



OPEN ACCESS

# Cognitive performance of military men and women during prolonged load carriage

Nicola C Armstrong <sup>1,2</sup>, S J R Smith,<sup>3</sup> D Risius,<sup>1</sup> D Doyle,<sup>1</sup> S L Wardle,<sup>4</sup>  
J P Greeves,<sup>4,5,6</sup> J R House,<sup>2</sup> M Tipton,<sup>2</sup> M Lomax<sup>2</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjilitary-2021-002000>).

<sup>1</sup>Human Sciences Group, Defence Science and Technology Laboratory, Salisbury, UK

<sup>2</sup>School of Sport, Health and Exercise Science, University of Portsmouth, Portsmouth, UK

<sup>3</sup>Human Sciences Group, Defence Science and Technology Laboratory, Fareham, UK

<sup>4</sup>Department of Army Health and Physical Performance Research, UK Ministry of Defence, Andover, UK

<sup>5</sup>Faculty of Medicine and Health Science, University of East Anglia Norwich Medical School, Norwich, UK

<sup>6</sup>Division of Surgery and Interventional Science, University College London, London, United Kingdom

## Correspondence to

Dr Nicola C Armstrong, Defence Science and Technology Laboratory, Salisbury, UK; [ncarmstrong@dstl.gov.uk](mailto:ncarmstrong@dstl.gov.uk)

Received 14 September 2021

Accepted 16 January 2022

Published Online First

7 April 2022

## ABSTRACT

**Background** This study evaluated cognitive workload in soldiers undertaking a long duration march wearing different loads.

**Methods** Military participants (n=12 men and n=10 women) performed four 3-hour loaded marches (12.25 km at 4.9 km/hour) wearing either 21 kg, 26 kg, 33 kg or 43 kg. During the march, accuracy and response time were measured using the verbal working memory n-back test (0, 1, 2 and 3) and two bespoke Go/No Go tests (visual/auditory) to assess inhibition of a pre-potent response.

**Results** The physical demands of the march increased with load and march duration but remained at moderate intensity. N-back test accuracy ranged from 74% to 98% in men and 62% to 98% in women. Reduced accuracy was observed as load and time increased. Accuracy during the visual Go/No Go test also reduced with load, accuracy ranged from 69% to 89% in men and 65% to 90% in women. No differences due to load or time were observed during completion of the auditory Go/No Go task; accuracy ranged from 93% to 97% in men and 77% to 95% in women. A number of participants were unable to complete the march due to discomfort. Reports of discomfort were more frequent in women, which may have contributed to the greater reductions in accuracy observed.

**Conclusion** These data provide further evidence that cognitive performance of military personnel can be affected during long duration loaded marching. Women reported discomfort from equipment more frequently than men, which may make them more susceptible to declines in cognitive performance. These findings highlight important considerations for equipment procurement.

## INTRODUCTION

The mass carried by soldiers during operational deployments has increased due to an increase in the volume of combat protective equipment, electronic devices (including batteries) and the amount of ammunition carried.<sup>1</sup> During operations in Afghanistan (2002–2014), high mean (57 kg) and peak (71 kg) patrol loads were recorded for UK soldiers.<sup>1</sup>

The physiological, biomechanical and clinical implications of carrying load have been studied,<sup>2–5</sup> but the effect of load carriage on cognition has received less attention. Several critical military tasks involve a cognitive component, including maintaining vigilance through attentional processes, perceiving and processing verbal communications and using executive processes to respond to threats and stimuli in the environment. The evolving demands for the future warfighter means there is likely to be increased

## Key messages

- ⇒ This study provides further evidence that cognitive performance is negatively affected during prolonged load carriage; preliminary data indicate that women may be more affected than men.
- ⇒ Increased cognitive decrement in women was attributed to increased discomfort and work rate in women and greater load carriage experience of men.
- ⇒ These data highlight important considerations for equipment procurement and future research of this nature.

requirement for high-levels of cognitive performance in the field, for example, as the requirement to integrate warfighter-worn technology to support situation awareness continues. As such, it is essential to consider both the physical and cognitive aspects of military tasks when assessing soldier performance, to understand any resulting decrements and consider how we train personnel to optimise performance under dual and multi-tasking conditions.<sup>6</sup>

Early work exploring the effect of load on cognition measured cognitive task performance before and after a loaded march.<sup>7</sup> This approach was limited as it allowed for recovery from the physical demands of the task and therefore potentially underestimated the impact of load on cognition. More recent load carriage studies have examined the physical and cognitive demands concurrently. Under these dual-task conditions, there is evidence that increased cognitive demands are placed on a soldier as the mass carried rises. Reductions in accuracy and decreased sensitivity have been observed during a vigilance task in participants wearing 40 kg in comparison with no load<sup>8</sup>; impaired target detection and memory recall reported with 61 kg loads<sup>9</sup> and performance declines (demonstrated by reductions in accuracy and decreased sensitivity) during an auditory Go/No Go task when loads  $\geq 40$  kg were carried in both laboratory<sup>10</sup> and field settings.<sup>11</sup>

The cognitive demands placed on the wearer during load carriage<sup>8–12</sup> are compounded by complex movements, such as navigating obstacles,<sup>8</sup> suggesting that physical movement and cognitive processing compete for shared neural resources. This is supported by findings showing that dual-task conditions, for example, combining running while completing a word recall task, result in impaired performance of one of the tasks compared



© Crown copyright (2023), Dstl and MOD.

**To cite:** Armstrong NC, Smith SJR, Risius D, et al. *BMJ Mil Health* 2023;**169**:37–45.

with single-task performance.<sup>13</sup> The effects of combining tasks that are competing for resources require further investigation to understand the implications for soldier task performance.

When considering the influence of physical exertion on cognition,<sup>14–16</sup> it is evident that the characteristics of the task and of participants influence performance outcomes, which may make certain individuals more susceptible to declines in cognitive performance. This can be explained in part by the arousal–performance relationship which is based on work by Yerkes and Dodson<sup>17</sup> and proposes an inverted-U relationship between arousal and performance. During familiar tasks, performance is greatest at moderate levels of arousal, but is reduced at the extremes. This relationship may shift left or right depending on the complexity of the task, demands of the environment or the characteristics of the individual.

Exercise is a stimulus that increases arousal<sup>18</sup> and there is evidence of declines in cognitive performance as exercise intensity increases.<sup>19</sup> As women experience greater physiological strain than men during load carriage,<sup>20</sup> this may move them towards the extremes of the arousal–performance relationship and result in greater decline in the performance of cognitive tasks. However, this may only become clear after performing the task for some time; previous evidence shows that decrements resulting from the effects of differing loads may only become apparent after 30 minutes of physical activity.<sup>11</sup> It may be that the differences between men and women are only observed with greater loads and after a longer period of sustained physical performance.

Women, who are now eligible to join ground close combat roles, will carry the same load as men in combat. There are no published studies examining the effect of load on cognition in men and women during a prolonged loaded march. Load carriage research must consider the performance of both men and women, to fully understand the implications of such tasks for both sexes.

The primary objective of this work was to compare the cognitive performance of military men and women during prolonged loaded marching with representative military loads. The experimental hypotheses were: (1) accuracy on cognitive tests would decrease with time during a 3-hour loaded march; (2) accuracy on cognitive tests would decrease with increased loads; and (3) accuracy on cognitive tests would decrease to the same extent in men and women.

## METHOD

### Participants

This study was part of a larger body of work exploring the response of Infantry soldiers to load carriage. At the time this study was undertaken, there were no women in Infantry roles. Rather than exclude women from this study, participants were recruited from the Royal Artillery as the physical demands associated with these roles were considered the closest match to the Infantry. Twelve men from the Infantry and ten women from the Royal Artillery volunteered to participate and provided written informed consent.

Study inclusion criteria included: (1) unit sign-off as fit to undertake loaded marching; (2) passing a study entrance medical examination and (3) lifting a minimum of 30 kg to a height of 1.45 m using the Army Power Bag single lift test. At the time of the study, this test was one of the physical selection standards for the British Army. In this context, it was used to assess participant suitability for load carriage tasks. Prior to each test session, participants were reviewed by the study medical officer to ensure they remained fit to take part.

### Study design

Participants performed four 3-hour loaded marches on a slatted belt treadmill (Woodway, model PPS70, Germany) over a 2-week period; each loaded march differed in the mass carried. Participants conducted one march per day with at least one rest day between marches; no physical activity was undertaken on rest days. During each loaded march, a cognitive test battery was performed to evaluate the effect of march duration and mass carried on overall accuracy. Response time was also measured during the cognitive tests to better understand accuracy data.

### Participant characteristics

Prior to commencing the load carriage element of the study, stature (Seca, model 225, Germany), body mass (Ian Fellows, model Lucid, UK) and body composition (from a whole body scan; dual-energy X-ray absorptiometry, GE Healthcare, model Lunar iDXA, USA) were assessed. An incremental exercise test was performed to determine maximal rate of oxygen uptake ( $\text{VO}_{2\text{max}}$ ) and gas exchange threshold (GET) (Metamax 3B (stationary mode), Cortex, Germany). Work rate (Metamax 3B (stationary mode), Cortex, Germany) was measured during each loaded march and compared to the GET to quantify the physical demands placed on the participants.

### Loaded march

The loaded march was developed in collaboration with subject matter experts, operational analysts and with reference to unpublished data collected by the Defence Science and Technology Laboratory (Dstl) during military training exercises/operations. Relevant published literature was also considered. The exercise profile represented a 3-hour march (4.9 km/hour; 0%) with a 10-minute water stop every 50 minutes. The total distance marched on each occasion was 12.25 km. Marches were performed in a fed state, and participants were allowed water *ad libitum* during the rest periods. Rating of perceived exertion<sup>21</sup> was measured at three points every hour. Oxygen uptake ( $\text{VO}_2$ ) was measured continuously during each loaded march. All testing was conducted at 19.8°C (0.7°C) and 48.9% (3.3%) relative humidity. At the end of each march, participants were interviewed by an investigator to capture subjective feedback including reports of pain, soreness and discomfort.

### Load configurations

Four load configurations were investigated (Table 1). These configurations were selected in collaboration with Dstl military advisors and designed to represent the mass typically carried during military tasks. While in-service equipment was used where possible, it was not the purpose of this study to provide an equipment evaluation. The order in which the load configurations were worn was counterbalanced using a Latin Square. Participants were issued with body armour that was self-reported as the most comfortable with the greatest range of movement. Investigators confirmed that the appropriate landmarks on the torso were covered.<sup>22</sup>

### Cognitive tests

Three cognitive tests were conducted each hour during the loaded march. These tests included visual and auditory Go/No Go tests (two per hour) to measure response inhibition and n-back working memory tests (once per hour). Unpublished work conducted by Dstl has developed a framework for research studies which maps military tasks to cognitive tests.<sup>23</sup> These cognitive tests were selected as

**Table 1** Load configurations

| Configuration   | Equipment   | Mass (kg)       |
|---|---|-----------------|
| A—personal armour (LOW)<br>(not a military dress state, but required to understand the impact of wearing protective equipment without additional load carriage) | Helmet (1.46 kg)<br>Enhanced Protection Under Body Armour Combat Shirt<br>Personal Clothing System trousers<br>Boots*<br>SA80 (dummy rifle matched for dimensions and mass 5.13 kg)<br>Commercially available body armour vest with front and rear plates (~10.9 kg available in 5 sizes XS to XXL) | 20.94 (0.56) kg |
| B—assault order (MED)<br>(includes equipment worn to fight through a position)  | Configuration A<br>Personal Load Carrying Equipment (PLCE) webbing pouches (×5) evenly distributed around the waist and weighted with 5 kg equipment required to achieve target mass  | 26.12 (0.49) kg |
| C—patrol order (HIGH)<br>(additional equipment carried to support longevity in the field)   | Configuration B<br>Daysack with chest belt fastened, weighted with 7 kg equipment required to achieve target mass   | 33.28 (0.59) kg |
| D—marching order (V-HIGH)<br>(associated with extended duration, overnight)   | Configuration A<br>PLCE short back bergen with waist belt fastened, weighted with 17 kg equipment required to achieve target mass<br>The body armour was carried in the bergen  | 43.37 (0.64) kg |

Mass represents dressed mass minus undressed mass and is provided as mean (SD). Variability in mass relates to differences in the size of the equipment worn.  
\*Participants were asked to wear boots that they would normally wear during loaded marching.

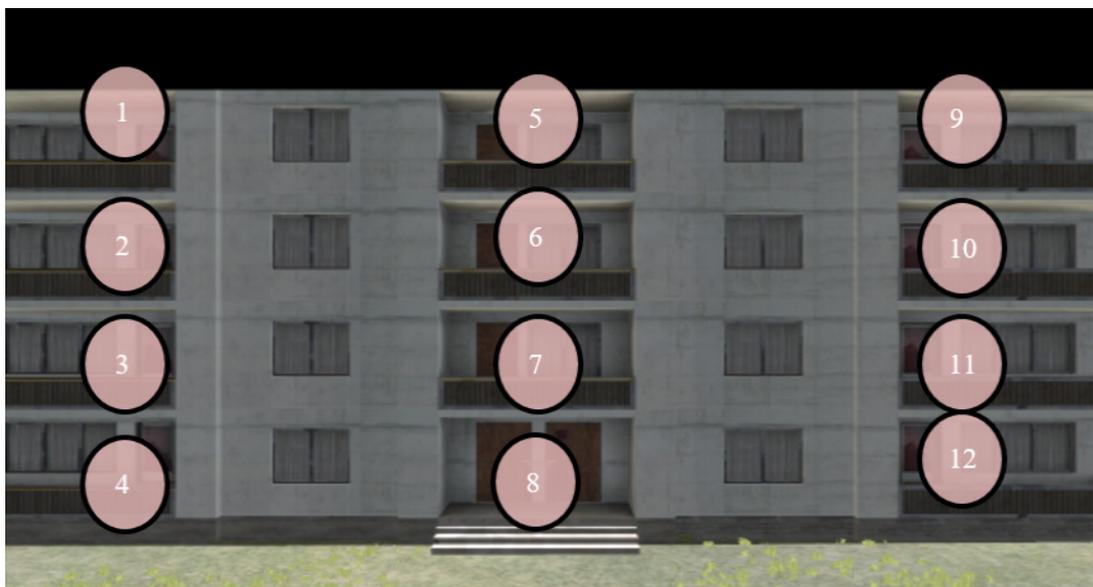
they represent important cognitive functions that are critical to the dismounted soldier's operational performance. These include attention, working memory, that is, the maintenance and manipulation of information, and response inhibition which allows a participant to withhold a prepotent response to a target.<sup>23</sup> Other published research has demonstrated that these tasks are reliable and valid<sup>24 25</sup> and have been implemented successfully in other load carriage studies.<sup>8 10 11</sup>

The visual cognitive tests were displayed on a 297×214 cm projector screen, 4.5 m from the treadmill and 118 cm from the ground to approximately eye level. The visual Go/No Go test was presented using the Virtual Battlespace 2 program. Familiarisation and training were undertaken on all cognitive tests to ensure participants understood the test, were responding correctly and had an opportunity to practise responding while walking on the treadmill.

### Go/No Go tests

Two Go/No Go response inhibition tests (one visual, one auditory) were administered at minutes 5 and 20 of each hour during the march. In total, each test was completed six times during the 3 hours and each element (visual/auditory) lasted approximately 3 minutes.

In both tests, participants were presented with 60 targets; 48 (80%) targets were enemy and 12 (20%) targets were friendly. Participants were asked to respond using a hand-held device whenever an enemy target was presented ('Go' stimuli) and not respond whenever a friendly target was presented ('No Go' stimuli). This test design results in a prevalent 'Go' response to the enemy target so that participants must slow their response to correctly inhibit their response when the friendly 'No Go' stimuli was presented. The order of friendly and enemy targets was randomised and the interstimulus interval was 1000–2000 ms. Stimuli were presented for 500 ms and participant responses were

**Figure 1** Visual Go/No Go balcony scene and target locations.



**Figure 2** Friendly target (left), enemy target (centre), enemy target appearing at position 12 (right).

recorded for 1000 ms from the start of stimulus onset. Response accuracy (%) and time to respond correctly (response time: ms) are reported.

The visual test comprised a static image of an apartment block with 12 possible target locations (Figure 1). Friendly targets were depicted as uniformed soldiers, enemy targets were depicted as non-uniformed individuals (Figure 2). The stimuli for the auditory test consisted of the sound of a single shot from a Dragonov sniper rifle as the enemy target, and the sound of a single shot from an L85 A2 as the friendly target. Sounds were controlled for volume.

### N-back verbal working memory test

The n-back test can be conducted at a number of different levels, usually 0, 1, 2 or 3, with the 'n' referring to the level of difficulty. In the easiest verbal version of the test, the 0-back, a series of letters are presented on the screen, participants have to respond to each letter by deciding whether it is the same or different to a target letter presented at the beginning of the test. At the next level, 1-back, instead of comparing the letter on the screen with a target letter, it is compared with the letter presented on the previous screen (*i.e.* one screen back). In the 2-back test, the letter is compared with the letter on the screen two screens back, and the 3-back involves seeing if the current letter matches the letter three screens back. The demands of the 'n-back' gradually increase such that the easiest versions of the test are placing cognitive demands in terms of attention and vigilance for targets, whereas the harder versions require continuous updating of working memory.

In the current study, the test required participants to respond to each letter by pressing a green button on a hand-held device whenever the letter presented matched the letter presented 'n' letters before. For letters that did not match, participants had to respond by pressing a red button on the same hand-held device. Each test consisted of 45 stimulus presentations (letters) and lasted approximately 3 minutes. The stimuli were presented centrally on a screen for 500 ms followed by a blank screen for 2500 ms. Of the stimuli presented, 15 (33%) were targets and 30 (67%) were non-targets.<sup>25</sup>

The n-back tests were administered at minutes 30–45 of each hour (three times in each load configuration) and lasted for approximately 12 minutes. Mean response time (ms) and mean accuracy (%) to targets and non-targets were recorded.

### Data analysis

Statistical analysis was performed using IBM SPSS Statistics V.26. Parametric statistics (analysis of variance (ANOVA), t-test) were applied to data that were normally distributed (Shapiro-Wilk  $>0.05$ ). Non-parametric analyses (Friedman, Wilcoxon, Mann-Whitney U) were performed on data that were not normally distributed. Data were accepted as significant at the level of  $p < 0.05$ . Due to a number of incomplete cognitive tests, data for men and women were pooled and differences between the sexes explored using t-tests or descriptive statistics where  $n < 6$ .

### Parametric data

A two-way repeated measures ANOVA was performed to investigate the interaction between load ( $\times 4$ ) and time ( $\times 3$  for n-backs;  $\times 6$  for Go/No Go) and the F-value reported. The Greenhouse-Geisser correction was applied where the assumption of sphericity was not met. When a significant main effect was observed, the Bonferroni *post-hoc* test was used to identify differences between the loads; for the effect of time, data from the first test undertaken were compared with all subsequent tests. Differences between men and women were explored using an unpaired t-test ( $n \geq 6$ ).

### Non-parametric data

When data were not normally distributed, Friedman's one-way ANOVA was used to investigate the effects of load and time without exploring the interaction effects. Where this test was used,  $\chi^2$  is reported. This test was followed by Wilcoxon's test when a main effect was detected and the U-value reported. The Mann-Whitney U test was used to compare men and women ( $n \geq 6$ ).

### Effect sizes

Partial eta squared ( $\eta_p^2$ ) was reported for the main ANOVA effects to inform future power calculations. For pairwise comparisons, Cohen's d ( $d$ ) was calculated for parametric data and interpreted as 0.2 (small), 0.6 (moderate), 1.2 (large), 2.0 (very large) and 4.0 (extremely large).<sup>26</sup> Pearson's r ( $r$ ) was calculated for non-parametric data and interpreted as 0.1 (small), 0.3 (medium) and 0.5 (large).

**Table 2** Participant characteristics

|                                | Men (n=12) |        | Women (n=10) |        |
|--------------------------------|------------|--------|--------------|--------|
|                                | Mean       | (SD)   | Mean         | (SD)   |
| Age (years)                    | 22.6       | (3.6)  | 22.8         | (3.7)  |
| Stature (m)                    | 1.80       | (0.1)  | 1.70         | (0.1)* |
| Body mass (kg)                 | 77.9       | (8.4)  | 71.4         | (9.1)  |
| Lean body mass (kg)            | 59.7       | (5.4)† | 47.2         | (7.2)* |
| Body fat (kg)                  | 15.9       | (4.3)† | 21.2         | (1.5)* |
| VO <sub>2max</sub> (mL/kg/min) | 55.2       | (5.4)  | 44.1         | (4.3)* |
| Smokers (n/total n)            | 4/12       |        | 5/10         |        |

Data are mean (SD).  
\*Indicates significant difference from men ( $p < 0.05$ ).  
†=11 men.  
VO<sub>2max</sub>, maximal rate of oxygen uptake.

## RESULTS

### Participant characteristics

Participant characteristics are presented in [Table 2](#). On average, men were taller and leaner, with a greater VO<sub>2max</sub>, although there was overlap between the groups for all parameters ([Table 2](#)). Women reported their hormonal contraceptive use as follows: combined contraceptive pill, n=4; contraceptive implant, n=2; intrauterine contraceptive device, n=1; no hormonal contraception, n=3.

### Loaded march

Only six men and one woman were able to complete the full exercise test under all conditions ([Table 3](#)). Test withdrawals were related to self-reported discomfort caused by the load carried (shoulder discomfort, blisters and chaffing around the hips and groin) or muscular discomfort. Reports of discomfort increased with load in both sexes, but the frequency of reporting was greatest in women (see online supplemental table 1). No withdrawals were self-reported as being related to work rate.

Participants reported discomfort when they were required to lift their head to look up at the projector screen in some load configurations. This discomfort caused some participants to withdraw from the exercise test altogether, while others withdrew from the cognitive tests but continued marching.

### Physical demands

VO<sub>2</sub> increased with load ( $F_{(2, 12)} = 49.316, p \leq 0.0001, \eta_p^2 = 0.851$ ) and time ( $F_{(8, 48)} = 12.193, p \leq 0.0001, \eta_p^2 = 0.680$ ), but there was no interaction between load and time ( $F_{(16, 96)} = 0.845, p > 0.05, \eta_p^2 = 0.134$ ). When VO<sub>2</sub> was expressed relative to lean body mass, differences between the sexes were observed for each load ( $t_{(16)} = -3.261, p = 0.005, d = 1.1-1.6$ ) indicating that physical work rate was greater in women. Men marched at 31%–41% of VO<sub>2max</sub> and women 36%–55% of VO<sub>2max</sub>. VO<sub>2</sub> did not exceed GET in either group during the march, regardless of the load

carried or duration of the march indicating that all participants worked at a moderate exercise intensity.

### N-back working memory test

#### Accuracy

Accuracy data are presented in [Figure 3](#). Visual inspection indicates that some participants were more susceptible to reductions in accuracy in the heaviest loads than others ([Figure 3B,C](#)). These data points represent those participants who attempted the cognitive tests but were unable to continuously look at the projector screen as they adjusted their backpacks or sought relief from the discomfort of lifting their heads. These data points could be considered outliers, but the authors opted to include them in the data as they represent the demands placed on the participants rather than measurement errors.

There was no effect of load ( $\chi^2_{(3)} = 6.155, p > 0.05$ ) or time ( $\chi^2_{(2)} = 0.913, p > 0.05$ ) during the 0-backs. During the 1-backs, there was a main effect of time in participants wearing HIGH ( $\chi^2_{(2)} = 9.160, p = 0.010$ ). Accuracy decreased during hour 3 when compared with hour 1 ( $p = 0.003, r = -0.7$ ) and hour 2 ( $p = 0.045, r = -0.5$ ). There was also a main effect of load during hour 1 ( $\chi^2_{(3)} = 11.541, p = 0.009$ ). Accuracy was reduced in V-HIGH when compared with LOW ( $p = 0.003, r = -0.8$ ), MED ( $p = 0.004, r = -0.7$ ) and HIGH ( $p = 0.007, r = -0.7$ ) ([Figure 3A](#)).

During the 2-backs, there was no effect of time ( $\chi^2_{(2)} = 1.849, p > 0.05$ ), but there was a main effect of load during hour 1 ( $\chi^2_{(3)} = 12.459, p = 0.006$ ) and hour 3 ( $\chi^2_{(3)} = 13.791, p = 0.003$ ). Accuracy was reduced in V-HIGH when compared with HIGH during hour 1 ( $p = 0.028, r = -0.6$ ) and LOW ( $p = 0.018, r = -0.8$ ), MED ( $p = 0.028, r = -0.8$ ) and HIGH ( $p = 0.017, r = -0.8$ ) during hour 2 ([Figure 3B](#)).

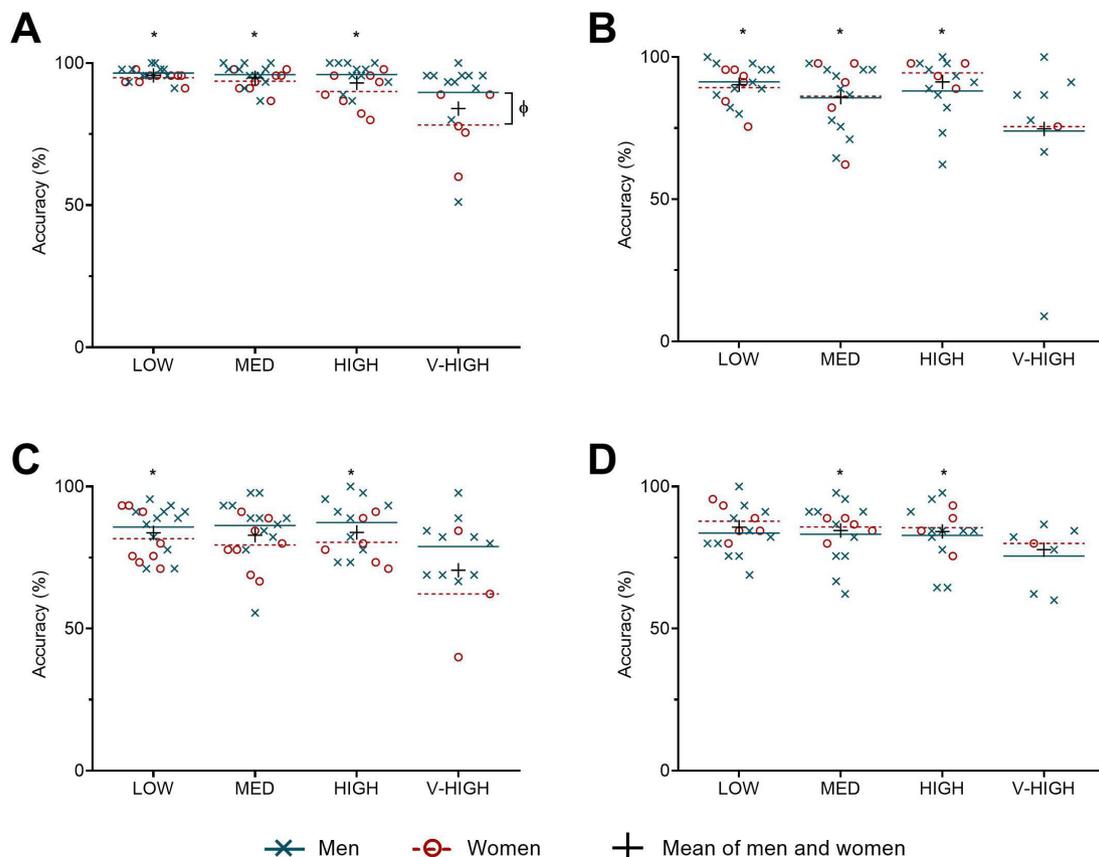
A similar pattern was observed during the 3-backs, with no effect of time ( $\chi^2_{(2)} = 2.769, p > 0.05$ ), but a main effect of load during hour 1 ( $\chi^2_{(3)} = 10.886, p = 0.012$ ) and hour 3 ( $\chi^2_{(3)} = 9.800, p = 0.020$ ). Accuracy was reduced in V-HIGH when compared with LOW ( $p = 0.018, r = -0.7$ ) and HIGH ( $p = 0.008, r = -0.8$ ) during hour 1 ([Figure 3C](#)) and MED ( $p = 0.028, r = -0.9$ ) and HIGH ( $p = 0.018, r = -0.9$ ) during hour 3 ([Figure 3D](#)).

Accuracy ranged from 74% to 98% in men and 62% and 98% in women (see online supplemental table 2). Accuracy was reduced in women during the 0-backs in HIGH (hour 2:  $U = 12.000, p = 0.015, r = -0.6$ ) and V-HIGH (hour 1:  $U = 13.500, p = 0.046, r = -0.5$ ). Descriptive statistics indicate that accuracy in women was reduced earlier in the march in the heavier loads. Women with a higher accuracy tended to march for longer; thus, the difference between men and women reduced as time progressed.

**Table 3** Number of participants who completed the march in each load configuration

| Configuration | Men          |              |             | Women      |             |            |
|---------------|--------------|--------------|-------------|------------|-------------|------------|
|               | Hour 1       | Hour 2       | Hour 3      | Hour 1     | Hour 2      | Hour 3     |
| LOW           | 12/12 (100%) | 12/12 (100%) | 11/12 (92%) | 9/10 (90%) | 8/10 (80%)  | 6/10 (60%) |
| MED           | 12/12 (100%) | 12/12 (100%) | 11/12 (92%) | 8/10 (80%) | 6/10 (100%) | 5/10 (50%) |
| HIGH          | 11/12 (92%)  | 11/12 (92%)  | 11/12 (92%) | 7/10 (70%) | 6/10 (100%) | 4/10 (40%) |
| V-HIGH        | 10/11 (91%)  | 8/11 (73%)   | 7/11 (64%)  | 6/10 (60%) | 3/10 (30%)  | 2/10 (20%) |

Six men and one woman completed each hour in every configuration. One man was withdrawn from the study as he developed a respiratory tract infection prior to completing his final test in V-HIGH. Two women were able to complete the V-HIGH configuration in full, one of these women withdrew from the march when wearing HIGH.



**Figure 3** Accuracy measured during the n-backs. (A) 1-back, hour 1; (B) 2-back, hour 3; (C) 3-back, hour 1; (D) 3-back, hour 3. Dashed line represents mean data for women, complete line represents mean data for men. \* indicates a difference from V-HIGH;  $\Phi$  indicates a difference between the sexes ( $p < 0.05$ ).

**Response time**

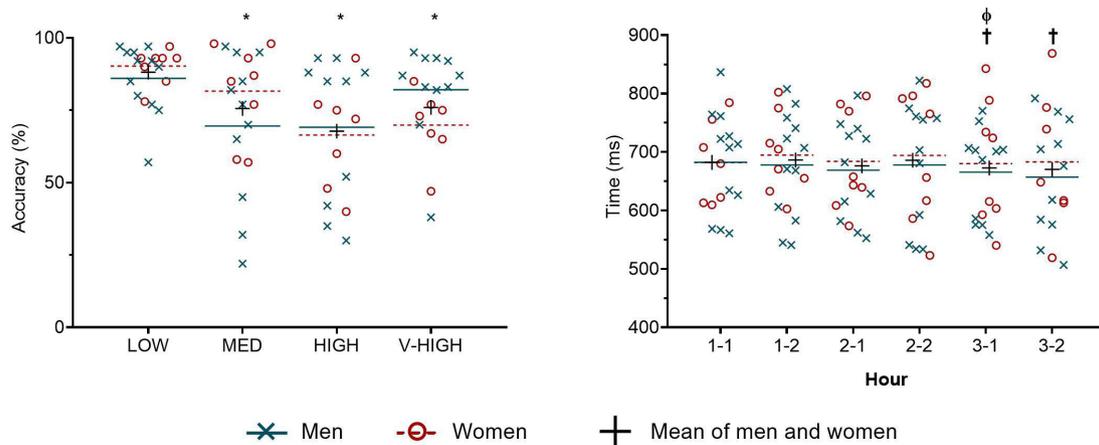
During the 0-back, there was a main effect of load ( $F_{(3, 12)} = 4.087, p = 0.033, \eta^2 = 0.505$ ), but not time ( $F_{(2, 8)} = 2.025, p \geq 0.05, \eta^2 = 0.336$ ) and there was no interaction between load and time ( $F_{(6, 24)} = 1.363, p \geq 0.05, \eta^2 = 0.254$ ). During hour 3, reaction time was reduced in V-HIGH when compared with MED ( $p = 0.010, d = 0.7$ ). No other differences in

reaction time were observed. Descriptive data indicated that women took longer to respond to a target than men.

**Visual Go/No Go test**

**Accuracy**

During hour 1, there was a main effect of load ( $\chi^2_{(3)} = 9.478, p = 0.024$ ). Accuracy was greatest in LOW when compared with



**Figure 4** Left: visual Go/No Go accuracy measured during hour 1. Right: visual Go/No Go response time in LOW. Dashed line represents mean data for women, complete line represents mean data for men. \* indicates a difference from LOW;  $\dagger$  indicates a difference from hour 1;  $\Phi$  indicates a difference between the sexes ( $p < 0.05$ ).

MED ( $p=0.010$ ,  $r=-0.5$ ), HIGH ( $p=0.002$ ,  $r=-0.6$ ) and V-HIGH ( $p=0.017$ ,  $r=-0.5$ ) (Figure 4). No other differences were observed between loads or over time.

Accuracy ranged from 89% to 69% in men and 90% to 65% in women (see online supplemental table 3). Descriptive data suggest that women tended to have a higher accuracy in lighter loads, but accuracy in men was higher in the heavier loads.

### Response time

There was a main effect of time when participants wore LOW ( $\chi^2_{(5)}=17.629$ ,  $p=0.003$ ). Response time was quicker in the final hour of the march when compared with the first hour ( $p=0.004$ ,  $r=-0.7$ ) (Figure 4). Men responded more quickly than women in the final hour of the march in LOW ( $p=0.028$ ,  $r=-0.5$ ) (Figure 4). No differences in response time were observed with load ( $p>0.05$ ).

### Auditory Go/No Go test

#### Accuracy

Accuracy data are provided in the online supplemental table 4. Due to a technical error, data were only analysed for eight men. No differences in accuracy were observed with load ( $F_{(3, 12)}=0.945$ ,  $p>0.05$ ,  $\eta_p^2=0.191$ ) or time ( $F_{(5, 20)}=2.135$ ,  $p>0.05$ ,  $\eta_p^2=0.348$ ) with no interaction between load and time ( $F_{(15, 60)}=1.209$ ,  $p>0.05$ ,  $\eta_p^2=0.232$ ). Accuracy ranged from 93% to 97% in men and 77% to 95% in women. Descriptive data indicated that accuracy was reduced in the heavier loads in women.

#### Response time

Overall response time was unaffected by load ( $F_{(3, 12)}=2.132$ ,  $p>0.05$ ,  $\eta_p^2=0.348$ ) and time ( $F_{(5, 20)}=1.040$ ,  $p>0.05$ ,  $\eta_p^2=0.206$ ) and there was no interaction between load and time ( $F_{(15, 60)}=0.927$ ,  $p>0.05$ ,  $\eta_p^2=0.1688$ ). No differences in response time were observed between the sexes.

## DISCUSSION

This study aimed to compare cognitive performance in service men and women conducting a long duration loaded march in doctrinal military loads. In spite of stringent medical and physical selection criteria, only six men and one woman were able to complete all conditions, with discomfort being a common reason for withdrawal. Due to incomplete data sets, it was not possible to fully test the study hypotheses; however, this work provides evidence of the impact of discomfort on military relevant physical and cognitive task performance. This work also highlights important issues that have implications for the procurement of soldier equipment as well as methodological considerations for future research.

Notwithstanding the high number of test withdrawals, decreased accuracy was observed during the n-back tests and visual Go/No Go tests with both load and time. This supports the findings of other studies which have reported reductions in cognitive performance during load carriage.<sup>8–12</sup> In addition, review of the descriptive statistics from the current study provides initial indication that reductions in accuracy occurred during less complex tests and with ‘lighter’ loads in women, although this observation should be confirmed with follow-on work.

Reports of discomfort increased in both sexes with load, but the frequency of these reports was greatest in women (see online supplemental table 1). Previous work has demonstrated that the configuration of equipment can impact on marching performance.<sup>7</sup> In the current study, participants attributed the

discomfort experienced to issues with equipment fit and integration. Despite five sizes of armour being available, 9 out of 10 women reported issues related to fit, compared with 0 out of 12 men. The mass and stature of the participants was in close agreement with data from the UK military anthropometric database. This highlights the subject that women should not be issued with smaller sizes of equipment that has been designed to fit men; equipment should be designed so that it is inclusive of different body shapes.

When considering body armour and load carriage systems, it is challenging to define ‘good fit’ beyond the subjective experience of the wearer. Stirling *et al*<sup>27</sup> have attempted to define fit based on three characteristics which include static, dynamic and cognitive fit. These characteristics consider how the equipment interacts with the wearer during static postures, dynamic movements and the impact of the equipment on cognitive capabilities, specifically somatosensation and executive function. In relation to cognitive fit, the authors consider how the equipment affects movement, posture and sensory feedback as well as cognitive processes. The current study supports the approach proposed by Stirling *et al*<sup>27</sup> as it highlights the impact that discomfort and poor fit can have on cognitive performance and the importance of assessing these characteristics during the equipment procurement process.

Anecdotal reports from investigators also indicate that more experienced soldiers were able to better mitigate against the discomfort caused by load. In the current study, the experience of the participants ranged from newly trained soldiers to those who had undertaken several operational tours as section commanders. As the women used in this study were not Infantry soldiers, a lack of load carriage experience may have contributed to some of the observed results.

VO<sub>2</sub> data were used to quantify work rate so that the results of the cognitive tests could be interpreted within the context of the physical demands. VO<sub>2</sub> was greater in women, which likely reflects the additional dead mass carried by this group (5 kg), where dead mass was the sum of load carried and fat mass. VO<sub>2</sub> remained below GET which indicates that the work rate remained moderate throughout the march regardless of the mass carried or sex. As moderate intensity exercise can be sustained for longer than 4 hours,<sup>28</sup> the work rate should not have influenced the participants’ ability to complete the march. However, perceived exertion was increased in women compared with men (see online supplemental tables 5 and 6) which may have reduced the available resource to devote to the cognitive tasks, and suggests that women found the marching task more challenging than men.

Our data suggest that increasing the mass carried led to increased conflict between the dual cognitive/physical task conditions, that is, there was a greater impact of the physical task on the behavioural ability to complete the cognitive task, and there are initial indications that this occurred to a greater extent in the women. These potential differences between the sexes may be related to a number of factors including discomfort caused by poor-fitting equipment, load carriage experience given that the women tested were not Infantry, and perceived exertion.

The preliminary findings suggest an interaction between the behavioural requirements of the visual task, specifically the requirement to look at the screen and the discomfort or physical demands of carrying load which affected participants’ posture. This observation illustrates the conflicts that may occur in a dual task involving both physical and cognitive demands and the effect it can have on performance. For example, Kobus *et al*<sup>9</sup> reported decreased accuracy during a visual choice reaction

time test after 85 min of marching at 3.2 km/hour in 61 kg. These authors presented eye-tracking data, which identified that visual focus was directed towards the lower part of the projector screen and threat detection was reduced towards the upper portion of the visual field as mass increased.

This study indicates that both perceived exertion and discomfort have the potential to add to the cognitive load and negatively affect the subjective experience of the soldier. There is a growing body of evidence exploring the relationship between physical exertion and cognition<sup>14–16, 29</sup>; however, few studies have considered discomfort.<sup>30</sup> In the current study, the participants experienced difficulties completing the cognitive tests as it was uncomfortable for them to lift their head to look up at the projector screen in some load configurations. This discomfort caused some participants to withdraw from the exercise test altogether, while others withdrew from the cognitive tests but continued marching. The implications of this behavioural response in an operational scenario may include reduced situational awareness, reduced vigilance and an increase in task errors. It is also interesting to note that reductions in accuracy were not observed during the auditory Go/No Go test which conflicts with the findings of others who have used auditory tests to assess cognitive performance during load carriage.<sup>10, 11</sup> In the current study, there was no requirement to look at the projector screen during the auditory tests which could have reduced the conflict between the posture needed to look at the screen and the limitations associated with the load configurations resulting in reduced demands placed on the participant during this test.

### Limitations

These findings should be considered preliminary given the small number of participants included in statistical analyses. Military participants were recruited to ensure the highest ecological validity, with the trade-off that they were only available for a limited time period and it was not possible to recruit participants until the desired power was achieved. This limitation also reduced the number of recovery days between tests and may have contributed to a number of test withdrawals due to the cumulative effects of blisters, chaffing and muscular discomfort. This schedule, however, could be considered representative of the demands placed on soldiers during operations. Future studies should consider fewer test configurations, longer recovery periods to mitigate against the risk of data loss resulting from test withdrawals or increased participant recruitment to allow for withdrawals if the research requirement necessitates testing on multiple successive days to better replicate demands in the field.

Women were recruited from the Royal Artillery as there were no women in the Infantry at the time of the study. Although the Royal Artillery have physical employment standards of similar demands to the Infantry, they do not routinely carry loads of this magnitude, which may have contributed to increased reports of discomfort and higher ratings of perceived exertion.

The test battery was designed to represent the demands of operational scenarios in terms of the loads carried and the cognitive tasks required by the soldier. Conducting this study in a more immersive simulated environment or in a field environment would further increase face validity. The test battery was developed for a controlled environment to reduce the impact of confounding factors such as varying weather and light levels on the findings. Practically, the participants who were trained soldiers, many with operational experience, may have suffered boredom from the monotony of treadmill-based exercise, which

prevented self-paced exercise. Future work should investigate how this test battery could be used in the field environment to evaluate, ecologically, the cognitive demands on the soldier, or a hybrid of simulated laboratory environment and field testing.<sup>6, 10, 11</sup>

### Summary

This study provides further evidence that reductions in cognitive performance will be observed during a long duration loaded march. Initial observations indicate that women may experience reductions in task accuracy earlier, in lighter loads and during less complex tasks than men, but this finding requires further confirmation. The differences observed between the sexes were attributed to increased discomfort in women, possibly caused by poor-fitting equipment and poor integration, greater load carriage experience of men and a greater perceived exertion experienced by women. These physical factors will have influenced the amount of reserve available for women to devote to the cognitive tasks. These data should be considered during future equipment procurement to ensure that equipment issues are not a limiting factor for the physical and cognitive performance of service personnel.

**Twitter** Nicola C Armstrong @nic\_c\_armstrong and J P Greeves @JulieGreeves

**Acknowledgements** The authors would like to thank the project management team Mrs Lorraine Beavis and Mr Jon Russell; the military advisors involved in the study design Lt Col Mike Potter and Lt Col Howard Long, and Dr Thomas O'Leary and Dr Charlotte Coombs for providing technical review. © Crown copyright (2022), Dstl. This material is licensed under the terms of the Open Government Licence except where otherwise stated. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gov.uk](mailto:psi@nationalarchives.gov.uk)

**Contributors** NCA, SLW, JPG, SJRS and JRH designed the study. Data collection was conducted by NCA, DR, DD and SLW. NA and DD analysed the data. NCA and SJRS prepared the manuscript. All authors reviewed the manuscript, MT, JRH and ML supervised the project (PhD supervisors). NCA is the guarantor for the study.

**Funding** This work was jointly funded by the UK Ministry of Defence, through the Dstl Land Integrated Survivability Programme, Dismounted Protection project (contract number STECH/0008) and the Women in Ground Close Combat Review Programme (Army Contract number: ARMYHQ2/00061).

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This study involves human participants and received favourable ethical opinion from the Ministry of Defence Research Ethics Committee (protocol 694/MODREC/15). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplemental information.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This material is licensed under the terms of the Open Government Licence except where otherwise stated. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gov.uk](mailto:psi@nationalarchives.gov.uk)

### ORCID iD

Nicola C Armstrong <http://orcid.org/0000-0001-8088-0279>

## REFERENCES

- 1 Lloyd-Williams RC, Fordy G. *The consequences of burden on the Dismounted soldier. (DSTL/CR68536/V1.0). Cited with permission from Dstl*, 2013.
- 2 Knapik JJ, Reynolds KL, Harman E. Soldier load carriage: historical, physiological, biomechanical, and medical aspects. *Mil Med* 2004;169:45–56.
- 3 Haisman MF. Determinants of load carrying ability. *Appl Ergon* 1988;19:111–21.
- 4 Drain J, Billing D, Neesham-Smith D, et al. Predicting physiological capacity of human load carriage - a review. *Appl Ergon* 2016;52:85–94.
- 5 Bossi LLM, Tack DW, Karakolis T. Rationalizing the approach to mitigate soldier physical burden: are Ironman or Captain America our Magic Bullet? In: Belanger SAH, Breede HC, von Hlatky S, eds. *Transhumanising war: performance implications for policy, society, and the soldier*. Montreal & Kingston: McGill-Queen's University Press, 2020.
- 6 Vine CAJ, Myers SD, Coakley SL, et al. Transferability of Military-Specific cognitive research to military training and operations. *Front Psychol* 2021;12:604803.
- 7 Knapik JJ, Ang P, Meiselman H, et al. Soldier performance and strenuous road marching: influence of load mass and load distribution. *Mil Med* 1997;162:62–7.
- 8 Mahoney CR, Hirsch E, Hasselquist L, et al. The effects of movement and physical exertion on soldier vigilance. *Aviat Space Environ Med* 2007;78:B51–7.
- 9 et alKobus DA, Brown CM, Wu L. Cognitive performance and physiological changes under heavy load carriage. Pacific science and engineering group report., 2011. Available: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a551468.pdf>
- 10 Eddy MD, Hasselquist L, Giles G, et al. The effects of load carriage and physical fatigue on cognitive performance. *PLoS One* 2015;10:e0130817.
- 11 Giles GE, Hasselquist L, Caruso CM, et al. Load carriage and physical exertion influence cognitive control in military scenarios. *Med Sci Sports Exerc* 2019;51:2540–6.
- 12 May B, Tomporowski PD, Ferrara M. Effects of backpack load on balance and decisional processes. *Mil Med* 2009;174:1308–12.
- 13 Epling SL, Blakely MJ, Russell PN, et al. Free recall and outdoor running: cognitive and physical demand interference. *Exp Brain Res* 2016;234:2979–87.
- 14 Brisswalter J, Collardeau M, René A. Effects of acute physical exercise characteristics on cognitive performance. *Sports Med* 2002;32:555–66.
- 15 Tomporowski PD. Effects of acute bouts of exercise on cognition. *Acta Psychol* 2003;112:297–324.
- 16 Brisswalter J, Arcelin R, Audiffren M, et al. Influence of physical exercise on simple reaction time: effect of physical fitness. *Percept Mot Skills* 1997;85:1019–27.
- 17 Yerkes RM, Dodson JD. The relation of strength of stimulus to rapidity of habit-formation. *J Comp Neurol Psychol*. 1908;18:459–82.
- 18 Kubitz KA, Mott AA. Eeg power spectral densities during and after cycle ergometer exercise. *Res Q Exerc Sport* 1996;67:91–6.
- 19 Smith M, Tallis J, Miller A, et al. The effect of exercise intensity on cognitive performance during short duration treadmill running. *J Hum Kinet* 2016;51:27–35.
- 20 Bhabhani Y, Maikala R. Gender differences during treadmill walking with graded loads: biomechanical and physiological comparisons. *Eur J Appl Physiol* 2000;81:75–83.
- 21 Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982;14:377??381–81.
- 22 Breeze J, Lewis EA, Fryer R, et al. Defining the essential anatomical coverage provided by military body armour against high energy projectiles. *J R Army Med Corps* 2016;162:284–90.
- 23 Smith SJ, Mason EA, Bowditch SC. *Dismounted close combat (DCC) cognitive burden: assessment framework. (Dstl/TR52376). Cited with permission from Dstl*, 2011.
- 24 Soveri A, Lehtonen M, Karlsson LC, et al. Test-Retest reliability of five frequently used executive tasks in healthy adults. *Appl Neuropsychol Adult* 2018;25:155–65.
- 25 Jaeggi SM, Buschkuhl M, Perrig WJ, et al. The concurrent validity of the N-back task as a working memory measure. *Memory* 2010;18:394–412.
- 26 Hopkins WG, Marshall SW, Batterham AM, et al. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 2009;41:3–12.
- 27 Stirling L, Kelty-Stephen D, Fineman R, et al. Static, dynamic, and cognitive fit of Exosystems for the human operator. *Hum Factors* 2020;62:424–40.
- 28 Burnley M, Jones AM. Oxygen uptake kinetics as a determinant of sports performance. *Eur J Sport Sci* 2007;7:63–79.
- 29 Dietrich A. Transient hypofrontality as a mechanism for the psychological effects of exercise. *Psychiatry Res* 2006;145:79–83.
- 30 Bell R, Cardello A. V, Schutz HG. Relationship between perceived clothing comfort and exam performance. *Fam Consum Sci Res J* 2005;33:308–20.

## ADDITIONAL RESULTS

### Symptoms

Only 6/12 men and 1/10 women were able to complete the full exercise test under all conditions (See Table 3 in main body of manuscript). Test withdrawals were related to self-reported discomfort caused by the load carried (shoulder discomfort, blisters and chaffing around the hips and groin) or muscular discomfort. Reports of discomfort increased with load in both sexes, but the frequency of reporting was greatest in women (See Table 1). No withdrawals were related to work rate.

Table 1: Symptoms reported by participants post exercise test

|                             | Men |     |      |        | Women |     |      |        |
|-----------------------------|-----|-----|------|--------|-------|-----|------|--------|
|                             | LOW | MED | HIGH | V-HIGH | LOW   | MED | HIGH | V-HIGH |
| Shoulder discomfort         | 1   |     | 5    | 8      | 2     | 5   | 7    | 6      |
| Chaffing                    |     | 2   |      | 2      | 1     | 4   | 3    |        |
| Blisters / hot spots (feet) |     | 2   |      | 2      |       | 2   | 5    | 2      |
| General discomfort          | 1   |     | 1    |        |       |     |      | 2      |
| Helmet discomfort           | 1   |     |      |        | 1     |     |      |        |
| Back discomfort             |     | 2   |      | 1      | 3     | 3   | 3    | 2      |
| Hip discomfort              |     | 1   | 2    |        | 1     | 3   | 1    |        |
| Neck discomfort             |     |     |      | 2      |       | 1   | 1    | 2      |
| Legs tired                  |     |     |      | 1      |       | 1   | 1    | 1      |
| Hamstring discomfort        |     |     |      | 1      |       |     |      |        |
| Feeling hot                 |     |     |      |        | 2     |     |      |        |
| Knee discomfort             |     |     |      |        | 1     | 1   |      |        |
| Arm discomfort              |     |     |      |        | 1     |     |      |        |
| Shin discomfort             |     |     |      |        | 1     |     | 1    |        |
| Calf discomfort             |     |     |      |        |       |     |      | 1      |
| Breathing discomfort        |     |     |      |        |       | 1   | 3    | 1      |
| Chest discomfort            |     |     |      |        | 1     |     | 1    | 1      |

*Note that not all participants marched for the same length of time as some withdrew from the test due to discomfort.*

## Cognitive Tests

Table 2: N-back accuracy data for men and women measured in four load configurations during each hour of the march.

|               |               | <i>Men</i>       |                  |                  | <i>Women</i>    |                 |                 |
|---------------|---------------|------------------|------------------|------------------|-----------------|-----------------|-----------------|
|               |               | <i>Hour 1</i>    | <i>Hour 2</i>    | <i>Hour 3</i>    | <i>Hour 1</i>   | <i>Hour 2</i>   | <i>Hour 3</i>   |
| <i>0-back</i> | <i>LOW</i>    | 97.0 (3.3) [12]  | 98.0 (2.7) [11]  | 97.0 (2.3) [11]  | 97.0 (1.9) [9]  | 96.1 (2.9) [8]  | 98.2 (2.2) [6]  |
|               | <i>MED</i>    | 96.7 (4.3) [12]  | 97.4 (2.4) [11]  | 97.2 (2.8) [11]  | 97.0 (2.9) [8]  | 94.6 (5.1) [7]  | 96.9 (2.5) [5]  |
|               | <i>HIGH</i>   | 96.6 (3.6) [11]  | 97.8 (2.4) [11]  | 91.3 (13.1) [11] | 94.1 (5.8) [6]  | 89.5 (8.3) [7]  | 97.8 (1.8) [4]  |
|               | <i>V-HIGH</i> | 96.4 (4.1) [11]  | 95.6 (3.4) [8]   | 96.8 (2.8) [7]   | 84.0 (14.1) [6] | 93.3 (-) [1]    | 97.8 (-) [1]    |
| <i>1-back</i> | <i>LOW</i>    | 96.5 (2.6) [12]  | 97.0 (2.3) [11]  | 96.9 (3.0) [10]  | 94.8 (1.9) [9]  | 92.5 (6.6) [8]  | 93.8 (4.0) [6]  |
|               | <i>MED</i>    | 95.9 (4.0) [12]  | 95.0 (6.1) [12]  | 93.1 (6.6) [11]  | 93.8 (3.9) [8]  | 91.1 (3.1) [7]  | 96.2 (2.0) [5]  |
|               | <i>HIGH</i>   | 96.0 (4.6) [11]  | 95.6 (4.0) [11]  | 92.7 (5.9) [11]  | 90.1 (6.7) [8]  | 89.8 (7.8) [6]  | 91.3 (6.1) [4]  |
|               | <i>V-HIGH</i> | 89.7 (13.7) [11] | 93.9 (5.0) [8]   | 92.4 (6.6) [7]   | 78.4 (11.9) [5] | 84.0 (-) [1]    | 91.0 (-) [1]    |
| <i>2-back</i> | <i>LOW</i>    | 92.4 (7.1) [12]  | 94.3 (4.9) [11]  | 91.3 (6.6) [11]  | 92.6 (5.3) [9]  | 89.7 (7.1) [8]  | 89.3 (7.9) [6]  |
|               | <i>MED</i>    | 88.7 (11.2) [12] | 90.7 (10.8) [11] | 85.7 (11.6) [11] | 90.0 (6.1) [8]  | 91.0 (9.7) [7]  | 86.2 (15.0) [5] |
|               | <i>HIGH</i>   | 92.0 (8.8) [10]  | 89.3 (11.3) [11] | 88.1 (11.6) [11] | 82.9 (11.1) [7] | 87.2 (17.0) [6] | 94.5 (4.4) [4]  |
|               | <i>V-HIGH</i> | 89.3 (8.6) [10]  | 87.2 (11.3) [8]  | 74.0 (30.5) [7]  | 74.0 (8.7) [3]  | 71.0 (-) [1]    | 76.0 (-) [1]    |
| <i>3-back</i> | <i>LOW</i>    | 85.7 (8.3) [12]  | 87.1 (6.7) [11]  | 83.6 (9.1) [11]  | 81.6 (9.3) [8]  | 88.7 (5.9) [7]  | 87.7 (6.1) [6]  |
|               | <i>MED</i>    | 86.3 (11.4) [12] | 82.6 (11.0) [11] | 83.2 (11.8) [11] | 79.5 (8.6) [8]  | 84.9 (5.1) [7]  | 85.8 (3.8) [5]  |
|               | <i>HIGH</i>   | 87.3 (10.0) [10] | 86.7 (9.9) [11]  | 82.8 (10.8) [11] | 80.3 (8.2) [6]  | 84.7 (8.7) [6]  | 85.5 (7.3) [4]  |
|               | <i>V-HIGH</i> | 78.9 (10.3) [10] | 81.1 (10.7) [8]  | 75.6 (11.6) [6]  | 62.0 (22.0) [3] | 76.0 (-) [1]    | 80.0 (-) [1]    |

Data presented are mean (SD), the number of participants who completed the test is shown in square brackets.

Table 3: Visual Go/No Go accuracy data for men and women measured in four load configurations during each hour of the march.

|              |               | <i>Hour 1</i>    |                  | <i>Hour 2</i>    |                  | <i>Hour 3</i>    |                  |
|--------------|---------------|------------------|------------------|------------------|------------------|------------------|------------------|
|              |               | <i>Minute 5</i>  | <i>Minute 20</i> | <i>Minute 5</i>  | <i>Minute 20</i> | <i>Minute 5</i>  | <i>Minute 20</i> |
| <i>Men</i>   | <i>LOW</i>    | 83.7 (13.8) [12] | 85.8 (12.0) [12] | 88.8 (7.4) [11]  | 86.8 (9.3) [11]  | 87.7 (9.1) [11]  | 84.5 (12.1) [11] |
|              | <i>MED</i>    | 81.3 (15.9) [10] | 69.4 (26.1) [11] | 69.4 (21.4) [11] | 73.8 (26.3) [10] | 75.7 (16.7) [10] | 77.2 (20.0) [10] |
|              | <i>HIGH</i>   | 75.2 (19.9) [9]  | 69.2 (26.2) [10] | 73.9 (21.8) [9]  | 82.2 (15.5) [9]  | 78.7 (16.9) [9]  | 76.5 (18.0) [9]  |
|              | <i>V-HIGH</i> | 81.8 (14.0) [11] | 82.1 (16.2) [11] | 83.0 (16.0) [10] | 87.8 (7.7) [9]   | 85.7 (8.9) [7]   | 83.8 (12.4) [7]  |
| <i>Women</i> | <i>LOW</i>    | 84.5 (7.4) [7]   | 90.4 (6.0) [8]   | 86.0 (7.7) [8]   | 82.5 (15.5) [8]  | 71.5 (118.2) [8] | 88.3 (8.2) [7]   |
|              | <i>MED</i>    | 70.5 (11.9) [7]  | 81.7 (16.6) [8]  | 72.5 (20.2) [6]  | 79.7 (20.8) [6]  | 77.8 (18.0) [6]  | 85.0 (18.3) [5]  |
|              | <i>HIGH</i>   | 72.4 (12.4) [7]  | 66.4 (18.2) [7]  | 86.0 (6.3) [8]   | 78.1 (13.7) [7]  | 76.7 (18.7) [5]  | 70.7 (23.2) [5]  |
|              | <i>V-HIGH</i> | 79.4 (10.8) [6]  | 69.8 (12.1) [7]  | 82.1 (7.0) [4]   | 75.0 (13.0) [3]  | 85.0 (7.1) [2]   | 65.0 (35.4) [2]  |

Data presented are mean (SD). The number of participants who completed the test is shown in square brackets.

Table 4: Auditory Go/No Go accuracy data for men and women measured in four load configurations during each hour of the march.

|              |               | <i>Hour 1</i>    |                  | <i>Hour 2</i>    |                  | <i>Hour 3</i>    |                  |
|--------------|---------------|------------------|------------------|------------------|------------------|------------------|------------------|
|              |               | <i>Minute 10</i> | <i>Minute 25</i> | <i>Minute 10</i> | <i>Minute 25</i> | <i>Minute 10</i> | <i>Minute 25</i> |
| <i>Men</i>   | <i>LOW</i>    | 96.5 (3.5) [8]   | 96.7 (6.8) [8]   | 95.6 (6.7) [8]   | 94.8 (6.8) [8]   | 95.4 (6.0) [8]   | 93.1 (11.8) [8]  |
|              | <i>MED</i>    | 96.0 (5.9) [8]   | 94.4 (6.6) [8]   | 95.7 (6.4) [7]   | 94.5 (6.0) [7]   | 95.0 (7.6) [8]   | 95.2 (5.5) [8]   |
|              | <i>HIGH</i>   | 97.3 (6.4) [8]   | 95.4 (5.0) [8]   | 96.2 (7.3) [8]   | 94.8 (9.4) [8]   | 93.5 (8.4) [8]   | 95.8 (6.7) [8]   |
|              | <i>V-HIGH</i> | 95.7 (8.5) [7]   | 95.5 (8.4) [7]   | 93.8 (8.0) [7]   | 94.2 (6.8) [6]   | 97.3 (2.5) [5]   | 94.3 (6.5) [5]   |
| <i>Women</i> | <i>LOW</i>    | 94.8 (4.6) [9]   | 96.7 (4.1) [9]   | 94.4 (4.6) [8]   | 95.2 (3.1) [8]   | 94.0 (6.0) [8]   | 95.8 (3.1) [6]   |
|              | <i>MED</i>    | 93.6 (6.6) [7]   | 92.4 (7.4) [7]   | 89.5 (5.8) [7]   | 89.6 (7.6) [5]   | 93.9 (4.2) [3]   | 91.3 (7.6) [4]   |
|              | <i>HIGH</i>   | 94.3 (4.6) [7]   | 93.6 (4.7) [7]   | 89.3 (7.3) [7]   | 83.9 (18.2) [6]  | 87.7 (5.8) [5]   | 86.3 (17.7) [4]  |
|              | <i>V-HIGH</i> | 95.7 (3.8) [5]   | 86.4 (12.8) [6]  | 88.0 (7.2) [5]   | 82.1 (14.9) [4]  | 82.8 (10.2) [3]  | 77.2 (15.5) [3]  |

Data presented are mean (SD). The number of participants who completed the test is shown in square brackets.

Table 5: Rating of Perceived Exertion in Men Measured During the Exercise Test.

|         | Hour 1 |    |                 |      | Hour 2         |                |                   |      | Hour 3         |                  |                 |    |
|---------|--------|----|-----------------|------|----------------|----------------|-------------------|------|----------------|------------------|-----------------|----|
|         | BA     | AO | PO              | MO   | BA             | AO             | PO                | MO   | BA             | AO               | PO              | MO |
| N       | 12     | 12 | 11              | 10   | 12             | 12             | 11                | 8    | 11             | 11               | 11              | 7  |
| Minimum | 6      | 6  | 6               | 7    | 6              | 6              | 6                 | 7    | 6              | 6                | 7               | 8  |
| Maximum | 11     | 13 | 12              | 15   | 11             | 13             | 13                | 15   | 12             | 16               | 14              | 19 |
| Median  | 6.5    | 7  | 9 <sup>ab</sup> | 11.5 | 7 <sup>†</sup> | 7 <sup>†</sup> | 10 <sup>ab†</sup> | 10.5 | 7 <sup>†</sup> | 8.5 <sup>†</sup> | 10 <sup>†</sup> | 13 |

Notes: Data presented is for the end of each hour. "a" indicates where a difference from BA was observed; "b" indicates where a difference from AO was observed; † indicates a difference from the start of hour 1 ( $p < 0.05$ ).

Table 6: Rating of Perceived Exertion in Women Measured During the Exercise Test.

|         | Hour 1 |     |                   |                   | Hour 2           |     |                    |      | Hour 3           |      |                    |      |
|---------|--------|-----|-------------------|-------------------|------------------|-----|--------------------|------|------------------|------|--------------------|------|
|         | BA     | AO  | PO                | MO                | BA               | AO  | PO                 | MO   | BA               | AO   | PO                 | MO   |
| N       | 9      | 8   | 7                 | 6                 | 8                | 6   | 6                  | 3    | 6                | 5    | 4                  | 2    |
| Minimum | 7      | 7   | 8                 | 12                | 7                | 7   | 8                  | 13   | 7                | 7    | 8                  | 13   |
| Maximum | 9      | 10  | 12                | 17                | 13               | 12  | 14                 | 20   | 12               | 14   | 14                 | 17   |
| Median  | 8.0    | 8.5 | 10.0 <sup>a</sup> | 14.0 <sup>*</sup> | 8.5 <sup>*</sup> | 9.5 | 12.0 <sup>a†</sup> | 15.0 | 9.0 <sup>†</sup> | 11.0 | 11.0 <sup>a†</sup> | 15.0 |

Notes: Data presented is for the end of each hour. "a" indicates where a difference from BA was observed; † indicates a difference from the start of hour ; \* indicates a difference from men ( $p < 0.05$ ).