



Refs Revealed: Many NBA Referees are “More Equal” (and Less Equal) Than Others

John Ball
San Francisco, CA

Email: john.ball@beyondboxscores.com

Abstract

This paper summarizes ongoing research to manually review video of almost every NBA referee call (99.5% over the past two NBA seasons) and logging the referees and the violations into a database so that a myriad of referee and foul analytics can be produced. Comparing each NBA referee’s average to the overall average across many types of violations, we find that a number of referees call violations at significant rates above and below the average for each violation type. Additional research has been conducted to determine if NBA teams can leverage referee analytics to gain a competitive edge. By avoiding unfavorable referee tendencies and making use of favorable ones, such analytics can help increase the efficiency of offensive possessions, produce more free throw opportunities, and reduce violations called when playing defense.

1 Introduction

Many NBA fans would think that the path to becoming an NBA referee – and the training involved – is so rigorous that all referees would call violations at nearly the same rate because of their extensive experience and training. However, given the many NBA referee call and “no call” controversies over the years where teams have lost games because of them, we wanted to find out just how similar (or different) NBA referees are in calling violations. In addition, we wanted to see if NBA teams could use this information to help them win more games.

There has been speculation by many observers that NBA referees do have tendencies that set them apart in the violations they call. NBA analyst and former player Greg Anthony stated that referees “really do (have tendencies)” and that “you will know when you’re going into a game what the tendencies of those officials are.” [1] As a former NBA player, Anthony said many players from his era “would keep books on officials,” although we believe this was most likely the anecdotal sharing of referee observations as opposed to formally recording every violation and distributing detailed statistics on each referee. (Anthony did not respond to our inquiries for elaboration).

Before the start of the 2013 NBA Conference Finals, Anthony also stated the ability of players to control the fouls they commit gives teams a winning edge. Specifically, Anthony cited the “four best defensive teams in the league” (Miami, Indiana, San Antonio and Memphis) were in the conference finals “because of their ability to defend without fouling so much.” [2]

NBA fans and media now have access to an immense amount of data about player-specific attributes, but the tendencies of referees have never been well understood or quantified objectively to the general public, or even to NBA coaches and players themselves. That’s due in part because of the tremendous amount of time and resources required to visually review and enter into a database tens of thousands of referee calls throughout an NBA season.

We decided to take on this formidable task of reviewing almost all violations called over the past two NBA seasons (and we continue to log 2013-14 season data to this day). After logging over 108,000 violations from the 2011-12 and 2012-13 NBA seasons, we have determined there are a number of referees who call violations significantly higher and lower than the mean for every violation type. In this paper we explain the degree that these differences exist, and how NBA teams can use this information to try to gain an advantage.

2 Methodology

We have visually reviewed video of almost every violation called by NBA referees (99.5%) over the past two full seasons (2,388 games, including 169 playoff games) as well as compared it to the NBA “play-by-play” (PbP) data. We have discovered that the NBA PbP is not always accurate because of interpretations by game scorekeepers that sometimes differ from the action on the court and/or the referees’ hand signals. Hence, our visual review of each violation instead of just relying solely on the PbP makes our dataset very accurate. The main reason for not being able to log the remaining .5% of violations is a result of television cameras prematurely cutting away from the action sometimes when a violation is called, such as a technical foul or delay of game violation.



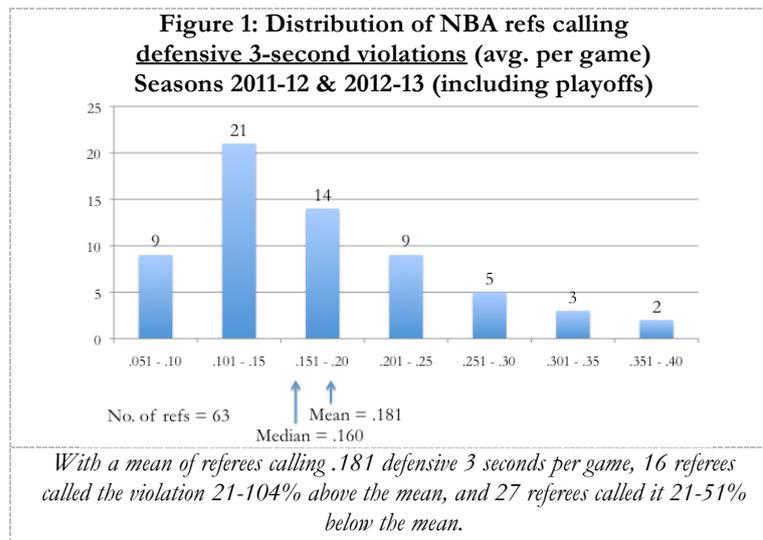
We have also catalogued referee calls that are intentional or extremely obvious so that we can more accurately identify each referee's tendency to call violations on plays requiring a higher level of a referee's discretion and subjectivity. For example, we have broken out intentional fouls from personal and shooting fouls so that the referees' calling of these violations does not unfairly increase their averages for personal and shooting fouls.

In addition, when comparing the number of calls made by each referee to the average number of calls made by all referees, we decided to remove from our calculations the data for seven referees who officiated a small number of games to ensure their averages did not skew the results. Sixty-three referees constituted the regular rotation assigned to games. The average number of games officiated by the 63 referees in this regular rotation for the two seasons was 102.3 games.

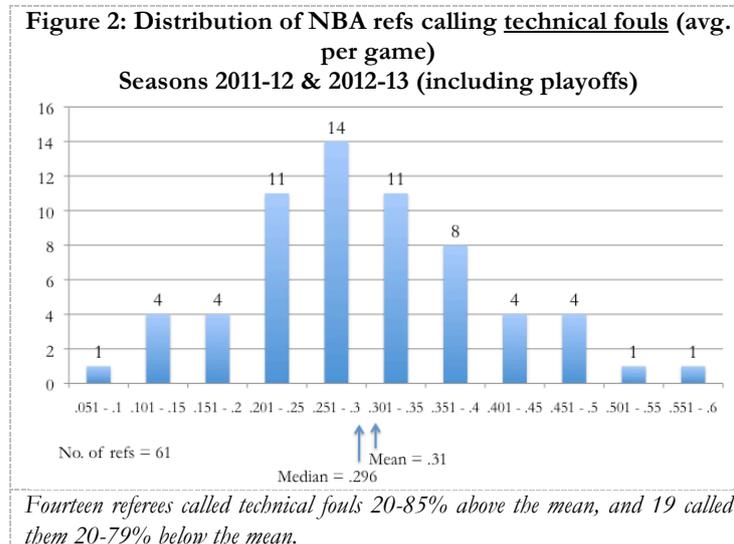
3 Referee Tendencies Quantified

We have determined that many NBA referees call violations with more and less frequency across all violation categories. Violations that require more subjective decision-making (such as non-contact violations compared to those involving obvious contact) make the dispersion of referees even wider. Below are some examples of violation types. (We could have included many others that are similar but was constrained by page limitations).

Defensive 3 seconds - We first start with defensive 3-second violations because of the complexity involved for referees in evaluating if a violation occurred. Referees not only have to count to three seconds "in their head," which can be highly subjective since many referees may count at a different pace, but they also have to observe if a defender has been in the lane for those 3 seconds and not actively guarding an offensive player (also being highly subjective). As you can imagine, these many variables result in a wide frequency distribution histogram as seen in Figure 1 below.



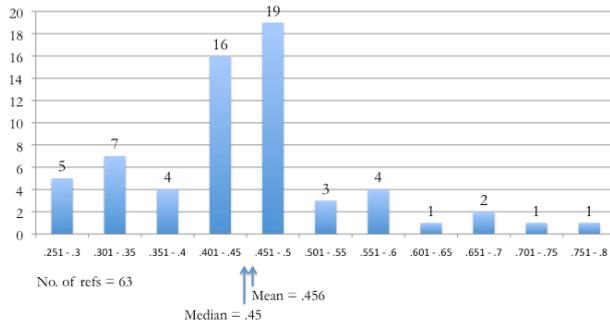
Technical fouls - Another violation that isn't consistent among referees is the calling of technical fouls on players and coaches as seen in Figure 2 below.





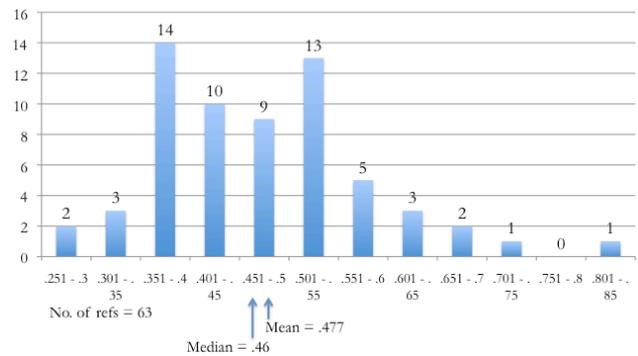
Block-charges - The block-charge scenario is considered by many to be one of the more difficult and subjective calls to make for a basketball official. Referees need to determine if a defensive player was “established” defensively (which can be very subjective), and often times if he was established before the offensive player was in his upward shooting motion. As seen with other violations, the frequency of the referees’ averages is widely dispersed for both blocking violations and offensive charges as depicted in Figures 3 and 4 below.

Figure 3: Distribution of NBA refs calling blocking violations (avg. per game) Seasons 2011-12 & 2012-13 (including playoffs)



Nine referees called blocking violations 22-66% above the mean, and 13 referees called them 20-42% below the mean.

Figure 4: Distribution of NBA refs calling offensive charges (avg. per game) Seasons 2011-12 & 2012-13 (including playoffs)

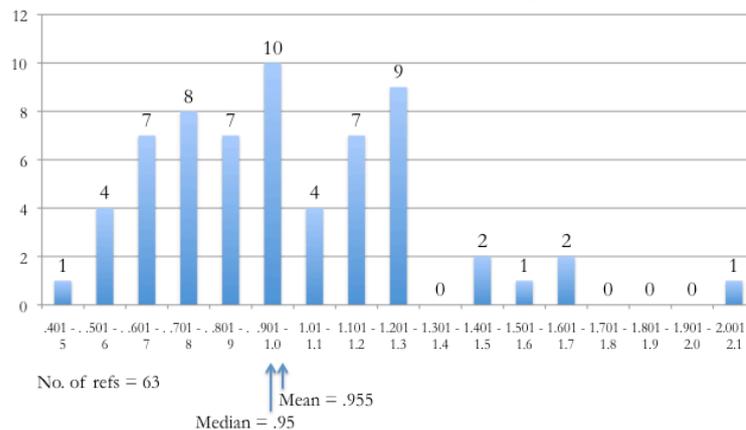


Eleven referees called offensive charges 21-76% above the mean, and 13 referees called them 20-40% below the mean.

NBA fans sometimes hear broadcasters say that the block-charge call can be a “50/50” call where the foul could have been called “either way” by the referee because it is typically very close. Although the average ratio is close to being 50/50 (.955 instead of 1.0), there is a wide disparity in the ratio for many of the referees individually.

Many of the “block-charge” ratios for individual referees do not come close to being 50/50 (at 1.0) as seen in Figure 5 below. Many referees call more charges than blocks, and vice versa.

Figure 5: Distribution of NBA refs’ block-to-charge ratios Seasons 2011-12 & 2012-13 (including playoffs)

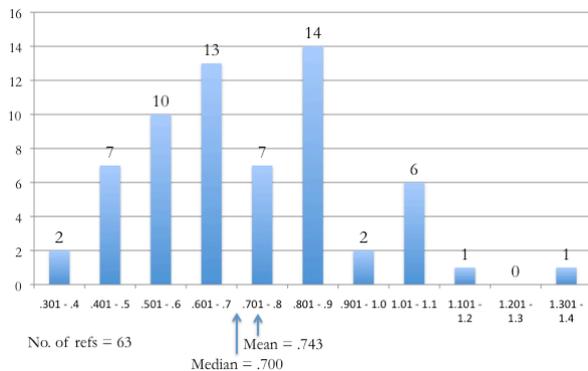


Sixteen referees had a block-charge ratio 25-119% above the mean, and 17 referees had a ratio 20-49% below the mean.

Traveling and illegal screens - We also studied the referees’ tendencies to call **traveling violations** as well as **illegal screens** as seen in Figures 6 and 7 below (we combined travels with double-dribbles and palmings since they are very similar violations).

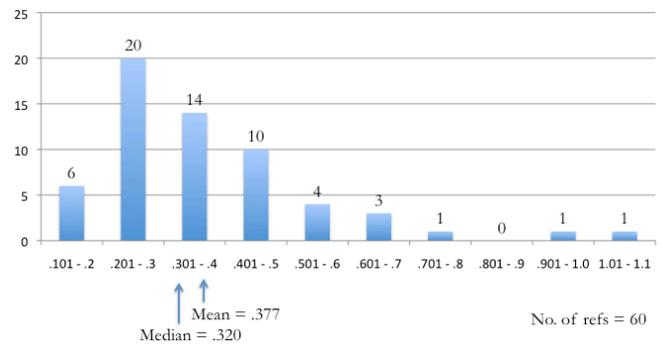


Figure 6: Distribution of NBA refs calling travels, double-dribbles, palmings (avg. per game) Seasons 2011-12 & 2012-13 (including playoffs)



Twelve referees called these violations 21-78% above the mean, and 17 referees called them 21-50% below the mean.

Figure 7: Distribution of NBA refs calling illegal screens (avg. per game) Season 2012-13 (including playoffs)



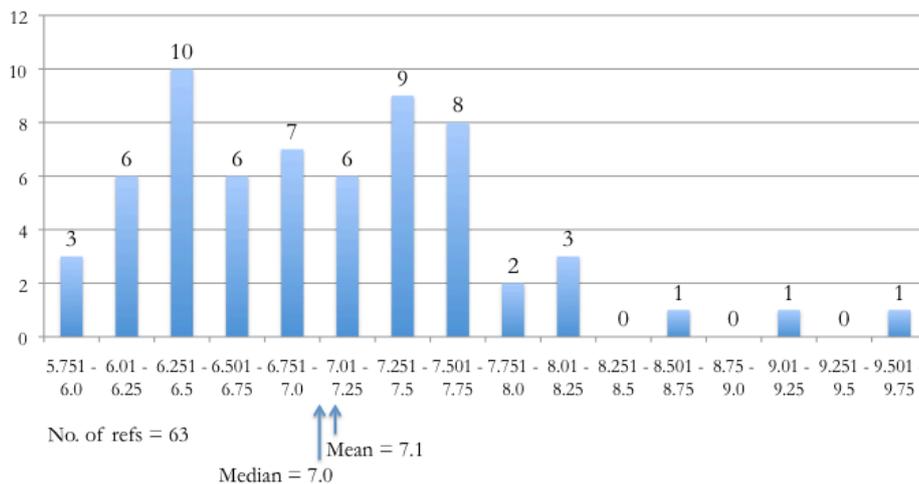
Thirteen referees called illegal screens 30-191% above the mean, and 18 referees called them 31-66% below the mean.

Shooting fouls, Personal fouls and Loose Ball fouls

The majority of violations called by NBA referees (approximately 70%) are shooting fouls, personal fouls and loose ball fouls (with contact from loose ball fouls being very similar to personal fouls). One would assume that judging these violations that entail more contact would be less subjective compared to violations where there is no contact, like defensive 3-second violations, travels, and many technical fouls. As a result, one would think that the referees' averages for these contact fouls would be much more similar to each other. However, we discovered this was not necessarily the case.

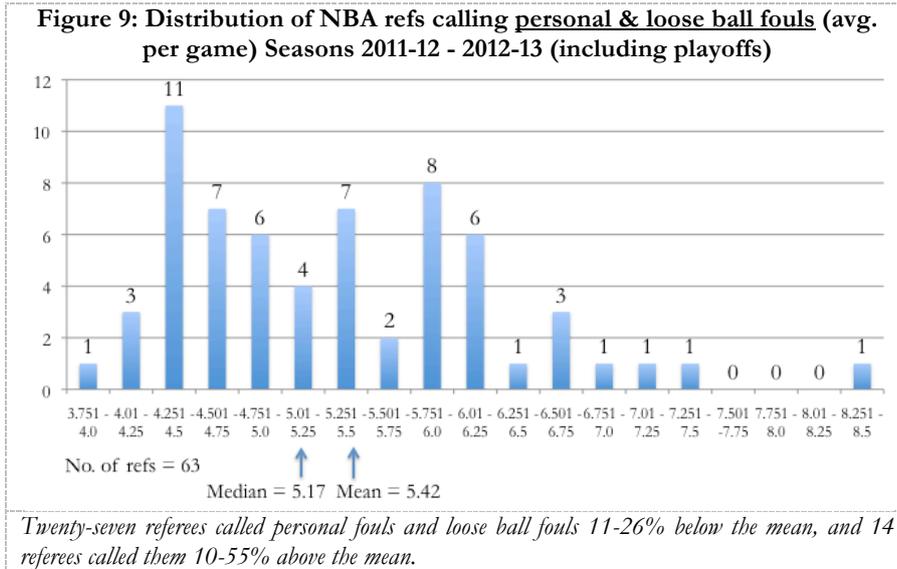
Shooting fouls - The distribution of referee averages for these violations involving more contact is tighter than the previously discussed violations, but not as tight as one would expect as seen in Figure 8 at right.

Figure 8: Distribution of NBA refs calling shooting fouls (avg. per game) Seasons 2011-12 & 2012-13 (including playoffs)



The distribution of referee averages for shooting fouls is tighter around the mean (7.1), but there is still dispersion among the referee averages. Eight referees called them 10-34% above the mean, and sixteen referees call them 10-17% below the mean.

Personal fouls and loose ball fouls - Similarly, the distribution for personal fouls and loose ball fouls is also tighter around the mean (5.42), but surprisingly there is a larger number of referees on the lower end of the distribution than other violation types. See Figure 9 below.



4 Can Referee Analytics Provide an Edge?

“Any coach that wants to do well is going to try to find what little advantages we can gain.” [3]

Brad Stevens, Head Coach - Boston Celtics, October 2013

Understanding that NBA teams are always looking for a competitive advantage, we set out to determine how much of a difference knowing each referee’s tendencies can help in winning more NBA games.

During the 2012-13 NBA season, we logged from one NBA team’s 50-game stretch every referee call and “no-call” that we designated as wrong or questionable and compared these calls to each referee’s tendencies (based on their averages for calling each type of violation). For example, if we observed a referee calling a wrong or questionable traveling call, we referred to his average for calling travels. If the referee’s average was significantly higher than the mean, we flagged it as a violation that might have been avoided if the player knew beforehand the referee’s tendency to call that violation was higher than average.

This review of every call during this 50-game stretch was a 4% sample from every game played during the 2012-13 NBA regular season. Extensive notes of interpretations for each call and no-call were recorded for future audit purposes.

This was a labor-intensive project taking on average 10-12 hours to analyze each game, requiring a total of approximately 700 hours of analysis and compilation. It also required an in-depth understanding of NBA rules to fairly interpret the rules when reviewing each play when a violation was called or not called.

During this 50-game stretch of consecutive games played for this “middle of the pack” NBA team, we discovered some associations between the referees’ questionable calls/no-calls and the team’s outcome in those games.

- The selected team’s opponents benefitted from 562 questionable and wrong calls/no-calls (an 11.24 average), whereas the selected team benefitted from 414 bad or questionable and wrong referee calls/no-calls (average of 8.28 per game). This resulted in a total deficit of 148 calls over the 50-game stretch, an average deficit of -2.96 per game. Interestingly, the selected NBA team had a losing record of 20 wins and 30 losses over the course of the 50 games studied.
- In the twenty games the team won from these 50 games, the questionable and wrong call/no-call disparity between the team and their opponents was an average deficit of -1.8. In the 30 games lost, the disparity was 3.7 per game, a 100% increase.
- The selected NBA team lost ten games by less than 4 points, which we designate as a relatively “close” loss. In all ten of these losses, the average deficit of bad calls/no-calls per game compared to the winning team was -4.4. On the other hand, in the twenty games lost by more than 4 points, the average deficit was -3.4, a 22% reduction. It is arguable that in games that are more closely contested, the number of questionable/wrong calls and no-calls increases because of the heightened intensity and activity in the game. One can also argue that a team is more likely to lose a game if its players cannot avoid or exploit the officials’ tendencies as much as their opponent, especially in close games.

A summary of every questionable/wrong call and no calls from this team’s 50-game stretch is provided in Figure 11 below.



Figure 11: Summary of wrong/questionable calls & no-calls for a select NBA team (50 game stretch: 2012-13 season)

| Violation | Wrong/questionable calls & no-calls | | | | Violation | Wrong/questionable calls & no-calls | | |
|---------------------|-------------------------------------|--------------------|----------------------------|--------------|---------------------------|-------------------------------------|--------------------|----------------------------|
| | Favoring select NBA team | Favoring opponents | Differential from opponent | % difference | | Favoring select NBA team | Favoring opponents | Differential from opponent |
| Shooting fouls | 125 | 157 | (32) | -26% | Flops (fouls were called) | 1 | 6 | (5) |
| Personal fouls | 78 | 130 | (52) | -67% | Defensive goaltendings | 0 | 5 | (5) |
| Travels* | 57 | 71 | (14) | -25% | Technical fouls | 2 | 4 | (2) |
| Offensive fouls | 21 | 38 | (17) | -81% | Kicked balls | 1 | 3 | (2) |
| Loose ball fouls | 22 | 31 | (9) | -41% | Delay of game | 1 | 2 | (1) |
| 3 second violations | 17 | 19 | (2) | | Flagrant fouls | 0 | 2 | (2) |
| Illegal screens | 9 | 17 | (8) | | Backcourt violations | 0 | 2 | (2) |
| Out of bounds call | 10 | 15 | (5) | | 8 second violations | 1 | 1 | 0 |
| Defensive 3 seconds | 13 | 15 | (2) | | Clear path violations | 0 | 1 | (1) |
| Offensive charges | 11 | 14 | (3) | | Jump ball violations | 1 | 1 | 0 |
| Blocking violations | 24 | 15 | 9 | | Away from the play fouls | 0 | 1 | (1) |
| Lane violations | 19 | 12 | 7 | | Shot clock violations | 1 | 0 | 1 |
| | | | | | Totals | 414 | 562 | (148) |

*The number of questionable and missed travels is most likely higher than these numbers, but would probably double the 10-12 hours it takes to break down game's video for every questionable and missed travel. These numbers reflect the most obvious ones.

We believe these findings reveal an association exists between losing games and incurring a deficit of questionable and wrong calls/no-calls that go against a team. A deficit of questionable and wrong calls/no-calls per game ranging between 2.96 to 4.4 can result in more turnovers lost, more lost free throw opportunities and more missed shots per game.

5 How NBA Teams Can Use Referee Analytics

Can a team do anything to reduce the number of questionable and wrong calls/no-calls that go against them each game? Our analysis shows that it's possible. For these 50 games, we compared every referee's questionable and wrong call/no-call to each referee's average for calling (or not calling) each type of violation. We found that 50% of the time the referee's tendencies were in line with making the questionable or wrong call/no-call. This 50% chance is higher than random outcomes that are roughly divided evenly at 33% between a match, no match at all, and a neutral outcome (the referee has no tendency at all for a violation since he calls them close to an average rate).

Although it's not realistic to expect a team can prevent all questionable and wrong calls/no-calls that go against them (an average around 10 per game per team in the 50-game sample described above), if a team knew the referees' tendencies beforehand, there is an opportunity to prevent at least a portion of them. They could start by focusing on those violations that don't require a player to be cognizant of a referee's tendencies in a spontaneous situation, such as a block-charge scenario.

However, if the players knew that all three referees in a game have the same tendency for a particular violation (our studies show this does occur), the players would not need to be aware of which referee has a particular tendency. Even if only two of the 3 assigned referees to a game have the same tendency, the player could decide to "play the percentages" all game long with the strategy that it could work to their benefit throughout the course of the game.

Furthermore, in our sample of 50 games, we determined there was an opportunity to win at least three more games if the referee tendency information had been used in late game situations where the outcome was decided on one possession involving a referee's wrong or questionable call/no-call, with tendencies associated with that referee's call.

6 Conclusions

Even if the most ardent skeptics believe that applying the use of referee tendency data in an NBA game is too challenging because of the amount of information that players already need to keep in mind about their opponent, there are some referee tendencies that can be more easily avoided (or made use of) in situations where players have more time to think about their actions that are less reactive and instinctive to other players' actions. Examples include the following:

| | | | |
|---------------------|-----------------|-----------------|-----------------|
| Defensive 3 seconds | Offensive fouls | Traveling | Lane violations |
| 3-second violations | Illegal screens | Technical fouls | |

Providing to players some of this information in easy-to-remember "chunks" before a game seems like a viable proposition at first without as much fear of overwhelming the players with too much information. As a result, this would seem to be in line with the "little advantage" that Boston head coach Brad Stevens referenced for "any coach that wants to do well" that desires to maximize the efficiency out of every possession.

Over the course of an 82-game regular season, this could add up to several hundred calls per season. With conference playoff positioning always tightly contested, this could mean the difference in getting a higher playoff seed or making the playoffs at all.



References

- [1][2] http://www.nba.com/video/channels/nba_tv/2013/05/22/20130522-delaney-game-time.nba/index.html, May 22, 2013, 4:28 mark
- [3] <http://nesn.com/2013/10/brad-stevens-values-eye-test-as-well-as-advanced-statistics-in-analyzing-basketball-games/>, October 3, 2013.