Studying the Combined Psychological Effect of Exercise, Music and Food Consumption Through Self-Experimentation

Grant Lee (grant_lee@brown.edu)

Introduction

A common motivation of self-experimentation (N = 1) is for personal informatics. In doing so, we hope to discover what parts of our lives matter to us the most (Daskalova et al., 2017). By studying data associated with these variables, we can attempt to identify causal relationships between variables in our lives.

Exercising, listening to music and eating are activities known to release “happy hormones”, specifically serotonin, dopamine and endorphins. For my self-experiment, I was interested in combining the feel-good effects of these activities. I seeked to explore how concurrently doing physical exercise and listening to music, followed by having breakfast can have an effect on my psychological state. I made the assumption that such an effect (if there was one) would be immediate and short, and would not last the rest of the day. As such, I came up with the following hypothesis:

> *On days where I do 60 pushups and situps while listening to my favorite playlist right before breakfast, my average mood for the next 2 hours will increase by 0.5 point on a 5-point happiness scale, but readiness will remain the same each day.*

Related work

The concept of doing physical exercise between waking up and having breakfast was inspired by a workout routine that I had done while previously serving in the Singapore Army (Choo, 2021). The aerobic workout, known as 5BX (five basic exercises) was developed in the late 1960s with the intended purpose of warming up a soldier’s body for the activities of the day. For my experiment, I was interested in replicating a process that was similarly structured and disciplined.

It has also been found that aerobic exercise that was self-paced increased motivation (Elliot et al., 2007). Following this, I decided to design my experiment such that the physical exercise involved was not a timed workout, but instead a self-paced one. This should allow me to maximize the effectiveness of the exercise on my psychological state.

I also applied some self-experiment guidelines proposed by Daskalova et al. (2017) to my experiment design. In particular, my experiment comprises the first two stages (Exploration and

Preliminary Hypothesis Testing), and the physical activity is designed to have minimal carryover effect. For statistical analysis, I used a difference-of-means approach by performing two variations of a paired T-test: a one-tailed test to test the *Mood* part of the hypothesis, and a two-tailed test to test for the *Readiness* variable.

Method

Experiment setup

The experiment lasted 12 consecutive days. The independent variable is a binary value – whether or not I do the activity:

- *Yes* = Concurrently doing 60 pushups and situps while listening to my favorite playlist, immediately before having breakfast
- *No* = Simply having breakfast

A coin flip is done at the start of every day to determine whether the activity is done or not.

Two dependent variables are measured:

- *Mood*, using a 5-point scale (0 being of worst mood, 5 being of best mood)
- *Readiness*, recorded as expected number of remaining productive hours

These dependent variables are measured at six different times in the day:

(A) Immediately before breakfast (and physical activity)
(B) Immediately after breakfast
(C) 1 hour after breakfast
(D) 2 hours after breakfast
(E) 4 hours after breakfast
(F) 8 hours after breakfast

Note that only readings (A) to (D) are technically of significance in relation to the hypothesis. Nevertheless, readings (E) and (F) are also recorded in order to study whether the effect (if there is one) also applies beyond the 2-hour threshold.

Hypothesis testing

The data collected was classified into two groups according to the independent variable: (1) the *Activity?* = *Yes* group, and (2) the *Activity?* = *No* group. Each group had 6 sets of readings (A) to (F). The readings for each of the dependent variables were then averaged across the group.

Given the small sample size of 6 for each group, I used a paired T-test to compare the means of both groups with a 95% confidence interval. For the *Mood* variable, this was a one-tailed test, since our hypothesis posits an increase in *Mood* value. For the *Readiness* variable, the test was two-tailed, since our hypothesis posits that there would be no change in the *Readiness* value.

Notably, to perform further statistical analysis, I also group and manipulate a subset of readings (A) to (F), thereby deriving a set of values, which can act as proxies for other meaningful variables:

- *Mean(A, B, C, D, E, F)*: Proxy for average mood/readiness in the day
- *Mean(B, C, D) – A*: Proxy for change in average mood/readiness in the following 2 hours, brought about by the physical activity
- *Mean(B, C, D, E, F) – A*: Proxy for change in average mood/readiness in the rest of the day, brought about by the physical activity

Assumptions

The following assumptions are made:

1. I am able to complete 60 pushups and situps.
2. The time taken to complete the workout does not affect the result.
3. How I complete the workout (type of pushups, how many sets of how many repetitions) does not affect the result.
4. The exact choice of songs I listen to, the order I listen to them in and the volume I listen to them at do not have an impact on the result.
5. I do not perform this workout at any other time of the day.
6. I eat breakfast every day, and the experience of doing so does not impact my mood or readiness in the following 2 hours.
7. The activities that I otherwise partake in within 2 hours of the exercise do not affect the mood or readiness.
8. My subjective measures of mood and readiness are generally reliable and unaffected by circumstances external to the experiment.

Results

Dependent variable: Mood

The following table describes the readings for *Mood*, using a 5-point scale, with 0 being of worst mood and 5 being of best mood:

Day	Activity?	(A)	(B)	(C)	(D)	(E)	(F)	*Mean(B,C,D)*	*Mean(B,C,D) – A*
1	No	3	3	3	3	3	3	3	0
2	Yes	3	4	4	4	4	4	4	+1
3	No	4	4	3	3	2	2	3.333	– 0.667
4	Yes	2	3.5	3	3	2	3	3.167	+ 1.167
5	No	3	3	4	4	4	3	3.667	+ 0.667
6	No	3	4	3	3	2	2	3.333	+ 0.333

7	No	3.5	3.5	4	4	3.5	3	3.833	+ 0.333
8	Yes	3.5	4	4	3	3	4	3.667	+ 0.167
9	Yes	3.5	4	4	4	4	4	4	+ 0.5
10	No	4	4	4	4	4	4.5	4	0
11	Yes	4	4	4	4	4	4	4	0
12	Yes	3.5	4	4	3	3	3	3.667	+ 0.167

Table 1: Readings for *Mood*

Using the readings and calculations of ***Mean(B, C, D) – A*** for *Mood*, we can perform a **one-tailed**, paired T-test with the following hypotheses:

- Null hypothesis: On days where I do the activity, my average mood stays the same for the next 2 hours.
- Alternative hypothesis: On days where I do the activity, my average mood for the next 2 hours **increases**.

The T-test derives the following values:

- P-value = 0.164 > 0.05 ***(statistically insignificant)***
- Cohen's D = 0.828 ***(large effect, but irrelevant since statistically insignificant)***

Performing a similar T-test using calculations of ***Mean(B, C, D, E, F) – A*** [i.e. change in average mood in the rest of the day, not just in the following 2 hours] derives a similar result:

- P-value = 0.131 > 0.05 ***(statistically insignificant)***
- Cohen's D = 0.92 ***(large effect, but irrelevant since statistically insignificant)***

Using the readings in Table 1, we can also average over the data points with *Activity? = Yes* (green rows) and *Activity? = No* (red rows) respectively, allowing us to generate following proxy values, which we explore in the Discussion section:

Calculated value of…	**Proxy for…**	***Activity? = Yes***	***Activity? = No***
A	Mood before breakfast/activity	3.25	3.42
Mean(A, B, C, D, E, F)	Average mood in the day	3.58	3.33
Mean(B, C, D)	Average mood in the following 2 hours	3.75	3.53
Mean(B, C, D) – A	**Change in average mood in the following 2 hours**	**+ 0.50**	**+ 0.11**

Mean(B, C, D, E, F)	Average mood in the rest of the day following breakfast/activity	3.65	3.32
Mean(B, C, D, E, F) – A	**Change in average mood in rest of the day following breakfast/activity**	**+ 0.40**	**– 0.10**

Table 2: Meaningful proxy values for *Mood*

Dependent variable: Readiness

The following table describes the readings for *Readiness*, recorded as expected number of remaining productive hours:

Day	Activity?	(A)	(B)	(C)	(D)	(E)	(F)	*Mean(B,C,D)*	*Mean(B,C,D) – A*
1	No	6	6	6	4	2	1	5.333	– 0.667
2	Yes	8	8	7	7	5	1	7.333	– 0.667
3	No	4	4	4	3	2	0	3.667	– 0.333
4	Yes	4	4	4	3	2	1	3.667	– 0.333
5	No	6	6	5	5	4	1	5.333	– 0.667
6	No	8	8	7	5	2	0	6.667	– 1.333
7	No	7	7	6	6	5	2	6.333	– 0.667
8	Yes	6	6	6	5	2	0	5.667	– 0.333
9	Yes	6	6	5	5	4	2	5.333	– 0.667
10	No	5	5	5	4	3	2	4.667	– 0.333
11	Yes	5	5	5	4	2	1	4.667	– 0.333
12	Yes	5	5	4	3	3	0	4	– 1

Table 3: Readings for *Readiness*

Using the readings and calculations of ***Mean(B, C, D) – A*** for *Readiness*, we can perform a **two-tailed**, paired T-test with the following hypotheses:

- Null hypothesis: On days where I do the activity, my average readiness stays the same for the next 2 hours.
- Alternative hypothesis: On days where I do the activity, my average readiness **changes**.

The T-test derives the following values:

- P-value = 0.288 > 0.025 ***(statistically insignificant)***

- Cohen's D = 0.345 ***(small effect, but irrelevant since statistically insignificant)***

Performing a similar T-test using calculations of ***Mean(B, C, D, E, F) – A*** [i.e. change in average readiness in the rest of the day, not just in the following 2 hours] derives a similar result:

- P-value = 0.346 > 0.025 ***(statistically insignificant)***
- Cohen's D = 0.244 ***(small effect, but irrelevant since statistically insignificant)***

Using the readings in Table 3, we can also average over the data points with *Activity? = Yes* (green rows) and *Activity? = No* (red rows) respectively, allowing us to generate following proxy values, which we explore in the Discussion section:

Calculated value of…	**Proxy for…**	***Activity? = Yes***	***Activity? = No***
A	Readiness before breakfast/activity	5.667	6
Mean(A, B, C, D, E, F)	Average readiness in the day	4.139	4.333
Mean(B, C, D)	Average readiness in the following 2 hours	5.111	5.333
Mean(B, C, D) – A	**Change in average readiness in the following 2 hours**	**– 0.556**	**– 0.667**
Mean(B, C, D, E, F)	Average readiness in the rest of the day following breakfast/activity	3.833	4
Mean(B, C, D, E, F) – A	**Change in average readiness in the rest of the day following breakfast/activity**	**– 1.833**	**– 2**

Table 4: Meaningful proxy values for *Readiness*

Discussion

A glance at the ***Mean(B, C, D) – A*** and ***Mean(B, C, D, E, F) – A*** values for both *Mood* and *Readiness* variables would suggest that the physical activity was associated with:

- A greater increase in mood in the following 2 hours (+ 0.50 with activity vs. + 0.11 without activity), as well as for the rest of the day (+ 0.40 with activity vs. – 0.10 without activity)
- A smaller decrease in readiness in the following 2 hours (– 0.556 with activity vs. – 0.667 without activity), as well as for the rest of the day (– 1.833 with activity vs. – 2 without activity)

However, the p-values derived from the T-tests suggest that the hypothesis was inconclusive. In other words, the physical activity:

- Does not necessarily improve my average mood, for the next 2 hours nor for the rest of the day
- Does not change my average readiness, for the next 2 hours nor for the rest of the day

This result was not unexpected, as self-experiments often involve many confounding variables, which make it difficult to control the environment. This is corroborated by Daskalova et al. (2017), who mentions that an inconclusive result probably suggests that the observed effects may be a consequence of randomness, rather than causation.

Two ways to mitigate this would be:

1. To conduct a longer study: A longer study could result in a more consistent pattern, steering the association between the independent variable and dependent variables away from one of randomness, towards one of correlation (or causation).
2. To create a more controlled environment: A more controlled study would reduce the effect of confounding variables that would otherwise introduce a degree of randomness and noise to the association.

That being said, the calculation of Cohen's D value does suggest that physical activity definitely has a greater effect size on the *Mood* variable (0.828) than the *Readiness* variable (0.345), as implied by the original hypothesis.

Our evaluation of the proxy values in Tables 2 and 4 also give us the following learning points:

- The effect of the physical activity on *Mood* and *Readiness* (though invalidated) do not simply last for 2 hours after the activity, but for the rest of the day as well. This was an unexpected finding and suggests that the initial assumption I made while formulating the hypothesis is wrong.
- Absolute values of the *Mood* and *Readiness* variables (i.e. without deducting the initial value *A*) do not carry much meaning alone.

References

1. Choo, D. (2021, Nov 23). Goodbye, 5BX: Army's new strength-training exercises look to halve soldier injuries, improve form. *Today*. Retrieved from https://www.todayonline.com/singapore/goodbye-5bx-armys-new-strength-training-exercises-look-halve-soldier-injuries-improve-form

2. Daskalova, N., Desingh, K., Papoutsaki, A., Schulze, D., Sha, H., & Huang, J. (2017). Lessons learned from two cohorts of personal informatics self-experiments. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies, 1*(3), 1-22.

3. Elliott, D., Carr, S., & Orme, D. (2005). The effect of motivational music on sub-maximal exercise. *European Journal of Sport Science, 5*(2), 97-106.