Running Title:
Load carriage assessment

Title:
Preloaded time-trial to assess load carriage performance
Abstract

Purpose:
The relevance and importance of load carriage in recreational and occupational tasks has stimulated a large body of research. Exercise protocols have been criticised for a lack of relevance to occupational activities; accordingly the aim of this study was to assess the reliability of a preloaded time-trial protocol for load carriage assessment.

Methods:
After full familiarisation, eight healthy males performed two trials separated by one week. Each trial comprised 60min walking at 6.5km·h\(^{-1}\) and 0\% gradient (LC), 15 min seated recovery followed by a 2.4km time-trial (LC\(_{TT}\)). All trials were performed wearing a 25 kg backpack.

Results:
Performance time was 16.71 ± 1.82 min and 16.37 ± 1.78 min for LC\(_{TT}\) one and two respectively; with a mean difference of -0.34 ± 0.89 min. Using log ratio limits of agreement the mean bias was 1.02 and random error component of the agreement ratio was 1.11. The intra-class correlation (ICC) was 0.85, coefficient of variation (CV) was 10.5\% and Cohen’s \(d\) was 0.35. The protocol demonstrated a very good level of reliability.

Conclusions:
We present a novel and reliable pre-loaded time-trial protocol that more closely reflects operational activities and can be used to quantify load carriage performance. This protocol provides greater ecologically validity with respect to physical demands of load carriage activities than those adopted previously and provides an excellent tool for the strength and conditioning practitioner to assess individual load carriage performance.
Key Words:
Load Carriage, Reliability, Ecological Validity, Limits of Agreement, Variability

A - INTRODUCTION

Load carriage is defined as locomotion while bearing a mass upon the torso supported by shoulder straps and/or a hip belt and is a constituent element in some occupational settings such as the military or emergency services and also physical activities such as hiking (14). Carrying a backpack remains one of the most convenient and economical ways of transporting an external load (2) particularly in some military contexts where land vehicles may be restricted (15).

Studies have previously investigated the physiological and performance consequences of load carriage however; protocols typically lack ecological validity for occupational load carriage activities. For example lab-based studies have utilised single trials (normally a time trial over a set distance ~2.4 to 20 km), for various durations (up to 30 mins) and with varying external loads (15 – 46 kg) (12,15). In a military context soldiers will often exercise at submaximal levels (e.g. whilst on patrol) and also at high-intensity levels (e.g. during engagement) (6). Although it is inherently difficult to devise a laboratory-based protocol which precisely reflects operational requirement as the load, intensity, duration, terrain and environment vary with each scenario; protocols should more closely reflect the physiological characteristics of activities encountered in military situations such as during training or pre-deployment assessments. Therefore the practical question examined here is whether a laboratory-based protocol that incorporates more realistic occupational demands, i.e., a combination of constant speed low intensity and self-paced high intensity exercise, can demonstrate sufficiently low between session variability and be of use to the strength and conditioning coach. Accordingly, this protocol will improve the relevance of future studies that investigate load carriage performance, and also providing practitioners with a useful tool to measure load carriage specific fitness, readiness for deployments and responses to relevant interventions. However to date a protocol of this nature has yet to be designed with their reliability yet to be determined.
Therefore the aim of this study was to determine the reliability of a preloaded treadmill time-trial protocol (and also selected physiological parameters) by combining a period of sub-maximal exercise (pre-loaded phase), a rest period and a performance phase (time trial). We hypothesised that the protocol detailed will demonstrate very good between trial reliability and be of great benefit to the strength and conditioning practitioner.
B - METHODS

Experimental Approach to the Problem

The importance of load carriage was outlined in an excellent review within this journal (15), discussing at length the implications of load carriage assessment within an occupational context. This review suggested several areas of further study including the improvement of methodologies used to quantify performance, noting that previously adopted models were confounded on methodical issues (see introduction). The novel approach outlined here aims to establish a method suitable of quantifying load carriage performance which could be used by the Strength and Conditioning coach for baseline testing of performance, determining performance improvements following specific interventions and assessment of the individuals load carriage capability prior to periods of operational activity or training.

Subjects

Following ethics approval from the host university 8 healthy, non-smoking males with experience of load carriage through regular recreational load carriage activities, provided written informed consent to participate in this study. The physical characteristics of the participants are shown in Table 1. All participants were engaged in recreational physical training (strength and endurance) which was constant in the two months leading up to the study. A series of additional control measures were adopted prior to testing: participants did not engage in any strenuous exercise on the day preceding and the day of an exercise trial. Each participant also completed a 24h nutrition log, which was replicated for all subsequent trials. All trials were completed at identical times in the day and participants abstained from alcohol and caffeine in the 24h prior to testing and arrived at the laboratory 2h post-prandial.

Procedures

Preliminary Trials

Participants were briefed individually on the experimental design; following this each completed three preliminary trials. The first preliminary trial consisted of a body composition assessment using dual
energy x-ray absorptiometry (Lunar iDXA, GE Healthcare, Hertfordshire, UK) followed by an incremental exercise test on a motorised treadmill (Desmo, Woodway, Germany) for the determination of \( \dot{V}O_2 \)max. Following a 5 min warm-up at 8 km·h\(^{-1}\) and 1% gradient, the gradient was subsequently increased to 4% and speed increased by 1km·h\(^{-1}\)·min\(^{-1}\) until the limit of volitional tolerance (11). Online breath by breath gas analysis (MetaLyser ll, Cortex Biophysik, Birmingham, UK) was used to determine \( \dot{V}O_2 \)max defined as the highest 30s \( \dot{V}O_2 \) recorded throughout the test.

During the second preliminary trial, participants were familiarised and fitted with the 25 kg backpack (Web Tex, Bedford, UK) and completed 20 min exercise at 0% gradient and 6.5km·h\(^{-1}\). Following 15 min of seated recovery participants then completed a self-paced 2.4 km time-trial. The mass of the load was evenly distributed within the backpack and worn in accordance with manufacturer’s guidelines. The backpack incorporated shoulder straps and a waist strap which were adjusted individually and recorded to the nearest mm for subsequent trials. The third and final preliminary trial provided a full familiarisation of the experimental trial detailed below.

**Experimental Trial**

Participants performed the experimental trial on two occasions separated by a minimum of seven days. Participants walked for 60 min, 0% gradient and 6.5km·h\(^{-1}\) carrying a 25 kg backpack (hereon referred to as LC) (5,13). Following 15 min of seated recovery participants then completed a 2.4 km time-trial whilst bearing the load (LC\(_{TT}\)) where the speed of the treadmill was manually adjusted by the individual to complete the distance in the quickest time possible (13). The elapsed time was masked from the participant during all trials. The walking speed and duration, time-trial distance and absolute mass were selected to more closely reflect realistic occupational requirements in line with previous recommendations for laboratory based studies (18,19). Throughout LC physiological parameters were measured at 15min intervals, immediately prior to, after 1.2 and 2.4km of LC\(_{TT}\). Heart rate was measured using short-range telemetry (Polar T31, Kempele, Finland), expired pulmonary gases were assessed using Douglas bags (Cranlea and Co, Birmingham, UK), and blood lactate concentration ([lac]\(_B\), Accu-Check,
Safe T-Pro, Birmingham, UK) was measured from arterialised-venous fingertip blood samples. Core body temperature was recorded using a tympanic thermometer (Braun, IRT 4520, Nottingham UK). Ratings of whole body perceived exertion (RPE) were measured using the Borg scale (8). Perceptions of effort were further separated for leg (RPE<sub>legs</sub>) and breathing (RPE<sub>breathing</sub>) discomfort using a visual analogue scale: where 0 = no exertion and 10 = maximal exertion (24). In addition, participants performed a number of respiratory muscle strength and pulmonary function tests. Maximal inspiratory (P<sub>Imax</sub>) and expiratory pressures (P<sub>Emax</sub>) were assessed using a hand-held mouth pressure meter (MicroRPM, Micro Medical, Kent, UK) to provide an index of inspiratory and expiratory muscle strength (13). Pulmonary function was also assessed using a pneumotachograph (MS03, Micro Medical, Buckinghamshire, UK). All manoeuvres were performed pre and post-LC and post-LC<sub>TT</sub> (13).

**Statistical Analyses**

Firstly mean differences between trials were calculated and paired t-tests were used to determine systematic bias between the experimental trials (a-priori α = 0.05) using SPSS for Windows (SPSS, Chicago, IL, USA). Secondly, the intra-class correlation coefficient (ICC), coefficient of variation (CV) and Cohen’s d were calculated as a general indicator of reliability between experimental trials. Following this absolute limits of agreement were calculated using methods detailed previously (7), however, due to the presence of heteroscedasticity in the data (a positive relationship between the absolute measurement differences and their mean) and in line with previous recommendations (1), 95% log ratio limits of agreement (LoA) were calculated for all variables using methods detailed previously (7) providing an average dimensionless measurement bias (e.g. general learning effect) and random measurement error (e.g. level of agreement) using Microsoft Excel (Microsoft Corporation, Redmond, WA, USA). In addition, standard error and 95% confidence intervals for the measurement bias and random error components of the limits of agreement were calculated. Finally the 95% ratio limits of agreement were used to estimate sample sizes required for 5%, 10%, 20%, and 30% effects for a repeated-measures methodological design. Sample size calculations were performed using Microsoft Excel according to published equations (25).
C - RESULTS

Time trial performance

Time trial performance was 16.71 ± 1.82 min in trial 1 and 16.37 ± 1.78 min in trial 2 (P>0.05) with a mean difference of 0.34 ± 0.89 min (Figure 1). Velocity was fixed during the pre-loaded phase at 1.8m·s⁻¹ but during LC strongly increased to 2.42 ± 0.24 m·s⁻¹ in trial one and 2.46 ± 0.23 m·s⁻¹ in trial 2 with no difference between trials in the magnitude of this increase (P>0.05).

Physiological responses

Physiological responses to both trials are shown in Table 3. All transient changes in physiological parameters both post-LC and post LC strongly increased significantly to 2.42 ± 0.24 m·s⁻¹ in trial one and 2.46 ± 0.23 m·s⁻¹ in trial 2 with no difference between trials in the magnitude of this increase (P>0.05). Relative to baseline HR increased post LC strongly increased by 84 ± 28 beats·min⁻¹ and 90 ± 24 beats·min⁻¹, respectively (P<0.05). VO₂ increased from 1.49 ± 0.18 L·min⁻¹ and 1.37 ± 0.38 L·min⁻¹ pre-LC by 0.54 ± 0.35 L·min⁻¹ and 0.70 ± 0.42 L·min⁻¹ post-LC and a further 0.27 ± 0.48 L·min⁻¹ and 0.28 ± 0.43 L·min⁻¹ post LC strongly increased (P<0.05). VCO₂ increased from 1.27 ± 0.17 to 1.81 ± 0.35 L·min⁻¹ (absolute increase: 0.54 ± 0.35 L·min⁻¹) in trial 1 and from 1.28 ± 0.23 to 1.98 ± 0.31 L·min⁻¹ (absolute increase: 0.70 ± 0.42 L·min⁻¹) in trial 2 post-LC (P>0.05). Post LC strongly increased a further 0.27 ± 0.48 L·min⁻¹ and 0.57 ± 0.55 L·min⁻¹ post LC strongly increased (P>0.05). Baseline VE was 35.65 ± 4.36 L·min⁻¹ in trial 1 and 35.80 ± 6.09 L·min⁻¹ in trial 2. Post-LC, VE increased to 49.54 ± 7.42 L·min⁻¹ and 52.94 ± 7.77 L·min⁻¹ in trials 1 and 2, respectively (P<0.05) and post LC strongly increased to 60.52 ± 8.93 L·min⁻¹ and 64.08 ± 18.17 L·min⁻¹ in trials 1 and 2, respectively (P<0.05).
Respiratory function

Relative to baseline $P_{\text{Imax}}$ was reduced by $13 \pm 11 \text{ cmH}_2\text{O}$ (11%) and $18 \pm 11 \text{ cmH}_2\text{O}$ (15%) post-LC ($P<0.05$) and reduced $20 \pm 8 \text{ cmH}_2\text{O}$ (15%) and $21 \pm 9 \text{ cmH}_2\text{O}$ (17%) post LC$_{\text{TT}}$ (Baseline to post-LC$_{\text{TT}}$, $P<0.05$) in trials 1 and 2 respectively. Relative to baseline $P_{\text{Emax}}$ was reduced $14 \pm 11 \text{ cmH}_2\text{O}$ and $1 \pm 10 \text{ cmH}_2\text{O}$ in trials 1 and 2 respectively; with further reductions of $1 \pm 9 \text{ cmH}_2\text{O}$ (1%) and $3 \pm 11 \text{ cmH}_2\text{O}$ (3%) in trials 1 and 2 respectively post LC$_{\text{TT}}$ ($P<0.05$). Baseline pulmonary was within normal limits (17). Relative to baseline; FVC was reduced by $0.35 \pm 0.24 \text{ L}$ in trial 1 and $0.42 \pm 0.26 \text{ L}$ in trial 2 (8% and 12%, respectively) post-LC and $0.31 \pm 0.027 \text{ L}$ in trial 1 and $0.46 \pm 0.31 \text{ L}$ in trial 2 (10% and 11%, respectively) post LC$_{\text{TT}}$. FEV$_1$ was reduced by $0.35 \pm 0.24 \text{ L}$ in trial 1 and $0.31 \pm 0.20 \text{ L}$ in trial 2 (9% and 8%, respectively) post-LC; further similar differences were observed post LC$_{\text{TT}}$, trial 1: $0.31 \pm 0.27 \text{ L}$, trial 2: $0.46 \pm 0.31 \text{ L}$ (8% and 11%, respectively). Increases in FEV$_1$/FVC were observed relative to baseline during both trials and to a similar magnitude.

Reliability measurements

The test-retest and mean difference data for time-trial performance and physiological measurements for LC and LC$_{\text{TT}}$ are shown in Table 2 and absolute 95% limits of agreement for the LC$_{\text{TT}}$ experimental trials one and two are shown in Figure 1. Correlation coefficients of the absolute differences vs. their mean in some variables between trials demonstrated heteroscedasticity therefore log ratio limits of agreement were calculated for all data sets in addition to ICC, CV and Cohen’s $d$. Tables 3 and 4 show the ratio limits of agreement for time trial performance and respiratory function tests post-LC and post-LC$_{\text{TT}}$, respectively and Tables 5 and 6 shows the ratio limits of agreement for physiological and perceptual responses post-LC and post-LC$_{\text{TT}}$, respectively. The estimated sample sizes are displayed in Table 7. Overall, for an alpha level of 0.05 and 10% effect, mean sample sizes for all variables were 20 and 22 for LC and LC$_{\text{TT}}$ respectively.
D - DISCUSSION

Main Findings

We are the first to present a reliable load carriage protocol which more closely reflects the intensity and physiological demands of some occupational activities (18). We also demonstrate that a range of physiological and perceptual responses can be assessed between trials, with acceptable between session variability (1). We calculated ratio LoA between each trial and the ICC, CV and Cohen’s $d$, for which combination of analyses provides empirical evidence supporting the inclusion of this performance assessment protocol within future studies and within future practice (1).

Performance measures

Time trial performance was similar to previous study using an identical protocol (13) and also similar to British Infantry recruits carrying the same mass in a backpack over the same distance however, without a 60 min prior-constant intensity bout (9,10). In this study mean difference between efforts was $0.34 \pm 0.89$ min demonstrating excellent agreement (measurement bias: 1.02; random error $\times/\div$ 1.11). This is supported by narrow confidence intervals as depicted in Table 3 and Figure 1. The ICC (0.85) and CV (10.5%) values demonstrated good reliability and Cohen’s $d$ (0.36) suggests that differences between trial variance was moderate and inconsequential demonstrating that this protocol is suitable for the Strength and Conditioning coach to utilise in order to assess load carriage performance.

Physiological responses

Between session variability of physiological and perceptual parameters post-LC and $LC_{TT}$ are presented in Tables 5 and 6, illustrating good agreement and small error ratios. The variation between trials is greater during $LC_{TT}$ than LC, most likely explained by the nature of the time-trial which was self-paced and hence sensitive to changes in the adopted pace. However, the very good agreement should give confidence to the strength and conditioning practitioner that the physiological parameters presented
herein can be reliably performed between trials and used to quantify the physiological responses to steady state and time trial load carriage activity. Data represented here show similar temporal changes as existing literature investigating the physiological responses to treadmill marching at 6.5 5km/h⁻¹ with 25 kg followed by a 2.4 km time trial (4,13) which are also key determinants of load carriage performance (9). Accordingly the protocol presented presents a useful mode of exercise which reflects the physiological demands of some operational tasks.

Respiratory muscle and pulmonary function

Baseline and changes in $P_{\text{Imax}}$ and $P_{\text{Emax}}$ were similar ($P >0.05$) between trials demonstrating excellent agreement (see Tables 3 and 4). There was low bias post-LC ($P_{\text{Imax}}$ 1.08, $P_{\text{Emax}}$ 0.91) and post LC$_{\text{TT}}$ ($P_{\text{Imax}}$ 1.09, $P_{\text{Emax}}$ 0.91) and small random error ratios post-LC ($x/\div P_{\text{Imax}}$ 1.11, $P_{\text{Emax}}$ 1.24) and post-LC$_{\text{TT}}$ ($x/\div P_{\text{Imax}}$ 1.30, $P_{\text{Emax}}$ 1.37) (see Figures 4 and 5). The agreement ratio’s for $P_{\text{Emax}}$ are larger than anticipated, however the data in Tables 3 and 4 demonstrate narrow 95% confidence intervals for both the upper and lower bound limits. It is likely that greater variability was witnessed in the post-LC$_{\text{TT}}$ measures due to the individual differences in recovery rate post LC, adopted running speeds and time taken to complete the time trial since reductions in respiratory muscle pressures are inversely related to exercise intensity. However, excellent ICC values post-LC ($P_{\text{Imax}}$ 0.93, $P_{\text{Emax}}$ 0.86) and post-LC$_{\text{TT}}$ ($P_{\text{Imax}}$ 0.86, $P_{\text{Emax}}$ 0.81); where >0.8 demonstrates a very good level of reliability (1) qualify their use in this protocol. We observed reductions in respiratory muscle pressure generation which is illustrative of respiratory muscle fatigue which are similar to reductions demonstrated recently from our laboratory where (13). The implications of this reduction are far reaching in an occupational context as respiratory muscle fatigue may exacerbate limb muscle fatigue and impair performance through a sympathetically-mediated reflex reduction in limb blood flow (13).

Reductions in pulmonary function under load carriage conditions have been previously observed (16). Thoracic restriction modifies the normal breathing mechanics of the exercise hyperpnoea response and
consequently heightens the work of breathing (22). Measurements of pulmonary function were similar between trials at baseline, post LC and post LC_{TT} \((P > 0.05);\) See Table 2) and demonstrated excellent agreement with narrow confidence intervals and trial bias (See Tables 3 and 4). Interestingly however, agreement of PEF was poor, probably due to the effort dependence of this parameter during the initial high flow-low volume phase of the manoeuvre. Large variations in PEF occur with only slight changes in the inspiratory muscle recruitment pattern during inspiration by the participants prior to expiration (20). If inspiration from functional residual capacity (FRC) to total lung capacity (TLC) occurs rapidly and is subsequently followed by a forceful exhalation a greater PEF will result compared with a controlled inspiration and longer time spent at TLC due to the elastic recoil properties of the thorax (23). Neither timing or inspiration strategy were controlled here and similar to existing research we offer this as a potential explanation for the heightened variability in PEF post-exercise (20). On this basis studies should control inspiratory flow rate and the time spent at TLC during initial phase of the manoeuvre when conducting this measurement within a load carriage setting.

E – PRACTICAL APPLICATION

We present a novel protocol that can assess load carriage performance through a pre-loaded time-trial and a host of physiological markers that are relevant to the analysis of load carriage performance. This protocol can be adopted in settings where it is necessary to quantify load carriage performance; prior to and following a training intervention where load carriage performance is the dependant variable and prior to an operational deployment where load carriage is considered to be a critical role-related task. Within physical selection tests, armed forces organisations do not currently employ a loaded time trial, rather, they employ unloaded running (typically 2.4 – 3.2 km distance). However, this is known to be a very poor predictor of occupational (i.e., load carriage) performance (3). The omission of a loaded time trial is likely to mitigate associated injury although our findings suggests that including this test would allow the strength and conditioning practitioner to reliably capture changes in military specific fitness. It is impossible to devise a laboratory-based protocol which precisely reflects operational requirements since the load, intensity, duration; terrain and environment vary with each deployment or training scenario.
However, the protocol we present here is similar to current British Army assessments used in pre-deployment training and fitness assessments. For example, the 2.4 km loaded time trial has been used to determine the physical and physiological responses to acute changes in British Army infantry training programmes (9,10). In addition, we present a protocol that is similar to the Infantry Basic Combat Fitness Test where soldiers complete an eight mile course, carrying a load of 25 kg at 15 min·mile$^{-1}$ (6.4 km·h$^{-1}$) and the Advanced Combat Fitness Test 1 whereby following an 800 m warm up at 15 min·mile$^{-1}$, participants perform a 2.4 km time trial carrying a 20 kg load (21). Thus our protocol provides the strength and conditioning coach with a useful tool in assessing responses to training interventions and readiness for operations. In addition, as shown by the sample size calculations (Table 7), controlled trials can be performed effectively typically with a sample of 20 and 22 for the LC and LC$_{TT}$ trials respectively during a repeated measures design offering a large effect size (≥10%).
F - CONCLUSION

The results from this study suggest that the variability in parameters during LC and LCTT are negligible and the load carriage protocol presented provides a reliable measure for assessing the performance and physiological effects of bearing an external load upon the thorax; whilst also containing improved ecological validity on submaximal and high intensity load carriage activities compared with previous measures. To the authors' knowledge, this is the first study that has attempted to address the reliability of pre-loaded time-trial performance and has important implications for future research design and practice by strength and conditioning professionals.
REFERENCES


21. The British Army land-mediacomms-webcontent@land.mod.uk. The British Army - Homepage [Internet], 2014. Available from: http://www.army.mod.uk/


Acknowledgments
None

Current Contact information:
Mr Mark Faghy, School of Science, University of Derby, Kedleston Road, Derby, DE22 1GB, Tel: +44 (0)1332 592109, Fax: +44 (0) 1332 597747, Email: M.Faghy@Derby.ac.uk

Conflict of interest: None

Grants: None

Funding: None

The results of the present study do not constitute endorsement of the product by the authors or the NSCA.
Figure 1, Bland-Altman plot for $\text{LC}_{\text{TT}}$ data achieved during the two experimental trials.