Bouncing Balls: Evaluating the Statistical Effect of Various Sports Surfaces and Balls on Total Bounce Time

Zachary Kobban, Emilia Natale, and Olive O'Riordan
November 13th, 2023

Introduction

Topic Background:
We are curious about why sports are played on different types of surfaces with different types of balls. Why is Squash played on a hardwood floor while Lacrosse is played on a field? In order to investigate this question we want to look at how different types of balls bounce on different surfaces. Specifically, how do ball type and bounce surface affect the amount of time it takes for a ball, dropped at a given height to stop bouncing?

Purpose of Experiment:
The objective of this study is to evaluate and compare the time it takes balls to stop bouncing on three different types of surfaces with the treatment of different types of sports balls. The goal of the experiment would be to use the ANOVA workflow we developed to prove which surface allows which ball to continue bouncing for the longest time.

Is ball type and floor surface additive factors for bounce time or is there some type of interaction between the factors? What surface and ball combination will lead to the longest and shortest bounce times?

Scientific Hypothesis:
We hypothesize that the ball and floor combination that will have the longest bounce time will be the tennis ball on the squash court. This is because we anticipate that balls with a lower density will bounce longer than balls with higher densities.
First, we determined the different diameters and weights of the balls we used ("Sports Ball Size Comparison"; "Sports Ball Weight Comparison").

Squash balls are 3.96 to 4.04 centimeters in diameter and weigh 23 to 25 grams.
Tennis balls are 6.54 to 6.86 centimeters in diameter and weigh 56 to 59 grams.
Lacrosse balls are 6.35 centimeters in diameter and weigh 142 to 156 grams.

$$Density = mass/volume = mass/(4/3\,\pi\, r^3) = mass/(4/3\,\pi\,(d/2)^3)$$

Let's assume the lower bound for ball weights and ball diameters. That gives us the following numbers:
The density of the squash ball is $23\ g/\ 32.73\ cm^3 \approx 0.703\ g/cm^3$
The density of the tennis ball is $56\ g/\ 147.05\ cm^3 \approx 0.381\ g/cm^3$
The density of the lacrosse ball is $142\ g/\ 135.23\ cm^3 \approx 1.050\ g/cm^3$

The tennis ball has the lowest density, the lacrosse ball has the highest density, and the squash ball's density is in between.

We additionally believe that balls (regardless of type) will bounce longer on harder surfaces. This means that we predict balls will bounce longer on the squash court, than the track surface or the turf field.

Materials Used in the Experiment

- Timer (phone)
- Tape measure
- Different surfaces (turf field, squash court, outdoor running track)
- Sports balls (tennis ball, squash ball, and lacrosse ball)

Experimental Design and Procedure:

In our experiment, we hope to find significant differences in mean bounce time given the type of floor and type of ball. We went to Kenyon's athletic center to the football field, outdoor track, and squash courts and borrowed lacrosse, tennis, and squash balls from the front desk. To determine the order in which we would conduct our trials, we had R create a random order for the numbers 1-9 (1 = Hardwood.Lacrosse, 2 = Hardwood.Tennis, 3 = Hardwood.Squash, 4 = Track.Lacrosse, 5 = Track.Tennis, 6 = Track.Squash, 7 = Turf.Lacrosse, 8 = Turf.Tennis, and 9 = Turf.Squash). For the sake of ease, we completed all 5 trials for each combination in direct succession. The order we ran our tests was 4, 2, 6, 7, 5, 8, 9, 3, and 1, with each number referring to the combination stated above. Otherwise, we would have had to switch locations between the 45 trials. Given our personal time constraints, we knew that if we had to move between locations for each trial, our data collection would be rushed, more impacted by human error, and result in more variability between trials. Because of this we collectively decided on a procedure that did satisfy complete randomization.

When testing, one person timed the bounce, one person dropped the ball, and one person recorded the data in our spreadsheet. We used a tape measure to measure 2 meters above the ground so that our drop height was the same across all trials. We started the timer when the ball left the hand of the dropper and stopped timing once the ball was no longer bouncing and recorded the time to the hundredth of a second.

Statistical Analysis

We began with some basic EDA of our data. A boxplot of our data (fig. 1) reveals variance in the variables. At first glance, the Hardwood and Track appear to have similar means but the amount of time the balls bounced on the Turf appears to be much shorter. This trend is reinforced by observing the favstats data for the different surface types (Table 1). There also appears to be a difference between each of the different ball types (Table 2). Each of the ball types appears to have a different meaning from the others.

We were initially concerned that the standard deviations of some of the balls appeared to be twice as great as others, calling into question constant variance. In conditions, we made sure to use Levene's test to check this condition.

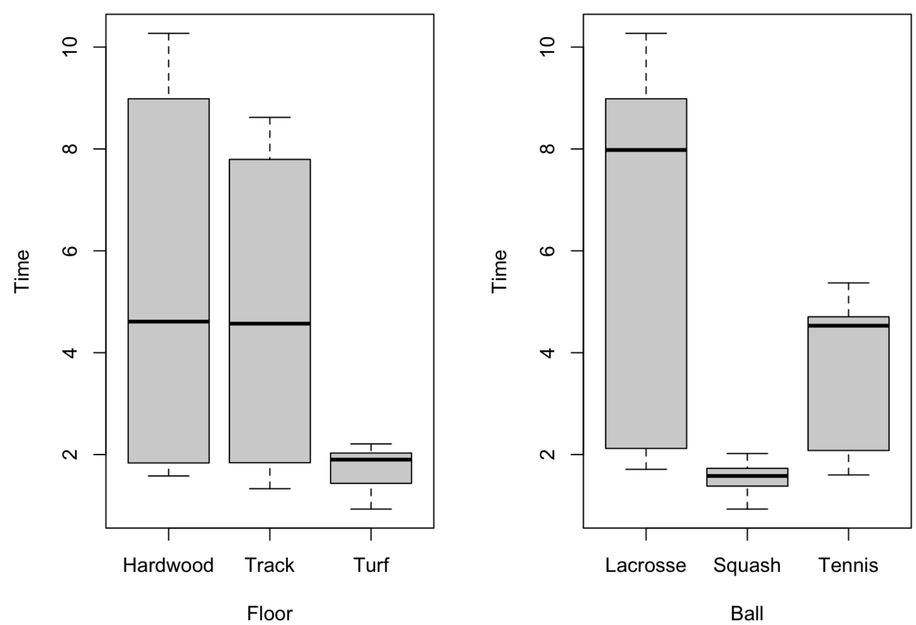

Figure 1: boxplots for floor type(left) and ball type(right) variables.

	Floor	min	Q1	median	Q3	max	mean	sd	n	missing
1	Hardwood	1.58	1.835	4.61	8.985	10.27	5.344667	3.3818293	15	0
2	Track	1.33	1.840	4.57	7.795	8.62	4.858000	2.7459066	15	0
3	Turf	0.93	1.435	1.90	2.030	2.21	1.738667	0.3771863	15	0

Table 1: favstats analysis of floor factor.

	Ball	min	Q1	median	Q3	max	mean	sd	n	missing
1	Lacrosse	1.71	2.12	7.98	8.985	10.27	6.559333	3.4412756	15	0
2	Squash	0.93	1.38	1.58	1.730	2.02	1.558000	0.2725593	15	0
3	Tennis	1.60	2.08	4.53	4.705	5.37	3.824000	1.4024000	15	0

Table 2: favstats analysis of ball factor.

To determine if an interaction between *Ball* and *Floor* would be necessary for our model, we created and analyzed interaction plots (fig. 2). Because these plots were not parallel, we decided to proceed with a model that included an interaction term rather than an additive model.

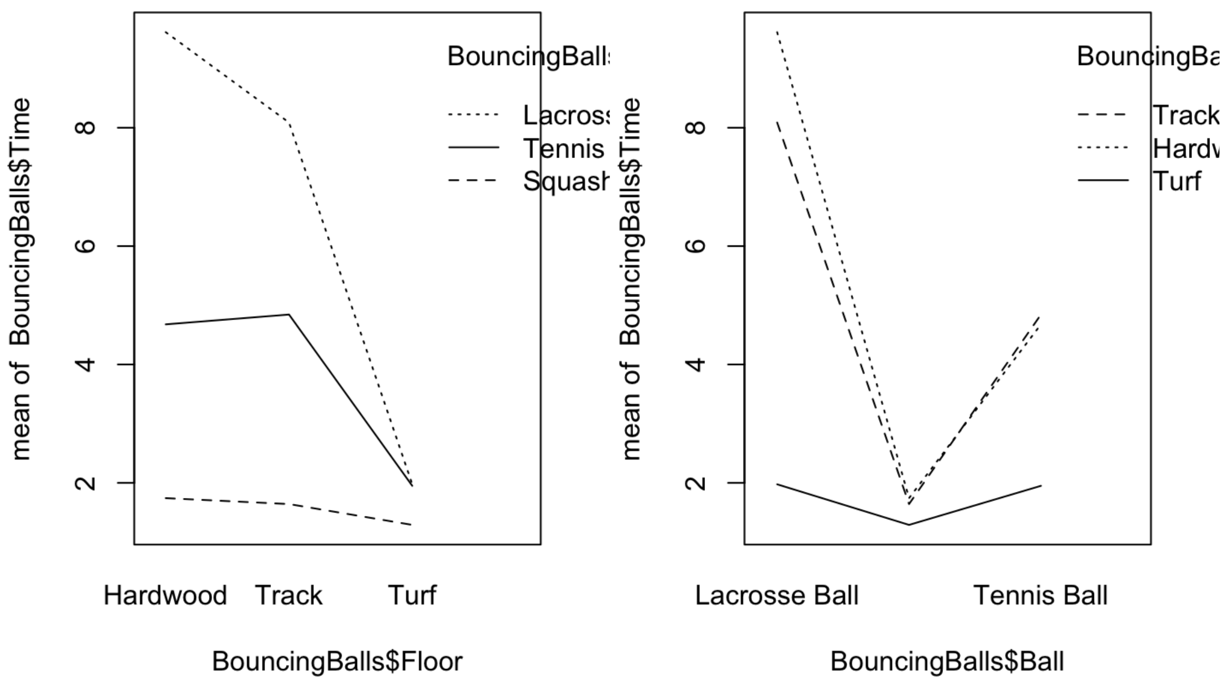

Figure 2: Interaction plots with Floor Type on the X-axis (left) Ball Type on the X-axis (right) and time on the Y-axis. Note: For the right plot, the middle X value represents the Squash ball but is unlabeled.

This leads us to believe we need an interaction in our model rather than using an additive model. We created the model with the interaction term and a null model. The full model with interaction is

$$BounceTime_{ijk} = \mu + BallType_i + FloorType_j + (BallType \times FloorType_{ij}) + \varepsilon_{ijk}$$

We then created the null model $BounceTime = 1$ which gives us just the overall mean. When running a nested F test with the full model and the null model, our hypotheses were H_0: we do not need any of the effects from our full model and we can just use the overall mean to determine bounce time, and H_a: we need something in our full model that we need to include. With df = (8, 36), F-statistic = 423.47, and p-value < 0.05 = significance level, we can reject the null hypothesis as we have significant evidence that something within our model is significant for determining mean bounce time.

Since our model from the nested F test proved to have significance, we will use the full model $Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$, where ε_{ijk} ~(iid)~$N(0,\sigma_\varepsilon)$ and i = 1(Lacrosse), 2(Tennis), 3(Squash), j = 1(Hardwood), 2(Track), 3(Turf), and k = 1, 2, 3, 4, 5 (for each trial of each interaction).

In this model,

μ = overall population mean
α_i = type of ball effect
β_j = type of floor effect
$\alpha\beta_{ij}$ = interaction effect
ε_{ijk} = observation error

To fit our model, we estimated our parameters μ, α_i , β_j , and $\alpha\beta_{ij}$. First, we looked at the means of the different interaction levels and the main effect means (table 3).

Table 3: all mean estimates for the data.

Means	Floor = Hardwood	Floor = Track	Floor = Turf	Overall Ball Means
Ball = Lacrosse	9.614	8.088	1.976	6.559333
Ball = Tennis	4.678	4.844	1.95	3.824
Ball = Squash	1.742	1.642	1.29	1.558
Overall Floor Means	5.344667	4.858	1.738667	3.980444

In table 3, we see that $\overline{Y} = 3.980444$. Using this information, we were then able to calculate the main and interaction effects (table 4) with the mean statistics:

$$Ball\,factor\,effect = \overline{Y_{i\bullet}} - \overline{Y}$$
$$Floor\,factor\,effect = \overline{Y_{\bullet j}} - \overline{Y}$$
$$Interaction\,effect = \overline{Y_{ij}} - \overline{Y_{i\bullet}} - \overline{Y_{\bullet j}} + \overline{Y}$$

Table 4: all effects for the data.

Effects	Floor = Hardwood	Floor = Track	Floor = Turf	Overall Ball Effect
Ball = Lacrosse	1.690444	0.651111	-2.341556	2.578889
Ball = Tennis	-0.510223	0.142444	0.367777	-0.156444
Ball = Squash	-1.180223	-0.793556	1.973777	-2.422444

Overall Floor Effect	1.364223	0.877556	-2.241777	3.980444

Conditions:

Before drawing any implications from our model, we checked the ANOVA conditions of zero-mean, independence, constant variance, additivity, and normality.

Zero mean is always true in this case but we can check just to be sure. If we take the mean of all of the residuals in the model we get $8.524238 * 10^{-18}$. This is very very close to zero and because of this, we can check off the zero mean condition.

For independence, we have to ask ourselves, is it possible for the outcomes of past trials to impact the outcomes of future trials? We made the reasonable assumption that the bouncing of each ball should not do any damage to either the ball or the floor to such an extent that it would impact future trials. However, we recognize that our trials are not fully independent, given that we did not achieve full randomization in our procedure. This is a limitation of our model that we recognize.

To check constant variance, we analyzed the Residuals vs. Fitted plot for the model and determined the condition is met as vertical spreads are generally equal for each group even though there are a few outliers like point 3 (fig. 3). To further confirm that the condition for constant variance is met, we used Levene's test for homogeneity of variances. The null hypothesis for the Levene test is that all variances are equal and the alternative hypothesis is that all variances are not equal. The test showed that, with df = (8, 36), F-statistic = 1.9841, and a p-value = 0.07697 > 0.05 = significance level, we can not reject the null hypothesis. This once again suggests that there is constant variance across groups for our model with the interaction between *Ball* and *Floor*.

With respect to additivity, we are not able to meet this condition as we use an interaction term in our model. Our interaction plots (fig. 2)suggest that additivity is not sufficient and interaction is necessary to construct an adequate model.

For normality, we checked the Normal Q-Q plot (fig. 3). and noticed some outliers and curvature, especially at the upper tail. Though this would make us wary about normality being met, since our data is balanced it is not a major concern that would lead us to believe this two-way model could not be used for the data.

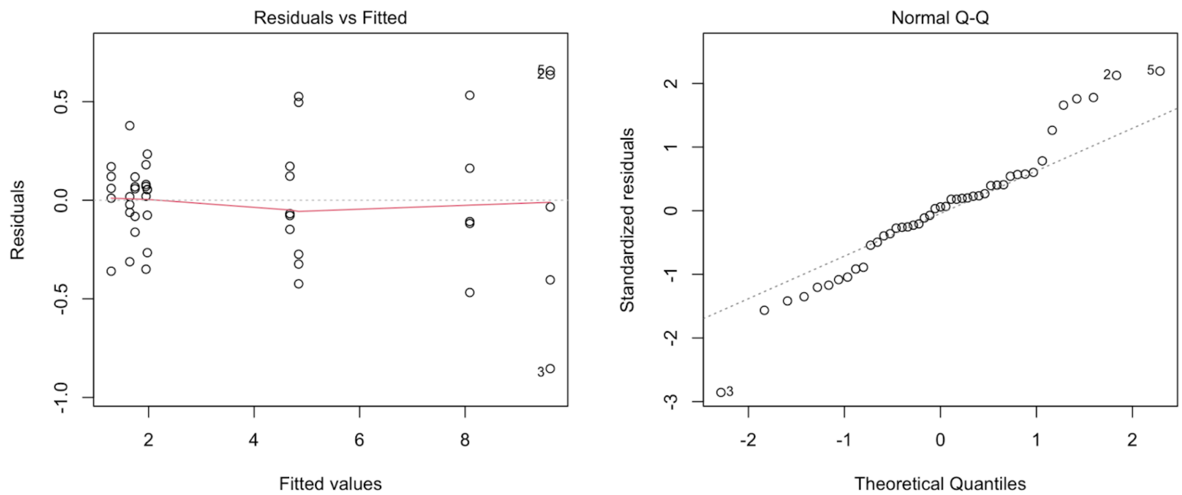

Figure 3: Residuals versus Fitted Plot (left) and Normal QQ Plot (right) for the final model.

Results

For this ANOVA model, we are testing the hypotheses for the interaction. The hypotheses are H_0: $\alpha\beta_{11} = \alpha\beta_{12} = \alpha\beta_{13} = \alpha\beta_{21} = \alpha\beta_{22} = \alpha\beta_{23} = \alpha\beta_{31} = \alpha\beta_{32} = \alpha\beta_{33} = 0$ and H_a: at least one $\alpha\beta_{ij} \neq 0$.

Response: Time

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Ball	2	188.151	94.075	842.03	< 2.2e-16
Floor	2	114.852	57.426	514.00	< 2.2e-16
Ball:Floor	4	75.494	18.873	168.93	< 2.2e-16
Residuals	36	4.022	0.112		

Table 5: ANOVA table for final model

Using the ANOVA table for our model (Table 5), we can determine if the interaction between the factors is significant. With df = (4,36), F-statistic = 168.93, and p-value < 0.05 = significance level, we can reject the null hypothesis since we have significant evidence that the mean bounce times for different combinations of ball and floor types are not equal.

Given that the interaction effect is significant, we can analyze the pairwise comparisons on two-way interactions using Tukey's HSD:

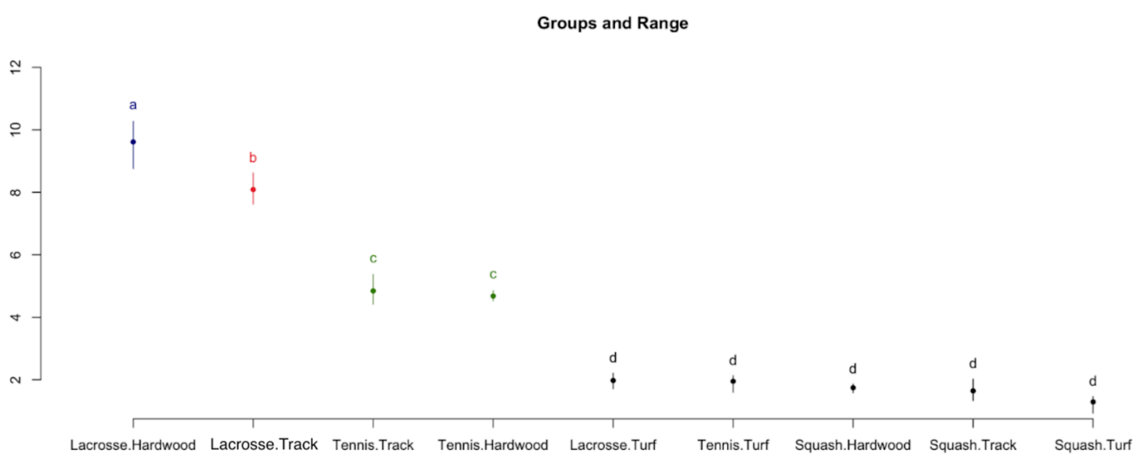

Figure 4: Tukey HSD plot for final ANOVA model

The Tukey HSD test (fig. 4) shows that Lacrosse.Hardwood had a bounce time that was the longest and was statistically different from all other combinations of Ball and Floor. In second place was Lacrosse.Track which was also statistically different from the other combinations. Tennis.Track and Tennis.Hardwood had statistically very similar bounce times and tied for the third-longest bounce time. Finally, Lacrosse.Turf, Tennis.Turf, Squash.Hardwood, Squash.Track, and Squash.Turf all came in last and had statistically similar bounce times as they were all a part of the same group.

This plot also shows us that the surface turf and the squash ball have a strong negative impact on overall bounce time. Because these lead to low bounce times, there is also less variability between the factors interacting with squash or turf. This leads to all of the factors in group d being statistically similar. The magnitude of the effects of ball and turf leads us to believe that the interaction we found in our interaction plots (fig. 2) is in part explained by the impact of the squash ball and turf on bounce time. This would be an interesting topic for further research.

The lacrosse ball bounce time was statistically significant on different floor types. The mean bounce time on hardwood is in group a, on the track is in group b, and on the turf is in group d. Therefore, all lacrosse ball bounce times are statistically different from each other for each floor type. Additionally, the mean bounce times of all balls on the hardwood floor are statistically different from each other as Lacrosse.Hardwood is in group a, Tennis.Hardwood is in group c, and Squash.Hardwood is in group d. Similarly, the mean bounce times on the track had statistically significant differences for all of the ball types with Lacrosse.Track in group b, Tennis.Track in group c, and Squash.Track in group d. The tennis ball's mean bounce time is statistically different when on turf (in group d) compared to when on the track and the hardwood (both in group c).

Addressing our Scientific Hypothesis:

Before conducting our experiment, we hypothesized that the tennis ball on the hardwood squash court would have the longest bounce time. Given our results, our hypothesis was not supported as the tennis ball on the hardwood had the fourth-longest bounce time out of all combinations with a mean time of 4.4678 seconds. The longest bounce time was the lacrosse ball on hardwood with a mean time of 9.614 seconds.

While we anticipated the tennis ball to have the longest bounce time of all three balls due to it being the least dense, our findings actually show that bounce time regardless of floor type is greatest for the lacrosse ball, followed by the tennis ball, and then the squash ball. We now know that greater density does not necessarily lead to greater mean bounce time.

We assumed correctly that when only focusing on floor type the hardwood squash court would have the longest bounce time. The track had the second-greatest bounce time and the turf field had the lowest.

Scope of the Results:

Since all balls we used were regulation size, our findings should be able to be applied to other squash, tennis, and lacrosse balls as long as they too are regulation size. We can also confidently say that the surfaces of the squash court, the outdoor track, and the turf football field are standard given the importance of consistency in college athletics. Wear and tear on the floors and balls could lead to slight deviations in bounce time but all in all our findings should be able to be applied to the wider population of sports balls and athletic fields and courts.

However, we are hesitant to make this generalization because of the flaws of our randomization. Because we cannot fully meet the conditions, we should not take our results to be a true representation of the population at large. Rather, these conclusions should only be generalized to this specific sample. If we had met these conditions and sufficiently controlled potential human error, we would expect that our results would be representative of the larger population stated above.

Further Discussion

There were several elements of our experiment that if we had more time, we would have done differently in order to create more accurate results. Firstly, we would have chosen the day to conduct our experiment based more on the weather. On the day that we chose, there was a slight breeze and the ground was a little wet. If we had more time, a day this past weekend would have been better to conduct our experiment as the weather was much nicer and there was no breeze. Furthermore, if we were to conduct this experiment again with more time we would have conducted more trials. More trials likely would have made the normality assumption easier to meet. More trails are always a good idea when conducting an experiment. In regard to conditions, human error is certainly a potential error in our experiment, especially given the flaws of our randomization process. If we could do this experiment again we would ensure that each trail was an individual in the process of randomization. The last issue we encountered had to do with how the ball bounced on the surface. We were not too specific about where we dropped the ball and the ball would bounce in different directions. This meant that the spot on the surface that the ball bounced on was not perfectly controlled. If we were to do this experiment again, we would have been more specific about where we dropped the ball and perhaps would have tried to control the angle a bit more and where the ball bounced after we dropped it. This would have prevented different areas of the surfaces, which may, over time, have been impacted differently by their use, impacting the results of our experiment. Ultimately, if we had had more time and resources to conduct our experiment, we likely would have been able to control for external effects a bit more. However, we are satisfied with the results of our experiment and feel confident that we did the best that we could given our constraints.

Conclusion

Our original hypothesis did not seem to hold completely. Specifically, our hypothesis failed because it carried with it the assumption that ball and floor would behave as additive factors. The interaction plot showed that this behavior was far more complex. The order of floor types that we hypothesized was correct and the order of ball types that we hypothesized was almost correct. Because of this, our experiment was most surprising in how floor type and ball type interacted. Unfortunately, the scope of our results is greatly hindered by our incomplete randomization process. However, we feel that our results present sufficient evidence to suggest there is some kind of interaction between sports surfaces and sports balls. We suspect research, taking into account our suggestions in the discussion section of our paper, would find similar results.

References

"Sports Ball Size Comparison." Topend Sports, https://www.topendsports.com/resources/equipment/ball-size.htm. Accessed 9 November 2023.

"Sports Ball Weight Comparison." Topend Sports, https://www.topendsports.com/resources/equipment/ball-weight.htm. Accessed 9 November 2023.