

VILLANOVA UNIVERSITY

DESIGN OF EXPERIMENTS
STAT 4416

Final Project Report

Author:
Richard Lim

May 8th, 2023



VILLANOVA
UNIVERSITY

I. Introduction

The objective of this experiment was to see which combination of the three factors would maximize free throw percentage. The three factors were Location (indoor and outdoor courts shown in the Appendices section titled Figure 1 and Figure 2), Ball (2 different balls shown in the Appendices section titled Figure 3 and Figure 4), and Music (music and no music). This experiment was done with two different people, which adds a blocking factor. With three two-level factors and two different people, there were a total of sixteen total runs. In each run, twenty shots were taken, and the response variable was the percentage of made free throws out of twenty. This experiment consisted of three two-level factors and a blocking factor making it a 2^3 Factorial Design with a blocking factor. Before completing the experiment, the hypothesis was that the combination of shooting on an indoor court and without music would maximize the free throw percentage. The reason why an indoor court was hypothesized to maximize free throw percentage was that there are minimal to no weather conditions affecting free throw shooting on an indoor court as opposed to an outdoor court. The lack of music was also hypothesized to maximize free throw shooting because the presence of music may lead to distractions, ultimately affecting the outcome.

II. Performing the Experiment

This experiment was performed on Saturday, April 15th, 2023 between 2:00 PM and 4:00 PM. The order of the runs to be completed was determined by assigning the eight different combinations of factors a number from one through eight and using a number randomizer to determine the order in which they were to be completed. The order to be determined along with the results can be found in Table 1 in the Appendices section. Once the order of the runs to be

completed was determined, both I and my friend would take a total of twenty shots with the determined combination of factors. As aforementioned, a total of sixteen runs were done with each person taking turns to minimize fatigue. Once each run was completed, the participant would also take a sip of water to attempt to minimize fatigue.

One of the confounding variables is the weather conditions when shooting free throws on an outdoor court. For one of the runs, the impact of the weather may be minimal, but for other runs, the impact of the weather may be more severe. But, the reason this factor was used was to show that there may be some difference between the two levels because of these conditions. One of the possible confounding variables may have been fatigue. Towards the end of the conduction of the experiment, both participants may have been getting worn out as both shot one hundred sixty free throws each. This could have skewed the results of the later runs. Another possible confounding variable could have been the surrounding noise when completing the runs that involved no music. These other noises or background noises may have distracted my friend and me, causing us to miss more shots. The final possible confounding variable is possible spectators. When there are spectators watching, free-throw shooters tend to miss more shots than they otherwise would have if there were no spectators. Thus, the acknowledgment of possible spectators may have skewed the results.

Finally, as there were only two possible outcomes of free throw shooting (made free throw or missed free throw) there were no issues with accurately collecting data.

III. Results

$$Y_{ijk} = \mu + \tau_i + \beta_j + \alpha_k + (\tau\beta)_{ij} \delta_1 + (\tau\alpha)_{ik} \delta_1 + (\beta\alpha)_{jk} \delta_1 + (\tau\beta\alpha)_{ijk} \delta_1 + \varepsilon_{ijkl}$$

μ = overall mean

τ_i = effect of the ith level of the Location factor, $i = 1, 2$

β_j = effect of the jth level of the Ball factor, $j = 1, 2$

α_k = effect of the kth level of the Music factor, $k = 1, 2$

$(\tau\beta)_{ij}$ = interaction effect due to the ith level of the Location factor and jth level of the Ball factor

$(\tau\alpha)_{ik}$ = interaction effect due to the ith level of the Location factor and kth level of the Music factor

$(\beta\alpha)_{jk}$ = interaction effect due to the jth level of the Ball factor and kth level of the Music factor

$(\tau\beta\alpha)_{ijk}$ = interaction effect due to the ith level of the Location factor, jth level of the Ball factor, and kth level of the Music factor

δ_l = effect of the lth level of the blocking factor, $l = 1, 2$

ε_{ijkl} = random error

Assumption: $\varepsilon_{ijkl} \sim \text{iid } N(0, \sigma^2)$

Errors are random and independent, and are normally distributed

The interaction plots for each 2-factor interaction can be found in the Appendices section titled Figure 5. Since plots 1, 2, 3, 4, and 6 have parallel lines (Additive model), it can be assumed that there is no interaction. However, plot 5 (Ball and Music Interaction) has slightly intersecting lines (Non-additive Disorderly), it can be assumed that there is an interaction but there is no guarantee.

From the Normal Q-Q plot in the Appendices section titled Figure 9, it can be assumed that this is a normally distributed model, however, there are not a lot of data points so it may look different with more points. This checks out with the model assumption of normality. From the residual plot under Figure 9, there seems to be no pattern in the points, thus a linear model is appropriate for this experiment.

The code for the ANOVA table and the resulting output can be found under the Appendices section under Figure 6. Since all of the P values are greater than 0.05, all of the main factors, 2 factor-interactions, and 3 factor-interaction are not significant. Since there are no significant factors, the first ones to be removed from the analysis are Location:Ball interaction and Location:Music interaction because both of these interactions share the lowest F value of 0.023. By default, the three-factor interaction of Location:Ball:Music must be removed as well because it involves the insignificant 2-factor interactions removed. The resulting ANOVA table can be found titled Figure 7 in the Appendices section. Again in this table, since the interaction of Ball:Music has a P value greater than 0.05, this interaction is to be removed as well. The resulting ANOVA table after removing all interactions can be found titled Figure 8 in the Appendices section. Even after removing all of the interactions, none of the main factors are found to be significant. However, the Location factor is very close to being significant with a P value of 0.0521. Thus, one of the factors hypothesized to be significant was found to be borderline significant.

IV. Conclusions

In summary, none of the interaction of factors or individual factors was found to be statistically significant, although one of the factors came close. In terms of my hypothesis, I was incorrect. I hypothesized that both Location and Music would be significant factors. Although the Location was close to being significant, neither was correct.

If I could do this experiment over again, I would add more levels to the Music factor. The different levels would represent different genres of music and no music. Classical music is

thought to assist in concentration, which is important when shooting free throws. Perhaps different genres can be found to be significant as well.

V. Appendices

Table 1:

| Person | Location | Ball | Music | Result | Order |
|--------|----------|------|-------|--------|-------|
| 1 | O | 1 | Y | 11/20 | 3 |
| 1 | O | 1 | N | 8/20 | 5 |
| 1 | O | 2 | Y | 4/20 | 8 |
| 1 | O | 2 | N | 9/20 | 6 |
| 1 | I | 1 | Y | 9/20 | 1 |
| 1 | I | 1 | N | 16/20 | 4 |
| 1 | I | 2 | Y | 10/20 | 7 |
| 1 | I | 2 | N | 13/20 | 2 |
| 2 | O | 1 | Y | 4/20 | 3 |
| 2 | O | 1 | N | 7/20 | 5 |
| 2 | O | 2 | Y | 8/20 | 8 |
| 2 | O | 2 | N | 4/20 | 6 |
| 2 | I | 1 | Y | 10/20 | 1 |
| 2 | I | 1 | N | 8/20 | 4 |
| 2 | I | 2 | Y | 9/20 | 7 |
| 2 | I | 2 | N | 4/20 | 2 |

Figure 1:



Figure 2:



Figure 3 (Ball 1):

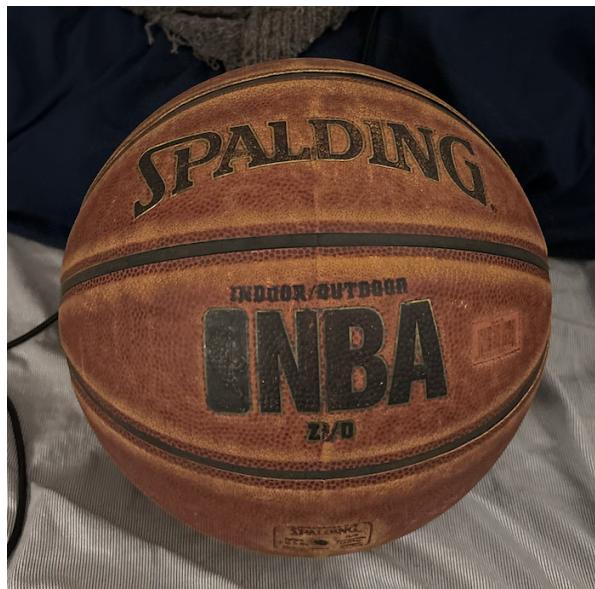


Figure 4 (Ball 2):

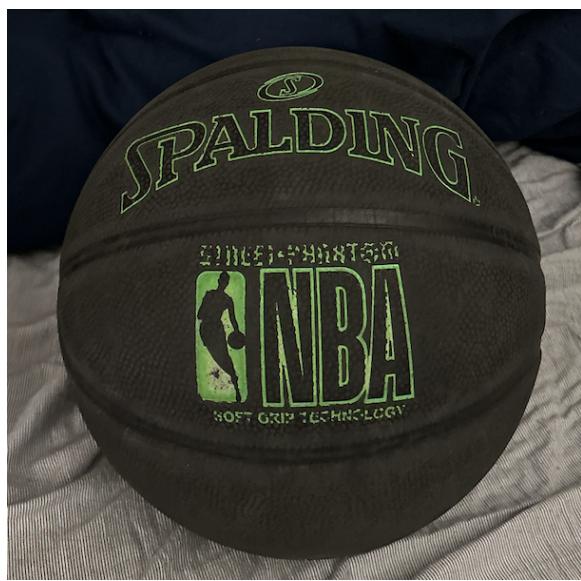


Figure 5:

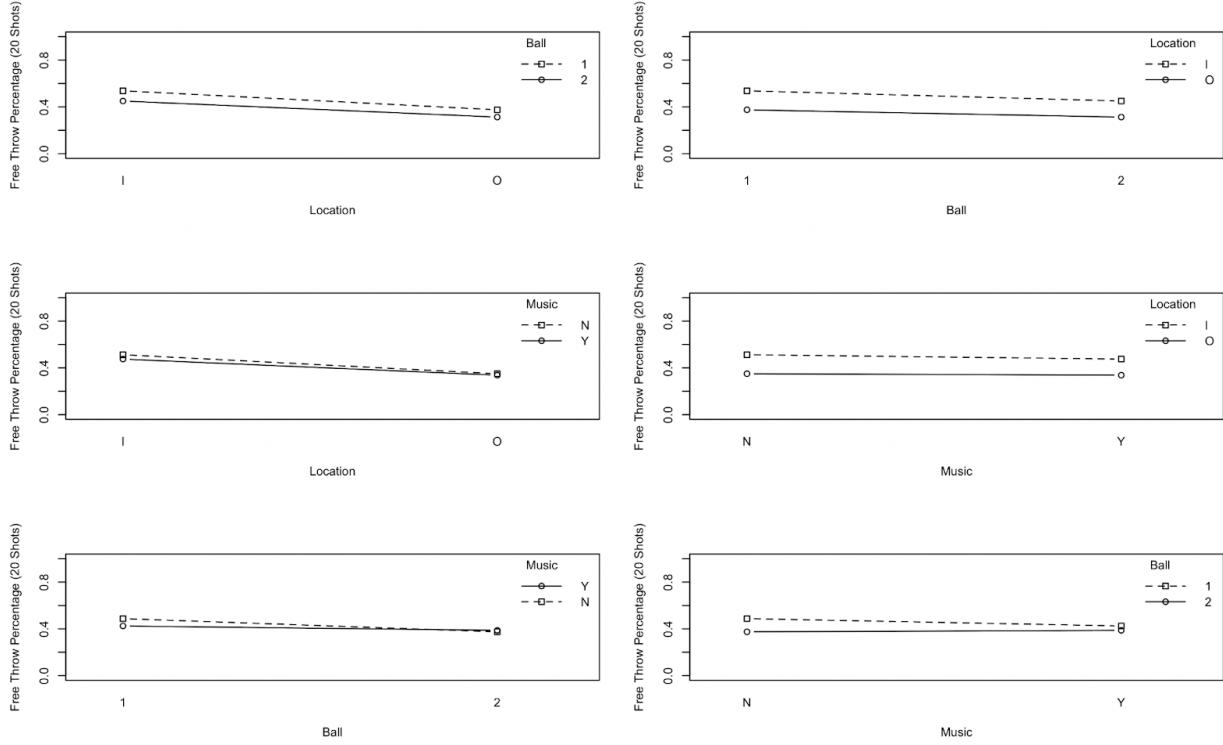


Figure 6:

```
> model <- aov(Result ~ Person+Location+Ball+Music+Location*Ball+Location*Music+Ball*Music+Location*Ball*Music)
> summary(model)

Df  Sum Sq Mean Sq F value Pr(>F)
Person          1 0.10562 0.10562  3.853 0.0904 .
Location         1 0.09000 0.09000  3.283 0.1129
Ball             1 0.02250 0.02250  0.821 0.3950
Music            1 0.00250 0.00250  0.091 0.7714
Location:Ball    1 0.00063 0.00063  0.023 0.8842
Location:Music   1 0.00062 0.00062  0.023 0.8842
Ball:Music       1 0.00563 0.00563  0.205 0.6643
Location:Ball:Music 1 0.01000 0.01000  0.365 0.5649
Residuals        7 0.19187 0.02741
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 7:

```
> model2 <- aov(Result ~ Person+Location+Ball+Music+Ball*Music)
> summary(model2)

   Df  Sum Sq Mean Sq F value Pr(>F)
Person      1 0.10562 0.10562  5.200 0.0458 *
Location    1 0.09000 0.09000  4.431 0.0616 .
Ball        1 0.02250 0.02250  1.108 0.3174
Music       1 0.00250 0.00250  0.123 0.7330
Ball:Music  1 0.00563 0.00563  0.277 0.6102
Residuals   10 0.20312 0.02031
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 8:

```
> new_model <- aov(Result ~ Person+Location+Ball+Music)
> summary(new_model)

   Df  Sum Sq Mean Sq F value Pr(>F)
Person      1 0.1056 0.10562  5.566 0.0379 *
Location    1 0.0900 0.09000  4.743 0.0521 .
Ball        1 0.0225 0.02250  1.186 0.2995
Music       1 0.0025 0.00250  0.132 0.7235
Residuals   11 0.2087 0.01898
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 9:

