CHANGES IN COGNITIVE PERFORMANCE DURING A 216 KILOMETER, EXTREME ENDURANCE FOOTRACE: A DESCRIPTIVE AND PROSPECTIVE STUDY 1,2

MICHAEL M. DOPPELMAYR
Department of Physiological Psychology
University of Salzburg

HOLGER FINKERNAGEL
Institute for Leistungsdiagnostik
Bad Berleburg

HILDEGARD I. DOPPELMAYR
Department Internal Medicine
Landeskrankenanstalten Salzburg

Summary.—Two subjects participated in a 216 km ultramarathon with outside temperatures above 50°C while several physiological and psychological parameters (cognitive performance assessed by a mental calculation task and an attentional task, subjective bodily experience, and lactate level) were evaluated throughout the race. Severe stress from dehydration, sleep deprivation, and total physical exhaustion are combined in a unique manner, allowing evaluation of their effects in a range far outside that obtainable in a laboratory setting. During the race the subjects answered a questionnaire about their actual bodily experiences, underwent 8 medical examinations, and performed two cognitive tests approximately every 35 kilometers. Analysis showed cognitive performance did not decrease steadily in a simple and gradual way but reached a peak in the morning of Day 2 after a short sleeping period and then decreased. In the early morning of Day 3, in general cognitive performance exhibited the worst results but increased differentially between the subjects again in the last test 1 km before the finish line.

The Badwater Ultramarathon is a footrace 216 km long which takes place in late July in the Death Valley, CA, where temperatures rise to more than 50°C. Participants have to run from Badwater (80 m below sea level) to Mt. Whitney. The finish line is at approximately 2500 m above sea level. Although the course record is 26 hours, most competitors need about 45 to 50 hours to complete the distance. The race is listed among the world’s toughest footraces and combines several highly interesting stressors related to cognitive performance, including extreme heat, dehydration, total physical exhaustion, and sleep deprivation.

Attentional processes and intact cognitive capacity under such challeng-

1The authors thank their respective racing teams for their encouraging support throughout the race and for collecting the data during these three days. Also, we thank the organizer of the Badwater Ultramarathon, Chris Kostman, for the possibility of performing this unique study during this event.
2Address enquiries to Michael Doppelmayr, Department of Physiological Psychology, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria or e-mail (michael.doppelmayr@sbg.ac.at).
ing conditions is a highly relevant issue for military situations and rescue operations (Amos, Hansen, Lau, & Michalski, 2000; Hocking, Silberstein, Lau, Stough, & Roberts, 2001), for research in space and aviation (Manzey & Lorenz, 1999; Faervik & Reinertsen, 2003; Sauer, 2004), for human performance under specific working conditions (Shalini, 2001), or for other sporting events in deserts or in hot cockpits (Mosier, Skitka, Flees, & Burdick, 1997; Lambert, Jan, & Charland, 1998). Although there is sufficient literature about individual stressors, there is no single paper describing a realistic combination, as done in this report. In most cases and for good reasons, ethical guidelines prevent researchers from using conditions that might strongly and negatively affect the subject's health and cognitive performance. Thus, such races provide an opportunity to study reactions to extreme physical stress. In this study two Badwater runners served as subjects.

Running in extreme heat can easily lead to dehydration, which is well known to affect cognitive performance negatively. Chian, Barraud, Melin, and Raphel (2001) have shown that dehydration affects cognitive performance similarly, whether induced by heat-stress or by exercise. In their experiment subjects either had to remain seated for approximately 2 h in a climatic chamber with temperatures between 45°C and 50°C or to exercise on a treadmill at 65% of VO2max for two hours with a temperature of 25°C. In both cases core temperature remained below 39°C. While for both dehydration conditions a similar decrease in performance for a simple reaction time task (pressing the left or right mouse button), a tracking task in which subjects had to follow a moving target with a joystick, and short-term memory controlling digit span, has been reported, there were no impairments for long-term memory or errors in the reaction time paradigm. Hocking, et al. (2001) investigated the effects of heat stress on the human electroencephalogram (EEG) and performance on a verbal learning test, inspection time, digit span, spatial working memory and a continuous performance task. To examine the effect of thermal stress three different experimental conditions have been compared. Subjects had to walk 40 minutes at 5 km per hour on a treadmill in either 25°C/65% relative humidity or in 35°C/65% relative humidity. In the third condition, again 35°C/65%, subjects additionally had to wear a Nuclear, Biological and Chemical protective suite (NBC Mk. IV). The core temperatures reached 37.3°C, 37.4°C and 38.7°C for the three experimental sessions. Steady state visual evoked potentials (SSVEP) were used to compare brain responses between different conditions. In this study design, a 13-Hz sinusoidal flickering light induced a 13-Hz EEG rhythm. During a working memory task as well as during a simple detection task, both the amplitude and the latency of this EEG response have been significantly different between the 25°C and the 35°C+NBC conditions, showing higher amplitudes and decreased latencies for the high temperature condition.

These differences have been observed in frontal as well as in parieto-occipital electrode sites. While for the behavioural data, there were no significant differences between the conditions for digit span forward, a significant decrease in digit span backwards has been observed for the 35°C+NBC condition. Inspection time too gave significant findings across these three conditions with the longest inspection times for 35°C+NBC.

Nielsen and Nybo (2003) reported that high core or high brain temperature might be a limiting factor for both cognitive and physical performance. They assume that fatigue, induced by hyperthermia either through heat or through exercise, is “located” in the cortex. As in the work of Hocking, et al. (2001), they stressed that hyperthermia can alter brain states and affect EEG and cognitive performance.

Collardeau, Brisswalter, and Audiffren (2001) showed differential effects of a 40 min. run on cognitive performance, indicating a decrease in performance at the beginning and a general improvement after prolonged exercise in a reaction time paradigm. In an extensive review Tomporowski (2003) analyzed three different types of studies and their respective effects on cognitive performance. The first was intense anaerobic exercise, which had differential results, showing both facilitation and impairment. The second type, submaximal aerobic exercise up to 60 min., in most cases led to either facilitation or a U-shaped improvement as the length of exercise increased. The third type, most relevant to our study, was a steady state aerobic exercise that too showed differential effects. Two of 15 studies reviewed by Tomporowski did not show a relation between aerobic exercise and cognitive function. Eleven of these studies provided evidence that this type of exercise, lasting below 60 min., exert a facilitative influence on selective cognitive functions, especially complex problem-solving and attentional processes. The remaining two studies showed that long lasting exercise that leads to depletion of physiological energy stores as the glycogen reserves in a muscle and liver, dehydration and the accompanying metabolic changes have negative effects on cognitive performance. These effects are visible in reduced short and long term memory or more complex reaction time tasks. Tomporowski (2003), therefore, concluded that bouts of exercise that do not deplete the physiological reserve of the respective subjects lead to facilitation; however, long lasting physical demands that lead to dehydration might well be sufficient to reduce cognitive performance.

Besides heat and dehydration the amount of physical exhaustion, as measured by lactate, is a relevant factor for cognitive performance during exercise. Several studies have shown that exercise can reduce reaction times in different cognitive tasks. Chmura, Nazar, and Kaciuba-Uscilko (1994) reported that the reaction time in a choice reaction-time paradigm decreased
with increasing workload until the participants reached approximately 75% VO$_2$ max. Exceeding this threshold led to a rapid increase in reaction times. In an additional study Chumla, Krysztofiak, Ziemba, Nazar, and Kaczyńska-Uscilko (1998) compared 20-min. and 60-min. exercises below and above lactate threshold. While the short exercise had to be performed at about 10% above the lactate threshold, the longer lasting exercise had to be performed at a level approximately 30% below this threshold. In both types of exercise reaction times decreased initially, but, while on the high intensity task reaction times decreased and heart rate increased throughout the 20 min., on the low intensity task reaction times reached a minimum after 40 min. and remained at this level.

Another interesting and highly relevant factor in this study is sleep deprivation. There is a wealth of literature discussing the effect of sleep deprivation on cognitive performance. The main finding of these studies (for a review and meta analysis, see Pilcher and Huxley, 1996) is a decrease in cognitive performance after a sustained period of sleep deprivation. With respect to the Badwater Ultramarathon the most relevant study has been performed by Edinger, Maush, McCall, Erwin, and Lininger (1990) reporting a decline in several performance measures in two subjects during a 146-hr. tennis match.

Besides partial or total sleep deprivation the usual circadian rhythm leads to variations in cognitive performance. Bonnecoe, Gisselbrecht, Hoefi, Eschenauer, Muzet, and Tass (2003) reported that the amount of impairment in cognitive performance is influenced both by time of day and task difficulty. They demonstrated a significant effect of time of day for a simple descending subtraction task. The shortest response times have been recorded in the evening and slowest during the night. The morning measures yielded intermediate results.

**Method**

**Badwater Ultramarathon**

Taking place in the Death Valley and in summer the Badwater Ultramarathon is the “hottest” and generally one of the most demanding footraces. Only experienced athletes (having finished several 100-km races) are allowed to participate. Nevertheless, usually 30% to 40% of the competitors have to quit because of heat stroke, dehydration, blisters, or exhaustion. During the race the runners are allowed to take breaks or sleep breaks. The race course runs along the main road, has several steep slopes and especially for the first 80 km, extreme temperatures up to and above 50°C. Runners are assigned to different starting times at either 6, 8, or 10 a.m., and finish after 60 hours.

**Tests**

The Badwater Ultramarathon is a highly emotional race, and participating in a study undisputably leads to a considerable loss of time. Because no other voluntary subjects could be recruited, two of the authors participated in this race. While Subject A (male, age 60, 10 a.m. starting group, finishing time 45 hr., 30 min.) had a very long running history (more than 150 marathons and 30 ultramarathons including the Badwater Ultramarathon), Subject B (male, age 41, 6 a.m. starting group, finishing time 54 hr., 12 min.) only marginally reached the necessary running experience (15 marathons, 5 ultramarathons).

Three times, during the two days preceding the race (baseline) and at seven stations, located approximately at km 30, 60, 95, 130, 160, 195, and 215 (one km before the finish line) test-sessions were performed, and the following psychological tests were conducted.

**KLT Konzentrations-Leistungstest** (Düker & Lienert, 1965).—We used simple mathematical tasks as $9-5+8=?$ taken from the KLT to evaluate the cognitive performance. In the original version of the KLT subjects are asked to perform a somewhat harder mathematical task, namely, adding two numbers and either adding or subtracting the result of a second addition, depending on whether the second result is higher or lower than the first result. We have found this task to be both too difficult and to last too long, to be performed several times during a race (the duration of the tests was a critical factor); therefore we only used the numerical examples provided by the KLT in the respective order. We decided to use these KLT calculation tasks instead of tasks with arbitrary numbers because these calculations are standardized by difficulty. While the subject walked, one member of the support crew accompanied him for exactly 3 min. reading the tasks aloud; the subject had to answer as fast and correctly as possible, and the respective result was noted. Because the number of performed calculations is dependent on speed of calculation as well as on the speed at which the task is read, we had to reduce differences in reading speed between the tests. To reduce the differences always the same reader, who was instructed to keep the reading speed as constant as possible, performed this task. Additionally, approximately 20 training tests (for reading speed) have been performed with each of the investigators to get an “average reading speed.” To reduce training and practice effects for the calculation tasks, both participants had trained on similar simple computations more than 20 times (with 40 single calculations in each training session).

**D2 Aufmerksamkeits-Belastungstest** (Brickenkamp, 1994).—The D2, an attention and concentration test, originally is a single page displaying 14 rows with 47 separate letters, either p or d. These letters are accompanied
by none, one, two or three little lines, similar to high or low apostrophes or quotations marks like 'p' or 'd' or 'p. All ds accompanied by exactly two such lines as "d or d" or â€œ, etc. served as targets and had to be marked, while all others combinations as "p 'd 'p, etc. were distractors and had to remain unmarked. Using the D2 several times within a few days would lead to memory effects in the first few rows. Therefore we modified this test and used a set of 10 different pages (signs in different sequences). On every test session the subject had 1 min. to mark all targets on one of these pages, working line by line. The number of correctly finished items (targets and distractors) was analyzed.

The results of the performance tasks, KLT (arithmetic) and D2 (attention), might have been influenced to some degree by the respective difficulty of the single tests. For example, the modified version of the D2 may not have had the exact difficulty for all lines. Although every line comprised (+1) the same number of correct and incorrect items, some might have accumulated in the first or second part of the line. This could have led to some inconsistencies in task difficulty. To account for this variation in difficulty 14 subjects (students of the local university) performed all training tests, within a longer period of 2 weeks as well as all of the actual race-tests within a time frame of 2 days. Had all tests been equally difficult, there should have been no variation in the performance of this control group. However, task difficulty did vary, so the results had to be adjusted accordingly. The respective test scores of Subjects A and B have been adjusted to the difficulty of the respective task by the following procedure. First, the actual data of the subjects were transformed into a percentage score according to the median for respective task of a control group. Then the percentages increase or decrease with respect to the baseline measures of Subjects A or B were calculated.

EKV Dimensionen der erlebten körperlichen Verfassung (Kleinert & Liesenfeld, 2001).—The EKV is designed to evaluate four dimensions of the subjective bodily experience (dimensions of experienced physical constitution) and comprises a set of 20 questions covering these four dimensions. While two of these dimensions, covering questions about Training and Flexibility are of little interest here, we concentrate on the dimensions of Health and Energy in this report. "Health" is evaluated by five items like pain, injury or suffering; the dimension of "Energy" covers items like power, energy or weakness. The subject has to rate questions by assigning a value from 0 to 5. As an example the question "How weak do you feel at the moment" had to be rated, with the anchor 0 indicating "not at all" and 5 "completely." While in the questionnaire the positive/negative scaling is changed several times to reduce left/right biases, for analysis the values had to be recalculated into a direction that higher values indicate more energy or a better physical health.

Including these tests, a single session comprised measures in the following order. While walking the ambient temperature was noted and the EKV followed by KLT were performed. Then the subject sat down and blood samples were taken to analyze lactate (Diaglobal Vario-Photometer). In the last step the subject had to perform the D2. Several additional parameters were evaluated, but due to temperature and methodological problems these values were not reliable; therefore only the described parameters are analyzed and discussed. A whole test session took 15 to 20 min.

RESULTS

Although performance data could be correlated with the EKV, lactate, mileage, or outside temperature, time of day or time out on the course, using the data of only two subjects means there are too many uncontrolled factors and therefore a descriptive presentation is more appropriate. Fig. 1a for Subject A and Fig. 1b for Subject B depict the percent of increase or decrease of the KLT and the D2 scores compared to a baseline measure and corrected for task difficulty. For the KLT, on the average of the three baseline tests, Subject A finished 35.3 calculations and Subject B only 26.1. Contrary to this, for the D2 Subject A only completed 65.6 items correctly while Subject B completed 93.8 (average of three baseline tests). Because A and B had quite different baseline values, the results are recalculated as a percentage score for a better comparison between the subjects as outlined in the Method section. The exact results of the respective tests are depicted in Table 1 in detail. The values depicted in Fig. 1a and Fig. 1b, therefore, represent an increase or decrease with respect to the baseline for the number of performed calculations in the KLT and the correctly finished items (ds with two lines) in the D2.

As depicted, both subjects show only marginal variations (±10%) in the KLT scores until km 160, but then performance drops to −15% at km 195. In more detail for Subject A, a small increase to 6.6% above the baseline, on the morning of Day 2 at km 95 can be observed, followed by a decrease to −8.4% at km 130. At km 160 the percentage score for completed calculations increased again to 6.6% above the baseline and then dropped to −15.8% for the last two sessions. For Subject B the KLT scores remained relatively stable until km 160 with only one exception at km 70 at which a decrease of 9% could be observed. Similar to Subject A at km 195, the KLT values were at a minimum with −13.8% below baseline but, contrary to Subject A, Subject B showed the best result immediately before the finish line reaching 8.2% above the baseline, which is an increase of approximately 22% from km 195 to km 215.

The results for the D2 show a much larger variation. The data of Subject A show an immediate drop to approximately −21%, compared to the
baseline during the first 12 hr. of the race, a time interval which corresponds to the distance until km 70. On the morning of Day 2, the values reach the maximum, marginally (4.9%) above baseline, and then decrease steadily until km 195 where the minimum of almost -25% is reached. During the last test the values return to baseline. For Subject B too an initial decrease to -11.6% can be observed, but then the values return to baseline until km 160, the evening of Day 2 and then decrease steadily until the end of the race, down to -16.6%.

| Table 1 | Percent of Changes of the KLT and D2 Values for Subjects A and B Compared to Baseline Measurement |
|---|---|---|---|---|---|
| Distance (km) | Subject A | % KLT Total | %D2 Correct | Subject B | % KLT Total | %D2 Correct |
| 30 | -2.93 | -21.63 | -7.7 | -11.59 |
| 70 | -1.81 | -21.31 | -9.01 | -5.13 |
| 95 | 6.60 | 4.89 | 2.24 | 1.8 |
| 130 | -8.38 | -2.35 | 0.24 | 0.55 |
| 160 | -6.60 | -14.32 | -1.53 | -8.36 |
| 215 | -15.84 | 1.03 | 8.20 | 16.65 |

Fig. 2 depicts the respective values for the EKV for the dimensions Health and Energy. As expected, for Subject B the values of the EKV continuously decreased; however, for Subject A after an initial decrease, lasting...
until km 70 for Energy and until km 95 for Health, the values increase again during Day 2 until km 160. Then both values decrease again until the end of the race but neither reach the minimum of km 70 or 95 nor the respective minimum of Subject B. In Fig. 3 the measure of lactate in mmol·l⁻¹ (mM) are depicted. These values indicate that Subject A maintained on a very low lactate level below 2 mM at the beginning of the race. These levels increased during Day 2 but remained below 4 mM until km 160. Approximately 6 hours before the end of the race at km 195 Subject A exhibited the highest lactate level reaching 8.5 mM. Subject B, on the other hand, exhibited the highest lactate levels right at the beginning at km 30 and km 70 marginally above 4 mM. Then this level decreased to below 2 mM and remained stable until the end of the race.

![Blood lactate concentration in mmol·l⁻¹ (mM) are shown for the respective race distance (test points). While Subject A (■) had low lactate levels until the end of the race, Subject B (○) had the highest values during the first part of the race.](image)

**DISCUSSION**

Beside the different ages and running histories a comparison of the scores of the two subjects leads to several additional problems because starting times, temperatures, or hours out on the course were different. Nevertheless, the results of the cognitive performance data show several similarities that let us conclude a comparison is justified if these factors are kept in mind. These similarities are a first decrease on Day 1, an increase to or partly above baseline on Day 2, followed by a decrease at night between Day 2 and Day 3 and the poorest scores in cognitive performance in the early morning of Day 3.

According to the results described by Tomporowski (2003), we expected an initial increase in cognitive performance until the glycogen reserves were almost depleted followed by a decrease after prolonged exercise. During Day 1 both subjects showed clear symptoms of total exhaustion, as depicted in Fig. 2 for Energy and known from the verbal report. The first measure during the race was done at km 30. Although most researchers reported an increase in cognitive performance after a short bout of exercise, we should keep in mind that both subjects had already been out on the course for about 4 hours before this first measure. Even this first section is a time interval exceeding the range of most other studies. Both subjects showed only marginal changes for the KLT-calculation test, but a remarkable decrease for the D2 (21% for Subject A and 12% for Subject B). While for Subject A the lactate levels were low during this time interval, for Subject B these values reached the maximum.

The enormous decrease for the D2 attentional task for Subject A, the older participant, might be explained a little bit more by the effect of heat compared to the total depletion of the glycogen reserves. Temperature was 48°C and 45°C for the test points at km 30 and 70, and the rating for Energy of Subject A was on the lowest level at km 70. From verbal report we know that Subject A was totally exhausted by the heat, not by the physical effort in completing this distance. With respect to this we stress that Subject A has a running history including about 150 marathons and 30 Ultramarathons, including a former participation in the Badwater Ultramarathon—so it seems very likely that he was able to organize the race in a way that he would not run out of his reserves.

Subject B, on the other hand, had the highest lactate levels at the distance of 30 to 70 km. Although these levels were below the threshold described by Chmura, et al. (1994) for an increased reaction time, the reserves might already have been depleted, because this high level of lactate was kept for a very long time interval. The data show that the lactate levels of Subject B were high at 30 and 70 km. However, we assume that the lactate values increased to this level right after the start and remained there until km 70, so Subject B had these high levels for more than 17 hours. The reason for this assumption is that all participants start at a relatively high pace, to complete the first marathon distance as fast as possible during the early, somewhat cooler, daytimes. Although lactate levels in this amount do not necessarily mean that reserves are depleted, it is clear that running for much longer than 4 hr. is in any case comparable to a type 3 exercise according to Tomporowski (2003), described as long lasting, low aerobic exercise. Compared to Subject A the changes in cognitive performance of Subject B were, in general, smaller during Day 1 and less sensitive to temperature; this might lead to the conclusion that for Subject B the depletion of the reserves, compared to temperature, might have been more relevant for the decrease in cognitive performance.

Between the two tests at km 70 and km 95 both subjects had a sleeping break (not longer than 2 hours) during the ascent to the Towness Pass (more
scores are the result of heavy blistering and the low Energy scores might be explained by the fact that Subject B had higher lactate values at the beginning of the race which might have led to a depletion of the energy stores within Day 1. For Subject B the EKV values remained low until the finish. The only remarkable change in Energy was a small increase, showing a little bit more Energy after the first sleeping period at about km 80. At km 195, where cognitive performance exhibited the poorest scores, those for Health and Energy were lowest (or equal to the lowest) too. In this case the weak cognitive results are easily to explain by the total exhaustion—obviously physical and psychological. In the last test session Subject B showed the lowest results for the D2 but the highest results for the KLT. Although the further reduction in the D2 can be explained by the still low scores of the EKV, the good result on the KLT is hard to explain. Possibly daytime might have influenced these results. At the checkpoint at km 195 it was 7 a.m. and the negative effects of sleep deprivation are more influential in the morning. With respect to daytime it would be interesting to evaluate the chronotype of the subjects. Natale, Martoni, and Cicogna (2003) demonstrated the differential effects of morning vs evening chronotype on cognitive and physical performance.

An additional factor for the partly increased performance data of both subjects during the last test sessions might be the motivational aspect of reaching the finish in a few minutes after this tremendous effort.

Taken together, we can conclude that cognitive performance varied differently between the subjects and was influenced, as expected by heat, exhaustion, and sleep deprivation. All in all after an initial decrease the performance increased again to almost baseline levels during Day 2, but dropped on Day 3, especially during the early morning. While during Day 1 the reduced cognitive performance of Subject A is more likely to be attributable to the heat, for Subject B these scores seem to be more an effect of exhaustion due to exercise.

The results suggest, well in line with other reports, that stress from heat and prolonged exercise leads to a considerable decrease in cognitive performance. Nevertheless in this study, a short nap of approximately one to two hours during the first night, rehybridation, and exposure to a cooler surrounding temperature helped re-establish high performance. Additional heat and exercise then led to a continuous 20% to 25% decrease in performance. The data, thus, showed that exposure to this extreme environment for only one day led to a decrease in performance that could be reversed, but longer lasting exposure yielded detrimental effects.

Nevertheless, we have to keep in mind that these are the results of two subjects, quite different in age and running experience. For a more precise
analysis of the cognitive changes during an extreme event like the Bachwater Ultramarathon, a larger sample of subjects and additional parameters like EEG for recording the drowsiness, and different blood parameters such as haematocrit for dehydration, blood glucose for nutrition, or cortisol as stress parameter, would be necessary.

REFERENCES


Accepted February 22, 2005.