COGNITIVE PERFORMANCE AFTER STRENUOUS PHYSICAL EXERCISE

Eef Högervorst, Wim Riedel, Asker Jeukendrup, Jelle Jolles

Department of Psychiatry and Neuropsychology
Department of Human Biology
Department of Psychiatry and Neuropsychology
University of Limburg, Maastricht

Summary.—Stimulating as well as detrimental effects of exercise on cognitive functioning have been reported. In the present study, 15 endurance-trained athletes (aged 18 to 42 years) performed a bicycle ergometer endurance test at 75% of their maximal work capacity (Wmax). Psychomotor and cognitive tests were administered before and immediately after exercise. These consisted of simple reaction time (RT), 3-choice RT and Stimulus-Response (S-R) incompatible RT tasks, a finger-tapping task, and the Stroop test. Simple RT tasks, but also the more complex S-R in compatible RT, and Color Word Interference in the Stroop test showed an increase in speed of performance after exercise relative to baseline. An enhanced activation was probably responsible for this better performance on psychomotor and cognitive tests. Since performance on the most complex task, the Interference subset of the Stroop, was especially improved after exercise, the expectancy of the subjects of a potential positive effect of exercise was thought to have been responsible.

Controversies exist over the effect of exercise on cognition. Several studies have described a negative effect of strenuous physical exercise on performance of cognitive tasks (Gutin, 1973; Isaacs & Pohlmant, 1991; Hancock & McNaughton, 1986; McMorris & Keen, 1994; Salmela & Ndoye, 1986). Fatigue has often been hypothesized to underlie the negative influence of intense exercise on cognition (Tomporowski, Ellis, & Stephens, 1987) but not all authors have found strenuous exercise had a negative effect on cognitive functions (Adam, Teeken, Ypelaar, & Verstappen, in press; Bard & Fleury, 1987; Côté, Salmela, & Papathanasopoulou, 1992; Fleury, Bard, Jobin, & Carriere, 1981; Paas & Adam, 1991; Tomporowski, et al., 1987). The various results are difficult to compare. Some studies tested subjects from different backgrounds and with different levels of fitness, and various cognitive tasks were used either during or after exercise of varying durations (McMorris & Keen, 1994; Tomporowski & Ellis, 1986). Often, exercise of a short duration (varying from 2 to 15 min.) was used and the amount of fatigue was not evaluated. Matters are further complicated since, as Tomporowski and Ellis (1986) stated, there is no agreement on the definition of fatigue.

According to Tatarkwa (1971), fatigue is defined as the combined out-

1Please address enquiries to Eef Högervorst, Department of Psychiatry and Neuropsychology, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands or e-mail (e.hogervorst@NFLaLimburg.NL).
put of mental activity and physiological functions. This definition was used by Hancock and McNaughton (1986), who found that fatiguing exercise (running on a treadmill at or above the anaerobic threshold) inhibited especially higher mental skills such as decision-making. Isaacs and Pohlmans (1991) also showed that exercise with a heavy workload (cycling at 100% of the maximal oxygen uptake \( \text{VO}_2\text{max} \) as compared to 75, 40, and 25% and rest) negatively affected performance on coincidence anticipation which task measures skilled timing performance. These studies supported the conclusion of Fleury, Bard, and Carriere (1981) that fatigue mainly affects complex cognitive functioning that requires a great deal of the resources of the central nervous system. However, Paas and Adam (1991) noted an unexpected beneficial effect of a substantial change in workload (cycling at 75-85% of the \( \text{VO}_2\text{max} \)) on a decisional task, while no effect was observed on a simpler perceptual task. It is thus unclear whether and how physical fatigue may affect simple and complex cognitive processing.

A possible flaw of former studies is the use of exercise of short duration (<15 min.) to induce fatigue. Thus, in the present study, subjects were tested before and immediately after prolonged (± 60 min.) endurance exercise. Endurance tests are thought to induce exhaustion or physical fatigue (Jeukendrup, Saris, Brouns, & Kester, 1996) which is then implicated in the decline of performance that occurs on any prolonged or repeated task (Kennedy, 1988). Subjects were not tested during exercise since earlier McMorris and Keen (1994) and Isaacs and Pohlmans (1991) hypothesized that divided attentional mechanisms rather than actual effects of fatigue caused the decreased performance which was seen during cycling on cognitive tasks. With higher arousal, subjects may focus rather on internal perceptions of discomfort (an hypothesis first described by Nideffer, 1979) than focus on cognitive tasks. The present study is similar to the study by Fleury, et al. (1981) in that fatigue is defined as the consequence of a physical workload on subsequent psychological performance. If fatigue affects the functioning of the central nervous system, fatiguing exercise should especially affect complex, decisional, cognitive tasks (Easterbrook, 1959; Gutin, 1973; Kennedy, 1988; Fleury, et al., 1981). Because subjects can compensate for a lack of energy by investing additional effort on a psychological task (Zijlstra & Meijman, 1989), the effort invested in task performance was taken into account as a modulating variable. Simple measures of speed (tapping and simple reaction time) were also taken into account. Performance on these tasks often improves after or during exercise, possibly due to enhanced activation (Easterbrook, 1959; Gutin, 1973; Salmela & Ndoye, 1986). Heart rate was measured as an indication of the intensity of performance and arousal or activation at the time of testing. We tested the hypotheses that (1) subjects would perform faster on simple tasks after an endurance test than before the test.
and (2) subjects would perform worse on complex cognitive tasks after an endurance test.

**Method**

**Subjects**

Fifteen healthy male triathletes and competitive cyclists participated. The subjects had enrolled in a validation study of endurance performance (Jeukendrup, et al., 1996). They trained on a regular basis (≥2 hr./day and ≥4 times/week). The mean age of the subjects was 24.9 yr. (SD = 7.9; range = 18 to 42), their mean weight was 73.5 kg (SD = 6.5), and their mean height was 183 cm (SD = 4.5).

**Design and Procedure**

The experiment was conducted according to a single-factor repeated-measures design. The factor “Level of exercise” consisted of three levels (pre-exercise, postexercise, baseline). Subjects came to the laboratory three times at 7-day intervals. The first visit was intended to familiarize the subjects with the tests and procedures. On the second visit, heart rate and cognitive tests were measured before and after the endurance test. Baseline measurements were taken on the third visit.

One week before the experiment, maximal workload was attained on a bicycle ergometer during an incremental exercise test until exhaustion. After a warm-up period of 5 min. at 100 W, the workload was increased every 2.5 min. until heart rate reached 160 beats/min. Then workload was increased with 25 W every 2.5 min. until the pedaling rate dropped below 60 rpm. The maximal work capacity ($W_{\text{max}}$) was determined by the following formula: $W_{\text{max}} = W_{\text{out}} + (t/150) \cdot 25$, where $W_{\text{out}}$ is the workload of the last completed step, $t$ is the time in the final step. Mean maximal workload of the subjects was 385 W (33), indicating that subjects were well-trained.

The exercise consisted of a short warm-up (5 min, 100 W) followed by a simulated time trial. In this time trial subjects were asked to perform a certain amount of work (equal to about 1 hour of cycling) as fast as possible. This total amount of work was based on the maximal workload ($W_{\text{max}}$) according to the formula:

$$\text{Total amount of work} = 0.75 \cdot W_{\text{max}} \cdot 3600.$$  

The ergometer was set in the linear mode according to the formula

$$W = L \cdot (\text{RPM})^2$$

in which RPM is the pedaling rate and $L$ is a linear factor. This factor was chosen in a way which would cause a pedaling rate of 90 RPM at 70% of the maximal work capacity. In other words, the linear factor was dependent
on a subject's maximal work capacity. This would mean that 75% of the maximal work capacity could be achieved at about 100 RPM which appeared to be the preferential pedaling rate of most of the cyclists. This type of exercise can be classified as 'maximal' or 'exhaustive' and subjects usually report maximal values on Ratings of Perceived Exertion (Salmela & Ndoye, 1986).

Apparatus

During the experiment, the subjects cycled on a Lode electronically braked ergometer (Lode Excaliber Sport®, Lode BV, Groningen, The Netherlands). Heart rate was measured electronically using a Dinamap® (Critikon, Type 8100) before and immediately after exercise.

Psychomotor and Cognitive Tasks

Subjects completed a short test battery (15 min.) which included complex cognitive and psychomotor tests. The following tests were included.

The Stroop Color-Word test.—This is a well-known test for the ease of shifting perceptual sets to conform to changing task requirements (Lezak, 1983). The test consists typically of three subtests (Bohnen, Jolles, & Twijnstra, 1992). In Subtest I, 10 rows by 10 columns of color names (red, blue, green, and yellow) are printed in black on white cardboard. In Subtest II, the same number of correspondingly colored patches are printed, whereas Subtest III contains a number of color names, printed in incongruously colored ink. For instance, the word “red” can be printed in green. The protocol for administering the test in this study was as follows. For Subtest I (‘Color Word Reading’), the subject was requested to “read the color names row by row, as fast as you can, without making any mistakes.” The time needed to complete the whole card was recorded with a stopwatch. For the second subtest (‘Color Naming’) the instruction was to “name the colored patches.” The third subtest (‘Color-Word Interference’) involved naming the color of the ink in which the names of colors were printed without paying attention to the word itself. For each subtest the time taken to finish the card was recorded. In this study the short version was used. The short version stops after 40 stimuli (whereas the whole test contains 100 stimuli). The correlation between the abridged version and the original version is remarkably high (Pearson \( r = .93, p < .001 \) on average for all cards) (Klein, Ponds, Houx, & Jolles, in press).

The Choice Reaction Time test.—The subjects pressed one button and were asked to press one of five other buttons located equidistant from the hold button when they onset. This yielded reaction times in three subsequent conditions of increasing task complexity. The response set consisted of pressing only one button that lit up (simple RT), pressing one of three buttons that lit up (3-choice RT), or pressing the button to the right of the
lighted button (incompatible 3-choice RT) as quickly as possible. Hence, on these three subtasks, two task factors are systematically varied within subjects: the number of response alternatives, and the compatibility of the stimulus and the required response (Houx, Vreeling, & Jolles, 1991).

The Finger Tapping Test.—Finger-tapping speed is one of the more simple aspects of psychomotor performance. The time elapsed between two single taps was registered with an IBM (Type 486) compatible computer at millisecond accuracy. In this experiment the number of taps per second (on a response button on a panel connected to the computer), using the preferred hand, was used as a dependent variable (Brand & Jolles, 1988).

Effort.—Effort was measured using the Rating Scale ('Beoordelings-
Schaal') for subjectively experienced Mental Effort ('Mentale Inspanning'). Apart from a global indication of the psychophysiological state of the subjects (fatigue), the amount of effort can also be seen as an indicator of the "cost" of performing a task. This scale was constructed by using the 'magnitude-estimation method' with several different population samples (bus drivers and students) and different rating modalities to estimate the scale values of the nine adjectives (not at all exhausting, hardly exhausting, slightly exhausting to extremely exhausting) (Zijlstra & Meijman, 1989). The data from the original 15-cm visual analog scale were transformed to percentages.

Statistical Analyses

The "Level of exercise" effect was tested for each task according to a repeated-measures design with three levels: low (preexercise), high (postexercise), and low (baseline). The main effect of "Level of exercise" was evaluated in conjunction with task factors in repeated-measures, within-subjects analysis of variance. The task factors were number of response alternatives (2 levels: 1 or 3 responses), and S-R Compatibility (2 levels: Compatible, Incompatible) in the Motor Choice Reaction Task, and Color Naming (2 levels: Words, Colors), and Color-Word Interference (2 levels: Response Conflict, No Response Conflict) in the Stroop task. Post hoc paired t tests were used to assess the differences between the pairs of experimental conditions. Probabilities of .05 or less are reported as significant. Nonsignificant probabilities are reported when they are between .05 and .10. All analyses were performed with SPSS 4.0 on an Apple MacIntosh computer.

Results

Heart rate was significantly different over conditions as multivariate analyses showed ($F_{2,11} = 36.65, p < .001$). Heart rate was significantly lower before exercise ($t_{11} = -8.72, p < .001$) and baseline ($t_{11} = 8.71, p < .001$) relative to that after exercise (see Table 1).

The subjects took significantly longer to complete the Stroop subtask Color Naming than they did the subtask Color Word Reading. Multivariate
analysis indicated a main effect of stimulus type (Words, Colors) \( (F_{1,11} = 57.88, p < .001) \). Also, a main effect of the factor “Level of exercise” \( (F_{2,11} = 15.97, p < .001) \) and a significant interaction between stimulus type and “Level of exercise” were detected \( (F_{2,11} = 5.78, p = .02) \). Post hoc paired \( t \) tests between conditions showed learning effects on both cards. Time needed to perform Color Word Reading was significantly longer before than after exercise \( (t_{14} = 2.36, p < .05) \) and at baseline \( (t_{14} = 2.74, p < .05) \) and equal after exercise to that at baseline \( (t_{14} = -.48, \text{ns}) \), signifying a bottom effect. On Color Naming the learning effect was even more clearly shown, since significantly more time was needed to perform this task before exercise than afterwards \( (t_{14} = 3.17, p < .01) \) and baseline \( (t_{14} = 4.95, p < .001) \) and after exercise relative to baseline \( (t_{14} = 2.79, p = .01) \). Response time on the Color-Word Interference subtask was significantly longer than that on the Color Naming subtask since multivariate analysis showed that a main effect of stimulus type (Response Conflict, No Response Conflict) was significant \( (F_{1,11} = 52.79, p < .001) \). Also, the “Level of exercise” effect \( (F_{2,11} = 7.14, p < .01) \) and the interaction of stimulus type and “Level of exercise” were significant \( (F_{2,11} = 10.34, p < .01) \). As can be seen in Table 1, exercise clearly affected Color-Word Interference. This was confirmed with post hoc paired \( t \) tests. Time needed to perform the Color-Word Interference was longer before exercise than afterwards \( (t_{14} = 3.42, p < .005) \) but equal to baseline \( (t_{14} = 1.58, p > .10) \). The baseline performance was significantly longer than the time needed to perform the task after exercise \( (t_{14} = -3.56, p < .005) \).

On the Motor Choice Reaction Time test, Choice RTs were longer than simple RTs. Multivariate analysis yielded a main effect of the number of response alternatives (2 levels: 1 or 3 responses) \( (F_{1,11} = 113.36, p < .001) \). A main effect of the factor “Level of exercise” was seen on multivariate analysis \( (F_{2,12} = 6.28, p = .01) \). Exercise did not affect the relation between the number of response alternatives and RT, since no interaction was seen between the number of response alternatives and “Level of exercise” \( (F_{2,12} = .71, \text{ns}) \). Post hoc paired \( t \) tests indicated that a lower mean response time on the simple reaction time task was seen after relative to before exercise \( (t_{14} = 1.99, p = .07) \). No difference was seen between the preexercise condition and baseline \( (t_{14} = -.17, \text{ns}) \) but response time after exercise was significantly lower than at baseline \( (t_{14} = -2.81, p = .01) \). Choice reaction time showed no effect of condition; response times were equal across conditions. Incompatible RTs took longer than compatible RTs. Multivariate analysis indicated a main effect of “S-R Compatibility” (2 levels: Compatible, Incompatible) \( (F_{1,11} = 57.32, p < .01) \). “Level of exercise” was nonsignificant \( (F_{2,12} = 2.87, p =

\footnote{Data of one of the subjects was missing due to apparatus failure.}
Table 1

Means and Standard Deviations of HR (beats/min.), Stroop (Time to Finish, Cards/sec.), Motor Choice Reaction Time: Reaction Time (msec.), Simple RT, Choice RT, and Incompatible RT; Finger Tapping Test (Number Taps/sec.), and Effort (0-100 Scale)* Preexercise, Postexercise, and at Baseline.

<table>
<thead>
<tr>
<th>Task</th>
<th>Preexercise</th>
<th>Postexercise</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Heart Rate, bpm</td>
<td>63.1</td>
<td>15.3</td>
<td>203.0</td>
</tr>
<tr>
<td>Stroop Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color Reading</td>
<td>13.3</td>
<td>2.0</td>
<td>12.6</td>
</tr>
<tr>
<td>Color Naming</td>
<td>17.6</td>
<td>3.5</td>
<td>16.8</td>
</tr>
<tr>
<td>Color Word Interference</td>
<td>23.3</td>
<td>6.3</td>
<td>20.8</td>
</tr>
<tr>
<td>Motor Choice Simple RT</td>
<td>378</td>
<td>37</td>
<td>369</td>
</tr>
<tr>
<td>Choice RT</td>
<td>429</td>
<td>53</td>
<td>417</td>
</tr>
<tr>
<td>Incompatible RT</td>
<td>494</td>
<td>61</td>
<td>485</td>
</tr>
<tr>
<td>Finger-tapping Test</td>
<td>7.0</td>
<td>5</td>
<td>6.9</td>
</tr>
<tr>
<td>Effort Score</td>
<td>53.1</td>
<td>3.9</td>
<td>54.5</td>
</tr>
</tbody>
</table>

*0 = not exhausting, 100 = extremely exhausting.

.09), indicating no effect of exercise in the Incompatible task. The interaction between "S-R Compatibility" and "Level of exercise" was also not significant ($F_{2,61} = .05$, ns). On the incompatible subtask, lower response time after exercise than before ($t_{19} = 1.91, p = .08$) was not significant and no differences were seen when pre- and postexercise conditions were compared with baseline conditions ($p > .50$). The average number of taps per second was 6.97 and did not differ significantly between conditions ($F_{2,61} = .18$, ns). The average subjectively experienced effort was 52.90 on a scale of 100 and did not significantly differ between conditions ($F_{2,61} = .17$, ns).

Discussion

The results of the experiment support the first hypothesis, namely, that exercise has a positive effect on performance speed in simple tasks. Simple RTs were significantly lower after fatiguing exercise than at baseline. However, simple manual motor speed which is mainly limited by peripheral factors (Tapping) was not influenced at all by exercise of high intensity. Further, the results did not support the second hypothesis, that endurance exercise would have a general detrimental effect on complex cognitive functions. In fact, the opposite was found. For instance, the RT for the most complex task that required the inhibition of a learned response, the Color-Word Interference subtask of the Stroop test, was decreased by endurance exercise. Important is that the subjectively experienced effort was not increased, and no more errors were made with the higher speed of performance.

The findings seem to indicate that there is no clear-cut relation between physical fatigue immediately after endurance exercise and mental fatigue as
was suggested by Tatakuwa (1971). However, it is possible that the subjects were not thoroughly fatigued, and it may be worthwhile to study fatigue after an extended period of endurance exercise.

The enhanced performance on psychomotor and cognitive tasks can be explained in the following fashion. Gutin (1973) proposed that various amounts of activation differently affect certain types of mental tasks. Low activation is best for performance of tasks which require a great deal of information processing. Low activation is reflected by a slight increase in heart rate during and after exercise (90 to 120 bpm). Higher activation, provoked by exercise of long duration (> ± 15 min., inducing a heart rate of more than 160 bpm), should diminish performance on tasks that require a great deal of information processing. In the study of Isaacs and Pohlmam (1991) in which exercise had a negative effect on complex tasks, heart rate was 150 to 175 bpm. Although we found that exercise of longer duration at heart rates of 160 to 180 bpm had a beneficial effect on performance of complex cognitive tasks, these effects were seen after exercise, which was contrary to Gutin's expectations. However, during the actual testing activation was still sufficiently enhanced (heart rate 90–120 bpm) to lead to an improved performance consistent with the first hypothesis. Thus, at the time of psychomotor and cognitive testing, the subjects were still in a highly aroused state. Hence, the enhanced performance on the tasks after exercise might reflect a generally increased activation of the central nervous system. The absence of this effect, as reflected by the means on the Tapping task, could reflect a ceiling effect.

The finding of a significant interaction between exercise and the results of the most complex task, Stroop Color-Word Interference, is interesting. This effect may be comparable to that found by Paas and Adam (1991) who also reported an unexpected beneficial effect on a decisional task during a substantial change in workload (cycling at 75-85% of the VO2 max). The authors concluded that subjects in their study apparently had decided to invest more processing capacity or resources into the more complex (decision) task. Similar findings were reported by Adam, et al. (in press). Although effort in our study was not elevated after exercise, it is possible that subjective scales of effort are not reliable or valid indicators of the actual effort invested. An unconscious decision made by the subjects actually to invest more effort in certain tasks is influenced by motivation and probably the effects of expectancy. Tomporowski and Ellis (1986) previously suggested that many studies may have been confounded by expectancy effects. The expectancy of the participants influences performance through motivational variables. Hence, the subjects may have compensated to overcome the possible effects of fatigue. This compensation effect is especially seen in highly trained subjects and is not mediated by fitness but rather, as mentioned by Tomporowski
and Ellis (1986), by the expectancy that exercise has a beneficial effect in general. Subjects who do not often engage in exercise of high intensity have a negative expectancy about the effect of exercise on cognition. For instance, Delignières, Brisswalter, and Legros (1994) showed that experts exerted additional effort to overcome the effects of fatigue when challenged. Their performance increased as exertion increased (cycling at 20, 40, 60, and 80% of maximal aerobic power). Nonexperts of similar fitness showed a decrease in performance on their 2- and 4-choice RT task. The error rate in both groups remained stable.

It is thus unclear whether the effects seen in the present study are mediated by the effects of activation or expectancy; however, it is clear that expectancy variables and valid measures of effort should be taken into account in studies into the effects of fatigue. Preferably, expectancy should be controlled while effort is measured in an experimental design to study the actual relations among fatigue, expectancy, and effort.

REFERENCES


Accepted July 17, 1996.