Shooters Shoot... But Should They Have Passed?

Do Assists Matter? A Position-by-Position Analysis Using NBA Data

Samantha Glick, Emily Wang, Grace Yim

1 Introduction

In basketball, the label of "selfish" is often used to criticize players who prioritize scoring and individual performance over passing and team-oriented play. These athletes may dominate possession, take difficult shots, and appear unwilling to pass to teammates. Stars like Luka Doncic, James Harden, and Russell Westbrook, and Carmelo Anthony have all been called selfish at some point in their careers for their scoring-heavy styles of play.

But is being selfish always a bad thing? What if creating your own shot — instead of relying on teammates — actually leads to better performance or more wins?

Our project aims to explore two main questions:

- Does shot creation — assisted or self-made — affect the success of that shot?
- Is this relationship dependent on the player's position?

To answer this question, we look specifically at shot success from the current 2024 - 2025 season for assisted vs. unassisted shots by using a Multilevel Logistic Regression model. By incorporating player position and using multilevel modeling techniques to account for player-level differences, our attempts to answer the question of if being "selfish" on the court can be a strategic move or a player's downfall.

2 Data

We used NBA play-by-play and player box score data from the hoopR R package, focusing on all regular-season games from the 2024–2025 season. The load_nba_pbp() function provided detailed event-level information for each game, including all shot attempts, while load_nba_player_box() provided player-level summaries such as player names, positions, and season-long statistics. Our primary dataset was constructed from the play-by-play data. We filtered events to retain only field goal attempts, removing free throws and other non-scoring plays that do not include assists in a shot's success. For each shot, we created several key binary variables from the data:

- **Field_goal**: Field goal attempt (1) vs Non Field Goal Attempt (0)
- **assisted_shot**: Assisted_shot (1) vs Self-created Shot (0)
- **shot_success**: Field Goal made (1) vs. Missed Field Goal Attempts (0)

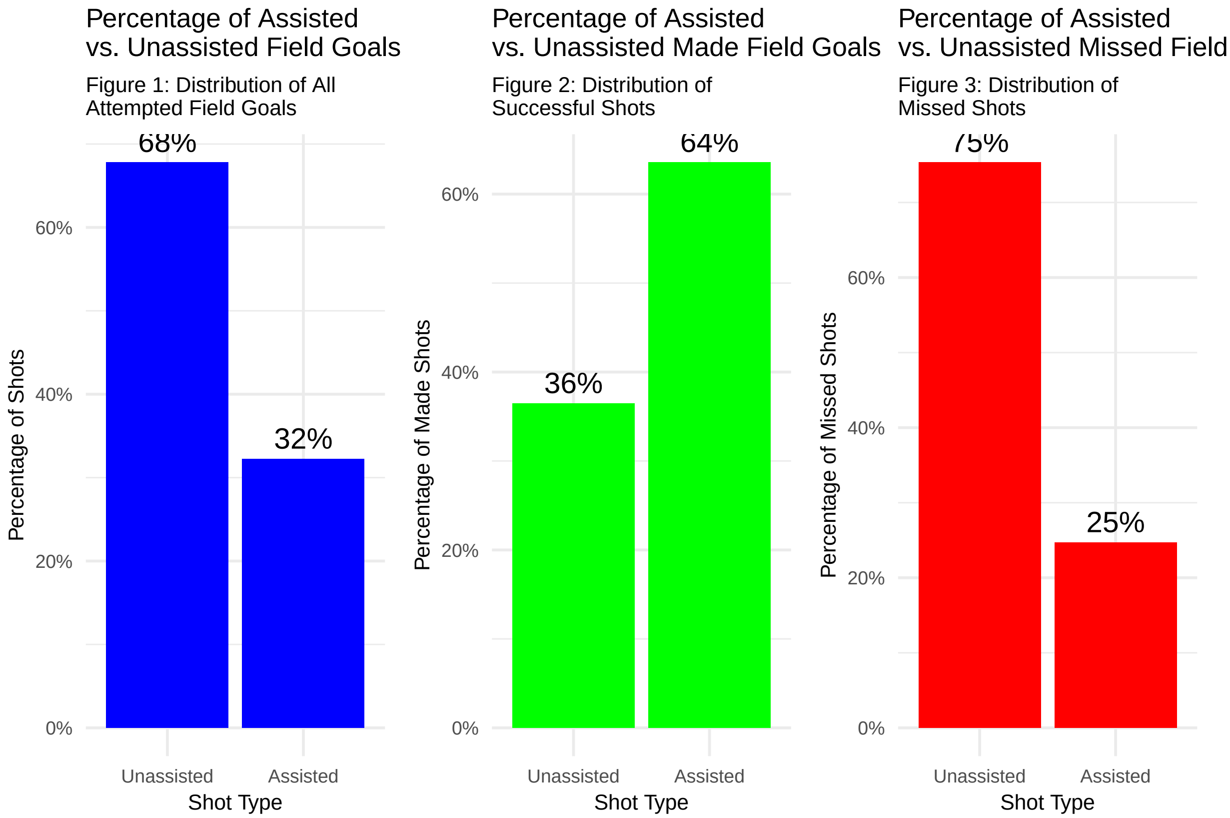
We then merged the shot-level data with player box score data, using the game ID (game_id) and player ID (athlete_id) as the key. This allowed us to link the athlete name (athlete_display_name) and their position (athlete_position_name) to each shot made in the play by play data.

Additional preprocessing steps included:

- Excluding any shot attempts for which assistance data was missing or ambiguous.
- Restricting the dataset to players with valid and non-missing position labels.

The first few lines of the dataset we use in the upcoming model is shown below:

athlete_display_name	athlete_position_name	Assisted_Shot	Shot_Success
Karl-Anthony Towns	Center	1	0
Karl-Anthony Towns	Center	1	1
Tobias Harris	Forward	1	0
Josh Hart	Shooting Guard	0	0
OG Anunoby	Small Forward	1	1
Tim Hardaway Jr.	Small Forward	0	0



To better understand the relationship between shot creation and shooting success, we performed an exploratory analysis of the distribution of assisted and unassisted field goals.

First, we examined the overall distribution of shot types. As shown in Figure 1, the majority of shots in the dataset were unassisted, with approximately 68% of field goal attempts being self-created and the remaining 32% being assisted. This suggests that solo shot taking is a common feature of NBA offensive play.

Next, we analyzed the success rates of assisted versus unassisted shots. Figure 2 displays the proportion of made field goals by shot type. Assisted shots had a noticeably higher success rate compared to self-created shots; furthermore, assisted shots accounted for a majority of the successful field goals, despite representing a smaller proportion of total attempts. We also explored missed shots separately. Figure 3 shows that unassisted shots make up the majority of missed field goals.

Comparing these figures highlight the clear differences in outcomes between assisted and self-created field goal shots. This difference motivates the question of how big of an impact assists are in shot success; furthermore, breaking this data down into position-specific roles can potentially expose other interesting relationships between the two.

3 Methodology

To investigate the relationship between shot creation and shot success, we used a multilevel logistic regression model. This approach is appropriate for our research question because the outcome variable we created - whether a shot was made or missed - is binary; furthermore, the play by play data we used exhibits a nested structure: each player attempts multiple shots throughout the season, and observations within a player are likely to be correlated.

We modeled the probability of a successful field goal (Shot_Success) as a function of whether the shot was assisted (Assisted_Shot), the player's primary position (athlete_position_name), and the interaction between the two variables to capture how the effect of being assisted on shot success differs across positions. We used the glmer() function from the lme4 R package to model the binary outcome, and included a random intercept for each player (athlete_display_name) to account for player-level differences in baseline shooting ability. Doing so allows each player to have their own baseline probability of making a shot.

Since player position is a categorical variable, we chose the reference level to be set to the Point Guard position. Point Guard was chosen as the baseline because our questions centers around assists and team play, and Point Guards start out with the ball and therefore, set most of the playmaking in the game; additionally, Point Guards take the most varied types of shots (drives, three pointers, etc.). Setting Point Guard as the baseline helps us to answer the question of "How does being a certain position change the relationship between assisted shots and success, compared to Guards?".

The model we created can be expressed as:

$$\text{logit}\left(\Pr(\text{Shot\_Success}_{ij} = 1)\right) = \beta_0 + \beta_1\,\text{Assisted\_Shot}_{ij} + \beta_2\,\text{Position}_{ij} + \beta_3\,(\text{Assisted\_Shot} \times \text{Position})_{ij} + u_j$$

- β_0 = overall intercept
- β_1 = effect of an assisted shot
- β_2 = effect of player position
- β_3 = interaction between assisted shot and position
- u_j = random intercept for player j

Below is the exact code we used given the written out model using the glmer library:

```
model_fg_interaction_test <- glmer(
  Shot_Success ~ Assisted_Shot * athlete_position_name + (1 | athlete_display_name),
  data = nba_pbp_mod,
  family = binomial
)
```

3.1 Multilevel Model Assumptions

Our multilevel logistic regression model relies on several key assumptions:

- Binary Outcome: The response variable, Shot_Success, is binary (made or missed), satisfying the fundamental assumption for logistic regression.
- Nested Data Structure: Observations (shots) are nested within players, meaning multiple shots come from the same player. Including a random intercept for each player accounts for this clustering and helps maintain the independence assumption across players.
- Conditional Independence: After accounting for player-level random effects, we assume that the remaining variation in shot success is independent across shots.
- Normality of Random Effects: The random intercepts, which capture player-specific baseline shooting ability, are assumed to follow a normal distribution around the global mean. Given our large dataset with hundreds of players, this assumption is reasonable.
- No Perfect Multicollinearity: The predictors (assist status, player position, and their interaction) must not be perfectly correlated. Based on exploratory analysis, no issues of severe multicollinearity were detected.

- Linearity in the Log-Odds: Logistic regression assumes a linear relationship between predictors and the log-odds of success. Since our predictors are either binary or categorical, this assumption is appropriate for our model.
- Sufficient Sample Size: Each random effect (player) requires a sufficient number of observations to be estimated reliably. With nearly 500,000 shot-level observations, our data easily satisfies this condition.

3.2 Quantifying Uncertainty in Model Estimates

To quantify uncertainty around our effect estimates, we computed 95% confidence intervals for all fixed effect coefficients. In the context of a logistic regression, this means that if we were to repeatedly sample new datasets under the same conditions, approximately 95% of the intervals constructed this way would contain the true parameter value.

Since this is a multilevel model, the confidence intervals account for uncertainty from both within-player shot-level variation and between-player differences through the random effects structure. The validity of these intervals relies on standard assumptions: correct model specification, approximate normality of estimates, and conditional independence. Given our large sample size and model structure, these assumptions are reasonably satisfied.

4 Results

The model reports coefficients as log-odds, so we first calculated them into odds ratios for ease of interpretation as shown below.

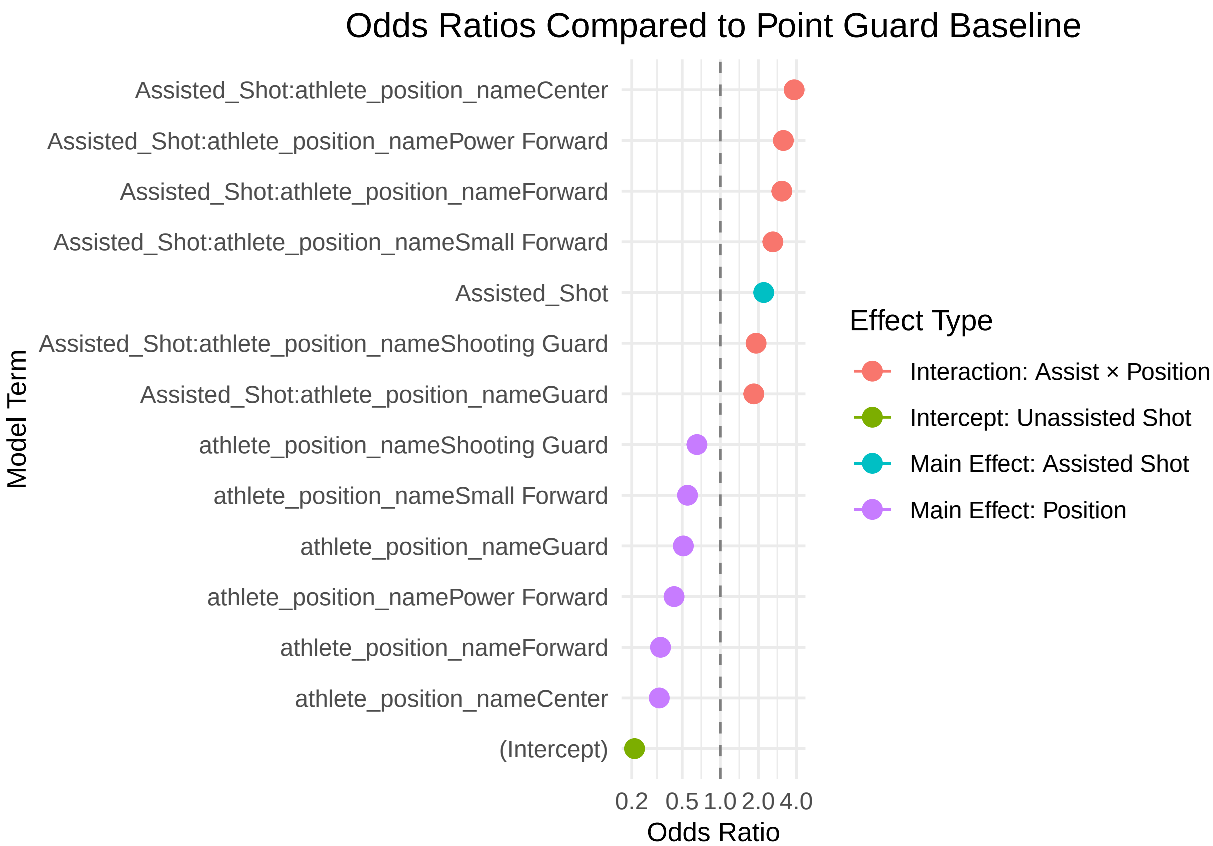


Table 1: Odds Ratios from Multilevel Logistic Model

Term	Odds Ratio	Std. Error	95% CI Lower	95% CI Upper	p-value
(Intercept)	0.2102620	0.0426811	0.1933885	0.2286078	<0.001
Assisted_Shot	2.2091160	0.0185558	2.1302168	2.2909375	<0.001
athlete_position_nameCenter	0.3297265	0.0577697	0.2944287	0.3692559	<0.001
athlete_position_nameForward	0.3373204	0.0575222	0.3013558	0.3775770	<0.001
athlete_position_nameGuard	0.5109267	0.0549835	0.4587294	0.5690634	<0.001
athlete_position_namePower Forward	0.4317553	0.0625507	0.3819394	0.4880686	<0.001
athlete_position_nameShooting Guard	0.6549757	0.0575901	0.5850654	0.7332397	<0.001
athlete_position_nameSmall Forward	0.5510781	0.0642210	0.4859013	0.6249973	<0.001
Assisted_Shot:athlete_position_nameCenter	3.8438628	0.0274606	3.6424480	4.0564150	<0.001
Assisted_Shot:athlete_position_nameForward	3.0724242	0.0295964	2.8992700	3.2559197	<0.001
Assisted_Shot:athlete_position_nameGuard	1.8444506	0.0275323	1.7475578	1.9467157	<0.001
Assisted_Shot:athlete_position_namePower Forward	3.1583842	0.0284173	2.9872815	3.3392871	<0.001
Assisted_Shot:athlete_position_nameShooting Guard	1.9225878	0.0251230	1.8302121	2.0196260	<0.001
Assisted_Shot:athlete_position_nameSmall Forward	2.6096218	0.0283064	2.4687842	2.7584937	<0.001

Table 2: Variance and Standard Deviation of Random Effects

Group	Effect	Variance	Std. Dev.
athlete_display_name	(Intercept)	0.101	0.318

Using point guards as the reference group and unassisted shots as the baseline, the intercept odds ratio of **0.210** indicates that the odds of a point guard successfully making an unassisted shot are relatively low. On the other hand, receiving an assist substantially improves shot success, with the odds of making a shot increasing by a factor of **2.209** when assisted compared to unassisted.

For unassisted shots, the other positions (Guard, Center, Forward, Power Forward, Shooting Guard, Small Forward) have significantly lower odds of making unassisted shots compared to Point Guards. All positions have odds ratios less than 1, indicating worse shot success compared to Point Guards when unassisted.

The interaction effects in the model show that assisted shots, while already beneficial for point guards, are even more beneficial for non-Point Guards. For instance, Centers gain the largest advantage from assists, with their odds of shot success increasing by a factor of **3.843** compared to Point Guards. Forwards **(3.072)** and Power Forwards **(3.158)** also benefit significantly from assists. For Shooting Guards and Guards, the odds ratio is smaller compared to Point Guards, but still meaningful with Shooting Guards odds of shot success relative to Points Guards increasing by **1.923** and Guards increasing by **1.844**.

The estimated variance of the random intercept for players was 0.101, corresponding to a standard deviation of 0.318. This indicates that, even after accounting for shot type and player position, there remains some variability in each player's baseline probabilities of making a shot.

5 Discussion

Our analysis investigated how shot creation - whether a shot was assisted or self-created - impacts shot success in the NBA. Using a multilevel logistic regression model, we found strong and statistically significant evidence that assisted shots are associated with a substantially higher probability of shot success compared to self-created shots. These results show that assists boost shot success across all positions, and the effect is particularly pronounced for certain positions like Centers and Forwards. This makes sense given that these players are larger and typically shoot closer to the net, so they most likely rely more heavily on assisted plays to score successfully. Point Guards and Shooting Guards, on the other hand, are relatively more capable of shooting successfully when unassisted, which suggests that much of their successful shots could be from far away when taken solo.

To summarize our findings into conclusive statements:

- Assists improve shot success across all positions
  - Centers and Forwards benefit the most from assisted plays
  - Shooting Guards behave similarly to Point Guards for assisted plays

- Out of all positions, Point Guards are best at creating and making unassisted shots

Overall, our findings reinforce the strategic value of assists for all positions in shot success, and especially highlights the advantage that assisted shot creation provides for certain players like Centers and Forwards.

However, several limitations should be considered when interpreting our findings. While we captured whether a shot was assisted, we did not control for other important contextual factors that may also influence shot success such as shot distance, type of shot, or the time left on the shot clock. Additionally, the binary classification of shots as assisted or unassisted oversimplifies the concept of shot creation, where degrees of help or defensive pressure vary.

Future work could extend this project by incorporating spatial tracking data to better control for shot context and difficulty. It would also be valuable to examine whether the advantage of assisted shots persists in high-leverage situations, such as in the final minutes of close games or whether there are quarter to quarter differences. Furthermore, expanding the analysis to include detailed positional subcategories (e.g., wings vs. true centers) could provide deeper insights into how different player roles interact with assisted scoring opportunities.