Refining Selection for Elite Troops by Predicting Military Training Outcome

Susan Vrijkotte; Romain Meeusen; Bart Roelands; Sabine Kubesch; Olivier Mairesse; Guy de Schutter; Nathalie Pattyn

INTRODUCTION: Paratrooper training courses are very demanding, leading to a high number of drop-outs, despite existing selection criteria. This study investigated physiological, neuropsychological, and subjective data of completers and drop-outs during paratrooper training to identify potential predictive indices.

METHODS: Tested were 24 paratrooper soldiers before (t_0), after 8 wk (t_1), and at the end of a 12-wk training camp (t_2). There were 11 soldiers who completed the course and 13 dropped out. The Training OPtimalisation test (TOPtest) uses two maximum exercise events to assess changes in measured parameters. The TOPtest was administered at t_0, t_1, and t_2; physiological [i.e., adrenocorticotrophic hormone (ACTH), cortisol, heart rate (HR)], neuropsychological (Stroop, Flanker, Go NoGo, Task Switch), and subjective data [Profile of Mood States (POMS)] were collected. Physiological and subjective raw data was gathered pre- and post-test from each of the two maximum exercise tests. The pre/post-test change of each parameter's raw values was calculated as the parameter's reactivity (or delta score).

RESULTS: At t_0, drop-outs showed a significantly smaller HR reactivity (117.9 $\pm$ 14.0 vs.107.7 $\pm$ 10.6). Delta scores of tension and fatigue values differed significantly between completers and drop-outs at t_0. Completers' physiological reactivity during the TOPtest at t_2 (HR: 105.91 $\pm$ 13.68 vs. 95.55 $\pm$ 10.28) was significantly reduced and became comparable to the drop-outs' reactivity at t_0. Delta scores of fatigue and tension values showed a similar pattern.

DISCUSSION: Reactivity of HR, tension, and fatigue parameter values were found to have predictive value in identifying completers vs. drop-outs of an elite paratrooper training course.

KEYWORDS: reactivity, predictive measures, military.

Vrijkotte S, Meeusen R, Roelands B, Kubesch S, Mairesse O, de Schutter G, Pattyn N. *Refining selection for elite troops by predicting military training outcome.* Aerosp Med Hum Perform. 2017; 88(9):850–857.

Sleep deficits, limited access to food, strenuous physical work, and psychological stress are all part of soldiers' environment during critical deployments. Despite these stressors, soldiers have to be able to operate in an accurate manner and have to make the right decisions at crucial moments. In order to prepare soldiers for these strenuous tasks and to select highly qualified personnel, specialized professional training courses are needed and have been developed within the different military services. An example of these specialized training courses is the paracommando training of the Belgian Army, part of the airborne force. It is believed to be one of the toughest military courses in Belgium because of the strain that is experienced by the participants. A common feature of these intensive military training courses is the high number of drop-outs. Only a limited number of participants complete this course despite the rigorous selection criteria that participants must meet beforehand. Being able to predict which participants are going to drop out during the course would reduce injury and overall costs.

Test batteries, like the Pilapt (pilot aptitude) computer-based system, have been developed for selection purposes.[14] These are useful in situations during which many participants are being tested. However, for elite troop training courses only a few participants apply. Therefore, a very precise and sensitive test is needed to predict soldiers' performance to avoid false negatives. Pattyn et al.[24] already showed that physiological testing can

From Vrije Universiteit Brussel, Brussels, Belgium; James Cook University, Queensland, Australia; the Royal Military Academy, Brussels, Belgium; and Institutet Bildung Plus, Heidelberg, Germany.

This manuscript was received for review in December 2016. It was accepted for publication in May 2017.

Address correspondence to: Romain Meeusen, Human Physiology Research Group, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium; rmeeusen@vub.ac.be.

DOI: https://doi.org/10.3357/AMHP.4818.2017

provide important predictive performance information with regard to stressful situations. Particularly the investigation of 'reactivity' was found to be a powerful tool. Reactivity provides information about the acute ability of a parameter to respond to a stressor in comparison to a resting situation, giving insight into the remaining functional reserve of the parameter.[25] Long-term stress and strain can negatively affect the psychophysiological functional reserve, leading to a diminished physiological reactivity and a decrease in performance.[25] The advantage of measuring reactivity is that participants cannot suppress their responses, whereas a single measurement only provides information about that specific moment in time, which can possibly be influenced by a short increase in effort.

Research on overreaching and the overtraining syndrome (OTS) in sports might be key literature in the search for appropriate markers in the military setting since encountered challenges[23] make it plausible that participants might drop out due to nonfunctional overreaching (NFO). Functional overreaching (FOR), or super compensation, is a temporary impairment of performance that later results in improved capabilities after days to weeks of continued effort. NFO is an undesirable state in which a temporary impairment of performance results in reduced capabilities for weeks or months. When OTS is being diagnosed, impaired performance will need even longer recovery. Based on the most recent insights on overreaching and OTS,[18] multiple markers, including hormone values, neuropsychological performance, and subjective information (self-evaluation), might help to predict soldiers' performance.

The stress-hormone adrenocorticotrophic hormone (ACTH) might be a suitable candidate to serve as a predictive marker. ACTH is produced by the pituitary gland, which normally reacts to exercise with an increase in ACTH serum concentration, mediated by the hypothalamus via corticotrophin releasing hormone.[20] Barron et al.[1] were the first authors to show hypothalamic dysfunction which caused decreased ACTH concentrations in athletes suffering from OTS. In addition, Meeusen et al.[20] developed a two-event exercise protocol (Training OPtimalization test; TOPtest) to determine whether an athlete might be suffering from OTS.[19] Normally, the ACTH serum concentration increases after the two maximal exercise tests. When the ACTH serum concentration decreases after the second maximal exercise test, this implies OTS. The drop of the ACTH serum concentration is thought to be caused by an exhaustion of the pituitary gland after reacting to the first exercise event.[18]

Self-evaluation is another important instrument that can be used to assess the physical and mental status of a subject. Athletes with OTS often report increased feelings of fatigue and depression,[18] which might also occur in soldiers suffering from NFO or OTS. For example, an 'iceberg-profile' with low scores on tension, depression, anger, and fatigue and a high score on vigor is a regular result for well-trained athletes when completing the Profile of Mood States (POMS).[22] This profile will diminish or be extremely reduced, depending on the severity of the OTS.[18]

In addition to physical impairment and affected subjective feelings, neuropsychological deterioration can also arise due to OTS.[13] Executive functions[21] are required to perform optimally in operational military settings: to control and rapidly switch attention, to plan actions and to suppress inappropriate responses, and to be able to focus on the most relevant information. The increased feelings of depression that have been reported in people suffering from OTS diminish a person's attention and executive function.[7] A decrease in executive function could accompany poor performance in military trainees, as well.

The aim of this study was to investigate whether physiological, neuropsychological and/or subjective data could provide information about the likelihood of a paratrooper soldier's ability to complete this advanced military training course. Possible predictive measures, including raw scores and reactivity (which is the change in a pre- to post- maximum exercise test parameter's score, or delta score) were investigated before the start of a training course, in parallel with Pattyn et al.[24] It was hypothesized that drop-outs would show physical, neuropsychological, and subjective responses that are different from completers and that these outcome measures would be similar to the responses that are being measured to identify NFO or even OTS.

METHODS

Subjects

Participating in this study were 24 healthy male army recruits. Of the subjects, 11 completed the training course while 13 of the subjects dropped out at some point during the training. Recruits dropped out due to injuries and problems coping with the severe demands and the rigor of the military discipline. Subjects did not have any medical history of psychiatric or neurological disorders and passed the medical army check. Participants had normal or corrected to normal vision. The study was approved by the Research Council of the Vrije Universiteit Brussels and the Bio-ethical Committee of the Belgian Army.

Procedure

Tests were conducted prior to, during, and at the end of a 12-wk training camp consisting of 8-wk paratrooper training (PT) and 4-wk specific commando training (ST). Participants were tested at three time points: at training onset (t_0), after 8 wk of PT (t_1), and after completion of the 4 wk of ST, which was the end of the 12-wk training camp (t_2). Testing took place at the Brussels Lab for Human Performance and Elite Sports from the Department of Human Physiology and Sports Medicine of the Faculty of Physical Education and Physiotherapy of the Vrije Universiteit Brussel in Brussels, Belgium. All tests were performed at the same time of the day within-subjects. Tests were conducted in a standardized order (**Fig. 1**).

Equipment/Materials

Blood was drawn at four moments throughout every experimental trial (see Fig. 1): as soon as the participants arrived, after the first maximal exercise test, shortly before, and then again after the second maximal exercise test. The blood was drawn in

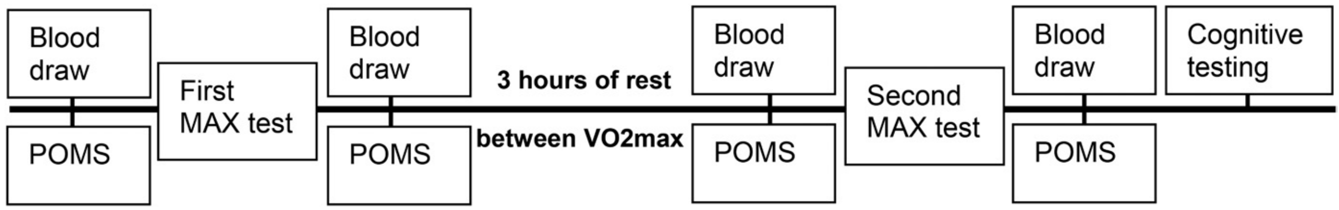

Fig. 1. Session design of the physical, subjective, and neuropsychological measurements.

precooled (−20°C) 4.5-ml K_3 EDTA vacutainer tubes (Beckton Dickinson Vacutainer System Europe, Plymouth, UK) and immediately centrifuged at 3000 rpm for 10 min (Minifuge 2, Heraeus, Germany). The plasma was pipetted in separate cups (Eppendorf, Wesselin-Berzdorf, Germany) and stored at −20°C until further analysis. Plasma was analyzed for ACTH (Nichols Institute Diagnostics, San Juan Capistrano, CA; VS), and free cortisol (DiaSorin, Stillwater, MN; 55,082).

The maximal exercise tests took place on a cycle ergometer (Lode Excalibur Sport, Groningen, Netherlands) in a climate chamber (Klimakamer, Weiss, Germany) with a constant temperature of 18°C and 60% humidity. The TOPtest consisted of two maximal exercise test protocols interspersed by 3 h of rest. Participants started with a workload of 80 W and a frequency between 80 and 90 rpm. Workload was increased by 40 W every 3 min until exhaustion. Heart rate (HR) was continuously measured (Polar Sporttester, Kempele, Finland). $\dot{V}O_2$ was measured throughout the test (Metamax, Cortex Biophysics, Leipzig, Germany) for the determination of maximum $\dot{V}O_2$ ($\dot{V}O_{2max}$). Blood samples (20 µl) were taken every 3 min during the test at the hyperemized (Forapin®, Heinrich Mack Nachf, Illertissen, Germany) earlobe for lactate determination (ESAT 6660, Medingen, Germany). Outcome measures for each $\dot{V}O_{2max}$ test were maximum workload ($Watt_{max}$), $\dot{V}O_{2max}$, maximum HR (HR_{max}), and maximum lactate (Lac_{max}).

All neuropsychological experiments were conducted by means of a standard PC (Apple Macintosh clone, 32 MB RAM, 180 MHz CPU) and a 17" color monitor (screen resolution 1024 × 768, 75 Hz; computer and monitor manufactured by Umax Systems GmbH, Willich, Germany) positioned at eye level and at a distance of 70 cm from the participant. The desktop microcomputer was equipped with PsyScope 1.2.5.[5] The keyboard was placed on a table in front of the participant at a comfortable distance. For ease of use the response keys were the SPACE bar or keys marked with different colored stickers. The overall computerized testing procedure, including short breaks, took about 30 min. Participants were instructed to respond as quickly and accurately as possible. Tests were conducted in the native language of the participants (either French or Dutch). Prior to the study participants were repeatedly trained to become familiar with the neuropsychological tests in order to avoid simple learning effects. Familiarization trials varied between one (Flanker Task), two (Stroop Task and GoNogo Task), and three test sessions (Task Switch).

The Stroop task[26] consisted of three conditions where color words were congruent with the presented color (e.g., the word "blue" in blue color), or incongruent (e.g., the word "blue" in red color). A neutral condition with colored strings of the letter x served as the control condition (e.g., "xxxx" in blue color). Stimuli were presented in the center of the screen on a black background (**Fig. 2A**). Five different colors were used (blue, red, yellow, green, and brown). The start of each trial was initiated by the participant by pressing the space bar, followed by a fixation cross in the middle of the screen for the duration of 700 ms. With appearance of the test stimulus participants had to press one of five response keys marked with the same color as the stimuli. A total of 120 trials were presented with equal probability of all three conditions. Reaction times and errors during the neutral, incongruent, and the congruent condition were defined as the dependent variables.

During the GoNogo task,[2] a centered fixation cross was presented for parametrically varied durations between 400 and 800 ms in steps of 10 ms. At the end of cross presentation, a black dot randomly appeared at 1 of 11 positions that were arranged in a clockwise fashion at a distance of between 4 and 8 cm from the center. For the Go condition, participants had to press a key whenever the central cross disappeared with presentation of the dot. During NoGo, participants were asked not to press any key when the cross remained visible on the screen (**Fig. 2B**). Responses were given by pressing a predefined key (the space bar) with the index finger of the dominant hand. The mean number of Go trials was 225. Average reaction times and errors during the task were defined as the dependent variable.

The Task Switch consisted of two alternating conditions with the same stimuli but different instructions.[3] In the center of the screen a string of between one and nine (except for five) identical digits appeared, with values ranging from one to nine with exception of the value of five (**Fig. 2C**). Digit strings were presented for a duration of 1000 ms. In the digit value condition (D) participants were asked to judge the value of the numbers and to press the left or right key when the string value was lower or higher than five, respectively. In the element number condition (E) participants had to indicate whether the number of digits was smaller or larger than five. One block consisted of five to eight trials per condition. The response-stimulus interval was 200 ms. Switches between blocks were indicated by a short cue on the PC monitor ("Digit value!!" and "Element number!!") for a duration of 1000 ms. Reaction times were recorded. Participants were informed that no within-block switches occurred. Altogether, 40 task switches had to be performed (D to E switches were of the same frequency as E to D switches). The dependent variable errors were defined as "switch costs": the increase in reaction times on switch trials compared to nonswitch trials.[15,17]

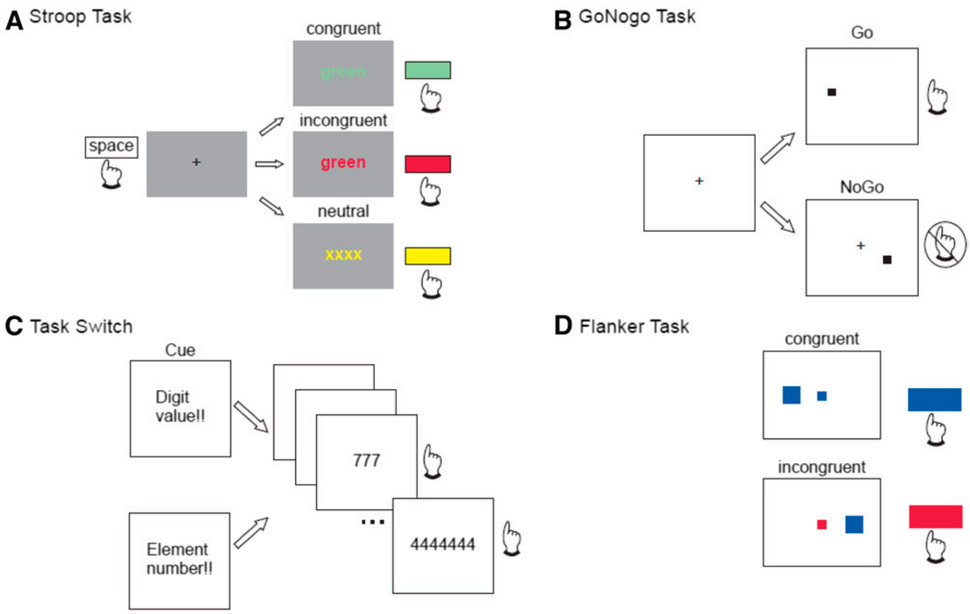

Fig. 2. Examples of the four neuropsychological tasks: A) Stroop Task, B) Go-NoGo Task, C) Task Switch, and D) Flanker Task.

During the modified Flanker task,[9] participants had to focus on the color of a small red or blue rectangle in the center of the screen (**Fig. 2D**). The center rectangle was flanked by a rectangle appearing 4.5 cm to the left or right. In the incongruent condition, the two rectangles were of opposite colors (blue and red). The width of the flanker was three times that of the center rectangle while the height was kept constant between the center and flanker. Participants had to respond depending on the color of the center rectangle by pressing one of two keys using the index and third finger of their dominant hand. To identify the same colors (congruent) or different colors (incongruent), center and flanker were displayed simultaneously until the participant responded. Appearance of the four combinations of center and flanker color was equiprobable with a probability of 25% (blue–blue, blue–red, red–blue, red–red). Consequently, the whole task comprised 50% congruent and 50% incongruent trials. The total number of trials was 220 (55 per condition), with a random order of the trials. The position of the flanker was balanced across subjects. Reaction times and errors during the congruent and incongruent trials were defined as the dependent variables.

The 32-item POMS consists of five subscales; tension, depression, anger, fatigue, and vigor. All items had to be scored from 0 (not at all) to 4 (extremely). The higher the score on a category, the more participants felt this mood state was present. The questionnaire was translated into the native language of the participants (French or Dutch).[10,28]

Statistical Analysis

Analyses were performed using IBM SPSS Statistics for Windows (SPSS, ver. 24, Chicago, IL). After checking for normality with a Shapiro-Wilk test, repeated measures ANOVAs were conducted in order to investigate differences between completers and drop-outs for the descriptive, physical, neuropsychological, and subjective data at t_0. When appropriate, *t*-tests were used. In order to measure reactivity, a 'delta score' was calculated for each parameter by subtracting the test outcome after the exercise event from the outcome before the exercise event. For the hormonal data, the test outcome after the exercise event was calculated as the percentage deviation from the test outcome before the exercise event, which was set to 100% to correct for interindividual variability. Delta scores of the first and second exercise event were analyzed using a repeated measures ANOVA with the delta scores being the dependent variables and the group the between factor. Delta responses of the completers at t_1 and t_2 were analyzed with *t*-tests. The confidence interval was set at 95% and alpha at 0.05.

RESULTS

No significant demographic differences were found between completers (age: 21.5 yr ± 2.4; height: 179.8 cm ± 6.7; weight: 75.9 kg ± 4.9; mean ± SD) and drop-outs (age: 20.4 yr ± 2.8; height: 178.1 cm ± 5.3; weight: 72.6 kg ± 6.6). There were also no significant differences found between completers and drop-outs for the raw ACTH [$F(3,57) = 0.87$, $P = 0.46$] and cortisol levels [$F(3,63) = 0.24$, $P = 0.87$] on t_0 (see **Table I**). Repeated measures of raw serum ACTH levels of the completers at t_0, t_1, and t_2 resulted in nonsignificant results [$F(6,42) = 2.16$, $P = 0.07$]. Raw serum cortisol levels were found to be significantly different between t_0, t_1, and t_2 for the completers [$F(2.88,28.81) = 4.85$, $P < 0.01$]. Post hoc analysis revealed that raw serum cortisol levels between t_0 and t_1 were significantly different before and after the first exercise event. At t_0, increased serum cortisol levels were found after the first exercise event, while at t_1 decreased serum cortisol levels were found after the first exercise event ($P < 0.05$). At t_0 and t_1 raw serum cortisol levels were found to be higher after the second exercise event compared to before, while at t_2 serum cortisol levels were found to be lower after the second exercise bout compared to the values before the exercise ($P < 0.05$) (detailed information in **Table A** can be viewed online at https://doi.org/10.3357/AMHP.4818sd.2017).

A postexercise increase in the variability of serum levels of ACTH and cortisol levels were found for all subjects. Visual inspection of the change from pre- to post-maximum exercise test scores (delta score) of ACTH serum concentration at t_0 revealed that the completers showed a small reactivity ACTH response after the first exercise event, whereas drop-outs

Table I. Mean (SD) of ACTH (Completers $N = 8$, Drop-Outs $N = 13$) and Cortisol (Completers $N = 11$, Drop-Outs $N = 12$) Levels at Different Recordings During t_0.

	ACTH ($pg \cdot mL^{-1}$)				CORTISOL ($nmol \cdot L^{-1}$)			
	FIRST $\dot{V}O_{2MAX}$		SECOND $\dot{V}O_{2MAX}$		FIRST $\dot{V}O_{2MAX}$		SECOND $\dot{V}O_{2MAX}$	
	BEFORE	AFTER	BEFORE	AFTER	BEFORE	AFTER	BEFORE	AFTER
Completers	46.2 (18.1)	51.9 (47.6)	18.8 (6.0)	46.6 (40.1)	37.4 (18.0)	42.3 (20.0)	16.0 (6.2)	31.8 (19.0)
Drop-outs	36.7 (17.8)	73.0 (58.7)	20.4 (21.9)	44.1 (20.0)	33.6 (19.3)	44.1 (24.7)	14.9 (6.4)	27.9 (12.4)

showed a large reactivity ACTH response after both exercise tests. However, these differences were not statistically significant [$F(1,19) = 0.313$, $P = 0.58$]. The delta scores of cortisol serum concentration when comparing completers to drop-outs was not found to be significantly different, either [$F(1,21) = 0.11$, $P = 0.74$], but showed a comparable pattern.

A repeated measures ANOVA for the delta score of ACTH showed nonsignificant results for the effect of time*delta [$F(2,14) = 1.02$, $P = 0.39$], but significant results were found for the effect of the deltas [$F(1,7) = 10.47$, $P = 0.01$]. At t_0 and t_1 the delta score of the second exercise event showed an increased response [t_0: $t(7) = -2.18$, $P = 0.07$; and t_1: $t(10) = -4.22$, $P < 0.01$], while hardly any changes were seen in the delta scores at t_2 [$t(10) = 0.67$, $P = 0.52$, **Fig. 3**].

For delta cortisol serum concentrations, significant differences were found for the effect of time*delta [$F(2,20) = 5.52$, $P = 0.01$]. A comparable pattern was found for delta cortisol serum concentrations as for the delta score of ACTH. At t_0 and t_1 the delta score of the second exercise event showed an increased response [t_0: $t(10) = -2.17$, $P = 0.06$; and t_1: $t(10) = -5.18$, $P < 0.01$], while hardly any changes were seen in the delta scores at t_2 [$t(10) = 0.22$, $P = 0.83$, Fig. 3]. In addition, the delta scores of ACTH and cortisol serum concentrations were found to show large interindividual variability at all times.

No significant differences were found between completers and drop-outs for raw data on $Watt_{max}$ [$F(1,22) = 0.23$, $P = 0.64$], $\dot{V}O_{2max}$ [$F(1,21) = 0.19$, $P = 0.67$], HR_{max} [$F(1,22) = 0.07$, $P = 0.80$], or Lac_{max} [$F(1,21) = 0.54$, $P = 0.47$] during the first and second exercise event (**Table II**). Although the delta scores of lactate values of the drop-outs was lower after the second exercise event compared to the completers at t_0, no significant differences were found (**Table III**) [$F(1,19) = 1.29$, $P = 0.27$]. However, the delta scores of HR values at t_0 were significantly different between completers and drop-outs [$F(1,21) = 7.26$, $P = 0.01$]. At t_0, the delta score of HR values of the completers remained stable, while the delta scores for HR values of the drop-outs declined after the second exercise event.

For completers, significant differences were found for the effect of time [$F(2,16) = 28.04$, $P < 0.01$] and delta [$F(1,8) = 5.85$, $P = 0.04$] for lactate. Post hoc analyses revealed significant overall higher delta lactate levels at t_0 compared to t_1 and t_2 ($P < 0.01$). Completers' delta scores for lactate values were significantly different at t_0 [$t(8) = 2.37$, $P = 0.05$], not significantly different at t_1 [$t(10) = 0.14$, $P = 0.89$], but significantly different again at t_2 [$t(10) = 2.52$, $P = 0.03$].

Repeated measures for delta scores of HR values showed a significant effect of time [$F(2,20) = 7.75$, $P < 0.01$]. At t_0 and t_1 delta scores of HR values revealed a comparable nonsignificant pattern [t_0: $t(10) = -1.51$, $P = 0.16$; and t_1: $t(10) = -0.48$, $P = 0.64$] with slightly higher delta scores after the second exercise event compared to the first. At t_2 nevertheless, a significantly [$t(10) = 2.27$, $P = 0.05$] different pattern was found with a significant decrease of the delta score of HR values after the second exercise event compared to the first, which is similar to the responses of the drop-outs at t_0.

Raw reaction times and error rates during the execution of the four neuropsychological tasks revealed no significant differences in values between completers and drop-outs. With the exception of the reaction times during the task switch [$t(20) = 3.27$, $P < 0.01$; detailed information in **Table B** can be viewed online at https://doi.org/10.3357/AMHP.4818sd.2017]. Drop-outs' reaction time values were significantly faster compared to completers' during this task at t_0. In addition, the error rate of the drop-outs during the task-switch test was higher compared to the amount of errors made by the completers, but this difference was not significant. Interindividual variability was found to be very large.

None of the raw scores of the five categories of the POMS (tension, depression, vigor, fatigue, and anger) were found to be significantly different between completers and drop-outs at t_0 (detailed information in **Table C** can be viewed online at https://doi.org/10.3357/AMHP.4818sd.2017). However, delta scores for tension values responses at t_0 were significantly different between completers and drop-outs [$F(1,22) = 7.34$, $P = 0.01$; **Fig. 4**]. At t_0, completers showed a large change in delta scores of tension values after the first exercise event and hardly any change after the second exercise event, whereas drop-outs showed hardly any change in delta scores of tension during the first and second exercise event. At t_0, delta scores of fatigue values showed a larger change in the delta scores of fatigue values for the drop-outs

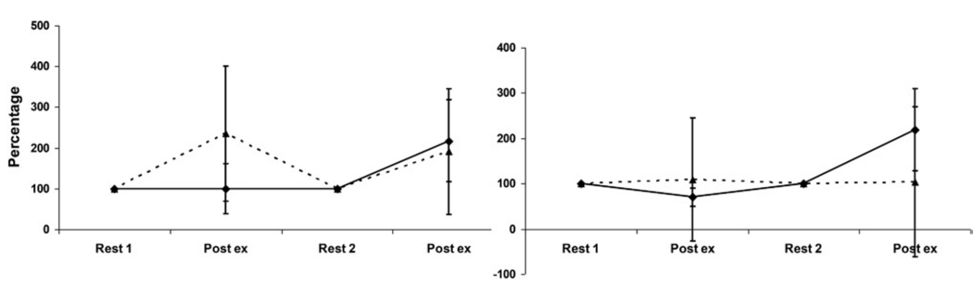

Fig. 3. ACTH (left) and cortisol (right) responses to the two exercise tests of the completers at t_1 (solid lines) and t_2 (dashed lines). Data are presented as percentage increase from both baseline values (SE) of the mean (ACTH and cortisol $N = 11$).

Table II. Mean (SD) of the Physiological Variables During t_0 (Completers $N = 11$, for Lacmax $N = 10$, drop-outs $N = 13$, for $\dot{V}O_{2max}$ $N = 12$).

	WATT$_{MAX}$ (W)		$\dot{V}O_{2MAX}$ (ml / min / kg)		HR$_{MAX}$ (bpm)		LAC$_{MAX}$ (nM)	
	FIRST MAX TEST	SECOND MAX TEST	FIRST MAX TEST	SECOND MAX TEST	FIRST MAX TEST	SECOND MAX TEST	FIRST MAX TEST	SECOND MAX TEST
Completers	259.6 (53.0)	260.9 (60.9)	53.0 (5.0)	51.7 (5.9)	191.4 (9.0)	189.9 (9.5)	8.8 (2.2)	7.5 (2.1)
Drop-outs	233.0 (29.5)	227.8 (32.8)	54.9 (8.0)	51.2 (10.2)	188.2 (10.6)	186.1 (11.3)	8.1 (1.8)	6.3 (1.0)

after the first exercise event compared to hardly any change for the completers [$F(1,22) = 4.17$, $P = 0.05$; Fig. 4]. Repeated measures of the POMS scores of the completers at t_0, t_1, and t_2 resulted in nonsignificant results. No significant differences in delta scores were found in the remaining POMS categories of depression, anger, and vigor between completers and drop-outs at t_0 [depression: $F(1,22) = 0.05$, $P = 0.82$; anger: $F(1,22) = 0.47$, $P = 0.50$; and vigor: $F(1,22) = 0.96$, $P = 0.34$].

For completers, repeated measures for the delta scores of the tension values revealed a significant effect of time*delta [$F(1.25,12.46) = 4.53$, $P < 0.05$]. Post hoc analysis revealed a significant difference between the delta scores of tension at t_0 compared to t_1 and t_2 ($P < 0.05$). Whereas at t_0 delta scores of tension showed a large negative reactivity during the first exercise event and nearly zero reactivity during the second, this pattern leveled off at t_1. At t_2 the pattern flipped and a small positive reactivity was found during the first exercise event followed by nearly zero reactivity during the second exercise bout. This pattern of reduced reactivity was found in the drop-outs at t_0 as well.

The delta scores of fatigue values revealed a significant effect of delta [$F(1,10) = 6.21$, $P = 0.03$]. At t_0 delta scores of fatigue values showed little reactivity, whereas at t_1 reactivity still is limited, the direction has been flipped compared to t_0. At t_2 the direction of the responses is comparable to t_1, but delta scores of fatigue values showed large reactivity during the first exercise event, whereas hardly any reactivity was seen after the second exercise event [$t(10) = 2.08$, $P = 0.06$]. The pattern found at t_2 is comparable to the pattern found at t_0 of the drop-outs. No significant differences were found for the delta scores of the other three POMS scales (detailed information in **Table D** can be viewed online at https://doi.org/10.3357/AMHP.4818sd.2017).

DISCUSSION

In this study physiological, neuropsychological, and subjective data were analyzed to investigate possible predictive measures to identify completers and drop-outs before the start of a paratrooper military training course. The main finding of this study is that the measures of reactivity of HR, tension, and fatigue values might be tools that could be used to identify completers vs. drop-outs before the start of a training course. HR reactivity at t_0 was found to be reduced in drop-outs, but not in completers of the training course. In addition, the reactivity of tension and fatigue values was found to be higher in drop-outs compared to completers at t_0. Contradictory to what was expected, no significant differences were found between completers and drop-outs for reactivity of ACTH or cortisol values. Neuropsychological performance did not show any predictive value either.

Although several studies have shown that reductions in physiological data can identify NFO or OTS,[20] no significant differences were found between completers and drop-outs in the raw data of the physiological variables measured during the two exercise test events. However, when investigating reactivity of a parameter (the change in a parameter's values between pre- and post-maximum exercise), differences were observed between completers and drop-outs. Drop-outs showed a significantly smaller reactivity of HR values at t_0 compared to the completers. This finding is in line with our expectations. For completers, during t_1 no significant differences were observed for delta scores of HR and lactate values, but during t_2 significant differences in the reactivity of these parameters were found. We surmise that completers' reactivity changed after having completed the strenuous course and being exhausted.

Vrijkotte et al.[27] monitored soldiers during a Dutch military training course in which participants went home over the weekends. Measurements of raw data revealed that participants started at baseline values on the Monday after the weekend that was spent at home. These findings seem to show that although military training courses can be very strenuous, rigorous, and tough, non-OTS participants need little time to recover. When comparing the results found in this study, it would be interesting to reinvestigate the data to determine the reactivity of HR values of the participants in the Vrijkotte et al. study.[27]

In addition, Gabbett[11] states that athletes who prepare for competition with high training loads reduce their injuries and achieve better results and resilience during competition. This might account for the performance of the completers during this military training course, as well. Drop-outs experienced

Table III. Mean (SD) of the Delta HR and Delta Lactate for Completers and Drop-Outs (HR Completers $N = 11$, Drop-Outs $N = 12$; Lactate t_0 Completers $N = 9$, Drop-Outs $N = 12$; t_1 and t_2 Completers $N = 11$).

		DELTA HR (bpm)		DELTA LACTATE (nM)	
		FIRST MAX TEST	SECOND MAX TEST	FIRST MAX TEST	SECOND MAX TEST
Drop-outs	t_0	117.9 (14.0)	107.7 (10.6)	7.4 (1.7)	5.3 (1.2)
Completers		112.6 (11.9)	114.6 (9.4)	7.6 (2.4)	6.5 (2.0)
Completers	t_1	105.1 (10.1)	108.0 (19.4)	4.4 (1.4)	4.3 (1.1)
Completers	t_2	105.9 (13.7)	95.5 (10.3)	5.0 (1.3)	4.1 (1.2)

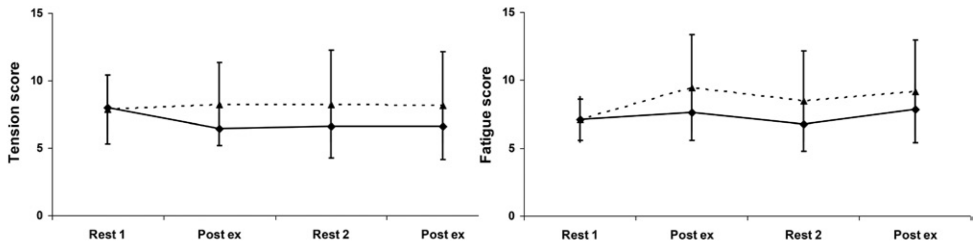

Fig. 4. Tension (left) and fatigue (right) ratings from the two exercise tests of the completers (solid lines) and the drop-outs (dashed lines) at t_0 (completers $N = 11$, drop-outs $N = 13$).

injuries or had problems coping with the severe demands and rigor of the course and were found to have smaller reactivity of HR values compared to completers. No data were collected about the amount of exercise that participants conducted when preparing for the training. For future studies it is recommended to gather this information and combine these results with the outcome measures from the TOPtest. This could provide a clearer picture of the participant's state before the start of a training course and may identify another potential parameter to assist in the selection process.

It was expected that hormonal responses of the completers at t_0 would reveal a normal pattern, whereas drop-outs would possibly show an altered response comparable to those found in athletes with NFO. This expected altered response of the drop-outs might be related to a dysfunction of the hypothalamus or the pituitary gland.[1,18,20] Nevertheless, no significant differences were found for the raw data or reactivity of ACTH and cortisol values between completers and drop-outs. The absence of significant differences might be due to the small sample size and large intra-individual variability.

The reactivity of subjective values of completers revealed a smaller reaction to the two-event exercise test. Ratings of tension and fatigue in the drop-outs, on the contrary, increased during the two-event exercise test, which is like the results found in athletes suffering from NFO.[18] This means that the reactivity of the tension and fatigue values did confirm our expectations: drop-outs experienced larger reactivity, while completers experienced smaller reactivity. Fatigue in this study was measured using the POMS. Another test that can be used to objectively measure arousal is the Critical Flicker Fusion Frequency. This test might add objective information about the arousal of a participant and, thus, the level of fatigue, while the test is easy to conduct, takes little time, and is noninvasive.[4]

No significant differences were found for neuropsychological task performance when comparing completers' results with the results of the drop-outs. Neither did task results differ for completers over time in the training course. Although it is known that executive functions are affected when soldiers become fatigued, our study was unable to find differences in the neuropsychological performance values of completers and drop-outs. This might be due to the fact that no reactivity for neuropsychological performance values could be calculated in our study. Cognition was tested only once, upon acute fatigue, while physiological and subjective tests were administered four times during a test day. In addition, there is a possibility that the neuropsychological tasks used in this study were not sensitive enough to measure neuropsychological changes. Dupuy et al.[8] used the Stroop to identify neuropsychological changes in overreaching athletes and their effects were small. Motivational aspects can affect task results and so might the complexity of the task, with too-easy tasks being unable to discriminate NFO or OTS.[18] Gardner et al.[12] showed that neuropsychological performance might be preserved by a change in neuropsychological strategy to mask the decline. This principle was observed in our study during the task switch. Reaction times of the drop-outs were significantly faster compared to completers, but the amount of errors made was higher as well. This means that the drop-outs used a different strategy during the task which resulted in a different speed/accuracy trade-off.

Raw physiological, neuropsychological, and subjective data did not reveal any significant differences between completers and drop-outs at t_0. Nevertheless, when reactivity scores were calculated, significant differences came to light between completers and drop-outs. When using reactivity scores, the range of resource mobilization of a participant is represented. Reactivity scores give information about the acute ability of a parameter to respond to an acute stressor in comparison to a resting state.[25] Raw data provides information only about a specific moment in time, which lacks information about the functional reserve of a participant. When the range of resource mobilization changes this is a sign of significant, subjective stress.[16] Reactivity scores can be calculated for physiological, neuropsychological, and subjective values. Pattyn et al.[24] reported physiological reactivity to be a promising early indicator for performance decrement. In our study, raw physiological data was not helpful in identifying drop-outs, but the reactivity of HR, tension, and fatigue values were significant. When considering the outcome of our study, it would be worthwhile to discuss the inclusion of reactivity scores for selection purposes. Reactivity appears to be more sensitive to change and thus give more information about the likelihood that a participant would succeed in a military training course.

To be able to calculate reactivity scores, the TOPtest is a valuable tool. The test is easy to administer, testing requires 1 d, and physiological, as well as subjective information, can be gathered. But above all, by exposing participants to a strenuous maximum exercise test regimen twice, the TOPtest can measure someone's functional reserve better than a single maximum exercise test. This repetitive testing makes it less possible to mask a reduced range of resource mobilization.

Although we did expect significant differences in neuropsychological, physiological, and subjective parameter values, most of the results did not separate completers and drop-outs. We noted that our measures showed large interindividual differences. This holds true for hormonal data as well as the neuropsychological and subjective data. Small group sizes in combination with large interindividual differences may have limited our measures' effectiveness. Similarly, because there were few drop-outs between t_1 and t_2, we could not identify any significant parameters to suggest who might succeed. In the context of elite paratrooper training, strong group pressure can influence the outcome of the subjective data by encouraging social desirable responses. It is recommended to use the social desirability questionnaire[6] in future studies, as this tool may be able to identify motivational influences that would interfere with accurate collection of subjective data. Participants in the current study were French and Dutch native speakers. Therefore, the POMS was presented in two languages. Differences in languages might have caused some differences in the subjective scores given and should be omitted in future studies.

In conclusion, the reactivity of HR, tension, and fatigue values in response to two maximum exercise testing events before the start of an elite paratrooper military training school (t_0) were found to provide significant differences that separated course completers from drop-outs. In addition, completers' reactivity scores at course completion (t_2) were found to be comparable to the reactivity scores of the drop-outs at the start of the course (t_0). These findings indicate that reactivity scores of HR, tension, and fatigue values are parameters that identified performance differences in these military training participants. Further studies using an individual's reactivity scores could build support for the use of these parameters as a selection tool before the start of elite military training.

ACKNOWLEDGMENTS

This study was supported by the Medical Component of the Belgian Defence. We would like to thank the paratrooper participants for their cooperation.

Authors and affiliations: Susan Vrijkotte, M.Sc., B.Sc., Romain Meeusen, Ph.D., M.Sc., Bart Roelands, Ph.D., M.Sc., and Guy de Schutter, M.D., M.Sc., Vrije Universiteit Brussel, Brussels, Belgium; Sabine Kubesch, Ph.D., M.Sc., University of Ulm, Ulm, Germany; and Olivier Mairesse, Ph.D., M.Sc., and Nathalie Pattyn, Ph.D., M.D., Royal Military Academy, Brussels, Belgium;

REFERENCES

1. Barron JL, Noakes TD, Levy W, Smith C, Millar RP. Hypothalamic dysfunction in overtrained athletes. J Clin Endocrinol Metab. 1985; 60(4):803–806.
2. Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature. 1999; 402(6758):179–181.
3. Casey BJ, Thomas KM, Welsh TF, Badgaiyan RD, Eccard CH, et al. Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. Proc Natl Acad Sci USA. 2000; 97(15):8728–8733.
4. Cavalade M, Papadopoulou V, Theunissen S, Balestra C. Heart rate variability and critical flicker fusion frequency changes during and after parachute jumping in experienced skydivers. Eur J Appl Physiol. 2015; 115(7):1533–1545.
5. Cohen JD, Mc Whinney B. PsyScope: A new interactive environment for designing psychology experiments. Behav Res Methods Instrum Comput. 1993; 25(2):257–271.
6. Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Psychol. 1960; 24(4):349–354.
7. Davidson RJ, Pizzagalli D, Nitschke JB, Putnam K. Depression: perspectives from affective neuroscience. Annu Rev Psychol. 2002; 53:545–574.
8. Dupuy O, Renaud M, Bherer L, Bosquet L. Effect of functional overreaching on executive functions. Int J Sports Med. 2010; 31(9):617–623.
9. Eriksen BA, Eriksen CW. Effects of noise letters upon identification of a target letter in a nonsearch task. Percept Psychophys. 1974; 16(1):143–149.
10. Fillion L, Gagnon P. French adaptation of the shortened version of the Profile of Mood States. Psychol Rep. 1999; 84(1):188–190.
11. Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter and harder? Br J Sports Med. 2016; 50(5):273–280.
12. Gardner RS, Uttaro MR, Fleming SE, Suarez DF, Ascoli GA, Dumas TC. A secondary working memory challenge preserves primary place strategies despite overtraining. Learn Mem. 2013; 20(11):648–656.
13. Hynynen E, Uusitalo A, Konttinen N, Rusko H. Cardiac autonomic responses to standing up and cognitive task in overtrained athletes. Int J Sports Med. 2008; 29(7):552–558.
14. Kokorian A, Valsler C, Burke E. International validation of a computerized testing suite for pilot selection. Paper presented at the Australian Aviation Psychology Symposium; Dec. 1, 2003. Victoria (Australia): Australian Aviation Psychology Association; 2003.
15. Leung HC, Skudlarski P, Gatenby JC, Peterson BS, Gore JC. An event-related functional MRI study of the stroop color word interference task. Cereb Cortex. 2000; 10(6):552–560.
16. Lovallo WR. Do low levels of stress reactivity signal poor states of health? Biol Psychol. 2011; 86(2):121–128.
17. MacDonald AW, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science. 2000; 288(5472):1835–1838.
18. Meeusen R, Duclos M, Foster C, Fry A, Gleeson M, et al. European College of Sport Science, American College of Sports Medicine Prevention, diagnosis and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. Med Sci Sports Exerc. 2013; 45(1):186–205.
19. Meeusen R, Nederhof E, Buyse L, Roelands B, de Schutter G, Piacentini MF. Diagnosing overtraining in athletes using the two-bout exercise protocol. Br J Sports Med. 2010; 44(9):642–648.
20. Meeusen R, Piacentini MF, Busschaert B, Buyse L, De Schutter G, Stray-Gundersen J. Hormonal responses in athletes: the use of a two bout exercise protocol to detect subtle differences in (over)training status. Eur J Appl Physiol. 2004; 91(2–3):140–146.
21. Monsell S, Driver J, editors. Control of cognitive processes. Attention and performance XVIII. Cambridge (MA): MIT Press; 2000.
22. Morgan WP, Brown DR, Raglin JS, O'Connor PJ, Ellickson KA. Psychological monitoring of overtraining and staleness. Br J Sports Med. 1987; 21(3):107–114.
23. Opstad K. Circadian rhythm of hormones is extinguished during prolonged physical stress, sleep and energy deficiency in young men. Eur J Endocrinol. 1994; 131(1):56–66.
24. Pattyn N, Migeotte PF, Morais J, Cluydts R, Soetens E, et al. Predictive performance assessment: trait and state dimensions should not be confused. J Gravit Physiol. 2008; 15(1):P-221–P-222.
25. Pattyn N, Migeotte PF, Neyt X, van den Nest A, Cluydts R. Comparing real-life and laboratory-induced stress reactivity on cardio-respiratory parameters: differentiation of a tonic and a phasic component. Physiol Behav. 2010; 101(2):218–223.
26. Stroop JR. Studies of interference in serial verbal reaction. J Exp Psychol. 1935; 18(6):643–662.
27. Vrijkotte S, Valk PJL, Raymann REM, Simons M, Veenstra BJ. Measuring physical strain and cognitive performance in the field during an Air Mobile Brigade training course. Soesterberg (Netherlands): TNO Defence, Security and Safety Department of Human Performance; 2010. Report No: TNO-DV 2010 A249.
28. Wald FDM, Mellenbergh GJ. De verkorte versie van de Nederlandse vertaling van de Profile of Mood States (POMS). Ned Tijdschr Psychol. 1990; 45:86–90.