

# Beep Test Performance Is Influenced by 30 Minutes of Cognitive Work

CLARE MACMAHON<sup>1</sup>, ZACKARY HAWKINS<sup>2</sup>, and LINDA SCHÜCKER<sup>3</sup>

<sup>1</sup>Department of Dietetics, Nutrition, and Sport, Latrobe University, Melbourne, AUSTRALIA; <sup>2</sup>Department of Health and Medical Sciences, Swinburne University of Technology, Melbourne, AUSTRALIA; and <sup>3</sup>Institute of Sport and Exercise Sciences, University of Münster, Münster, GERMANY

## ABSTRACT

MACMAHON, C., Z. HAWKINS, and L. SCHÜCKER. Beep Test Performance Is Influenced by 30 Minutes of Cognitive Work. *Med. Sci. Sports Exerc.*, Vol. 51, No. 9, pp. 1928–1934, 2019. **Purpose:** This study explored conflicting findings in the literature on the influence of perceived cognitive fatigue on physical performance by testing the effect of the Stroop task (high cognitive load) on an intermittent running test (beep test). **Methods:** In a within-subjects repeated-measures experiment, 13 active athletes performed the beep test on two occasions, in a randomized, counterbalanced order. In each session, a preceding cognitive task was completed for 30 min, with the incongruent Stroop task in the high load condition, and the congruent Stroop task in the low load condition. Perceived cognitive fatigue was measured before testing (baseline) and at 10, 20, and 30 min of the cognitive load manipulation. Perceived effort on the cognitive task and general motivation for the physical task (beep test) were measured before the beep test, and motivation-related perception of the beep test and ratings of perceived exertion were measured after completion of the test. Heart rate and beep test performance (completion stage and time) were also recorded. **Results:** The incongruent Stroop task was perceived as more fatiguing and effortful. Participants also withdrew from the beep test significantly earlier in the high load condition ( $M = 8:48$  min,  $SD = 2:32$  min) compared to the low load condition ( $M = 9:20$  min,  $SD = 2:28$  min),  $F(1,11) = 21.76$ ,  $P < 0.01$ ,  $\eta^2 = 0.67$ . There were no differences in heart rate or general motivation between the two conditions. **Conclusions:** Whereas previous research shows that active athletes can maintain performance on the beep test after 10 min of the incongruent Stroop task, this study shows that performance is impaired after 30 min. Variables in need of exploration in future investigations include experience with both the physical and cognitive task. **Key Words:** COGNITIVE FATIGUE, INTERMITTENT RUNNING, SELF-REGULATION

An expanding area of research is devoted to understanding the effects of cognitive and self-control tasks on physical performance. Indeed, the growth of interest in this topic is evidenced by a number of recent review articles (e.g., [1–4]) and a boom in the empirical work from a number of different subdisciplines (e.g., [5,6]). Concomitant with the increasing amount of research in this topic are some inconsistencies, as researchers look to understand the complex interplay of factors, including the specific cognitive task used, the target physical task, and individual difference

variables, such as training background and motivation (e.g., [7]). In particular, this study sought to explore the influence of the length of the manipulation task, in follow-up from a previous study (8).

Support for the finding that cognitive exertion has a negative impact on physical performance is exemplified in two key studies. Marcora et al. (9) tested performance in a cycling time trial after a 90-min cognitive task compared with a control condition in which participants watched a 90-min documentary. They found that participants withdrew from the physical task significantly earlier in the exertion condition compared with the control condition. Similarly, MacMahon et al. (10) showed that participants in a self-paced running task (3-km time trial) ran significantly slower in the exertion condition compared with the control condition.

The inconsistent effects in this area are exemplified by Schücker and MacMahon (8), who found that active participants were not impeded in running the beep test, an externally paced 20-m shuttle run (11), after completing a demanding cognitive task. One key factor is that the length of the cognitive task is inconsistent throughout the literature in this area and has ranged from 3 min (12) to 90 min (9,10), with different results depending on the study's design. For example, Brown and Bray (6) found a threshold effect for the Stroop task; after 4–6 min of the Stroop task, there were negative effects on a

Address for correspondence: Clare MacMahon, B.A., M.A., Ph.D., Health Sciences Building 3, Room 215, La Trobe University, Melbourne 3086, VIC, Australia; E-mail: c.macmahon@latrobe.edu.au.

Submitted for publication January 2019.

Accepted for publication March 2019.

0195-9131/19/5109-1928/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Copyright © 2019 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American College of Sports Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1249/MSS.0000000000001982

handgrip endurance task. In comparison, Schücker and MacMahon (8) found that 10 min of the Stroop task did not have a negative effect on beep test performance. In line with this finding, Van Cutsem et al. (2) suggest that for complex, full body endurance tasks, the cognitive task must have a minimum duration of 30 min to influence physical performance. This suggestion is further supported by Smith et al. (13), who used a physical task similar to the beep test, the Yo-Yo intermittent running task. Smith et al. (13) found that the participants' performance declined after 30 min of the Stroop task, compared with the performance after 30 min of leisurely reading of emotionally neutral magazines.

It is thus notable that the research shows that not only does the Stroop task increase perceived fatigue and impair subsequent full body physical performance, it can do so after only a few minutes of the task (12,14–16). Thus, the main goal of this study was to retest the manipulations of Schücker and MacMahon (8) using a 30-min Stroop task instead of a 10-min Stroop task. In addition, the key feature of this study was the intentional sampling of actively competing recreational athletes. This sampling allowed us to address that comparison studies incorporating shorter fatiguing tasks have found an effect on physical performance, but have not focused particularly on the experience of the participants, who were often defined as inactive/inexperienced (12,16), untrained (6), or at best, sport students (14). Intentionally sampling actively competitive participants allowed us to isolate the effect of the task length on the original results of Schücker and MacMahon (8).

Related to performance level, the secondary aim was to examine motivation of participants. Specifically, Inzlicht et al. (17) argue that acts of cognitive exertion or self-control lead to a refractory or recovery period; people prefer subsequent activities that they deem more immediately enjoyable or gratifying (intrinsically motivating, “want to” tasks) to activities they feel they ought to do due to external pressures (extrinsically motivating “have to” tasks). Conceptually, then, even when a subsequent task may be relatively intrinsically motivating for a participant, like the case of a competitive physical test like the beep test, a preceding self-regulation task will lower preference for engaging in this task due to cognitive exertion. This leads to the prediction that when participants have high perceived cognitive fatigue, they will report greater levels of feeling they “had to” engage in the task and lower levels of feeling they “wanted to,” compared with when they have lower perceived levels of cognitive fatigue.

In this study, we hypothesized that participants would perceive the incongruent 30-min Stroop task to be more cognitively fatiguing, and that beep test performance would be worse in this condition, and perceived as less intrinsically motivated.

## METHODS

**Participants.** Using a repeated-measures ANOVA statistical test in a G\*Power calculation based on effects in previous research (e.g.,  $\eta^2 = 0.31$ , see 10), a sample size of 10 was estimated to have 95% power. Fourteen participants ( $n = 11$  male,

$n = 3$  female) were recruited for the study; however, one male failed to complete the second testing session, and was thus removed from further analyses, for a total of 13 participants. The average age of the participants was 19.92 yr ( $SD = 1.75$ ). The training load of the group was an average of three sessions per week ( $M = 3.55$ ;  $SD = 1.44$ ) with an average of  $6.72 \text{ h} \cdot \text{wk}^{-1}$  ( $SD = 4.03$ ), and all competed in registered recreational leagues or higher. Because the range of weekly training hours per week was 3 to 15, and two of the participants were recreational athletes who played sport and competed twice a week, but did not train, the group as a whole was conservatively classified as recreational (active) athletes (see (18,19), for further discussion of participant classification). On average, participants were involved in their sport for 10.8 yr ( $SD = 4.48$ ) and were all taking part in a competition. Most of the samples reported participation in a team sport ( $n = 11$ ), with the exception of two individual athletes (endurance and esthetic/form-based sports).

Before participation, athletes were required to fill out a medical questionnaire. To eliminate any health or injury risks, only healthy athletes participated. Participation in this study was encouraged by a US \$40 gift voucher given to participants upon completion of their second testing session. After application, the study was approved by the Swinburne University Human Research Ethics committee, and all athletes provided informed consent before participation.

**Procedure.** The study was a repeated-measures within-subjects design, with participants tested on two separate occasions each lasting approximately 70 min. The same researcher was used for all testing, but was not blind to the condition being tested. During both sessions, participants worked on a cognitive task (high load or low load) for 30 min followed by the physical performance task. The order of sessions (high load or low load) was counterbalanced as much as possible (seven in high load–low load; six in low load–high load, accounting for the eliminated participant who did not complete the second testing session). Before reporting to the laboratory for testing, participants received instructions about exercise, sleep, and food intake. Specifically, participants were asked to avoid hard training the day before testing, sleep for at least 7 h the night before, and to eat a meal no closer than 2 h before testing. Participants were also asked to refrain from caffeine and nicotine consumption at least 3 h before testing. Efforts were made to test participants at the same time of day in each session. For the majority of participants, the testing time of day differed by no more than 60 min (e.g., time 1 at 10:00 AM, time 2 at 11 AM). Two participants had a time of day difference of 90 min, and two of over 4 h (e.g., time 1, 10:30 AM; time 2, 2:45 PM). Most testing took place in sessions between 10:00 AM and 12:00 PM for both sessions. Two participants were tested after 4:00 PM for both sessions, with one of these in each condition order. After the demographics were completed and the study procedure was explained, participants were fitted with a heart rate monitor. They then began the 30-min cognitive task, followed by the physical performance task. The participants answered a number of questionnaires (see below). After the second testing

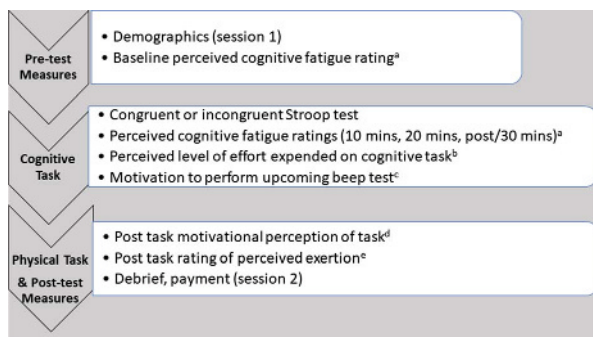


FIGURE 1—Testing procedure. <sup>a</sup>On a scale from 1 (not at all) to 7 (very much). <sup>b</sup>On a scale from 1 (no effort) to 7 (a lot of effort). <sup>c</sup>Items: “I am concerned about not doing as well as I can in the beep test”; “I want to be successful in the beep test”; “I am going to give my best in the beep test,” rated on a scale from 1 (no agreement) to 7 (full agreement). <sup>d</sup>“How much did you feel like you *had* to complete the beep test”; “How much did you feel like you *wanted* to complete the beep test,” rated from 1 (not at all) to 10 (very much so). <sup>e</sup>Rated from 6 (no effort at all) to 20 (maximum effort).

session, participants were debriefed, given a US \$40 gift voucher and were informed of the study aims and hypotheses.

**Cognitive task.** In both conditions, participants completed the Stroop task, in which the names of the colors red, blue, yellow, and green were presented in colored font. Participants were required to indicate the color of the font in which the word was presented. In the congruent condition (low load condition), the font color and name of the word was the same. In the incongruent condition (high load condition), the font color and name of the word was different. There was also an exception rule in the incongruent condition; if the word was presented in the color red, participants were to indicate the name of the word not the color in which it was presented. Participants worked on the task on a computer for 30 min per condition and were required to press labeled keys on the keypad as quickly as possible.

**Physical performance task.** The physical performance task was the 20-m shuttle run known as the beep test, which is an externally paced running fitness test (11). Participants run back and forth between two lines 20 m apart, reaching each line in synch with the corresponding beep. The interval between beeps decreases steadily as the task progresses so that running speed increases accordingly. In this test, athletes can miss the line on one shuttle, if they are able to catch up on the very next shuttle. If two shuttles in a row are missed, however, athletes are eliminated from the test. Athletes can thus withdraw themselves from the test when fatigued, or after missing the line once, ceasing to attempt to reach the line on the next beep, or they can be eliminated by the experimenter after attempting to reach the line but missing two beeps in succession. Both of these methods of test elimination identify the point after which two shuttles in a row were incomplete (e.g., level 8.2), as the standard method of administering the test. The time point of test elimination was then used as the performance score in minutes and seconds.

**Questionnaires.** As depicted in Figure 1, and detailed in Table 1, a number of measures were taken before, during, and after the cognitive task, and before and after the physical

task. In session 1, participants provided background information on their sporting experience related to sport and playing position, level of competition, and frequency of competition and training. Perceived level of cognitive fatigue (“How cognitively fatigued do you currently feel?”) was recorded on a seven-point scale ranging from 1 (not at all) to 7 (very much). This was collected before the cognitive task (as a baseline), then at 10, 20, and 30 min (post) of the cognitive task. These measures were presented to participants and collected on a laptop using Inquisit software. For the 10- and 20-min assessments during the Stroop task, the question appeared on the screen, interrupting the Stroop task briefly. Participants clicked on the appropriate scale number, which they were familiar with at that point, and then returned to the Stroop task. The 30-min assessment used the question, “How cognitively fatigued do you now feel after completing the troop task?”

After the cognitive task and before commencement of the beep test, a two-item questionnaire, again presented on a laptop through Inquisit, was used to measure perceived level of effort expended on the cognitive task (“How much effort do you perceive you expended on the cognitive task?” with 1, no effort; 7, a lot of effort). In addition, motivation to perform the upcoming physical task was collected on a seven-point scale ranging from 1 (no agreement) to 7 (full agreement), for three statements, as used in the previous study (8). The specific statements can be seen in Figure 1. After completing the beep test, additional items were completed on the laptop. Specifically, a two-item questionnaire was used to assess motivation: (a) the degree to which participants felt like they *had* to complete the beep test and (b) the degree to which they felt they *wanted* to complete the beep test (1, not at all; 10, very much so), see Figure 1 for wording). Finally, a 15-point scale ranging from 6 (no effort at all) to 20 (maximum effort) was used to assess RPE (20) after the beep test.

**Statistical analysis.** All analyses were conducted with IBM SPSS statistics 25. Before the main analyses, data were screened for aberrant values using two standard deviations above or below the mean and tested against assumptions of normality. Assumptions were met, and no outliers were identified. **Repeated-measures ANOVA were used to test the differences in performance and related variables between the two conditions.** Degrees of freedom were Greenhouse–Geisser

TABLE 1. Physical performance and motivation variables for the beep test compared by condition, showing no statistically significant differences.

Variables	Fatigue, M (SD)	Control, M (SD)	F(1,12)	P
Average HR during beep test	155.31 (20.34)	146.23 (17.79)	2.93	0.113
Maximum HR during beep test	193.00 (6.89)	186.77 (15.42)	1.75	0.210
Final beep test RPE	17.69 (1.25)	17.85 (0.80)	0.270	0.613
Pretest motivation 1 <sup>a</sup>	4.85 (1.41)	4.31 (1.55)	1.55	0.237
Pretest motivation 2 <sup>b</sup>	6.31 (0.95)	6.15 (1.07)	0.480	0.502
Pretest motivation 3 <sup>c</sup>	6.85 (0.38)	6.69 (0.63)	1.00	0.337
“Have-to” motivation <sup>d</sup>	7.31 (2.87)	8.00 (1.96)	2.03	0.179
“Want-to” motivation <sup>e</sup>	6.54 (2.73)	6.77 (2.17)	0.118	0.737

<sup>a</sup>“I am concerned about not doing well.”

<sup>b</sup>“I want to be successful.”

<sup>c</sup>“I am going to give my best.”

<sup>d</sup>“How much did you feel like you *had* to complete the beep test?”

<sup>e</sup>“How much did you feel like you *wanted* to complete the beep test?”

TABLE 2. Means and standard deviations comparing manipulation check variables by condition.

Variables	Fatigue, M (SD)	Control, M (SD)	F (1,12)	Partial $\eta^2$
Mean reaction time (ms)*	1225.69 (244.21)	817.28 (107.06)	51.81	0.812
Percentage correct responses*	87.6 (0.05)	98.56 (0.12)	118.00	0.908
Effort in cognitive task*	4.54 (1.27)	2.54 (1.56)	24.00	0.667
Cognitive fatigue: 10 min*	3.38 (0.96)	2.38 (1.12)	6.50	0.351
Cognitive fatigue: 20 min*	4.62 (1.04)	2.85 (1.14)	34.13	0.740
Cognitive fatigue: 30 min*	5.31 (1.25)	3.31 (1.43)	20.80	0.634

\*Significant difference in variables by condition at  $P < 0.05$ .

adjusted when sphericity assumptions were violated. A significance level of  $P$  less than 0.05 was set for all tests.

## RESULTS

**Manipulation check.** Analyses of manipulation check data show that there were no differences in baseline reported cognitive fatigue levels for the two testing sessions,  $t(12) = 0.822$ ,  $P = 0.427$ . However, a  $2 \times 3$  (condition, time) ANOVA compared fatigue ratings associated with performing the Stroop test, showing a main effect of condition,  $F(1,12) = 22.41$ ,  $P < 0.01$ , partial  $\eta^2 = 0.651$ , a main effect of time,  $F(2,24) = 31.89$ ,  $P < 0.01$ , partial  $\eta^2 = 0.727$ , and an interaction,  $F(2,24) = 5.51$ ,  $P < 0.05$ , partial  $\eta^2 = 0.314$ . Specifically, participants reported significantly increasing cognitive fatigue over time, particularly for the fatigue condition. Notably, follow-up paired samples  $t$ -tests showed that participants reported a statistically significantly higher level of subjective cognitive fatigue after the 30-min Stroop task in the high load condition compared with the 30-min Stroop task in the low load condition ( $t(12) = -4.56$ ,  $P < 0.01$ ). Data comparing the two conditions are presented in Table 2, with the main effects and interaction for ratings illustrated in Figure 2. Further analyses of Stroop performance data were used to compare the three time points (10, 20, and 30 min). These analyses showed

main effects of condition for latency ( $F(1,12) = 31.26$ ,  $P < 0.001$ , partial  $\eta^2 = 0.72$ ), with longer latencies in the incongruent condition, as well as percent correct ( $F(1,12) = 117.91$ ,  $P < 0.001$ , partial  $\eta^2 = 0.91$ ), with fewer correct answers in the incongruent condition. However, there was no effect of time, and no interaction of time and condition for both latency and percent correct (all  $P$ 's  $> 0.05$ ), showing that performance was stable throughout the 30 min. These analyses provide confidence that participants perceived themselves to be more fatigued in the high load condition compared with the low load condition.

**Performance.** The exact time participants stopped running (seconds) was recorded as the performance score for each condition. Participants withdrew from the beep test statistically significantly earlier in the high load condition ( $M = 8:48$  min,  $SD = 2:32$  min) compared with the low load condition ( $M = 9:20$  min,  $SD = 2:28$  min;  $F[1,11] = 21.76$ ,  $P < 0.01$ ,  $\eta^2 = 0.67$ ). These times translate to a difference of one half beep test level between the two conditions (level 8.5, beep 2 for the high load condition vs level 9, beep 4 for the low load condition). Table 1 shows that average heart rate and maximum heart rate during the beep test did not differ for the low or high load conditions, and that similarly high RPE were found in both conditions.

**Task perception and motivation.** The perceived level of effort invested in the cognitive task was explored using a two-condition (high load, low load) ANOVA. A significant effect of condition showed statistically higher ratings for the high load condition compared with the low load condition ( $F[1,12] = 24.00$ ,  $P < 0.001$ , partial  $\eta^2 = 0.667$ ). Table 1 shows additional motivation data related to completion of the cognitive tasks and perception of the physical task. There were no statistically significant differences in the levels of agreement for the three motivation statements as compared by

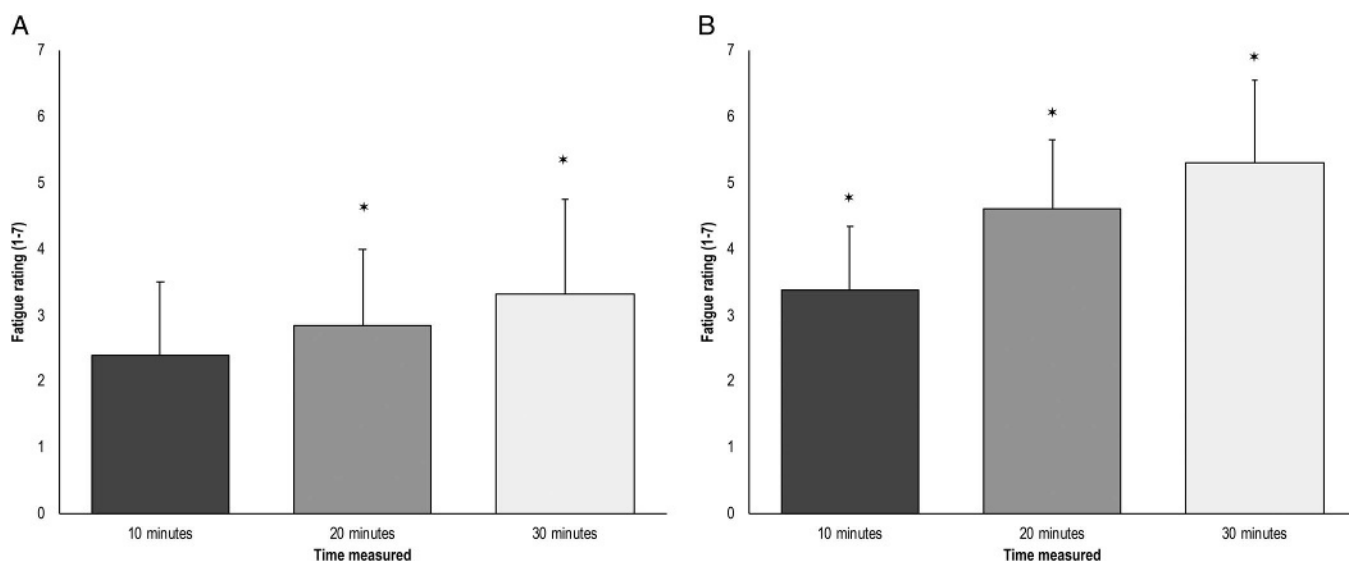


FIGURE 2—Fatigue ratings for the congruent (A) and incongruent (B) Stroop task, indicating significant increases in fatigue over time, and differences between all time periods, with the exception of the difference between 10 and 20 min in the congruent condition,  $t(12) = -2.14$ ,  $P = 0.053$ . Error bars represent 95% confidence intervals.



condition. Overall, participants reported moderate levels of concern with performing well and a desire to be successful and invest effort in the beep test across the two conditions. Paired samples *t* tests comparing perceived motivation for the beep test after its completion also showed no differences in feelings of “having to” or “wanting to” complete the task by condition.

## DISCUSSION

The effect of cognitive fatigue on physical performance is currently a quickly growing area of research in sports science, with implications for advancing both theoretical and applied knowledge. This study sought to resolve conflicting findings by replicating the methods used in Schücker and MacMahon (8) and testing performance on the beep test after a 30-min Stroop task. The results support the findings of previous studies, such as those conducted by Marcora et al. (9), MacMahon et al. (10), and Smith et al. (13) in relation to the negative influence of cognitive fatigue on physical performance in general. Specifically, in the high load condition, participants withdrew from the beep test significantly earlier than when in the low load condition.

In line with the performance results, participants reported that they invested more effort in the 30 min of the incongruent Stroop task as compared with the 30 min of the congruent Stroop task. Perceived cognitive fatigue levels were also found to increase significantly at each measurement in the incongruent Stroop task. Thus, although Brown and Bray (6) reported increasing levels of reported perceived mental fatigue with increasing length of time spent on the incongruent Stroop task, we confirmed this pattern using a within-subjects, repeated-measures design, and extending beyond their end point of 10 min. Moreover, measurement of baseline fatigue levels allowed us to eliminate the possibility of differences due to this variable. It is also notable that Stroop task performance was maintained, although subjective perceived cognitive fatigue increased. This underscores the importance of measuring perceived fatigue, rather than relying solely on task performance as an indicator, as also shown in previous studies (9,10). This is also in line with Ackerman et al.’s findings that fatigue is a subjective cognitive state that may not always be reflected in impaired performance with increasing time spent on a cognitive task (21). Altogether, these findings resolve the previously conflicting results; although the cognitive fatigue ratings after 10 min of the incongruent Stroop task (high load condition) were relatively high and indeed higher than those after 30 min of the congruent Stroop task (low load condition), Schücker and MacMahon’s (8) use of 10 min of the high load task (incongruent Stroop) may not have induced enough fatigue to negatively influence performance in an athletically active sample.

It should be noted, however, that there are differences in the versions of the Stroop task that are used throughout the literature. For instance, Brown and Bray (6) incorporated six color options, and no exception rule, whereas this study, Schücker and MacMahon (8), and Pageaux et al. (22) used a

computerized, four-option version with an exception task, and Martin Ginis and Bray (12) and Smith et al. (13) used a paper stimulus, verbal-response version of the four-option, exception rule Stroop task. Moreover, Martin et al. (7) used 30 min of the incongruent Stroop task with only 10 min of the congruent Stroop task in the low load condition. These comparisons suggest the possibility that small variations in the cognitive task and its demand may be enough to determine whether there is an influence on the performance of a subsequent physical task. They also suggest the influence of extraneous variables and particularly individual differences in factors, such as response to cognitive manipulations.

The majority of research in this area uses subjective ratings of fatigue. Hockey (23) argues that this is appropriate, given that the subjective feeling of cognitive fatigue is the primary marker of the state. Additionally, there are methodological barriers, given that most of the measures used to determine if a participant is in a state of cognitive fatigue involve completion of another cognitive task, such as working memory tests. These tests are themselves fatiguing and thus inappropriate when the goal is to understand the effects of cognitive fatigue on subsequent task performance. Thus, although there are developments in tasks used to assess cognitive fatigue, including more objective measures (24), we do not yet have appropriately sensitive and easily applied continuous measures of cognitive fatigue that can be used throughout a cognitive task for individual and threshold control of the manipulation.

This study supports the findings of Schücker and MacMahon (8) with regard to RPE. In line with previous studies (10), there was no difference in the final RPE scores between conditions. Because participants performed a shorter beep test in the fatigue condition, this is interpreted as an indicator that the high load condition was perceived to be more difficult, with similarly high ratings as the low load condition, for less physical work.

Based on the process model as put forward by Inzlicht et al. (17), we hypothesized that participants would report higher levels of feeling like they *had* to complete the beep test in the high demand condition compared with the low demand condition and, similarly, that they would report lower levels of feeling like they *wanted* to complete the beep test in the high demand compared with the low demand condition. This hypothesis was not supported, because there were no significant differences between the two conditions for ratings of “having to” and “wanting to” complete the beep test. Measurement of this variable needs further examination, however, because this is the first attempt to use this specific terminology.

We acknowledge that there are several limitations within this work. First, the researcher collecting data was not blind to the condition that participants were completing within each testing session. Although this is difficult to avoid, given view of the cognitive task and its congruence or incongruence on the computer screen, there are ways to remedy this, including a different researcher testing the physical tests alone. Regardless, the researcher in this study was trained to objectively

assess beep test performance and to refrain from making any motivational comments during performance, which mitigated the lack of blinding. Second, although efforts were made to control for effects of the time of day on performance in both cognitive and physical tasks, this was not possible given logistical difficulties. Therefore, although participants were generally tested between the hours of 10:00 AM and 2:30 PM, there were variations and the exact testing time may have also varied within this range between the two testing sessions. Finally, ratings of perceived exertion were only collected at the end of each beep test performance, and not throughout testing as in other studies (8). This is also not the traditional method of collecting RPE and may introduce measurement error.

## DIRECTIONS FOR FUTURE RESEARCH

The findings of this study present several directions and areas for future research. First, continuing to explore the effect of the length of the cognitive task, as well as variations of the demands of this task is important; the duration point for fatigue, as well as individual differences in this duration point, is still unclear. Indeed, it would be beneficial to investigate whether this length is universal to alternative types of cognitive tasks used in this area (e.g., crossing out letters, counting backwards), or whether this potential minimum duration of 30 min is exclusive to the Stroop task, and whether individual differences can be attributed to characteristics, such as working memory, self-control (25), and action orientation (26). Moreover, as Brown and Bray (6) used the handgrip task, and this study used a whole body externally paced task, it is worth exploring the nature of the physical task and possible interactions with the cognitive task.

Second, further exploration of the influence of motivation on performance is needed. Although several studies have measured different aspects of motivation (e.g., 1, 9, 10, 15), this has been in a limited manner. Developing and enhancing appropriate measures would be fruitful for future investigations, particularly given increased focus on this variable within the self-regulation literature (17). This is the first attempt we are aware of to use the specific “want to” versus “have to” terminology. Development and exploration of changes in motivation might incorporate stronger measures of intrinsic and extrinsic task motivation given that typical measurement—as is the case here—is through a limited number of self-report items. In addition, experimental designs may prompt the development of implicit expectations of impaired performance after a harder cognitive task, which may in turn influence subsequent behaviors and subjective responses. Indeed, will-power beliefs have been shown to both influence outcomes and be vulnerable to manipulation (27). So far, research designs have not been mindful of potential expectation effects.

**Implications.** This study supports the call to identify the conditions under which declines in performance are present as empirically as well as practically valuable (5). The results show a difference of one half beep test level between the two conditions: level 8.5 for the high load condition and level

9 for the low load condition. This is an important finding: given use of the beep test in many sports (e.g., the Australian Rules Football League draft), the difference in performance could determine whether an athlete is selected for their sport or not. There are similar implications for areas, such as fire services and police training, which also use this and similar physical tasks for testing and selection (28). Thus, even in simple testing, being aware of the influence of cognitive fatigue and being able to cope with these effects can provide an advantage or minimize a potential decrement.

The theoretical implications of this study are clear: they provide a greater understanding of the effects of cognitive tasks on subsequent physical performance. As discussed, we resolved the conflict between Schücker and MacMahon (8) and Smith et al. (13) and the issue of the length of the Stroop task as relevant for subsequent performance on an externally paced intermittent running task. Specifically, we discovered that, for a recreationally active group participating in this research design, a Stroop task duration of 30 min was long enough to impair physical performance. These findings imply that the training of a physical task may influence whether performance will decline after a perturbation, such as higher perceived cognitive fatigue. However, given that different subject pools have been used across similar studies, and that experience and competition level have not always been controlled or a main variable of interest, more work with clearer designs is needed. Moreover, generalizing this result to all cognitive tasks is not simple, and questions remain to understand the effects of different tasks (e.g., counting backward) and the duration surrounding these different tasks, for both more or less experienced participants. Similarly, questions also remain around the nature of the physical task, particularly given that Brown and Bray (6) show a threshold of 4-min engagement in the Stroop task for interference in performance of a handgrip endurance task, albeit, using a different version of the Stroop task (six options).

This study clearly provides a strong platform for continued work, with exploration of different target and manipulation tasks, and individual differences and levels of experience and expertise as key topics for further focus. There is also strong potential to study interventions based on these findings, and the results provide several practical implications. The key among these implications is the need to consider the cognitive demands for performing athletes. Indeed, sporting teams spend much of their training time working on the physical side of their respective sport, whether that be skills, endurance, strength, and so on. This study and the literature on the influence of cognitive tasks on physical performance highlight that cognitive exertion may also be a factor in performance and worthy of consideration within training.

The authors have no conflicts of interest, source of funding, or financial ties to disclose and no current or past relationship with companies or manufacturers who could benefit from the results of the present study. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation, and the results of the present study do not constitute endorsement by ACSM. The authors also acknowledge Stijn Mentzel for his kind help in improving the figures.

## REFERENCES

- Englert C. The strength model of self-control in sport and exercise psychology. *Frontiers in Psychol.* 2016;7:314.
- Van Cutsem J, Marcora S, De Pauw K, Bailey S, Meeusen R, Roelands B. The effects of mental fatigue on physical performance: a systematic review. *Sports Med.* 2017;47(8):1569–88.
- McMorris T, Barwood M, Hale BJ, Dicks M, Corbett J. Cognitive fatigue effects on physical performance: a systematic review and meta-analysis. *Physiol Behav.* 2018;188:103–7.
- Taylor IM, Boat R, Murphy SL. Integrating theories of self-control and motivation to advance endurance performance. *Int Rev Sport Exerc Psychol.* 2018;23:1–20.
- Silva-Cavalcante MD, Couto PG, Azevedo RA, et al. Mental fatigue does not alter performance or neuromuscular fatigue development during self-paced exercise in recreationally trained cyclists. *Eur J Appl Physiol.* 2018;118(11):2477–87.
- Brown DM, Bray SR. Graded increases in cognitive control exertion reveal a threshold effect on subsequent physical performance. *Sport Exerc Perform Psychol.* 2017;6(4):355–69.
- Martin K, Staiano W, Menaspà P, et al. Superior inhibitory control and resistance to mental fatigue in professional road cyclists. *PLoS One.* 2016;11(7):1–15.
- Schücker L, MacMahon C. Working on a cognitive task does not influence performance in a physical fitness test. *Psychol Sport Exerc.* 2016;25:1–8.
- Marcora SM, Staiano W, Manning V. Mental fatigue impairs physical performance in humans. *J Appl Physiol (1985).* 2009;106(3):857–64.
- MacMahon C, Schücker L, Hagemann N, Strauss B. Cognitive fatigue effects on physical performance during running. *J Sport Exerc Psychol.* 2014;36(4):375–81.
- Ramsbottom R, Brewer J, Williams C. A progressive shuttle run test to estimate maximal oxygen uptake. *Br J Sports Med.* 1988;22(4):141–4.
- Martin Ginis KA, Bray SR. Application of the limited strength model of self-regulation to understanding exercise effort, planning and adherence. *Psychol Health.* 2010;25(10):1147–60.
- Smith M, Coutts A, Merlini M, Deprez D, Lenoir M, Marcora S. Mental fatigue impairs soccer-specific physical and technical performance. *Med Sci Sports Exerc.* 2016;48(2):267–76.
- Englert C, Wolff W. Ego depletion and persistent performance in a cycling task. *Int J Sport Psychol.* 2015;46:137–51.
- Bray S, Martin Ginis K, Hicks A, Woodgate J. Effects of self-regulatory strength depletion on muscular performance and EMG activation. *Psychophysiology.* 2008;45(2):337–43.
- McEwan D, Martin Ginis K, Bray S. The effects of depleted self-control strength on skill-based task performance. *J Sport Exerc Psychol.* 2013;35(3):239–49.
- Inzlicht M, Schmeichel B, Macrae C. Why self-control seems (but may not be) limited. *Trends Cogn Sci.* 2014;18(3):127–33.
- MacMahon C, Parrington L. Not all athletes are equal, but don't call me an exerciser: response to Araujo and Scharhag. *Scand J Med Sci Sports.* 2017;27(8):904–6.
- McKinney J, Velghe J, Fee J, Issero S, Drezner J. Defining athletes and exercisers. *Am J Cardiol.* 2019;123(3):532–5.
- Borg G. *Borg's Perceived Exertion and Pain Scales.* Champaign, IL, US: Human Kinetics; 1998.
- Ackerman PL, Kanfer R, Shapiro SW, Newton S, Beier ME. Cognitive fatigue during testing: an examination of trait, time-on-task, and strategy influences. *Hum Perform.* 2010;23(5), 381–402.
- Pageaux B, Lepers R, Dietz K, Marcora S. Response inhibition impairs subsequent self-paced endurance performance. *Eur J Appl Physiol.* 2014;114(5):1095–105.
- Hockey R. *The Psychology of Fatigue: Work, Effort and Control.* Cambridge: Cambridge University Press; 2013. 272 p.
- Tanaka M, Ishii A, Watanabe Y. Neural effects of mental fatigue caused by continuous attention load: a magnetoencephalography study. *Brain Res.* 2014;1561:60–6.
- Tangney JP, Baumeister RF, Boone AL. High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *J Pers.* 2004;72(2):271–324.
- Gröpel P, Baumeister RF, Beckmann J. Action versus state orientation and self-control performance after depletion. *Personal Soc Psychol Bull.* 2014;40(4):476–87.
- Francis Z, Job V. Lay theories of willpower. *Soc Personal Psychol Compass.* 2018;2(4):1–13.
- Victoria State Government Website [Internet]. Victoria (Aus): Victoria Government; [cited 2019 Jan 3] Available from: <https://firefighter.vic.gov.au/fitness-preparation/>.