From Ball Pressure to Distance: A Statistical Slam Dunk

Christian Koutsandreas[1], Yuika Takeuchi[1], Seokjin An[1], Aleah Ashlee[1], Rylan Smalbrugge[1†]

[1]Faculty of Mathematics and Statistics, University of Victoria, 3800 Finnerty Road, Victoria, V8W2Y2, BC, Canada.

Contributing authors: christiank@uvic.ca; yuikat@uvic.ca; seokjinan@uvic.ca; aleahashlee@uvic.ca; rylansmalbrugge@uvic.ca;
†These authors contributed equally to this work.

Abstract

This study investigates the effects of PSI, rebounder type, and shooting distance on the shooting accuracy of female varsity basketball players. The experiment used a randomized factorial design with three players as blocks to control for individual differences. Each player completed all eight trials in a randomized order, resulting in a total of 24 trials. Shot percentage was recorded for each trial.

1 Introduction

Basketball is a popular sport that requires players to perform a variety of physical and cognitive tasks, such as shooting, dribbling, passing, and defending. One of the most important skills in basketball is shooting accuracy, which can significantly impact a team's performance and ultimately determine the outcome of a game[1]. Despite the importance of shooting accuracy, there is limited research on the factors that affect it, especially in the context of training at different shooting distances and rebounding conditions.

The present study aims to investigate the effects of three factors, namely, PSI (5, 6), rebounder type (human vs machine), and shooting distance (three-point vs free throw), on the shooting accuracy of varsity basketball players. We aim to answer which of these factors have a significant effect on shooting accuracy, and what is the magnitude and direction of these effects.

This research problem is important for several reasons. First, it addresses a gap in the literature on basketball shooting performance during practice, especially in relation to female players who are often underrepresented in sports research [2]. Second, it has practical implications for coaches, players, and equipment manufacturers who can use the findings to optimize training programs, shooting strategies, and equipment design. Finally, it can lead to the greater importance and theoretical implications of other works investigating motor control factors such as shooting form or cognitive aspects on performance [3].

In this report, we will describe the experimental design and methodology, present and analyze the results, and discuss the implications and limitations of the findings.

2 Experimental Design

The experiment was designed as a 2^3 factorial design with blocks being UVIC Women's varsity basketball players. The three primary factors were PSI (5, 6), rebounder type (human vs machine), and shooting distance (three-point vs free throw), resulting in eight possible factor combinations or treatment levels. The three players were considered as blocks to control for individual differences in skill, motivation, and fatigue.

The experiment adhered to the three principles of experimental design, namely, randomization, replication, and blocking. Randomization refers to the random assignment of the factor combinations to the trials to control for the effects of extraneous variables and to ensure that each treatment level is equally represented in the sample. The ordering of the factor combinations was randomized using a computer program, and each player completed all eight trials in a randomized order. Replication refers to the repetition of the same experiment with different participants to estimate the variability of the treatment effects and to increase the generalizability of the findings. The three players were replicated in the experiment, which resulted in a total of 24 trials. Blocking refers to the grouping of similar experimental units (e.g., players) to control for the effects of extraneous variables that vary within blocks but not between blocks. Blocking reduces the variability of the treatment effects and increases the precision of the estimates. In this experiment, the three players were considered as blocks, and each player completed all eight trials in a randomized order.

To minimize the effects of nuisance factors that could affect the shooting accuracy, such as noise, temperature, lighting, warm-up, and non-independent trials, we controlled the environment and the procedures. We conducted the experiment in the primary UVIC basketball gym where the players were familiar with the court, the baskets, and the surroundings. We used the same basketballs for all trials, and we monitored the temperature and lighting conditions to ensure consistency across trials. We also gave the players a 3-minute break between trials to prevent fatigue and to reset their focus.

Overall, the experimental design followed the three principles of experimental design and minimized the effects of nuisance factors, which increases the internal validity and generalizability of the findings.

3 Data Collection

To conduct the experiment, three varsity basketball players from the UVIC Women's Basketball team volunteered to shoot 80 shots each and have their shot percentage recorded. The experiment was conducted on a weekday afternoon in the CARSA Performance Gym.

3.1 Procedure

1. To begin, the gear needed for the experiment was set up. Two shooting machines were set up on the side hoops and ball pressure was recorded using a PSI gauge for each ball used in the experiment. The balls with low pressure were marked with tape for easy identification during the operation. Because the design was being conducted in CARSA Performance Gym, regulation basketball lines were already in place for free throw and three point line distance.
2. While the gear for the experiment was being set up, the players were asked to do the regular dynamic warm up they would usually do for practice, along with shooting on a hoop without a shooting machine and shooting on a hoop with one, for about five minutes. The players' warm up time totalled to around ten minutes.
3. Each statistician was assigned a player and data sheet*. The statistician guided the player to each distance, rebounder type and ball pressure combination that was randomly assigned on their data sheet. The same 4 hoops were used throughout the experiment.
4. Once each trial began, the statistician recorded the number of successful shots out of ten. In total, each player took ten shots for each factor combination. This resulted in a total of 80 shots for each player and a grand total of 240 shots (3 players x 8 factor combinations x 10 shots).
5. Shot percentage was then recorded in each trial by dividing the total number of shots made by total shots taken.

3.2 Concerns

The main issue that arose during the experiment was the anticipated ball pressure did not match the actual ball pressure recorded during the experiment. Before conducting the experiment, we found that the regulation ball pressure for a basketball in the NBA was between 7.5 and 8.5 PSI, but the recorded ball pressure ended up being 5 and 6 PSI for the low and high levels respectively. After consulting the varsity basketball players, we concluded that the 5 and 6 PSI levels were appropriate, as the players expressed that those were the pressure levels they were used to. In fact, the players also informed us that the way they usually determined ball pressure was "ball drop method", which involves dropping the ball from shoulder height and if it lands somewhere near the hips, it is deemed to have an appropriate pressure level. This method is supported by the NCAA basketball rules on ball inflation as well. There were no other issues that emerged throughout and despite the original plan being changed, the experiment was conducted successfully.

4 Data analysis

4.1 Anova

We used a 2^3 factorial design which gives us the statistical model:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{jk} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \delta_l + \epsilon_{ijkl} \quad (1)$$

where τ_i denotes the i^{th} effect of rebounder (machine or player), β_j denotes the j^{th} effect of ball pressure (5 psi or 6 psi), γ_k denotes the k^{th} effect of distance (free-throw or 3-point line), and δ_l denotes the block effect (replicates 1,2,3). We use an ANOVA table as well as a Normal Q-Q plot to determine significant factors. This gives us a reduced model where only the distance (factor C) is significant with the block effect being near significant.

Table 1 Reduced ANOVA model summary.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
C	1	0.45	0.45	25.21	0.0001
factor(blocks)	2	0.09	0.05	2.57	0.1016
Residuals	20	0.36	0.02		

As we had initially expected we have that the distance from which the player throws being significant to shot accuracy. Blocking will require additional investigation to determine whether or not it is significant.

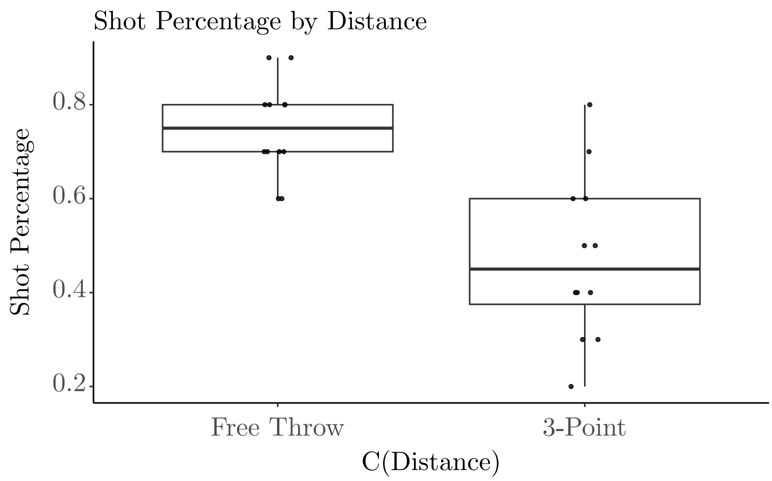

Fig. 1 Boxplot of shot percentage by distance level.

A quick look at the means and quantiles for shot accuracy by distance shows a very obvious difference between free throw and 3-point shots. This agrees with both our results and what one would expect.

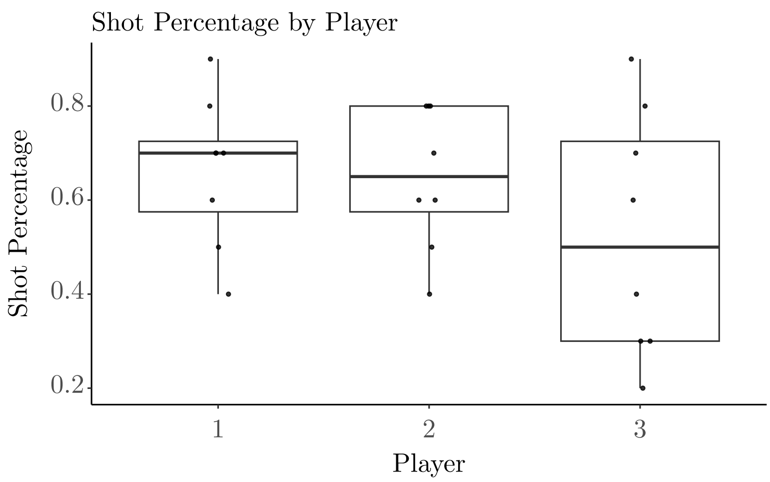

Fig. 2 Boxplot of shot percentage by player.

While it would seem obvious that a difference in player would affect shooting accuracy it is difficult to determine from our plots and ANOVA with certainty. This could be due to our 3 players being similar in ability or due to a low sample size. We investigate further through a linear model.

A residual analysis of our ANOVA confirms our assumptions of independence, constant variance, and normally distributed errors are reasonable with no cause for concern as displayed in figures B1, B2, and B3. The shapiro test gives us a p-value of 0.3207.

4.2 Linear Model

We now fit the data to a linear model:

Table 2 Reduced linear model summary

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5250	0.0474	11.07	0.0000
C	-0.1375	0.0274	-5.02	0.0001
factor(blocks)2	0.1250	0.0671	1.86	0.0772
factor(blocks)3	0.1375	0.0671	2.05	0.0537

	R.squared	Adj.R.squared
1	0.60	0.54

With a R^2 of 0.6028 the linear model appears as a good fit for our data. However we see in figure B4 that the cook's distance of point 14 is high enough to deem it an influential point. Removing this point gives us a new model:

Table 3 Reduced linear model summary with observation 14 removed.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5250	0.0414	12.67	0.0000
C	-0.1519	0.0245	-6.20	0.0000
factor(blocks)2	0.0819	0.0607	1.35	0.1935
factor(blocks)3	0.1375	0.0586	2.35	0.0299

	R.squared	Adj.R.squared
1	0.70	0.65

This improves our R^2 value significantly to 0.7001 giving us reason to believe that our data is linear. A residual analysis on this model shows no issues with our assumptions of independence, constant variance, and normally distributed errors (See Figures B5, B6, and B7).

5 Conclusion

In conclusion, our experiment aimed to investigate the effects of ball pressure, rebounder type, and distance on basketball shot percentage. Our results suggest that ball pressure and rebounder type did not have a significant impact on basketball shot percentage. The insignificance of rebounder type may be due to the familiarity of the players with both human and machine rebounders, which are commonly used in training. Thus, the large obstructive appearance of machine rebounders should not be avoided in training strategies as it does not affect shot percentage. The insignificance of ball pressure may be due to the relatively small difference in factor levels. Implying the factor levels lied within an acceptable threshold and hence was not enough to affect shot accuracy. As the other expectation, if we consider the shot during the game, the players will dribble before shooting, so the ball pressure should be one of the significant factors toward the shot percentage.

However, we found that distance from the basket was a significant factor, with a higher shot percentage for free throws compared to three-point shots. This significance is attributed to the increasing level of difficulty the further away a shot is attempted. Our results suggest that other factors, such as shot form and mental state [4], may have a greater impact on shot percentage than ball pressure and rebounder type.

Overall, our experiment provides valuable insights into the effects of different factors on basketball shot percentage during training. It also highlights the importance of carefully controlling for nuisance factors in experimental designs. Our findings underscore the importance of continuing research into other factors that may affect basketball shot accuracy, such as form, cognitive aspects, and environmental conditions. Further research in this area could lead to more effective training strategies and improved player performance.

Appendix A Code

```
```{r}
# import libraries
library(readxl)
library(tidyverse)

# load data
df = read_excel("Final_Project_Data.xlsx", range = "A1:H9")

# pre-process data
df_long <- df %>%
  pivot_longer(
```
```

```r
    cols = c(yi, yj, yk),
    names_to = "player",
    values_to = "value",
    names_prefix = "y"
  ) %>%
  mutate(player_num = case_when(
    player == "i" ~ 1,
    player == "j" ~ 2,
    player == "k" ~ 3
  )) %>%
  select(-Runs, -player) %>%
  rename("Shot Accuracy" = value, Player = player_num)

# set factor columns
df_long$`A(Rebounder)` = as.factor(df_long$`A(Rebounder)`)
df_long$`B(Ball Pressure)` = as.factor(df_long$`B(Ball Pressure)`)
df_long$`C(Distance)` = as.factor(df_long$`C(Distance)`)
df_long$Label = factor(df_long$Label,
                       levels = c("1", "a", "b", "c",
                                  "ab", "ac", "bc", "abc"))
df_long$Player = as.factor(df_long$Player)

# box plot of players
df_long_modified = df_long
df_long_modified$`C(Distance)`
   <- factor(ifelse(df_long$`C(Distance)` == "-", "Free Throw", "3-Point"),
    levels = c("Free Throw", "3-Point"))

distance_box = df_long_modified %>%
  ggplot(aes(x=`C(Distance)`, y=`Shot Percentage`)) +
    geom_boxplot() +
    geom_jitter(color="black", size=1, alpha=0.8, width = .05, height = 0) +
    #scale_fill_manual(values = c("#FF2E2E", "#D10000", "#750000")) +
    theme(legend.position="none") +
    ggtitle("Shot Percentage by Distance Level") +
    theme_classic() +
    theme(axis.text.x = element_text(vjust = -5)) +
    theme(axis.title.x = element_text(margin = margin(t = 25))) +
    theme(axis.title.y = element_text(margin = margin(r = 25)))

distance_box

# 3 players boxplot
player_box_wjit = df_long %>%
  ggplot(aes(x=Player, y=`Shot Percentage`)) + #, fill=Player
```

```
    geom_boxplot() +
    geom_jitter(color="black", size=1, alpha=0.8, width = .05, height = 0) +
    #scale_fill_manual(values = c("#FF2E2E", "#D10000", "#750000")) +
    theme(legend.position="none") +
    ggtitle("Shot Percentage by Player") +
    theme_classic() +
    theme(axis.text.x = element_text(vjust = -5)) +
    theme(axis.title.x = element_text(margin = margin(t = 25))) +
    theme(axis.title.y = element_text(margin = margin(r = 25)))

player_box_wjit
```

```{r cleaning}
A = rep(rep(c(-1,1),4),3)
B = rep(rep(c(rep(-1,2),rep(1,2)),2),3)
C = rep(c(rep(-1,4),rep(1,4)),3)

y1 = c(0.7,0.9,0.8,0.6,0.3,0.2,0.3,0.4)
y2 = c(0.6,0.7,0.8,0.8,0.6,0.8,0.4,0.5)
y3 = c(0.8,0.9,0.7,0.7,0.4,0.6,0.7,0.5)
y = append(y1,y2)
y = append(y,y3)

blocks = c(rep(1,8),rep(2,8),rep(3,8))
data = data.frame(y,A,B,C, blocks)

res.aov = aov(y~A*B*C + factor(blocks),data = data)
summary(res.aov)
res.lm = lm(y~A*B*C + factor(blocks), data = data)
summary(res.lm)

library(xtable)

library(daewr)
fullnormal(coef(res.lm)[-1],alpha=.05, refline = T) #Figure

res.aov2 = aov(y~C + factor(blocks), data = data)
xtable(res.aov2) #Table 1
res.lm2 = lm(y~C + factor(blocks), data = data)
xtable(res.lm2) #Table 2
r_sq = summary(res.lm2)$r.squared
adj.r = summary(res.lm2)$adj.r.squared
xtable(data.frame(R.squared = r_sq, Adj.R.squared = adj.r)) #Table 2
summary(res.aov2)
summary(res.lm2)
```

```
```

```{r plots}
res.aov2 = aov(y~C + factor(blocks), data = data)
res.lm2 = lm(y~C + factor(blocks), data = data)

plot(res.aov2, which = c(1,2,4))
shapiro.test(res.aov$residuals)
plot(1:1:24,res.aov2$residuals, ylab="Residuals",
        xlab="Order", main = "Independence Check")
abline(h=0)

res.aov2 = update(res.aov2,.~.,subset=(1:length(y)!=14))
summary(res.aov2)
res.lm2 = update(res.lm2,.~.,subset=(1:length(y)!=14))
summary(res.lm2)
r_sq = summary(res.lm2)$r.squared
adj.r = summary(res.lm2)$adj.r.squared
xtable(data.frame(R.squared = r_sq, Adj.R.squared = adj.r)) #Table 2
xtable(res.lm2) #Table 2
plot(res.aov2, which = c(1,2,4))
shapiro.test(res.aov$residuals)
plot(1:1:23,res.aov2$residuals, ylab="Residuals",
    xlab="Order", main = "Independence Check")
abline(h=0)

```

```{r fun plots}
library(ggplot2)
y1 = c(0.7,0.9,0.8,0.6,0.3,0.2,0.3,0.4)
y2 = c(0.6,0.7,0.8,0.8,0.6,0.8,0.4,0.5)
y3 = c(0.8,0.9,0.7,0.7,0.4,0.6,0.7,0.5)
y = append(y1,y2)
y = append(y,y3)
A = rep(rep(c(-1,1),4),3)
B = rep(rep(c(rep(-1,2),rep(1,2)),2),3)
C = rep(c(rep(-1,4),rep(1,4)),3)
ymean = (y1+y2+y3)/3

close = y[C == -1]
far = y[C == 1]
machine = y[A == -1]
player = y[A == 1]
```

```
lower = y[B == -1]
high = y[B == 1]

data = data.frame(
  Factor = factor(c(rep("Machine",12),rep("Human",12),rep("5 psi",12),
    rep("6 psi",12),rep("Free Throw",12),rep("3-Point",12)),
    levels = c("Machine","Human","5 psi","6 psi","Free Throw","3-Point")),
  ShotPercent = c(machine,player,lower,high,close,high))
  ggplot(data ,aes(x=Factor, y=ShotPercent, fill = Factor)) +
    geom_boxplot() +
    scale_fill_manual(values=c("#ff9933","#ff9933",
        "#ffff33","#ffff33","#ff3333","#ff3333")) +
    geom_jitter(color="black", width = 0.1,size=0.4, alpha=0.9) +
    ggtitle("Shot Percentage by Factor")
    + theme(legend.position = "none") +
    xlab("") + ylab("Shot Accuracy")
#Plot for Presentation
boxplot(machine,player,lower,high,close,far,
    main = "Shooting Accuracy by Factor",
    names = c("Machine","Human","5 psi","6 psi","Free throw","3-Point"),
    col = c("#ff9933","#ff9933","#ffff33","#ffff33","#ff3333","#ff3333"))

boxplot(close,far, main = "Boxplots of Shooting Accuracy by Distance",
    names = c("Free Throw","3-Point Line"), col = c("#ffccff","#993333"))

boxplot(y1,y2,y3, main = "Boxplots of Shot Percentages by Player",
    names = c(1,2,3), col = c("#ff9933", "#ffff33", "#ff3333"))

mid = barplot(ymean, main = "Mean Shot Percentage",
    names = c("(1)","a","b","ab","c","ac","bc","abc"),
    col = rgb(0.8,0.1,0.1,0.6), ylim = c(0,1))
arrows(x0 = mid, y0 = ymean+sd(ymean), x1 = mid, y1 = ymean-sd(ymean),
    code = 3, angle = 90, length = 0.05)


plot(y1, xaxt = "n", ylab = "Shot Percentage",
    ylim = c(0,1), col = "#ff9933", pch = 5, lwd = 2,main = "Data")
points(y2,col = "#ffff33", pch = 2, lwd = 2)
points(y3, col = "#ff3333", pch = 6, lwd = 2)
legend(1,0.4, legend = c("Player 1", "Player 2", "Player 3"),
    col = c("#ff9933","#ffff33","#ff3333"),
    pch = c(5,2,6), lty = c(NA,NA,NA), lwd = 2)
axis(1, at=1:8,labels = c("(1)","a","b","ab","c","ac","bc","abc"))
```
```
```

Appendix B Figures

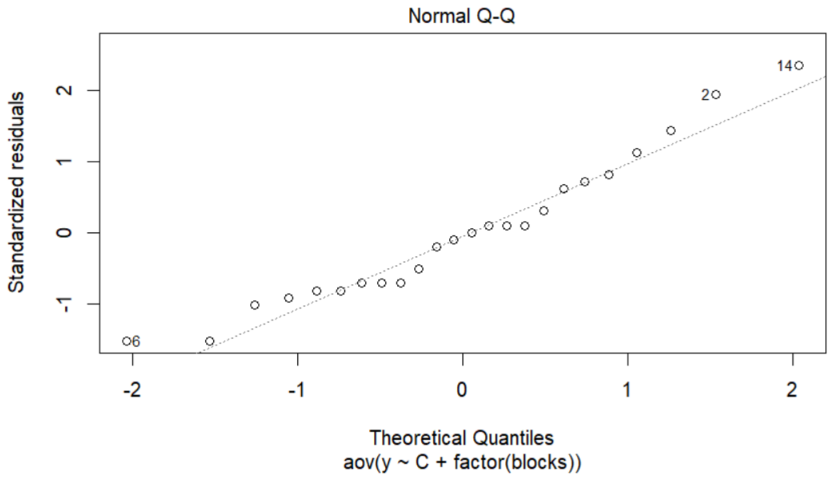

Fig. B1 Normal Q-Q reduced model residuals.

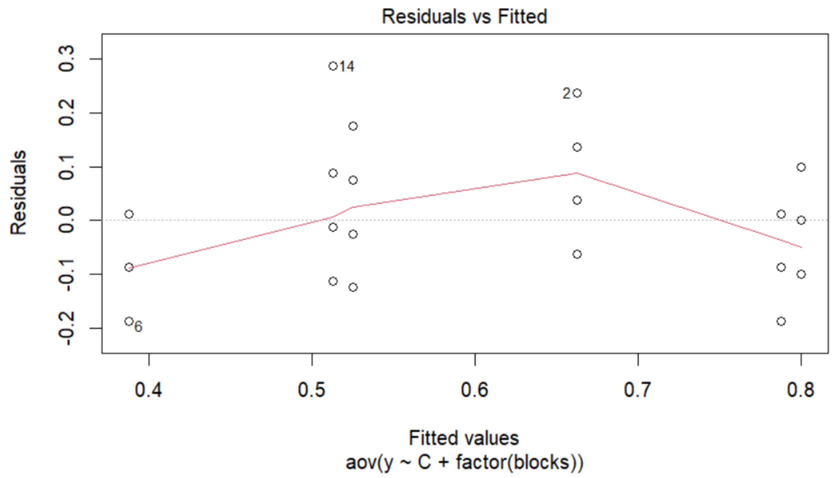

Fig. B2 Reduced model residuals by fitted value scatter.

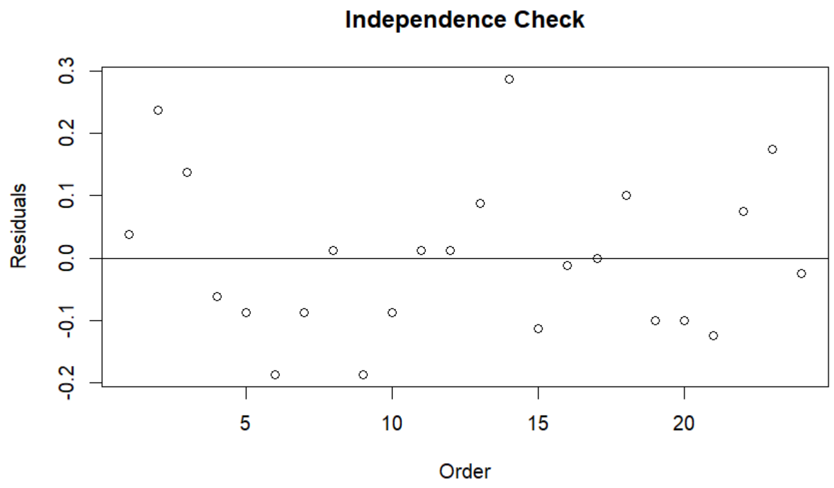

Fig. B3 Independence Check for reduced ANOVA model.

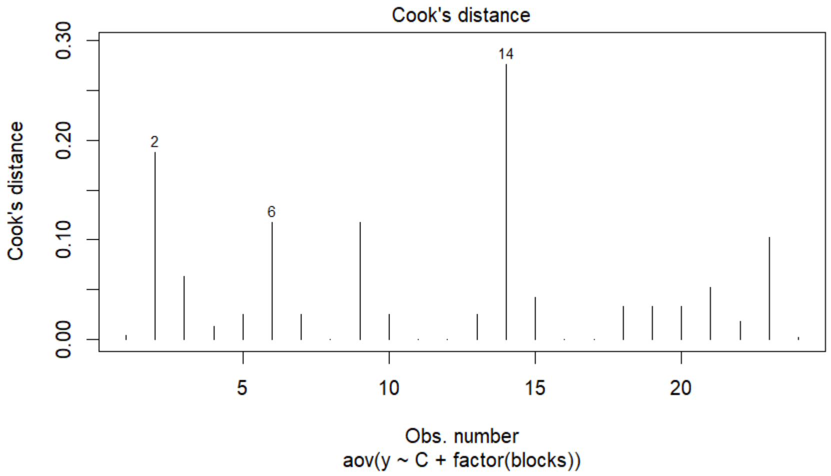

Fig. B4 Cook's Distance for reduced Model.

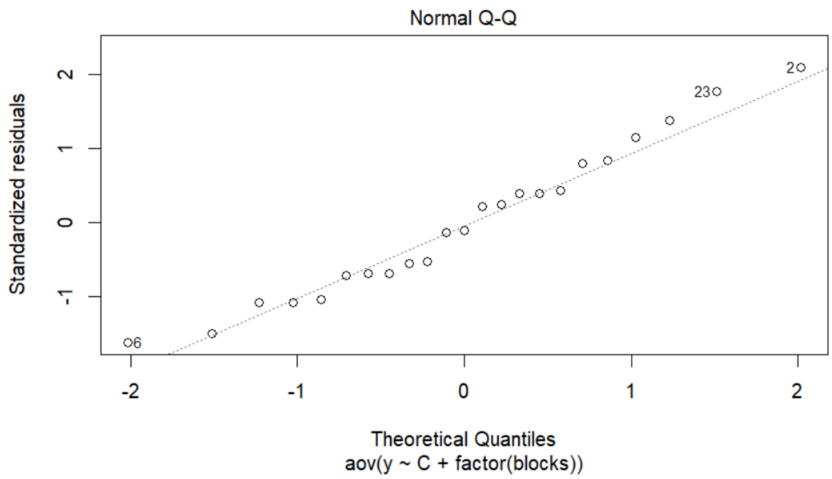

Fig. B5 Q-Q Normal plot for model with observation 14 removed.

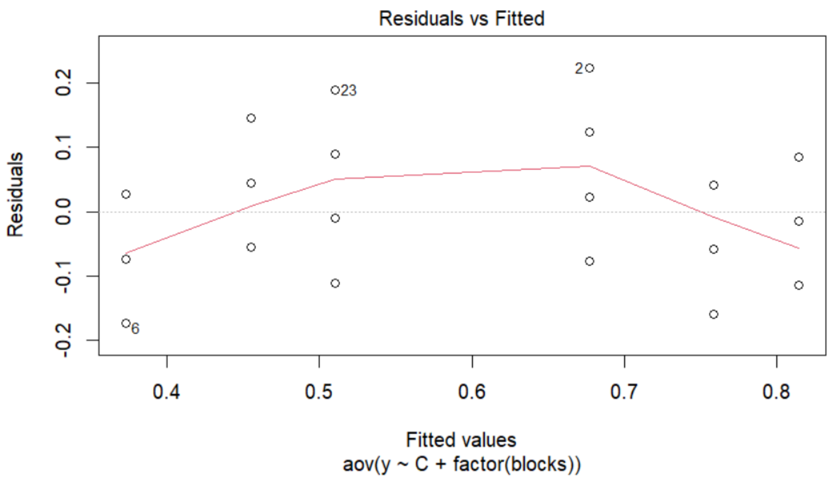

Fig. B6 Constant variance check for model with observation 14 removed.

Fig. B7 Independence check for model with observation 14 removed.

Table B1 Collected dataset.

	P1	P2	P3	Label
1	0.7	0.6	0.8	(1)
2	0.9	0.7	0.9	a
3	0.8	0.8	0.7	b
4	0.6	0.8	0.7	ab
5	0.3	0.6	0.4	c
6	0.2	0.8	0.6	ac
7	0.3	0.4	0.7	bc
8	0.4	0.5	0.5	abc

References

[1] Kozar, B., Vaughn, R.E., Whitfield, K.E., Lord, R.H., Dye, B.: Importance of free-throws at various stages of basketball games. Perceptual and Motor skills **78**(1), 243–248 (1994)

[2] Kim, S.: Train like a girl: Why we need more studies focused on female athletes (2021). https://hiseye.org/9547/his-eye/features/train-like-a-girl-why-we-need-more-studies-focused-on-female-athletes/

[3] Okazaki, V., Rodacki, A., Satern, M.: A review on basketball jump shot. Sports Biomechanics **14**, 1–16 (2015) https://doi.org/10.1080/14763141.2015.1052541

[4] Miyaguchi, S., Inukai, Y., Hashimoto, I., Otsuru, N., Onishi, H.: Sleep affects the motor memory of basketball shooting skills in young amateurs. Journal of Clinical Neuroscience **96**, 187–193 (2022) https://doi.org/10.1016/j.jocn.2021.11.016