What Makes Blocks Good? Revisiting Blocks in the NBA in a New Light

Analysis Completed by Kenny Huang and Brian Lin Report Written by Kenny Huang

May 3, 2022

1 Introduction

The field of sports analytics is rapidly growing, and with it comes a direct challenge to traditional statistics. In basketball, there are seven basic individual metrics: points, rebounds, assists, steals, blocks, player fouls, and turnovers. Every game has its box score, a detailed table of how many of each metric each player recorded that game. Until the introduction of advanced metrics in the mid-1990s, most of player evaluation was done by eye, but what little statistical evaluation that was done was based entirely on these seven numbers. In recent years, analysts have created proprietary, complicated, and frankly confusing metrics like RAPTOR from combinations of the seven as well as a trove of additional data collected for this purpose. However, in this paper, we seek to take a different approach and instead examine the game more closely through the lens of an existing statistic: blocks.



Figure 1: A defensive player (6) blocks a shot by an offensive player (22)

By definition, a block is a legal deflection of a shot from an offensive player by a defensive player. As one of the two strictly defensive metrics (along with steals), it is commonly used as a gauge for defensive ability. However, we argue that blocks are not a good unit of measurement because not all blocks are created equal; as the metric is defined today, it does not take into account what happens the rest of the current possession— before the other team gets the ball— or the effect on future possessions. In this analysis, we will address both of these limitations and provide an analytical but approachable look at NBA defenses by building on the established statistic of blocks.

2 Related Work

This project was directly inspired by a previous article titled "Bad Blocks", researched and written by myself. While the analysis is comprehensive, I felt that there were a few major shortcomings as well as interesting areas for future research. When Brian and I were given the opportunity to complete a sports analytics project in class, we jumped on the opportunity to build on "Bad Blocks" and create an even more interesting and clear analysis.

3 Data Acquisition

For the purposes of our possession-based analysis, we needed data at the possession level. Unfortunately, there is no publicly available dataset that fine-grained, so we used the Python BeautifulSoup package to scrape the dataset from a website called Basketball Reference. As can be seen in Figure 2, Basketball Reference (BR) has detailed play-by-play records for every game dating back to 1996.



Figure 2: Sample play-by-play data from Basketball Reference

Using the descriptions, scores, and timestamps, we were able to generate a dataset of the 2021-22 NBA season with our desired features, including the following:

- GameID Unique Game ID generated by BR of the game the possession belongs to
- away_home Indicator of whether team in possession was away team
- poss_time Possession time in seconds
- points_scored Number of points scored during possession
- num_orebs_team Number of team offensive rebounds in possession

- num_orebs_ind Number of offensive rebounds by individual players in possession
- num_1shots Number of free throws in possession
- num_2shots Number of two-point shots in possession
- num_3shots Number of three-point shots in possession
- distance Distance from basket of first shot in feet
- num_blocks Number of blocks that occurred in possession
- block_time Time from block to end of possession in seconds; NA if no block occurred in possession
- blocked Unique Player ID generated by BR of player whose shot was blocked in this possession; refers to first block if multiple blocks occurred in possession; NA if no block occurred in possession
- blocker Unique Player ID generated by BR of blocking player; refers to first block if multiple blocks occurred in possession; NA if no block occurred in possession
- block_def Indicator of whether defense secured the ball after the block; refers to first block if multiple blocks occurred in possession; NA if no block occurred in possession

	Α	В	С	D	E	F	G	H	1	J	K	L	M	N	0	Р	
1	GameID	away_hom	ne team	oppt	time	poss_time	block_time	points_scor	num_blocks	num_orebs	num_orebs	num_1shots	num_2shots r	num_3shots	distance	turnover k	bloc
2	202204100N		1 NOP	GSW	48	26		2	0	0	0	0	1	0	14	0	
3	202204100N	(0 GSW	NOP	47.5666667	16		3	0	0	0	0	0	1	27	0	
4	202204100N		1 NOP	GSW	47.3	16		0	0	0	0	0	0	1	23	0	
5	202204100N	(0 GSW	NOP	47.0333333	18		0	0	0	0	0	0	1	25	0	
6	202204100N		1 NOP	GSW	46.7333333	16		0	0	0	1	0	0	2	26	0	
7	202204100N	(0 GSW	NOP	46.4666667	13		0	0	0	0	0	0	1	27	0	
8	202204100N		1 NOP	GSW	46.25	4		2	0	0	0	1	1	0	2	0	
9	202204100N		0 GSW	NOP	46.1833333	10		1	0	0	0	0	0	0		1	
10	202204100N		1 NOP	GSW	46.0166667	5		2	0	0	0	0	1	0	0	0	
11	202204100N	(0 GSW	NOP	45.9333333	20		0	0	0	0	0	0	1	23	0	
12	202204100N		1 NOP	GSW	45.6	1		3	0	0	0	1	1	0	2	0	
13	202204100N	(0 GSW	NOP	45.5833333	13		2	0	0	0	0	1	0	15	0	
14	202204100N		1 NOP	GSW	45.3666667	15		3	0	0	0	0	0	1	23	0	
15	202204100N	(0 GSW	NOP	45.1166667	19		2	0	0	0	0	1	0	13	0	
16	202204100N		1 NOP	GSW	44.8	15		0	0	0	1	0	0	1	25	1	
17	202204100N		0 GSW	NOP	44.55	13		2	0	0	0	0	1	0	17	0	
18	202204100N		1 NOP	GSW	44.3333333	21		0	0	0	0	0	0	1	22	0	
19	202204100N		0 GSW	NOP	43.9833333	10		3	0	0	0	0	0	1	25	0	
20	202204100N		1 NOP	GSW	43.8166667	24		2	0	0	2	0	3	0	2	0	
21	202204100N	(0 GSW	NOP	43.4166667	19		0	0	0	0	0	0	1	25	0	
22	202204100N		1 NOP	GSW	43.1	6		2	0	0	0	2	0	0		0	
23	202204100N		0 GSW	NOP	43	16		0	0	0	0	0	0	1	24	0	
24	202204100N		1 NOP	GSW	42.7333333	11		0	0	0	0	0	0	1	26	0	
25	202204100N	(0 GSW	NOP	42.55	12		3	0	0	0	0	0	1	25	0	
								-	-	-	-	-	-	-			

Figure 3: Sample possession data scraped using BeautifulSoup

3.1 Limitations of Dataset

For simplicity and consistency, we did not include overtime in our dataset. We admit that it would be interesting to look at overtime possessions specifically and compare the results to regulation possessions, but because the scope of this analysis is only the 2021-22 season, the sample size would be too small for us to gain any robust insights. In addition, we would point out that our dataset draws entirely from the BR database. As amazing at BR is, there are some accuracy concerns for such a large open-source project, although the inconsistencies that we found were scarce and uncorrelated with each other.

4 Analysis: Redefining Block Evaluation for Individuals

Our primary basis for this analysis is that the goal of defense is to prevent points. As such, "good" blockers should not be defined as players who accumulate large numbers of blocks but rather as those who minimize the number of points scored by the opposing team. Currently, defensive rating, defined as the average number of points scored by the opposing team per 100 possessions, is already commonly used, but in this analysis we will be modifying this by conditioning on possessions containing a block. That is, on possessions that Player A blocked a shot, how many points did the opposing team score anyway? If the opposing team still scores often, then Player A would not be considered a "good" blocker, even if they block a large number of shots.

But what is the point? What does conditioning on blocked possessions actually do for us? The answer is largely perception correction. Through the "Bad Blockers" article, we realized that many players traditionally viewed as great defenders and blockers turned out to have given up unacceptable numbers of points anyway. In contrast, underrated defenders like Greg Ostertag and Chris Webber excelled in this area and deserve more praise for their defensive efforts. Renowned greats like Tim Duncan reinforced their images by passing both the eye test and numbers test with flying colors. By generating detailed evaluation metrics specifically for the domain of blocks, we provide a tool to find undervalued gems and avoid empty stats. While this analysis on its own does not have many in-game applications, we hope that our work can contribute to the body of work of defensive evaluation, an area that is steadily gaining visibility but is still far too murky.

4.1 Some New Metrics

For this analysis, we defined a player's points per block (PPB) as the expected points scored in possessions where they recorded a block. This can be broken down into two new metrics: offensive rebounding rate (ORR)— the proportion of their blocks that were rebounded by the offense— and points per offensive rebound (PPO)— the expected points scored in possessions after an offensive rebound. Mathematically, because points can only be scored if the block is followed by an offensive rebound, PPB is always equal to the product of ORR and PPO. In addition, we also computed the points blocked (PBK) by each player, defined as the expected points scored by the shots that they blocked had they not been blocked.

4.2 Blocking Efficiency: ORR and PPO

As a whole, the league had an average ORR of about 41% and an average PPO of 1.07. In other words, after 41% of blocks the ball ended up back with offense, who scored an average of 1.07 points in those possessions. For reference, offenses scored an average of 1.12 points overall. Intuitively, these numbers make sense: a majority of blocks are by primary rim protectors, who are taller and after the block now out of position to get the defensive rebound, which explains why the offensive rebounding number is higher than the normal average of 27%. After a block, there is also less time on the shot clock for the offense to shoot if the ball leaves the paint area, resulting in less efficient offense for the remainder of the possession that offsets any benefit from the defender being out of position. In fact, if the ball goes out of bounds and everything resets, the offense with the additional time pressure only scores 0.96 points, compared to 1.19 if the ball is controlled by an offensive player.

Notably, there is only a slim relationship between the two factors with a correlation coefficient of 0.053, meaning that we can treat them as orthogonal when ranking players with them later. When

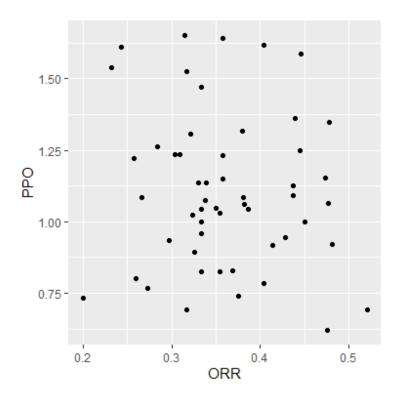


Figure 4: Plot of ORR vs PPO for players with minimum of 50 blocks in 2021-22 season

looking at the 53 players with at least 50 blocks this past season, a strong negative correlation exists between height and both ORR (-0.36) and PPO (-0.59), with slightly weaker correlations of weight with ORR (-0.34) and PPO (-0.27). This aligns with intuition, as height generally helps rebounding more than weight. In terms of blocking efficiency, Isaiah Stewart (DET), Kristaps Porzingis (WAS), and JaVale McGee (PHX) stand out. They are by no means unknown players, but they rarely come up in the conversation for elite rim protection.

Player	Count	Pts	Off_Reb	ORR	PPO
Isaiah Stewart	75	13	15	0.20	0.73
Kristaps Porzingis	82	27	26	0.32	0.69
JaVale McGee	77	18	20	0.26	0.80
Evan Mobley	110	36	30	0.27	0.77
Nikola Jokic	61	19	29	0.48	0.62
Jonas Valanciunas	56	27	13	0.23	1.54
Nic Claxton	50	26	26	0.52	0.69
Hassan Whiteside	101	35	30	0.30	0.93
Darius Bazley	70	26	18	0.26	1.22
Ivica Zubac	74	30	18	0.24	1.61

Figure 5: Most efficient blockers

4.3 Blocking Volume: PBK

Thus far, the metrics that we discussed were purely from an efficiency standpoint. However, even after accounting for outliers by only including players with a minimum of 50 blocks, we can not

definitively say that the "best" blockers are simply the most efficient. In our opinion, volume matters as well, so to more accurately gauge the scale of a player's impact, we computed a new metric called Points Blocked that is the expected points scored by the shots they blocked had they not been blocked. For the sake of simplicity, we made the reasonable assumption that the expected points scored is primarily determined by the distance from the basket, and then computed the average points scored by shots from each distance.

Distance			
0	1.832	18	1.006
1	1.739	19	1.000
2	1.536	20	0.952
3	1.354	21	0.922
4	1.220	22	1.154
5	1.138	23	1.353
6	1.110	24	1.347
7	1.137	25	1.270
8	1.092	26	1.241
9	1.128	27	1.235
10	1.099	28	1.231
11	1.102	29	1.148
12	1.095	30	1.167
13	1.103	31	1.092
14	1.052	32	1.029
15	1.031	33	0.955
16	1.030	34	0.759
17	1.015	35	0.825

Figure 6: Expected points scored in possession by distance of shot

For each blocked shot, we then found the average points scored with similar shots, and a player's points blocked is the sum of all of those values. This gives a better indication of the volume impact than the raw number of blocks. Most of the rankings stay the same, but Robert Williams (BOS) and Myles Turner (IND) were leap-frogged in the rankings by Jakob Poeltl (SAS) and Evan Mobley (CLE), who had less blocks but higher numbers of points blocked. Overall, Jaren Jackson Jr. widely leads the field, followed by Rudy Gobert and Mitchell Robinson.

Player	PBK	Count
Jaren Jackson Jr.	228.54	166
Rudy Gobert	177.02	136
Mitchell Robinson	175.47	131
Jakob Poeltl	157.05	112
Evan Mobley	156.43	110
Myles Turner	155.95	117
Robert Williams	155.08	126
Mo Bamba	139.83	112
Joel Embiid	130.9	96
Hassan Whiteside	130.89	101

Figure 7: Highest volume blockers

As a final note, we also checked the relationship between the efficiency metrics and PBK. Among the players with at least 50 blocks, the correlation coefficients were all 0.20 or less, indicating a weak at best relationship and confirming that it made sense to create these two separate criteria.

4.4 Overall Rankings

Combining our efficiency and volume rankings, we can obtain our overall rankings. Although there are many ways to approach this, we decided on a geometric mean approach. To compute this, we first found all 53 players who had at least 50 blocks during the 2021-22 season and ranked them in each of ORR, PPO, and PBK. Next, we took the geometric mean of the ORR and PPO ranks to obtain the efficiency rank score, and then took the geometric mean of that and the PBK rank, representing the volume criterion, to obtain the final rank score. We thus deemed the players with the lowest final rank scores to be the best overall blockers during this season.

Rank	Player	ORR_Rank	PPO_Rank	PBK_Rank	Overall
1	Jaren Jackson Jr.	41	16	1	5.06
2	Evan Mobley	7	6	5	5.69
3	Rudy Gobert	16	20	2	5.98
4	Isaiah Stewart	1	4	22	6.63
5	Mitchell Robinson	36	25	3	9.49
6	Jakob Poeltl	15	42	4	10.02
7	Kristaps Porzingis	14	2	19	10.03
8	Hassan Whiteside	9	15	10	10.78
9	JaVale McGee	5	8	21	11.52
10	Myles Turner	26	24	6	12.24

Figure 8: Best overall blockers

In our analysis, Jaren Jackson Jr. comes up on top, barely edging out Evan Mobley and Rudy Gobert who round out our top 3. Although Jackson gave up too many offensive rebounds after his blocks, the absurd number of shots he blocked was high enough to eke out Mobley and cement JJJ in our eyes as the blocking champion of the 2021-22 season.

4.5 Alternative: Points Averted

One main drawback of the overall ranking method described above is that it is based on rankings with other players. While this works well enough for finding the best players for a given year like in this article, this poses a few issues. First, it is difficult to compare two players directly. Every time a new set of players is compared, all the calculations have to be recomputed, which is inconvenient and inefficient. More importantly, the rankings themselves are not robust to the comparison set. In other words, adding a single player to an existing ranking set will almost surely shuffle some of the players' rankings, which makes it impossible for this method to look at two similarly rated players and definitively point to which one is better. However, we posit that there is no objective answer key of how to rank such multi-dimensional players in the first place, so in our opinion this is not a problem; if anything, this method is able to separate the best and the worst from the rest of the group, which accomplishes our aforementioned goal of creating "a tool to find undervalued gems and avoid empty stats". In addition, it is simple enough and accessible for the general public to understand, more so than other possibilities, like normalizing the ORR, PPO, and PBK values themselves.

Even so, we created another singular metric that alone accounts for the efficiency and volume metrics. Each player's Points Averted (PAV) is defined as the Points Blocked minus the total points allowed in blocked possessions. In a very simple sense, this is the expected additional number of points that the team would have allowed if the player had not blocked any shots throughout the

season.

Using this metric, we can get another set of rankings that starts remarkably similarly to the previous ranking. It appears that Points Averted favors efficiency less than the overall ranking method, but Jaren Jackson Jr., Evan Mobley, and Rudy Gobert are still the best of the best.

Player	count	pts	orb	PAV
Jaren Jackson Jr.	166	77	71	151.54
Evan Mobley	110	36	30	120.43
Rudy Gobert	136	59	44	118.02
Mitchell Robinson	131	62	50	113.47
Myles Turner	117	47	41	108.95
Jakob Poeltl	112	49	36	108.05
Hassan Whiteside	101	35	30	95.89
Mo Bamba	112	50	37	89.83
Daniel Gafford	95	32	35	89.54
Giannis Antetokounmpo	90	33	24	89.53

Figure 9: Points averted rankings

5 Analysis: Quantifying Effect on Future Possessions

One of the most commonly asked questions about the "Bad Blockers" article was about momentum. With game-changing blocks like LeBron's in the 2016 NBA Finals crystal clear in our memory, we hypothesized that a block positively impacts the game in favor of the team that recorded the block. In other words, blocks shift the momentum in the favor of the defending team in a way that can be numerically detected in the following possessions.

To test this hypothesis, we computed each team's average points per possession for each game and for each block the average points scored per possession in the next x possessions for both teams. Finally, we computed the following values for x = 1, 2, 3, 4, 5:

blocking team: average PPP over next x possessions – average PPP over entire game blocked team: average PPP over entire game – average PPP over next x possessions

We expected both metrics to be negative: the momentum shifts to the blocking team at the expense of the blocked team. However, once we computed them, we found the exact opposite to be true: not only were the averages for both metrics positive for all x, it was significantly positive at the 95% level for x=5. Contrary to expectations, teams scored an average of 1.6% higher after being blocked. To explain why, we came up with hypotheses about shot selection, three-point shooting, and free throw shooting.

5.1 Shot Selection

Our first hypothesis was that after being blocked, players might be more hesitant to attack the rim again and will instead shoot more midrange or three-point shots. After examining the shot distributions of five possessions immediately preceding and following the blocked shot, we concluded that this was almost surely not the cause. Aside from there being little difference visually (Figure 10), we also tested the hypothesis using t-tests and found no significant differences for any distance rounded to the nearest foot.

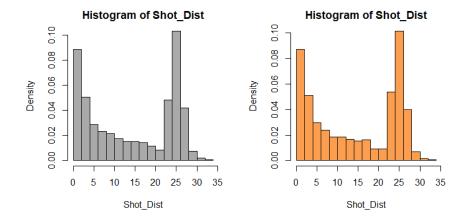
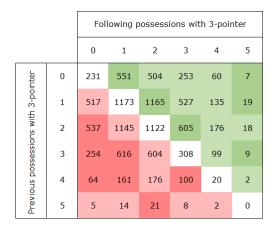


Figure 10: Two shot distributions

5.2 Three-Point Shooting

Another hypothesis more specifically targeted three-point shooting. However, the results were similar to the first hypothesis. In the five possessions before blocks, an average of 1.673 of them involved a three-point shot, compared to 1.650 of the five possessions immediately after, a negligible difference. We also visualized this using a two-step matrix, where the row indicates the count in the previous possessions and the column indicates the count in the following possessions. If the hypothesis were true, this would be reflected in the matrix as having larger numbers in the top right half than the bottom left half of the first matrix in Figure 11, which is not the case.



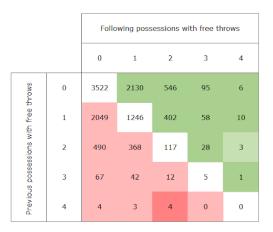


Figure 11: Two-step matrix of three-point and free throw outcomes

5.3 Free Throw Shooting

Our last hypothesis to test was that after blocks, players draw more fouls due to a combination of more aggressive play, crowd pressure on referees, and possibly other factors. Surprisingly, when we computed how many of the five possessions before or after the block involved free throws, the numbers rose from 0.5534 before to 0.5879 after. Statistically, our t-test yielded a z-score of 5.2, corresponding with a very low p-value of 4.8×10^{-7} . A two-step matrix analogous to the one made for three-pointers verifies this finding; the numbers above the diagonal are indeed higher than their counterparts underneath.

This suggests that there could be some truth to this theory. One additional nugget that we found worth noting was that there was an extremely significant difference in blocks between the home and away teams: away teams were blocked on 4.8% of their possessions compared to just 4.5% of home team possessions, amounting to 360 more blocks over the course of the season. The most plausible reason for this discrepancy is that referees subconsciously feel pressure from the fans and allow the home team to be more aggressive in defending shots without being charged with a foul, thus converting what would have been fouls for the away teams into electrifying defensive plays for the home teams. It would be insightful to look further into this to see if the home-court advantage and fan pressure is what is causing this change, but this is outside the scope of this analysis.

6 Conclusion

In this analysis, we looked at the established statistic of blocks in a new light. We re-evaluated who we believe are good and bad blockers based on the direct results of their blocks, and we created two ranking methods that agreed that Jaren Jackson Jr., Evan Mobley, and Rudy Gobert were the best blockers in the 2021-22 season. We explored the numerical effect that blocks had on the flow of the game and examined the changes in play style that resulted regarding shot selection and fouls. As mentioned in the paper, we hoped to improve the way that players are perceived and to share our subjective criteria system of what an ideal defender looks like. As analytics continues to advance in the sport, we hope that this can be a point of inspiration for future analysis.

7 Future Work

As much as we believe in our analysis, there are some limitations or avenues of future research. First, we wanted to investigate the impact of blocks specifically in clutch scenarios (i.e. scoring margin is within five points with five or fewer minutes remaining in the game). As Professor Kornhauser pointed out, these are the most important moments within any game, so momentum would shift the most when teams feel the pressure and are the most focused. Unfortunately, because this is a small sample size, there was too much noise for us to derive anything substantial. However, this is where including the overtime possessions when we initially scraped the dataset would have helped, so this is definitely an area to be revisited in the future. In addition, the expected points value used in the Points Blocked and Points Averted metrics currently depend only on the distance from the basket. Using more features, like the closest defender, catch-and-shoot vs pull-up, or data on the shooter, would improve the model and allow it to more accurately reflect the scale of a player's impact. Finally, there are so many different ways that the final rankings could have be produced from the ORR, PPO, and PBK values, so there absolutely could be a better way that we did not consider yet.

8 Acknowledgements

We would like to thank Professor Alain Kornhauser for giving us this opportunity and providing us with feedback and suggestions. In addition, we would like to thank Basketball Reference for the amazing work they do for sports nerds like ourselves and creating the possibility to complete this project in the first place.