, 1987); similar, (50 – 70%; Duncan *et al.*, 1997;

Mayhew and Gross, 1975); and lower, (> 70%; Davies *et al.*, 1984; Taylor *et al.* 1955) than pervious studies. Such a wide variation is unsurprising given that different test protocols (Rivera-Brown *et al.*, 1993), and $\dot{V}O_2$-plateau criterion were applied (Freedson *et al.*, 1986). If the popular $\dot{V}O_2$-plateau criterion of 2.1 $ml \cdot kg^{-1} \cdot min^{-1}$ (Taylor *et al.*, 1955) were applied to the present study a higher $\dot{V}O_2$-plateau incidence would have been observed (70% vs. 80%). As the Taylor criterion is reputedly too generous (Howley *et al.*, 1995) the criteria devised in the present study was more likely identify a 'true' $\dot{V}O_2$-plateau.

In conclusion, a loaded marching protocol elicited a maximal oxygen uptake 5.3% smaller than a generic running protocol. This demonstrated that loaded marching $\dot{V}O_2max$ is specific in nature, but not as specific as previously thought (Rayson *et al.*, 1995). It is possible the lower $\dot{V}O_2max$ attained during maximal loaded marching may be explained by an impairment in pulmonary function brought about by the addition of a Bergen, which may have restricted the mechanical movement of the thoracic cage. The identical $\dot{V}O_2$-plateau incidence between protocols (70%) suggested that peripheral limitation may not have been responsible for the lower values attained in the loaded marching protocol. Finally, the subsequent use of running $\dot{V}O_2max$ to express relative loaded marching exercise intensity (Chapters 7, 8, and 9) will systematically underestimate the actual exercise intensity of loaded marching by 5.3%.

CHAPTER 6

THE REPEATABILITY OF ACCEPTED AND POTENTIAL DETERMINANTS OF LOADED MARCHING PERFORMANCE.

6.1 Introduction

Classically, the determinants of loaded marching performance (LMP) include; absolute $\dot{V}O_2max$, maximal running performance, dynamic and isometric strength, and anthropometric measures (Rayson *et al.*, 1993, 2000, 2002a; Frykman and Harman, 1995; Harman and Frykman, 1995; Williams and Rayson, 2006). Thus, a physical training programme that improves these measures should improve LMP. A criticism of these cited studies is that none considered specific measures of loaded marching (LM), which is surprising given that research performed on other exercise modalities (i.e. running) understood the importance of specificity on performance outcome (Joyner, 1991; Brandon, 1995).

These accepted determinants of LMP indicate that LMP is the partial product of a high maximal oxygen uptake, and a large body size (i.e. structure), or strength (i.e. function) (Haisman, 1988). Absolute $\dot{V}O_2max$ is a combination of oxygen delivery capacity and body size, and is the single most important determinant of LMP (Rayson *et al.*, 2000). Therefore the precision with which $\dot{V}O_2max$, strength, and anthropometric measures can be tracked over a training programme may have implications on inferential analyses that appraise physical training for LMP.

Potential measures that may determine LMP include maximal and sub-maximal indices of LM ability. To date, two studies have highlighted the importance of maximal LM in determining LMP (Simpson *et al.*, 2006; Griggs *et al.*, 2008). In addition, the specific nature of maximal LM physiology was also confirmed in Chapter 5. However, the importance of sub-maximal LM measures on LMP is yet to be reported, which is surprising given that sub-maximal exercise economy is a well established determinant of endurance exercise performance (Joyner, 1991; Brandon, 1995; Fallowfield and Wilkinson, 1999). Furthermore, improvements in exercise economy through training will cause concomitant improvements in exercise performance (Jones and Carter, 2000). The responses of sub-maximal loaded marching economy (LME) to a military style training intervention was investigated (Gutekunst *et al.*, 2006). Here the training response approximated 6%, but the absence of a criterion LMP test meant the associated link between LME and LMP could not be made. As such, further research is required on LME, however in the first instance it would be prudent to quantify the repeatability of this novel measure.

Fundamentally, a measurement approach should be repeatable as this will allow it to be valid, discriminatory, and sensitive (Abernethy and Wilson, 2000). Knowledge of the repeatability of a test measure is essential in determining the correct sample size required in future studies. Choosing the correct sample size means adequate statistical power for an anticipated effect size will be generated for a pre-determined level of alpha, which is necessary to reject the null hypothesis (Thomas and Nelson, 1996). Thus, the more repeatable a measure is, a smaller sample size or effect size will be required to detect a statistical significance, which may be an important consideration in understanding how best to improve physical training for LMP.

The purpose of this study was to quantify the repeatability of accepted and potential determinants of loaded marching performance. These measures will subsequently track responses in British Army infantry recruits for the purpose of understanding physical training prescription for the improvement of LMP.

6.2 Methods

6.2.1 Subjects

Twenty-two male post graduate students and staff (Table 6.1) volunteered to participate in the research, which had been approved by the University of Chichester Ethics Committee. All were actively involved in sport and exercise and had prior experience of physiological laboratory testing.

Table 6.1. Physiological characteristics of the subjects (mean ± SD).

	Age (years)	Stature (m)	Body Mass (kg)	$\dot{V}O_2$max ($ml \cdot kg^{-1} \cdot min^{-1}$)
	30.6 ± 6.6	1.81 ± 0.06	82.5 ± 10.9	51.1 ± 4.9

6.2.2 *Experimental design*

The accepted and potential determinants of LMP were evaluated for repeatability, and presented here (Table 6.2 to 6.5), or elsewhere (Appendix A). The repeatability of test measures was determined from two identical visits, separated by at least 2 days, but no more than 14 days. All tests were conducted at the same approximate time of day within-subjects. The order test measures were undertaken is listed:

1. Body mass (section 3.2.1);
2. Estimates of body composition (section 3.2.5);
3. Dynamic strength (chest press, back pull, leg press) (section 3.4);
4. Isometric strength (hand grip, static lift strength) (section 3.3);
5. Sub-maximal loaded marching economy (LME) (section 3.5.2.2);
6. Maximal running test ($\dot{V}O_2$max) (section 3.5.2.4).

6.2.3 *Test battery*

The test battery listed (section 6.2.2) is described in full in Chapter 3. Of note a Douglas bag gas analysis system was built for this study (section 3.6), which was based on the approach used in Chapters 4 and 5. In later studies this system will be applied to a field laboratory (Chapters, 7, 8, and 9).

6.2.4 *Statistical analysis*

Bland and Altman's 95% Limits of Agreement (LoA) quantified the repeatability of the test measures (Altman and Bland, 1983; Bland and Altman, 1986), and 95% Confidence Intervals (CI) were calculated to estimate the precision of the 95% LoA (Bland and Altman, 1986; 2003). A positive Pearson's correlation coefficient indicated test re-test differences were heteroscedastistic (Nevill and Atkinson, 1997). See section 3.11 for a detailed discussion.

6.3 Results

Table 6.2 presents the mean of test 1 and test 2, and correlation coefficients to identify data heteroscedasticity (i.e. a positive correlation when absolute test re-test differences are plotted against the pooled mean of test 1 and 2). An example of data heteroscedasticity is presented (Figure 6.1). The high incidence of heteroscedasticity (Table 6.2) meant data were more appropriately expressed as Ratio LoA (Nevill and Atkinson, 1997). Ratio LoA are expressed as, **systematic bias (x/÷ 95% LoA).**

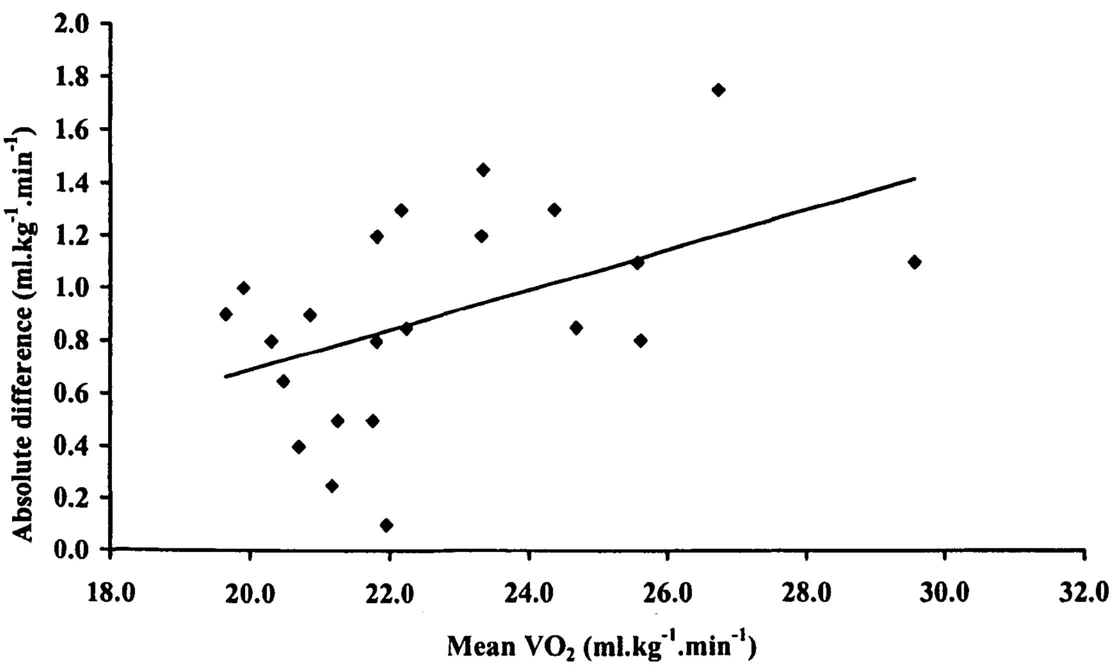


Figure 6.1. An example of measurement error proportional to the magnitude of the mean for loaded marching economy, otherwise referred to as heteroscedasticity. Here, absolute (test-retest) differences increase in proportion to the mean of test 1 and 2 ($r=0.47$, $P=0.027$). The correlation coefficient decreased to $r=0.22$, $P=0.316$ when data were logarithmically transformed, thereby decreasing the magnitude of heteroscedasticity. This indicated ratio limits of agreement should be applied.

Table 6.2. Descriptive data for sub-maximal loaded marching, maximal running, dynamic strength, isometric strength, and anthropometric measures. Data includes sample size, the mean of two observations, and checks for heteroscedasticity (i.e. absolute differences plotted against the mean of test 1 + test 2).

Measure (Units)	Sample size	Mean 1	Mean 2	Absolute Correlation (abs (diff) vs. mean 1 + 2)	Log Correlation (abs (diff) vs. mean 1 + 2)
Loaded march economy					
LM $\dot{V}O_2$ ($l \cdot min^{-1}$)	22	1.87 ± 0.22	1.84 ± 0.20	0.27	0.09
LM $\dot{V}O_2$ ($ml \cdot kg^{-1} \cdot min^{-1}$)	22	22.9 ± 2.6	22.5 ± 2.5	0.47*	0.22
Running $\dot{V}O_2max$					
RM $\dot{V}O_2max$ ($l \cdot min^{-1}$)	18	4.11 ± 0.41	4.13 ± 0.41	0.04	-0.12
RM $\dot{V}O_2max$ ($ml \cdot kg^{-1} \cdot min^{-1}$)	18	51.1 ± 4.9	51.3 ± 4.9	0.08	-0.08
Dynamic strength					
Chest Press Peak (kg)	20	72 ± 12	71 ± 12	0.14	0.07
Chest Press Average (kg)	20	66 ± 12	66 ± 11	0.29	0.24
Back Pull Peak (kg)	20	76 ± 9	76 ± 9	0.19	0.05
Back Pull Average (kg)	20	69 ± 9	69 ± 9	0.03	-0.16
Leg Press Peak (kg)	20	207 ± 56	213 ± 52	-0.22	-0.42
Leg Press Average (kg)	20	191 ± 54	200 ± 52	-0.24	-0.56*
Isometric strength					
SLS Peak (kg)	14	131.5 ± 34.5	131.5 ± 31.0	-0.06	-0.30
SLS Average (kg)	14	124.0 ± 34.0	122.0 ± 29.0	0.23	-0.12
Hand Grip Peak (kg)	19	55.0 ± 4.5	55.5 ± 5.5	-0.08	-0.21
Hand Grip Average (kg)	19	52.5 ± 5.0	53.0 ± 5.5	-0.15	-0.24
Anthropometry					
Mass (kg)	19	82.5 ± 10.9	82.3 ± 10.6	0.28	0.01
Sum of 8 SF (mm)	19	103.3 ± 34.2	102.5 ± 32.9	0.57*	0.28
Estimated FM (kg)	19	13.5 ± 4.8	13.4 ± 4.7	0.05	-0.07
Estimated FFM (kg)	19	69.0 ± 7.2	68.8 ± 7.0	0.15	0.04
Estimated Body Fat (%)	19	15.9 ± 4.3	16.0 ± 4.2	-0.13	-0.18

* statistically significant correlation coefficients (P<0.05).
LM = Loaded Marching, RM = Running Maximum.

6.3.1 *Loaded marching economy*

An example of a Bland and Altman plot is presented (Figure 6.2). Both absolute $\dot{V}O_2$ (mean ± SD) 1.87 ± 0.22 $l{\cdot}min^{-1}$, and relative $\dot{V}O_2$ 22.9 ± 2.6 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ had ratio LoA that approximated 10% to 11% for LME. Ratio LoA for absolute LME were 0.98 (×/÷ 1.09). Thus 95% of the ratios were between the systematic bias (0.98), multiplied or divided by the random error ratio (×/÷ 1.09). Consequently, the ratio LoA were (0.98 × 1.09 = **1.06**, to 0.98 ÷ 1.09 = **0.90**). Translated into absolute units of measurement absolute LME varied by +0.11 to –0.19 $l{\cdot}min^{-1}$ around the mean value of 1.87 $l{\cdot}min^{-1}$ (from **1.98 $l{\cdot}min^{-1}$ (upper LoA)** (i.e. 1.87 $l{\cdot}min^{-1}$ × 1.06) to 1.68 **$l{\cdot}min^{-1}$ (lower LoA)** (i.e. 1.87 $l{\cdot}min^{-1}$ × 0.90)). Relative $\dot{V}O_2$ had LoA of ~10% (or 0.98 (×/÷ 1.08)) (Table 6.3).

6.3.2 *Running $\dot{V}O_2max$*

The absolute $\dot{V}O_2max$ of 4.11 ± 0.41 $l{\cdot}min^{-1}$ had ratio LoA of ~7% (or 1.01 (×/÷ 1.07)), whilst the relative $\dot{V}O_2max$ of 51.1 ± 4.9 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ had LoA of ~6% (or 1.00 (×/÷ 1.06)) (Table 6.3).

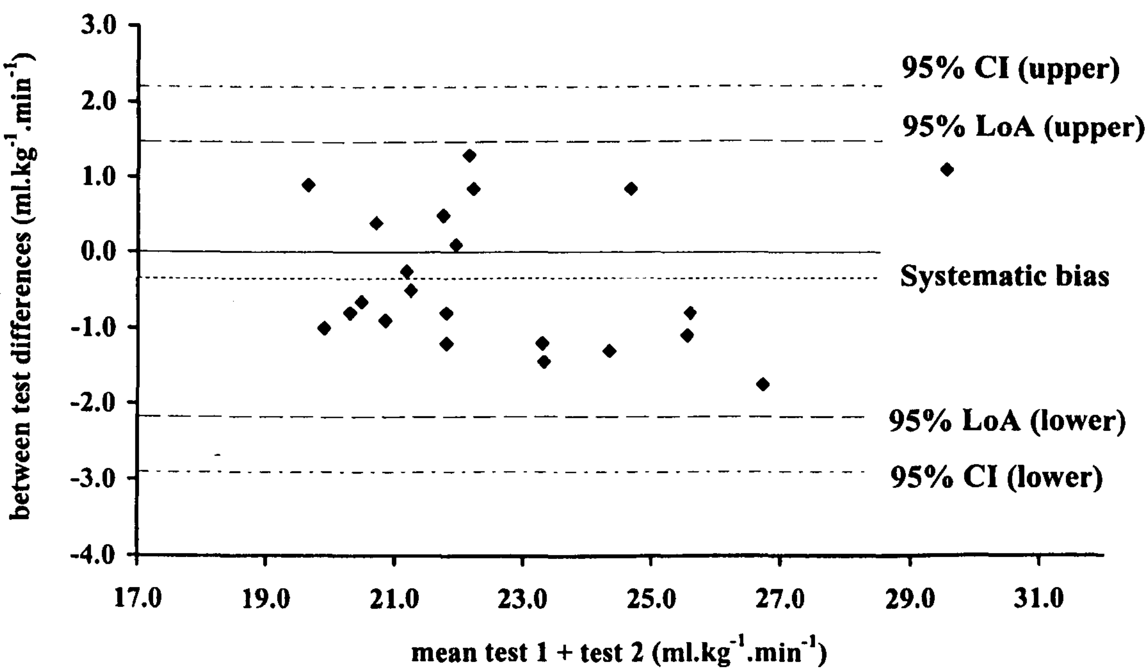


Figure 6.2. An example of a Bland and Altman 95% LoA plot. Showing systematic bias, random error (95% LoA), and 95% Confidence Intervals for loaded marching economy.

Table 6.3. The repeatability of loaded marching economy and running $\dot{V}O_2max$. Systematic bias and (95% LoA) are expressed in ratio and absolute terms. Ninety-five percent Confidence Intervals (CI) are presented for absolute 95% LoA.

	Ratio LoA	Absolute LoA	95% CI (95% LoA)	
	Bias (×/÷ 95% LoA)	Bias (± 95% LoA)	Upper ↓	Lower ↓
Loaded march economy				
LM $\dot{V}O_2$ ($l\cdot min^{-1}$)	0.98 (×/÷ 1.09)	-0.03 (± 0.15)	**0.18** – 0.06	-0.13 – **-0.25**
LM $\dot{V}O_2$ ($ml\cdot kg^{-1}\cdot min^{-1}$)	0.98 (×/÷ 1.08)	-0.4 (± 1.8)	**2.2** – 0.8	-1.5 – **-2.9**
Maximal running				
RM $\dot{V}O_2max$ ($l\cdot min^{-1}$)	1.01 (×/÷ 1.07)	0.02 (± 0.27)	**0.41** – 0.17	-0.13 – **-0.37**
RM $\dot{V}O_2max$ ($ml\cdot kg^{-1}\cdot min^{-1}$)	1.00 (×/÷ 1.06)	0.2 (± 2.9)	**4.4** – 1.8	-1.5 – **-4.0**

Values in bold (highlighted with arrows) indicate the outer limits of the 95% confidence intervals for absolute LoA.

6.3.3 *Dynamic Strength*

Systematic bias was present in one dynamic strength measure (average leg press strength) and approximated +6% (P=0.002).

Back pull strength had the smallest LoA. The average back pull strength of 69 ± 9 kg had LoA of ~8% (or 1.00 (×/÷ 1.08)), and the peak back pull strength of 76 ± 9 kg had LoA of ~9% (or 1.00 (×/÷ 1.09)). Marginally worse ratio LoA were observed for chest strength. The average chest press strength of 66 ± 12 kg had LoA of ~12% (or 1.01 (×/÷ 1.11)), and the peak chest strength of 72 ± 12 kg had LoA of ~10% (or 0.99 (×/÷ 1.09)). Leg press strength had excessive LoA. The average leg press strength of 191 ± 54 kg had LoA of ~20% (or 1.06 (×/÷ 1.14)), and the peak leg press strength of 207 ± 56 kg had LoA of ~ 19% (or 1.03 (×/÷ 1.16)) (Table 6.4).

6.3.4 *Isometric Strength*

Static Lift Strength (SLS) had the largest LoA of the isometric strength measures. The average SLS of 124.0 ± 34.0 kg had LoA of ~17% (or 0.99 (×/÷ 1.16)), and the peak SLS of 131.5 ± 34.5 kg had LoA of ~15% (or 1.01 (×/÷ 1.14)). The average Hand Grip (HG) strength of 52.5 ± 5.0 kg had LoA of ~11% (or 1.01 (×/÷ 1.10)), and the peak HG strength of 55.0 ± 4.5 kg had LoA of ~12% (or 1.00 (×/÷ 1.12)) (Table 6.4).

6.3.5 *Anthropometry*

The body mass of 82.5 ± 10.9 kg had LoA of ~2% (or 1.00 (×/÷ 1.02)). The four derived skinfold (SF) measures of, sum of 8 SF (103.3 ± 34.2 mm), estimated Fat Mass (FM) (13.5 ± 4.8 kg), estimated Fat Free Mass (FFM) (69.0 ± 7.2 kg), and estimated body fat percentage (15.9 ± 4.3%) had LoA of between 2% to 8%. The sum of eight SF had LoA of ~6% (or 1.00 (×/÷ 1.06)), estimated FM ~8% (or 1.00 (×/÷ 1.08)), estimated body fat percentage ~7% (or 1.00 (×/÷ 1.07), and estimated FFM had the smallest LoA of ~2% (or 1.00 (×/÷ 1.02)) (Table 6.5).

Table 6.4. The repeatability of dynamic and isometric strength measures. Systematic bias and (95% LoA) are expressed in ratio and absolute terms. Ninety-five percent Confidence Intervals (CI) are presented for absolute 95% LoA.

	Ratio LoA	**Absolute LoA**	**Absolute 95% CI (95% LoA)**	
	Bias (x/÷ 95% LoA)	**Bias (± 95% LoA)**	**Upper**	**Lower**
Dynamic strength				
Chest Press Peak (kg)	0.99 (x/÷ 1.09)	-1 (± 6)	**8** – 3	-4 – **-9**
Chest Press Average (kg)	1.01 (x/÷ 1.11)	0 (± 7)	**10** – 4	-4 – **-10**
Back Pull Peak (kg)	1.00 (x/÷ 1.09)	0 (± 7)	**9** – 4	-4 – **-9**
Back Pull Average (kg)	1.00 (x/÷ 1.08)	0 (± 5)	**7** – 3	-3 – **-7**
Leg Press Peak (kg)	1.03 (x/÷ 1.16)	6 (± 24)	**40** – 20	-8 – **-28**
Leg Press Average (kg)	1.06**(x/÷1.14)	9** (± 20)	**38** – 21	-3 – **-19**
Isometric strength				
SLS Peak (kg)	1.01 (x/÷ 1.14)	0.0 (± 17.0)	**26** – 8.5	-8.5 – **-26.0**
SLS Average (kg)	0.99 (x/÷ 1.16)	-2.0 (± 18.0)	**25.5** – 7	-11.0 – **-29.5**
Hand Grip Peak (kg)	1.00 (x/÷ 1.12)	0.5 (± 6.0)	**9.0** – 4.0	-3.0 – **-8.5**
Hand Grip Average (kg)	1.01 (x/÷ 1.10)	0.5 (± 4.5)	**7.5** – 3.5	-2.0 – **-6.0**

** statistically significant systematic bias ($P<0.01$).

Values in bold (highlighted with arrows) indicate the outer limits of the 95% confidence intervals for absolute LoA.

Table 6.5. The repeatability of body mass, the sum of 8 skinfolds, estimated fat mass, estimated fat free mass, and estimated body fat percentage. Systematic bias and (95% LoA) are expressed in ratio and absolute terms. Ninety-five percent Confidence Intervals (CI) are presented for absolute 95% LoA.

	Ratio LoA	Absolute LoA	Absolute 95% CI (95% LoA)	
	Bias (×/÷ 95% LoA)	Bias (± 95% LoA)	Upper	Lower
Mass (kg)	1.00 (×/÷ 1.02)	-0.2 (± 1.8)	**2.4** – 0.8	-1.3 – **-2.8**
Sum of 8 SF (mm)	1.00 (×/÷ 1.06)	0.8 (± 7.0)	**10.8** – 4.8	-3.2 – **-9.2**
Estimated FM (kg)	1.00 (×/÷ 1.08)	0.0 (± 1.0)	**1.5** – 0.6	-0.6 – **-1.4**
Estimated FFM (kg)	1.00 (×/÷ 1.02)	0.2 (± 1.4)	**2.2** – 1.0	-0.6 – **-1.8**
Estimated Body Fat (%)	1.00 (×/÷ 1.07)	0.0 (± 1.0)	**1.5** – 0.6	-0.5 – **-1.4**

Values in bold (highlighted with arrows) indicate the outer limits of the 95% confidence intervals for absolute LoA.

6.4 Discussion

6.4.1 Main findings

The repeatability of accepted and potential determinants of loaded marching performance was mixed. The repeatability of loaded marching economy, running $\dot{V}O_2max$, dynamic upper body strength, and anthropometric measures was smaller than the large repeatability observed in dynamic leg strength, and isometric strength.

6.4.2 Loaded marching economy

The repeatability of LME approximated 10 to 11%, which was comparable to Chapter 4. This confirmed that the new Douglas Bag gas analysis system that was built for this study, and subsequent application to the field (Chapters 7, 8, and 9) was fit for purpose. Data that quantified the repeatability of LME is absent from the literature, which is no surprise given that LME's role in determining LMP has been ignored (Rayson *et al.*, 1993, 2000, 2002a; Pandorf *et al.*, 2002; Simpson *et al.*, 2006; Williams and Rayson, 2006; Griggs *et al.*, 2008). Consequently, data on the repeatability of running economy (RE) is the closest exercise modality to that of LM, given that it is sub-maximal, rhythmical, cyclical and continuous exercise. Thus, the repeatability of running economy will be discussed to provide a context to which LME repeatability can be compared.

Over a wide range of running speeds (9 $km \cdot h^{-1}$ to 18 $km \cdot h^{-1}$), and at absolute (Armstrong and Costill, 1985; Morgan *et al.*, 1991; 1994; Williams *et al.*, 1991), and relative exercise intensities (Pereira *et al.*, 1994; Pereira and Freedson, 1997) the repeatability of RE was ~1 % to 5 %, which was better than the present study. However, the coefficient of variation (CV) was the statistic reported, which accounts for 68% of the differences in the sample (Atkinson and Nevill, 1998). Thus, if 95% of the differences within the sample were calculated (by multiplying the CV by the *z*-score of 1.96) the CVs from these studies would have been ~2% to 10%, which (at the upper end) approximates the repeatability LME in the present study. A similar criticism can be attributed to the small (1.6 $ml \cdot kg^{-1} \cdot min^{-1}$) Typical Error (TE) observed during repeated running bouts (Saunders *et al.*, 2004). As TE is calculated by the standard deviation of the difference divided by the square root of 2 ($SD_{diff}/\sqrt{2}$), the repeatability of this statistic is always systematically smaller (by a factor of 2.77) compared to 95% LoA (Hopkins, 2000). Thus, multiplying the data of Saunders *et al.*, (2004) by 2.77 (to make these directly comparable to 95% LoA) indicates the repeatability of LME from the present study is better than some observations of RE.

Tight repeatability of LME may be required when tracking LME changes over a physical training programme, as the one study to investigate this question observed small improvements of 5.9% (~1 $ml \cdot kg^{-1} \cdot min^{-1}$) in response to 8 weeks of military style training (Gutekunst *et al.*, 2006). This change was approximately half the magnitude of LME repeatability observed in the present study. Small adaptations in exercise economy, in response to a physical training programme are nothing new. Running economy improved by 2.8% to 7.2% (1.4 to 3.1 $ml \cdot kg^{-1} \cdot min^{-1}$) to endurance training (Billat *et al.*, 1999; Franch *et al.*, 1998; Jones *et al.*, 1999; Sjödin *et al.*, 1982), and 4.1% to 6.7% (1.8 $ml \cdot kg^{-1} \cdot min^{-1}$) to concurrent strength and endurance training programmes (Johnston *et al.*, 1997; Spurs *et al.*,2003). Whilst the day-to-day variability of LME is comparable to previous running economy observations, the magnitude of LME's adaptation to a physical training programme is small (Gutekunst *et al.*, 2006), and within the repeatability of LME observed in the present study, which will have implications on the sample size estimations in subsequent studies.

6.4.3 Running $\dot{V}O_2max$

$\dot{V}O_2$max repeatability was ~6% to 8%, which was better than that reported in; running $\dot{V}O_2$max (i.e. ± 11.2%) (Katch *et al.*, 1982), cycle ergometry (i.e. -3.5 (± 4.0) $ml \cdot kg^{-1} \cdot min^{-1}$) (Bingisser *et al.*, 1997), and arm crank ergometry (i.e. 0.00 (± 0.50) $l \cdot min^{-1}$) (Price and Campbell, 1997). One study that can be directly compared to the present study, due to the similar; subjects (males, age 38 ± 7.1 years vs. 30 ± 6.6 years); $\dot{V}O_2$max (3.87 ± 0.50 $l \cdot min^{-1}$ vs. 4.11 ± 0.41 $l \cdot min^{-1}$); sample size (n = 22 vs. 18); $\dot{V}O_2$max protocol (continuous incremental running); and statistical analysis (95% LoA) also reported larger 95% LoA (~10%) in comparison to the present study (Midgley *et al.*, 2007a). The running $\dot{V}O_2$max repeatability in the present study is encouraging as indices of running $\dot{V}O_2$max (both direct measurements and indirect estimations) are consistently the most cited determinants of LMP (Dziados *et al.*, 1987; Knapik *et al.*, 1990; Rayson *et al.*, 1993, 2000, 2002a, 2002b, Frykman and Harman, 1995; Harman and Frykman, 1995; Pandorf *et al.*, 2002, Williams and Rayson, 2006). This may mean that the ability to detect a $\dot{V}O_2$max training response in order to explain concomitant improvements observed in LMP in subsequent studies may be enhanced, assuming the same sample size is used for all measures.

The direct measurement of $\dot{V}O_2$max increased 1.3% to 3.5% (0.7 ml·kg^{-1}·min^{-1} to 1.9 ml·kg^{-1}·min^{-1}) in response to military training (Daniels *et al.,* 1979; Gordon *et al.,* 1986; Drystad, *et al.,* 2006). This small increase in $\dot{V}O_2$max was smaller than the repeatability observed in the present study (6% to 8%), and was explained by an inverse relationship between baseline $\dot{V}O_2$max and the magnitude of the $\dot{V}O_2$max training response. Larger training related changes of 5.7% to 16.2% (or 2.9 ml·kg^{-1}·min^{-1} to 7.8 ml·kg^{-1}·min^{-1}) have been reported in military training studies that have estimated $\dot{V}O_2$max responses to training (Stacy *et al.,* 1982; Williams *et al.,* 1999; 2002; Williams, 2005), which did approximate the repeatability of this measure in the present study. Other military training studies have reported even greater improvements in $\dot{V}O_2$max of 12.5% to 14.0% (Lim and Lee, 1994; Harman *et al.,* 1998). However, the magnitude of these responses could be a function of the low initial fitness of these female subjects and obese males. The repeatability of running $\dot{V}O_2$max in the present study is better than that previously reported. This is encouraging given that the magnitude of the $\dot{V}O_2$max adaptation during military training has the potential to be small.

6.4.4 Body mass

Although, for heavier loads absolute $\dot{V}O_2$max is regarded as the most important $\dot{V}O_2$max parameter to determine LMP (Haisman, 1988), relative $\dot{V}O_2$max (ml·kg^{-1}·min^{-1}) also underpins LMP (Rayson *et al.,* 2002a; 2002b). However, it should be noted that fluctuations in body mass cause variability in the calculation of $\dot{V}O_2$ when body mass is scaled using an exponent of 1 (e.g. ml·**kg**$^{-1}$·min^{-1}). Thus, $\dot{V}O_2$ will artificially increase with decreases in body mass, and decrease with increases in body mass (Berg, 2003). Increases in body mass of 1.0 ± 4.7 kg occur during British Army infantry training (Carter *et al.,* 2006), which will further increase the variability of relative $\dot{V}O_2$ measurement. In the present study, day-to-day variability of body mass was, -0.2 (± 1.8) kg. Therefore, body mass fluctuated between + 1.6 to – 2.0 kg. For a 70 kg person this would cause a LME of 22.9 ± 2.6 ml·kg^{-1}·min^{-1} to vary by ± 1.4 ml·kg^{-1}·min^{-1}, and a $\dot{V}O_2$max of 51.1 ± 4.9 ml·kg^{-}

$^{1}\cdot min^{-1}$ to vary by ± 3.0 $ml\cdot kg^{-1}\cdot min^{-1}$. It is important to be aware of how both daily variation, and longer term training adaptations in body mass cause artefact changes in relative $\dot{V}O_2$ calculation. Allometric scaling is one approach that can correct for the artefact of body mass change in relative $\dot{V}O_2$ calculation (Winter and Nevill, 2001). However, in the context of load-carriage where the combination of oxygen uptake and body size (i.e. absolute $\dot{V}O_2$) is key to determining LMP (Dziados *et al.*, 1987; Knapik *et al.*, 1990; Rayson *et al.*, 1993, 2000, 2002a, 2002b; Harman and Frykman, 1995; Pandorf *et al.*, 2002) it is appropriate that the composite measure of absolute $\dot{V}O_2$ (i.e. $\dot{V}O_2$ $l\cdot min^{-1}$) is used during load-carriage research.

6.4.5 Dynamic strength

Weak to strong correlations have been reported between dynamic strength and LMP (Dziados *et al.*, 1987; Mello *et al.*, 1988; Harman and Frykman, 1995). The repeatability of dynamic upper body strength (i.e. chest press and back pull) approximated 8% to 12%, which was better than the ~19% to 20% observed for dynamic leg press strength. The repeatability of the dynamometer used in the present study is yet to be quantified in the literature.

A systematic bias of +6% was observed for average leg press strength, which indicated the presence of a learning effect. Further support for this apparent learning effect is presented elsewhere (section 3.4.2.3). Average leg press strength is therefore not a stable measure (Bland and Altman, 1986), hence increases over a training programme should be interpreted tentatively, or compared to control group changes.

Concurrent resistance and endurance training models have improved LMP, whereas endurance training, and strength training alone has not (Kraemer *et al.*, 1987; 2001; 2004). The efficacy of this physical training model on LMP improvement is exacerbated further when applied to civilian females (Harman *et al.*, 1998). Dynamic occupational measures of strength (i.e. a military material handling strength test) have (Brock and Legg, 1997; Harwood *et al.*, 1999) and have not increased (Jetté *et al.*, 1989; Williams *et al.*, 1999) during military training. These mixed reports are not surprising given that little or no strength based resistance training was performed in these military training programmes.

The Combat Infantryman's Course at the Infantry Training Centre (Catterick) also has an absence of resistance based strength training within the physical training programme. Therefore detecting increases in strength that may influence LMP gains in subsequent studies may be beyond the sensitivity of this dynamic field based measure of strength (assuming a realistic sample size). However, larger and more consistent increases in material handling strength (12% to 14%) (Williams *et al.,* 2002) and 16% (Knapik and Gerber, 1996), as well as 1RM (7.7%) (Mouneimneh *et al.,* 2007) and 39.4% (Knapik and Gerber, 1996) occur if structured resistance training is embedded into a military training programme, which may be a detectable training adaptation.

6.4.6 Isometric Strength

The repeatability of SLS approximated 15% to 17%, which was worse than the 11% to 12% observed for HG strength. This was at best equivalent, but generally worse than dynamic strength indices (section 6.4.5). Whilst isometric indices of strength have consistently been cited as determinants of LMP (Rayson *et al.,* 2000, 2002a, 2002b; Williams and Rayson, 2006), isometric dynamometry is reputedly unable to discriminate between populations, and lacks the sensitivity to track training responses (Wilson, 2000). This should be taken into account in subsequent studies.

The repeatability of isometric strength measures in the literature have mostly applied intra class correlation (Wilson and Murphy, 1996), which should not constitute the sole repeatability statistic used (Atkinson and Nevill, 1998). More appropriate statistics of absolute reliability, reported CVs of around 12 % (Hattori *et al.,*1998), 4% (Thorstensson *et al.,* 1977), and 18-20% (Wilson and Murphy, 1995). Furthermore, the TE for an isometric squat test (using similar muscle groups to SLS) was 69 N (Blazevich *et al.,* 2002). When multiplied by 2.77 to make this statistic equivalent to 95% LoA (section 6.4.2) the TE of 69 N became 191.1 N or 19.5 kg, which approximates the SLS repeatability in the present study.

Static Lift Strength has been tracked over military training (Brock and Legg, 1997; Harwood *et al.,* 1999; Williams *et al.,* 1999, 2002). Just one study reported a significant improvement (5%) (Harwood *et al.,* 1999), whilst others reported a trend for SLS to decrease by 4.9% (Williams *et al.,* 2002). The insensitivity of SLS to a dynamic physical training programme that included structured resistance training (Williams *et al.*,2002) is no surprise considering a 200% increase in the size of weights lifted was accompanied by a 15% to 20% increase in maximal isometric force development over a dynamic strength training programme (Rutherford and Jones, 1986).

The repeatability of HG strength approximated 11% to 12%, which is better than the 95% LoA of 21% reported in 171 healthy hospital workers (Haidar *et al.*, 2004). Six weeks of British Army training increased female HG strength by 15.9% (Brock and Legg, 1997), but not in the Canadian Army (Jetté *et al.*, 1989). Regardless, the repeatability and training response of HG strength to military training seems to lack relevance, as weak correlations have been reported between HG strength and LMP (Knapik *et al.*, 1990).

6.4.7 Anthropometry

The repeatability of estimated FFM (1%), the sum of eight SF (6%), estimated FM (8%), and percent body fat (9%) was quantified. As the repeatability of FM, FFM, and percentage body fat derive from raw SF measurements, it is logical to appraise the repeatability of the sum of 8 SF. Verifying the repeatability of this measure, 95% ratio LoA for the sum of seven SF were similar to the sum of eight SF measures in the present study (~4% vs. 6%) (Woolford and Gore, 2004).

In addition to the measure of maximal aerobic fitness, anthropometric measures of, percent fat, FFM and body mass were all included in a model of LMP, which explained moderate levels of variance in LMP (Adjusted R^2=0.57) (Williams and Rayson *et al.*, 2006). Anthropometric measures (i.e. body structure), and strength measures (i.e. body function) typically accompany high $\dot{V}O_2max$ scores in LMP models (Haisman, 1988; Rayson *et al.*, 2000). Absolute loads carried in most LMP tests (e.g. 25kg) means that larger individuals carry a smaller percentage of their body mass. Thus the combined effect of body size and a good oxygen transport and delivery system (i.e. absolute $\dot{V}O_2max$) justifies the role of anthropometry in determining LMP (all be it secondary to $\dot{V}O_2max$).

Theoretically, of the four anthropometric measures, FFM (the propulsive compartment of body composition) should influence LMP greatest, although this is far from unequivocal.

The repeatability of FFM in the present study was within the magnitude of the observed FFM response during British Army recruit training (Williams *et al.*, 1999; 2002; Wilkinson *et al.*, 2003; Carter *et al.*, 2006).

In conclusion, loaded marching economy varied by 11%, which bared some resemblance to the running economy literature. Running $\dot{V}O_2max$ varied by 8%, which was better than previously reported. However, the measurement of $\dot{V}O_2$ comes with the caveat that changes in body mass further increase variability in $\dot{V}O_2$ should $\dot{V}O_2$ be expressed relative to body mass. Dynamic leg strength (20%) and isometric strength (17%) had poor repeatability compared to dynamic upper body strength (12%). Anthropometric measures varied by up to 9%. A lack of consistency in the statistical analysis of previous repeatability studies made comparisons to the present study challenging. The repeatability of these accepted and potential determinants of loaded marching performance will influence subsequent sample size calculations, which is important for ensuring adequate statistical power (i.e. ≥0.8) can be generated to minimise the risk of Type II errors during inferential analyses throughout the remainder of this thesis.

CHAPTER 7

A MATHEMATICAL MODEL TO PREDICT 2.4 KM LOADED MARCH PERFORMANCE

7.1. Introduction

Mathematical models can predict loaded marching performance (LMP), and classically have identified general measures of aerobic fitness, dynamic and isometric muscular strength & endurance, and anthropometry as determinants of performance (Rayson *et al.*, 2000; Pandorf *et al.*, 2002; Rayson *et al.*, 2002a; Rayson *et al.*, 2002b; Wilkinson *et al.*, 2003; Williams and Rayson, 2006). Principally, the combination of maximal oxygen uptake and body size makes absolute $\dot{V}O_2max$ ($\dot{V}O_2$ l·min^{-1}) an important determinant of LMP for heavy loads (Rayson *et al.*, 2000). Indices of aerobic fitness have therefore consistently underpinned LMP. However these measures of aerobic fitness only encompassed running performance (e.g. 2.4 km run), or indirect running $\dot{V}O_2max$ measures (e.g. multi stage fitness test). As the determinants of endurance performance include, $\dot{V}O_2max$, $\dot{V}O_2$ kinetics, maximal sustainable % of $\dot{V}O_2max$, and exercise economy (Fallowfield and Wilkinson, 1999), the existing LMP models have scientifically failed to include potentially key maximal and sub-maximal measures of endurance performance in LMP models.

Specific measures of loaded marching (LM) are in general absent from LMP models, which is surprising given the importance of modelling specificity in the exercise modality of running (Joyner, 1991). Just two studies have included LM measures within the LMP models (Simpson *et al.*, 2006; Griggs *et al.*, 2008). Simpson *et al.*, (2006) reported both LM $\dot{V}O_2max$ and 3.2 km LMP explain 48% and 70% of the variance in 3.2 km and 29 km LMP, respectively. Although similar amounts of explained variance were reported in other LMP models that included general, and not LM specific measures of fitness (Rayson *et al.*, 2000; Williams and Rayson, 2006). However, Griggs *et al.*, (2008) highlighted the importance of LM specificity in LM modelling as a 23% increase in explained variance (74% vs. 51%), and a 50% reduction in prediction error (124 s vs. 224 s) occurred when

2.4 km LMP (compared to 2.4 km run time and body mass) was included in a model to predict 12.8 km LMP.

In addition to the paucity of studies that have included maximal effort LM measures in LMP models (Simpson *et al.*, 2006; Griggs *et al.*, 2008), sub-maximal LM measures have been completely overlooked in LMP models. 'Strong' bivariate correlations were reported between sub-maximal running blood lactate and ventilatory indices, and 12.8 km LMP in a small sample (n=13) of elite soldiers (Graham *et al.*, 2007). The sub-maximal oxygen demand during a LM task (i.e. loaded marching economy (LME)) is a form of exercise economy (Jones and Carter, 2000). Theoretically LME may be an important predictor of LMP as exercise economy is important in the performance of the exercise modality of running (Joyner, 1991; Brandon, 1995). Running economy can compensate for a comparatively low $\dot{V}O_2max$, in order to produce comparable 10 km performances (Weston *et al.*, 2000), and can explain 64% of the variation in 10 km running performance (Conley and Krahenbuhl, 1980). Consequently, LME may be an important determinant of LMP, which remains to be determined.

The purpose of this study was to develop a mathematical model to elucidate the determinants of loaded marching performance. Subsequently, this model will identify the most influential components of fitness that should be emphasised to improve physical training prescription for loaded marching performance in British Army infantry recruits.

7.2 Methods

7.2.1 Subjects

Forty-four new entry recruits and 1 'back-squadded' recruit (*n*=45) formed Rifles 3 Platoon and undertook the 24-week Combat Infantryman's Course (CIC) at the Infantry Training Centre (Catterick) (ITC(C)). Ethical clearance was provided by the Ministry of Defence (Naval) Personnel Research Ethics Committee (MOD(N) PREC). A full Initial Medical Examination (IME) was conducted prior to participation in the study, and a resting 12-lead electrocardiogram (ECG) was obtained to screen for cardiac abnormalities. Two recruits failed the IME and were excluded from the trial, the remaining 43 recruits presented normal ECG traces and were deemed fit to participate in the study (Table 7.1).

Table 7.1. The 43 subject characteristics (mean ± SD).

Age (years)	Stature (cm)	Body mass (kg)	Body fat (%)	$\dot{V}O_2max$ (ml·kg^{-1}·min^{-1})	2.4 km run time (min:s)
19 ± 2	174.9 ± 5.8	69.0 ± 8.5	14.2 ± 4.2	53.1 ± 4.2	10:17 ± 0:36

7.2.2 Experimental design

A test battery (section 8.2.3) collected performance data at the approximate beginning, middle and end of training. These time points were coincident with weeks 1/2 (T1), 11/12 (T2), and 20/21 (T3) of training. Here, 2.4 km LMP (dependent variable (DV)) was predicted from the remaining test battery measures (independent variables (IV)). The experimental design was affected due to a decreasing subject data pool during training, as a consequence of training attrition (section 8.3.1), absence, sicknesses, and missing data points. The available sample size approximated n=29 (T1), n=17 (T2), n=13 (T3). The small sample sizes at T2 and T3 meant that, an individual case could exert a disproportionately large influence over the strength of calculated relationships (Fallowfield *et al.*, 2005), the generalisability of findings to other samples would be poor (Pallant, 2007), and spuriously high coefficient of determinations may occur due to an increase in the IV to subject ratio (Thomas and Nelson, 1996). Consequently, the mathematical modelling of LMP at T2 and T3, as well as modelling of delta changes in LMP between T1 and T3 was deemed inappropriate. Thus, a LMP model was produced from T1 data. A schematic representation of the study design is presented (Figure 7.1).

7.2.3 Test battery

Independent variables that were entered into the model included LME, 2.4 km running performance, running $\dot{V}O_2max$, dynamic strength, isometric strength, 2 minutes press-ups, 2 minutes sit-ups, body mass, body stature, 6-site circumferential girths and 8-site skinfolds. The DV (LMP) consisted of a 2.4 km 'best effort' LM test. As over 60% of subjects were loaded march naïve, the first 1.2 km of the LMP test was squadded (performed as a group) and the second 1.2 km was performed as an individual best effort (section 8.2.3).

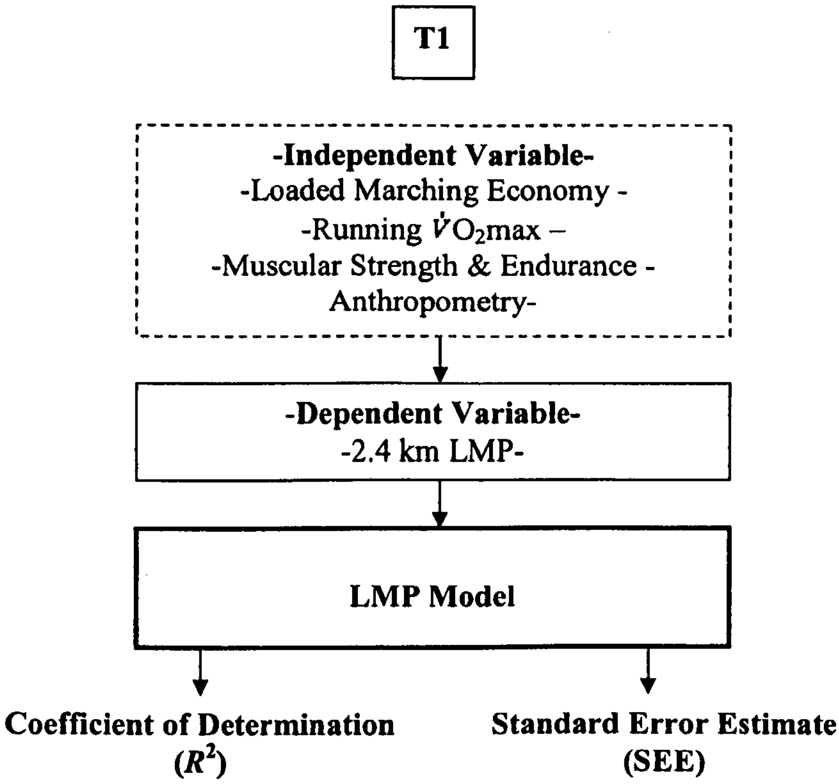


Figure 7.1. A schematic representation of the study design.

7.2.4 Statistical analysis

Least squares multiple linear regression generated a mathematical model from the test battery (IVs) to explain variance in LMP (DV), at T1. Data were screened to reduce the initial number of IV (44) included in the regression analysis. Only significant bivariate Pearson Product moment correlations (i.e. IV vs. LMP) were included ($P<0.05$). A total of 29 cases at T1 meant up to 3 IV could be entered at anyone time into the regression model by the *forced entry* method (Thomas and Nelson, 1996). Independent variables were entered into the model on a trial and error basis, which was underpinned by previous empirical observations and/or physiological rationale. The coefficient of multiple correlation (R), the coefficient of determination (R^2), the adjusted coefficient of determination (*Adj* R^2), and the standard error of the estimate (SEE) explained the fit of the model. The criteria for selecting the best model was the minimisation of the SEE, and maximisation of R^2 (Winter *et al.*, 2001). The unstandardised coefficients in the LMP model calculated the theoretical importance of a change (i.e. training response) in each IV on subsequent changes to LMP. Five assumptions were checked in the LMP model. The standardised residuals plotted against standardised predicted values revealed assumptions were met for, normality (bell shaped histogram and straight line normal probability plot), linearity (no curve in data), homoscedasticity (random and even distribution of data) (Field, 2001), and outliers (standardised residual < ± 3.3) or (Cook's distance value < 1.0) (Tabachnick and Fidell, 2001). Multicollinearity was also met (Tolerance statistic close to 1.0) (Ntoumanis, 2001).

7.3 Results

7.3.1 Bivariate correlates of LMP

A total of 19 out of 44 variables significantly correlated with 2.4 km LMP (P<0.05), and were deemed practically and theoretically relevant. With the exception of one measure (LM fractional utilisation of $\dot{V}O_2max$ r=0.70), the strength of the correlations were moderate and ranged from r=-0.35 to 0.65. These included, sub-maximal LM (r=0.70 to r=0.39); maximal running (r=0.42 to r=0.38); muscular strength and endurance (r=-0.48 to r=-0.35); and anthropometric measures (r=-0.41 to r=0.38) (Table 7.2).

7.3.2 LMP multiple linear regression model

From the initial 43 subjects, 29 complete data sets were available to develop the regression model. The multivariate model that best predicted 20 kg, 2.4 km LMP included; LME (t=4.11, P<0.01), 2.4 km run time (RT) (t=2.48, P=0.020), and peak static lift strength (SLS) (t=-2.21, P=0.036), which was mathematically explained as, y = mx1 + mx2 – mx3 + c (Equation 7.1), and graphically represented with a linear trend line (Figure 7.2). The model fit was R=+0.81, R^2=0.65, *Adjusted* R^2=0.61, SEE ± 51 s, $F_{(3,25)}$=15.41, P<0.01. At the 95% level of confidence the equation had a SEE of ± 100 s (Table 7.3). Squared part correlations revealed LME explained the most (22%), RT the second most (11%), and SLS the least (6%) of the variance in LMP. The sum of the part squared correlations indicated 39% of the explained variance in LMP was independently contributed by each IV, and the remaining variance (26%) was shared between the IVs.

$$LMP(s) = (13.13 \cdot LME\ ml{\cdot}kg^{-1}{\cdot}min^{-1}) + (0.81 \cdot RT\ s) - (0.92 \cdot SLS\ kg) + 90.91$$

(Equation 7.1)

LMP = Loaded March Performance; LME = Loaded March Economy; RT = 2.4 km Run Time; SLS = Static Lift Strength.

Table 7.2. Significant correlations with LMP at T1.

Test measure	Pearson's r	
Loaded march fractional utilisation (% $\dot{V}O_2$max)	0.70	Sub-maximal loaded marching
Loaded march economy $\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	0.65	
Loaded march RER	0.46	
Loaded march HR (beats·min^{-1})	0.46	
Loaded march [BLa] (mmol·l^{-1})	0.39	
Loaded march $\dot{V}_E$ (l·min^{-1})	0.45	
2.4 km run time (s)	0.42	Maximal running
Running $\dot{V}O_2$max (l·min^{-1})	-0.56	
Peak dynamic chest strength (kg)	-0.35	Muscular strength & endurance
Peak dynamic back strength (kg)	-0.40	
Peak isometric HG strength (kg)	-0.40	
Peak isometric SLS strength (kg)	-0.48	
Sit ups (N°)	-0.35	
Body mass (kg)	-0.39	Anthropometry
Chest girth (cm)	-0.41	
Relative load carried (% BM)	0.38	

All correlations were significant ($P<0.05$). HG = Hand Grip; SLS = Static Lift Strength. Note: Average Chest, Back, HG, and SLS strength were also significantly correlated, but to a lesser degree than peak strength measures.

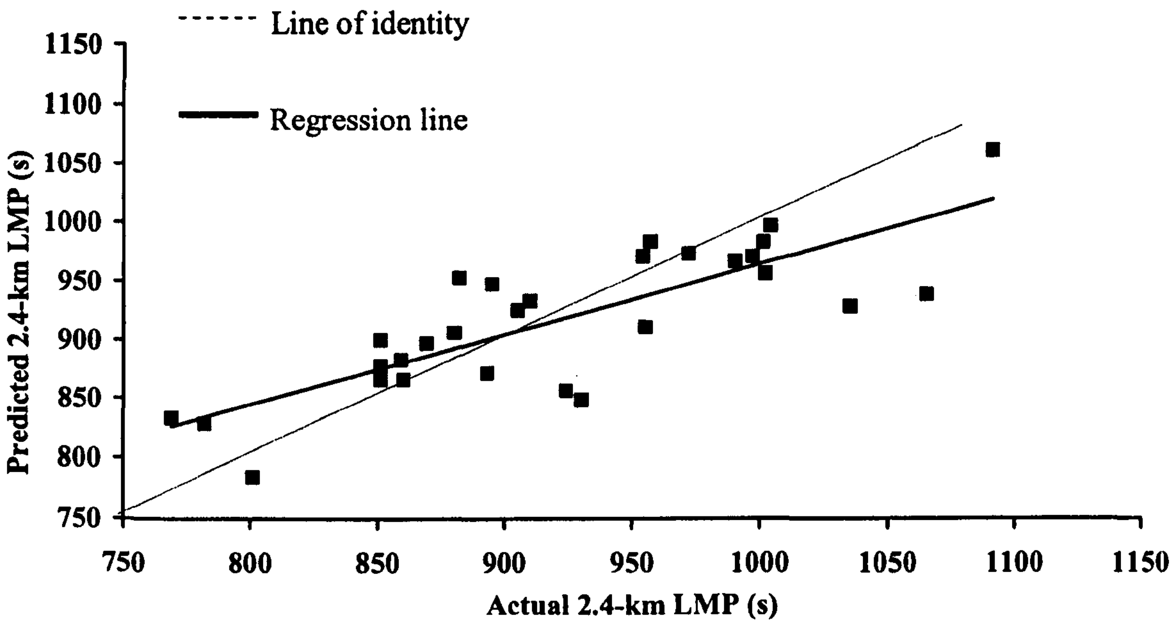


Figure 7.2. Predicted *versus* actual LMP at T1.

Table 7.3. The fit of the LMP model at T1.

R	R^2	*Adjusted* R^2	SEE (s)	95% SEE (s)	n
0.81	0.65	0.61	51	100	29

7.4 Discussion

7.4.1 Main findings

The main findings of this study were that sub-maximal loaded marching, maximal running, dynamic & isometric strength, and anthropometric variables had moderate correlations with loaded marching performance. One measure (loaded march fractional utilisation of $\dot{V}O_2max$) was strongly associated with loaded marching performance. The three independent variables of loaded marching economy, 2.4 km run time, and static lift strength best predicted loaded marching performance. Loaded marching economy is the first sub-maximal measure of loaded marching to be cited as a determinant of loaded marching performance. Unexpectedly, absolute $\dot{V}O_2max$ was not included in the loaded march model, despite having a stronger bivariate relationship with loaded marching performance than 2.4 km run time. Furthermore, despite having the strongest bivariate relationship, loaded march fractional utilisation did not produce the best loaded marching performance model.

7.4.2 Correlates of LMP

Five sub-maximal LM measures significantly correlated with LMP. The LM fractional utilisation, and LME had the strongest relationships that were both positive in direction (Table 7.2). Thus the lower the oxygen demand of a simulated LM task, the better the LMP. Similar relationships (but not as strong) were observed for minute ventilation ($\dot{V}_E$), heart rate (HR), RER, and whole blood lactate concentration ([BLa]). Correlations between physiological LM measures and LMP have been reported in just one study (Simpson *et al.* 2006). Here test duration of a LM $\dot{V}O_2$max test correlated with 3.2 km (r=-0.57) and 29 km (r=-0.66) LMP, however LM $\dot{V}O_2$max did not. Strong correlations between various sub-maximal blood lactate concentrations, the ventilatory equivalent of oxygen, and 12.8 km LMP were reported, however these sub-maximal measures were taken from an incremental running (not LM) test (Graham *et al.*, 2007). Others observed strong correlations (r=-0.76) between the fractional utilisation of LM (i.e. %$\dot{V}O_2$max) and absolute $\dot{V}O_2$max (not LMP) (Lyons *et al.*, 2005). This indirectly supported observations in the LMP literature (Rayson *et al.*, 2000), highlighting that individuals with a lower fractional utilisation of $\dot{V}O_2$max in a 40 kg LM task were predisposed to a higher absolute $\dot{V}O_2$max. However, the strong correlation reported may have been spurious due to the existence of a common divisor in the bivariate correlation (i.e. $\dot{V}O_2$max and %$\dot{V}O_2$max) (Atkinson *et al.*, 2003). The present study is the first to identify the relationship between sub-maximal (metabolic) measures of LM and LMP, and to identify LME as a determinant of LMP.

Two measures of running ability (i.e. absolute $\dot{V}O_2$max and 2.4 km RT) correlated with LMP (Table 7.2). This indicated that a higher absolute $\dot{V}O_2$max, and faster 2.4 km RT were associated with better LMP. A similar observation (but stronger correlations) was reported between absolute $\dot{V}O_2$max (r=-0.61 to r=-0.70), 3.2 km run time (r=0.67 to r=0.80) and 3.2 km LMP (Pandorf *et al.*, 2002). It is generally agreed that absolute (not relative) $\dot{V}O_2$max best explains the relationship between maximal oxygen uptake and LMP

(Haisman, 1988; Rayson *et al.*, 2000), however LM specific tests of absolute $\dot{V}O_2max$ do not support this (Simpson *et al.*, 2006). A high absolute $\dot{V}O_2max$ is coincident with a large, body / fat free mass (Haisman, 1988), and anthropometric dimension (Frykman and Harman, 1995). Thus when carrying absolute loads (e.g. 25 kg) larger individuals carry a smaller percentage of body mass, which would be advantageous in a LMP task. In support of this notion, strong (Harman and Frykman, 1995), moderate (Pandorf *et al.*, 2002), or weak (Knapik *et al.*, 1990) absolute $\dot{V}O_2max$ correlations with LMP all became absent, or weaker when relative $\dot{V}O_2max$ was instead correlated with LMP. This indicated that the interaction between a high capacity to deliver oxygen, coupled with a large body size are fundamental determinants in LMP.

Five strength measures were inversely correlated with LMP (Table 7.2), suggesting stronger individuals are better load-carriers (assuming all other physiological and anthropometric factors are equal). Muscle cross-sectional area is positively correlated (r=0.56) with muscle function (Mauhghan *et al.*, 1984), and accounts for 40% of the improvement in training induced muscle function (Narici *et al.*, 1989). Thus, strength's association with LMP may simply be reflection of body size (i.e. structure) influencing strength (i.e. function). In the present study, the moderate relationship between indices of strength and LMP approximate that reported between isokinetic dynamometry and longer loaded march distances of 8 km to 12 km (Mello *et al.*, 1988) and 16 km (Dziados *et al.*, 1987). The strength of this relationship was reported to weaken as the march distance increased to 20 km. The reason for this was suggested to be variable levels of subject motivation over longer march distances, which contaminate the involvement of physical / physiological measures (Knapik *et al.*, 1990).

Three anthropometric measures correlated with LMP, which included body mass and chest girth (both in a negative direction), and relative load size as a percentage of body mass (positive direction) (Table 7.2). These data indicated that a larger body size / dimension (i.e. body mass and chest girth), and a smaller relative load size (load % body mass) are associated with better LMP. This is likely due to the fact the LMP task had a prescribed absolute load of 20 kg, thus the larger individuals carried a smaller percentage of body mass. Previously, the total number, and magnitude of anthropometric correlations with LMP increased as external load size increased from 0 kg to 61 kg (Frykman and Harman, 1995), and 14 kg to 41 kg (Pandorf *et al.*, 2002). These data suggested that absolute load-carriage tests benefit larger individuals, which is further exacerbated as the size of an absolute load increases. Furthermore, these correlations were stronger than the 'moderate' correlations in the present study, which may be due to the larger loads applied, hence may have placed greater emphasis on the importance of body size in the LMP task.

7.4.3 The LMP model

Individually, LME (r=0.65), 2.4 km RT (r=0.42) and SLS (r=-0.48) moderately correlated with LMP, and formed the constituents of a multiple regression model at T1. This explained 65% (61% *Adjusted* R^2) of the variation in LMP (Equation 7.1), which was better, similar and worse than previously reported (Rayson *et al.*, 2000; Pandorf *et al.*, 2002; Rayson *et al.*, 2002a; Simpson *et al.*, 2006; Williams and Rayson, 2006; Griggs *et al.*, 2008). Within each of the studies cited marked variation existed in coefficient of determinations, which may have been a function of load size. For example R^2 decreased from 75% to 40% as load size increased from 15 kg to 25 kg (Rayson *et al.*, 2000), and from 82% to 69% as load increased from 14 kg to 41 kg (Pandorf *et al.*, 2002). Although the novel inclusion of LME as a determinant of LMP produced the best LMP model within the present study, these previous studies have produced both better and worse models, without the inclusion of specific sub-maximal measures of LM.

The LMP model had 95% SEE of 100 s, which was 31% of variation in LMP between-subjects at T1 (12:49 min:s to 18:11 min:s). Just one other study used a criterion LM task of similar distance (3.2 km) to the present study, and reported 95% limits of agreement (analogous to 95% SEE) (Williams and Rayson, 2006). Here the 95% limits of agreement were 171 s to 192 s for a 15 kg, and 243 s for a 25 kg LMP test, which were larger than the present study (100 s). When expressed in absolute units of measurement the smaller error observed in the present study may in part be due to the shorter test distance (2.4 km vs. 3.2 km), as LM models over longer distances (12.8 km) consistently have a larger error when expressed in absolute units (Rayson *et al.*, 2002a). Although little comparative data are available, the accuracy of the LMP model seems broadly similar to that previously reported.

This was the first study to report the sub-maximal metabolic measure of LME as a determinant of LMP. Others also have included LM variables in LMP models, and explained 48% to 71% (Simpson *et al.*, 2006) and 74 % (Griggs *et al.*, 2008) of the variance in LMP. However, in both instances these were maximal / performance measures of LM, not sub-maximal measures as per the present study. Griggs *et al.*, (2008) demonstrated LMP models that contain IVs specific to LM can explain 23% more variance, and have almost half of the prediction error of models that contain IV that are not specific to LM (124 s vs. 224 s).

The fact that LME emerged as a determinant of LMP was no surprise given that exercise economy is an established determinant of exercise performance (Coyle, 1995). However, this discovery is a new and novel finding. Previously, insights have been provided on the importance of sub-maximal LM measures, however these were related to absolute $\dot{V}O_2max$ and not LMP (Bilzon *et al.*, 2001; Lyons *et al.*, 2003, 2005). The total absence of literature pertaining LME means the potential influence it may have upon LMP is presently unknown. Whilst LME is known to improve during military style training (Gutekunst *et al.*, 2006), the concomitant effect it has upon LMP is yet to be established. Due to the dearth of knowledge on the LME-LMP interaction, the running literature must be consulted to gain an idea of how exercise economy interacts with exercise performance.

Running is continuous, repetitive and cyclical exercise, and although it is biased against heavier individuals that would ordinarily be competent load-carriers (Vanderburgh and Flanagan, 2000), it is the closest available comparison to load-carriage. Unlike the modelling of LMP, running performance has typically included physiological running measures (not unspecific fitness measures as LMP models have) to explain the determinants of running performance. Previously, marathon running performance was explained by; $\dot{V}O_2max$, $\%\dot{V}O_2max$ at lactate threshold, and running economy (Joyner, 1991). Brandon (1995) added to this, explaining middle distance performance with the addition of running velocity. The velocity attained at $\dot{V}O_2max$ (i.e. $v\dot{V}O_2max$) is a composite of $\dot{V}O_2max$ and running economy (Jones and Carter, 2000), and is the strongest individual correlate of 10 km to 90 km running performance (Noakes *et al.*, 1990). Together these models of running performance highlight the importance of exercise economy in determining exercise performance. Furthermore, a spread in running economy of ~8 $ml \cdot kg^{-1} \cdot min^{-1}$ differentiates poor from good running economy (Wood, 1999). Therefore, assuming two runners are matched for $\dot{V}O_2max$, poor *versus* good running economy would mean a higher velocity (hence better performance) can be achieved for the same oxygen demand / $\dot{V}O_2max$.

The importance of running economy was highlighted when individuals with relatively low $\dot{V}O_2max$ scores, but superior running economy maintained the same level of 10 km running performance compared to those with high $\dot{V}O_2max$ scores (Weston *et al.*, 2000). Such a phenomenon may be caused by Type IIa fibre distribution, which is inversely related to exercise economy (i.e. the greater number of Type IIa fibres, the poorer economy) (Hunter *et al.*, 2005), or an increase in Type I fibres (Coyle, 2005). Furthermore, running economy explained 65% of the variation in 10 km performance in individuals homogenous for $\dot{V}O_2max$ (Conley and Krahenbuhl, 1980). The running literature suggests running economy is a fundamental parameter in determining running performance, which provides a physiological rationale for the presence and inclusion of LME in the LMP model.

Loaded marching economy, 2.4 km RT and SLS produced the best LMP model despite LM fractional utilisation having the strongest bivariate correlation with LMP. A possible explanation for this may have been multicollinearity between LM fractional utilisation and other IVs. This may have been a consequence of LM fractional utilisation being a constituent of two measures (i.e. running $\dot{V}O_2max$ and LME), which may explain why the correlation of 0.65 between LM fractional utilisation and 2.4 km RT approached a level indicative of multicollinearity (Pallant, 2007).

As an inadequate sample size was available to mathematically model the change in LMP over infantry training (section 7.2.2), an alternative approach was undertaken. The individual importance each IV exerted on LMP (in terms of a change to LMP in seconds) was quantified to develop an understanding of how a PT prescription may improve LMP. In the present study the three determinants (IV) of LMP all contributed varying amounts to the LMP model. The squared part correlations (section 7.3.2) and standardised coefficients revealed that LME, RT, and SLS explained the largest to smallest amount of variation in LMP, respectively. Using the model's unstandardised coefficients (i.e. the m value in $y = mx + c$) it is possible to project the effect an improvement each IV (i.e. LME, RT, and SLS) has upon LMP. A hypothetical example is provided for the IV of LME (Equation 7.2). The individual effect a 10% improvement in LME, RT, or SLS has upon LMP can be calculated by **(m·(x·0.1)**, and when added to the LMP model gives:

$$LMP = (13.13 \cdot (LME \cdot 0.1)) + (0.81 \cdot RT) - (0.92 \cdot SLS) + 90.91$$

(Equation 7.2)

Based on this mathematical simulation, LMP would improve 42 ± 5 s if LME improved by 10 %, 47 ± 4 s for a 10 % improvement in RT, and 8 ± 6 s for a 10% improvement in SLS. These data suggest that from baseline fitness at the start of infantry training, an improvement in LME and RT would markedly improve LMP (42 s to 47s). However, a small improvement in LMP (8 s) was calculated from a 10% increase in SLS, which questions the efficacy of strength development at eliciting large improvements in LMP. This simulation suggests that during infantry training, PT that prioritises the development of LME and 2.4 km RT (i.e. aerobic training) will have the greatest influence on LMP gains.

In conclusion, sub-maximal loaded marching, maximal running, dynamic & isometric strength and anthropometric measures moderately correlated with 2.4 km loaded marching performance at the beginning of British Army infantry training. Loaded marching economy, 2.4 km run time, and static lift strength produced the best model of loaded marching performance. Mathematically, improvements in loaded marching economy, and 2.4-km run time would improve loaded marching performance by the greatest margin. Therefore, the prioritisation of physical training towards the development of loaded marching economy, and 2.4 km run time may be the most effective strategy to improve 2.4 km loaded marching performance.

CHAPTER 8

THE PHYSICAL AND PHYSIOLOGICAL RESPONSES TO 24 WEEKS OF BRITISH ARMY LINE INFANTRY TRAINING.

8.1 Introduction

The British Army has a duty of care to develop recruits to a minimum physical standard during initial training that is commensurate with the physical requirements of a particular career Arm or Service. These Career Employment Group standards vary between Arms or Services (British Army, 2002). The physical output standard for an infantry recruit is a 12.8 km loaded march (LM), carrying 25 kg, in a time no longer than 2 hours. This Basic Combat Fitness Test (BCFT) is one of the most demanding physical output standards in British Army recruit training, and is a categorical pass or fail test. As such the purpose of physical training (PT) in the British Army is very different to individuals involved in athletic competition. The British Army appraise the efficacy of PT by the incidence with which the minimum physical output standard (i.e. BCFT) is attained by recruits. This is fundamentally different to athletic populations that strive to become the fastest, longest, highest, or strongest.

Whilst there is not an organisational requirement to deliver 'optimal' PT, but rather 'adequate' PT during British Army training, training smart through the systematic application of the principles of training makes intuitive sense (Bompa, 1999). Raising the physiological, hence performance ceiling of a recruit will mean greater physical reserves can be called upon in a combat situation. However, training 'smart' during recruit training is confounded by the training environment, and the style of military PT prescription. First, the training environment is designed to test physical and mental robustness, as well as skill at arms and field craft (British Army website, 2007). The physical and mental stress imposed over the long waking hours of military training can cause energy deficits of up to 1000 kcal, decreases in testosterone and insulin-like growth factor-I, increases in cortisol, and losses in strength and power (Nindl *et al.*, 2007). Such sub-optimal conditions are often a function of inadequate rest and recovery cycles (McCafferty and Horvath, 1977). Consequently, a paradox exists. The arduous nature of recruit training is central to the evolution of a physically, and mentally robust soldier, but these conditions are not conducive to the development of physical performance.

Second, recruits are typically prescribed PT on a group (not individual) level. Thus, the same absolute PT load is performed by the most and least fit recruits. As a result an inverse relationship between initial fitness level and the magnitude of physical progression is observed (Vogel *et al.*, 1978; Stacy *et al.*, 1982; Gordon *et al.*, 1986; Trank *et al.*, 2001), which may in part explain low fitness as a risk factor for injury in military training (Blacker *et al.*, 2005). The prescription of PT on an individual basis appears unrealistic due to large recruit numbers. However, a possible solution has proposed to stream recruits into ability groups based on aerobic fitness (Bilzon, 2003; Knapik *et al.*, 2006), or into endurance or strength focused groups based on the weakest aspect of fitness (i.e. endurance or strength) (Williams *et al.*, 2004). Currently, the Infantry Training Centre (Catterick) (ITC(C)) does neither, and prescribes PT on a group level.

Accepted measures of LMP include aerobic fitness, anthropometry and strength (Rayson *et al.*,1993; Frykman and Harman, 1995; Harman and Frykman, 1995; Rayson *et al.*, 2000; Bilzon *et al.*, 2001; Pandorf *et al.*, 2002; Rayson *et al.*, 2002a; Wilkinson *et al.*, 2003; Lyons *et al.*, 2005; Williams and Rayson, 2006). As such the extent these parameters adapt under the constraints of British Army recruit training should influence the magnitude of LMP improvement.

Recent work has highlighted both maximal (Simpson *et al.*, 2006; Griggs *et al.*, 2008), and sub-maximal measures of LM (Chapter 7) as determinants of LMP. In the running literature task specific measures are common to running models (Joyner, 1991; Brandon, 1995). Furthermore, within these models, exercise economy is essential (i.e. the oxygen uptake at an absolute exercise intensity (Jones and Carter, 2000)). For example good running economy can compensate for an inferior $\dot{V}O_2max$ without an observed decrease in running performance (Weston *et al.*, 2000), and can explain 65% of running performance in individuals homogenous for $\dot{V}O_2max$ (Conley and Krahenbuhl, 1980). As such an equivalent parameter for LM (i.e Loaded Marching Economy, LME) could help to better understand the LMP training response. The only study to investigate the training response of LME reported a 5.9% improvement (i.e. a decrease in O_2 demand) in LME over 8-weeks of military style training (Gutekunst *et al.*, 2006). However, a criterion test of LMP was not performed in this seminal work, meaning the actual influence LME changes could exert on changes to LMP was not quantified.

Several studies have assessed training responses during British Army recruit training (Legg and Duggan, 1996; Brock and Legg, 1997; Williams *et al.*, 1999; Williams *et al.*, 2002; Wilkinson *et al.*, 2004; Williams, 2005). However, none have used the direct measurement of maximal $\dot{V}O_2max$ or LME, which may provide a greater insight into the physiological mechanisms of adaptation during recruit training, and how these influence LMP.

The purpose of this study was to track the training responses of infantry recruits over the duration of the Combat Infantryman's Course (CIC) at the Infantry Training Centre (Catterick), with the aim of determining the adequacy of the current physical training programme at developing aspects of fitness that are important to the representative military task of loaded marching performance.

8.2 Methods

8.2.1 Subjects

Forty-four new entry recruits and 1 'back-squadded' recruit (*n*=45) formed Rifles 3 Platoon, and undertook the 24-week CIC at ITC(C). All recruits volunteered to participate in the study, which was approved by the Ministry of Defence (Naval) Personnel Research Ethics Committee (MOD(N) PREC). A full Initial Medical Examination (IME) was conducted prior to participation in the study, and a resting 12-lead electrocardiogram (ECG) was obtained to screen for cardiac abnormalities. Two recruits failed the IME and were excluded from the trial, the remaining recruits presented normal ECG traces and were deemed fit to participate in the study. The characteristics of the 25 subjects that completed the study are presented (Table 8.1). All analyses were conducted on the 25 successful subjects.

Table 8.1. Subject characteristics (median (range)).

Age (years)	Stature (m)	Body mass (kg)	Body fat (%)	$\dot{V}O_2max$ ($ml \cdot kg^{-1} \cdot min^{-1}$)	2.4 km run time (min:s)
19 (8)	1.74 (0.25)	67.6 (26.1)	12.7 (12.3)	54.7 (12.0)	10:17 (2:00)

8.2.2 Experimental design

The study consisted of a within-subject, repeated measures design. This was characterised by three discrete testing periods at the approximate beginning, middle and end of the 24-week training programme. Testing periods equated to weeks 1/2 (T1), 11/12 (T2), and 20/21 (T3). On each of the three test occasions subjects performed a test battery that had previously undergone a repeatability assessment (Chapter 6, Appendix A). The test battery included; 2.4 km LMP, sub-maximal LME, 2.4 km run time, dynamic upper and lower body strength, isometric static lift strength (SLS) strength, isometric hand grip (HG) strength, muscular endurance (2 minutes of press ups and sit ups), body mass, circumferential girths (6-site), and skinfold thickness (8-site). Formal PT was delivered between weeks 1 to 21. A schematic representation of the study design is presented (Figure 8.1).

8.2.3 Test battery

A more detailed explanation of the test battery is presented in Chapter 3.

8.2.3.1 Loaded Marching Tests

After an 800 m warm-up a 2.4 km LMP test, with an external load of 19.7 ± 0.7 kg was conducted around an all weather athletics track. The first 1.2 km was conducted as a group, and the second 1.2 km was an individual 'best effort'.

The Loaded March Economy (LME) test was conducted in a field laboratory that was established within the ITC(C) Medical Centre. Temperature (22.6 ± 1.0 °C) and relative humidity (45.6 ± 8.7%) were controlled with a portable air conditioning unit (PLM

15KEH-410, Amcor, UK). Wearing boots and CS95 clothing the test (section 3.5.2.2) was performed on a motorised treadmill (PPS 70Med-I, Woodway, Germany) between the hours of 1800 and 2300. Expired gases were collected with the Douglas Bag method previously described (section 3.6.1).

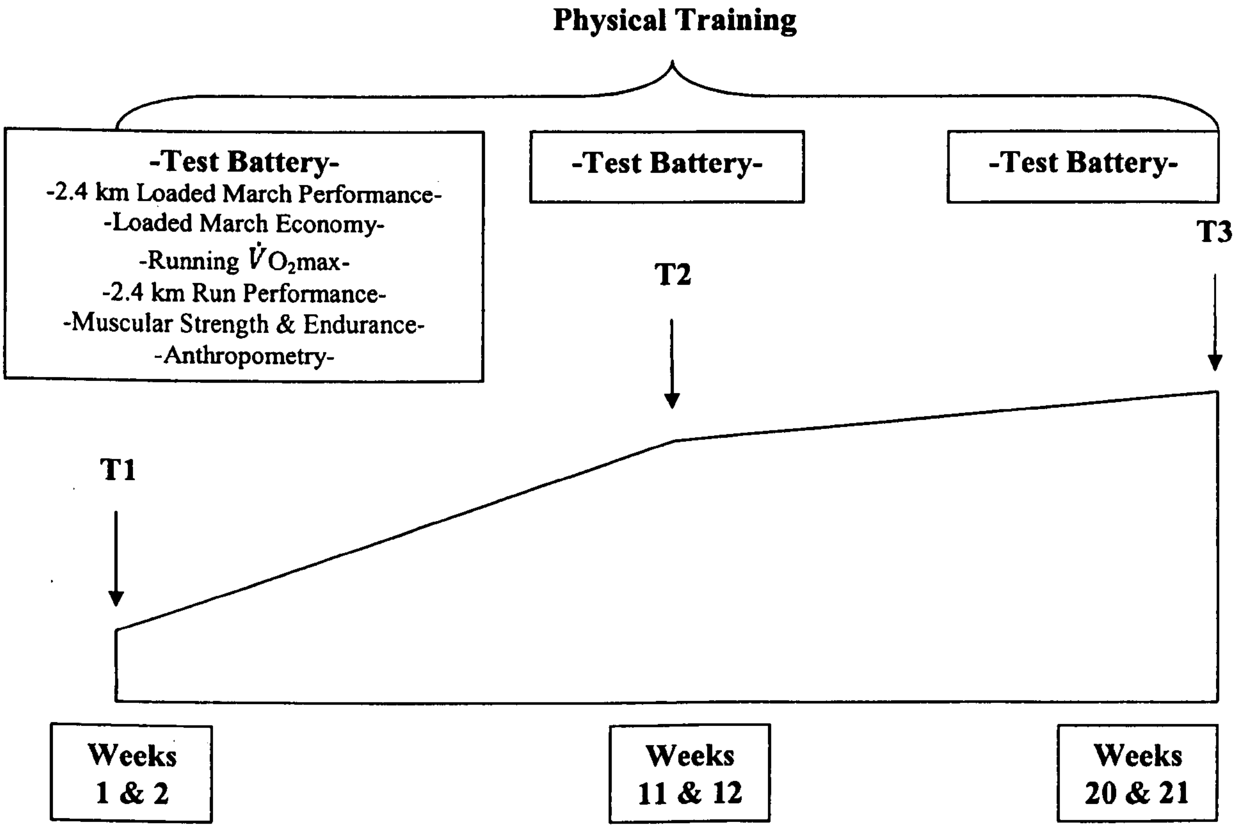


Figure 8.1. Schematic representation of the study design.

8.2.3.2 Running Tests

Following an 800 m warm-up a 2.4 km individual 'best effort' running test was performed around an all weather athletics track.

A running $\dot{V}O_2$max test in standard PT kit was performed on a motorised treadmill as previously described (section 3.5.2.4) between the hours of 1800 and 2300 hours. Throughout the test an Independent Medical Officer monitored for cardiac abnormalities with a 5-lead ECG.

8.2.3.3 Muscular Strength and Endurance Tests

Muscular strength tests were conducted in the Platoon's accommodation block between the hours of 1800 and 2200. Dynamic strength was evaluated using a portable dynamometer (DYNO, Concept 2, Nottingham, UK). Following three warm-up repetitions each subject was instructed to perform a chest press, back pull, and leg press exercise as hard and as fast as possible. The mean and peak forces of the 5 upper body, and 10 lower body maximal repetitions were recorded (section 3.4).

Isometric strength tests consisted of the Static Lift Strength (SLS) and Hand Grip (HG) strength. SLS was performed from the squat position with the dynamometer's handle 0.38 m above the platform (Knapik *et al.,* 1981) (Model 5402, Takei, Japan). HG strength was performed from a standing position with a HG dynamometer (Model 5001, Takei, Japan), and the dominant arm was held at the side of the body. In both instances one warm-up repetition at ~50 % was allowed (British Army, 2002). Subsequently, 3 maximal repetitions were performed, separated by 30 s, and the mean and peak forces were recorded (section 3.3).

Muscular endurance tests were conducted during formal PT lessons, and preceded the 2.4 km run. These consisted of the maximum number of press-ups (section 3.10.2.1) and sit-ups (section 3.10.2.2) performed over a duration of 2 minutes.

8.2.3.4 Anthropometric Tests

Anthropometric measures were conducted in the Platoon's accommodation block between the hours of 1800 and 2200 and consisted of: body mass; body stature; 6-site circumferential girths (i.e. relaxed upper arm (mid-point), chest (level with the nipple line), waist (narrowest measure), gluteals (most posterior protuberance), thigh (mid-point) and calf (largest measure)); and 8-site skinfold (i.e. triceps, biceps, subscapular, iliac crest, supraspinale, abdominal, front thigh and medial calf). The sum of the eight skinfolds (ISAK, 2001), and estimated fat mass, fat free mass and percentage body fat (Durnin and Womersley, 1974) were also derived (section 3.2).

8.2.4 Physical training programme

Over the 24 weeks of infantry training, the PT programme was delivered between weeks 1 to 21. PT lessons consisted of blocked periods, with each period 40 minutes in duration. A total of 130 periods (which equated to 60 PT lessons) were conducted over the 21 weeks. The 130 periods were, loaded marching (37 periods), continuous runs (15), circuit training (25), Representative Military Tasks (RMTs) (e.g. material handling tasks) and Personal Fitness Test (PFT) (e.g. 2.4 km run, 2 minutes of press ups and sit ups) (12), battle PT (e.g. obstacle course and steeple chase) (18), sports (15), swimming (6), and gym agility (2) (Figure 8.2). On average 6.2 periods (4.1 hours) of PT were delivered each week. The PT volume peaked in weeks 2 and 17, where the total number of periods reached 11 (7.3 hours) (Figure 8.3). The total loaded march distance was 119 km (5.7 km·week^{-1}), with the largest weekly volume of 12.8 km occurring in week-21. The three Introduction Exercises (Intro Ex) and three Tactical Exercises (Tac Ex) involved additional loaded marching, and were separate from the PT programme (Figure 8.4). External loads during loaded marches increased from 8.6 kg (week 2), to 25 kg (weeks 19, 20, and 21) (Figure 8.4). The total continuous run distance was 50 km (2.4 km·week^{-1}) (Figure 8.5). Additional physical exertion (not part of the formal PT programme) occurred during exercises (Figure 8.4), drill, bayonet lessons, rifle ranges and adventurous training.

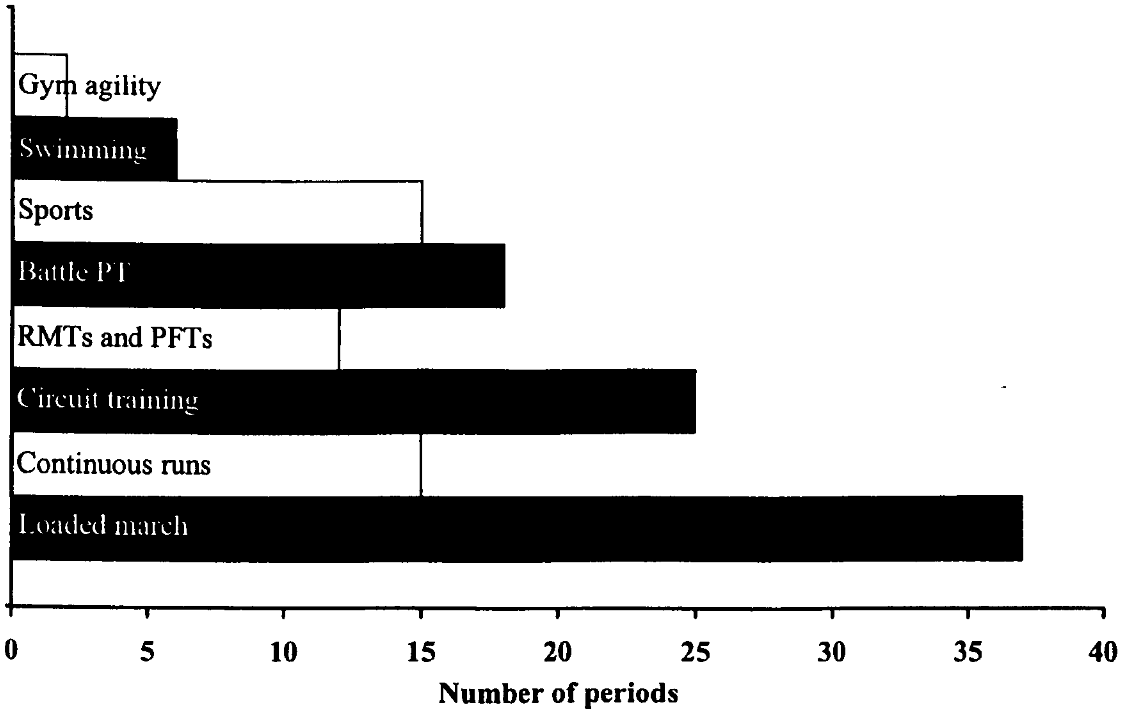


Figure 8.2. The modality and number of formal PT periods during infantry training.

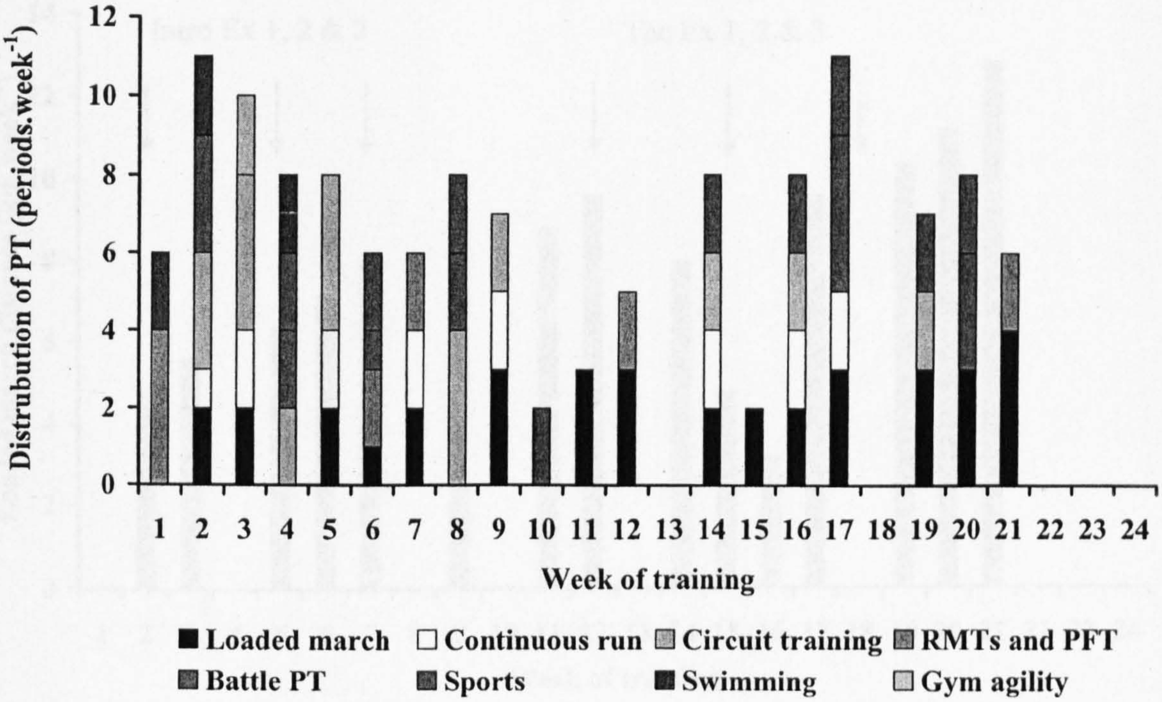
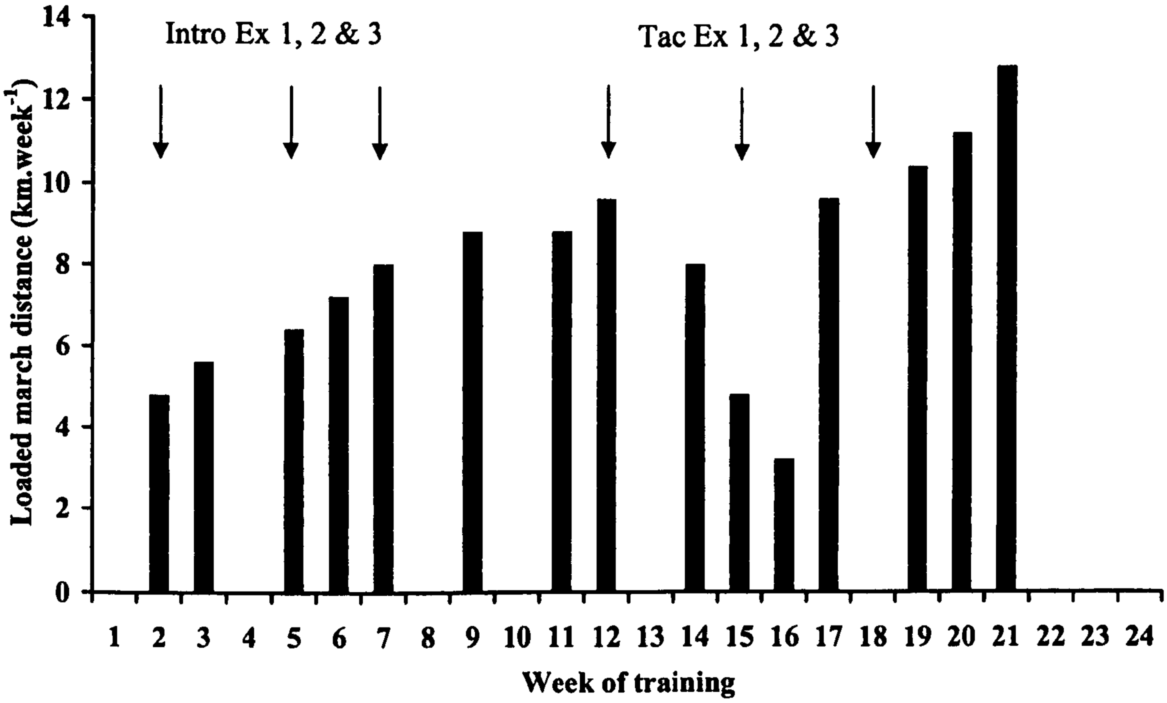


Figure 8.3. The distribution and modality of formal PT periods during infantry training. PT = Physical training; RMT = Representative Military Tasks; PFT = Personal Fitness Test.



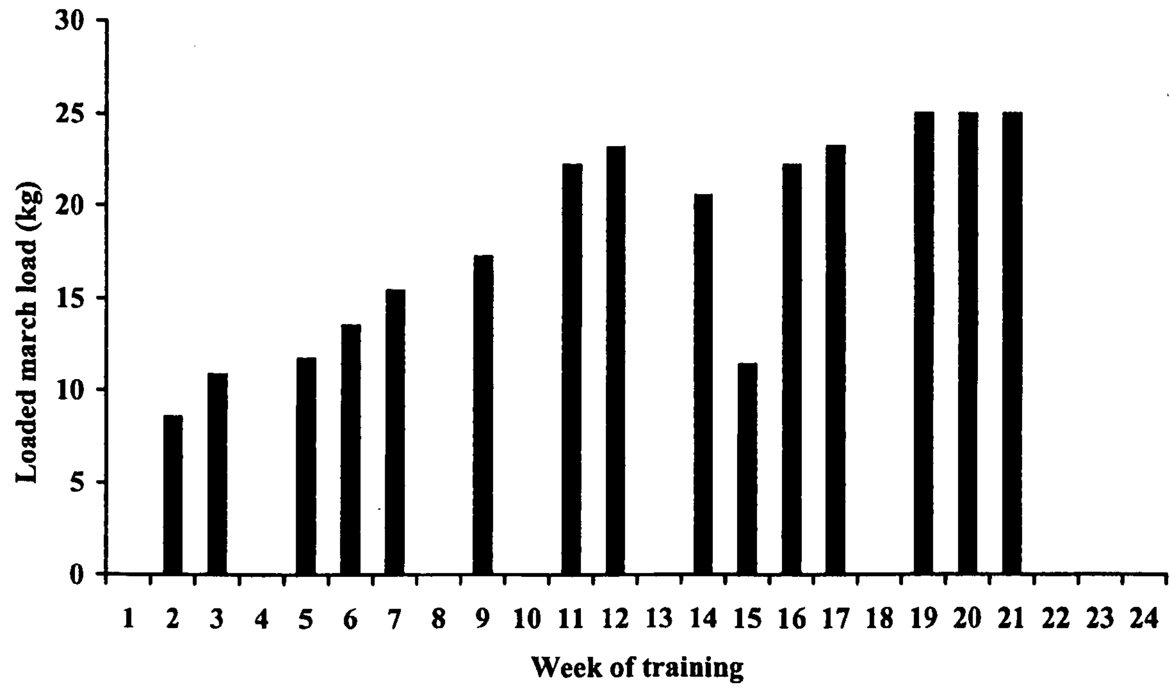


Figure 8.4. (Top) The weekly loaded march distances, and additional loaded marches during field exercises. (Bottom) The external load size of loaded marches during infantry training.

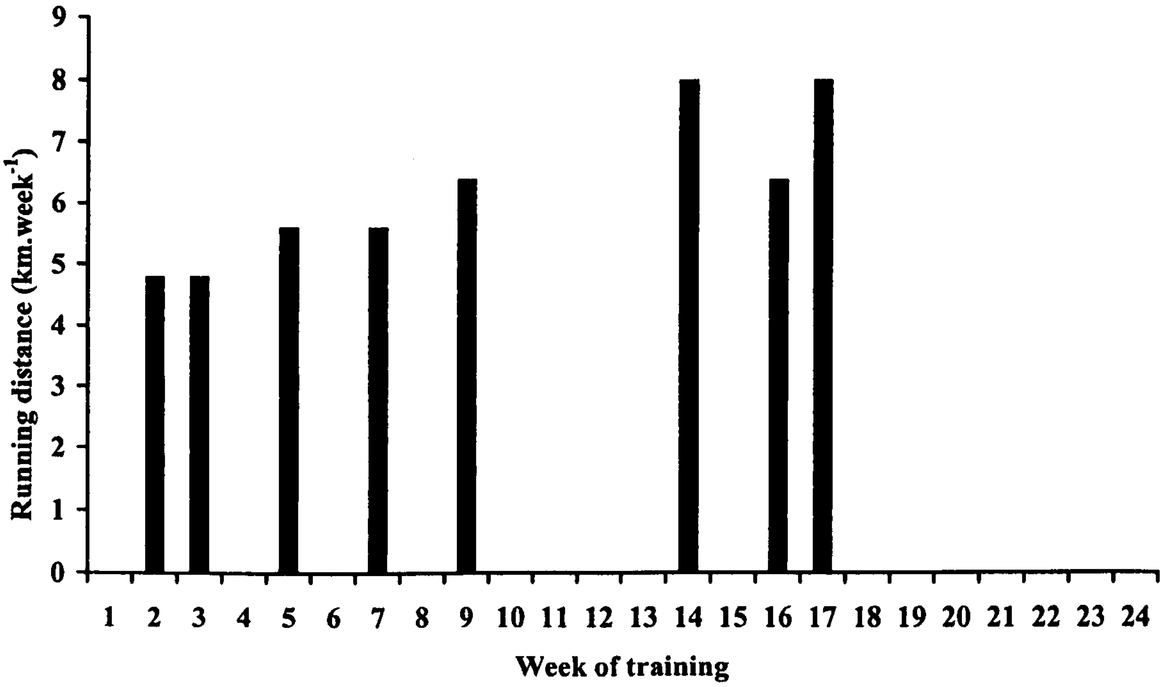


Figure 8.5. The distribution and volume of continuous run distance over the duration of infantry training.

8.2.5 Statistical analysis

Four variables were used to calculate the prospective sample size required to detect a significant change in LME over training (PC Size, Dallal, 1985). These were: the level at which the null hypothesis would be tested ($P<0.05$); the minimum anticipated effect size of LME over a training intervention (1.7 ml·kg^{-1}·min^{-1} (Sjödin *et al.*, 1982; Johnston *et al.*, 1997)); the random error of LME measurement (± 2.7 ml·kg^{-1}·min^{-1} (Chapter 4)); and the ability (power) with which the null hypothesis could be rejected (0.8). The minimum sample size required was n=22. As infantry first time pass rates approximate 40-50% (Carter *et al.*, 2006) an initial sample size of 44 was required to achieve adequate statistical power to reject the null hypothesis.

Data were checked for the assumption of Normality (Fallowfield *et al.*, 2005). Due to the high prevalence of non-Normally distributed data, the central tendency and spread of data were reported as Median and (Range) (Thomas and Nelson, 1996). For visual purposes the middle two quartiles of the data (and not the range) constituted the error bars present in all figures, as this highlighted trends more effectively in the data. A non-parametric inferential statistical approach was chosen, the rationale for this decision is provided elsewhere (section 3.11.3). A non-parametric Friedman ANOVA with 3 levels (T1, T2, and T3) examined the effect of time on physical performance ($P<0.05$). A Wilcoxon Signed Rank Test *post hoc* test, with Bonferroni adjustment (Winter *et al.,* 2001) identified which time points were different. As such *post hoc* statistical significance was accepted as $P<0.0167$. Effect size (r) was reported for *post hoc* comparisons using equation 3.8 (section 3.11.3).

8.3 Results

8.3.1 Subject Attrition

Of the initial 45 subjects, 20 did not complete the 21-week research trial. Of these 20 recruits, 2 recruits were transferred for medical rehabilitation, 2 were 'back-squadded' and 16 were discharged (i.e. 2 recruits medically discharged, 12 discharged as of right, and 2 recruits were discharged with services no longer required). A total of 25 recruits successfully completed the study, which equated to a training attrition of 44%.

8.3.2 Loaded Marching Fitness

8.3.2.1 2.4 km loaded marching performance

Data are presented as *Median (Range)*. Effect size (r) is also provided. The 2.4 km LMP improved ($\chi^2_{(2)}=17.38$, $P<0.01$) by 63 (239) s over training (i.e. T1 to T3) (from 900 s to 837 s) ($T=1$, $P=0.001$, $r=-0.59$). There was a tendency for performance to improve 12 (201) s at T2 ($T=5$, $P=0.026$, $r=-0.39$). The main improvement in LMP of 51 (88) s occurred in the second half of training (from 889 to 837 s, $T=2$, $P=0.001$, $r=-0.58$) (Figure 8.6).

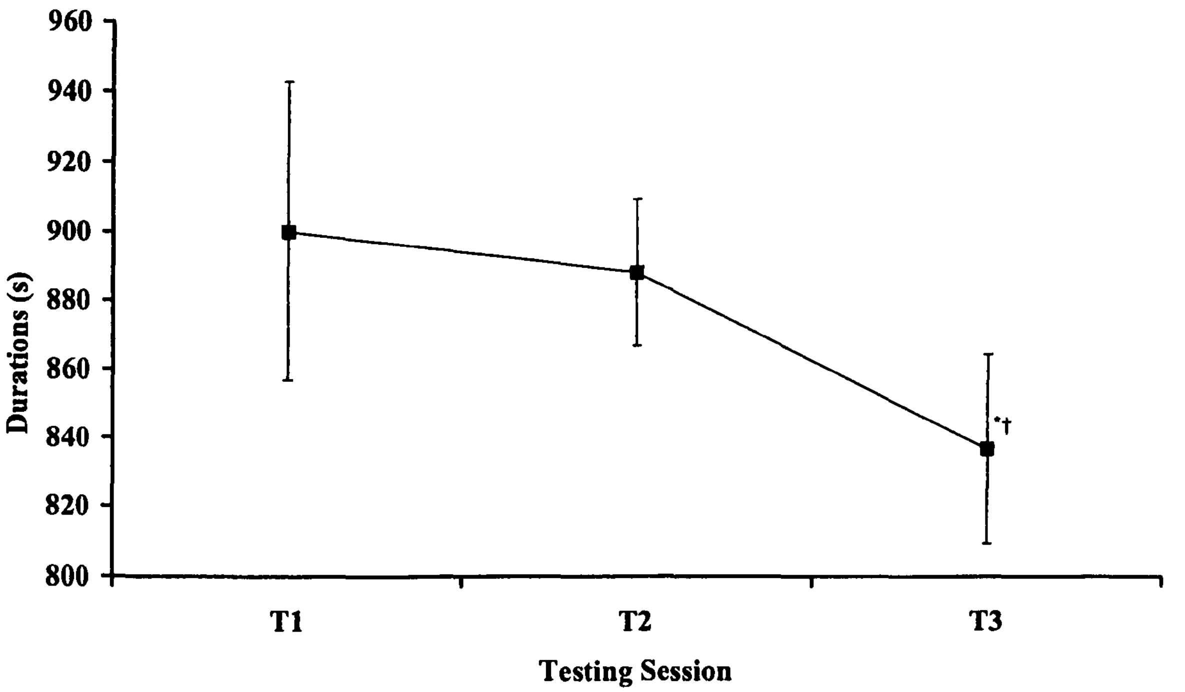


Figure 8.6. 2.4-km LMP time at T1, T2 and T3. * = significantly different from T1. † = significantly different from T2.

8.3.2.2 Sub-maximal treadmill loaded marching

Expressed relative to body mass, LME significantly improved ($\chi^2_{(2)}$=27.55, *P*<0.01) (i.e. the oxygen demand of loaded marching decreased) by 4.9 (7.5) ml·kg^{-1}·min^{-1} between T1 to T3 (*T*=0, *P*<0.01, r=-0.60). The entirety of this response occurred in the first half of training (i.e. by T2) (*T*=0, *P*<0.01, r=-0.62) (Table 8.2). Absolute LME also improved ($\chi^2_{(2)}$=23.47, *P*<0.01) however this was less marked, as $\dot{V}O_2$ decreased by 0.22 (0.62) l·min^{-1} between T1 to T3 (*T*=1, *P*<0.01, r=-0.60). Most of the response occurred by T2 (*T*=1, *P*<0.01, r=-0.61). The fractional utilisation of $\dot{V}O_2$max during the LM task decreased ($\chi^2_{(2)}$=21.70, *P*<0.01) by 8.6 (23.56) %$\dot{V}O_2$max between T1 to T3 (*T*=0, *P*<0.01, r=-0.62). The adaptation of 9.8% $\dot{V}O_2$max occurred by T2 (*T*=1, *P*<0.01, r=-0.59). Minute ventilation ($\dot{V}_E$) decreased over training ($\chi^2_{(2)}$ = 19.05, *P*<0.01) by 14.1 (25.2) l·min^{-1} (*T*=2, *P*<0.01, r=-0.58), most of this decrease had occurred by T2 (*T*=2, *P*<0.01, r=-0.59). The fraction of expired oxygen (F_EO_2) decreased between T1 to T3 ($\chi^2_{(2)}$=18.01, *P*<0.01) by 0.43 (1.85) % (*T*=2, *P*<0.01, r=-0.52). This adaptation occurred by T2 (*T*=2, *P*=0.012, r=-0.41). The fraction of expired carbon dioxide (F_ECO_2) increased between T1 to T3

($\chi^2_{(2)}$=21.15, P<0.01) by 0.21 (1.35) % (T=0, P=0.010, r=-0.54). This increase occurred by T2 (T=2, P=0.007, r=-0.43) (Table 8.2). Respiratory Exchange Ratio (RER) did not change over training ($\chi^2_{(2)}$=4.82, P=0.090, r=-0.26). Blood lactate concentration decreased over training ($\chi^2_{(2)}$ = 29.66, P<0.01) by 2.0 (4.9) mmol·l^{-1} (T=0, P<0.01, r=-0.60), with most of the decrease occurring by T2 (T=0, P<0.01, r=-0.60). Heart rate decreased ($\chi^2_{(2)}$=20.46, P<0.01) by 22 (14) b·min^{-1} over training (T=0, P=0.010, r=-0.62). All adaptation occurred by T2 (T=0, P<0.01, r=-0.62). Loaded march Stride Frequency (SF) decreased ($\chi^2_{(2)}$=8.30, P=0.016), by 2.0 (10.0) strides·min^{-1}, however this did not occur until the second half of training (i.e. T2 to T3) (T=2, P=0.002, r=-0.47) (Table 8.2).

Table 8.2. Loaded marching physiology at T1, T2 and T3 (median (range)).

Measure	T1	T2	T3
$\dot{V}O_2$(l·min^{-1})	2.28 (0.81)	2.12 (0.75)*	2.06 (0.79)*
$\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	35.0 (13.4)	30.1 (13.0)*	30.1 (11.5)*
Fractional Utilisation (%$\dot{V}O_2$max)	62.1 (26.5)	52.3 (22.7)*	53.5 (24.9)*
RER	0.99 (0.17)	0.96 (0.13)	0.96 (0.13)
$\dot{V}_E$ (l·min^{-1})	56.9 (28.9)	43.9 (20.1)*	42.8 (23.1)*
F_EO_2 (%)	16.69 (2.25)	16.22 (1.45)*	16.26 (1.84)*
F_ECO_2 (%)	4.36 (1.88)	4.58 (1.00)*	4.57 (1.40)*
La (mmol·l^{-1})	3.4 (5.5)	1.6 (1.8)*	1.4 (1.4)*
HR (b·min^{-1})	164 (50)	142 (43)*	142 (19)*
Stride Frequency (strides·min^{-1})	67.0 (13.0)	68.0 (11.0)	66.0 (8.0)†

* = Significantly different from T1
† = Significantly different from T2

8.3.3 Running Fitness

8.3.3.1 2.4 km running performance

Run time improved ($\chi^2_{(2)}$=13.47, *P*=0.001), decreasing 22 (121) s between T1 to T3 (from 617 to 595 s, *T*=3, *P*=0.002, r=-0.48). The majority of this improvement (21 s) occurred by T2, but was only a tendency (*T*=6, *P*=0.027, r=-0.33). Statistically no improvement occurred between T2 to T3 (*T*=10, *P*=0.685, r=-0.06), however 2.4 km run time was different from baseline at T3 (*T*=3, *P*=0.001, r=-0.48) (Figure 8.7).

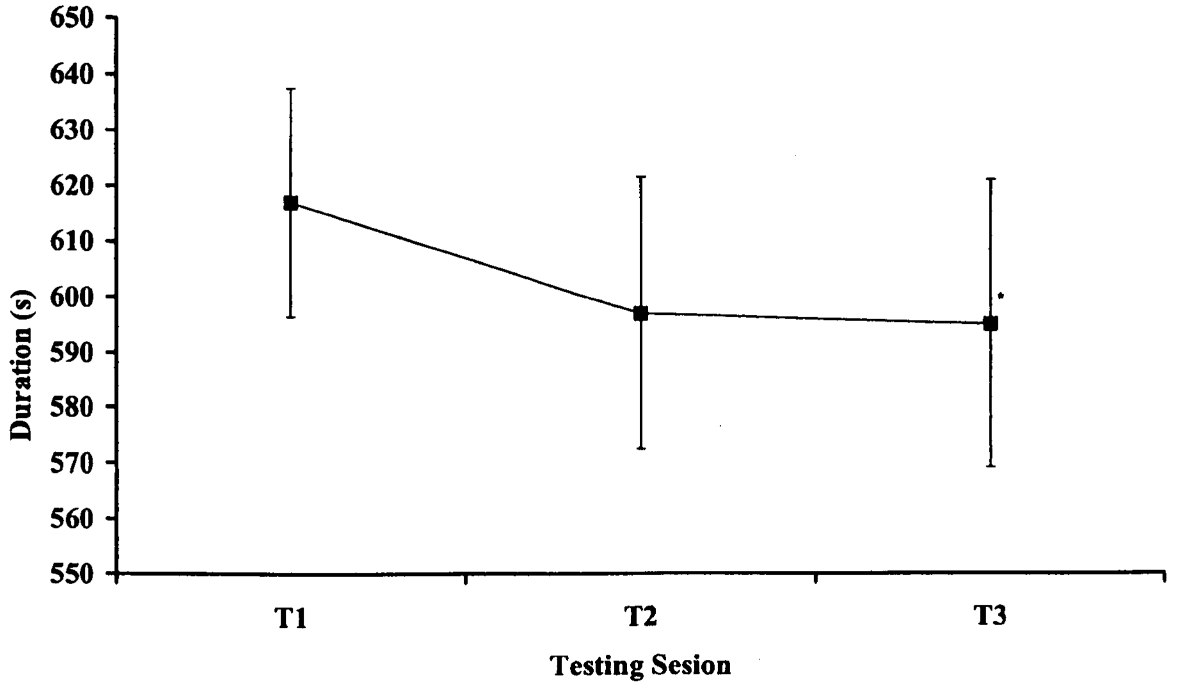


Figure 8.7. 2.4-km run time at T1, T2 and T3. * = significantly different from T1.

8.3.3.2 Running $\dot{V}O_2$max

Relative $\dot{V}O_2$max did not change ($\chi^2_{(2)}$=1.80, *P*=0.407), however a tendency for absolute $\dot{V}O_2$max to change over training was observed ($\chi^2_{(2)}$ = 5.78, *P*=0.056). The tendency for absolute $\dot{V}O_2$max to increase (*P*=0.056) was significant in the *post hoc* test. Here absolute $\dot{V}O_2$max significantly increased by 0.10 (0.88) $l \cdot min^{-1}$ between T1 to T2 (*T*=4, *P*=0.007, r=-0.45).

Weak associations between initial $\dot{V}O_2max$ at T1 and the resultant change in $\dot{V}O_2max$ over training were observed for absolute (R=0.31) and relative (R=0.38) $\dot{V}O_2max$ (Figure 8.8). The velocity at $\dot{V}O_2max$ (v$\dot{V}O_2max$) did change significantly over training ($\chi^2_{(2)}$=16.30, *P*<0.001). However, at T2 v$\dot{V}O_2max$ increased by 1.0 (3.4) km·h^{-1} (*T*=1, *P*<0.01, r=-0.60), which subsequently decreased (0.2 km·h^{-1} (2.8)) back to baseline between T2 and T3 (*T*=5, *P*=0.003, r=-0.47).

Minute ventilation ($\dot{V}_E$) did not change over training ($\chi^2_{(2)}$=4.33, *P*=0.115). However, the *post hoc* revealed a significant decrease of 1.3 (32.8) l·min^{-1} between T2 and T3 (*T*=4, *P*=0.008, r=-0.44). Peak RER changed over training ($\chi^2_{(2)}$=6.68, *P*=0.036). A trend increase of 0.03 (0.14) occurred by T2 (*T*=5, *P*=0.024, r=-0.38), with a significant decrease of 0.03 (0.18) between T2 to T3 (*T*=4, *P*=0.004, r=-0.48). Peak blood lactate concentration significantly changed ($\chi^2_{(2)}$=26.47, *P*<0.01). This decreased 4.2 (8.3) mmol·l^{-1} between T1 to T3 (*T*=0, *P*<0.01, r=-0.62), which was significant in both the first (*T*=2, *P*=0.004, r=-0.50) and second halves of training (*T*=2, *P*=0.006, r=-0.48). Similarly, peak heart rate decreased below baseline ($\chi^2_{(2)}$ = 24.087, *P* = 0.000) by 13 (23) b·min^{-1} over training (*T*=0, *P*<0.01, r=-0.62). The majority of the decrease occurred between T2 to T3 (*T*=2, *P*=0.001, r=-0.57)) (Table 8.3).

Table 8.3. Maximal running treadmill physiology at T1, T2 and T3. (median (range)).

Measure	T1	T2	T3
$\dot{V}O_2$(l·min^{-1})	3.74 (1.41)	3.84 (1.13)*	3.76 (1.27)
$\dot{V}O_2$ (ml·kg^{-1}·min^{-1})	54.7 (12.0)	55.4 (14.2)	55.0 (15.9)
v$\dot{V}O_2$max (km·h^{-1})	14.2 (5.2)	15.2 (3.6)*	14.4 (4.2)†
$\dot{V}_E$ (l·min^{-1})	101.6 (61.4)	102.9 (41.5)	101.2 (42.8)†
RER	1.19 (0.21)	1.22 (0.15)	1.19 (0.24)†
La (mmol·l^{-1})	10.2 (10.6)	8.3 (7.9)*	6.0 (8.0)*†
HR (b·min^{-1})	199 (37)	195 (32)	186 (37)*†

* = Significantly different from T1
† = Significantly different from T2

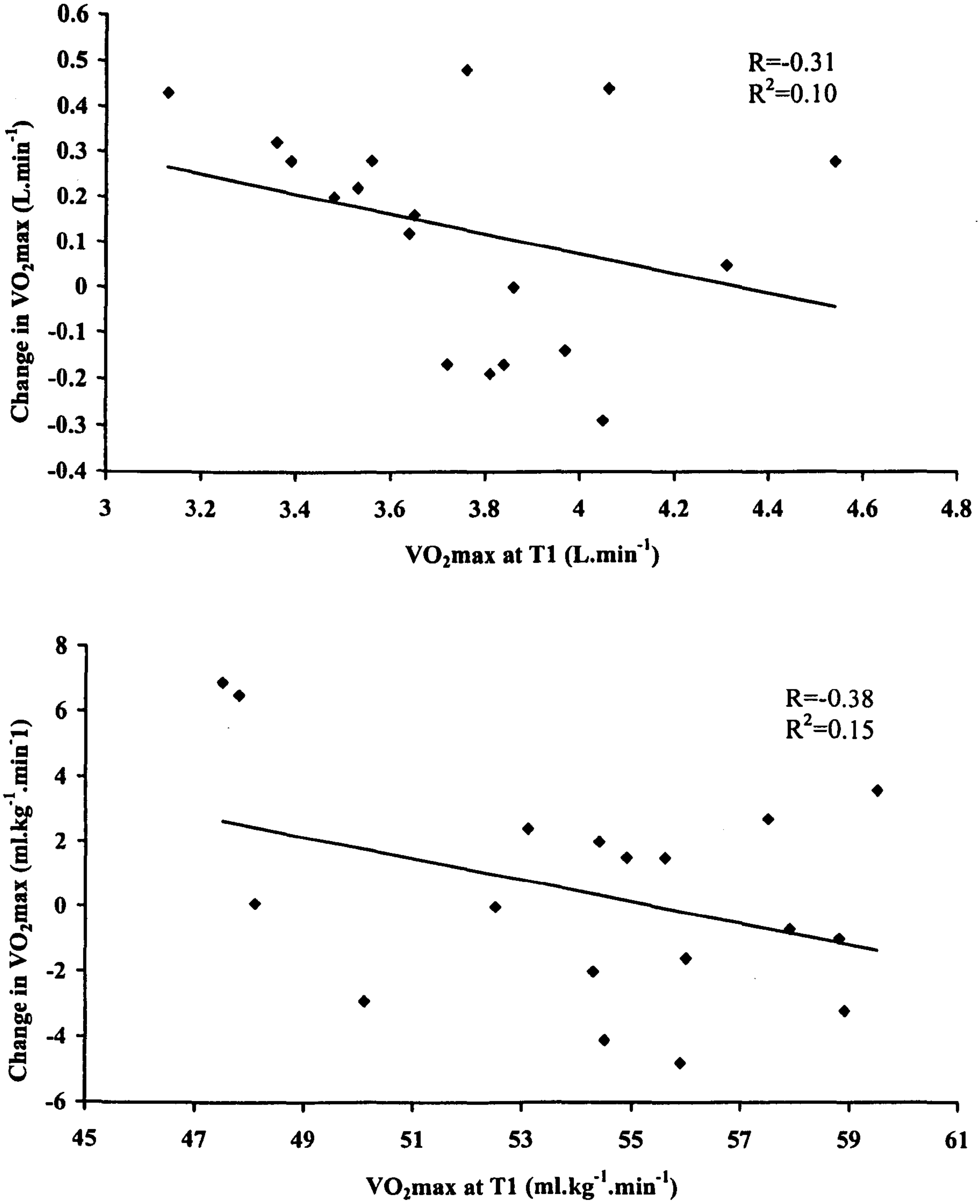


Figure 8.8. Changes in absolute (top) and relative (bottom) $\dot{V}O_2$max when related to $\dot{V}O_2$max at baseline.

8.3.4 *Muscular Strength and Endurance*

8.3.4.1 Dynamic strength

Changes in dynamic strength for chest press, back pull and leg press are presented (Table 8.4). The Friedman ANOVA revealed both average and peak upper body strength did not change (P>0.05). However, despite an absent initial increase in average chest strength (T=10, P=0.589, r=-0.07), the *post hoc* revealed a decrease of 2 (14) kg between T2 and T3 (T=5, P=0.010, r=-0.36). There was a tendency for peak chest strength to decrease 2 (22) kg between T1 to T3 (T=9, P=0.042, r=-0.28). Neither average ($\chi^2_{(2)}$=4.42, P=0.109; T=11, P=0.667, r=-0.06) or peak back pull strength changed ($\chi^2_{(2)}$=1.58, P=0.454; T=11, P=0.703, r=-0.05). Both average ($\chi^2_{(2)}$=17.71, P<0.01) and peak ($\chi^2_{(2)}$=8.75, P=0.013) leg press strength increased between T1 to T3. Average leg strength increased 23 (128) kg by T3 (T=4, P=0.002, r=-0.43), with most of the increase occurring by T2 (T=3, P=0.001, r=-0.48). There was a tendency for peak leg strength to increase 13 (126) kg by T3 (T=7, P=0.036, r=-0.30), which was significant in the first half of training (T=6, P=0.007, r=-0.38).

8.3.4.2 Isometric strength

Changes were observed in average ($\chi^2_{(2)}$=12.09, P=0.002), but not peak SLS ($\chi^2_{(2)}$=1.18, P=0.554). Between T1 to T3 average SLS was unchanged (T=8, P=0.372, r=-0.13), which was attributable to an increase of 4.2 (44.3) kg between T1 to T2 (T=4, P=0.005, r=-0.42), and subsequent decrease of 4.6 (36.8) kg between T2 to T3 (T=5, P=0.002, r=-0.47). No change in peak SLS was observed by T3 (T=11, P=0.249, r=-0.17), however the *post hoc* test reported a tendency for peak SLS to decrease 1.5 (37.0) kg between T2 to T3 (T=8, P=0.024, r=-0.34) (Table 8.4).

Changes were observed in average ($\chi^2_{(2)}$=6.80, P=0.033), but not peak HG strength ($\chi^2_{(2)}$=4.021, P=0.134). Over training average hand grip strength did not change (T=11, P=0.289, r=-0.15), however there was a tendency for average HG strength to decrease 2

(20.3) kg between T2 to T3 (T=5, P=0.036, r=-0.30). No change occurred in peak HG strength over training (T=11, P=0.596, r=-0.07) (Table 8.4).

8.3.4.3 Muscular endurance

The number of press-ups ($\chi^2_{(2)}$=8.27, P=0.016) and sit-ups ($\chi^2_{(2)}$=17.89, P<0.01) completed over a 2-minute period did change over training. The number of press-ups did not change between T1 to T3 (T=10, P=0.807, r=-0.04), which can be explained by the initial tendency for the repetition number to increase 3 (21) repetitions (T=5, P=0.017, r=-0.35) and subsequent decrease by 5 (46) repetitions between T2 to T3 (T=6, P=0.008, r=-0.39). The number of sit ups increased by 6 (27) repetitions between T1 to T3 (T=5, P=0.006, r=-0.40), which had manifested by T2 (T=3, P<0.01, r=-0.55) (Table 8.4).

Moderate to strong inverse correlations were observed between initial strength at baseline, and the subsequent change in strength by the end of training (Table 8.5).

Table 8.4. Indices of muscular strength and endurance at T1, T2 and T3. (median (range)).

	T1	T2	T3
Dynamic Strength			
A Chest Strength	61 (36)	63 (27)	61 (24)†
P Chest Strength	65 (36)	64 (25)	63 (22)
A Back Strength	62 (31)	62 (28)	64 (30)
P Back Strength	66 (32)	66 (30)	67 (30)
A Leg Strength	138 (135)	153 (111)*	161 (143)*
P Leg Strength	165 (135)	172 (124)*	178 (134)
Isometric Strength			
A Static Lift Strength	116.1 (81.2)	120.3 (76.0)*	115.7 (75.0)†
P Static Lift Strength	126.0 (77.0)	123.5 (76.5)	122.0 (76.5)
A Hand Grip Strength	43.3 (25.0)	45.0 (22.0)	43.0 (18.0)
P Hand Grip Strength	46.0 (26.0)	46.0 (24.0)	45.0 (20.0)
Muscular Endurance			
Press-ups	52 (41)	55 (42)	50 (30)†
Sit-ups	53 (40)	60 (45)*	59 (42)*†

* = Significantly different from T1; † = Significantly different from T2.
A = average strength, P = peak strength.

Table 8.5. Correlation between, changes in various peak strength indices and baseline strength.

	R	R^2
Chest Strength	-0.79	0.63
Back Strength	-0.48	0.23
Leg Strength	-0.50	0.25
Static Lift Strength	-0.50	0.25
Hand Grip Strength	-0.60	0.36

8.3.5 Anthropometry

8.3.5.1 Body composition

Body mass increased over training ($\chi^2_{(2)}$=13.18, *P*=0.001). The increase of 1.7 (9.1) kg between T1 to T3 (from 67.6 to 69.3 kg, *T*=7, *P*=0.003, r=-0.42) occurred by T2 (*T*=4, *P*<0.01, r=-0.49) (Figure 8.9). Estimated fat free mass also changed ($\chi^2_{(2)}$=29.52, *P*<0.01). The 0.9 increase (11.3) kg between T1 to T3 (from 59.1 to 61.0 kg, *T*=2, *P*<0.01, r=-0.53) also occurred by T2 (*T*=1, *P*<0.01, r=-0.54) (Figure 8.9). Estimated fat mass did not change over training ($\chi^2_{(2)}$=3.68, *P*=0.159) (Figure 8.10). Percentage body fat decreased ($\chi^2_{(2)}$=17.34, *P*<0.01) by 0.4 (7.2) % between T1 to T3 (from 12.7 to 12.3 %, *T*=2, *P*=0.003, r=-0.42). The adaptation had occurred by T2 (*T*=2, *P*=0.001, r=-0.46), but did not change thereafter (*T*=6, *P*=0.754, r=-0.05) (Figure 8.11).

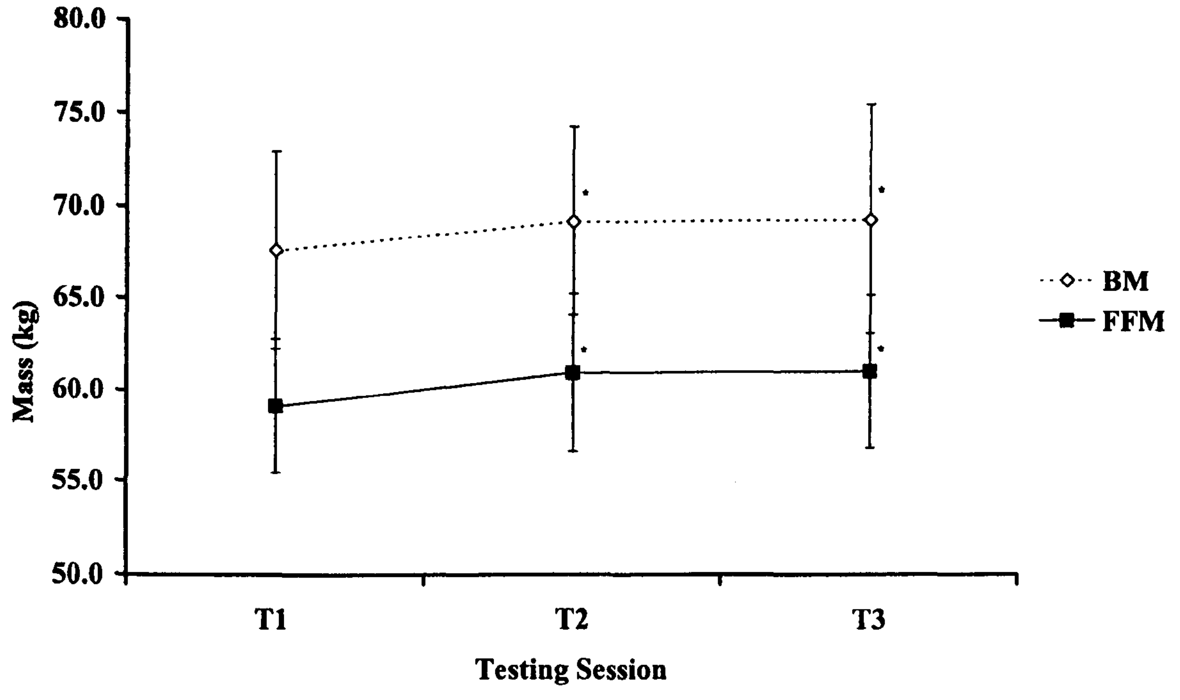


Figure 8.9. Body mass (BM) and estimated fat free mass (FFM) at T1, T2 and T3. * = significantly different from T1.

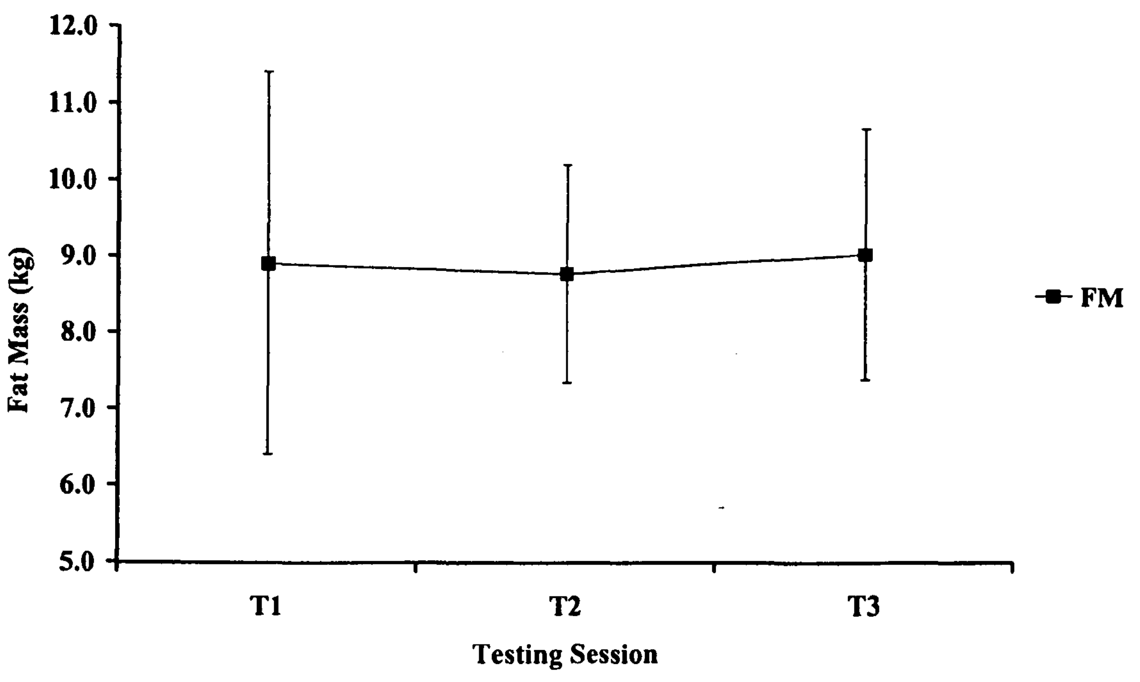


Figure 8.10. Estimated fat mass (FM) at T1, T2 and T3.

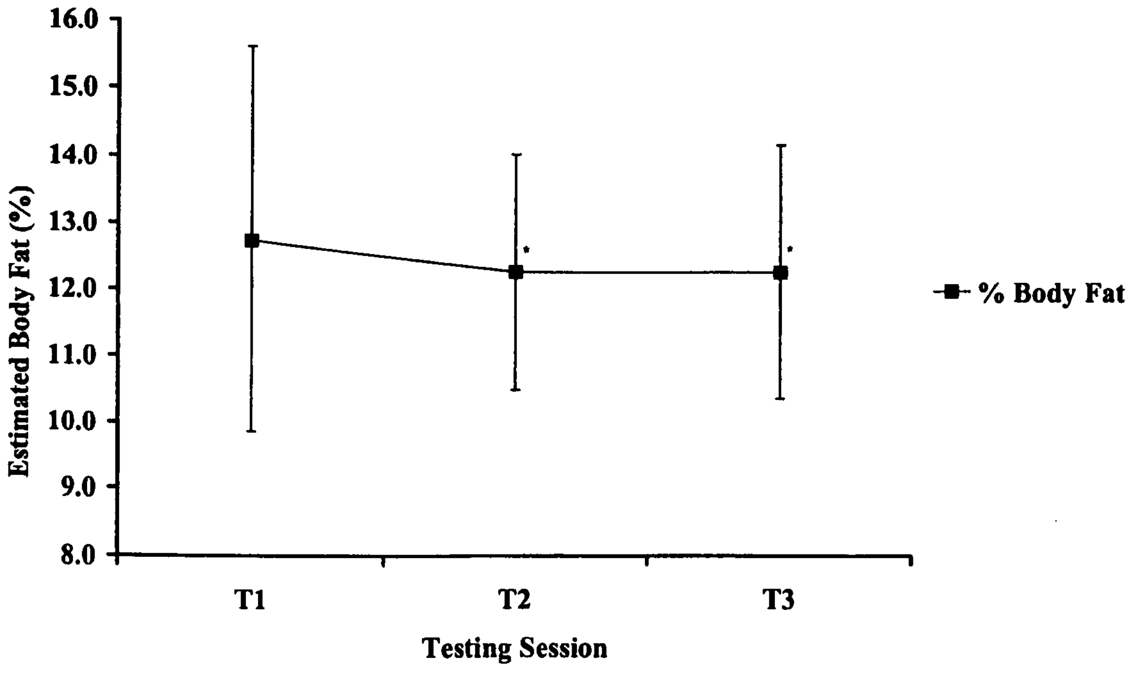


Figure 8.11. Percent body fat at T1, T2 and T3. * = significantly different from T1.

8.3.5.2 Skinfold measurements

The sum of 8-site and individual skinfold (SF) thicknesses are presented (Table 8.6). The sum of 8-site SF decreased over training ($\chi^2_{(2)}$=13.23, P=0.001). The decrease of 8.5 (88.6) mm between T1 to T3 (T=7, P=0.002, r=-0.42) occurred by T2 (T=5, P<0.01, r=-0.49), but did not change subsequently (T=9, P=0.258, r=-0.16).

With the exception of the triceps SF ($\chi^2_{(2)}$=3.02, P=0.221), all other SF sites decreased by T2 (P<0.05). One SF site (thigh) continued to decrease between T2 to T3 (T=7, P=0.009, r=-0.36). The mid-torso area (i.e. the iliac crest, supraspinale and abdomen) had the largest decreases in SF thickness. The total decrease for the iliac crest SF was 1.2 (26.7) mm (T=5, P=0.004, r=-0.40), supraspinale 1.0 (10.2) mm (T=3, P<0.01, r=-0.56), and abdomen 2.8 (20.8) mm (T=5, P=0.001, r=-0.48). These all occurred by T2 (Table 8.6).

Table 8.6. Skinfold thicknesses (mm) at T1, T2 and T3. (median (range)).

	T1	T2	T3
Sum of 8 SF	76.6 (103.4)	68.5 (73.0)*	68.1 (62.7)*
Triceps	9.7 (13.9)	10.1 (11.8)	10.0 (9.1)
Biceps	4.3 (6.1)	4.0 (4.4)*	4.2 (3.8)*
Subscapular	9.1 (8.0)	8.7 (6.5)*	9.0 (7.0)
Iliac Crest	12.4 (26.4)	11.2 (14.2)*	11.2 (13.4)*
Supraspinale	6.7 (15.3)	6.1 (10.2)*	5.7 (8.6)*
Abdomen	13.7 (21.1)	11.6 (15.4)*	10.9 (17.3)*
Thigh	11.0 (13.1)	10.2 (9.6)*	10.5 (9.1)†
Calf	8.7 (16.5)	8.9 (10.3)*	8.6 (9.7)

* = Significantly different from T1
† = Significantly different from T2

8.3.5.3 Circumferential girths

Adaptations in girth circumferences are presented (Table 8.7). Relaxed arm circumference changed ($\chi^2_{(2)}$=11.94, *P*=0.003). Gluteus circumference increased ($\chi^2_{(2)}$=13.39, *P*=0.001) 2.5 (6.0) cm between T1 to T3 (*T*=5, *P*=0.003, r=-0.41). The adaptation was evident by T2 (*T*=5, *P*=0.001, r=-0.45). Calf girth increased ($\chi^2_{(2)}$=8.67, *P*=0.013) 0.5 (1.8) cm over training (*T*=6, *P*=0.005, r=-0.39). Chest girth decreased over training ($\chi^2_{(2)}$=8.10, *P*=0.017), which was significant in the second half of training (*T*=6, *P*=0.001, r=-0.45). Both waist ($\chi^2_{(2)}$=2.58, *P*=0.275), and thigh ($\chi^2_{(2)}$=1.86, *P*=0.395) girth did not change over training.

Table 8.7.Girth circumferences (cm) at T1, T2 and T3. (mean (range)).

	T1	T2	T3
Relaxed Arm	28.1 (6.9)	28.2 (6.5)	28.1 (5.9)*†
Chest	89.2 (19.5)	88.5 (16.9)	87.9 (17.7)†
Waist	76.0 (18.3)	76.9 (14.5)	76.8 (14.8)
Gluteus	93.3 (18.1)	95.7 (16.9)*	95.8 (16.5)*
Mid-Thigh	50.6 (9.5)	49.9 (10.4)	49.2 (10.5)
Calf	35.8 (7.8)	36.0 (8.2)	36.3 (8.0)*

* = Significantly different from T1
† = Significantly different from T2

8.4 Discussion

8.4.1 Main findings

A 7.0% improvement was observed in 2.4 km loaded marching performance over infantry training. Concurrently, the extent with which the determinants of loaded marching performance (Chapter 7) improved would have influenced the magnitude of the observed improvement in loaded marching performance. Of these determinants, an improvement of 9.6% was observed for loaded marching economy, and 3.6% for 2.4 km running performance. However, peak static lift strength decreased by 3.2%.

8.4.2 The model of loaded marching performance

The model of 2.4 km LMP developed in Chapter 7 reported LME, 2.4 km run time and peak SLS to explain 65% of the variance in LMP. Independently, LME explained the most (22%), 2.4 km run time the second most (11%), and peak SLS the least (6%) amount of variance. The relative importance of these measures in determining LMP coupled with the magnitude these determinants respond to a training programme will establish whether (or not) a ceiling effect in the training adaptation of LMP has occurred, and which aspects of physical fitness to prioritise in subsequent PT programmes that are designed to further improve LMP.

8.4.2.1 Loaded marching performance

Loaded marching performance improved by 7.0% by T3, which can be classified as a *large* effect size (Cohen, 1992). Most of the training response (i.e. 51 s of the 63 s improvement) occurred in the second half of training (T2 to T3). This was an unexpected finding that was not consistent with the general pattern of adaptation in this study (i.e. a steep improvement between T1-T2, followed by a slowing down or reversal between T2-T3). As the volume of loaded marching was similar in the latter half of training (56.8 km) compared to the initial half (49.6 km), an increase in the intensity of LM specific training in the second half of training may explain this anomaly. The higher LM intensity during the second half of training (Figure 8.4) was a function of both speed (5.3 km·h^{-1} to 6.4 km·h^{-1}) and load increases (20 kg to 25 kg). Furthermore, it was not until week 11 that the load carried during PT lessons equated to, and then exceeded that carried in the criterion LMP test (20 kg). As both external load, and to a greater extent ambulatory speed increase the metabolic demand of a loaded march task (Christie and Scott, 2005), the progression in LM intensity between T2 to T3 may explain the continued improvement in LMP in the second half of training. This is feasible for two reasons. First, no other measures further improved in the second half of training, despite the progressive nature of the PT programme. Second, high intensity LM training is twice as effective as duration based LM training at improving LMP (Visser *et al.*, 2005).

One other study has monitored LMP at similar time points (i.e. beginning, middle and end of training) to the present study (Harman *et al.*,1998). Here, a military style training programme (modified with a resistance training intervention) elicited a 23.8% improvement in 3.2 km LMP over 24 weeks of training. In contrast to the present study most of the improvement (18.5%) occurred in the first 14 weeks of training and subsequently slowed. However, there was a different training modality (concurrent strength and endurance training), and subjects acted as volunteers (female civilians with no baseline fitness requirements). As initial fitness level is inversely related to the magnitude of a training adaptation (Vogel *et al.*, 1978; Trank *et al.*, 2001), the 23.8% improvement in LMP appears to represent the upper limit of adaptation in LMP, that previously trained individuals are unlikely to achieve. In a comparative population to the present study, a 15.7% (Williams *et al.*, 1999), and 16.7% (Williams *et al.*, 2002) improvement in 3.2 km (25 kg) LMP was reported, following 10-weeks of British Army recruit training. Whilst this was a two-fold greater training response in the LMP than the present study, a LM test with 15 kg in these same studies only improved LMP by 3.6% (Williams *et al.*, 1999) and 8.9% (Williams *et al.*, 2002). The generally smaller training response in the present study may be explained by the way the LMP test was conducted. As over 60% of the subjects were loaded march naïve, the first half (1.2 km) of the LMP test was conducted as a 'squadded' effort, and it was not until the second half of the test that subjects were instructed to give an individual 'best effort'. It was believed this strategy may have reduced the pacing / learning effect. Thus, if the subjects had been permitted to give an individual 'best effort' from the start of the 2.4 km LMP test (similar to the studies cited), it is possible that larger improvements would have been observed.

8.4.2.2 Loaded marching economy

In the model of LMP, LME independently explained the largest amount of variance in LMP (Chapter 7). This was a novel finding that highlighted the importance of exercise economy as a determinant of exercise performance, which is common knowledge in other exercise modalities (Joyner, 1991; Coyle, 1995). Absolute and relative LME improved by 9.6% and 14.0%, respectively, which was a *large* effect size (Cohen, 1992) (Table 8.2). In contrast to LMP, most of this change had occurred by T2. The range in LME values was 13.4 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ (from 27.1 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ to 40.5 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$), which was larger than the 8 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ that is said to differentiate *good* from *poor* running economy (Wood, 1999). As such this is the first study to suggest that a range in sub-maximal oxygen demand of ~13 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ can differentiate *good* from *poor* LME, within a population of infantry recruits. Furthermore, this spread in LME (in part) accounted for the 4 minute range in 2.4 km LMP.

As the relative exercise intensity of sub-maximal LM was below 80-85%$\dot{V}O_2max$, reliable estimates of tissue CO_2 production can be assumed (hence a stable bicarbonate pool) (Jeukendrup and Wallis, 2005). Thus, the energy cost of LM could be derived from the non-protein RQ (Zuntz, 1901 cited in McCardle *et al.*, 2001). This calculation revealed the caloric cost of sub-maximal LM (which reflected the minimum intensity required to pass the week-21 BCFT) decreased over training. When these data were extrapolated to a duration equivalent to the BCFT (i.e. 2 hours) the caloric cost of the task decreased from 1378 to 1235 kcals between T1 to T3. Such a marked decrease in energy expenditure (~10%) could substantially improve aerobic capacity (i.e. physical endurance) and reduce the impact of fatigue during more prolonged activities such as field exercises and/or operations.

Five classifications of relative LM exercise intensity have been proposed (Christie and Scott, 2005). In accordance with these classifications the intensity of the LM task at T1 would have imposed a 'very heavy' physical demand on the subjects over extended march distances (e.g. BCFT). However, the physiological adaptation in LME decreased the intensity classification to 'tolerable–heavy' limits, which coincides with the 100% pass rate in the 12.8 km BCFT in the present study. This confirms that the British Army selection standards (PSS(R)) that predict physical output standards at the end of recruit training (i.e. a BCFT pass), and the process of infantry training appear to work.

Just two studies have monitored changes in sub-maximal exercise economy over the duration of military training programmes, and only one of these investigated changes in LME. Running economy improved by 5.5% at moderate exercise intensities, but not at higher intensities (80-85%$\dot{V}O_2$max) over 6 weeks of military training (Daniels *et al.*,1979). More specifically, an automated gas analysis system revealed LME improved 5.9 % (~1 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$) over 8 weeks of military style training (Gutekunst *et al.*, 2006). The LME adaptation in the present study (4.9 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$) was markedly greater than these previous observations. The longer duration of infantry training alone in the present study does not explain the reason for this difference, as the majority of adaptation in LME took place in the first 11 weeks of training, which was just 3 weeks longer than the training programme of Gutekunst *et al.*, (2006). The 4.9 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ improvement in LME was also larger in magnitude than similar exercise modalities such as running over; short (6-8-week) (Franch *et al.*, 1998; Billat *et al.* 1999; Jones *et al.*, 1999; Spurrs *et al.*, 2003), medium (14-18-week) (Sjödin *et al.*, 1982), and long training periods (up to 1-year) (Conley *et al.*, 1984; Svedenhag and Sjödin, 1985; Berg *et al.*, 1995). Whilst an improvement of 4.9 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ was a comparably large improvement, this may not be a large enough training adaptation to make the transition from *poor* to *good* LME (i.e. a change of ~13 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$). Although, some individual responses did approach this magnitude of adaptation (7.5 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$).

According to Cohen (1992), the effect size of LME was *large*, however the complex nature of exercise economy means that this adaptation may not be solely attributable to physiology, but could also be attributable to non-physiological factors (Saunders *et al.*, 2004), which are discussed.

The physiological mechanisms responsible the improvement in LME are alluded to (Table 8.2). Over training $\dot{V}_E$ and F_EO_2 decreased 24.8% and 2.6%, respectively. This is consistent with running economy training responses (Franch et al., 1998). In addition, circulating whole blood lactate concentration [BLa] and heart rate decreased 58.9%, and 14.5%, respectively. The improvement in LME (a decrease in calculated $\dot{V}O_2$) was coincident with a decrease in F_EO_2 and minute ventilation, which are two of the three primary variables used to calculate $\dot{V}O_2$ (Howley *et al.*, 1995). The decrease of F_EO_2 suggested a greater fraction of inspired O_2 (20.93%) was perfused and/or extracted at the alveolar and/or tissue levels, and may have afforded the decrease in minute ventilation.

Physiologically the body adapts centrally (i.e. oxygen delivery), and peripherally (i.e. oxygen extraction to produce and use ATP) in response to an endurance training intervention (Laursen and Jenkins, 2002). Adaptation at the periphery is led by metabolic adaptations (Green *et al.*, 1991), where increases in intra muscular glycogen, and an attenuation in [BLa], adenosine diphosphate (ADP), adenosine monophosphate (AMP), and inorganic phosphate (P_i) precede mitochondrial enzymatic adaptation (Green *et al.*, 1972). A reduction in these products is important as they are known to stimulate glycogenolysis and the glycolytic rate-limiting enzyme (phosphofructokinase - PFK), thereby leading to a reduction in glycolytic flux (Houston, 2001). As per the present study the attenuation in [BLa] has been reported over training interventions (Hurley *et al.*, 1984; Acevedo and Goldfarb, 1989). Lower circulating [BLa] have been attributed to a reduced rate of La appearance at low intensities, and increased rate of clearance at higher intensities (MacRae *et al.*, 1992). Furthermore, lower levels of glycogen depletion, and increased *β*-oxidation are concomitant sub-maximal adaptations to endurance training (Holloszy and Coyle, 1984), however this could not be implied in the present study as RER did not significantly decrease. Other adaptations that may explain the physiological mechanisms of adaptation in LME include changes in mitochondrial density (Holloszy and Coyle, 1984), a greater O_2 extraction due to an increased a-vO_2 difference (Henriksson, 1977; Van Handel *et al.*, 1976), and an increased muscle capillary density, succinate dehydrogenase (SDH) and cytochrome oxidase (CO), and fibre type response (Andersen and Henriksson, 1977; Ingjer, 1979).

The non-physiological factors of scaling, anthropometry, kinematics, kinetics, flexibility, and ground reaction forces can explain changes in exercise economy (Saunders *et al.*, 2004). In the present study the non-physiological factors of training specificity (i.e. motor learning), scaling (i.e. changes in body mass), and the condition of issued boots (i.e. kinematics) may partly explain the marked changes observed in LME:

Specificity and overload are the two most important principles of training (Tanaka, 1994). In the present study over 60% of the subjects were loaded-carriage naïve, and exposed to 15 periods (49.6 km) of LM by week 11 (T2) (Figure 8.4). Neuromuscular adaptations occur with training, which improve the central nervous system's ability to co-ordinate movements and develop force (Carroll *et al.,* 2001). Thus it is possible some form of motor-learning occurred in the present study that may have led to an optimisation of gait pattern during the LM task. In running tasks, experienced individuals intuitively adopt an optimal stride length (Heinert *et al.,* 1988), which may also be the case during LM. Although between T1 and T2 stride frequency (analogous to stride length) did not change, the range of stride frequency values narrowed over training (13.0 to 8.0 strides·min^{-1}). This suggested the group became more homogenous in their LM gait pattern, possibly shifting towards an optimal gait. In addition, it emerged in Chapter 4 that following two familiarisation sessions there was an apparent learning effect in LME of ~2-3% (that may have been kinematic in nature) and reflect changes in gait in the present study.

The second non-physiological factor that may explain the marked change in LME are changes in body mass. Body mass significantly increased by 2.4% by T2. The use of ratio standards to normalise $\dot{V}O_2$ (i.e absolute $\dot{V}O_2$ ÷ body mass = $\dot{V}O_2$ in ml·kg^{-1}·min^{-1}) has been criticised, as body mass changes cause an artefact change in $\dot{V}O_2$ (Winter and Nevill, 2001). Assuming absolute $\dot{V}O_2$ remained constant, an increase in body mass will decrease normalised $\dot{V}O_2$ and *vice versa* (Berg, 2003). Thus it is possible that the 4.9 ml·kg^{-1}·min^{-1} (14.0%) decrease in LME may be partly attributed to an increase in body mass. For example LME was 2280 ml·min^{-1} at T1, which equated to a normalised $\dot{V}O_2$ of 33.7 ml·kg^{-1}·min^{-1}. When normalised to the T3 body mass (2280 ml·min^{-1} ÷ 69.3 kg) $\dot{V}O_2$ decreased to 32.9 ml·kg^{-1}·min^{-1}. Thus, 0.9 ml·kg^{-1}·min^{-1} of the adaptation was due to changes in body mass and not qualitative structural / functional changes. For this reason it may be more appropriate to refer to the $\dot{V}O_2$ training response in absolute, not relative units of measurement.

The final potential non-physiological cause for the marked improvement in LME could be attributed to the British Army issue boots worn. At T1 boots had been issued 1 to 4 days prior to testing, and were not fully 'broken in'. At T2 and T3 these boots were 'broken in', which may have produced changes in gait kinematics, thereby contributing to the improvement in LME.

The 9.6% improvement in absolute LME was similar to the 7.0 % improvement in LMP. Others have reported the adaptation of LME to military style training programmes (Gutekunst *et al.*, 2006), however this is the first study to concurrently track the adaptation of LME and LMP. Exercise economy is an important determinant of endurance performance (Joyner, 1991; Coyle, 1995; Brandon, 1995; Ingham *et al.*, 2008), which applies to LM (Chapter 7). However, establishing a targeted LME PT intervention to improve LMP would be both, challenging (due to the complex interaction of physiologic, anthropometric and biomechanic sub-disciplines (Morgan *et al.*, 1989; Anderson, 1996; Saunders *et al.*, 2004)), and unwarranted (based on the current magnitude of the LME response to training).

8.4.2.3 Running performance

The 2.4 km run time was an influential determinant of 2.4 km LMP. Mathematically a 10 % improvement in 2.4 km run time would improve LMP by 47 ± 4 s, which was a similar effect LME had upon LMP (Chapter 7). The actual improvement over training was 3.6%, which was a *medium-large* effect size (Cohen, 1992), and smaller than previously reported during British Army recruit training (Harwood *et al.*, 1999; Rayson *et al.*, 2002c; Wilkinson *et al.*, 2003; Carter *et al.*, 2006). One of these studies suggested that even larger improvements of 5.6% were sub-optimal (Rayson *et al.*, 2002c). Furthermore, consistently larger improvements of 8.9% in Officer Cadets (Harwood *et al.*, 1999), 10.0% in infantry Foot Guards (Carter *et al.*, 2006), and 5.8% in Parachute Regiment recruits were reported (Wilkinson *et al.*, 2003). Data from Carter *et al.*, (2006) are interesting as the only aspect of infantry training (on paper) that differentiated the infantry Foot Guard from the Line Infantry recruits in the present study were 2 additional weeks towards the end of Foot Guards training where drill was practiced. The comparatively poor training response in the present study was similar (2.9%) to US Navy recruits that were classified with excellent-superior 2.4 km run times at baseline (09:42 min:s) (Trank *et al.*, 2001), but was worse than individuals with 'good' (10:42 min:s), and 'poor-fair' (12:15 min:s) 2.4 km run times.

Part of the reason for the small improvement in 2.4 km run time may have been the similarly small improvement in absolute $\dot{V}O_2max$. Middle distance events (i.e. 2.4 km) are performed at an intensity that approximates the velocity attained at $\dot{V}O_2max$ ($v\dot{V}O_2max$) (Billat, 2001), hence $\dot{V}O_2max$ is central to middle distance events (Foster, 1983). Absolute $\dot{V}O_2max$'s influence on LMP is also well documented (Harman and Frykman, 1995; Rayson *et al.*, 1993, 2000, 2002a, 2002b, Pandolf *et al.*, 2002), and some consider it the most important determinant of LMP when carrying heavier loads (Rayson *et al.*, 2000). However, the fact that the improvement in 2.4 km run time (3.6%) was almost proportional to $\dot{V}O_2max$ (2.6%) was surprising given that the 2.4 km run not only reflects improvements in $\dot{V}O_2max$ but also, anaerobic threshold, running economy, run velocity (Brandon, 1995), as well as pacing strategy. This observation can be illustrated with the composite measure of $\dot{V}O_2max$ and running economy otherwise known as the velocity attained at $\dot{V}O_2max$ ($v\dot{V}O_2max$)) (Jones and Carter, 2000). This measure is regarded as the strongest correlate of endurance running performance (Noakes *et al.*, 1990). Here the 6.6% increase in $v\dot{V}O_2max$ at T2 was not accompanied by increases in peak $\dot{V}_E$, peak [BLa] and peak HR, indicating a lower metabolic demand (hence an improvement in running economy) at the same absolute intensity (Table 8.3). Previously, absent increases in $\dot{V}O_2max$ were accompanied by improvements in other measures of running fitness such as, running economy and onset of blood lactate accumulation (Sjödin *et al.*, 1982), anaerobic threshold (Denis *et al.*, 1982), 10 km run time and run to exhaustion (Acevedo and Goldfarb, 1989), and lactate threshold, $v\dot{V}O_2max$, running economy and 3 km run time (Jones, 1998). Thus, the $\dot{V}O_2max$ may not be the most sensitive measure to track the efficacy of a training intervention (Daniels *et al.*, 1979).

Absolute $\dot{V}O_2max$ increased 2.6% by T2 (Table 8.3), which was a *medium-large* effect (Cohen, 1992). Military PT is often conducted as a single group, hence the same absolute stress causes the least fit individuals to work at a relatively higher intensity compared to the most fit (Rayson *et al.*, 2002c). This results in an inversely proportional increase in $\dot{V}O_2max$ in relation to baseline fitness (Vogel *et al.*, 1978; Stacy *et al.*, 1982; Gordon *et al.*, 1986). However, the *weak* associations between initial $\dot{V}O_2max$ and the magnitude of the training response observed in the present study did not support this (Figure 8.8).

$\dot{V}O_2$max did not increase during 10 weeks of South African Defence Force training, where the total volume of run training was 160 km (3.2 km, 5 days per week, 10 weeks) (Gordon *et al.*, 1986a), some 2.5-fold greater than that recommended (Knapik *et al.*, 2006). With a similar volume of running (one 30-minute run, 5 days per week (~168 km)), 6 weeks of U.S. Military Academy training also reported no increase in $\dot{V}O_2$max, even when subjects were streamed into 3 ability groups (Daniels *et al.*, 1979). In this latter study an improvement in running economy and decrease in sub-maximal [BLa] occurred despite no increase in $\dot{V}O_2$max. The total running mileage of these studies was far greater than the 8 run sessions (49.6 km), over the 21 weeks of PT in the present study (Figure 8.5). However, it should be noted there was likely to be significant aerobic training stimulus in the present study from other exercise modalities such as loaded marching, battle PT, circuit training, sports, and RMTs / PFTs, which may have been responsible for the development of absolute $\dot{V}O_2$max (Figure 8.2).

It may be unrealistic to increase $\dot{V}O_2$max during military training programmes, to the same extent as civilian training studies. First, this may be due to the higher initial training status compared some civilian studies (Hickson *et al.*, 1977; Hardman *et al.*, 1986). Second, military training is physically and mentally arduous, therefore the training adaptation will be sub-optimal, due to inadequate periods of recovery (McCafferty and Horvath, 1977). This was apparent in 8 weeks of U.S. Army Ranger training where total testosterone and insulin-like growth factor-I decreased, cortisol increased, and decreases in body mass, explosive power and maximum strength were observed (Nindl *et al.*, 2007). Although this (in part) was different to the present study, where an increase in body mass was observed.

In addition, $\dot{V}O_2$max is optimally developed by training at exercise intensities that approximate $\dot{V}O_2$max (Billat, 2001; Midgley *et al.*, 2006), as this challenges the factors that limit oxygen transport (Wagner, 1996). Although contemporary hypotheses on the limitations to $\dot{V}O_2$max are proposed (Noakes, 1988; 1997; 1998), the classical argument cites the rate of oxygen delivery (e.g. maximal cardiac output) as the major limiting factor to $\dot{V}O_2$max (Bassett and Howley, 1997; 2000; Wagner, 2000). As the intensity of the PT was not monitored in the present study, the adequacy of the exercise intensity during PT for the development of $\dot{V}O_2$max is presently unknown.

The sub-optimal improvements in 2.4 km run time and absolute $\dot{V}O_2max$ during infantry training, coupled with the marked influence this aspect of fitness has on LMP (Chapter 7; Rayson *et al.*, 2000; Pandorf *et al.*, 2002; Rayson *et al.*, 2002a; Williams and Rayson, 2006; Griggs *et al.*, 2008) would suggest that an intervention to develop 2.4 km run time should be prioritised in the future. As the 2.4 km run will be conducted at $v\dot{V}O_2max$ (Billat, 2001), and training at an intensity that approximates $\dot{V}O_2max$ is the optimal intensity to develop $\dot{V}O_2max$ (Midgley *et al.*, 2006), a logical intervention would be the implementation of high intensity interval training (HIIT), which will ensure adequate training time is spent at this intensity to elicit the maximum training response (Laursen and Jenkins, 2002). Currently, structured HIIT is absent from infantry PT (section 8.2.4), and the inclusion of this training technique may further improve LMP. However, on a cautionary note, the potential performance benefits of a shift to higher intensity PT may be offset by the possible increased risk of musculo-skeletal injury.

8.4.2.4 Strength

There was a general lack of strength development in the present study (Table 8.4), which may be explained by the absence of strength based resistance training in the current infantry PT programme (Figure 8.2). One exception to this observation was average and peak dynamic leg strength that increased 14.3% and 7.3%, respectively (Table 8.4). Based on the fact that peak leg strength is typically attained on repetition number 8 out of 10 (section 3.4.2.3; Figure 3.9), and 95% limits of agreement revealed a systematic bias in leg press repeatability of up to 6% (Table 6.4), the increase in leg press strength over infantry training should be viewed with caution as this may partly be due to a learning effect.

Considering, peak SLS emerged as a determinant of LMP (Chapter 7), and isometric and dynamic indices of strength are known to influence LMP (Mello *et al.*, 1988; Knapik *et al.*, 1990; Rayson *et al.*, 2002a, 2002b; Williams and Rayson, 2006) the generally poor magnitude of strength development in the present study may have blunted the LMP training response. One solution would be to conduct concurrent strength and endurance training, which seems to be more effective at improving LMP (Kraemer *et al.*, 1987, 2001, 2004). However, Chapter 7 calculated that a 10% increase in SLS mathematically improved LMP by a markedly smaller amount (8 ± 6 s), compared to LME (42 ± 5 s), and

2.4 km run time (47 ± 4 s). Thus, assuming the LMP model was robust over infantry training, and SLS had substantially increased in the present study, the subsequent improvement in LMP may have been modest.

During the PT programme 25 periods of circuit training were the closest type of training to a strength stimulus. These lessons relied upon an individual's body mass to provide a resistance overload. It is suggested that in order to develop strength, a resistance equivalent to 75% to 100% of repetition maximum needs to be overcome, which is considerably greater than resistances of 25% to 65% that are typically overcome in circuit training lessons (Bompa, 1999). The muscular endurance emphasis here may explain the increase in the maximal number of sit-ups performed (Table 8.4). The *moderate* to *strong* inverse associations between initial strength and the magnitude of strength change suggested that there was variation in strength responses to current infantry training (Table 8.5). This indicated at baseline weaker individuals improved more over training compared to stronger individuals.

In agreement with the present study, a military training programme absent of structured resistance training reported just 1 out of 8 upper and lower body measures of dynamic strength to increase over 10 weeks of South African Defence Force training (Gordon *et al.*, 1986b). Furthermore, isokinetic elbow flexion, and knee extension strength did not increase during 12 weeks of standard Royal Artillery training (Legg and Duggan, 1996). Some (Vogel *et al.*, 1978; Legg and Duggan, 1996; Brock and Legg, 1997; Williams *et al.*, 1999; 2002), but not others (Brock and Legg, 1997; Harwood *et al.*, 1999) also reported no increase in isometric strength during military training, which is not surprising given isometric tests lack sensitivity at detecting improvements in dynamic strength training programmes (Rutherford and Jones, 1986). However, there is a suggestion that if structured resistance training programmes are implemented into military/military style training strength increases in, one-repetition maximum (Knapik and Gerber, 1996), dynamic military specific strength (Harman *et al.*, 1998), and job specific measures of material handling strength can occur (Williams *et al.*, 2002).

The present study highlighted that strength development is a current weakness of British Army infantry training. Even if strength had increased a negligible change in LMP should have been observed (Chapter 7). Furthermore, the influence of strength development on

LMP is far from conclusive. Twelve weeks of concurrent strength and high intensity endurance training have improved LMP by up to 14%, compared to just 4% and 0% in resistance only, and high intensity endurance only controls, respectively (Kraemer *et al.*, 1987; 2004). However, the small sample size (n=8/9) reduced the generalisability of these findings. Elsewhere mixed findings have been reported. Both marked (24%) (Harman *et al.*, 1998; Reynolds *et al.*, 2001), and modest improvements (4%) (Knapik and Gerber, 1996) in LMP coincided with substantial increases in strength, in females. In both instances the absence of control groups makes the interpretation of these mixed findings difficult. Furthermore, 3.2 km LMP was 5.3% faster in resistance training compared to non-resistance training groups during 15 kg LMP, but no difference was reported between-groups during a 25 kg LMP task during initial military training (Williams *et al.*, 2002).

It has been suggested that LMP training response might be optimised if the PT modality emphasised (e.g. aerobic fitness or resistance) corresponded with the physical deficiency of an individual (Williams *et al.*, 2004). For example, individuals with good aerobic fitness and poor strength should be prescribed a training programme that emphasises resistance training, to improve LMP. Although in such a scenario the LMP model (Equation 7.2) suggested changes in peak SLS have little influence on LMP. As empirical observations lack clarity on the efficacy of resistance training for LMP, further investigation in this area is warranted.

8.4.2.5 Anthropometry

Anthropometric measures have been associated with LMP (Frykman and Harman, 1995; Rayson *et al.*, 2000, Pandorf *et al.*, 2002, Williams and Rayson, 2006). This aspect of fitness had *weak-moderate* correlations with 2.4 LMP in Chapter 7, and were not included in the LMP model. Of note, body mass significantly increased 2.5%, which was a *medium-large* effect (Cohen, 1992). Such an adaptation would have contributed to the marked decrease in LME (relative $\dot{V}O_2$), and the absent increase in relative $\dot{V}O_2max$ in the present study (Berg, 2003). The increase in body mass was explained by decreases in estimated percentage body fat that were offset by increases in estimated fat-free mass. The individual SF sites (Table 8.6) demonstrated that males undergoing infantry training predominately lose sub-cutaneous fat from the mid-torso area (i.e. iliac crest, supraspinale and abdomen). Upper body circumferential girth measures were unchanged or decreased, whereas lower body measures increased over training. The directionally different loss in percentage body fat and gain in estimated FFM make the cause of the change in circumferential girths speculative (Table 8.7). However, assuming the lower body experienced an increase in fat-free mass; this may in part explain the significant increase in leg strength observed in the present study (Narici *et al.,* 1989).

In conclusion the physical and physiological trend of adaptation during British Army infantry training was characterised by an improvement in the first half of training that slowed down or returned towards baseline values during the second half of training. Loaded marching performance and loaded marching economy underwent large improvements, however, loaded marching performance may have improved more if other aspects of physical fitness had developed adequately (i.e. 2.4 km run time, $\dot{V}O_2$max and strength). It is suggested that the current physical training programme during British Army infantry training is sub-optimal at developing these aspects of fitness. Thus, structured high intensity interval training and strength-based resistance training should subsequently be implemented into infantry physical training in an attempt to further improve loaded marching performance.

CHAPTER 9

THE PHYSICAL AND PHYSIOLOGICAL RESPONSES TO A MODIFIED PHYSICAL TRAINING PROGRAMME DURING 26 WEEKS OF BRITISH ARMY LINE INFANTRY TRAINING.

9.1 Introduction

The development of occupational fitness to a minimum standard commensurate with Career Employment Group (CEG) requirements ensures British Army recruits receive appropriate physical preparation prior to joining the Field Army. Loaded marching is one example of an essential Representative Military Task (RMT) that must be developed during recruit training. Previously, we demonstrated that both accepted (i.e. 2.4 km run time and static lift strength), and novel factors (i.e. loaded marching economy) best explain LMP (Chapter 7). However, maximal running performance (i.e. 2.4 km run time), $\dot{V}O_2$max, dynamic strength, and isometric strength improved little or not at all during infantry training (Chapter 8). Thus, if physical training (PT) was modified to specifically develop these fitness parameters, further improvements in LMP may be achieved.

Various modifications to PT programmes have been suggested for the development of LMP. The manipulation of LM training alone may improve LMP (Knapik *et al.*, 1990), though this may lead to an increase in injury incidence (Visser *et al.*, 2005). Concurrent resistance and endurance training was proposed as the most effective training prescription for LMP (Kraemer *et al.*, 1987; 2001; 2004), which was both substantiated (in 15 kg LM) and unsubstantiated (in 25 kg LM) (Williams *et al.*, 2002). It appears that the upper ceiling of LMP improvement, from a concurrent resistance and endurance training programme is 24% (Harman *et al.*, 1998).

A novel method of PT prescription for LMP has been suggested (Williams *et al.*, 2004). Here the deficient components of fitness (e.g. endurance or strength) were calculated by the size of the multi-stage shuttle run test-to-isometric strength ratio, on an individual basis. A low ratio value (e.g. 0.55) would indicate a relative deficiency in aerobic fitness, whereas a high ratio value (e.g. 0.64) indicated a deficiency in strength. Therefore, PT emphasised the area of relative physical deficiency to optimise the development of LMP.

Within the present thesis it appears that the development of LME may be one of the most effective strategies to improve 2.4 km LMP (Chapter 7). However, the already large response of LME to infantry training (Chapter 8), and the complex interaction of physiological, anthropometric and biomechanical measures underpining exercise economy (Saunders *et al.*, 2004) makes such an intervention both unnecessary, and unfeasible. Other key determinants of LMP (i.e. 2.4 km run time and $\dot{V}O_2$max) (Chapter 7) responded less well to infantry training (Chapter 8), hence a PT intervention should prioritise the development of these measures to further improve LMP.

Aerobic fitness is determined by $\dot{V}O_2$max, exercise economy, the velocity at $\dot{V}O_2$max (v$\dot{V}O_2$max), lactate / ventilatory threshold, oxygen uptake kinetics (Jones and Carter, 2000), and the maximal sustainable percentage of $\dot{V}O_2$max (Wood, 1999). Whilst the recommended dose of physical activity to develop and maintain general cardio respiratory fitness is broad (i.e. 3-5 $d \cdot week^{-1}$, of 20-60 minutes, between 50-85% Heart Rate Reserve (HRR)) (ACSM, 1998), the 'optimal' intensity for increasing aerobic fitness is narrow (i.e. 90-100% of $\dot{V}O_2$max) (Wenger and Bell, 1986). To ensure sufficient time is spent at this optimal intensity to create adequate training overload repeated, 10-second to 5-minute bouts of exercise above the anaerobic threshold should be performed (Laursen and Jenkins, 2002). This is commonly referred to as High Intensity Interval Training (HIIT). Aerobic HIIT can be classified as, short (e.g. velocities > v$\dot{V}O_2$max), long (e.g. velocities around v$\dot{V}O_2$max), and very long (e.g. velocities < v$\dot{V}O_2$max) (Midgely *et al.*, 2006). As middle distance running events (e.g. 2.4 km run) are typically performed at the v$\dot{V}O_2$max, HIIT around this speed will benefit from both an optimal central (i.e. cardiovascular) and peripheral (i.e. muscle recruitment) stimulation (Billat, 2001b). Alternatively, very long intervals conducted between the velocity at maximum lactate steady state and the v$\dot{V}O_2$max result in the manifestation of the $\dot{V}O_2$ slow component, which is responsible for further increasing $\dot{V}O_2$ towards $\dot{V}O_2$max (Gaesser and Poole, 1996). Thus, whilst very long intervals may achieve the same central stimulation as short or long intervals, the quality of the peripheral stimulation (i.e. muscle recruitment) will be reduced (Laursen and Jenkins, 2002).

The absolute external loads applied during most LMP tests means that larger individuals carry a smaller percentage of body mass. Whilst a larger body mass is a disadvantage during unloaded running tasks (Vanderburgh and Mahar, 1995), it is not during load-carriage tasks (Harman and Frykman, 1995; Rayson *et al.*, 2002a). It is possible that the importance of strength to LMP (Harman and Frykman, 1995; Rayson *et al.*, 2002a, 2002b; Williams and Rayson, 2006) simply reflects the inter-relationship between a biological system's structure dictating its function (Jones and Round, 1990), as muscle cross-sectional area is positively correlated with muscular strength (Maughan *et al.*, 1984). Resistance training (RT) is a well established training technique that through hormonal, metabolic, and mechanical stimuli (Jones *et al.*, 1989) elicits muscle hypertrophy (i.e. structure) and improved neural drive that collectively increase strength (i.e. function) (Narici *et al.*, 1989).

The manipulation of acute RT variables (i.e. muscle action, load size, volume, repetition velocity, exercise selection, frequency, and rest) in a PT programme will influence the training outcome (i.e. muscular endurance, hypertrophy, maximal strength, or power) (Bird *et al.*, 2005). In healthy populations strength development requires the participation in two resistance training sessions per week, at an intensity of 8 – 12 repetition maximum (RM) (ACSM, 1998). In order to specifically target strength development a relative load intensity of 75% to 100% of one repetition maximum should be lifted (Bompa, 1999). To date few studies have quantified the influence a structured RT programme has to LMP improvements in military recruits undergoing initial training (Williams *et al.*, 2002, 2004).

It was therefore the purpose of this study to implement a modified physical training programme that prioritised the development of $\dot{V}O_2max$ and 2.4 km running performance (through high intensity interval training), and strength (through resistance training), with the aim of further improving loaded marching performance in British Army recruits during infantry training.

9.2 Method

9.2.1 Subjects

Ninety-two new entry recruits (n=92), forming the Prince of Wales (PoW) 3 and 4 Platoons undertook the recently extended 26-week CIC at ITC(C). All recruits volunteered to participate as subjects in the study, which was approved by the Ministry of Defence Research Ethics Committee (MOD REC). Upon presentation at the Initial Medical Examination a total of six recruits were excluded from the study (four were deemed unfit for army service and medically discharged, one chose not to participate in the study but continue Army training, and one had delayed doctors notes hence temporarily could not conduct physical activity). Thus n=86 commenced data collection.

To control for the time of year both Platoons began training 2 weeks apart. The PoW 3 Platoon (n=46) were assigned to the control group, and the PoW 4 Platoon (n=40) were assigned to the modified (i.e. experimental) group. The control group (CON) underwent standard infantry training, whilst the modified group (MOD) completed an altered Physical Training (PT) programme (section 9.2.4.1; 9.2.4.2). Other than the PT programme, all other aspects of training were programmed to be identical between Platoons. The physical and racial characteristics of the subjects that successfully completed the research trial are presented (Table 9.1). All analyses were conducted on the 32 CON, and 33 MOD subjects.

9.2.2 Experimental design

The effect of Group (i.e. control vs. modified) and Time (i.e. beginning, middle and end of training) constituted the mixed experimental design. The effect of Group consisted of between-subject comparisons at each of the three time points during training (T1, T2, T3). The effect of Time consisted of within-subject comparisons at weeks 1/2 (T1), 12/13 (T2), and 21/22 (T3) of training. On each test occasion subjects performed an identical test battery that underwent interrogation in Chapters 4, 5 and 6, and application in Chapters 7 and 8.

Table 9.1. Subject characteristics (median (range)).

Physical characteristics	Control Platoon n = 32	Modified Platoon n= 33
Age (years)	20 (15)	19 (9)
Stature (m)	1.75 (0.25)	1.75 (0.18)
Body mass (kg)	68.8 (39.5)	73.6 (30.6)
Body fat (%)	11.4 (16.1)	12.7 (12.8)
$\dot{V}O_2$max (ml·kg^{-1}·min^{-1})	55.1 (16.6)	54.7 (15.0)
2.4 km run time (min:s)	10:21 (3:18)	10:28 (2:05)
Racial characteristics		
White (British)	25	26
Mixed race (British)	0	1
Fijian	5	6
Black (African)	2	0

9.2.3 *Test battery*

The test battery consisted of, 2.4 km loaded marching performance with 20 kg ($LMP_{2.4}$), 6.4 km LMP with 25 kg, which was preceded by a 6.4 km squadded march ($LMP_{12.8}$), 2.4 km running performance, running $\dot{V}O_2$max, peak dynamic strength (i.e. chest, back and leg), peak isometric strength (static lift strength, SLS), loaded marching economy (LME), body mass, estimated fat mass, estimated fat free mass, and estimated percentage body fat (Chapter 3; section 7.2.3). Due to a mechanical fault with the treadmill, LME speeds were different between-groups at T1 (i.e. 6.5 km·h^{-1} control group vs. 6.8 km·h^{-1} modified group), which was kept identical within each group for T2, and T3 testing sessions.

9.2.4 Physical training programmes

Structured PT lessons were delivered over the 26 weeks of infantry training. Week 22 was deemed to be the logical end point of the study as this coincided with the completion of the infantry physical output standard (BCFT - 12.8 km loaded march in under 2 hours, with a load of 25 kg). During the 22 weeks both CON and MOD Platoons each undertook 42 PT lessons. Many of the PT lessons were double periods, therefore the 42 PT lessons consisted of 85, 40-minute periods. On average 2.6 hours of PT were conducted each week.

9.2.4.1 Control group physical training

The CON group conducted 42 PT lessons, which constituted the existing infantry PT programme. This was similar in content to the PT programme undertaken in Chapter 8, but smaller in volume (42 vs. 60 PT lessons). The PT programme was broken down into periods of 40 minutes, and consisted of; loaded marching (24 periods), continuous runs (11), circuit training (13), Representative Military Tasks (RMTs) and Personal Fitness Test (6), battle PT (22), swimming (6) and other (3) lessons.

9.2.4.2 Modified group physical training

The total number (42) and distribution of PT lessons over the 22 weeks of training was identical to that of the CON group. Thus, the CON and MOD training programmes were solely differentiated by the content of the PT lessons. The two PT programmes were differentiated by a total of 13 HIIT sessions, and 15 RT sessions over the 22 weeks of training.

The HIIT replaced continuous runs and battle PT lessons (i.e. obstacle course, steeple chase etc.). Interval distances were progressively increased from 0.2 km (week 3) to 1.2 km (week 20), and were separated by active recovery that approximated the duration of the work interval (e.g. 1:1). Shorter intervals (i.e. 0.05 to 0.1 km) were also included in each interval session to elicit powerful contractions, which has relevance to middle distances running (Billat *et al.* 2001). The total volume of work intervals during each interval session was at least 2.4 km.

A RT intervention replaced circuit training lessons, or was added on to the end of standard PT lessons. Strength development in the chest, back, and legs were the focus of the intervention. Three core compound exercises were adapted to progress the absolute load intensity during training. The progressive order of these core exercises were; press-up (PU), decline PU, crossed thumb PU, one-legged PU, marine PU (for the chest region); assisted heaves, rope climbs, under-grasp heaves, over-grasp heaves (for the back region); body weight squat, power bag squat (25 kg), ¼ piggy back (person of similar size), one-legged squat, ½ piggy back squat (person of similar size) (for the leg region). In each session at least 3 sets were performed for each core exercise. As progression in absolute load intensity was limited, training progression was mostly manipulated by increasing the number of repetitions, which were from 10 to 30 (marine press-ups), 5 to 15 (over-grasp heaves), 5 to 20 (½ piggy back squats).

9.2.5 The quantification of relative exercise intensity during physical training

During each PT lesson subjects wore a heart rate monitor (Polar Team, Polar Electro Oy, Finland). Heart rate was logged at 5 second intervals and the data were collated at the end of each lesson (Polar Precision Performance, Finland). Individual HR responses were visually inspected after each lesson and excluded from analysis if data artefact was present.

The relative intensity of PT was quantified as percentage of heart rate reserve (%HRR) (American College of Sports Medicine, 1998; Howley, 2001), which is equivalent to the %$\dot{V}O_2$Reserve (Swain and Leutholtz, 1997). Maximal HR (HR_{max}), classified as the maximal HR attained during the laboratory running $\dot{V}O_2max$ test (Chapter 3), and resting HR (HR_{rest}) taken from the lowest rolling 30 s mean during sleep (Rayson *et al.*, 2002c), were used to calculate HRR (Equation 9.1). Percentage HRR was calculated (Equation 9.2) and the time (minutes) and percentage of time (%) spent in five exercise intensity zones was calculated (i.e. very light (<20% HRR), light (20-39% HRR), moderate (40-59% HRR), hard (60-84% HRR), and very hard ($\geq$85% HRR) (Howley, 2001)).

$$HRR = HR_{max} - HR_{rest}$$

(Equation 9.1)

$$\%HRR = ((\text{Exercising HR} - HR_{rest})/HRR) \cdot 100$$

(Equation 9.2)

Percentage HRR data were only analysed if CON and MOD groups had data from corresponding PT lessons (e.g. both CON and MOD data present for PT lesson number 1 etc.).

9.2.6 The quantification of absolute exercise intensity during physical training

Absolute exercise intensity was quantified with a Global Positioning System (GPS) (Garmin, eTrex® Legend) during selected continuous runs (control group), and corresponding HIIT sessions (modified group). This system was previously proven to be valid (Wilkinson *et al.*, 2006). The GPS was secured with duck-tape to the strap of a Camelbak hydration pack, worn by the subject. As only one GPS was available, an individual subject was identified that had a 2.4 km run time within one standard deviation of the group mean. Between-groups, subjects with identical 2.4 km run times were selected for GPS comparison (e.g. PT lesson number 1 for control vs. modified group). Following a stationary warm-up, the GPS track log was set up to record time (hr:min:sec), altitude (m), horizontal distance (m), and speed ($km \cdot h^{-1}$) at 5 s intervals. Data were extracted with the relevant software (Garmin Mapsource Metroguide Europe v6, International Inc, Kansas, USA).

9.2.7 Statistical analysis

A prospective effect size calculation (in absolute units of measurement e.g. $ml \cdot kg^{-1} \cdot min^{-1}$) was performed to determine the adequacy of a sample size equivalent to a single Platoon of subjects, at detecting a change in $\dot{V}O_2max$ (a targeted measure of the modified PT intervention, for which repeatability data existed) (PC Size, Dallal, 1985). The four known variables: *n* (23), *Alpha* (0.05), *Power* (0.80), and $\dot{V}O_2max$ measurement error – *95% Standard Deviations of the Difference* (2.9 $mlO_2 \cdot kg^{-1} \cdot min^{-1}$) (Chapter 6), calculated the *Effect Size* required to detect a significant adaptation in $\dot{V}O_2max$.

Approximately 50% of a Platoon (n=~23) undergoing CIC at ITC(C) can expect to reach the end of training at the first attempt (Carter *et al.*, 2006). Thus, within-subjects (i.e. the effect of Time) an *n* of 23 would require an *Effect Size* of 1.8 $ml \cdot kg^{-1} \cdot min^{-1}$ in $\dot{V}O_2max$ to detect statistical significance. Between-subjects (i.e. the effect of Group) an *n* of 23 would require an *Effect Size* of 2.5 $ml \cdot kg^{-1} \cdot min^{-1}$ in $\dot{V}O_2max$ to detect statistical significance. Thus, for the PT intervention to be successful, the modified PT group must increase $\dot{V}O_2max$ by 2.5 $ml \cdot kg^{-1} \cdot min^{-1}$ to 3.6 $ml \cdot kg^{-1} \cdot min^{-1}$, which was feasible (Gutekunst *et al.*, 2006).

Data were checked for the assumption of Normality (Fallowfield *et al.*, 2005). Due to the high prevalence of non-Normally distributed data, the central tendency and spread of data were reported as Median and (Range) (Thomas and Nelson, 1996). For visual purposes the middle two quartiles of the data (and not the range) constituted the error bars in the figures, as this highlighted trends better in the data. A non-parametric inferential statistical approach was chosen, the rationale for this decision is provided elsewhere (section 3.11.3). The factors of Group (CON vs. MOD) and Time (T1, T2, T3) were investigated separately. Between-subjects a Mann-Whitney Test ($P<0.05$) compared both the effect of Group at each of the three discrete points of training (i.e. T1 (beginning), T2 (middle), T3 (end)), and differences in each of the %HRR exercise intensity zones (i.e. <20%, 20-39%, 40-59%, 60-84%, >85% HRR). Within-subjects the effect of Time was assessed with a Friedman's ANOVA. Here a *post hoc* Wilcoxon Signed Rank Test with Bonferroni adjustment changed the level of alpha ($0.05 \div 3$, $P<0.0167$). Effect sizes (r) are reported both for between and within-group comparisons (section 3.11.3).

9.3 Results

9.3.1 Subject attrition

Of the 50 CON subjects 18 did not reach week-22 of CIC training at the first attempt. Of these 18, 1 was transferred to medical rehabilitation, 6 were 'back squadded', and 11 were discharged (i.e. 4 medically discharged, 2 discharged as of rights, 1 services no longer required, 4 unfit for army service). Of the 42 MOD subjects 8 did not reach week-22 at the first attempt. Of these 8, 4 were 'back squadded', and 4 were discharged (i.e. 1 discharged as of rights, 3 unfit for army service). A total of 32 CON, and 34 MOD recruits completed the study, which equated to a training attrition of 36% and 19%, respectively.

9.3.2 Baseline fitness

Between-subjects, peak [BLa] was the only measure significantly different at baseline. Here, MOD peak [BLa] was 1.6 $mmol \cdot l^{-1}$ greater than CON (U=184.5, P=0.036, r=-0.30) (Table 9.2). Due to differences in LME treadmill speeds between-group comparisons were not possible for LME (section 9.2.3).

Measures that were not different at baseline included, 2.4 km Loaded Marching Performance ($LMP_{2.4}$) (MOD 886 (187) s vs. CON 917 (295) s, U=304.00, P=0.358, r-0.13) (Figure 9.1), and 2.4 km running performance (MOD 628 (125) s vs. CON 621 (198) s, U=426.50, P=0.909, r=-0.01) (Figure 9.2). In the maximal treadmill running tests, absolute $\dot{V}O_2max$ (U=307.00, P=0.748, r=-0.04), relative $\dot{V}O_2max$ (U=265.00, P=0.265, r=-0.16), peak $\dot{V}O_2max$ RER (U=314.50, P=0.857, r=-0.03), and peak $\dot{V}O_2max$ HR (U=228.50, P=0.786, r=-0.04) were not different (Table 9.2). Of the strength measures, peak chest strength (U=393.50, P=0.407, r=-0.11), peak back strength (U=381.50, P=0.425, r=-0.10), peak leg strength (U=440.00, P=0.888, r=-0.02), and peak SLS (U=334.00, P=0.256, r=-0.15) were not different (Table 9.3). And of the body composition measures, body mass (U=408.00, P=0.539, r=-0.08), estimated body fat percentage (U=391.50, P=0.390, r=-0.11), estimated fat mass (FM) (U=401.00, P=0.473, r=-0.09), and estimated fat free mass (FFM) (U=410.00, P=0.559, r=-0.08) were not different (Table 9.4).

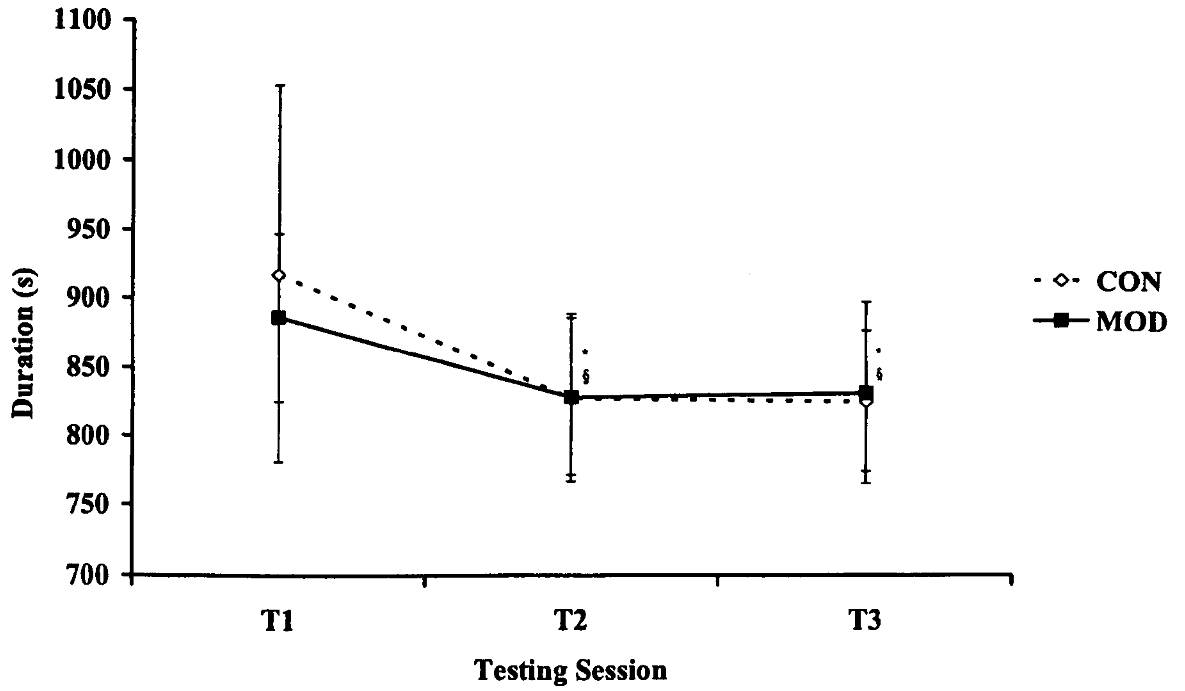


Figure 9.1. 2.4 km loaded marching performance at T1, T2, and T3 (median (range – 1st to 3rd quartile)). Within-group differences were * = CON different from T1, § = MOD different from T1.

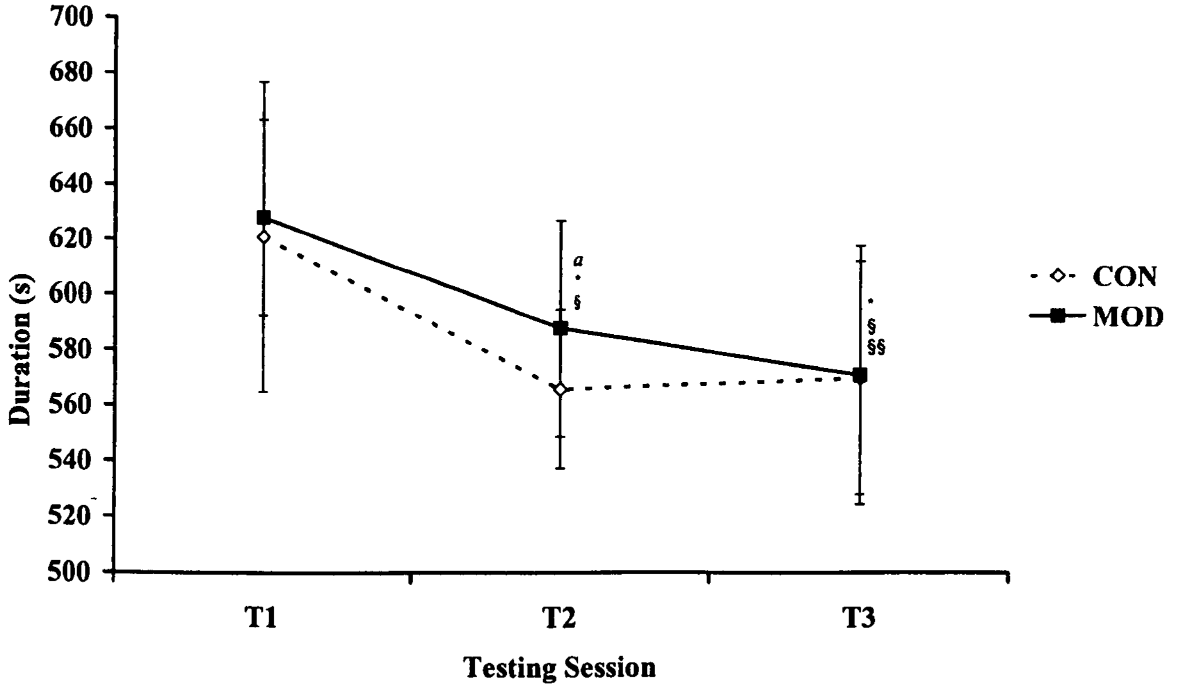


Figure 9.2. 2.4 km run time (median (range – 1st to 3rd quartile)). [a] = between-group differences at T2. Within-group differences were * = CON different from T1, § = MOD different from T1, §§ = MOD different from T2.

Table 9.2. Maximal running treadmill measures showing, between-subject, and within-subject differences at T1, T2, and T3 (median (range)).

Measure		*n*	T1	%T2	T2	%T2 toT3	T3	%T3
$\dot{V}O_2$	CON	24	3.89 (1.56)	+1.5	3.95 (1.69)	+2.7	4.06 (1.77)*	+4.2
($l \cdot min^{-1}$)	MOD	27	3.91 (1.68)	+1.8	3.98 (1.36)*	+0.7	4.01 (1.56)	+2.5
$\dot{V}O_2$	CON	24	55.1 (16.6)	+0.2	55.2 (13.3)	+3.0	56.9 (16.2)	+3.2
($ml \cdot kg^{-1} \cdot min^{-1}$)	MOD	27	54.7 (15.0)	+0.4	54.9 (18.9)	-1.1	54.3 (19.4)	-0.3
RER	CON	24	1.23 (0.17)	0.0	1.23 (0.22)	0.0	1.23 (0.19)	0.0
	MOD	27	1.22 (0.18)	+2.4	1.25 (0.19)*	0.0	1.25 (0.16)	+2.4
BLa	CON	24	8.4 (6.3)	+5.6	8.9 (11.8)	-13.5	7.7 (6.3) †	-8.3
($mmol \cdot l^{-1}$)	MOD	27	9.6 (9.0)[a]	+17.2	11.6 (8.9)[b]*	-17.2	9.6 (7.3)[c]†	0.0
Heart Rate	CON	24	198 (37)	-0.5	197 (33)	-4.1	189 (32)*†	-4.5
($beats \cdot min^{-1}$)	MOD	20	197 (31)	+0.5	198 (27)	-3.0	192 (24)* †	-2.5

* = significantly different from T1, † = significantly different from T2.

[a, b, c] = significantly different from control group at T1, T2 and T3, respectively.

%T2 = % change in first half of training; **%T2 to T3** - % change over second half of training; **%T3** = % change over all training.

Table 9.3. Dynamic and isometric strength measures, between-subject, and within-subject differences at T1, T2, and T3 (median (range)).

Measure		n	T1	%T2	T2	%T2 toT3	T3	%T3
Dynamic Strength								
Peak Chest Strength (kg)	CON	31	64 (41)	-1.6	63 (40)	+4.5	66 (40)	+3.0
	MOD	29	65 (53)	+3.0	67 (48)*	-1.5	66 (43)	+1.5
Peak Back Strength (kg)	CON	31	67 (41)	-1.5	66 (37)	0.0	66 (41) †	-1.5
	MOD	28	67 (38)	+4.3	70 (37)	-1.4	69 (37)	+2.9
Peak Leg Strength (kg)	CON	31	179 (154)	+7.7	193 (145)	-1.0	191 (150)* †	+6.3
	MOD	29	184 (125)	+10.3	205 (144)*	+1.9	209 (109)*	+12.0
Isometric Strength								
Peak Static Lift Strength (kg)	CON	30	116.8 (89.5)	-2.0	114.5 (101.5)	+3.8	119.0 (89.0)	+1.8
	MOD	27	124.0 (96.0)	-1.2	122.5 (101.0)	+3.2	126.5 (78.0)	+2.0

* = significantly different from T1, † = significantly different from T2.

%T2 = % change in first half of training; **%T2 to T3** - % change over second half of training; **%T3** = % change over all training.

Table 9.4. Body composition responses between-subjects and within-subjects at T1, T2, and T3 (median (range)).

Measure		*n*	T1	%T2	T2	%T2 toT3	T3	%T3
Body Mass	CON	31	68.8 (39.5)	+5.5	72.8 (36.4)	+0.1	72.9 (35.7)	+5.6
(kg)	MOD	29	73.6 (30.6)	+0.1	73.7 (27.9)	+2.4	75.7 (29.0)*†	+2.8
Percentage Body Fat	CON	31	11.4 (16.1)	0.0	11.4 (10.4)	-3.5	11.0 (9.8)	-3.5
(%)	MOD	29	12.7 (12.8)	-10.2	11.4 (13.2)*	0.0	11.4 (15.0)*	-10.2
Estimated Fat Mass	CON	31	8.7 (16.4)	0.0	8.7 (9.9)	-3.4	8.4 (9.8)	-3.4
(kg)	MOD	29	8.7 (13.0)	-6.9	8.1 (12.6)*	+3.6	8.4 (14.2)*	-3.4
Estimated Fat Free Mass	CON	31	61.0 (27.9)	+3.9	63.5 (29.3)*	+0.6	63.9 (30.8)*	+4.5
(kg)	MOD	29	63.5 (20.0)	+3.1	65.5 (22.8)*	+1.9	66.8 (24.5)* †	+4.9

* = significantly different from T1, † = significantly different from T2.

%T2 = % change in first half of training; **%T2 to T3** - % change over second half of training; **%T3** = % change over all training.

9.3.3 *The cardiovascular stimulus of physical training*

Differences in the average %HRR (i.e. MOD minus CON) during each individual PT lesson are presented (Figure 9.3). A positive value indicated that MOD had a higher mean %HRR during a PT lesson compared to CON. From a single PT lesson, a typical example of the relative cardiovascular stress (%HRR) and the corresponding absolute stress (i.e. running speed in $km \cdot h^{-1}$) in two individuals matched for 2.4 km run time is presented (Figure 9.4 and 9.5).

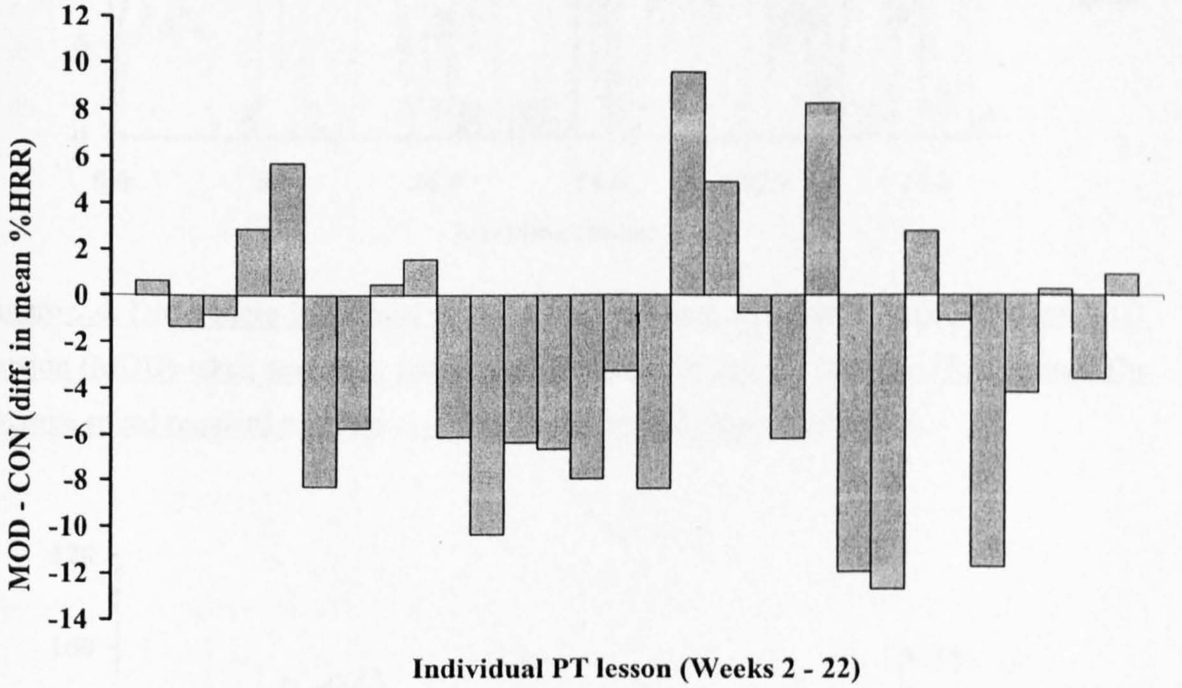


Figure 9.3. The difference in MOD group average %HRR for each of the 30 PT lessons over 22 weeks of infantry training, compared to the corresponding CON group PT lesson.

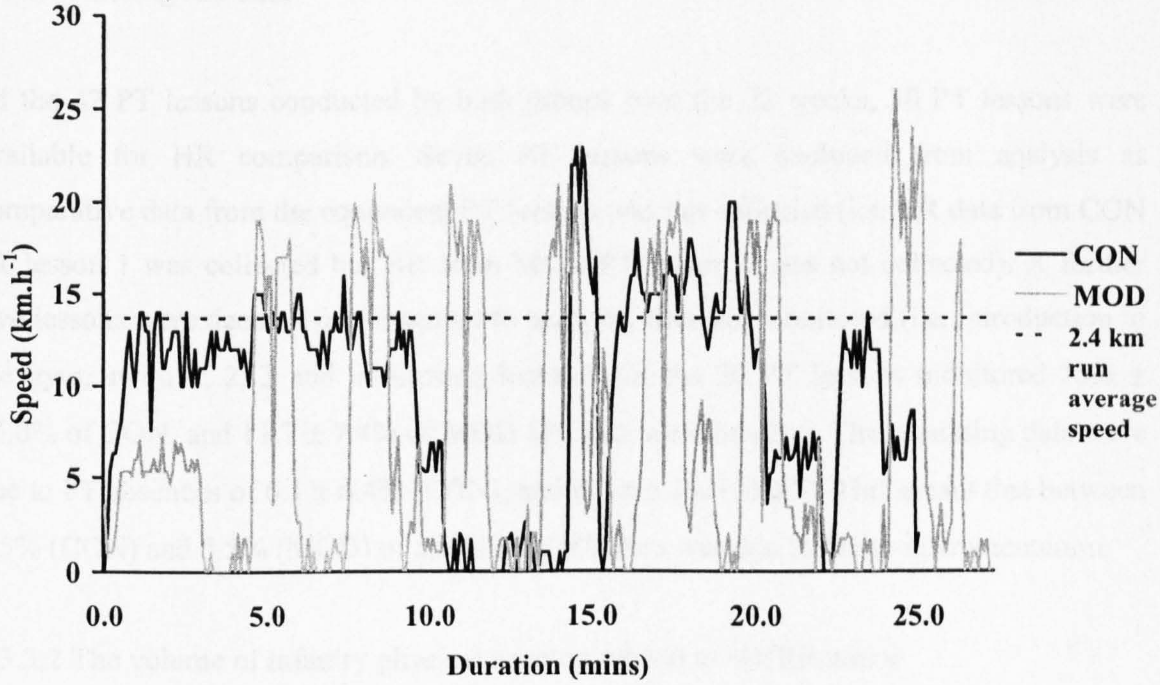


Figure 9.4. Differences in running speed during a typical continuous run (CON) and HIIT session (MOD) taken from two individuals matched for 2.4 km run time (10.5 mins). The average speed required to achieve 10.5 mins 2.4 km run time is annotated.

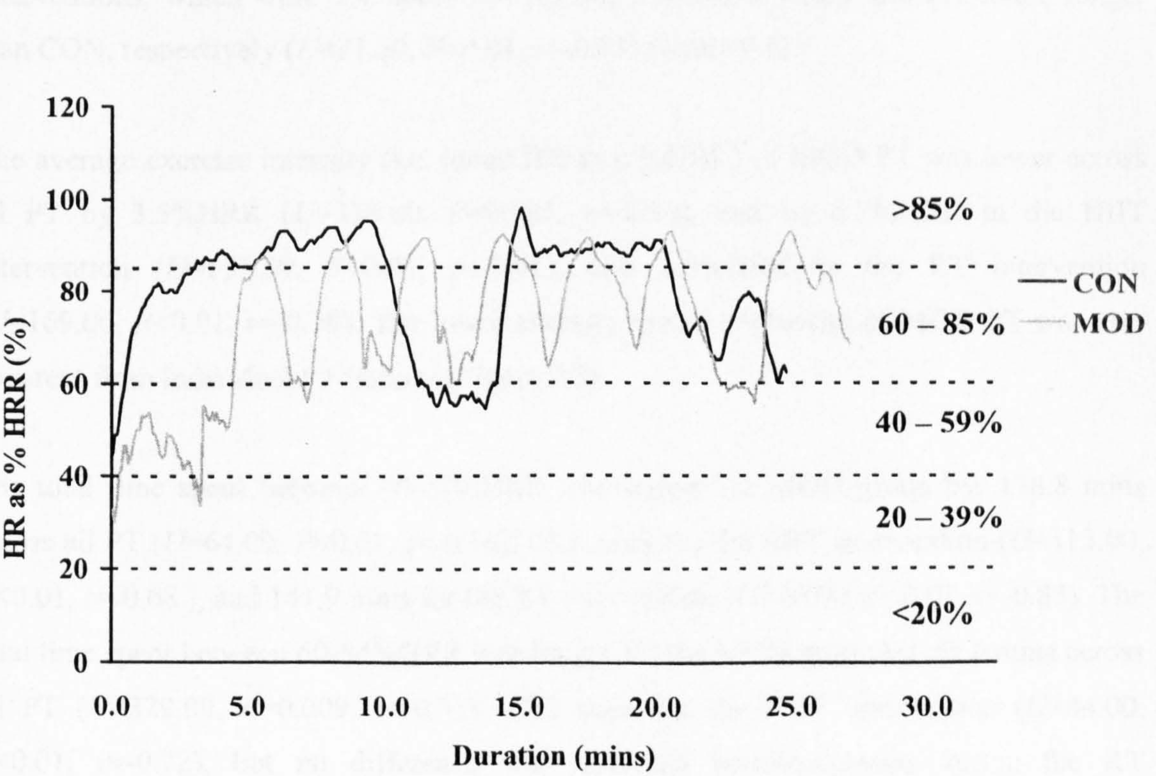


Figure 9.5. From the same session and subjects (Figure 9.4), HR expressed as a %HRR. Percentage of HRR zones are annotated.

9.3.3.1 Missing HR data

Of the 42 PT lessons conducted by both groups over the 22 weeks, 30 PT lessons were available for HR comparison. Seven PT lessons were excluded from analysis as comparative data from the equivalent PT lessons was not collected (i.e. HR data from CON PT lesson 1 was collected but HR from MOD PT lesson 1 was not collected). A further five lessons were deemed non-essential to analysis, thus not monitored (i.e. introduction to the gym, swim 1, 2, 3 and a nutrition lecture). Of the 30 PT lessons monitored 15.6 ± 15.0% of CON, and 11.7 ± 7.4% of MOD HR data were missing. These missing data were due to PT absences of 6.1 ± 6.4% (CON), and 6.2 ± 5.3% (MOD). This meant that between 9.5% (CON) and 5.5% (MOD) of all missing HR data was due to faulty instrumentation.

9.3.3.2 The volume of infantry physical training related to %HRR zones

The total volume (i.e. duration) of MOD PT was 3.7 hours longer than CON over 22 weeks of training (U=137.00, P<0.01, r=-0.64). This was also reflected in the HIIT, and RT interventions, which were 1.4 hours (U=128.50, P<0.01, r=-0.65) and 2.4 hours longer than CON, respectively (U=21.50, P<0.01, r=-0.83) (Table 9.5).

The average exercise intensity (i.e. mean HR as a %HRR) of MOD PT was lower across all PT by 3.5%HRR (U=316.00, P=0.005, r=-0.35), and by 5.7%HRR in the HIIT intervention (U=151.50, P<0.01, r=-0.61), and 5.9%HRR in the RT intervention (U=169.00, P<0.01, r=-0.58). The lower average exercise intensity of MOD PT was also apparent from individual PT lessons (Figure 9.3).

The total time spent between 40-59%HRR was longer for MOD group by, 178.8 mins across all PT (U=64.00, P<0.01, r=-0.76), 60.2 mins for the HIIT intervention (U=113.00, P<0.01, r=-0.68), and 141.9 mins for the RT intervention (U=16.00, P<0.01, r=-0.83). The total time spent between 60-84%HRR was longer for the MOD group by, 83.6 mins across all PT (U=329.00, P=0.009, r=-0.32), 63.2 mins for the HIIT intervention (U=84.00, P<0.01, r=-0.72), but no difference was observed between-groups during the RT intervention (U=416.00, P=0.142, r=-0.18). The total time spent above 85%HRR across all PT was not different between-groups (U=431.00, P=0.203, r=-0.16). However, the MOD group spent significantly less time (-33.5 mins) above 85%HRR during the HIIT intervention (U=234.00, P<0.01, r=-0.48), and the RT (-28.8 mins) intervention (U=133.00, P<0.01, r=-0.64) (Table 9.5).

9.3.3.3 The percentage of time spent in the %HRR zones during infantry physical training

The percentage of time spent between 40-59%HRR was longer for the MOD group by, 8.8% across all PT (U=110.00, P<0.01, r=-0.68), 10.4% for the HIIT intervention (U=186.50, P<0.01, r=-0.56), and 16.5% for the RT intervention (U=71.00, P<0.01, r=-0.74). The percentage of time spent between 60-84%HRR was not different across all PT (U=437.50, P=0.235, r=-0.15). However, the MOD group spent a greater percentage of time (8.5%) between 60-84%HRR during the HIIT intervention (U=215.50, P<0.01, r=-0.51), but less percentage of time (-9.9%) in this zone during the RT intervention (U=331.00, P=0.010, r=-0.32). The percentage of time spent above 85%HRR was shorter for the MOD by, -3.7% across all PT (U=285.00, P=0.001, r=-0.40), -18.0% during the HIIT intervention (U=74.00, P<0.01, r=-0.74), and -9.2% during the RT intervention (U=89.5, P<0.01, r=-0.71) (Table 9.5).

Table 9.5. The relative intensity (%HRR) of all 30 PT lessons, the high intensity interval training (HIIT), and resistance training (RT) interventions, for MOD and CON groups (median (range)).

		Total PT	HIIT PT	RT PT
Total Time (hours)	CON	25.0 (18.4)	5.2 (4.6)	5.8 (5.9)
	MOD	28.7 (10.8)*	6.6 (3.1)*	8.2 (3.0)*
Mean HR as %HRR (%)	CON	68.2 (15.8)	74.0 (18.5)	60.5 (20.9)
	MOD	64.7 (17.6)*	68.3 (17.0)*	54.6 (23.0)*
Time <20% HRR (mins)	CON	3.8 (25.9)	0.0 (21.5)	0.3 (4.8)
	MOD	1.8 (41.8)	0.3 (6.7)*	0.3 (3.6)
Time 20 – 39% HRR (mins)	CON	107.3 (263.5)	9.6 (48.8)	40.7 (130.7)
	MOD	114.3 (314.2)	13.1 (42.0)	42.3 (191.8)
Time 40 - 59% HRR (mins)	CON	343.3 (364.2)	60.2 (84.9)	131.9 (173.3)
	MOD	522.1 (383.7)*	111.9 (114.9)*	273.8 (201.3)*
Time 60 – 84% HRR (mins)	CON	661.3 (578.2)	128.9 (155.2)	120.5 (146.4)
	MOD	744.9 (578.7)*	192.1 (145.3)*	130.9 (315.4)
Time >85% HRR (mins)	CON	325.8 (553.8)	95.5 (128.1)	38.8 (80.3)
	MOD	291.7 (474.4)	62.0 (132.5)*	10.0 (22.7)*
Time <20% HRR (%)	CON	0.3 (1.8)	0.0 (6.9)	0.1 (1.3)
	MOD	0.1 (2.4)	0.1 (1.5)*	0.1 (0.8)
Time 20 – 39% HRR (%)	CON	7.6 (14.5)	3.0 (13.7)	10.9 (33.2)
	MOD	6.7 (18.6)	3.3 (11.5)	11.4 (40.2)
Time 40 - 59% HRR (%)	CON	22.8 (23.8)	19.2 (24.0)	38.6 (29.7)
	MOD	31.6 (20.6)*	29.6 (27.7)*	55.1 (43.8)*
Time 60 – 84% HRR (%)	CON	44.4 (23.7)	42.6 (37.8)	36.5 (52.7)
	MOD	43.3 (20.6)	51.1 (24.3)*	26.6 (57.9)*
Time >85% HRR (%)	CON	21.1 (33.4)	35.1 (45.7)	11.2 (28.1)
	MOD	17.4 (25.8)*	17.1 (28.1)*	2.2 (4.2)*

Significantly different from CON * = $P<0.05$.

9.3.4 The differences and adaptations in loaded marching performance

The external load was the same for both groups during $LMP_{2.4}$ (19.5 ± 0.6 vs. 19.8 ± 1.0 kg, $P>0.05$). Between-subjects, $LMP_{2.4}$ was not different at T2 (MOD 829 (184) s vs. CON 828 (173) s, $U=329.00$, $P=0.630$, r=-0.07), or T3 (MOD 832 (195) s vs. CON 826 (145) s, $U=281.00$, $P=0.187$, r=-0.18) (Figure 9.1). Over the longer distance, $LMP_{12.8}$ was 3.7 % slower in the MOD group (MOD 51.5 (10.1) mins vs. CON 49.7 (13.5) mins, $U=282.00$, $P=0.005$, r=-0.35) (Figure 9.6). No difference was observed in the percentage of time spent between 60-85%HRR (MOD 61.6 %time vs. CON 66.8%time, $U=267.00$, $P=0.330$), and above 85%HRR during $LMP_{6.4}$ (MOD 29.5%time vs. CON 26.5%time, $U=271.00$, $P=0.341$) (Figure 9.7).

Within-subjects, $LMP_{2.4}$ improved in both MOD (χ^2 (2)=36.40, $P<0.01$) and CON over training (χ^2 (2)=31.55, $P<0.01$). MOD $LMP_{2.4}$ decreased by 6.1% between T1 to T3 (from 886 s to 832 s, $T=0$, $P<0.01$, r=-0.62), which occurred by T2 (from 886 s to 829 s, $T=0$, $P<0.01$, r=-0.62). CON $LMP_{2.4}$ decreased 9.9% between T1 to T3 (from 917 s to 826 s, $T=3$, $P<0.01$, r=-0.58), which also occurred by T2 (917 to 828 s, $T=3$, $P<0.01$, r=-0.55) (Figure 9.1).

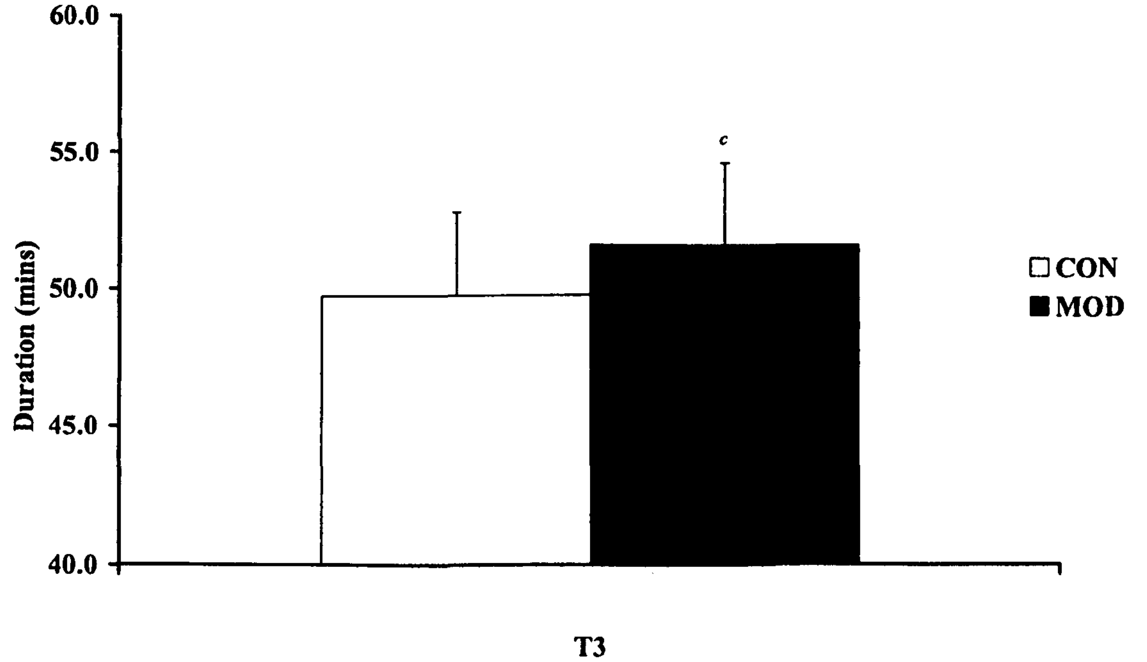


Figure 9.6. The 6.4 km best effort LMP (during the 12.8 km LMP test) at T3 (median (range – 1st to 3rd quartile)). c = MOD significantly different from CON ($P<0.05$).

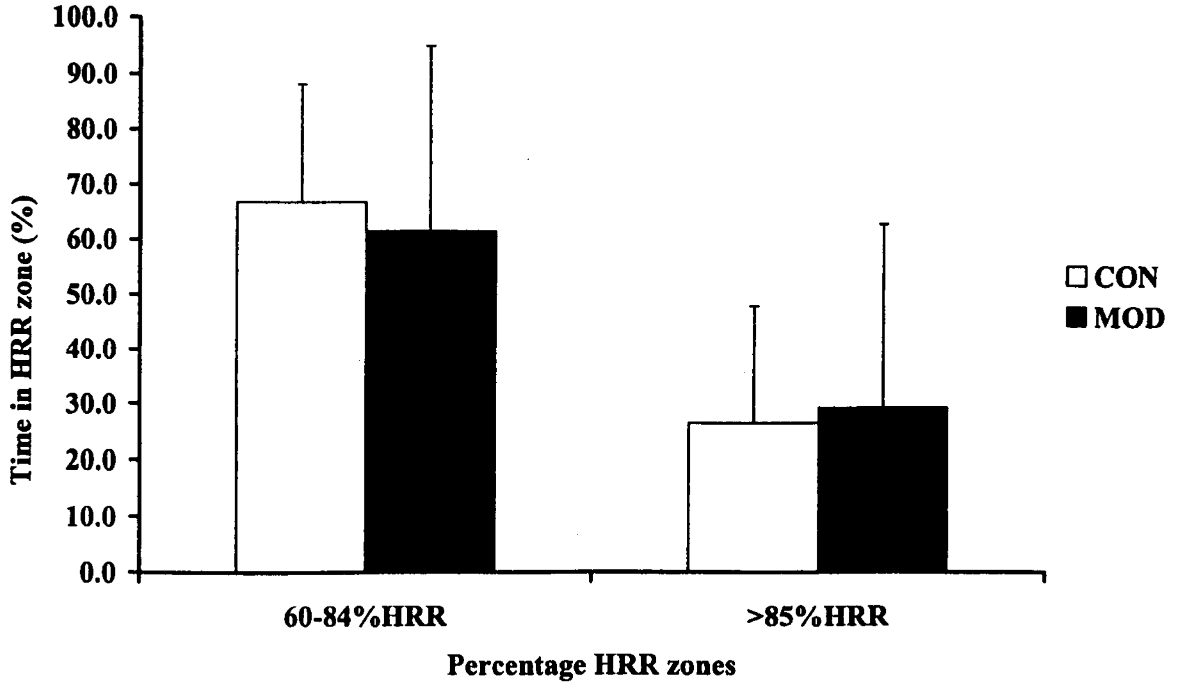


Figure 9.7. Percentage of time spent in two HRR zones during the 6.4 km best effort component (of the 12.8 km LMP test) at T3 (median (range – 1st to 3rd quartile)).

9.3.5 *The differences and adaptations in the determinants of loaded marching performance, targeted by the modified physical training programme*

9.3.5.1 Two point four km running performance

Between-subjects, MOD 2.4 km run time was 3.7% slower than CON at T2 (588 (166) s vs. 566 (141) s, U=219.00, P=0.001, r=-0.34), but not different at T3 (571 (81) s vs 570 (126) s, U=351.00, P=0.208, r=-0.18) (Figure 9.2).

Within-subjects, 2.4 km run time improved for both MOD (χ^2 (2)=44.27, P<0.01) and CON groups over training (χ^2 (2)=47.67, P<0.01). MOD run time decreased 9.1% between T1 to T3 (from 628 to 571 s, T=1, P<0.01, r=-0.61). Most of the adaptation (6.4%) occurred by T2 (T=1, P<0.01, r=-0.59), and a continued improvement of 2.9% was observed between T2 and T3 (T=3, P<0.01, r=-0.57). CON run time decreased 8.2% between T1 to T3 (from 621 s to 570 s, T=0, P<0.01, r=-0.62). No further improvement was observed after T2 (Figure 9.2).

9.3.5.2 Running $\dot{V}O_2max$ test

Between-subject differences were observed in one out of the six $\dot{V}O_2max$ measures. Here, MOD peak [BLa] was greater than CON by, 2.7 mmol·l^{-1} at T2 (U=124.00, P=0.001, r=-0.48), and 1.9 mmol·l^{-1} at T3 (U=128.50, P=0.001, r=-0.47). Absolute $\dot{V}O_2max$ was not different at T2 (U=285.00, P=0.462, r=-0.10), or T3 (U=312.50, P=0.828, r=-0.03). Also relative $\dot{V}O_2max$ was not different at T2 (U=305.50, P=0.727, r=-0.05), or T3 (U=231.00, P=0.079, r=-0.25) (Table 9.2).

Within-subjects, absolute $\dot{V}O_2max$ increased for both MOD (χ^2 (2)=7.63, P=0.022), and CON over training (χ^2 (2)=8.14, P=0.017). The increase of 0.07 l·min^{-1} in MOD absolute $\dot{V}O_2max$ manifest between T1 to T2 (T=6, P=0.005, r=-0.39), whereas the 0.17 l·min^{-1} increase in CON group absolute $\dot{V}O_2max$ only became significant at T3 (T=5, P=0.001, r=-0.49). Relative $\dot{V}O_2max$ did not increase over training for both MOD (χ^2 (2)=5.20, P=0.074), and CON groups (χ^2 (2)=1.08, P=0.582). Peak RER increased in MOD (χ^2 (2)=6.63, P=0.036), but did not change in CON over the training (χ^2 (2)=1.11, P=0.573). The 0.03 increase in the MOD group was apparent by T2 (T=7, P=0.010, r=-0.35). Peak [BLa] significantly changed for MOD (χ^2 (2)=10.37, P=0.006), and CON groups (χ^2 (2)=7.18, P=0.028). The MOD peak [BLa] increased 2.0 mmol·l^{-1} by T2 (T=7, P=0.002, r=-0.41), and subsequently decreased back to baseline by T3 (T=4, P=0.001, r=-0.45). Whereas CON peak [BLa] decreased 1.2 mmol·l^{-1} between T2 to T3 (T=5, P=0.011, r=-0.38). Peak heart rate (HR) decreased in MOD (χ^2 (2)=26.00, P<0.01) and CON groups (χ^2 (2)=20.86, P<0.01). In both groups peak HR decreased between T2 to T3, by 6 beats·min^{-1} in the MOD (T=0, P<0.01, r=-0.62), and 8 beats·min^{-1} in the CON (T=3, P=0.001, r=-0.49). Consequently, peak HR was below baseline values at T3 in the MOD (T=1, P<0.01, r=-0.57), and CON groups (T=2, P<0.01, r=-0.51) (Table 9.2).

9.3.5.2 Strength

Between-subjects, none of the four strength indices were different at T2, and T3 (P>0.05). No difference was observed for peak static lift strength (SLS) between MOD and CON at

T3 (U=328.00, P=0.218, r=-0.16). There was a tendency for MOD peak back strength to be greater at T2 (U=310.00, P=0.059, r=-0.25). Despite, numerically an 8.6% greater MOD peak leg strength at T3, this did not approach statistical significance (U=358.00, P=0.176, r=-0.17) (Table 9.3).

Within-subjects, peak SLS did not change over training for both MOD (χ^2 (2)=0.92, P=0.633, r=-0.16), and CON (χ^2 (2)=1.88, P=0.390, r=-0.05). Peak chest strength increased in the MOD (χ^2 (2)=13.02, P=0.001), but not the CON group (χ^2 (2)=3.52, P=0.172). The 2 kg increase in MOD chest strength was manifest by T2 (P<0.01, r=-0.48), but did not remain above baseline by T3 (T=8, P=0.063, r=-0.24). Peak back strength remained unchanged in MOD (χ^2 (2)=3.92, P=0.141), and decreased during CON training (χ^2 (2)=6.17, P=0.046). Although MOD back strength did not increase it did approach significance (T=9, P=0.048, r=-0.26). The decrease in CON back strength of 1 kg occurred between T2 to T3 (T=7, P=0.015, r=-0.31). Peak leg strength increased in both MOD (χ^2 (2)=18.82, P<0.01) and CON training (χ^2 (2)=17.12, P<0.01). The percentage change in MOD leg strength was almost two-fold that of CON (12.0% vs. 6.3%). MOD leg strength increased 25 kg over training (T=6, P<0.01, r=-0.51), with most of this adaptation occurring by T2 (T=4, P<0.01, r=-0.54). A 12 kg increase in CON leg strength was observed over training (T=6, P=0.001, r=-0.42) (Table 9.3).

9.3.6 The differences and adaptations of measures not specifically targeted by the modified physical training programme

9.3.6.1 Loaded marching economy

Within-subjects, ten out of ten MOD LM measures were significantly different by T2, with five continuing to change significantly between T2 to T3. Six out of 10 CON measures were significantly different by T2, and seven continued to change significantly between T2 to T3 (Table 9.6).

Absolute LME improved during training for MOD (χ^2 (2)=45.92, P<0.01) and CON (χ^2 (2)=40.55, P<0.01). During training absolute $\dot{V}O_2$ decreased 0.42 $l \cdot min^{-1}$ for MOD (T=1, P<0.01, r=-0.61), and 0.28 $l \cdot min^{-1}$ for CON (T=1, P<0.01, r=-0.60). In both groups LME continued to improve between T2 to T3 ($P<0.0167$), however most of the change had occurred by T2 (MOD $T=2$, $P<0.01$, r=-0.56; CON $T=2$, $P<0.01$, r=-0.58). Relative $\dot{V}O_2$ also decreased in MOD (χ^2 (2)=44.51, $P<0.01$) and CON groups by T3 (χ^2 (2)=42.49, $P<0.01$), by some 5.7 $ml \cdot kg^{-1} \cdot min^{-1}$ in MOD ($T=1$, $P<0.01$, r=-0.61), and 4.7 $ml \cdot kg^{-1} \cdot min^{-1}$ in CON ($T=1$, $P<0.01$, r=-0.57). Changes were observed between T2 and T3 ($P<0.0167$), however most of the response was evident by T2 (MOD $T=1$, $P<0.01$, r=-0.58; CON $T=1$, $P<0.01$, r=-0.57) groups. The fractional utilisation of $\dot{V}O_2max$ (%$\dot{V}O_2max$) during the LM task decreased in both groups over training (MOD χ^2 (2)=34.84, $P<0.01$; CON χ^2 (2)=33.68, $P<0.01$), by 12.6% $\dot{V}O_2max$ in the MOD ($T=0$, $P<0.01$, r=-0.62), and 8.5%$\dot{V}O_2max$ in the CON group ($T=1$, $P<0.01$, r=-0.61). The pattern of adaptation was different between-groups. Most of the improvement in MOD occurred by T2 ($T=3$, $P<0.01$, r=-0.59), whereas for the CON the magnitude of improvement was comparable between T1 to T2 ($T=3$, $P<0.01$, r=-0.54) and T2 to T3 ($T=4$, $P<0.01$, r=-0.53) (Table 9.6).

In the MOD group, minute ventilation ($\dot{V}_E$) (χ^2 (2)=32.51, $P<0.01$), F_EO_2 (χ^2 (2)=16.52, $P<0.01$), and F_ECO_2 (χ^2 (2)=17.56, $P<0.01$) changed over training. $\dot{V}_E$ decreased by 11.8 $l \cdot min^{-1}$ ($T=1$, $P<0.01$, r=-0.60), F_EO_2 0.10% ($T=1$, $P<0.01$, r=-0.46), and F_ECO_2 0.17% ($T=5$, $P<0.01$, r=-0.46). In all instances adaptation occurred between T1 to T2 ($P<0.0167$), did not improve further between T2 to T3, but remained different to baseline. In CON $\dot{V}_E$ decreased over training (χ^2 (2)=24.64, $P<0.01$), but F_EO_2 (χ^2 (2)=2.13, $P=0.343$), and F_ECO_2 did not (χ^2 (2)=3.53, $P=0.172$). Here CON $\dot{V}_E$ decreased 8.2 $l \cdot min^{-1}$ by T3 ($T=3$, $P<0.01$, r=-0.55), with decreases by T2 ($T=9$, $P<0.01$, r=-0.46), and between T2 and T3 ($T=5$, $P<0.01$, r=-0.41) (Table 9.6).

The MOD Respiratory Exchange Ratio (RER) decreased 0.03% by T2 but did not decrease further ($T=5$, P=0.003, r=-0.41). The CON RER did not change over training (χ^2 (2)=0.13, $P=0.93$). Circulating whole blood lactate concentration ([Bla]) decreased in MOD (χ^2 (2)=24.64, $P<0.01$), and CON groups (χ^2 (2)=17.37, $P<0.01$). A two-fold greater decrease in MOD [BLa] of 2.0 $mmlo \cdot l^{-1}$ was observed over training ($T=1$, $P<0.01$, r=-0.59), compared to 0.6 $mmol \cdot l^{-1}$ in CON ($T=3$, P<0.01, r=-0.56). In both groups [BLa] continued to decrease in the second half of training ($P<0.0167$), although the majority of the adaptation occurred by T2 (MOD $T=5$, $P=0.006$, r=-0.39; CON $T=6$, $P=0.006$, r=-0.38). Heart rate decreased over training in MOD (χ^2 (2)=36.40, $P<0.01$) and CON groups (χ^2

(2)=33.84, *P*<0.01) by 25 beats·min^{-1} (MOD *T*=0, *P*<0.01, r=-0.57) and 22 beats·min^{-1} (CON *T*=0, *P*<0.01, r=-0.61). Stride frequency (SF) decreased over training in MOD (χ^2 (2)=24.77, *P*<0.01), and CON groups (χ^2 (2)=18.43, *P*<0.01). The decrease was 4 strides·min^{-1} (MOD *T*=3, *P*<0.01, r=-0.57), and 3 strides·min^{-1} (CON groups *T*=5, *P*=0.002, r=-0.42). The decrease was evident by T2 in MOD (*T*=4, P=0.001, r=-0.46), but not until T3 in CON (*T*=2, *P*<0.01, r=-0.51) (Table 9.7).

9.3.6.2 Body Composition

Between-subjects, none of the MOD body composition measures (i.e. body mass, estimated percentage body fat, estimated fat mass and estimated fat free mass) were different to CON at T2 and T3 (*P*>0.05) (Table 9.4).

Within-subjects, significant increases in MOD (χ^2 (2)=7.00, P=0.030) and CON body mass were observed (χ^2 (2)=6.78, *P*=0.034). A 2.1 kg increase in MOD body mass occurred by T3 (*T*=9, *P*=0.015, r=-0.32), with most of this increase occurring between T2 to T3 (*T*=8, *P*=0.004, r=-0.38). Although CON body mass increased 4.1 kg by T3 this was not significant in the *post hoc* test (*T*=10, *P*=0.098, r=-0.21). Percentage body fat changed in MOD (χ^2 (2)=14.4, *P*=0.001), but not CON training (χ^2 (2)=4.26, *P*=0.119). A marked decrease of 1.3% was evident in MOD at T3 (*T*=4, *P*<0.01, r=-0.47), which occurred between T1 to T2 (*T*=6, *P*=0.001, r=-0.45). The smaller, 0.4 % decrease in CON body fat approached significance (*T*=7, *P*=0.022, r=-0.29). Estimated fat mass (FM) changed in MOD (χ^2 (2)=6.02, *P*=0.049), but not CON training (χ^2 (2)=0.155, *P*=0.925). MOD FM decreased by 0.3 kg (*T*=10, *P*=0.012, r=-0.33), but no change was observed in CON (*T*=13, *P*=0.153, r=-0.18). Estimated Fat Free Mass (FFM) increased in MOD (χ^2 (2)=28.07, *P*<0.01) and CON training (χ^2 (2)=23.76, *P*<0.01). The 2.0 kg increase in MOD FFM was distributed between the first (*T*=6, *P*=0.001, r=-0.45), and second halves of training (*T*=7, *P*=0.002, r=-0.41). A similar increase of 2.9 kg in FFM occurred during CON training (*T*=5, *P*<0.01, r=-0.50), however most of the change occurred between T1 to T2 (*T*=4, *P*<0.01, r=-0.47) (Table 9.4).

Table 9.6. Treadmill loaded marching responses, within-subjects at T1, T2, and T3 (median (range)).

Measure		*n*	T1	%T2	T2	%T2 toT3	T3	%T3
$\dot{V}O_2$	CON	29	2.40 (1.77)	-7.9	2.21 (0.55)*	-4.1	2.12 (0.63)* †	-11.7
($l \cdot min^{-1}$)	MOD	27	2.86 (1.61)	-11.2	2.54 (1.24)*	-3.9	2.44 (1.32)* †	-14.7
$\dot{V}O_2$	CON	29	34.1 (22.2)	-10.0	30.7 (13.8)*	-4.2	29.4 (12.1)* †	-13.8
($ml \cdot kg^{-1} \cdot min^{-1}$)	MOD	27	38.6 (18.3)	-9.6	34.9 (16.2)*	-5.7	32.9 (19.0)* †	-14.8
Fractional Utilisation	CON	25	61.1 (48.4)	-6.5	57.1 (23.7)*	-7.4	52.9 (21.5)* †	-13.4
(% VO_2max)	MOD	26	74.6 (30.5)	-13.8	64.3 (32.7)*	-3.6	62.0 (33.1)* †	-16.9
$\dot{V}_E$	CON	29	54.9 (70.0)	-9.7	49.6 (21.8)*	-5.8	46.7 (14.9)* †	-14.9
($l \cdot min^{-1}$)	MOD	27	64.4 (60.4)	-7.7	54.3 (26.2)*	-3.1	52.6 (32.8)*	-8.3
F_EO_2	CON	29	16.35 (2.1)	+0.2	16.39 (1.7)	-0.9	16.25 (1.22)	-0.6
(%)	MOD	27	16.57 (1.98)	-2.0	16.24 (1.27)*	+0.4	16.30 (1.71)*	-1.6
F_ECO_2	CON	29	4.66 (1.57)	-2.4	4.55 (1.36)	+2.4	4.66 (1.13)	0.0
(%)	MOD	27	4.49 (1.70)	+4.9	4.72 (1.23)*	-1.3	4.66 (1.40)*	+3.6

* = significantly different from T1, † = significantly different from T2, within subjects.

%T2 = % change in first half of training; **%T2 to T3** - % change over second half of training; **%T3** = % change over all training.

Table 9.7. Treadmill loaded marching responses, within-subjects at T1, T2, and T3 (median (range)).

Measure		*n*	T1	%T2	T2	%T2 toT3	T3	%T3
RER	CON	29	0.99 (0.21)	0.0	0.99 (0.21)	-1.0	0.98 (0.12)	-1.0
	MOD	27	1.02 (0.14)	-2.9	0.99 (0.13)*	+1.0	1.00 (0.13)	-2.0
BLa	CON	26	2.5 (7.1)	-20.0	2.0 (3.6)*	-5.0	1.9 (2.9)* †	-24.0
($mmol{\cdot}l^{-1}$)	MOD	25	4.6 (5.3)	-30.4	3.2 (7.5)*	-18.7	2.6 (4.6)* †	-43.5
Heart Rate	CON	25	157 (68)	-5.1	149 (57)*	-9.4	135 (39)* †	-14.0
($beats{\cdot}min^{-1}$)	MOD	20	176 (35)	-9.7	159 (54)*	-5.0	151 (60)* †	-14.2
Stride Frequency	CON	29	67 (22)	-1.5	66 (11)	-4.0	64 (8)* †	-4.5
($strides{\cdot}min^{-1}$)	MOD	27	70 (20)	-4.3	67 (9)*	-1.5	66 (6)*	-5.7

* = significantly different from T1, † = significantly different from T2.

%T2 = % change in first half of training; **%T2 to T3** - % change over second half of training; **%T3** = % change over all training.

9.4 Discussion

9.4.1 Main findings

The modified and control group subjects were not physically or physiologically different at baseline. The modified physical training intervention did not further improve 2.4 km loaded marching performance compared to controls. This may have been attributable to the inability of modified physical training to enhance the training responses of 2.4 km running performance, $\dot{V}O_2max$, and indices of muscular strength to a greater magnitude than the control group. The lack of difference in 2.4 km running performance and $\dot{V}O_2max$ (between-groups) may be attributed to the lack of difference in time spent above 85% of heart rate reserve during physical training lessons. At the end of training the modified physical training group were slower over the longer loaded marching performance distance of 12.8 km, which was an unexpected finding.

9.4.2 Differences and adaptations in loaded marching performance

9.4.2.1 Loaded marching performance

Modified $LMP_{2.4}$ was not different from CON at any time point during training. A programme of research similar in design to the present study (i.e. British Army recruits undergoing modified and normal PT) reported mixed results on the efficacy of concurrent aerobic and strength training on the subsequent improvement in LMP (Williams *et al.*, 1999; 2002). In agreement to the present study no difference was observed in training response of 25 kg, 3.2 km LMP (16.7% vs. 15.7%), however in a mixed gender group, 15 kg LMP improved more in the modified PT group (8.9% vs. 3.6%) (Williams *et al.*, 2002). In contrast to the present study, others have also reported the benefits of concurrent endurance and resistance training programmes for improving 3.2 km LMP, in comparison to endurance training, or resistance training alone (Kraemer *et al.*, 1987; 2001; 2004).

Chapters 7 and 8 evidenced that 2.4 km running performance (hence by default $\dot{V}O_2max$), and SLS were important determinants of $LMP_{2.4}$, however these responded sub-optimally during infantry training. These determinants of LMP are well established (Harman and

Frykman, 1995; Rayson *et al.*, 2000; Pandorf *et al.*, 2002; Williams and Rayson, 2006) but very different (i.e. endurance vs. strength). The dichotomy between these two aspects of fitness may explain why little difference has been observed between gains in LMP, when PT that emphasises endurance is compared to PT that emphasises strength (Williams *et al.*,2004). In this study Williams *et al.*, (2004) observed that the magnitude of LMP training adaptation (i.e. high or low responders) was dependent upon the relative physical deficiency an individual had (i.e. endurance or strength training status). Thus a training programme that emphasised strength development (Williams *et al.*, 2002) may cause a large LMP training response in individuals with poor strength, and a small response in individuals with good strength. Alternatively, an endurance training emphasis (Williams *et al.*, 1999) may cause a large LMP training response in individuals with poor endurance, and a small response in individuals with good endurance. Thus the effect of high and low responders in both PT groups may have cancelled out the LMP training response between-groups (Williams *et al.*, 2004). However, it is doubtful that this notion explains the lack of difference in $LMP_{2.4}$ between-groups in the present study, as the purpose of the MOD PT programme was to improve both aerobic fitness and strength. Thus, the potential poor responders in the MOD group should have received an increased PT stimulus in both endurance and strength, meaning the aspect of fitness they were deficient in would have been targeted. Therefore, this would imply that emphasising endurance or strength training deficiencies, to improve LMP as suggested (Williams *et al.*, 2004) is either ineffective, and/or the endurance and strength stimulus of MOD PT in the present study was inadequate.

Within-subjects, $LMP_{2.4}$ improved during training 6.6% (MOD) and 9.9% (CON), which were classified as *large* effects (Cohen, 1992). These were smaller (up to 23%) and larger (as low as 0.1%) than previously reported (Kraemer, 1987, 2004; Harman, 1998; Williams *et al.*, 1999, 2002). In light of the fact that the first 1.2 km of the $LMP_{2.4}$ in the present study was conducted as a squadded effort, and the second 1.2 km as an individual 'best effort' it is not surprising that the response of LMP in the present study was smaller than previously reported (Kraemer, 1987, 2004; Harman, 1998; Williams *et al.*, 1999, 2002).

The time to complete the $LMP_{12.8}$ at T3 was slower in the MOD group, which was an unexpected finding. This was somewhat of an anomaly as between-group differences in the determinants of $LMP_{2.4}$ that were established in Chapter 7 (i.e. 2.4 km run time and SLS) were no different at T3. Although $\dot{V}O_2max$ was not a determinant of $LMP_{2.4}$, it is strongly linked to middle distance (i.e. 2.4 km) running performance (Billat, 2001). In individuals homogenous for $\dot{V}O_2max$, 65% of the difference reported in running performance is a consequence of differences in running economy (Conley and Krahenbuhl, 1980). A similar causal link in the present study (i.e. differences in LME) may have explained these between-group differences $LMP_{12.8}$, however this could not be verified for reasons already stated (section 9.2.3). One indirect method of assessing between-group differences in LME might be to compare the %HRR data during the $LMP_{12.8}$ test. As exercise economy is the oxygen uptake required for a given absolute exercise intensity (Jones and Carter, 2000), and the %HRR is equivalent to %$\dot{V}O_2$reserve (Swain and Leutholtz, 1997), the %HRR data collected during $LMP_{12.8}$ are analogous to LME, making it a feasible surrogate measure of LME. Between-group comparisons revealed that the percentage of time spent between 60-84%HRR, and above 85%HRR during $LMP_{12.8}$ was not different between-groups (Figure 9.7). As relative LM exercise intensity (i.e. % of predicted $\dot{V}O_2max$) increases with increasing speeds of LM (Christie and Scott, 2005), the fact that the CON $LMP_{12.8}$ time was significantly faster than the MOD, despite no difference between the time spent in the upper %HRR zones may imply LME was better in the CON. It is tentatively suggested that surrogate LME measure of % HRR may explain the slower $LMP_{12.8}$ in the MOD group, as LME is known to have a marked effect on LMP.

Another possible explanation for the slower $LMP_{12.8}$ time in the MOD group could be the longer test distance, compared to $LMP_{2.4}$. The $LMP_{12.8}$ constituted the BCFT, thus subjects completed a squadded 6.4 km LM immediately prior (~2 mins) to the start of a 6.4 km best effort LM. The weak correlations between LMP and physical / physiological measures over LM distances of 20 km was suggested to have been a consequence of poor subject motivation (Knapik *et al.*, 1990). Long march distances are known to exacerbate symptoms of fatigue, discomfort, alertness and well-being (Johnson *et al.*, 1995), which may influence an individual's motivation. In addition, social factors may have also contaminated the data set in the present study. As after being instructed to provide an 'individual best effort' it was noticeable that subjects marched in small groups, hence may not have given a 'best effort'.

9.4.3 The determinants of loaded marching performance targeted by the modified physical training programme.

As 2.4 km running performance and absolute $\dot{V}O_2max$ improved inadequately during infantry training, and are influential in determining LMP (Chapter 7 and 8, Rayson *et al.*, 2000; Pandorf *et al.*, 2002) they were prioritised in the MOD PT intervention. In addition, indices of strength did not improve during infantry training and SLS emerged as a determinant of LMP (Chapter 7 and 8), which also provided a rationale for a strength training intervention in the MOD PT programme. Furthermore, a mathematically simulated increase in SLS elicited small improvements in LMP (Chapter 7), and the efficacy of concurrent resistance and endurance training for the purposes of improving LMP are mixed (Williams *et al.*, 2002). This suggested that further investigation on the causal link between strength development and LMP was required.

9.4.3.1 $\dot{V}O_2max$

Absolute and relative $\dot{V}O_2max$ was not different between-groups in the present study. The MOD PT was designed to improve 2.4 km run (a key determinant of $LMP_{2.4}$ - Chapter 7), which is the product of various individual components of fitness (Wilkinson, 1999), one being $\dot{V}O_2max$ (Midgley *et al.*, 2006). In a mathematical simulation a 10 % improvement in $\dot{V}O_2max$ improved middle distance running performance by 5-6% (Wood, 1999), which would indicate that response of $\dot{V}O_2max$ to a PT intervention is relevant to concomitant improvements in 2.4 km run time.

A possible explanation for the absent increase in MOD $\dot{V}O_2max$ compared to CON could have been the same total volume of training (i.e. duration) spent above 85%HRR, across all PT (Table 9.5). This meant there was a lack of differentiation between MOD and CON PT at the highest exercise intensities. The %HRR variable represented the relative stress of PT in the present study, and is equivalent to %$\dot{V}O_2$reserve (Swain and Leutholtz, 1997), or the %$\dot{V}O_2max$ at exercise intensities of ≥75%HRR (Lounana *et al.*, 2007). The time spent above 85%HRR is relevant in the appraisal of the appropriateness of a PT stimulus that targets $\dot{V}O_2max$ as the optimal adaptation in $\dot{V}O_2max$ is achieved at the exercise intensity of 90-100% of $\dot{V}O_2max$ (Wenger and Bell, 1986). The rationale being that the individual components of the oxygen transport chain (Sutton, 1992) are the determinants of $\dot{V}O_2max$, which are challenged severely at exercise intensities approaching $\dot{V}O_2max$ (Bassett and Howley, 1997; 2000).

The specific aim of the HIIT in the present study was to increase $\dot{V}O_2max$ (and by default other parameters of aerobic fitness - Jones and Carter, 2000), as well as 2.4 km running performance. However, $\dot{V}O_2max$ is reputed to be not as sensitive as sub-maximal aerobic measures at responding to PT programmes (Daniels *et al.*, 1979; Jones, 1998). The appropriateness of MOD PT in stimulating $\dot{V}O_2max$ can be appraised by the amount of time spent above 85%HRR, which would indicate the extent with which the central, and peripheral determinants of $\dot{V}O_2max$ were stressed (Midgley *et al.*, 2006). Although the shorter duration and percentage of time spent above 85%HRR in MOD PT could be interpreted as a failure of the MOD HIIT to achieve its goal compared to CON (i.e. stimulate $\dot{V}O_2max$), this may simply be a reflection of the intermittent nature of the HIIT that were not steady-state in nature. Typically interval duration ranged from ~7 s to 360 s, which was markedly shorter than the continuous runs performed in the CON group. These continuous runs would have been of sufficient duration to allow the oxygen uptake kinetics of the $\dot{V}O_2$ slow component further elevate the %HRR (Gaesser and Poole, 1996) (Figure 9.5). Therefore, the %HRR may not be the most appropriate measure to appraise the adequacy of the HIIT intensity due to some interval durations being ≤180 s.

The volume (i.e. duration), frequency, and periodisation of PT were fixed in the present study, thus the manipulation of exercise intensity and exercise modality were the only methods of differentiating PT between-groups. Over the 22 weeks of training, 6.6 hours of HIIT was specifically designed and implemented to be different to CON PT, with the aim of further developing $\dot{V}O_2max$ in the MOD group. This equated to 13 HIIT sessions over the 22 weeks of training. As one to two HIIT sessions are recommended per week for the optimal development of $\dot{V}O_2max$ (Midgely *et al.*, 2006) (any more can produce markers of overtraining (Billat *et al.*, 1999)), the HIIT programme in the present study may have been of inadequate frequency.

A further explanation for the absent $\dot{V}O_2max$ difference between-groups may be that the CON group in the present study were different to controls in previous research that has investigated interval training. Typically the literature show favourable improvements in $\dot{V}O_2max$ as a result of a HIIT intervention, however the efficacy of these findings are often judged against $\dot{V}O_2max$ adaptations of control groups that are trained at 'moderate' intensities (Tabata *et al.*, 1996; Helgerud *et al.*, 2007; Gormley *et al.*, 2008). The CON group in the present study were not 'moderately' trained, as much of the PT programme required subjects to perform either an individual, section, or platoon 'best effort'. For the CON group the standard approach of performing a 'best effort' in British Army PT, coupled with the modest frequency of HIIT in the MOD group may explain the lack of difference observed in $\dot{V}O_2max$ between-groups.

Within-subjects absolute $\dot{V}O_2max$ increased by 1.8% in MOD, which was a *medium* effect, and 4.2% in CON, which was a *medium-large* effect (Cohen, 1992). Relative $\dot{V}O_2max$ did not change in both groups, which might be explained by the increase in body mass over training artificially lowering $\dot{V}O_2max$ (Berg, 2003). The direct assessment of absolute running $\dot{V}O_2max$ in males undergoing British military training is lacking. Seven weeks of US army training significantly increased absolute $\dot{V}O_2max$ by a similar magnitude (3.7%) to the present study (Patton *et al.*, 1980), however 6 weeks of US (Daniels *et al.*, 1979) and 10 weeks of South African training (Gordon *et al.*, 1986) did not increase $\dot{V}O_2max$. Improvements in absolute running $\dot{V}O_2max$, in male regular British Army recruits were as good (Williams *et al.*, 1999), and better (Williams *et al.*, 2002, Williams, 2005) than the present study, however these were indirect estimations based on the composite measure of running performance, not direct physiological measurements.

9.4.3.2 Running performance over 2.4 km

Middle distance running performance is the product of various physiological measures, one of which is $\dot{V}O_2max$ (Joyner, 1991; Brandon, 1995; Wood, 1999; Ingham *et al.*, 2008). This performance measure is an important determinant of $LMP_{2.4}$, as a 10% improvement will mathematically improve $LMP_{2.4}$ by 47 s, which equates to a 5.2% improvement in a 15.0 min $LMP_{2.4}$ time (Chapter 7). However, the small 3.6 % improvement in 2.4 km run time during British Army infantry training previously observed was deemed sub-standard

(Chapter 8), thus HIIT was implemented to target this performance measure of aerobic fitness.

Generally, interval training distance in the present study increased from 0.05 km (week 3) to 1.2 km (week 20), although interval distance was also varied in individual PT lessons to replicate the varying physical demands of middle distance races (Bompa, 1999). This prescription served the three functions of, the creation of progressive overload (Bompa, 1999), the stimulation of specific physiological systems to a greater extent than required during a 2.4 km running test (Laursen and Jenkins, 2002), and the stimulation of anaerobic metabolism, which is a known determinant of middle distance running performance (Billat, 2001b). However, this clear rationale did not deliver as MOD 2.4 km run time was 3.7% slower than CON at T2, and no different at T3.

Middle distance running events are performed at an intensity equivalent to the $v\dot{V}O_2max$ (Billat, 2001a). As running performance is a composite of several different aerobic parameters (Wood, 1999), the efficacy of a training intervention for the development of 2.4 km run time should not solely be appraised by the time (or % time) spent above 85%HRR, as previously discussed (section 9.4.3.1). As $v\dot{V}O_2max$ is a characteristic of 2.4 km running performance (Billat, 2001a) it would seem logical that the velocity at which HIIT is conducted, in relation to an individual's average velocity over 2.4 km would be a practical method (in combination with %HRR) to appraise the appropriateness of HIIT in the present study (Figure 9.4; 9.5).

Whilst, the MOD group spent less time above 85%HRR during the HIIT than CON, the HIIT was primarily designed to develop 2.4 km running time, which is a performance not physiological measure (Wilkinson, 1999). As such all intervals in the present study were conducted with 2.4 km performance in mind, hence were performed at an intensity equal, if not faster than the average velocity an individual completed 2.4 km (Figure 9.4). This would indicate that the intensity of the HIIT programme in the present study was of an appropriate intensity due to the fact that running at or above the $v\dot{V}O_2max$ (or 2.4 km run time pace) stimulates both central (i.e. oxygen transport mechanisms) and peripheral mechanisms (i.e. muscular recruitment) that are coincident with race intensity. This is a well established training technique for improving middle distance running performance

(Billat, 2001b). An example is presented from the GPS data collected (Figure 9.4; 9.5). Here an individual in the MOD group spent a longer duration above 2.4 km race pace compared to CON, despite appearing to spend a shorter duration above 85%HRR. Consequently, the greater absolute stress of these intervals (i.e. time spent above 2.4 km race pace) would seem to indicate that the HIIT was of an appropriate intensity to develop 2.4 km running performance. This would suggest that an inadequate volume (i.e. 6.6 hours vs. 5.2 hours) and/or frequency (i.e. an average of ~1 HIIT session every two weeks) of HIIT explained the lack of improvement in MOD 2.4 km run time.

In retrospect, the absence of a difference in 2.4 km run time between-groups (at the end of training) is of little surprise. Although the HIIT appeared to be conducted at an appropriate intensity (e.g. Figure 9.4, 9.5; Billat, 2001a, 2001b), it has been suggested the frequency of intervals should constitute 1 to 2 session per week (Midgley *et al.*, 2006). This was not the case in the HIIT intervention. Furthermore, the standard infantry PT (i.e. the CON group) was very intermittent. For example, the continuous runs and loaded marches were interval based in nature due to the hilly terrain at ITC(C). Furthermore, the battle PT events such as the steeplechase and the obstacle course were essentially military specific interval training sessions. Therefore, the CON group were exposed to a significant volume of high intensity, intermittent activity. Consequently, this would explain the lack of differentiation in the cardiovascular / physiological demands between-groups, which may explain the lack of difference in performance gains during training. However, this does not explain the unexpected observation that 2.4 km run time was slower in the MOD group at T2, which was something of an anomaly.

Almost two-thirds of subjects (61% MOD) and (58% CON) completed the 2.4 km run within the infantry pass standard at baseline (i.e. ≤10.5 mins). By the end of training 100 % of subjects achieved this infantry pass standard. This was categorical (i.e. pass or fail) endorsement that infantry PT has adequate progression to develop running performance (i.e. aerobic fitness) to the minimum fitness standard by the end of training. Within-subjects, the improvements in run time of 9.1% for MOD, and 8.2% for CON were better than single entry (Rayson *et al.*, 2002c) and Parachute regiment recruits (Wilkinson *et al.*, 2003), but comparable to officer Cadets (Harwood *et al.*, 1999) and infantry foot guards (Carter *et al.*, 2006). These changes in 2.4 km run time indicate that the gains made in running performance (in both the CON and the MOD infantry PT) meet with the infantry standards. Furthermore, these improvements are as good as those previously reported (Harwood *et al.,* 1999; Rayson *et al.,* 2002c; Wilkinson *et al.,* 2003; Carter *et al.,* 2006). Chapter 8 reported small improvements of 3.6% in run time, which appears to have been addressed. However, the purpose of improving 2.4 km running performance was to further improve $LMP_{2.4}$, which did not happen in comparison to Chapter 8 (i.e. 6.1% to 9.9% vs. 7.0%). A possible explanation for this could be that the model of LMP, which was based on week-1 and 2 of infantry training (Chapter 7) was unable to model the training response of LMP (section 7.2.2) as others have (Williams and Rayson, 2006).

9.4.3.3 Dynamic and isometric strength

None of the strength measures were different between-groups, which may have blunted the hypothesised LMP gains in the MOD group (Chapter 7; Kraemer *et al.,* 1987, 2001, 2004; Rayson *et al.,* 2002a, 2002b; Williams and Rayson, 2006).

As with HIIT programme, the lack of strength difference between-groups may have been a result of inadequate differentiation between the muscular strength and endurance training stimuli. Although ITC(C) has a RT gymnasium, it does not have the capacity to deal with a high number of users. Furthermore, it was important that the RT intervention in the present study could be applied to the wider recruit population at ITC(C) if found to be beneficial. Typically, the core of MOD RT exercises consisted of (i.e. press-ups, heaves and squats) (section 9.2.4.2). These exercises were chosen because intensity could be increased (to an extent – see section 9.2.4.2), and these were simple field exercises that required few resources. The CON group's muscular strength and endurance stimulus equated to traditional style circuit training that is reputed to create resistances of between 25% to 65% of one repetition maximum (Bompa, 1999).

As a result it is likely that the application of the principles of training (i.e. overload, specificity, progression, individualisation, adaptation, and maintenance), and the acute training programme variables (i.e. muscle action, rest periods, loading and volume, repetition velocity, exercise selection and order, and frequency) lacked differentiation between-groups (Bird *et al.,* 2005). A lack a variation in the application of 'acute programme variables' can explain the lack of between-group strength differences as it is these that create an 'upstream' stimulus, which creates an ordered cascade of 'downstream' events that include; system responses (e.g. hormonal, neural, cardiovascular), signalling processes (e.g. receptor interaction, cellular signalling), genetic responses (e.g. transcription, translation), protein metabolism (i.e. myofibrillar protein production), and finally functional adaptation (e.g. muscle growth) (Spiering *et al.*, 2008). Additionally, 15 RT sessions were conducted over the 22 weeks of training, which equated to 8.2 hours of training. This frequency of RT was below the two sessions per week recommended for strength development (ACSM, 1998). Hence, it appears that the combination of an inadequate RT intensity and frequency are the most likely explanations of the absent between-group differences observed in strength in the present study.

In addition, strength development is inhibited during a concurrent resistance and endurance training programme (like the present study), which may be a function of overtraining, a conflict in the chronic adaptation mechanisms, and the acute effect of fatigue on subsequent training quality (Leveritt *et al.*, 1999). However, concurrent strength and endurance training studies have reported muscular strength increases to coincide with LMP improvements (Kraemer *et al.*, 1987, 2001, 2004; Knapik and Gerber, 1996; Harman *et al.*, 1998, Reynolds *et al.*, 2001). Unfortunately, these studies were not conducted on military recruits during initial training, which is known to increase catabolic, decrease anabolic hormone secretion, and cause a reduction in muscular strength and power (Nindl *et al.*, 2007).

In a similar experimental design to the present study, Williams *et al.*, (2002) observed strength and LMP gains in British Army recruits. This study demonstrated that it was possible to increase boxlift and incremental dynamic lift strength by a markedly greater magnitude than controls, which is in contrast to the findings of the present study. A possible explanation for this could be the more structured, focused and intense nature of the Williams *et al.*, (2002) RT intervention compared to the present study. Williams *et al.*, (2002) individually prescribed loads at the intensity of six repetition maximum, two times per week. Although the intensity was progressed in the present study, this was limited. The only exercise to approach the intensity of Williams *et al.*, (2002) would have been heaves, in which subjects had a 5 to 15 repetition maximum. In the present study, progression was typically achieved by an increase in volume (i.e. number of sets and repetitions), which was an inadequate stimulus.

9.4.4 Determinants of loaded marching performance not targeted by the modified physical training intervention.

9.4.4.1 Loaded marching economy

Whilst LME was established as an influential determinant of LMP (Chapter 7), the marked response of this measure to British Army infantry training (Chapter 8), and the complex nature of the mechanisms that underpin exercise economy (Saunder *et al.,* 2004) meant a targeted intervention to improve LME was both unnecessary, and unrealistic. Furthermore, for reasons previously stated between-group comparisons could not be made on LME (section 9.2.3). Nonetheless, within-group comparisons revealed improvements of 14.7% for MOD, and 11.7% for CON groups over training, which were both *large* effect sizes (Cohen, 1992). Although these were numerically larger than the 9.6% improvement previously observed (Chapter 8), they were of an equivalent effect size (Table 10.1).

In conclusion, British Army infantry recruits are physically and physiologically homogenous at the start of training, which indicates that the British Army physical selection standards for recruits appear to work. A modified physical training programme designed to increase $\dot{V}O_2max$, 2.4 km run time and muscular strength was no better than existing infantry physical training at improving the determinants of loaded marching performance, and 2.4 km loaded marching performance. A possible explanation of this finding may have been a lack of differentiation in the intensity, frequency and volume of the two physical training programmes. Overall the results of the present study are positive as on paper the total volume of infantry physical training decreased 30 % in comparison to Chapter 8, without an observed reduction in physical gains (within-subjects). However, with the exception of leg strength, there is still a lack of development in indices of strength.

CHAPTER 10

GENERAL DISCUSSION

10.1 Improving loaded marching performance

The purpose of this thesis was to improve physical training (PT) for loaded marching performance (LMP) during British Army infantry training. This process is schematically represented (Figure 10.1). Recruits at the Infantry Training Centre (Catterick) (ITC(C)) were selected as volunteers because the loaded marching physical output standard is the most physically demanding in this group, across the British Army. As such, the issue of improving PT for LMP was most pertinent to infantry recruits.

The literature revealed a number of determinants of LMP (Rayson *et al.*, 1993, 2000, 2002a; Harman and Frykamn, 1995; Pandorf *et al.*, 2002; Williams and Rayson, 2006). However, a gap in the understanding of LMP emerged, as both maximal (Simpson *et al.*, 2006) and sub-maximal physiological measures of aerobic fitness were scarce in relation to loaded marching (LM). This was surprising considering other exercise modalities such as running recognise the importance of $\dot{V}O_2max$, and economy on performance (Joyner, 1991; Brandon, 1995; Fallowfield and Wilkinson, 1999; Ingham *et al.*, 2008).

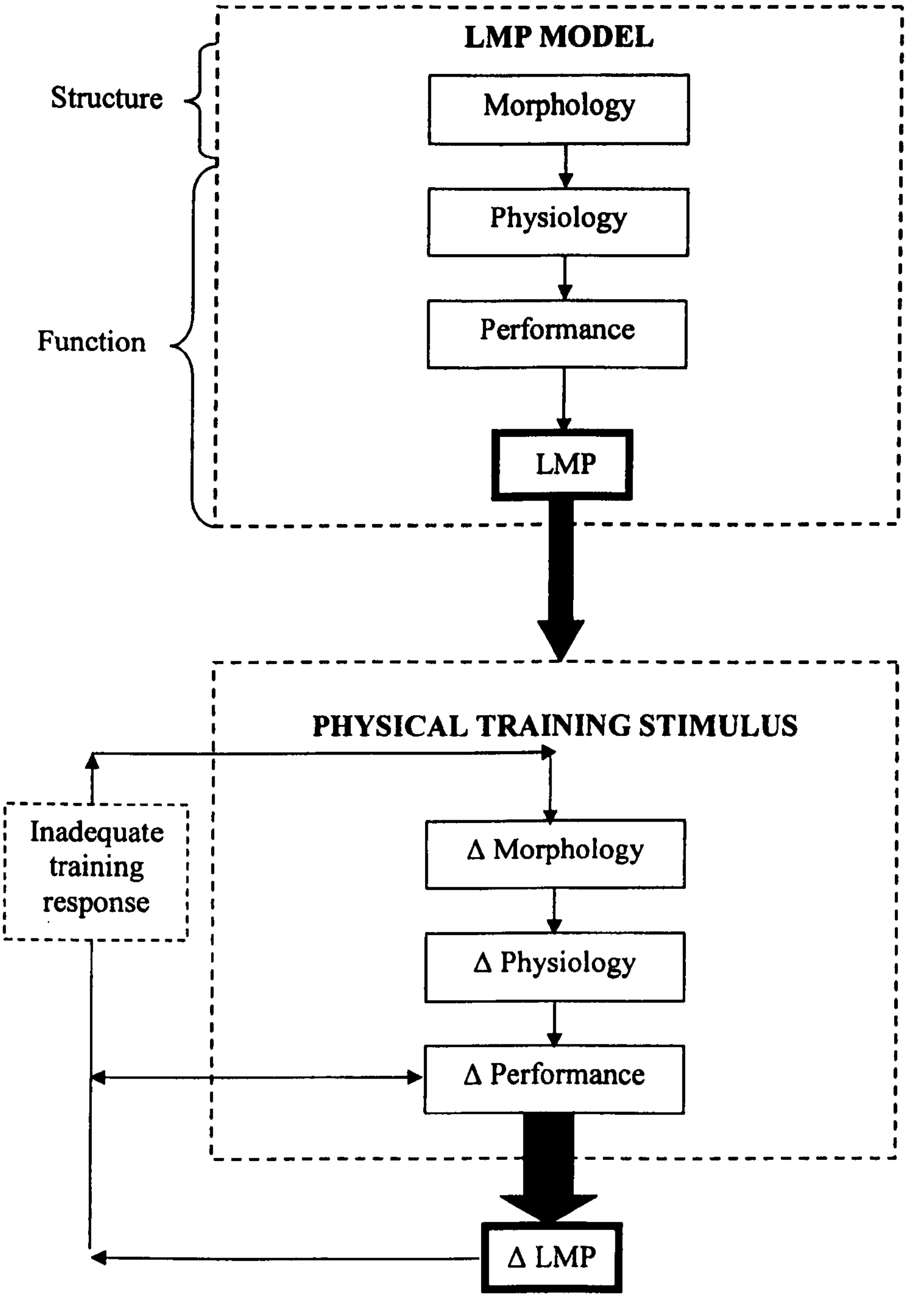


Figure 10.1. The process of the development, appraisal, and re-adjustment of a physical training stimulus for the purposes of improving loaded marching performance.

Step 1 was to develop a model of LMP (Figure 10.1). Chapters 4, 5 and 6 interrogated the validity and reliability of the accepted and potential determinants of LMP. The test battery was then applied to the field were the actual determinants of LMP were identified in a sample of British Army infantry recruits (Chapter 7) (Figure 10.1). In the development of the LMP model it should be understood that the structure of a biological organism (e.g. body size, muscle morphology, myocardium size etc.) will dictate its function (e.g. strength, running performance etc.) (Coyle, 1995). Thus, performance measures (e.g. 2.4 km run time) will be the product of individual components of fitness (Wilkinson, 1999), such as morphological (e.g. fat free mass etc.) and/or physiological measures (e.g. $\dot{V}O_2max$ etc.) (Figure 10.1). This hierarchy of increasing importance from structural to functional measures during the modelling of LMP was supported in the LMP literature, which reported 2.4/3.2 km LMP as a determinant of both 29 km (Simpson *et al.*, 2006) and 12.8 km LMP (Griggs *et al.*, 2008). This was also partially supported in the present thesis, as the performance measures of 2.4 km run time emerged as a determinant of 2.4 km LMP.

Step 2 was to quantify the extent both LMP, and the individual determinants of LMP responded to a PT stimulus (Figure 10.1). First, it was important that the criterion LMP test was valid. The most valid LMP test for infantry recruits is the Basic Combat Fitness Test (BCFT) (12.8 km, 25 kg march in no longer than 2 hours), which is conducted as a squadded effort in week-22 of training. However, tracking the LMP response during infantry training required a baseline test at the very beginning of training that could be repeated subsequently. Conducting a baseline BCFT would have been unacceptable due to the potential for injury among naïve load-carriers. As such the Advanced Combat Fitness Test 1 speed march was chosen (2.4 km, carrying 20 kg - British Army, 2002). The magnitude of improvement in LMP, and each determinant of LMP could then be appraised over training.

Step 3 involved appraising the adequacy of LMP and the determinants of LMP training responses, and suggesting relevant modifications to the PT stimulus if inadequate responses were observed (Figure 10.1). Smaller than anticipated changes were observed in $\dot{V}O_2max$, 2.4 km run time, and muscular strength in Chapter 8, hence the physical training stimulus was modified to further improve LMP in Chapter 9, as per Figure 10.1.

Table 10.1. The effect sizes of the three infantry platoons followed during this thesis. LMP and the determinants of LMP are presented. The effect sizes (i.e. training responses) are from the beginning (T1) to the end of training (T3).

	CH 8 PT	CH 9 PT (Control)	CH 9 PT (Modified)
Loaded marching fitness			
2.4 km loaded marching performance	-0.59*	-0.58*	-0.62*
Loaded marching economy (l·min^{-1})	-0.60*	-0.60*	-0.61*
Running fitness			
2.4 km run performance	-0.48*	-0.62*	-0.61*
$\dot{V}O_2$max (l·min^{-1})	-0.31	-0.49*	-0.24
Strength			
Peak chest	-0.28	-0.22	-0.24
Peak back	-0.05	-0.09	-0.26
Peak leg	-0.30	-0.42*	-0.51*
Peak static lift strength	-0.17	-0.05	-0.16
Anthropometry			
Body mass	-0.42*	-0.21	-0.32*
Estimated percentage fat	-0.42*	-0.29	-0.47*
Estimated fat mass	-0.25	-0.18	-0.33*
Estimated fat free mass	-0.53*	-0.50*	-0.59*

* = statistical significance ($P<0.0167$)
r = 0.10 (small effect)
r = 0.30 (medium effect)
r = 0.50 (large effect) (Cohen, 1992)

The magnitude of the change in LMP, and the determinants of LMP can be compared between the three different infantry PT programmes observed in this thesis by standardising all training responses with the effect size statistic (Field, 2005) (section 3.11.3) (i.e. standard PT 2006/07 from Chapter 8; standard PT 2008 from Chapter 9 control group; and modified PT 2008 from Chapter 9 modified group). This statistic allowed both the different PT programmes, and different measures (i.e. loaded marching fitness, running fitness, strength and anthropometry) to be compared using a standardised unit. All comparisons were made between baseline (T1) and end of infantry training (T3) (Table 10.1).

An interesting finding from these data are that the effect sizes reported in Chapter 9 are equivalent, or slightly larger than those reported in the Chapter 8, despite 30 % fewer PT lessons delivered in both the Chapter 9 PT programmes. The volume (meaning the duration of timetabled PT) was on average 4.1 $h \cdot week^{-1}$ in Chapter 8 compared to 2.6 $h \cdot week^{-1}$ in Chapter 9. A cost:benefit appraisal would suggest the PT dose in Chapter 8 may have been unnecessarily high, especially when combined with all other physical aspects of infantry training. In support of this finding, reduced running mileage in US Army Basic Combat Training produced higher pass rates in Army Physical Fitness Test (which is similar to the British Army's PFT), and reduced injury rates (Knapik *et al.*, 2002; 2005). Furthermore, an increase in training volume of 50%, from 1.5 $h \cdot day^{-1}$ to 3.0 $h \cdot day^{-1}$ does not further improve swimming performance (Costill *et al.*, 1991).

It is possible the volume of PT in Chapter 8 was too high. Whilst a training volume of 4.1 $h \cdot week^{-1}$ may seem minimal compared to the 90-100 $miles \cdot week^{-1}$ an elite endurance athlete may run (Conley *et al.*, 1984; Jones, 1998; Wilkinson, 1999), it should be noted that PT constitutes just one-tenth of the training programme at ITC(C). Hence, significant training stimuli exist in other aspects of training, which include drill, bayonet lessons, rifle range, adventurous training and the field exercises (Rayson *et al.*, 2002c; Carter *et al.*, 2006). As such military training is generally regarded as arduous, causing an increase in catabolic and decrease in anabolic hormone secretion, as well as a decrease in physical performance (Nindl *et al.*, 2007). Such disruption to rest recovery cycles is not optimal for physical development (McCafferty and Horvath, 1977). Therefore, the 30% extra PT conducted in Chapter 8 may have been enough to disrupt physical development.

An alternative explanation for equivalent training responses in Chapter 9 (compared to Chapter 8), despite a decrease in the volume of PT could be the changes that were made to the content of the ITC(C) PT programme between Chapters 8 and 9. Although the total number of PT lessons reduced by 30% between Chapters 8 and 9, the content of the lessons in Chapter 9 was used more effectively. For example the full 1.3 hours of a typical double PT lesson was used in Chapter 9, instead of just 40-50 minutes that typically occurred in Chapter 8. This was due to the implementation of concurrent training methods in Chapter 9 (e.g. run lessons were proceeded with circuit training lessons etc.). This basic change meant that it was feasible that the total volume (i.e. time) of formal PT in Chapter 9 was in fact similar to Chapter 8. Hence this may explain the similar effect sizes, despite the apparent larger training volume (on paper) in Chapter 8.

In relation to the LMP model, the largest effect size was observed in measures of LM fitness (i.e. LMP and LME) (Table 10.1). These were classified as *large* effects (Cohen, 1992). Considering the importance of LM to the infantry this was a positive finding. However, the LMP effect size did not increase as PT evolved over the duration of this thesis, which indicated the process of experimental investigation outlined and adhered to in Figure 10.1 was unsuccessful. For example, the original PT programme (Chapter 8) was modified locally at ITC(C) from the findings of Chapter 8, to produce the Chapter 9 Control group PT programme. Subsequently, the PT intervention (Chapter 9 Modified group) was designed to improve LMP to a greater extent than Chapter 9 Control group. Theoretically, LMP effect size should have increased during this process, however it did not.

The *large* effect size observed in LME was proportional to LMP. Furthermore, LME was the single most influential determinant of LMP to be included in LMP model (Chapter 7). These are the first reported findings to track training responses, and mathematically model both LMP and LME concurrently. At the two extremes of the performance continuum, LME (i.e. 25 kg carried at 6.5 $km{\cdot}h^{-1}$) varied from 41.2 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ to 27.4 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ at the start of infantry training, which may arbitrarily be classified as poor and good LME, respectively. This difference of 13.8 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ coincided with a 5:09 min:s difference in 2.4 km LMP time (18:11 min:s vs 13:02 min:s) (Chapter 8). This may indicate that marked improvements of 4.9 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ in LME (Chapter 8) may be insufficient to allow an individual with poor LME perform outstandingly well in a LMP test. The novel findings on LME adds much needed attention to the paucity of research in this area (Gutenkunst *et al.*, 2006; Simpson *et al.*, 2006; Griggs *et al.*, 2008), and provides a fresh and contemporary understanding of LMP. However, whilst LME maybe an influential determinant of LMP, the large effect size observed over training is likely to be a combination of physiological and non-physiological adaptations (i.e. kinematics, kinetics, flexibility, ground reaction forces) (Morgan *et al.*, 1989; Anderson, 1996; Saunders *et al.*, 2004). This would mean that interventions designed to further improve LME for LMP would be hard to identify.

A *large* effect size was also observed in 2.4 km running performance in Chapter 9, but not Chapter 8, which may reflect changes in the ITC(C) PT programme following Chapter 8 recommendations. However, the effect sizes were not different between the control and modified PT groups in Chapter 9, which indicated the High Intensity Interval Training (HIIT) was unsuccessful. Furthermore, the large effect size ($r=-0.49$) in the control group's absolute $\dot{V}O_2max$, but not the modified group was contrary to the modified PT intervention aim. These findings were unexpected as HIIT is effective at improving parameters of maximal running performance (Billat, 2001; Midgely *et al.*, 2006), which is known to underpin LMP (Chapter 7; Rayson *et al.*, 2000; Pandorf *et al.*, 2002; Rayson *et al.*, 2002a; Rayson *et al.*, 2002b; Williams and Rayson, 2006). Thus, whilst the targeted HIIT sessions were of an adequate intensity (Table 9.4; Figure 9.4; 9.5), the frequency (i.e. 13 sessions over 22 weeks) and volume (i.e. 6.6 hours over 22 weeks) were inadequately different from controls to further improve parameters of aerobic fitness, hence LMP.

The effect size for strength development ranged from *small* to *medium* during infantry training. The implications of this finding for improving PT for LMP remain unclear. For example, strength was a determinant of LMP, but its overall influence on LMP was small compared to LME and 2.4 km run time (Chapter 7). Furthermore, a limited amount of research has attempted to quantify the efficacy of concurrent resistance and endurance PT for LMP (Kraemer *et al.*, 1987, 2001, 2004; Williams *et al.*, 2002, 2004), hence the importance of strength development for LMP requires further confirmation. The suggested rationale for strength's importance to LMP has logic. Loaded marching is typically conducted with absolute loads (e.g. 25 kg) therefore larger individuals carry a smaller percentage of their body mass. As muscle cross- sectional area correlates moderately to muscle function (Maughan *et al.*, 1984), as well as improvements in muscle function (Narici *et al.*, 1989), strength's relationship with LMP may simply reflect that larger (hence possibly stronger) individuals are better load-carriers. This also would explain the association between anthropometric measures, absolute $\dot{V}O_2max$ and LMP (Rayson *et al.*, 1993; Frykman and Harman, 1995; Rayson *et al.*, 2000; Pandorf *et al.*, 2002; Williams and Rayson, 2006; Griggs *et al.*, 2008). Consequently, the increases in body mass and fat free mass in the present study are positive training outcomes for improving LMP (Table 10.1).

The *small* to *medium* effect size observed in strength during infantry training seems inadequate (Table 10.1). A possible explanation for such a small effect could be the body weight circuit training that is typically prescribed during infantry training. The resistance training intervention in Chapter 9 addressed this inadequacy with 15 targeted sessions, totalling 8.2 hours over the 22 weeks of infantry training. However, in general this was unsuccessful. A possible explanation for this could be that the intensity of loads lifted progressed little (i.e. % of 1 repetition maximum). As such the training volume (i.e. number of repetitions) was relied upon to apply the training principle of progression. This seems a feasible explanation for the lack of strength gain in the Chapter 9 intervention as load size is known to influence hormonal, neural and hypertrophic responses (Bird *et al.*, 2005), which are mechanisms of strength development (Jones *et al.*, 1989). Additionally, the concurrent endurance and resistance training delivered in military training, as well as increased cortisol and decreased testosterone secretions, would only exacerbate the problem of strength development in a military training context (Leveritt *et al.*, 1999; Nindl *et al.*, 2007).

10.2 Problems improving physical training for loaded marching performance in the context of military training

Improving PT for LMP in military training has a valid rationale (Chapter 1). In this thesis it was suggested that LMP might be improved further if a more focused PT programme was implemented that developed, $\dot{V}O_2$max and 2.4 km run time through HIIT, and muscular strength through a simple resistance training intervention. Theoretically, maximal gains in these components of fitness might be achieved if the correct volume, overload, progression, specificity, and periodisation of PT were prescribed (Bompa, 1999). This does not seem to be feasible during military recruit training.

There is a relevant, philosophical and pragmatic argument to claim that maximising physical development during military training is unnecessary, as there are no requirements to be at peak fitness for job related roles (British Army, 2002). Rather an individual is expected to attain a specified minimum fitness standard for their chosen Career Employment Group. This is different to an athlete, who by trial and error over many years will learn how to maximise the interaction of genetic endowment and environment. British

Army recruits are first and foremost training to become soldiers, not athletes. The ability to 'soldier' is assessed with physical tests, but also includes other core skills such as, skills at arms, field craft, and mental robustness. However, whilst the maximisation of physical development is not a requirement of military training it does have a role to play. Improving PT will develop greater physical robustness, and as loaded marches typically precede an enemy contact, recruits / soldiers will be 'fitter to fight' upon arrival at an operational destination (Washington DC, 1990), which could be potentially life saving.

Physical training will alter morphological and physiological characteristics, which should improve LMP (Coyle, 1995; Figure 2.1), however this is confounded by two basic factors in military training. First, the requirement to test essential soldiering skills throughout the duration of a training course, and second, the large recruit numbers that go through training.

First, the essential infantry skills tested during the training at ITC(C) are: skills at arms; field craft; physical robustness and mental robustness (British Army website, 2007). The arduous nature of training tests physical and mental robustness, as recruits are likely to feel under constant pressure, fatigued, and in sleep / nutrition deprived states. This environment elicits marked hormonal changes that are linked to losses in body mass and explosive power and strength (Nindl *et al.*, 2007). Such conditions are far from optimal for physical development, which requires adequate rest and recovery cycles (McCafferty and Horvath, 1977). Consequently, testing the core soldiering principles during infantry training is a less than ideal environment for these morphological, physiological, and physical adaptations to manifest.

Second, is the scale of British Army infantry training. The ITC(C) trains in the order of 2900 recruits every year, and has the capacity to intake 140 new recruits every 2 weeks (British Army website, 2007). This scale of training necessitates that PT lessons are conducted at platoon level (40 to 50 recruits). Even though individual training prescription may help prioritise the correct type of training for improving LMP (Williams *et al.*, 2004) the associated time and resource costs make this unlikely. A compromise would be to stream recruits into ability groups based on aerobic fitness (Bilzon 2003; Knapik 2006). This would standardise training dose in relative terms, however it has not been implemented at ITC(C). Group training prescription means that variations in relative stress occur on an individual level, which may translate in to variations in training responses. For example grouped PT elicits a larger training response in the least fit recruits, and a smaller (sometimes even detraining) response in the most fit (Vogel *et al.*, 1978; Stacy *et al.*, 1982; Gordon *et al.*, 1986; Trank *et al.*, 2001), and may also explain why 2.4 km run time is an injury risk factor in military training (Blacker *et al.*, 2005).

Consequently, the nature and logistical constraints of military training means adapting the PT stimulus to optimally develop morphologically, physiologically, and physically for LMP is challenging. Furthermore, the excessive noise created by a standardised absolute PT stress, and the additional aspects of infantry training that are physically demanding, but not part of formal PT (i.e. exercise, drill etc.) make this a difficult area of research to control, hence accurately evaluate.

10.3 Conclusion

In conclusion, this thesis developed a valid and repeatable test battery that subsequently evaluated the baseline fitness, and training response of new entry British Army infantry recruits. Mathematical modelling revealed loaded marching economy, 2.4 km run, and static lift strength to be important determinants of loaded marching performance. The improvements of $\dot{V}O_2max$, 2.4 km run time, and strength during infantry training were sub-optimal. Subsequently, high intensity interval training and a simple body weight resistance training programme was prescribed to further develop these determinants of loaded marching performance. However, this intervention was no better at improving the determinants of loaded marching performance, or loaded marching performance itself than the standard physical training programme. Interestingly, this thesis revealed a specific, sub-maximal loaded marching measure (i.e. loaded marching economy) was an important determinant of loaded marching performance, which provided new and novel information to the scientific community. Finally, attempting to optimally develop physical fitness within the context of military training is confounded by the absolute training stress prescribed, the sub-optimal conditions for physical development, and a significant proportion of training that is very demanding, but not part of the formal physical training programme. Consequently, if this hypothesis were to be tested in the future a holistic approach should be taken (i.e. monitor all physically demanding aspects of training), and an intervention should be devised that is markedly different (either in content and/or volume) than the existing physical training programme. Such a step would need to be taken to more effectively attempt to improve physical training for loaded marching performance.

CHAPTER 11

REFERENCES

Abernethy, P. & Wilson, G. (2000). Introduction to the assessment of strength and power. In. C. J.Gore (Ed.), *Physiological Tests for Elite Athletes: Australian Sports Commission* (pp. 147-154). Champaign, IL: Human Kinetics.

Acevedo, E. O. & Goldfarb, A. H. (1989). Increased training intensity effects on plasma lactate, ventilatory threshold, and endurance. *Medicine and Science in Sports and Exercise*, 21(5), 563-568.

Altman, D. G. & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *The Statistician*, 32, 307-317.

American College of Sports Medicine (1998). The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Medicine and Science in Sports and Exercise*, 30(6), 975-991.

Anderson, T. (1996). Biomechanics and running economy. *Sports Medicine*, 22(2), 76-89.

Andersen, P. & Henriksson, J. (1977). Capillary supply of the quadriceps femoris muscle of man: Adaptive response to exercise. *Journal of Physiology*, 270, 677-690.

Armstrong, L. E. & Costill, D. L. (1985). Variability of respiration and metabolism: Responses to submaximal cycling and running. *Research Quarterly for Exercise and Sport*, 56(2), 93-96.

Atkinson, G. & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26(4), 217-238.

Atkinson, G. & Nevill, A. (2000). Typical error versus limits of agreement. *Sports Medicine*, 30(5), 375-377.

Atkinson, G., Davison, R., Passfield,L., & Nevill, A.M. (2003). Could the correlation between maximal oxygen uptake and "economy" be spurious? *Medicine and Science in Sports and Exercise*, 35(7), 1242-1243.

Atkinson, G., Davison, R. C. R., & Nevill, A. M. (2005). Performance characteristics of gas analysis systems: What we know and what we need to know. *International Journal of Sports Medicine*, 26(Suppl 1), S2-S10.

Åstrand, P. O. (1952). *Experimental Studies of Physical Working Capacity in Relation to Sex and Age*. Copenhagen: Enjar Munksgaard.

Åstrand, I. (1960). Aerobic capacity in men and women, with special reference to age. *Acta Physiologica Scandinavica*, 192, 1-92.

Åstrand, P. O. & Saltin, B. (1961). Maximal oxygen uptake and heart rate in various types of muscular activity. *Journal of Applied Physiology*, 16(6), 977-981.

Babineau, C., Léger, L., Long, A., & Bosquet, L. (1999). Variability of maximum oxygen consumption in various metabolic systems. *Journal of Strength and Conditioning Research*, 13(4), 318-324.

Balke, B. & Ware, R. W. (1959). An experimental study of physical fitness of Air Force personnel. *U.S. Armed Forces Medical Journal*, 10, 675-688.

Bassett, D. R. & Howley, E. T. (1997). Maximal oxygen uptake: "Classical" versus "contemporary" viewpoints. *Medicine and Science in Sports and Exercise*, 29(5), 591-603.

Bassett, D. R. & Howley, E. T. (2000). Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Medicine and Science in Sports and Exercise*, 32(1), 70-84.

Bassett, D. R., Howley, E. T., Thompson, D. L., King, G. A., Strath, S. J., McLaughlin, J. E., & Parr, B. B. (2001). Validity of inspiratory and expiratory methods of measuring gas exchange with a computerized system. *Journal of Applied Physiology*, 91, 218-224.

Beaver, W. L., Wasserman, K., & Whipp, B. J. (1986). Bicarbonate buffering of lactic acid generated during exercise. *Journal of Applied Physiology: Respiratory Environmental and Exercise Physiology*, 60(2), 472-478.

Berg, K., Latin, R. W., & Hendricks, T. (1995). Physiological and physical performance changes in female runners during one year of training. *Sports Medicine Training and Rehabilitation*, 5, 311-319.

Berg, K. (2003). Endurance training and performance in runners: Research limitations and unanswered questions. *Sports Medicine*, 33(1), 59-73.

Billat, V. L., Flechet, B., Petit, B., Muriaux, G., & Koralsztein, J. (1999). Interval training at VO_{2max}: Effects on aerobic performance and overtraining markers. *Medicine and Science in Sports and Exercise*, 31(1), 156-163.

Billat, V. L. (2001a). Interval training for performance: A scientific and empirical practice. Special recommendations for middle- and long distance running. Part I: Aerobic interval training. *Sports Medicine*, 31(1), 13-31.

Billat, V. L. (2001b). Interval training for performance: A scientific and empirical practice. Special recommendations for middle- and long-distance running. Part II: Anaerobic interval training. *Sports Medicine*, 31(2), 75-90.

Bilzon, J. L. J., Allsopp, A. J., & Tipton, M. J. (2001). Assessment of physical fitness for occupations encompassing load-carriage tasks. *Occupational Medicine*, 51(5), 357-361.

Bilzon, J. L. J. (2003). 'Streaming' in British Army recruit training: Options for change. *HQ ATRA Report*, 2003.001, 1-19.

Bingisser, R., Kaplan, V., Scherer, T., Russi, E. W., & Bloch, K. E. (1997). Effect of training on repeatability of cardiopulmonary exercise performance in normal men and women. *Medicine and Science in Sports and Exercise*, 29(11), 1499-1504.

Birch, K., Sinnerton, S., Reilly, T., & Lees, A. (1994). The relation between isometric lifting strength and muscular fitness measures. *Ergonomics*, 37(1), 87-93.

Blacker, S., Wilkinson, D., Bilzon, J., & Rayson, M. (2005). Risk factors for the development of training injuries among British Army recruits. *Optimal Performance Ltd.*

Bland, J. M. & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 8(1), 307-310.

Bland, J. M. & Altman, D. G. (1995). Comparing two methods of clinical measurement: A personal history. *International Journal of Epidemiology*, 24(3), S7-S14.

Bland, J. M. & Altman, D. G. (1996a). Transforming data. *British Medical Journal*, 312, 770.

Bland, J. M. & Altman, D. G. (1996b). Logarithms. *British Medical Journal*, 312, 700.

Bland, J. M. & Altman, D. G. (1996c). Measurement error. *British Medical Journal*, 312.

Bland, J. M. & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8, 135-160.

Bland, J. M. & Altman, D. G. (2003). Applying the right statistics: Analyses of measurement studies. *Ultrasound Obstet Gynecol*, 22, 85-93.

Bland, J. M. & Altman, D. G. (2007). Agreement between methods of measurement with multiple observations per individual. *Journal of Biopharmaceutical Statistics*, 17, 571-582.

Blazevich, A. J., Gill, N., & Newton, R. U. (2002). Reliability and validity of two isometric squat tests. *Journal of Strength and Conditioning Research*, 16(2), 298-304.

Bompa, T. O. (1999). *Periodization: Theory and Methodology*. Champaign, IL: Hunam Kinetics.

Brandon, L. J. (1995). Physiological factors associated with middle distance running performance. *Sports Medicine*, 19(4), 268-277.

British Army (2002). Test protocols and administrative instructions for Physical Selection Standards (Recruits) fitness tests. *Fit to Fight: Pamphlet Two.*

British Army (2007). Infantry Training Centre Catterick. http://www.army.mod.uk/hqschinf/itc_catterick/itc_information/combat_infantryman_s_course/index.htm.

Brock, J. R. & Legg, S. J. (1997). The effects of 6 weeks training on the physical fitness of female recruits to the British Army. *Ergonomics*, 40(3), 400-411.

Buchfuhrer, M. J., Hansen, J. E., Robinson, T. E., Sue, D. Y., Wasserman, K., and Whipp, B. J., (1983). Optimizing the exercise protocol for cardiopulmonary assessment. *Journal of Applied Physiology* 55, 558-564.

Bygrave, S., Legg, S. J., Myers, S., & Llewellyn, M. (2004). Effect of backpack fit on lung function. *Ergonomics*, 47(3), 324-329.

Byrd, R., Chandler, T. J., Conley, M. S., Fry, A. C., Haff, G. G., Hatfield, F., Kirksey, K. B., McBride, J., McBride, T., Newton, H., OBryant, H. S., Pierce, K. C., Plisk, S., Ritchie-Stone, M., & Wathen, D. (1999). Strength training: Single versus multiple sets. *Sports Medicine*, 27(6), 409-412.

Carter, J., & Jeukendrup, A. E. (2002). Validity and reliability of three commercially available breath-by-breath respiratory systems. *European Journal of Applied Physiology and Occupational Physiology*, 86, 435-441.

Carter, J. M., Izard, R. M., Blacker, S. D., Wilkinson, D. M., Rayson, M. P., Bilzon, J. L. J. (2006). The physical demands of the Combined Infantryman's Course for the Foot Guards at ITC Catterick. *Optimal Performance Ltd.Report.*

Christie, C. J. & Scott, P. A. (2005). Metabolic responses of South African soldiers during simulated marching with 16 combinations of speed and backpack load. *Military Medicine*, 170(7), 619-622.

Cohen, J. (1992). "A power primer." *Psychological Bulletin*, 112(1), 155-159.

Coldwells, A., Atkinson, G., & Reilly, T. (1994). Sources of variation in back and leg dynamometry. *Ergonomics*, 37(1), 79-86.

Conley, D.L., & Krahenbuhl, G. S. (1980). Running economy and distance running performance of highly trained athletes. *Medicine and Science in Sports and Exercise*, 12(5), 357-360.

Conley, D. L., Krahenbuhl, G. S., & Burkett, L. N. (1981). Training for aerobic capacity and running economy. *Physician and Sports Medicine*, 9(4), 107-115.

Conley, D. L., Krahenbuhl, G. S., Burkett, L. N., & Millar, A. L. (1984). Following Steve Scott: Physiological changes accompanying training. *Physician and Sports Medicine*, 12(1), 103-106.

Costill, D. L., Thomas, R., Robergs, R. A., Pascoe, D., Lambert, C., Barr, S., & Fink, W. J. (1991). Adaptations to swimming training: Influence of training volume. *Medicine and Science in Sports and Exercise*, 23(3), 371-377.

Coyle, E. F. (1995). Integration of the physiological factors determining endurance performance ability. *Exercise and Sport Science Reviews*, 23, 25-63.

Crapo, R. O., Hankinson, J. L., Irvin, C., MacIntyre, N. R., Votel, K. Z., & Wise, R. A. (1995). Standardisation of spirometery - 1994 update. *American Journal of Respiratory Critical Care Medicine*, 152, 1107-1136.

Crouter, S. E., Antczak, A., Hudak, J. R., Della Vale, D. M., & Haas, J. D. (2006). Accuracy and reliability of the ParvoMedics TrueOne 2400 and MedGraphics VO2000 metabolic systems. *European Journal of Applied Physiology*, 98, 139-151.

Cullum, M. G., Welch, H., and Yates, J. W. (1999). Evalation of an automated metabolic cart compared to the Douglas bag measurement of VO_2. *Medicine and Science in Sports and Exercise* 31[Suppl 5], S302.

Dallal (1985). ftp://ftp.simtel.net/pub.simtelnet/msdos/statstcs/size102.zip.

Daniels, W. L., Kowal, D. M., Vogel, J. A., & Stauffer, R. M. (1979). Physiological effects of a military training program on male and female cadets. *Aviation, Space and Environmental Medicine*, 50(6), 562-566.

Davies, B., Daggett, A., Jakeman, P., & Mulhall, J. (1984). Maximal oxygen uptake utilising different treadmill protocols. *British Journal of Sports Medicine*, 18(2), 74-79.

Douglas, C. G. (1911). A method for determining the total respiratory exchange in man. *Journal of Physiology*, 42, 1-2.

Draper, S. B., Wood, D. M., & Fallowfield, J. L. (2003). The VO2 response to exhaustive square wave exercise: influence of exercise intensity and mode. *European Journal of Applied Physiology*, 90(1-2), 92-99.

Duncan, G. E., Howley, E. T., & Johnson, B. N. (1997). Applicability of VO_2max criteria: Discontinuous versus continuous protocols. *Medicine and Science in Sports and Exercise*, 29(2), 273-278.

Durnin, J. V. G. A., & Womersley, J. (1974). Body fat assessed from the total body density and its estimation from skinfold thickness: Measurements on 481 men and women aged from 16 to 72 years. *British Journal of Nutrition*, 32, 77-97.

Dyrstad, S. M., Soltvedt, R., & Hallén, J. (2006). Physical fitness and physical training during Norwegian military service. *Military Medicine*, 171(8), 736-741.

Dziados, J. E., Damokosh, A. I., Mello, R. P., Vogel, J. A., & Farmer, K. L. (1987). Physiological determinants of load bearing capacity. *US Army Research Institute of Environmental Medicine.* T19-87, 1-20.

Elia, M., McDonald, T., & Crisp, A. (1986). Errors in measurements of CO_2 with the use of drying agents. *Clinica Chimica Acta*, 158, 237-244.

Engebretson, J. E. (1998). Validity of breath-by-breath gas exchange analysis system. *Medicine and Science in Sports and Exercise* 30[Suppl 5], S330.

Fallowfield, J. L., & Wilkinson, D. M. (1999). *Improving Sports Performance in Middle and Long Distance Running.* Chichester: John Wiley & Sons Ltd.

Fallowfield, J. L., Hale, B. J., & Wilkinson, D. M. (2005). *Using Statistics in Sport and Exercise Science Research.* Chichester: Lotus Publishing.

Field, A. (2001). *Discovering Statistics: Using SPSS for Windows.* London: Sage Publications.

Field, A. (2005). 2nd edition. *Discovering Statistics: Using SPSS for Windows.* London: Sage Publications.

Franch, J., Madsen, K., Djurhuus, M. S., & Pedersen, P. K. (1998). Improved running economy following intensified training correlates with reduced ventilatory demands. *Medicine and Science in Sports and Exercise*, 30(8), 1250-1256.

Freedson, P., Kline, G., Porcari, J., Hintermeister, R., McCarron, R., Ross, J., Ward, A., Gurry, M., & Rippe, J. (1986). Criteria for defining VO_2max: A new approach to an old problem. *Medicine and Science in Sports and Exercise* 18[2 Suppl], S36.

Froelicher, V. F., Brammell, H., Davis, G., Nguera, I., Stewart, A., & Lancaster, M. C. (1974). A comparison of three maximal treadmill exercise protocols. *Journal of Applied Physiology*, 36(6), 720-725.

Frykman, P. N., & Harman, E A. (1995). Anthropometric correlates of maximal locomotion speed under heavy backpack loads. *Medicine and Science in Sports and Exercise* Suppl 27[5], S136.

Gaesser, G. A., & Poole, D. C. (1996). The slow component of oxygen uptake kinetics. *Exercise Sports Science Reviews* 24, 35-71..

Gordon, N. F., Van Rensburg, J. P., Moolman, J., Kruger, P. E., Russell, H. M. S., Grobler, H. C., & Cilliers, J. F. (1986). The South African Defence Force physical training programme. Part I. Effect of 1 year's military training on endurance fitness. *South African Medical Journal*, 69(8), 477-482.

Gordon, N. F., Moolman, J., Van Rensburg, J. P., Russell, H. M. S., Kruger, P. E., Grobler, H. C., & Cilliers, J. F. (1986). South African Defence Force physical training programme. Part II. Effect of 1 year's military training on muscular strength, power, power-endurance, speed and flexibility. *South African Medical Journal*, 69(8), 483-490.

Gore, C. J., Clark, R. J., Shipp, N. J., Van der Ploeg, G. E., & Withers, R. T. (2003). CPX/D underestimates VO_2 in athletes compared with an automated Douglas bag system. *Medicine and Science in Sports and Exercise*, 35(8), 1341-1347.

Gormley, S. E., Swain, D. P., High, R., Spina, R. J., Dowling, E. A., Kotipalli, U. S., & Gandrakota, R. (2008). Effect of intensity of aerobic training on VO_{2max}. *Medicine and Science in Sports and Exercise*, 40(7), 1336-1343.

Graham, S. M., Simpson, R. J., Clement, R., and Florida-James, G. D. (2008). Physiological and anthropometrical indicators of backpack running performance in 'Elite' British soldiers. *Medicine and Science in Sports and Exercise*, S199.

Grayson, D. (2004). Some myths and legends in quantitative psychology. *Understanding Statistics*, 3(1), 101-134.

Green, H. J., Jones, S., Ball-Burnett, M. E., Smith, D., Livesey, J., & Farrance, B. W. (1991). Early muscular and metabolic adaptations to prolonged exercise training in humans. *Journal of Applied Physiology*, 70(5), 2032-2038.

Green, H. J., Helyar, R., Ball-Burnett, M. E., Kowalchuck, N., Symon, S., & Farrance, B. (1992). Metabolic adaptations to training precede changes in mitochondrial capacity. *Journal of Applied Physiology*, 72(2), 484-491.

Griggs, K. E., Bilzon, J. L. J., Wilkinson, D. M., Richmond, V. L., and Rayson, M. P. (2008). Direct versus indirectprediction of prolonged loaded-march performance. *Medicine and Science in Sports and Exercise*.

Gutekunst, D. J., Frykman, P. N., Nindl, B. C., Rarick, K. R., Mello, R. P., and Harman, E. A. (2006). Training-induced improvements in submaximal gait economy with and without a backpack/military load. *Medicine and Science in Sports and Exercise.*

Haidar, S. G., Kumar, D., Bassi, R. S., & Deshmukh, S. C. (2004). Average versus maximum grip strength: Which is more consistent? *Journal of Hand Surgery*, 29(1), 82-84.

Haisman, M. F. (1988). Determinants of load carrying ability. *Applied Ergonomics*, 19(2), 111-121.

Hardman, A. E., Williams, C., & Wooton, S. A. (1986). The influence of short-term endurance training on maximum oxygen uptake, sub maximum endurance and the ability to perform brief, maximal exercise. *Journal of Sports Sciences*, 4, 109-116.

Harman, E. A., & Frykman, P. N. (1995). Heavy load carriage performance correlates: Backpack Vs Individual towed trailer. *Medicine and Science in Sports and Exercise,* 27[5], S136.

Harman, E. A., Frykman, P. N., Lammi, E. R., & Palmer, C. J. (1996). Effects of physically demanding training program on women's heavy work task performance. *Medicine and Science in Sports and Exercise*, 28, S128.

Harman, E. A., Frykman, P. N., Palmer, C., Lammi, E., Reynolds, K., and Backus, V. (1998). Effects of specifically designed physical condtioning program on the load carriage and lifting performance of female soldiers. *US Army Research Institute of Environmental Medicine,* T98-1, 1-84.

Harwood, G. E., Rayson, M. P., & Nevill, A. M. (1999). Fitness, performance, and risk of injury in British Army Officer cadets. *Military Medicine*, 164(6), 428-434.

Hattori, Y., Ono, Y., Shimaoka, M., Hiruta, S., Kamijima, M., & Takeuchi, Y. (1998). Test-retest reliability of isometric and isoinertial testing in symmetric and asymmetric lifting. *Ergonomics*, 41(7), 1050-1059.

Headquarters Department of the Army. (1990). Foot Marches. *United States Army.*21-18.

Heinert, L. D., Serfass, R. C., & Stull, G. A. (1988). Effect of stride length variation on oxygen uptake during level running and positive grade treadmill running. *Research Quarterly for Exercise and Sport*, 59(2), 127-130.

Helgerud, J., Hoydal, K., Wang, E., Karlsen, T., Berg, P., Bjerkaas,M., Simonsen, T., Helgesen, C., Hjorth, N., Bach, R., & Hoff, J. (2007). Aerobic high-intensity intervals improve VO_2max more than moderate training. *Medicine and Science in Sports and Exercise*, 39(4), 665-671.

Helgerud, J. (2007). Prescribing, quantifying, and monotoring exercise intensity during interval training. *Medicine and Science in Sports and Exercise*, 39(10), 1886.

Henriksson, J. (1977). Training induced adaptation of skeletal muscle and metabolism during sub maximal exercise. *Journal of Physiology*, 270, 661-675.

Hermansen, L. & Saltin, B. (1969). Oxygen uptake during maximal treadmill and bicycle exercise. *Journal of Applied Physiology*, 26(1), 31-37.

Hickson, R. C., Bomze, H. A., & Holloszy, J. O. (1977). Linear increase in aerobic power induced by a strenuous program of endurance exercise. *Journal of Applied Physiology*, 42(3), 372-376.

Hiilloskorpi, H., Manttari., A, Fogelholm, M., Pasanen, M., & Laukkanen, R. (2000). The comparison between three different respiratory gas-analysers. *Medicine and Science in Sports and Exercise*, 31[S5], S354.

Hill, A. V. & Lupton, H. (1923). Muscular exercise, lactic acid, and the supply and utilization of oxygen. *Quarterly Journal of Medicine*, 16, 135-171.

Hodges, L. D., Brodie, D. A., & Bromley, P. D. (2005). Validity and reliability of selected commercially available metabolic analyzer systems. *Scandinavian Journal of Medicine and Science in Sports*, 15, 271-279.

Holloszy, J. O., & Coyle, E. F. (1984). Adaptations of skeletal muscle to endurance exercise and their metabolic consequences. *Journal of Applied Physiology*, 56(4), 831-838.

Hopkins, W.G. (2000a). Measures of reliability in sports medicine and science. *Sports Medicine*, 30(1), 1-15.

Houston, M.E. (2001). *Biochemistry Primer for Exercise Science*. Champaign, IL: Human Kinetics.

Howley, E. T., Bassett, D. R., & Welch, H. G. (1995). Criteria for maximal oxygen uptake: Review and commentary. *Medicine and Science in Sports and Exercise*, 27(9), 1292-1301.

Howley, E. T. (2001). Type of activity: Resistance, aerobic and leisure versus occupational physical activity. *Medicine and Science in Sports and Exercise*, 33(6 Suppl), S364-S369.

Hunter, G. R., Bamman, M. M., Larson-Meyer, D. E., Joanisse, D. R., McCarthy, J. P., Blaudeau, T. E., & Newcomer, B. R. (2005). Inverse relationship between exercise economy and oxidative capacity in muscle. *European Journal of Applied Physiology and Occupational Physiology*, 94, 558-568.

Hurley, B. F., Hagberg, J. M., Allen, W. K., Seals, D. R., Young, J. C., Cuddihee, R. W., & Holloszy, J. O. (1984). Effect of training on blood lactate levels during submaximal exercise. *Journal of Applied Physiology*, 56(5), 1260-1264.

Ingjer, F. (1979). Effects of endurance training on muscle fibre ATP-ase activity, capillary supply and mitochondrial content in man. *Journal of Physiology*, 294, 419-432.

Ingham, S. A., Whyte, G. P., Pedlar, C., Bailey, D. M., Dunman, N., & Nevill, A. M. (2008). Determinants of 800-m and 1500-m running performance using allometric models. *Medicine and Science in Sports and Exercise*, 40(2), 345-350.

International Society for the Advancement of Kinanthropometry (2001). *International Standards for Anthropometric Assessment*. International Society for the Advancement of Kinanthropometry.

Issekutz, B., Birkhead, N. C., & Rodahl, K. (1962). The use of respiratory quotients in assessment of aerobic work capacity. *Journal of Applied Physiology*, 17, 47-50.

Jensen, K., Jorgensen, S., & Johansen, L. (2002). A metabolic cart for measurement of oxygen uptake during human exercise using inspiratory flow rate. *European Journal of Applied Physiology*, 87, 202-206.

Jette, M., Sidney, K., & Kimick, A. (1989). The effects of basic training on Canadian forces recruits. *Canadian Journal of Sports Science*, 14(3), 164-172.

Jeukendrup, A. E., & Wallis, G. A. (2005). Measurements of substrate oxidation during exercise by means of gas exchange measurements. *International Journal of Sports Medicine*, 26(Suppl 1), S28-S37.

Johnston, R. E., Quinn, T. J., Kertzer, R., & Vroman, N. B. (1997). Strength training in female distance runners: Impact on running economy. *Journal of Strength and Conditioning Research*, 11(4), 224-229.

Jones, D. A., Rutherford, O. M., & Parker, D. F. (1989). Physiological changes in skeletal muscle as a result of strength training. *Quarterly Journal of Experimental Physiology*, 74, 233-256.

Jones,D.A. & Round,J.M. (1990). *Skeletal Muscle in Health and Disease.* Manchester: Manchester University Press.

Jones, A. M. (1998). A five year physiological case study of an Olympic runner. *British Journal of Sports Medicine*, 32, 39-43.

Jones, A. M., Carter, H., & Doust, J. H. (1999). Effect of six weeks of endurance training on parameters of aerobic fitness. *Medicine and Science in Sports and Exercise* 31[5], S280.

Jones, A. M., & Carter, H. (2000). The effect of endurance training on parameters of aerobic fitness. *Sports Medicine*, 29(6), 373-386.

Joyner, M. J. (1991). Modelling: Optimal marathon performance on the basis of physiological factors. *Journal of Applied Physiology*, 70(2), 683-687.

Kang, J., Chaloupka, E. C., & Mastrangelo, M. A. (1999). Effect of testing protocol on maximal oxygen uptake in trained and untrained individuals. *Medicine and Science in Sports and Exercise,* 31[5 Suppl], S303.

Katch, V. L., Stanley, S. S., & Freedson, P. (1982). Biological variability in maximum aerobic power. *Medicine and Science in Sports and Exercise*, 14(1), 21-25.

Kilpatrick, W. G. (1989). Selection of soldiers. Physical build as an entry factor. *Journal Royal Army Medical Corps*, 135(2), 95.

Knapik, J. J., Vogel, J. A., & Wright, J. E. (1981). Measurement of isometric strength in an upright pull at 38 cm. *US Army Research Institute of Environmental Medicine,* T-3/81, 1-26.

Knapik, J. J., Staab, J., Bahrke, M., O'Connor, J., Sharp, M., Frykman, P. N., Mello, R., Reynolds, K., & Vogel, J. A. (1990). Relationship of soldier load carriage to physiological factors, military experience and mood states. *United States Army Research Institute of Environmental Medicine,* T17-90, 1-30.

Knapik, J. J., & Gerber, J. (1996). The influence of physical fitness training on the manual material-handling capability and road-marching performance of female soldiers. *ARL-TR-*1064, 1-97.

Knapik, J. J., Reynolds, K., & Harman, E. A. (2004). Soldier load carriage: Historical, physiological, biomechanical and medical aspects. *Military Medicine*, 169(1), 45-56.

Knapik, J. J., Darakjy, S., Hauret, K., Canada, S., Scott, S., Rieger, W., Marin, R., & Jones, B. H. (2006a). Increasing the physical fitness of low-fit recruits before basic combat training: An evaluation of fitness, injuries, and training outcomes. *Military Medicine*, 171(1), 45-54.

Knapik,J.J., Scott,S.J., Sharp,M.A., Hauret,K.G., Darakjy,S., Rieger,W.R., Palkoska,F.A., VanCamp,S.E., & Jones,B.H. (2006b). The basis for prescribed ability group run speeds and distances in U.S. Army basic combat training. *Military Medicine*, 171(7), 669-677.

Kraemer, W. J., Vogel, J. A., Patton, J. F., Dziados, J. E., & Reynolds, K. L. (1987). The effects of various physical training programs on short duration, high intensity load bearing performance and the Army Physical Fitness Test. *United States Army Research Institute of Environmental Medicine,* 30/87, 1-32.

Kraemer, W. J., Mazzetti, S. A., Nindl, B. C., Gotshalk, L. A., Volek, J. S., Bush, J. A., Marx, J. O., Dohi, K., Gómez, A. L., Miles, M., Fleck, S. J., Newton, R. U., & Häkkinen, K. (2001). Effect of resistance training on women's strength/power and occupational performances. *Medicine and Science in Sports and Exercise*, 33(6), 1011-1025.

Kraemer, W. J., Vescovi, J. D., Volek, J. S., Nindl, B. C., Newton, R.U., Patton, J. F., Dziados, J. E., French, D. N., & Hakkinen, K. (2004). Effects of concurrent resistance and aerobic training on load-bearing performance and the army physical fitness test. *Military Medicine*, 169(12), 994-999.

La Mere, V. J., Brown, K., Wigglesworth, J. K., & Edwards, J. E. (1993). Reproducibility between three metabolic systems and validation by Douglas Bag method. *Medicine and*

Science in Sports and Exercise 25[S5], S9.

Laursen, P.B., & Jenkins, D. G. (2002). The scientific basis for high-intensity interval training. Optimising training programmes and maximising performance in highly trained endurance athletes. *Sports Medicine*, 32(1), 53-73.

Leamon, S., & Scott, J. (2005). A reference guide for commanders: Military load carriage. *QinetiQ report*, 1-79.

Legg, S. J., & Duggan, A. (1996). The effects of basic training on aerobic fitness and muscular strength and endurance of British Army recruits. *Ergonomics*, 39(12), 1403-1418.

Legg, S. J., Perko, L., & Campbell, P. (1997). Subjective perceptual methods for comparing backpacks. *Ergonomics*, 40(8), 809-817.

Legg, S. J., & Cruz, C. O. (2004). Effect of single and double strap backpacks on lung function. *Ergonomics*, 47(3), 318-323.

Leveritt, M., Abernethy, P. J., Barry, B. K., & Logan, P. A. (1999). Concurrent strength and endurance training. *Sports Medicine*, 28(6), 413-427.

Lim, C. L., & Lee, L. K. H. (1994). The effects of 20 week basic military training program on the body composition, VO_2max and aerobic fitness of obese recruits. *Journal of Sports Medicine and Physical Fitness*, 34(3), 271-278.

Lyons, J. A., Allsopp, A. J., & Bilzon, J. L. J. (2003). Predicting the relative metabolic and cardiovascular demands of simulated military load-carriage tasks. *Contemporary Ergonomics*.

Lyons, J. A., Allsopp, A. J., & Bilzon, J. L. J. (2005). Influences of body composition upon the relative metabolic and cardiovascular demands of load-carriage. *Occupational Medicine*.

Macfarlane, D. J. (2001). Automated metabolic gas analysis system: A review. *Sports Medicine*, 31(12), 841-861.

MacRae, H. S. H., Dennis, S. C., Bosch, A. N., & Noakes, T. D. (1992). Effects of training on lactate production and removal during progressive exercise in humans. *Journal of Applied Physiology*, 72(5), 1649-1656.

Marieb, E. N. (1992). *Human Anatomy and Physiology*. Redwood: The Benjamin/Cummings Publishing Company.

Maughan, R. J., Watson, J. S., & Weir, J. (1984). Muscle strength and cross-sectional area in man: A comparison of strength-trained and untrained subjects. *British Journal of Sports Medicine*, 18(3), 149-157.

Mayhew, L. J., & Gross, P. M. (1951). Comparison of grade-incremented versus speed-incremented maximal exercise tests in trained men. *British Journal of Sports Medicine*, 9(4), 191-195.

Maxwell, S. E,. & Delaney, H. D. (1990). *Designing Experiments and Analysing Data*. California: Wadsworth Inc.

McCardle, W. D., Katch, F. I., & Katch, V. L.(2001). *Exercise Physiology*. Lippincott, Williams & Wilkins, Philadelphia, PA.

McCafferty, W. B., & Horvath, S. M. (1977). Specificity of exercise and specificity of training: A subcellular review. *Research Quarterly for Exercise and Sport*, 48(2), 358-371.

McCaig, R. H., & Gooderson, C. Y. (1986). Ergonomics and physiological aspects of military operations in a cold wet climate. *Ergonomics*, 29(7), 849-857.

McNaughton, L. R., Thompson, D., Philips, G., Backx, K., & Crickmore, L. (2002). A comparison of the Lactate Pro, Accusport, Analox GM7 and Kodak Ektachem lactate analysers in normal, hot and humid conditions. *International Journal of Sports Medicine*, 23, 130-135.

Mello, R. Damokosh, A. I., Reynolds, K., Witt, C. E., & Vogel, J. A. (1988). The physiological determinants of load bearing performance at different march distances. *United States Army Research Institute of Environmental Medicine,* T15-88, 1-25.

Meterology conversion table.(2007). http://www.csgnetwork.com/meteorologyconvtbl.html.

Midgley, A. W., McNaughton, L. R., & Wilkinson, M. (2006). Is there an optimal training intensity for enhancing the maximal oxygen uptake of distance runners? Empirical research findings, current opinions, physiological rationale and practical recommendations. *Sports Medicine*, 36(2), 117-132.

Midgley, A. W., McNaughton, L. R., & Carroll, S. (2007a). Reproducibility of time at or near VO_2max during intermittent treadmill running. *International Journal of Sports Medicine*, 28, 40-47.

Midgley, A. W., McNaughton, L. R., Polman, R., & Marchant, D. (2007b). Criteria for determination of maximal oxygen uptake: A brief critique and recommendations for future research. *Sports Medicine*, 37(12), 1019-1028.

Midownik, S., Carlon, V. A., Ferri, E., Burda, B., & Melendez, J. A. (2000). System of automated gas-exchange analysis for the investigation of metabolic processes. *Journal of Applied Physiology*, 89, 373-378.

Miles, D. S., Cox, M. H., & Verde, T. J. (1994). Four commonly utilized metabolic systems fail to produce similar results during submaximal and maximal exercise. *Sports Medicine Training and Rehabilitation*, 5, 189-198.

Morgan, D. W., Martin, P. E., & Krahenbuhl, G. S. (1989). Factors affecting running economy. *Sports Medicine*, 7(5), 310-330.

Morgan, D. W., Martin, P. E., Krahenbuhl, G. S., & Baldini, F. D. (1991). Variability in running economy and mechanics among trained male runners. *Medicine and Science in Sports and Exercise*, 3(3), 378-383.

Mouneimneh, L. A., Lester, M. E., Sharp, M. A., Werling, W. C., Cohen, B. S., Abbass, P., Dingman, J., Sweeney, B. P., Harman, E., and Ruediger, T. M. (2007). Effects of two training programmes on military tasks and physical fitness in infantry soldiers. *Medicine and Science in Sports and Exercise.*

Muza, S. R., Latzka, W. A., Epstein, Y., & Pandolf, K. B. (1989). Load carriage induced alterations of pulmonary function. *International Journal of Industrial Ergonomics*, 3, 221-227.

Narici, M. V., Roi, G. S., Landoni, L., Minetti, A. E., & Cerretelli, P. (1989). Changes in force, cross-sectional area and neural activation during strength training and detraining of the human quadriceps. *European Journal of Applied Physiology*, 59, 310-319.

Nevill, A. (1997a). Why the analysis of performance variables recorded on the ratio scale will invariably benefit from a log transformation. *Journal of Sport Sciences*, 15, 457-458.

Nevill, A. M. & Atkinson, G. (1997b).Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. *British Journal of Sports Medicine*, 31, 314-318.

Niemelä, K., Palatsi, I., Linnaluoto, M., & Takkunen, J. (1980). Criteria for maximum oxygen uptake in progressive bicycle tests. *European Journal of Applied Physiology and Occupational Physiology*, 44, 51-59.

Nindl, B. C., Barnes, B. R., Alemany, J. A., Frykman, P. N., Shippee, R. L., & Friedl, K. E. (2007). Physiological consequences of U.S. Army Ranger training. *Medicine and Science in Sports and Exercise*, 39(8), 1380-1387.

Noakes, T. D. (1988). Implications of exercise testing for prediction of athletic performance: A contemporary perspective. *Medicine and Science in Sports and Exercise*, 20(4), 319-330.

Noakes, T. D., Myburgh, K. H., & Schall, R. (1990). Peak treadmill running velocity during the VO_2max test predicts running performance. *Journal of Sports Sciences*, 8, 35-45.

Noakes, T. D. (1997). Challenging beliefs: *Ex Africa semper aliquid novi. Medicine and Science in Sports and Exercise*, 29(5), 571-590.

Noakes, T. D. (1998). Maximal oxygen uptake: "Classical" versus "contemporary" viewpoints: A rebuttal. *Medicine and Science in Sports and Exercise*, 30(9), 1381-1398.

Ntoumanis, N. (2001). *A Step by Step Guide to SPSS for Sport and Exercise Studies*.

Pallant, J. (2007). *SPSS Survival Manual*. New York: McGraw-Hill.

Pandorf, C. E., Harman, E. A., Frykman, P. N., Patton, J. F., Mello, R. P., & Nindl, B. C. (2002). Correlates of load carriage and obstacle course performance among women. *Work*, 18(2), 179-189.

Patton, J. F., Daniels, W. L., & Vogel, J. A. (1980). Aerobic power and body fat of men and women during Army basic training. *Aviation, Space and Environmental Medicine*, 51(5), 492-496.

Pereira, M. A., Freedson, P. S., & Maliszewski, A. F. (1994). Intraindividual variation during inclined steady-state treadmill running. *Research Quarterly for Exercise and Sport*, 65(2), 184-188.

Pereira, M. A., & Freedson, P. S. (1997). Intraindividual variation of running economy in highly trained and moderately trained males. *International Journal of Sports Medicine*, 18(2), 118-124.

Powers, S. K., & Howley, E. T. (1997). *Exercise Physiology: Theory and Application to Fitness and Performance*. Boston: WCB McGraw-Hill.

Price, M. J., & Campbell, I. G. (1997). Determination of peak oxygen uptake during upper body exercise. *Ergonomics*, 40(4), 491-499.

Pyne, D. B., Boston,T ., & Martin, D. T. (2000). Evaluation of the Lactate Pro blood lactate analyser. *European Journal of Applied Physiology*, 82, 112-116.

Rayson, M. P., Davies, A., & Stroud, A. (1993). The physiological predictors of maximum load carriage capacity in trained females. *UK Sport: Partners in Performance* (pp. p.212). London: Sports Council.

Rayson, M. P., Davies, A., Bell, D. G., & Rhodes-James, E. S. (1995). Heart rate and oxygen uptake relationship: A comparison of loaded marching and running in women. *European Journal of Applied Physiology and Occupational Physiology*, 71, 405-408.

Rayson, M. P., Holliman, D., & Belyavin, A. (2000). Development of physical selection procedures for the British Army. Phase 2: Relationship between physical performance tests and criterion tasks. *Ergonomics*, 43(1), 73-105.

Rayson, M., Wilkinson, D., & Nevill, A. (2002a). Physical selection standards for single entry recruits: Development and validation study. *Optimal Performance Ltd.*, 1-40.

Rayson, M., Wilkinson, D., & Nevill, A. (2002b). Physical selection standards for junior entry recruits: Development and validation study. *Optimal Performance Ltd.*, 1-48.

Rayson, M. P., Wilkinson, D. M., & Blacker, S. (2002c). The physical demands of CMS(R): An ergonomic assessment. *Optimal Performance Ltd.*, 1-63.

Reynolds, K., Harman, E. A., Worsham, R. E., Sykes, M. B., Frykman, P. N., & Bakus, V. L. (2001). Injuries in women associated with a periodized strength training and running program. *Journal of Strength and Conditioning Research*, 15(2), 136-143.

Rietjens, G. J. W. M., Kuipers, H., Kester, A. D. M., & Keizer, H. A. (2001). Validation of a computerised metabolic measurement system (Oxycon-Pro) during low and high intensity exercise. *International Journal of Sports Medicine*, 22, 291-294.

Roberts, M. J., & Russo, R. (1999). *A Student's Guide to Analysis of Variance*. New York: Routledge.

Ross, W. D., Kerr, D. A., Carter, J. E. L., Ackland, T. R., & Back, T. M. (1994). Anthropometric techniques: Precision and accuracy. In. J.E.L.Carter & T.R.Ackland (Eds.), *Kinanthropometry in Aquatic Sports. A Study of World Class Athletes* (pp. 158-173). Champaign, IL: Human Kinetics.

Rowland, T. W., & Cunningham, L. N. (1992). Oxygen uptake plateau during maximal treadmill exercise in children. *Chest*, 101, 485-489.

Ruiz, L., Colley, J. R. T., & Hamilton, P. J. S. (1971). Measurement of triceps skinfold thickness. An investigation of sources of variation. *British Journal of Preventive and Social Medicine*, 25, 165-167.

Rutherford, O. M., & Jones, D. A. (1986). The role of learning and coordination in strength training. *European Journal of Applied Physiology and Occupational Physiology*, 55, 100-105.

Saunders, P. U., Pyne, D. B., Telford, R. D., & Hawley, J. A. (2004). Factors affecting running economy in trained distance runners. *Sports Medicine*, 34(7), 465-485.

Sheehan, J. M., Rowland, T. W., & Burke, E. J. (1987). A comparison of four treadmill protocols for determination of maximum oxygen uptake in 10- to 12- year-old boys. *International Journal of Sports Medicine*, 8(1), 31-34.

Shephard, R. J. (1984). Tests of maximal oxygen intake: A critical review. *Sports Medicine*, 1(2), 99-124.

Simpson, R. J., Gray, S. C., & Florida-James, G. D. (2006). Physiological variables and performance markers of serving soldiers from two "elite" units of the British Army. *Journal of Sports Sciences*, 24(1), 1-8.

Sjödin, B., Jacobs, I., & Svedenhag, J. (1982). Changes in onset of blood lactate accumulation (OBLA) and muscle enzymes after training at OBLA. *European Journal of Applied Physiology and Occupational Physiology*, 49, 45-57.

Skinner, J. S., Wilmore, K. M., Jaskolska, A., Jaskolski, A., Daw, E. A., Rice, T., Gagnon, J., Leon, A. S., Wilmore, .H., Rao, D. C., & Bouchard, C. (1999). Reproducibility of maximal exercise test data in the HERITAGE Family Study. *Medicine and Science in Sports and Exercise*, 31(11), 1623-1628.

Sloniger, M. A., Cureton, K. J., Prior, B. M., & Evans, E. M. (1997). Anaerobic capacity and muscle activation during horizontal and uphill running. *Journal of Applied Physiology*, 83(1), 262-269.

Soule, R. G., Pandolf, K. B., & Goldman, R. F. (1978). Energy expenditure of heavy load carriage. *Ergonomics*, 21(5), 373-381.

Spiering, B. A., Kraemer, W. J., Anderson, J. M., Armstrong, L. E., Nindl, B. C., Volek, J. S., & Maresh, C. M. (2008). Resistance exercise biology: Manipulation of resistance exercise programme variables determines the responses of cellular and molecular signalling pathways. *Sports Medicine*, 38(7), 527-540.

Spurrs, R. W., Murphy, A. J., & Watsford, M. L. (2003). The effect of plyometric training on distance running performance. *European Journal of Applied Physiology* 89, 1-7.

Stacy, R. J., Hungerford, R. L., & McMahon, B. T. (1982). The effect of basic training on aerobic capacity and body fat in New Zealand Army recruits. *New Zealand Medical Journal*, 95(722), 876-878.

Sutton, J. R. (1992). Limitations to maximal oxygen uptake. *Sports Medicine*, 13(2), 127-133.

Svedenhag, J., & Sjödin, B. (1985). Physiological characteristics of elite male runners in and off-season. *Canadian Journal of Applied Sport Science*, 10(3), 127-133.

Swain, D. P., & Leutholtz, B. C. (1997). Heart rate reserve is equivalent to %VO_2Reserve, not to %VO_2max. *Medicine and Science in Sports and Exercise*, 29(3), 410-414.

Swain, D. P., & Franklin, B. A. (2002). VO_2 reserve and the minimal intensity for improving cardiorespiratory fitness. *Medicine and Science in Sports and Exercise*, 34(1), 152-157.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics*. New York: HarperCollins College Publisher.

Tabata, I., Nishimura, K., Kouzaki, M., Hirai, Y., Ogita, F., Miyachi, M., & Yamamoto, K. (1996). Effects of moderate-intensity endurance and high-intensity intermittent training on anaerobic capacity and VO_{2max}. *Medicine and Science in Sports and Exercise*, 28(10), 1327-1330.

Tanaka, H. (1994). Effects of cross-training: Transfer of training effects on VO_2max between cycling, running and swimming. *Sports Medicine*, 18(5), 330-339.

Taylor, H. L., Buskirk, E., and Henschel, A. (1955). Maximum oxygen intake as an objective measure of cardio-respiratory performance. *Journal of Applied Physiology* 8, 73-80.

Trank, T. V., Ryman, D. H., Minagawa, R. Y., Trone, D. W., & Shaffer, R. A. (2001). Running mileage, movement mileage, and fitness in male U.S. Navy recruits. *Medicine and Science in Sports and Exercise*, 33(6), 1033-1038.

Thoden, J. S. (1991). Testing aerobic power. In. J.D.Macdougall, H.A.Wenger, & H.S.Green (Eds.), *Physiological Testing of the High-Performance Athlete* (pp. 107-173). Champaign, IL: Human Kinetics.

Thomas, J. R., & Nelson, J. K. (1996). *Research Methods in Physical Activity*. Champaign,IL: Human Kinetics.

Thorstensson, A., Larsson, L., & Tesch, P. (1977). Muscle strength and fibre composition in athletes and sedentary men. *Medicine and Science in Sports and Exercise*, 9, 26-30.

Unnithan, V. B., Wilson, J., Buchanan, D., Timmons, J. A., & Paton, J. Y. (1994). Validation of the Sensormedics (S2900Z) metabolic cart for pediatric testing. *Canadian Journal of Applied Physiology*, 19(4), 472-479.

Vanderburgh, P. M., & Mahar, M. T. (1995). Scaling of 2-mile run times to body weight and fat-free weight in college-age men. *Journal of Strength and Conditioning Research*, 9(2), 67-70.

Vanderburgh, P. M., & Flanagan, S. (2000). The backpack run test: A model for a fair and occupationally relevant military fitness test. *Military Medicine*, 165(5), 418-421.

Van Handel, P. J., Costill, D. L., & Getchell, L. H. (1976). Central circulatory adaptations to physical training. *Research Quarterly for Exercise and Sport*, 47(4), 815-823.

Visser, T., van Dijk, M. J., Collee, H., and van der Loo, H. (2005). Is intensity or duration the key factor in march training. *Conference proceedings.*

Vogel, J. A., Crowdy, J. P., Amor, A. F., & Worsley, D. E. (1978). Changes in aerobic fitness and body fat during army recruit training. *European Journal of Applied Physiology and Occupational Physiology*, 40(1), 37-43.

Wagner, P. D. (1996). Determinants of maximal oxygen transport and utilization. *Annual Review of Physiology*, 58, 21-50.

Wagner, P. D. (2000). New ideas on limitations to VO_{2max}. *Exercise and Sports Science Reviews*, 28(1), 10-14.

Wenger, H. A., & Bell, G. J. (1986). The interactions of intensity, frequency and duration of exercise training in altering cardiorespiratory fitness. *Sports Medicine*, 3, 346-356.

Weston, A. R., Mbambo, Z., & Myburgh, K. H. (2000). Running economy of African and caucasian distance runners. *Medicine and Science in Sports and Exercise*, 32(6), 1130-1134.

Whipp, B. J., Ward, S. A., & Lamarra, N. (1982). Parameters of ventilatory and gas exchange dynamics during exercise. *Journal of Applied Physiology*, 52, 1506-1513.

Wilkinson, D. M. (1999). Physiological assessment of the middle and long-distance runner. In. J. L. Fallowfield & D. M. Wilkinson (Eds.), *Improving Sports Performance in Middle and Long-Distance Running Performance* (pp. 99-121). Chichester, UK: John Wiley and Sons Ltd.

Wilkinson, D. M. (1999). Training for middle and long-distance running. In. J. L. Fallowfield & D. M. Wilkinson (Eds.), *Improving Sports Performance in Middle and Long-Distance Running Performance* (pp. 69-98). Chichester, UK: John Wiley and Sons Ltd.

Wilkinson, D. M., Rayson, M. P., & Bilzon, J. L. J. (2003). The physical demands of the combined Parachute Regiment training syllabus. *Optimal Performance Ltd.*, 1-52.

Wilkinson, D., Blacker, S., and Rayson, M. (2006).The validity of the Garmin eTrex Legend for monitoring troop movement in the field. Optimal Performance Limited report .

Williams, A. G., Rayson, M. P., & Jones, D. A. (1999). Effects of basic training on material handling ability and physical fitness of British Army recruits. *Ergonomics*, 42(8), 1114-1124.

Williams, A. G., Rayson, M. P., & Jones, D. A. (2002). Resistance training and the enhancement of the gains in material-handling ability and physical fitness of British Army recruits during basic training. *Ergonomics*, 45(4), 267-279.

Williams, A. G., Rayson, M. P., & Jones, D. A. (2004). Training diagnosis for a load carriage task. *Journal of Strength and Conditioning Research*, 18(1), 30-34.

Williams, A. G. (2005). Effects of basic training in the British Army on regular and reserve army personnel. *Journal of Strength and Conditioning Research*, 19(2), 254-259.

Williams, A. G., & Rayson, M. P. (2006). Can simple anthropometric and physical performance tests track training-induced changes in load-carriage ability? *Military Medicine*, 171, 1-7.

Williams, T. J., Krahenbuhl, G. S., & Morgan, D. W. (1991). Daily variation in running eceonomy of moderately trained male runners. *Medicine and Science in Sports and Exercise*, 23(8), 944-948.

Wilson, G. (2000). Limitations to the use of isometric testing in athletic assessment. In. C. J. Gore (Ed.), *Physiological Tests for Elite Athletes* (pp. 151-154). Champaign, IL: Human Kinetics.

Wilson, G. J., & Murphy, A. J. (1995). The efficacy of isokinetic, isometric and vertical jump tests in exercise science. *Australian Journal of Science Medicine and Sport*, 27, 62-66.

Wilson, G. J., & Murphy, A. J. (1996). Strength diagnosis: The use of test data to determine specific strength. *Journal of Sports Sciences*, 14, 167-173.

Wilson, G. J., & Murphy, A. J. (1996). The use of isometric tests of muscular function in athletic assessment. *Sports Medicine*, 22(1), 19-37.

Winter, E. M., Eston, R. G., & Lamb, K. L. (2001). Statistical analyses in the physiology of exercise and kinanthropometry. *Journal of Sports Sciences*, 19, 761-775.

Withers ,R., Gore, C., Gass, G., & Hahn, A. (2000). Determination of maximal oxygen consumption (VO_2max) or maximal aerobic power. In. C. J. Gore (Ed.), *Physiological Tests for Elite Athletes: Australian Sports Commission* (pp. 114-127). Champaign, IL: Human Kinetics.

Wood, D. M. (1999). Physiological demands of middle distance running. In. J. L. Fallowfield & D. M. Wilkinson (Eds.), *Improving Sports Performance in Middle and Long-Distance Running* (pp. 14-38). Chichester: John Wiley & Son Ltd.

Wood, D. M. (1999). Physiological demands of long distance running. In. J. L. Fallowfield & D. M. Wilkinson (Eds.), *Improving Sports Performance in Middle and Long Distance Running* (pp. 39-68). Chichester: John Wiley & Sons Ltd.

Wood, D. M. (1999). Assessment of Maximal Oxygen Uptake in Runners: New Concepts on an Old Theme. PhD Thesis. University of Southampton.

Woolford, S. M., & Gore, C. J. (2004). Interpreting skinfold sums. Use of absolute or relative typical error? *American Journal of Human Biology*, 16(1), 87-90.

Yule, E., Kaminsky, L. A., Sedlock, D. A., King, B. A., and Whaley, M. H. (1996). Inter-laboratory reliability of VO_2max and submaximal measurements. *Medicine and Science in Sports and Exercise,* 28[S5], S15.

Zuntz, N. (1901). Ueber die bedeutung der verschiedenen nährstoffe als erzeuger der muskelkraft. *Pflügers Arch. Physiol.*, 83, 557.

APPENDIX A

Table A.1. The repeatability of body stature, and circumferential girth measures. Systematic bias and (95% LoA) are expressed in ratio and absolute terms. Ninety-five percent Confidence Intervals (CI) are presented for absolute 95% LoA.

	Ratio LoA	Absolute LoA	Absolute 95% CI (95% LoA)	
	Bias (×/÷ 95% LoA)	Bias (± 95% LoA)	Upper	Lower
Stature (cm)	1.00**(×/÷1.00)	0.2** (± 0.5)	**1.0** – 0.5	-0.1 – **-0.5**
Arm Girth (cm)	1.00 (×/÷ 1.02)	0.0 (± 0.8)	**1.1** – 0.5	-0.5 – **-1.1**
Calf Girth (cm)	1.00 (×/÷ 1.02)	0.0 (± 0.8)	**1.1** – 0.5	-0.5 – **-1.2**
Chest Girth (cm)	1.00 (×/÷ 1.02)	0.0 (± 2.4)	**3.4** – 1.4	-1.4 – **-3.4**
Waist Girth (cm)	1.00 (×/÷ 1.03)	0.1 (± 2.1)	**3.1** – 1.3	-1.1 – **-2.9**
Hip Girth (cm)	1.00 (×/÷ 1.03)	-0.1 (± 2.7)	**3.8** – 1.5	-1.6 – **-4.0**
Thigh Girth (cm)	1.00 (×/÷ 1.03)	-0.1 (± 1.6)	**2.2** – 0.8	-1.0 – **-2.4**

* symbolises statistical significance ($P<0.05$), ** ($P<0.01$).

Values in bold (highlighted with arrows) indicate the outer limits of the 95% confidence intervals for absolute LoA.

Table A.2. The repeatability of skinfold measures. Systematic bias and (95% LoA) are expressed in ratio and absolute terms. Ninety-five percent Confidence Intervals (CI) are presented for absolute 95% LoA.

	Ratio LoA	Absolute LoA	Absolute 95% CI (95% LoA)	
	Bias (×/÷ 95% LoA)	**Bias (± 95% LoA)**	**Upper** ↓	**Lower** ↓
Biceps SF (mm)	1.00 (×/÷ 1.25)	0.0 (± 1.0)	**1.4** – 0.6	-0.6 – **-1.5**
Triceps SF (mm)	1.03* (×/÷ 1.12)	0.4** (± 1.2)	**2.0** – 1.1	-0.3 – **-1.2**
Subscapular SF (mm)	1.00 (×/÷ 1.10)	0.1 (± 0.9)	**1.3** – 0.6	-0.5 – **-1.2**
Iliac Crest SF (mm)	1.01 (×/÷ 1.18)	0.3 (± 2.4)	**3.6** – 1.7	-1.2 – **-3.0**
Supraspinale SF (mm)	1.01 (×/÷ 1.15)	0.1 (± 1.1)	**1.6** – 0.8	-0.6 – **-1.4**
Abdomen SF (mm)	0.98**(×/÷1.06)	-0.3** (± 0.8)	**0.8** – 0.2	-0.8 – **-1.4**
Thigh SF (mm)	0.99 (×/÷ 1.12)	-0.1 (± 1.3)	**1.7** – 0.7	-0.8 – **-1.8**
Calf SF (mm)	0.97 (×/÷ 1.15)	-0.3 (± 1.4)	**1.7** – 0.6	-1.1 – **-2.2**

* symbolises statistical significance ($P<0.05$), ** ($P<0.01$).

Values in bold (highlighted with arrows) indicate the outer limits of the 95% confidence intervals for absolute LoA.

Table A.3. The repeatability of sub-maximal loaded marching and maximal running ($\dot{V}O_2max$). Systematic bias and (95% LoA) are expressed in ratio and absolute terms. Ninety-five percent Confidence Intervals (CI) are presented for absolute systematic 95% LoA.

	Ratio LoA	Absolute LoA	95% CI (95% LoA)	
	Bias (×/÷ 95% LoA)	Bias (± 95% LoA)	Upper ↓	Lower ↓
LM Heart Rate ($b \cdot min^{-1}$)	0.98 (×/÷ 1.11)	-3 (± 13)	**17** – 4	-10 – **-23**
LM Bla ($mmol \cdot L^{-1}$)	0.97 (×/÷ 1.65)	-0.1 (± 0.7)	**0.9** – 0.3	-4.0 – **-1.0**
LM RER	1.01 (×/÷ 1.06)	0.01 (± 0.05)	**0.08** – 0.04	-0.02 – **-0.06**
LM SF ($steps \cdot min^{-1}$)	1.00 (×/÷ 1.02)	0.0 (± 1.5)†	**2.0** – 1.0	-1.0 – **-2.0**
RM Bla ($mmol \cdot L^{-1}$)	0.98 (×/÷ 1.44)	-0.1 (± 3.7)	**5.2** – 1.8	-2.1 – **-5.5**
RM RER	1.00 (×/÷ 1.04)	-0.01 (± 0.04)	**0.06** – 0.02	-0.03 – **-0.07**
RM v$\dot{V}O_2$max ($km \cdot h^{-1}$)	0.98* (×/÷ 1.06)	-0.2* (± 0.6)	**0.7** – 0.1	-0.5 – **-1.1**

* symbolises statistical significance (P<0.05), ** (P<0.01).

† symbolises differences non-Normally distributed.

Values in bold (highlighted with arrows) indicate the outer limits of the 95% confidence intervals for absolute LoA.