Original Article

The effect of self-selected music on endurance running capacity and performance in a mentally fatigued state

HUI KWAN NICHOLAS LAM[1], HARRY MIDDLETON[2], SHAUN M PHILLIPS[1] ✉

[1]Human Performance Science Research Group, The University of Edinburgh, Edinburgh, United Kingdom
[2]Edinburgh Medical School, The University of Edinburgh, Edinburgh, United Kingdom

ABSTRACT

We investigated the effects of listening to self-selected music on intermittent running capacity (study 1) and 5 km time-trial (TT) performance (study 2) in a mentally fatigued state. In study 1, nine physically active males performed a 30-minute incongruent Stroop test (IST) followed by the Yo-Yo intermittent recovery test level 1 (YYIRT1) with (MF+MUSIC) and without (MFONLY) music. They also completed a baseline trial (BL). Study 2 repeated these trials with nine recreational runners. Mental fatigue (MF) showed large increases following IST in both studies (d_{unb} = 1.44 – 2.0). Intermittent running capacity was moderately greater in MF+MUSIC (564 ± 127 m; d_{unb} = 0.52) and BL (551 ± 106 m; d_{unb} = 0.51) vs. MFONLY (496 ± 112 m). Time-trial performance showed small improvements in MF+MUSIC (23.1 ± 2.4 min; d_{unb} = 0.28) and BL (23.4 ± 3.5 min; d_{unb} = 0.20) vs. MFONLY (24.1 ± 3.2 min). Differences in ratings of perceived exertion between trials were trivial to small in both studies (d_{unb} = 0-0.47). Listening to self-selected music in a mentally fatigued state negates the negative impact of MF on endurance running capacity and performance, potentially due to altered perception of effort when listening to music.
Keywords: Cognitive load; Mental fatigue; Running performance; Running capacity; Mental exertion; Perceived exertion; Sports performance.

Cite this article as:
Lam, H.K.N., Middleton, H., & Phillips, S.H. (2021). The effect of self-selected music on endurance running capacity and performance in a mentally fatigued state. *Journal of Human Sport and Exercise*, *in press*. doi:https://doi.org/10.14198/jhse.2022.174.16

INTRODUCTION

Mental fatigue (MF) is a negative change in psychobiological state caused by prolonged periods of demanding cognitive activity (Martin et al., 2018). The negative influence of MF on endurance capacity and performance is well established (Van Cutsem et al., 2017), and often associated with greater ratings of perceived exertion (RPE) for a given exercise intensity (Marcora et al., 2009; Pageaux et al., 2015). In contrast, physiological parameters such as heart rate (HR) and blood lactate concentration appear unaffected by MF (Van Cutsem et al., 2017).

Increased neural activity during a MF task may increase extracellular cerebral concentrations of adenosine, an inhibitory neurotransmitter (Martin et al., 2018). One of the brain regions that is heavily activated during cognitive activity, and where localised adenosine accumulation is likely to occur, is the anterior cingulate cortex (ACC) (Martin et al., 2018). The ACC is implicated in generating the perception of effort in humans (Williamson et al., 2001), and the working hypothesis is that mental exertion leads to adenosine accumulation within the ACC, exerting an inhibitory influence and thereby requiring greater stimulatory input to produce the required force output, resulting in increased RPE (Martin et al., 2018).

Listening to music during submaximal exercise can positively influence perceptions of the exercise task and exercise performance (Karageorghis and Priest, 2012a). The benefit of music during exercise may be due to a distraction effect which decreases RPE and fatigue perception (Karageorghis and Priest, 2012a). Interestingly, however, the brain may process music differently during exercise compared to passive listening (Hutchinson and Karageorghis, 2013), with the ACC implicated in the processing of music during exercise (Chanda and Levitin, 2013). The shared involvement of the ACC in adenosine accumulation during cognitive tasks, generating RPE during exercise, and processing of music during exercise provides an intriguing plausible mechanism whereby music could counteract the negative effect of MF on endurance exercise capacity and performance.

It is difficult to determine objectively the optimal characteristics of music required to induce an ergogenic effect during endurance exercise. Dyer and McKune (2013) found no significant effect of fast (140 bpm) vs slow (120 bpm) tempo music on 20 km cycle completion time but reported significantly greater measures of tension in participants during the fast tempo music trial. Conversely, Buhmann et al. (2018) reported a positive relationship between music tempo and preferred running cadence, with subsequent improvements in self-pacing strategies when listening to high tempo music. Since it is hard to determine the optimal music tempo, and other characteristics of music are likely to influence its ergogenic effect (Karageorghis and Priest, 2012a), self-selection of music may be a more effective approach (Buhmann et al., 2018). However, no studies have investigated the effect of listening to self-selected music following a MF task on endurance running capacity or performance.

This paper reports two studies which investigated the effects of listening to self-selected music on intermittent endurance running capacity (study 1) and 5 km running performance (study 2) in a mentally fatigued state. We hypothesised that MF would meaningfully increase subjective sensations of MF and meaningfully reduce subjective sensations of vigour in both studies. We also hypothesised that distance covered in the intermittent running task would be meaningfully greater and 5 km running performance meaningfully faster when listening to music in a mentally fatigued state vs. not listening to music in a mentally fatigued state.

MATERIALS AND METHODS

Participants
In study 1, nine males participated (mean ± SD: age: 22 ± 2.6 years; height: 1.73 ± 0.02 m; body mass: 70.0 ± 7.5 kg). In study 2, seven males and two females participated (age 21.1 ± 1.2 years; height 1.79 ± 0.04 m; body mass: 74.0 ± 8.5 kg). For both studies, inclusion criteria were aged between 18 and 30 years, habitually listen to music during exercise, injury free in the past three months, and with no known health conditions. Participants in study 1 were recreationally active in a variety of sport and exercise past times, and in study 2 were recreational runners. All participants were informed about the testing procedures and written informed consent and medical screening were completed before participation. Both studies received approval from the institutional research ethics sub-committee.

Procedures
Participants visited the laboratory on four occasions at a similar time (± 2 hours) with ≥ 48 h between sessions (Dyer and McKune, 2013; Hutchinson and Karageorghis, 2013). The first visit was a familiarisation where participants completed anthropometric measurements (BM: Seca Portable Scales, Model 803, UK; height: Seca Stadiometer, Model 213, UK) and performed the entire test procedure once without music. Participants subsequently completed three trials: MF and music (MF+M), MF and no music (MFONLY), and a baseline trial that did not include a MF protocol or the use of music (BL). Trial orders were determined via simple randomisation (https://www.randomizer.org/). The efficacy of using an emotionally neutral documentary as a control condition in MF research is debated (Smith et al., 2015). Therefore, the BL trial did not include a passive period in place of the MF protocol, reflecting participants' performance with no potential interference effects. Participants were instructed to avoid caffeine and alcohol consumption and to limit physically and cognitively demanding activity ≥ 24 h before each trial.

In the MF+M and MFONLY trials of both studies, the MF protocol was a 30-minute incongruent Stroop Test (IST; Stroop Test for Research and Teaching, Version 1.2, Mac-App Store). Participants sat comfortably at a desk with a 13-inch laptop placed 60 cm away from eye level. Earplugs (3M, Model 1100, UK) were worn during the test to reduce sound disturbance. Four words (blue, green, red and yellow) were consecutively displayed in a random order in the centre of the screen on a black background. Each word was presented in a particular colour in a random fashion (for example, the word red presented in a green font). Participants were required to extract the visual information and correctly select the colour of the word instead of the meaning of the word (in the previous example, the correct answer would be green). There was no time limit for each response, but participants were instructed to respond as quickly as possible. The next word was presented as soon as the participant selected their answer to the current word. No verbal communication was available during the IST, but total responses and remaining time were displayed at the top right corner of the screen. This protocol was sufficient to elicit MF and impair endurance performance in previous research (Clark et al., 2019; Pageaux et al., 2014; Smith et al., 2016).

One minute after completion of the IST, participants performed a standardized 7 min warm up as described elsewhere. Three min after completing the warm-up, participants began the relevant exercise, as described below:

Study 1: Yo-Yo Intermittent Recovery Test Level 1 (YYIRT1)
Participants performed 2 x 20-metre shuttle runs at increasing velocities interspersed with 10 sec active recovery in a climate controlled laboratory (20-22°C, 60-65% relative humidity). Initial running speed was 10 km.h^{-1}, with speed controlled by audio signals from a portable speaker (Beyond acoustic, Model RX12, UK).

The test ended when the participant failed to maintain the required speed for two consecutive shuttles. Total distance covered was calculated as the accumulated distance up to and including the last completed shuttle. The YYIRT1 demonstrates high reliability in evaluating the endurance capacity of recreationally active individuals (r = 0.95, coefficient of variation 4.9% to 8.7%) (Bangsbo et al., 2008; Schmitz et al., 2018). The BL trial followed the same procedures described above, beginning from the standardized warm-up.

Study 2: 5 km treadmill time-trial
A 5km time trial (TT) was completed on a treadmill (Woodway, Model ELG 55) in a climate controlled laboratory (20-22°C, 60-65% relative humidity). The treadmill was set at a 1% gradient, and participants were asked to complete the distance as quickly as possible. The TT began at 9 km.h^{-1}, but participants had full control over their speed from the outset. Feedback on distance covered was provided every 500 metres during the TT, but speed and elapsed time was withheld.

Music
We used the BMRI-2 to assist participants in selecting songs that motivated them the most from a playlist they habitually listened to while exercising. The BMRI-2 is a 7-point numerical scale rating from 1 (strongly disagree) to 7 (strongly agree) for six questions about the rhythm, style, melody, tempo, sound and beat of the music (Karageorghis et al., 2006). Although previous research suggests a score $\geq$ 36 is considered motivational music (Karageorghis et al., 2006; Terry et al., 2011), self-selected music rated < 36 was still considered, as motivation is entirely subjective. However, the BMRI-2 provided a framework for participants to carefully consider the motivational quality of their music. In the MF+M trial the selected songs were played in a random order through a portable speaker (JBL, Model Charge 3, UK) beginning at the start of the 3 min post-warm up period. This time period was selected as Nakamura et al. (2010) state that exercise with music significantly reduces RPE from 120 s after it begins to play, and Stork et al. (2015) report that music played for 2.5 minutes pre-exercise increases motivation for high-intensity exercise. The music played until termination of the exercise. Sound intensity of the music was measured by a decibel (dB) meter (Meterk, Model MK09, UK) and standardized to 75 dB at ear level (Karageorghis et al., 2009).

Measures
The perceptual measurements were adapted from previous research (Pageaux et al., 2015; Smith et al., 2016). Subjective perception of MF was assessed using a 100-mm visual analogue scale (VAS). Participants responded to a standardized question (''what is your current level of MF?'') by marking the VAS between zero ('no fatigue') and 100 ('severe fatigue') immediately before and after the IST. The 100-mm VAS demonstrates high internal consistency (α = 0.94 - 0.96) (Lee et al., 1991). The Brunel Mood Scale (BRUMS) was used to quantify mood states (Moss et al., 2018; Terry et al., 2011). The BRUMS contains six subscales (depression, anger, tension, vigour, fatigue and confusion) each comprising four items scored on a Likert scale (0 – not at all to 4 – extremely) (Brandt et al., 2016). Participants completed the BRUMS immediately before and after the IST in response to the question "*How do you feel now?*" Total scores for each subscale ranged from zero to 16. We report the vigour and fatigue subscales (Martin et al., 2015; Pageaux et al., 2014; Pageaux et al., 2015). The BRUMS has good reliability and internal consistency (α = 0.85) in healthy populations (Brandt et al., 2016). Rating of perceived exertion was measured using the 6-20 Borg RPE scale immediately following termination of the YYIRT1 in study 1 and at the beginning and every kilometre of the TT in study 2, in response to the question ''how hard is the task?' (Pageaux et al., 2014; Pageaux et al., 2015). The RPE scale is a reliable tool (r = 0.83 to 0.93) in assessing healthy participants' perceived exertion during exercise (Scherr et al., 2013; Terry et al., 2011). In study 2, heart rate (HR) was recorded at the start of the TT and every kilometre using a chest monitor (Polar, Model Accurex Plus).

Analysis

Null hypothesis significance testing (NHST) readily yields false conclusions about the existence of an effect and the practical meaning of data; *P* values are also subject to large variation due to sampling variability (Wasserstein et al., 2019). As a result, eminent statistical organisations have recently published extensively on moving away from NHST (Wasserstein et al., 2019). This guidance recommends that researchers do not conclude anything about the practical or scientific importance of data based on statistical significance (Wasserstein et al., 2019). Alongside words of caution about NHST, researchers are recommended to analyse data in a way that provides meaningful information about precision and uncertainty in the data, and the likely population effect based on the data (Calin-Jageman and Cumming, 2019). We take this approach in our analysis.

Data normality was assessed using the Shapiro-Wilk test. For all outcome variables, pairwise comparisons were made by calculating the mean difference with 95% confidence limits (95%CL) between the two time points or trials as appropriate. Cohen's *d* effect size for the mean difference was calculated using the equation:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{s_{mean}}$$

Where s_{mean} = mean of the standard deviations of the two numerator data sets and was calculated using the equation:

$$s_{mean} = \sqrt{\frac{s_1^2 + s_2^2}{2}}$$

Mean standard deviation represents the best estimate of the population standard deviation in within-participants designs, and is therefore the recommended standardiser for *d* (Lakens, 2013).

Effect size (ES) estimates based on sample data overestimate the true population effect (Lakens, 2013; Thompson, 2007). Therefore, we applied a correction to all ES to mitigate this bias, as described by Lakens (2013) and Cumming (2012). All ES in this study are presented as unbiased *d* (d_{unb}), in line with the convention recommended by Cumming (2012). We calculated 95%CL for the ES using the procedure described by Algina and Keselman (2003). The magnitude of the ES was defined as trivial (*d* < 0.2), small (*d* ≥ 0.2, <0.5), medium (*d* ≥ 0.5, <0.8), and large (*d* ≥ 0.8) (Cohen, 1992; Smith et al., 2015), expressed in units of standard deviation. Differences between trials are reported in the text in the following manner:

[mean difference, 95%CL for that difference followed by units of measurement]; [Cohen's d_{unb} ES for the difference, 95%CL for that ES]

Worked example:

2, 95%CL 0,4 beat.min^{-1}; *d* = 0.23, 95%CL 0.06, 0.41

RESULTS

For study 1, mean BMRI-2 score of the self-selected music was 33.9 ± 3.6. For study 2, mean BMRI-2 score was 36.9 ± 3.2.
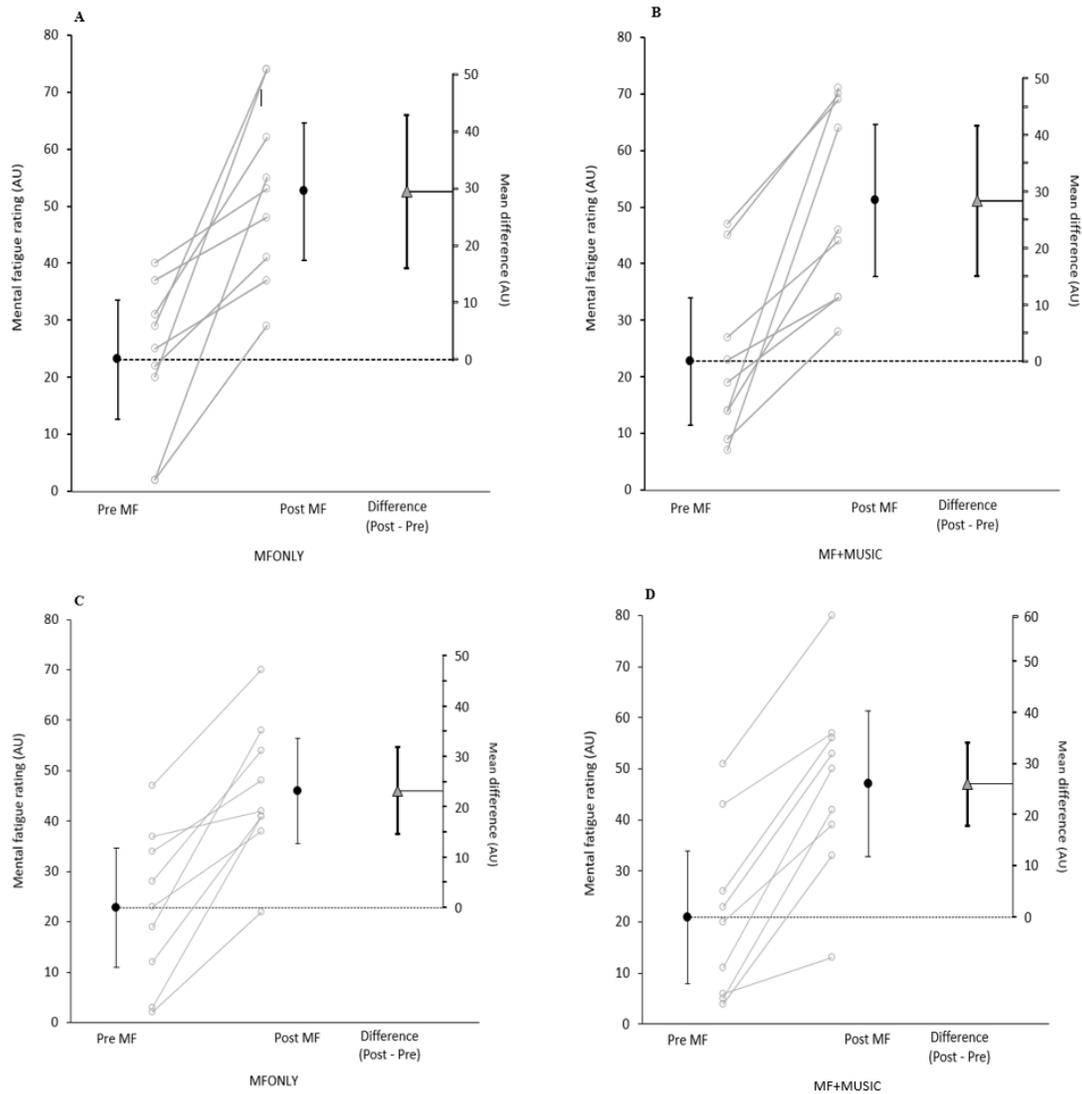
Mental fatigue (study 1 and 2)

Figure 1. Visual analogue ratings of mental fatigue pre- and post-mental fatigue protocol. Study 1: MFONLY (A) and MF+MUSIC (B). Study 2: MFONLY (C) and MF+MUSIC (D). Black circles with error bars are mean with 95%CI; grey lines represent individual participants; grey triangle is the mean difference between trials with 95%CI.

Mental fatigue VAS data from both studies is in Figure 1. In study 1, pre-MF VAS scores were similar between MFONLY and MF+MUSIC (mean difference 0.3, 95%CI -13.2 to 13.9 AU; d_{unb} = 0.02, 95%CL -0.79 to 0.84). A large increase in MF VAS scores occurred in the MFONLY (d_{unb} = 2.01, 95%CI 0.75 to 3.22) and MF+MUSIC (d_{unb} = 1.76, 95%CL 0.64 to 2.84) trials. The difference in post-MF VAS ratings between the two

trials was trivial (mean difference 1.4, 95%CL -8.2 to 11.1 AU, d_{unb} = 0.08, 95%CL -0.41 to 0.58). In study 2, pre-MF VAS scores were similar between MFONLY and MF+MUSIC (mean difference 1.8, 95%CI -8.9 to 5.2 AU; d_{unb} = 0.01, 95%CL -0.48 to 0.26). A large increase in MF VAS scores occurred in the MFONLY (d_{unb} = 1.44, 95%CI 0.67 to 2.5) and MF+MUSIC (d_{unb} = 1.32, 95%CL 0.65 to 2.26) trials. The difference in post-MF VAS ratings between the two trials was trivial (mean difference 1.0, 95%CL -6.6 to 8.6 AU, d_{unb} = 0.06, 95%CL -0.34 to 0.46).
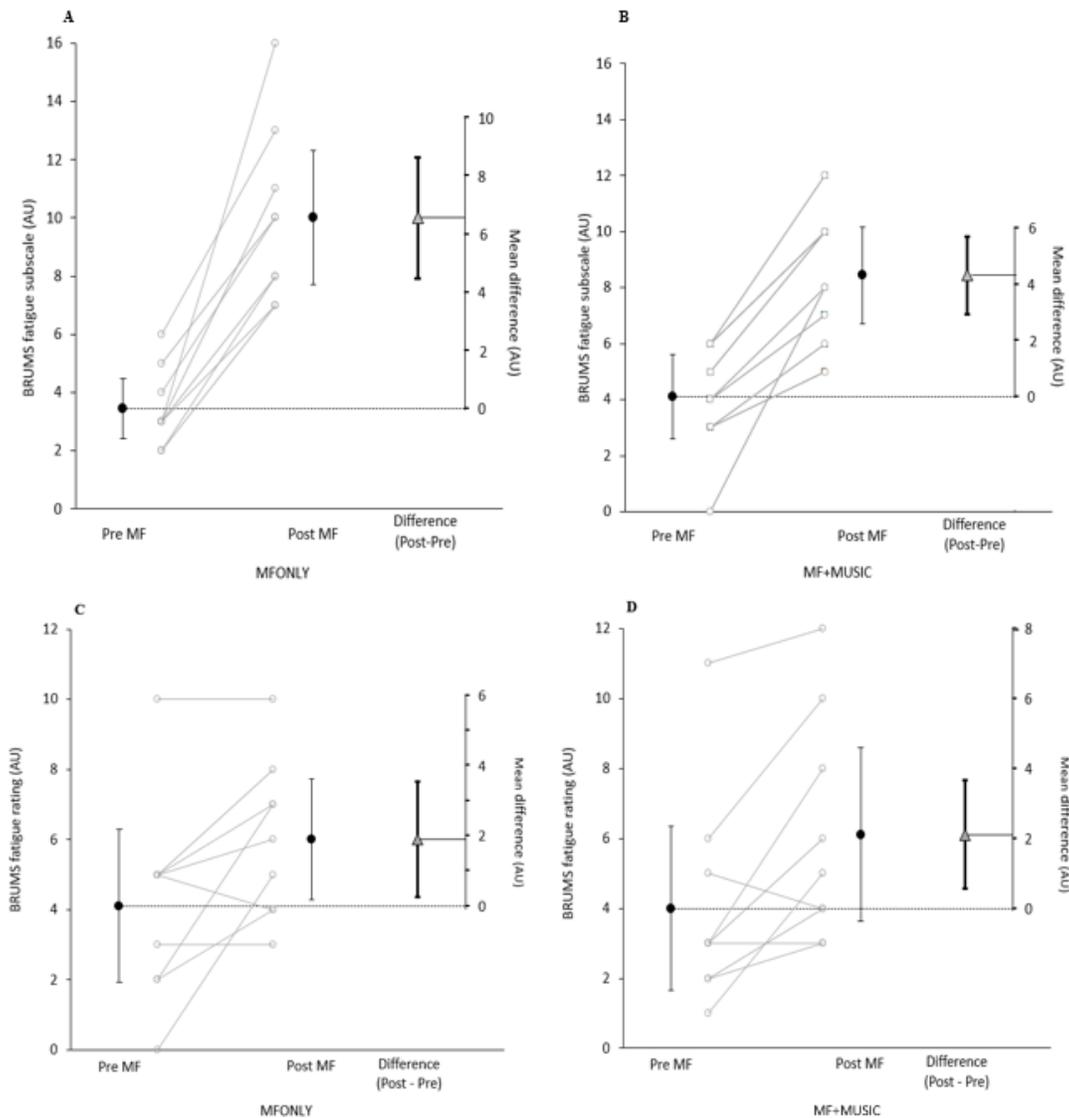

Figure 2. Brunel Mood Scale fatigue ratings pre- and post-mental fatigue protocol. Study 1: MFONLY (A) and MF+MUSIC (B). Study 2: MFONLY (C) and MF+MUSIC (D). Black circles with error bars are mean with 95%CI; grey lines represent individual participants; grey triangle is the mean difference between trials with 95%CI.

Data from the BRUMS fatigue subscale for both studies is in Figure 2. In study 1, pre-MF fatigue scores were greater to a small extent in MF+MUSIC vs. MFONLY (mean difference 0.7, 95%CI -0.77 to 2.1; d_{unb} = 0.36,

95%CL -0.37 to 1.10). A large increase in fatigue scores occurred in the MFONLY (d_{unb} = 2.55) and MF+MUSIC (d_{unb} = 1.86) trials. Post-MF fatigue ratings were moderately greater in MFONLY vs. MF+MUSIC (mean difference 1.6, 95%CI -0.6 to 3.7 AU, d_{unb} = 0.53, 95%CL -0.17 to 1.32). In study 2, pre-MF fatigue scores were similar between trials (mean difference 0.1, 95%CI -1.8 to 1.5; d_{unb} = 0.03, 95%CL -0.51 to 0.44). A medium increase in fatigue scores occurred in the MFONLY (d_{unb} = 0.67, 95%CI 0.07 to 1.38) and MF+MUSIC (d_{unb} = 0.61, 95%CI 0.13 to 1.20) trials. Post-MF fatigue ratings were similar between trials (mean difference 0.1, 95%CI -1.3 to 1.6; d_{unb} = 0.04, 95%CL -0.41 to 0.49).
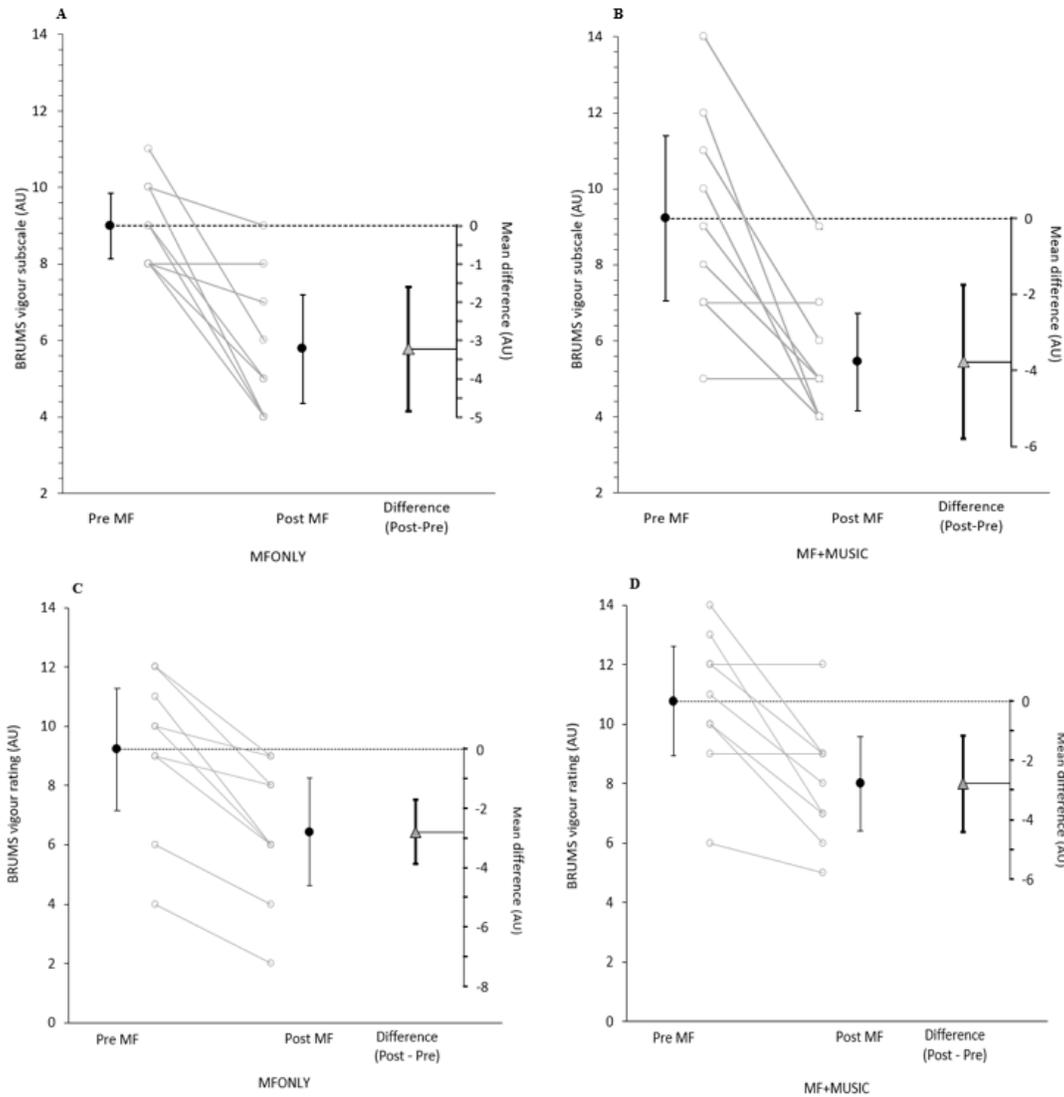

Figure 3. Brunel Mood Scale vigour ratings pre- and post-mental fatigue protocol. Study 1: MFONLY (A) and MF+MUSIC (B). Study 2: MFONLY (C) and MF+MUSIC (D). Black circles with error bars are mean with 95%CI; grey lines represent individual participants; grey triangle is the mean difference between trials with 95%CI.

Data from the BRUMS vigour subscale is in Figure 3. In study 1, pre-MF vigour scores were similar between MFONLY and MF+MUSIC (mean difference 0.2, 95%CI -2.2 to 1.8 AU; d_{unb} = 0.09, 95%CL -0.91 to 0.70). A

large decrease in vigour scores occurred in the MFONLY (d_{unb} = -1.90, 95%CL -0.72 to -3.43) and MF+MUSIC (d_{unb} = -1.47, 95%CL -0.53 to -2.69) trials. The difference in post-MF vigour ratings between the two trials was trivial (mean difference 0.3, 95%CI -1.2. to 0.5 AU, d_{unb} = 0.17, 95%CL -0.61 to 0.24). In study 2, pre-MF vigour scores were moderately greater in MF+MUSIC vs. MFONLY (mean difference 1.6, 95%CI -0.14 to 3.21 AU; d_{unb} = 0.55, 95%CL -0.04 to 1.23). A large decrease in vigour scores occurred in the MFONLY (d_{unb} = -0.99, 95%CL -0.45 to -1.73) and MF+MUSIC (d_{unb} = -1.12, 95%CL -0.37 to -2.10) trials. Vigour scores post-MF were moderately lower in MFONLY vs. MF+MUSIC (mean difference -1.6, 95%CI -0.01 to -3.09 AU; d_{unb} = 0.64, 95%CL 0.0 to -1.37).
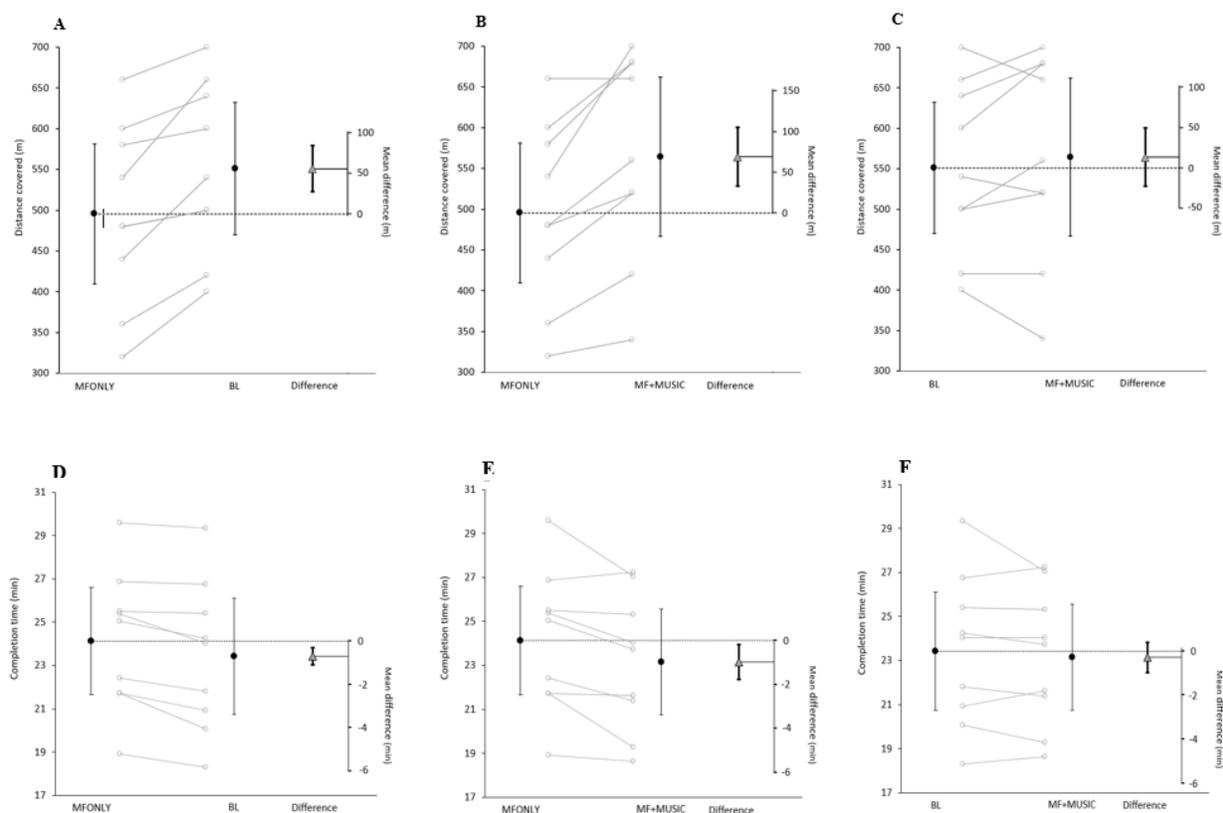
Exercise capacity and performance

Figure 4. Distance covered in the Yo-Yo Intermittent Recovery Test Level 1 in the MFONLY and BL trials (A), MFONLY and MF+MUSIC trials (B), and BL and MF+MUSIC trials (C) of study 1; 5 km time-trial completion time in the MFONLY and BL trials (D), MFONLY and MF+MUSIC trials (E), and BL and MF+MUSIC trials (F) of study 2. Black circles with error bars are mean with 95%CI; grey lines represent individual participants; grey triangle is the mean difference between trials with 95%CI.

For study 1, distance covered in the YYIRT1 is in Figure 4A-C. Mean distance covered in BL was moderately greater than MFONLY (+55.6 ± 37.1 m, d_{unb} = 0.51, 95%CL 0.17 to 0.84; Figure 4A). Similarly, distance covered in MF+MUSIC was moderately greater than MFONLY (+68.9 ± 47.0 m d_{unb} = 0.52, 95% CL 0.19 to 0.95; Figure 4B). There was a trivial difference in distance covered between BL and MF+MUSIC (-13.3 ± 46.9 m, d_{unb} = 0.10, 95%CL -0.16 to 0.38; Figure 4C). In study 2 (Figure 4D-F), TT completion time showed a small decrease in BL vs. MFONLY (-0.70 ± 0.52 min, d_{unb} = -0.20, 95%CL -0.06 to -0.35; Figure 4D).

Completion time in MF+MUSIC also showed a small reduction vs. MFONLY (-0.99 ± 1.03 min, d_{unb} = -0.28, 95%CL -0.04 to -0.56; Figure 4E). There was a trivial difference in completion time between BL and MF+MUSIC (-0.29, ± 0.90 min, d_{unb} = -0.08, 95%CL 0.10 to -0.27; Figure 4F).

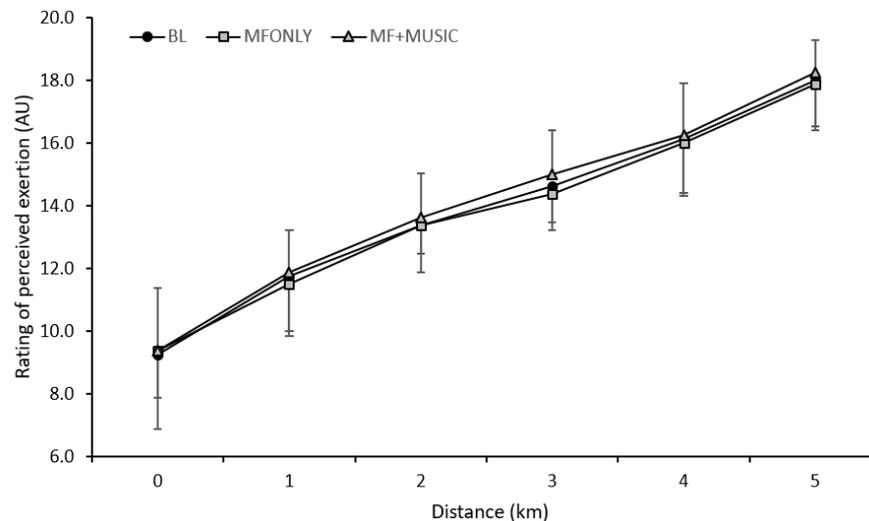Figure 5. Mean (± SD) ratings of perceived exertion at each kilometre of the 5 km time-trial in study 2.

In study 1, post-exercise RPE was higher in MFONLY vs. BL (18 ± 1.1 vs 17.6 ± 1.3, mean difference 0.4, 95%CL 0.0 to 0.8, d_{unb} = 0.33, 95%CL 0.02 to 0.68) and MF+MUSIC (17.7 ± 1.2, mean difference 0.3, 95%CL -0.33 to 1.0, d_{unb} = 0.26, 95%CL -0.23 to 0.78). There was a trivial difference in post-exercise RPE between BL and MF+MUSIC (mean difference 0.1, 95%CL -0.60 to 0.82, d_{unb} = 0.08, 95%CL -0.39 to 0.56). Ratings of perceived exertion for study 2 are in Figure 5. In all trials, RPE showed moderate to large increases as the TT progressed (d = 0.76 – 1.33). The difference in RPE at each measurement point was trivial between BL and MFONLY (d = 0 to -0.19) and trivial to small between BL and MF+MUSIC (d = 0.05 – 0.24) and MFONLY and MF+MUSIC (d = 0 – 0.47).
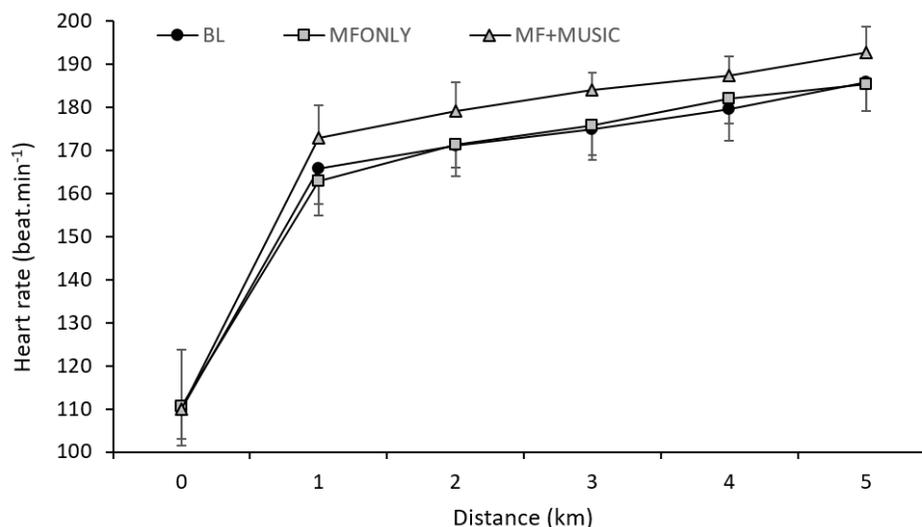

Figure 6. Mean (± SD) heart rate at each kilometre of the 5 km time-trial in study 2.

Heart rate during the 5km TT in study 2 is in Figure 6. In all trials, HR showed a large increase from 0-1km (d = 5.08-6.10) followed by moderate to large increases from 1-5km (d = 0.52-0.98). From 1-5km, differences in HR at each measurement point were trivial to small between BL and MFONLY (d = 0.02 to 0.32). However, differences were large between BL and MF+MUSIC (d = 0.80 – 1.38) and MFONLY and MF+MUSIC (d = 1.0 – 1.27).

DISCUSSION

We report that a 30-min IST elicited moderate to large increases in subjective sensations of MF and large reductions in subjective sensations of vigour. Distance covered in the intermittent running task (study 1) was moderately greater, and 5km TT performance (study 2) showed small improvements, when listening vs. not listening to music in a mentally fatigued state.

We saw moderate to large increases in subjective sensations of fatigue and large reductions in subjective sensations of vigour following the IST, is in agreement with Smith et al. (2016) but contrary to Pageaux et al. (2014). Our participants wore earplugs and were not given verbal feedback during the IST, reducing possible distraction and potentially contributing to a larger and more consistent effect of the IST. However, we must also consider potential inter-individual variability in response to MF protocols (Van Cutsem et al., 2017), which may be modulated by factors such as training status (Martin et al., 2016). Other potential modulators of the inter-individual response to MF have not been elucidated; therefore, any comparison of responses to a MF protocol across different samples should be done with caution.

We found similar increases in mean subjective sensations of fatigue and reductions in mean subjective sensations of vigour following the IST in the MF+MUSIC and MFONLY trials. Almost all individuals demonstrated the expected direction of change in these ratings pre- to post-MF in both trials. Mean ratings of fatigue and vigour were also similar at pre-MF and post-MF in both trials. Taken together, these data indicate that the MF protocol had a similar impact on participants in both trials. Therefore, changes in physical performance can be associated with the presence or absence of music rather than differences in response to the MF protocol.

In study 1, distance covered in MF+MUSIC was moderately greater than MFONLY and very similar to BL. Similarly, in study 2 TT completion time showed a small decrease in MF+MUSIC vs. MFONLY, and was very similar to BL. This improvement in endurance running capacity and performance when listening to self-selected music is consistent with the literature (Buhmann et al., 2018; Terry et al., 2011). However, a novel finding of our study is that listening to music while mentally fatigued negates the negative impact of MF on endurance running capacity and performance.

For mean distance covered during the YYIRT1, our participants fall below the 20% quantile for recreationally active males (Schmitz et al., 2018). However, it is difficult to accurately compare test outcomes, as "*recreationally active*" can be defined differently between studies(Schmitz et al., 2018). Thomas et al. (2006) reported a coefficient of variation (CV) of 8.7% for the YYIRT1 in recreational team sport players. If we apply this CV to the BL distance covered in the current study, then a minimum observable change in performance would be ± 47.9 m. This suggests that the reduction in distance covered in MFONLY vs. BL and improvement in distance covered in MF+MUSIC vs. MFONLY was meaningful. Driller et al. (2017) reported a largest CV for 5 km treadmill TT performance in well-trained runners of 3.9%. Using this figure in the current study, a minimum observable change in TT performance would be ± 0.91 min. This would indicate that while the improvement in TT performance in MF+MUSIC vs. MFONLY was meaningful, the difference between BL and

MFONLY may have been due to normal test variability. However, this is difficult to confirm without having a CV specific to the population sampled from in the current study. That every participant performed the TT more slowly in MFONLY vs. BL suggests there was a negative influence of MF on TT performance.

The impact of MF on endurance performance was smaller than that of endurance capacity, and smaller than the 5.5% decrement in 5 km treadmill performance reported by Pageaux et al. (2014). However, Pageaux et al. (2014) used a modified version of the IST, which may have been more mentally taxing. The authors also recruited recreationally active participants, whereas we recruited recreational runners. More experienced and/or trained individuals may resist the negative effects of MF during their sport more effectively than less experience/trained individuals (Martin et al., 2016).

In study 1 end-exercise RPE in MF+MUSIC was lower compared with MFONLY and was trivially different to BL, despite participants exercising for longer in MF+MUSIC vs. MFONLY. Similarly, in study 2, RPE in MF+MUSIC was similar to BL and MFONLY, despite participants running faster and maintaining a greater physiological load (HR) in the MF+MUSIC trial. Taken together, our data indicates that when listening to music in a mentally fatigued state, participants are able to A) run for longer, maintain a faster speed and greater physiological load for a given RPE compared to a mentally fatigued state with no music, and B) display endurance capacity and performance almost identical to that observed when no MF is present. Rating of perceived exertion is positively correlated with HR during treadmill running (Scherr et al., 2013). Therefore, the data from study 2 and to an extent study 1 support the suggestion that listening to music during exercise can distract from the interoceptive signals arising from exercise, dissociating the participant from perception of these signals and thereby reducing RPE (Karageorghis and Priest, 2012a). To our knowledge, this is the first study to demonstrate such a role for music during exercise in a mentally fatigued state.

However, the distraction explanation is more plausible at intensities <70% VO_{2max} (Karageorghis and Priest, 2012b). While we did not measure VO_2, it is safe to assume that participants in both studies would have exceeded 70% VO_{2max}. The involvement of the ACC in cognitive impairment following MF (Martin et al., 2018), generation of RPE during exercise (Williamson et al., 2001), and music processing during exercise (Chanda and Levitin, 2013) suggests that music may have elicited its positive effect on endurance capacity and performance in a mentally fatigue state by improving affect and/or motivation.

Our findings should be interpreted in light of limitations. In study 1, it would have been useful to record RPE and HR during the YYIRT1. Both studies could have also assessed motivation to exercise following the MF task, both before and after music began to play, to elucidate the effect of MF and music on motivation for exercise.

CONCLUSION

We present novel data demonstrating that listening to music in a mentally fatigued state can restore endurance running capacity and performance to a similar level as observed when no MF is present. This effect appears to be mediated by attenuations in RPE which may allow maintenance of a given exercise intensity for longer, and the attainment of higher exercise intensity for a given RPE.

AUTHOR CONTRIBUTIONS

Conceptualization: Shaun M Phillips. Methodology: Hui Kwan Nicholas Lam, Harry Middleton, Shaun M. Phillips. Formal analysis and investigation: Hui Kwan Nicholas Lam, Harry Middleton, Shaun M. Phillips. Data

collection: Hui Kwan Nicholas Lam, Harry Middleton. Writing – original draft preparation: Hui Kwan Nicholas Lam, Harry Middleton, Shaun M. Phillips Writing – review and editing: Shaun M Phillips.

SUPPORTING AGENCIES

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

Algina, J., & Keselman, J. (2003). Approximate confidence intervals for effect sizes. Educ Psychol Meas, 63, 537-553. https://doi.org/10.1177/0013164403256358

Bangsbo, J., Iaia, F. M., & Krustrup, P. (2008). The Yo-Yo intermittent recovery test : a useful tool for evaluation of physical performance in intermittent sports. Sports Med, 38(1), 37-51. https://doi.org/10.2165/00007256-200838010-00004

Brandt, R., Herrero, D., Massetti, T., Crocetta, B., Guarnieri, R., Monteiro, C., . . . Andrade, A. (2016). The Brunel Mood Scale Rating in Mental Health for Physically Active and Apparently Healthy Populations. Health, 8, 125-132. https://doi.org/10.4236/health.2016.82015

Buhmann, J., Moens, B., Van Dyck, E., Dotov, D., & Leman, M. (2018). Optimizing beat synchronized running to music. PLoS One, 13(12), e0208702. https://doi.org/10.1371/journal.pone.0208702

Calin-Jageman, R., & Cumming, G. (2019). The new statistics for better science: ask how much, how uncertain, and what else is known. Am Stat, 73(73), 271-280. https://doi.org/10.1080/00031305.2018.1518266

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. Trends Cogn Sci, 17(4), 179-193. https://doi.org/10.1016/j.tics.2013.02.007

Clark, I., Goulding, R., DiMenna, F., Bailey, S., Jones, M., Fulford, J., . . . Vanhatalo, A. (2019). Time-trial performance is not impaired in either competitive athletes or untrained individuals following a prolonged cognitive task. Eur J Appl Physiol, 119, 149-161. https://doi.org/10.1007/s00421-018-4009-6

Cohen, J. (1992). A power primer. Psychol Bull, 112(1), 155-159. https://doi.org/10.1037//0033-2909.112.1.155

Cumming, G. (2012). Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis. New York: Routledge. https://doi.org/10.4324/9780203807002

Driller, M., Brophy-Williams, N., & Walker, A. (2017). The reliability of a 5 km run test on a motorized treadmill. Meas Phys Educ Exerc Sci, 21(3), 121-126. https://doi.org/10.1080/1091367X.2016.1264402

Dyer, B., & McKune, A. (2013). Effects of music tempo on performance, psychological, and physiological variables during 20 km cycling in well-trained cyclists. Percept Mot Skills, 117, 484-497. https://doi.org/10.2466/29.22.PMS.117x24z8

Hutchinson, J. C., & Karageorghis, C. I. (2013). Moderating influence of dominant attentional style and exercise intensity on responses to asynchronous music. J Sport Exerc Psychol, 35(6), 625-643. https://doi.org/10.1123/jsep.35.6.625

Karageorghis, C. I., Mouzourides, D. A., Priest, D. L., Sasso, T. A., Morrish, D. J., & Walley, C. J. (2009). Psychophysical and ergogenic effects of synchronous music during treadmill walking. J Sport Exerc Psychol, 31(1), 18-36. https://doi.org/10.1123/jsep.31.1.18

Karageorghis, C. I., & Priest, D. L. (2012a). Music in the exercise domain: a review and synthesis (Part I). Int Rev Sport Exerc Psychol, 5(1), 44-66. https://doi.org/10.1080/1750984X.2011.631026

Karageorghis, C. I., & Priest, D. L. (2012b). Music in the exercise domain: a review and synthesis (Part II). Int Rev Sport Exerc Psychol, 5(1), 67-84. https://doi.org/10.1080/1750984X.2011.631027

Karageorghis, C. I., Priest, D. L., Terry, P. C., Chatzisarantis, N. L., & Lane, A. M. (2006). Redesign and initial validation of an instrument to assess the motivational qualities of music in exercise: the Brunel Music Rating Inventory-2. J Sports Sci, 24(8), 899-909. https://doi.org/10.1080/02640410500298107

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol, 4, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lee, K. A., Hicks, G., & Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. Psychiatry Res, 36(3), 291-298. https://doi.org/10.1016/0165-1781(91)90027-M

Marcora, S. M., Staiano, W., & Manning, V. (2009). Mental fatigue impairs physical performance in humans. J Appl Physiol (1985), 106(3), 857-864. https://doi.org/10.1152/japplphysiol.91324.2008

Martin, K., Meeusen, R., Thompson, K. G., Keegan, R., & Rattray, B. (2018). Mental Fatigue Impairs Endurance Performance: A Physiological Explanation. Sports Med, 48(9), 2041-2051. https://doi.org/10.1007/s40279-018-0946-9

Martin, K., Staiano, W., Menaspa, P., Hennessey, T., Marcora, S., Keegan, R., . . . Rattray, B. (2016). Superior Inhibitory Control and Resistance to Mental Fatigue in Professional Road Cyclists. PLoS One, 11(7), e0159907. https://doi.org/10.1371/journal.pone.0159907

Martin, K., Thompson, K., Keegan, R., Ball, N., & Rattray, B. (2015). Mental fatigue does not affect maximal anaerobic exercise performance. Eur J Appl Physiol, 115, 715-725. https://doi.org/10.1007/s00421-014-3052-1

Moss, S., Enright, K., & Cushman, S. (2018). The influence of music genre on explosive power, repetitions to failure and mood responses during resistance exercise. Psychol Sport Exerc, 37, 128-138. https://doi.org/10.1016/j.psychsport.2018.05.002

Nakamura, P. M., Pereira, G., Papini, C. B., Nakamura, F. Y., & Kokubun, E. (2010). Effects of preferred and nonpreferred music on continuous cycling exercise performance. Percept Mot Skills, 110(1), 257-264. https://doi.org/10.2466/pms.110.1.257-264

Pageaux, B., Lepers, R., Dietz, K. C., & Marcora, S. M. (2014). Response inhibition impairs subsequent self-paced endurance performance. Eur J Appl Physiol, 114(5), 1095-1105. https://doi.org/10.1007/s00421-014-2838-5

Pageaux, B., Marcora, S., Rozand, V., & Lepers, R. (2015). Mental fatigue induced by prolonged self-regulation does not exacerbate central fatigue during subsequent whole-body endurance exercise. Front Hum Neurosci, 9, 1-12. https://doi.org/10.3389/fnhum.2015.00067

Scherr, J., Wolfarth, B., Christle, J., Pressler, A., Wagenfeil, S., & Halle, M. (2013). Associations between Borg's rating of perceived exertion and physiological measures of exercise intensity. Eur J Appl Physiol, 113, 147-155. https://doi.org/10.1007/s00421-012-2421-x

Schmitz, B., Pfeifer, C., Kreitz, K., Borowski, M., Faldum, A., & Brand, S. M. (2018). The Yo-Yo Intermittent Tests: A Systematic Review and Structured Compendium of Test Results. Front Physiol, 9, 870. https://doi.org/10.3389/fphys.2018.00870

Smith, M. R., Coutts, A. J., Merlini, M., Deprez, D., Lenoir, M., & Marcora, S. M. (2016). Mental Fatigue Impairs Soccer-Specific Physical and Technical Performance. Med Sci Sports Exerc, 48(2), 267-276. https://doi.org/10.1249/MSS.0000000000000762

Smith, M. R., Marcora, S. M., & Coutts, A. J. (2015). Mental Fatigue Impairs Intermittent Running Performance. Med Sci Sports Exerc, 47(8), 1682-1690. https://doi.org/10.1249/MSS.0000000000000592

Stork, M. J., Kwan, M. Y., Gibala, M. J., & Martin Ginis, K. A. (2015). Music enhances performance and perceived enjoyment of sprint interval exercise. Med Sci Sports Exerc, 47(5), 1052-1060. https://doi.org/10.1249/MSS.0000000000000494

Terry, P., Karageorghis, C. I., Saha, A., & D'auria, S. (2011). Effects of synchronous music on treadmill running among elite triathletes. J Sci Med Sport, 15, 52-57. https://doi.org/10.1016/j.jsams.2011.06.003

Thomas, A., Dawson, B., & Goodman, C. (2006). The yo-yo test: reliability and association with a 20-m shuttle run and VO(2max). Int J Sports Physiol Perform, 1(2), 137-149. https://doi.org/10.1123/ijspp.1.2.137

Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. Psychol Schools, 44(5), 423-432. https://doi.org/10.1002/pits.20234

Van Cutsem, J., Marcora, S., De Pauw, K., Bailey, S., Meeusen, R., & Roelands, B. (2017). The Effects of Mental Fatigue on Physical Performance: A Systematic Review. Sports Med, 47(8), 1569-1588. https://doi.org/10.1007/s40279-016-0672-0

Wasserstein, R., Schirm, A., & Lazar, N. (2019). Moving to a World Beyond "p < 0.05". Am Stat, 73, 1-19. https://doi.org/10.1080/00031305.2019.1583913

Williamson, J. W., McColl, R., Mathews, D., Mitchell, J. H., Raven, P. B., & Morgan, W. P. (2001). Hypnotic manipulation of effort sense during dynamic exercise: cardiovascular responses and brain activation. J Appl Physiol (1985), 90(4), 1392-1399. https://doi.org/10.1152/jappl.2001.90.4.1392