ICA 2 Investigation

Student Numbers: 19095174, 19076794

1 Abstract

To investigate the effect of music and time-of-day on running speeds and their potential interaction effect. Runners were given randomised schedules to run for 2km with or without fast music (>120 beats-per-second), during the morning or evening (8am or 4pm) and to record the time per kilometre (TPK). Each operator ran two instances of each combination. A fully crossed linear ANOVA model was used to estimate treatment and interaction effects. The initial ANOVA analysis deemed interaction effect to be statistically insignificant (p = 0.372), and hence was dropped. Subsequently, the reduced ANOVA found significant blocking effect (F = 204.034, p = 2e-16) and music effect (- 0.615 minutes, 95% CI: -0.873231, -0.3563524), while morning effect (+ 0.171 minutes, 95% CI: -0.08781429, 0.4290643) was insignificant at $\alpha = 0.05$.

2 Introduction

Running speeds may be influenced by a number of factors, such as time-of-day and listening to music concurrently. It is of athlete's and casual runner's interest to see how one could potential improve running performance.

The benefits of music on running speeds have been demonstrated in multiple observational and experimental studies. Music can increase exercise intensity [1], lower fatigue by reducing perceived exertion [2] and spur motivation [3]. 'Fast Music' defined as a tempo of 125-145 beats/minute [4] is theorised to increase performance through mechanisms such as inducing arousal or increasing pacing [1].

Physiological functions follow a 24h circadian rhythm including body temperature and lung function [5]. It is conceivable that time of day impacts running speeds, although there is scant evidence in the literature around this. One study showed that running in the morning (0700-0900) significantly elevated the rating of perceived exertion compared to running in the evening (1800-2100) [6].

Through a fully crossed experimental design, we aim to investigate and quantify how music and the morning effect affect running performance, particularly the less researched "morning effect", as well as a possible interaction that we believe remains unstudied in existing literature.

3 Design and Data

A three-way crossed factorial design with two factors and a blocking variable was implemented, where interactions between the two factors can be studied. For simplicity of this smaller scale experiment, we assume no interaction between the blocker and the two factors, preventing the strain of estimating too many parameters for given sample size.

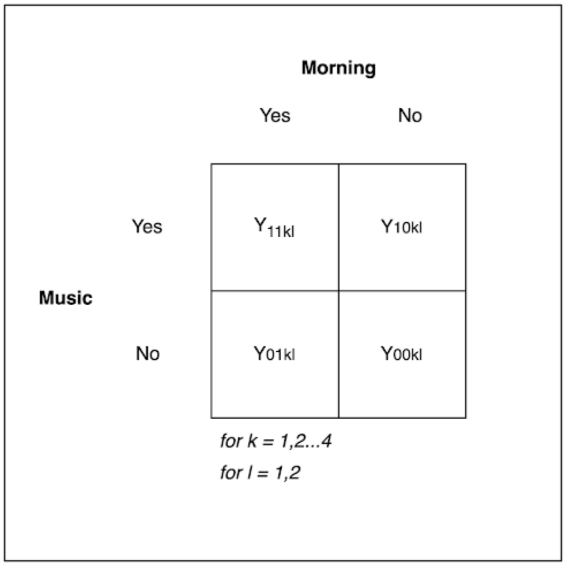


Figure 1: A Visualisation of design: The four types of data that the experiment aims to collect

3.1 Treatments and Outcome

Treatment 1: ‘Music’

Table 1: ‘Music’ Treatment Levels

Fast music [4]	Music above 120 BPM (defined as Music = 1)
No Music	Running without music (defined as Music = 0)

Spotify provided music used in the experiment for instance: “High Temp - Run” playlist created by “Running150BPM” . The BPMs of the songs were cross-verified using jog.fm (https://jog.fm/workout-songs/). Volume was standardised by making operators play music at full volume, preventing music volume having a confounding effect.

Treatment 2: ‘Morning’

Table 2: ‘Morning’ Treatment Levels

Morning	8:00 hr (defined as Morning = 1)
Evening	16:00 hr (defined as Morning = 0)

Definitions were borrowed from a similar study [7], focusing instead on resistance-exercise performance. For flexibility we allowed for a 1.5 hours window to complete the run.

Our outcome of interest is running pace, measured in time per kilometre (TPK in minutes).

3.2 Model

The statistical model is given as follows:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + \epsilon_{ijkl}$$

Where:

$$\alpha_i = \begin{cases} \alpha_1, & \text{if music listened to} \\ \alpha_0, & \text{otherwise} \end{cases} \qquad \beta_j = \begin{cases} \beta_1, & \text{if ran during morning} \\ \beta_0, & \text{otherwise} \end{cases}$$

- $\gamma_i = \gamma_1, \gamma_2...$ for operator i, representing operator effects
- indices : $i, j = 0, 1, k = 1, 2, 3, 4, l = 1, 2$ (repetition number)
- μ = Grand Mean
- ϵ_{ijkl} = Error term for operator k, music = i, morning = j, corresponding repitition = l
- $(\alpha\beta)_{ij}$ = interaction term between α_i and β_j
- assume $\epsilon_{ijkl} \sim (iid)\ N(0, \sigma^2)$

Table 3: Defining Running Types

Running Type	Treatment
(1)	(Music, Morning)
(2)	(No Music, Morning)
(3)	(Music, Evening)
(4)	(No Music, Evening)

The model is a 3-way linear ANOVA model with one set of interactions. To include treatment effects for all levels we use the "sum to 0" constraint i.e. $\alpha. = \beta. = 0$. We interpret the effects at "deviations from grand mean" and allows identifiability of the coefficients.

Unless we reject the interaction effect, 'main effects' of music and morning are interpreted as the 'effect of one factor, averaged over the other'.

3.3 Method: Detailing Experimentation

Operators were each instructed to run 2km. To mitigate potential psychological bias from awareness of treatments, we instructed operators to run as fast as possible. We also instructed that operators had consistent amount of sleep, same running routes and gear (fixed variables).

Data was collected using Runkeeper or Strava, which have the same functionality of starting and stopping the time, keeping track of their route, distance and speed [8]. Timings are made using fitness apps of the operator's choice. After running, the participant reported their running pace (TPK) see figure 2:

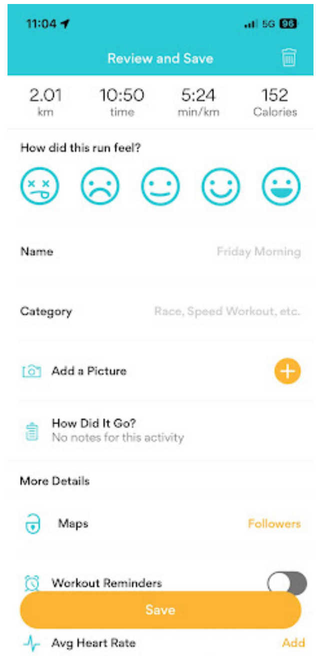


Figure 2: Using Runkeeper mobile application to log running pace - TPK = 5:24 is directly measured on the interface

3.4 Randomising Over Fitness

The main potential confounding variable was fitness levels which may vary over the week depending on schedule intensity for a given day. We solved this with the help of R to pre-determine running days by randomising intervals between runs using `round(runif(8, min = 1, max = 3))` (Table 2), requiring participants to decide on their starting date to not clash with their schedules. We then randomised treatment levels by `c(sample((1:4), replace = F), sample((1:4), replace = F))` (sampling without replacement) to avoid clustering as much as possible. These layers of randomisation helped with other confounders such as weather and temperature. Environmental conditions may affect running performance, for example a paper on marathon running performance [9] showed temperature increases above optimal temperatures led to slower running times and higher marathon withdrawal rates.

3.5 Blocking Variable

Fitness will vary from operator to operator, thus performance will vary according to background variables like running frequency, age, height and weight. We used blocking to account for said differences. Blocking for operators adsorbs other operator-specific traits such as running routes. Pacing strategies and motivation are also accounted for through operator blocking. A study of the impact of pacing strategies on five kilometre running time trial performance [10] showed a change in strategy can optimise running times, thus the need for blocking. Using such blocks can increase precision in estimating music and morning effects.

3.6 Fixed Effects:

We assumed fixed effects for our variables, due to the trivial nature of generalising from treatment levels of our factors, making out ANOVA analysis more directly interpretable.

3.7 Interaction Terms:

We initially included an interaction effect between music and morning variables. This is not only for simplicity of analysis but we are also interested in the general/average effects of morning or music, and how these may interact.

3.8 Sample size determination:

To estimate the models variance of error term σ^2, we need at least m=1 replicates per cell. We aimed for a statistically significant difference in TPK of 30 seconds. We used estimated variance in calculations from a paper [11] examining impact of training regiments on 2000m running performance. The paper found a standard deviation of 59 seconds for level training, a viable estimate for our experiment.

Table 4: Defining Sources of Variation and Degrees of Freedom

Source of Variation	Calculation	Degrees of Freedom
Music (Treatment 1) r1 levels	r1-1	1
Morning (Treatment 2) r2 levels	r2-2	1
Blocking (Operators)	b-1	3
Interaction between Music:Morning	(r1-1)(r2-1)	1
Within Cell (Error)	(by subtraction)	25
Total: m replicates per cell	(m)(r1)(r2)(b) - 1	31

3.8.1 Determining half confidence interval of measure effect when m=2

For 2 groups (Music vs No Music): Half confidence interval $= t_{\frac{\alpha}{2}} * s * \sqrt{\frac{2}{b'}}$, where s =error variance estimate, b' = no. samples for music only (balanced design)

Let significance level be $\alpha = 0.1$ $b' = 4 * 2 * 2 = 16$ (4 runners, morning, evening, 2 replicates) s = 1 (minutes, based on studies on similar distances). Result of half interval is 36.24. Thus m = 2 captures a difference of 36.24 seconds between music and no music with statistical significance. Due to cost and time constraints we settled on m = 2. We opted for a balanced design; all operators running same 8 runs to maximise power and estimate music and morning effects with equal precision, resulting in a total sample size of n=32.

Table 2: Running schedule

Below is the running schedule determined from spacing intervals to randomise over fitness. Before the experiment, we had trial runs that were initially used to get a realistic sense of times to expect, also highlighting potential problems.

Table 5: Running schedule: Intervals and Running Type

		Runs							
Operator	Interval / Run Type	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8
Operator 1	**Interval(days)**	**1**	**2**	**2**	**1**	**2**	**1**	**1**	**0**
Operator 1	Running Type	2	1	4	3	2	4	1	3
Operator 2	**Interval(days)**	**3**	**2**	**2**	**2**	**1**	**2**	**2**	**0**
Operator 2	Running Type	1	4	3	2	3	2	1	4
Operator 3	**Interval(days)**	**2**	**3**	**1**	**3**	**1**	**2**	**3**	**0**
Operator 3	Running Type	4	2	3	1	1	3	4	2
Operator 4	**Interval(days)**	**3**	**1**	**1**	**2**	**3**	**2**	**2**	**0**
Operator 4	Running Type	2	3	4	1	4	3	1	2
Operator 5	**Interval(days)**	**2**	**3**	**2**	**3**	**2**	**2**	**2**	**0**
Operator 5	Running Type	4	3	1	2	1	2	3	4

Note:

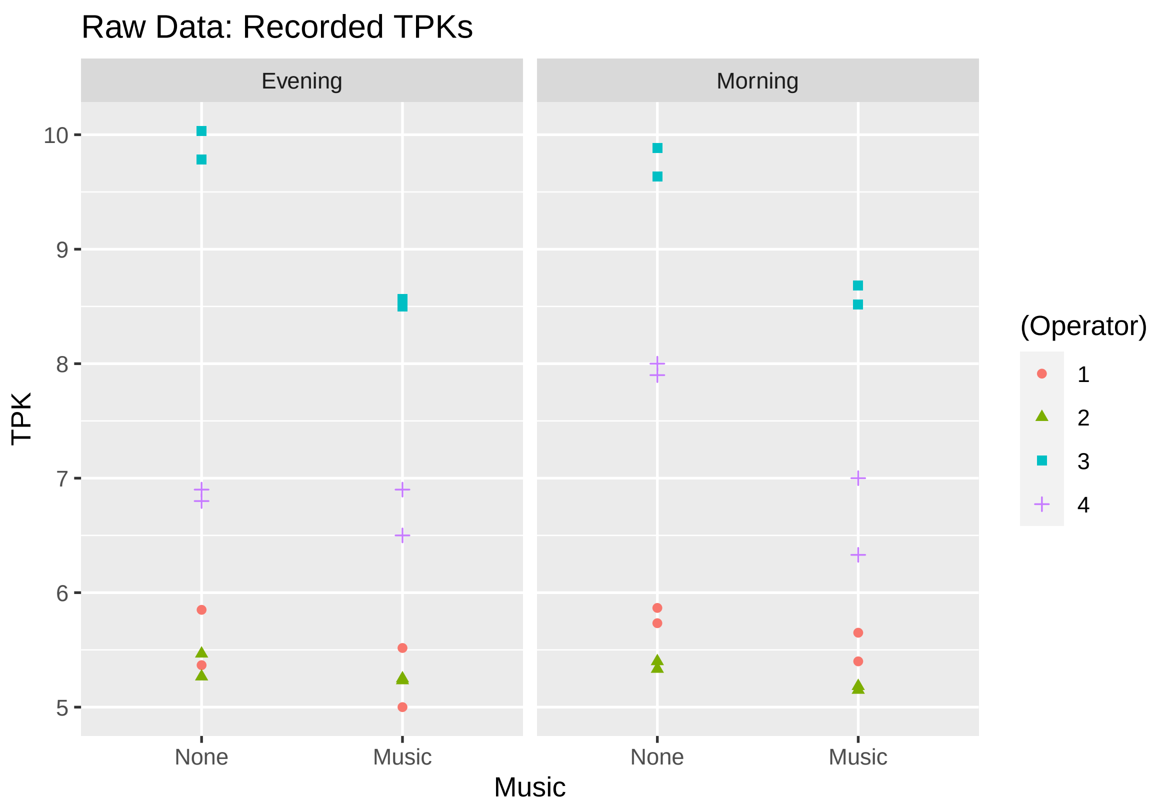
Runners have to decide on the date of their first run

Interval = number of days of rest between runs, e.g. 'Run 1 = 2' means 2 days of rest between Run 1 and Run 2

Running Type: 1,2,3,4 correspond to (Music, Morning), (No Music, Morning), (Music, Evening), (No Music, Evening) respectively

4 Analysis and Discussion

4.1 Initial Data Analysis



'A brief loop at TPKs across `Music` shows that timings are generally lower with Music, but no clear visual relationship can be spotted for `Time-of-day`. A jitterplot for operator run times shows clear disparity of running speeds for operators across all 4 combinations. Each operator has a different spread of points (variance), violating the constant variance ANOVA assumption, but mitigated as we have equal observations across each categorical type.

```
##               Df Sum Sq Mean Sq F value   Pr(>F)    
## Operator       3  77.41  25.803 202.685  < 2e-16 ***
## Music          1   3.02   3.024  23.752 5.17e-05 ***
## Morning        1   0.23   0.233   1.830    0.188    
## Music:Morning  1   0.11   0.105   0.828    0.372    
## Residuals     25   3.18   0.127                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## Call:
## aov(formula = TPK ~ Operator + Music * Morning, data = data1)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -0.4318 -0.3000 -0.0001  0.2495  0.6686 
## 
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                 5.7126     0.1669  34.233  < 2e-16 ***
## Operator2                  -0.2625     0.1784  -1.471  0.15365    
## Operator3                   3.6521     0.1784  20.472  < 2e-16 ***
## Operator4                   1.4933     0.1784   8.371 1.02e-08 ***
## MusicMusic                 -0.5000     0.1784  -2.803  0.00965 ** 
## MorningMorning              0.2854     0.1784   1.600  0.12219    
## MusicMusic:MorningMorning  -0.2296     0.2523  -0.910  0.37152    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.3568 on 25 degrees of freedom
## Multiple R-squared:  0.9621, Adjusted R-squared:  0.953 
## F-statistic: 105.7 on 6 and 25 DF,  p-value: < 2.2e-16
```

Using results above, the hypothesis test p-value for the presence of an interaction is 0.372, thus insufficient evidence from our dataset to reject the null hypothesis that an interaction does not exist. We proceed with the analysis without said interaction term, also improving interpretability of the main effects Blocking is significant as expected, a large F value of 202.685 and a p-value effectively 0 from the F(3,15) distribution. Hence blocking fulfils the purpose of 'removing some sum of squares error', improving precision of the other factor estimates.

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## Operator     3  77.41  25.803 204.034  < 2e-16 ***
## Music        1   3.02   3.024  23.910 4.49e-05 ***
## Morning      1   0.23   0.233   1.842    0.186    
## Residuals   26   3.29   0.126                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
```

```
## Call:
## aov(formula = TPK ~ Operator + Music + Morning, data = data1,
##     contrasts = list(Operator = "contr.sum", Music = "contr.sum",
##         Morning = "contr.sum"))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4892 -0.2822 -0.0001  0.2990  0.6112
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.76865    0.06286 107.671  < 2e-16 ***
## Operator1   -1.22073    0.10888 -11.211 1.86e-11 ***
## Operator2   -1.48323    0.10888 -13.622 2.40e-13 ***
## Operator3    2.43135    0.10888  22.330  < 2e-16 ***
## Music1       0.30740    0.06286   4.890 4.49e-05 ***
## Morning1    -0.08531    0.06286  -1.357    0.186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3556 on 26 degrees of freedom
## Multiple R-squared:  0.9608, Adjusted R-squared:  0.9533
## F-statistic: 127.6 on 5 and 26 DF,  p-value: < 2.2e-16

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = TPK ~ Operator + Music + Morning, data = data1, contrasts = list(Operator = "co
##
## $Operator
##          diff        lwr        upr     p adj
## 2-1 -0.262500 -0.7502826  0.2252826 0.4655823
## 3-1  3.652083  3.1643008  4.1398659 0.0000000
## 4-1  1.493333  1.0055508  1.9811159 0.0000000
## 3-2  3.914583  3.4268008  4.4023659 0.0000000
## 4-2  1.755833  1.2680508  2.2436159 0.0000000
## 4-3 -2.158750 -2.6465326 -1.6709674 0.0000000
##
## $Music
##                  diff       lwr        upr    p adj
## Music-None -0.6147917 -0.873231 -0.3563524 4.49e-05
##
## $Morning
##                     diff         lwr       upr     p adj
## Morning-Evening 0.170625 -0.08781429 0.4290643 0.1864187
```

Estimated factorial effect of music (twice the treatment effect alpha estimate from the model) is a decrease of 0.615 minutes in TPK. (95% CI: -0.873231, -0.3563524), statistically significant for $\alpha = 0.05$. This supports the current stance in literature that music improving running performance. Estimated factorial effect of morning is a 0.171 minute increase, although statistically insignificant for the same confidence level. This is likely due to the small sample size which results in a large standard error relative to the calculated morning effect.

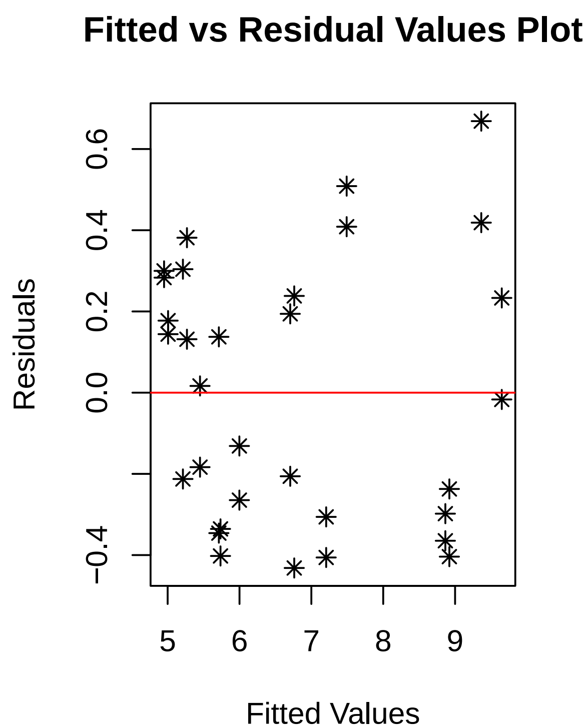
As the distance recommended and ran by participants was only 2km, the main factor effects may be less pronounced due to relatively short distance. The effects could be more noticeable using a larger distance, providing potentially larger, more significant estimates.

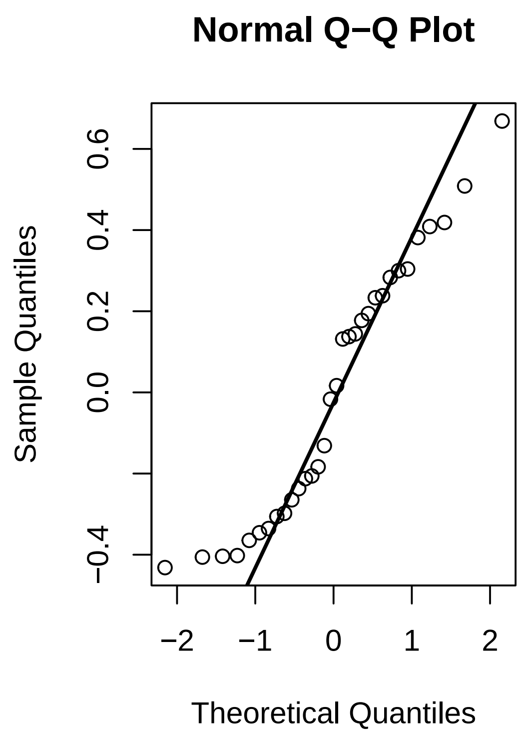
4.2 Conclusions

Factorial effect estimates, calculated as contrast of treatment means, allow interpretation of main effects as the overall effect averaged across all other independent variables. Operator variables have the greatest values in magnitude as expected due to initial fitness levels. Music has a relatively large value in magnitude compared to the morning variable. Morning1 has a negative coefficient, indicating that runs in the evening result in lower TPK compared morning runs. Music1 has a positive coefficient, thus suggesting implementing runs without music result in higher TPK. Outcomes suggest improved performance with music, and runs later in the day.

4.3 Model Assumptions / Diagnostic Plots

```
## integer(0)
```





We see from the fitted v residual plot there is no obvious pattern or correlation, exhibiting preferred behaviour with no residuals standing out as clear outliers, and points 'bouncing randomly' around the 0 line. Observing the QQ plots, given that we are working with a small sample, randomness tends to obscure interpretability, however we can observe light tailling off in values indicating some departure from normality.

4.4 Study Limitations

It is necessary to consider that some of the difference in TPK may be due to unobserved placebo or psychological factors. Runners aware of the existing literature on benefits of music on running may be primed to run faster, biasing the estimates for the factor effects.
The study can benefit from larger sample size, to enhance power and precision of factor effect estimates, allowing to narrow down standard error and detect smaller differences in TPK significantly, especially in our case for the Morning effect.

A 3 way interaction including blockers, or 2 way interaction involving the blocking variable may also improve the precision of factor effects estimation and model effectiveness. Further studies may also look into the interaction between music/morning and initial fitness levels, or focus on operators with closer baseline fitness levels.

Some factors were not accounted for in study design, such as different route elevation gain/loss [11], contributing in running performance variability, as downhill routes require less energy resulting in faster routes. Not all operators ran in the same country; some in Singapore, others in England. The different temperatures could have significant unobserved impacts on the performance. Controlling for these environmental factors, or including in the model would improve study design.

Word Count: 1959

5 Bibliography

Note: For statistical theory mentioned throughout article, we referenced [0].

[0]. Van Den Hout, A. (2022). STAT0029. Statistical Design of Investigations – Lecture Notes. UCL. Accessed from: https://moodle.ucl.ac.uk/course/view.php?id=29634.

[1]. Edworthy, J., & Waring, H. (2006). The effects of music tempo and loudness level on treadmill exercise. Ergonomics, 49(15), 1597–1610. https://doi.org/10.1080/00140130600899104

[2]. Yamashita, S., Iwai, K., Akimoto, T., Sugawara, J., & Kono, I. (2006). Effects of music during exercise on RPE, heart rate and the autonomic nervous system. The Journal of sports medicine and physical fitness, 46(3), 425–430.

[3]. Priest, D. L., Karageorghis, C. I., & Sharp, N. C. (2004). The characteristics and effects of motivational music in exercise settings: the possible influence of gender, age, frequency of attendance, and time of attendance. The Journal of sports medicine and physical fitness, 44(1), 77–86.

[4]. Lee, Sam & Kimmerly, Derek. (2014). Influence of music on maximal self-paced running performance and passive post-exercise recovery rate. The Journal of sports medicine and physical fitness. 56.

[5]. Medarov, B. I., Pavlov, V. A., & Rossoff, L. (2008). Diurnal variations in human pulmonary function. International journal of clinical and experimental medicine, 1(3), 267–273.

[6]. Martin, L., Doggart, A. L., & Whyte, G. P. (2001). Comparison of physiological responses to morning and evening submaximal running. Journal of sports sciences, 19(12), 969–976. https://doi.org/10.1080/026404101317108471

[7]. Blazer, H. J., Jordan, C. L., Pederson, J. A., Rogers, R. R., Williams, T. D., Marshall, M. R., & Ballmann, C. G. (2021). Effects of Time-of-Day Training Preference on Resistance-Exercise Performance. Research quarterly for exercise and sport, 92(3), 492–499. https://doi.org/10.1080/02701367.2020.1751032

[7]. Halbeisen C., (2022). Runkeeper: The Basics. Accessed from: https://runkeeper.com/cms/app/runkeeper-the-basics/

[8]. Helou N, Tafflet M, Berthelot G, Tolaini J, Marc A, Guillaume M, et al. (2012) Impact of Environmental Parameters on Marathon Running Performance. PLoS ONE 7(5): e37407. https://doi.org/10.1371/journal.pone.0037407

[10]. Gosztyla, Amy & Edwards, David & Quinn, Timothy & Kenefick, Robert. (2006). The Impact of Different Pacing Strategies on Five-Kilometer Running Time Trial Performance. Journal of strength and conditioning research / National Strength & Conditioning Association. 20. 882-6. 10.1519/R-19275.1.

[11]. Gallagher, Jacquilyn L., “The Effect of Downhill Running Training on 200m and 2000m Running Performance” (1991). Masters Theses. 2235. https://thekeep.eiu.edu/theses/2235