

Incorrect Non-call Risk and Review Priority in NBA Last Two Minute Reports

Shenlin Tu

*Department of Electrical and Electronic Engineering, University of Nottingham Ningbo China,
Ningbo, China
Zephyr_Xiaote@outlook.com*

Abstract. Incorrect non-calls late in close games may affect possession or free-throw opportunities, but their distribution across event categories and their implications for review priority remain unclear. This study examines which event categories derived from official NBA event names have higher NBA-assessed Incorrect Non-Call (INC) risk. It also explores how modeled risk, rule-constrained impact and event frequency support review prioritization. The analysis uses 42,114 non-call events from 2,760 NBA Last Two Minute Reports covering seven regular seasons from 2019-20 through 2025-26. Logistic regression adjusts category comparisons for remaining time and season. Rule-constrained impact mapping assigns a scoring or possession proxy when the available official information is sufficient to quantify an event. Defensive three seconds, travel/dribble violations and timing/throw-in/lane violations had the largest adjusted odds ratios relative to personal fouls. Shooting fouls contributed the largest share of observed INCs because of their high frequency, showing that within-category risk and total INC contribution capture different dimensions. Report-grouped cross-validation yielded an AUC of 0.712, while the mean leave-one-season-out AUC was 0.699, indicating moderate and broadly stable discrimination. Among quantified events, travel/dribble violations had the highest mean priority score, whereas shooting fouls had the largest priority share. These findings suggest focusing more closely on travel/dribble violations when prioritizing categories, while giving shooting fouls greater attention in overall review planning.

Keywords: NBA officiating, NBA Last Two Minute Reports, incorrect non-calls, logistic regression, review prioritization

1. Introduction

Officiating decisions are especially important in the final two minutes of a close basketball game. An incorrect call or missed call can change possession or lead to free throws, which may affect the game outcome. The NBA's Last Two Minute (L2M) reports make these decisions available for systematic study. For qualifying close games, the reports review calls and material non-calls during the final two minutes of the fourth quarter or overtime and classify each event as correct or incorrect [1].

Sport-officiating decisions are made under contextual and cognitive constraints. A recent review identifies game location, score, previous fouls and remaining time as relevant influences on these judgments [2]. In the field of basketball, quantitative research using officiating records has mainly focused on bias and situational differences. Deutscher compared on-court decisions with the NBA's postgame assessments on a call-by-call basis [3]. Using L2M reports, Belasen *et al.* examined the relationship between betting lines and correct and incorrect calls and non-calls [4]. Pelechrinis used L2M data to study home-team and player/team bias and play-by-play technical-foul data to examine racial bias [5]. In the Chinese Basketball Association, Shu *et al.* found that foul-calling patterns were associated with score differential, foul differential and differences in team quality [6].

Together, these studies show the value of officiating records at the event level, but their primary focus is on bias, team or player advantages and situational differences. There has been little research on how NBA-assessed "Incorrect Non-call (INC)" events are distributed across event categories derived from official event names, or on how these differences translate into review priorities. Therefore, this study examines how INC risk varies across event categories, how risk and the scoring or possession consequence of a missed call can be combined at the event level, and how event frequency affects category-level review priority.

This study analyzes 42,114 non-call events, including 2,854 INCs, from 2,760 regular-season L2M reports covering 2019-20 through 2025-26 NBA seasons. This study maps official event names to 12 categories and compares their observed INC proportions. A logistic model compares INC odds across event categories after accounting for remaining time and season. Rule-constrained impact mapping then uses NBA rules and season-level league statistics to assign a primary impact proxy when the available information is sufficient. For each quantified event, the fitted INC probability is multiplied by the primary impact proxy to obtain an event-level review-priority score. The mean priority scores summarize average event-level priority. The priority shares also account for event frequency and measure each category's contribution to the total score among quantified events. Model performance is evaluated across reports and held-out seasons. The analysis describes associations and review priorities within the close-game settings covered by L2M reports.

2. Methodology

2.1. Data collection

This study collected official NBA L2M report records covering the 2019-20 through 2025-26 regular seasons [7, 8]. Regular-season status was determined from the NBA game identifier. Play-off, play-in and pre-season games were excluded. For the same seven seasons, NBA Stats regular-season team totals for free throws made, free throws attempted, points and possessions were used to calculate the season-specific impact parameters [9, 10].

The collection comprised 2,761 regular-season L2M reports and 55,430 event-level records. The analysis retained events labelled by the NBA as Correct Non-call (CNC) or Incorrect Non-call (INC), excluding 13,257 called events and 59 records without a usable decision label. Because one report contained only a single Correct Call (CC) event, the resulting analysis sample comprised 42,114 non-call events from 2,760 reports: 39,260 CNCs and 2,854 INCs.

The primary outcome was a binary INC indicator, coded 1 for an NBA-assessed Incorrect Non-call and 0 for a Correct Non-call. Remaining time was calculated from the official period and game clock and grouped into 0-30, 31-60 and 61 or more seconds, corresponding to the final 30 seconds, the preceding 30 seconds, and the earlier portion of the L2M interval. Fifteen non-call records fell

outside the nominal 120-second boundary and were retained in the 61-or-more-second group. Season was represented as a seven-level categorical variable.

The primary event-category variable was constructed from the NBA's official event name. This study grouped the 47 non-missing official event names and one missing-name entry into 12 event categories by exact matching (Table 1). One reviewer audited all 48 mappings against the prespecified definitions without access to event frequencies, CNC/INC outcomes, comments or model results. All 48 mappings were approved. Fourteen events without an official event name were assigned to the "Administrative/technical/other" category. Distinct official entries were retained as separate events when multiple assessments occurred at the same game clock.

Table 1. Exact-match mapping of official NBA event names to event categories

Event category	Official event names
Administrative/technical/other	<MISSING>; Foul: Technical; Free Throw Technical; Instant Replay: {Overturn Ruling, Support Ruling}; Stoppage: {Clock, Inadvertent Whistle, TimeOut}; Turnover: Jump Ball Violation; Violation: {Delay of Game, Jump Ball}
Away-from-play foul	Foul: Away from Play
Defensive three seconds	Foul: {Defense 3 Second, Defensive 3 Second}
Goaltending/basket interference	Turnover: Offensive Goaltending; Violation: Defensive Goaltending
Kicked ball/backcourt violation	Turnover: {Backcourt Turnover, Kicked Ball Violation}; Violation: Kicked Ball
Loose-ball foul	Foul: Loose Ball
Offensive foul	Foul: Offensive; Turnover: {Illegal Screen, Offensive Foul}
Out of bounds	Stoppage: Out-of-Bounds; Turnover: {Lost Ball Out of Bounds, Out of Bounds, Stepped out of Bounds}
Personal foul	Foul: {Clear Path, Double Personal, Personal, Personal Block, Personal Take}; Turnover: Double Personal
Shooting foul	Foul: Shooting
Timing/throw-in/lane violation	Turnover: {10 Second Violation, 24 Second Violation, 3 Second Violation, 5 Second Inbound, 5 Second Violation, 8 Second Violation, Inbound Turnover, Lane Violation}; Violation: {Double Lane, Lane}
Travel/dribble violation	Turnover: {Discontinue Dribble, Double Dribble, Palming, Traveling}

Notes: "MISSING" denotes records without an official event name. Braces list official event names that share the prefix before the colon. Clear-path, personal-take and double-personal labels are grouped under personal foul in the event-category analyses, while their rule-specific consequences are retained in the impact mapping. Administrative/technical/other is a heterogeneous residual category.

2.2. Descriptive analysis

For each event category, the observed INC proportion was the INC count divided by all non-call events in that category. The share of observed INCs was the category's INC count divided by all 2,854 INCs. Percentile 95% CIs for the overall and category-specific proportions were obtained from 5,000 season-stratified bootstrap samples, with complete L2M reports resampled to preserve dependence among events in the same report. Proportions were also summarized by remaining-time group and season.

2.3. Incorrect non-call risk model

Let Y_i denote the INC indicator for event i . The primary model was a logistic regression:

$$\text{logitPr}(Y_i = 1) = \alpha + \beta^T G_i + \gamma^T T_i + \delta^T S_i \quad (1)$$

where G_i , T_i and S_i are vectors of indicator variables for event category, remaining-time group and season. Personal foul, 61-or-more-seconds and 2019–20 season were the reference levels respectively. Exponentiating the event-category coefficients produced adjusted odds ratios. For each event category, the odds ratio compares its model-estimated INC-to-CNC odds with those for the personal-foul reference category, holding the remaining-time group and season constant.

Report-level cluster-robust standard errors with a finite-sample correction were used to account for dependence among events in the same L2M report [11].

In season-stratified 10-fold cross-validation, all events from the same report were retained in the same fold. This prevented events from the same report from appearing in both the training and validation samples. This study also left out one season at a time to evaluate model performance on that season. Season indicators were excluded from these leave-one-season-out models because the omitted season had no observations in the training data. These validation schemes preserved the report-level and season-level structure of the data [12].

Out-of-fold performance was summarized by AUC, the Brier score and the calibration slope. AUC measures how well the model ranks INC events above CNC events by predicted risk. The Brier score is the average squared difference between predicted probabilities and observed outcomes, with smaller values indicating better accuracy. The calibration slope indicates whether predicted risks are too extreme or too moderate, with 1 as the target value [13].

2.4. Rule-constrained impact mapping

Rule-constrained impact mapping was used to assign a primary impact proxy when NBA rules supported a scoring or possession consequence. Season-specific free-throw and possession values were used where applicable. Using official NBA Stats team totals [9, 10], the season-specific values of one free throw and one team possession were defined as follows, where j index teams and s index seasons:

$$F_s = \frac{\sum_j \text{FTM}_{js}}{\sum_j \text{FTA}_{js}}, \quad P_s = \frac{\sum_j \text{PTS}_{js}}{\sum_j \text{POSS}_{js}} \quad (2)$$

In Eq. (2), FTM_{js} and FTA_{js} are the numbers of free throws made and attempted by the team j in season s respectively. PTS_{js} is the team's points scored and POSS_{js} is its team-possession count. Here, F_s is the expected point value of one free throw, numerically equal to the league free-throw success rate because a made free throw contributes one point. P_s represents the league-average points scored per team possession. Equivalently, P_s is the aggregate league offensive rating divided by 100, following the standard possession-based measure of basketball efficiency [14].

Official event names identified the applicable rule family. NBA Rules 5, 8, 10, 11 and 12 were used to determine the corresponding scoring or possession consequence [15]. Within each official rule family, comments were used to resolve four event-specific details: whether the text identified a made basket or shot zone, whether an out-of-bounds event explicitly transferred possession, whether

a defensive loose-ball foul was followed by a successful result, and whether an excessive-timeout consequence applied. Loose-ball determinations also used the official possession-team identifier.

The primary impact proxy was used as a rule-constrained, league-average ranking measure of this consequence. Among 12,183 shooting-foul events, comments identified a made basket in 10 events, each assigned F_s , and a three-point context without made-basket evidence in four events, each assigned $3F_s$. The remaining 12,169 events were assigned $2F_s$ in the primary analysis. In the proxy-value sensitivity analysis, these events were assigned F_s and $3F_s$ as the lower and upper values. Defensive three seconds and ordinary single technical fouls were assigned F_s . Events with an explicitly identified excessive-timeout consequence were assigned $F_s + P_s$. Offensive fouls and events labelled as turnovers or possession transfers were assigned P_s .

Among the 177 defensive-goaltending events, five were identified as a two-point shot according to the comments and each of them received a primary impact proxy of 2 points. The remaining 172 events lacked shot-zone information were also assigned 2 points in the primary analysis, while their lower and upper values in the proxy-value sensitivity analysis were 2 and 3 points, respectively. Offensive goaltending events were assigned P_s .

Because the offended team already had possession of the ball, the primary impact proxy for away-from-play and clear-path fouls only included the additional free throws: F_s and $2F_s$ respectively. The upper proxy-value sensitivity scenarios also included the retained possession, giving $F_s + P_s$ and $2F_s + P_s$.

An event was quantified when the available official information was sufficient to identify its main scoring or possession consequence. Comments or possession-team identifiers were used when additional context was needed. The impact proxy was left missing when this consequence remained unclear, mainly for personal fouls with unresolved penalty or special-foul context, loose-ball fouls with unresolved penalty context, and out-of-bounds stoppages with unresolved possession effects.

2.5. Review-priority score

To construct review-priority scores, this study fitted the primary logistic model to all 42,114 non-call events and obtained the full-sample fitted INC probability, $\hat{p}_i$, for each event. For each quantified event with primary impact proxy I_i , the event-level review-priority score was defined as:

$$Q_i = \hat{p}_i I_i \quad (3)$$

The event score Q_i was used to rank quantified events. Higher values reflected a higher fitted INC probability, a larger primary impact proxy, or both. For category g , let $n_g^{(q)}$ denote its number of quantified events. The mean priority score $\bar{Q}_g$ and priority share R_g were defined as:

$$\bar{Q}_g = \frac{1}{n_g^{(q)}} \sum_{i \in g, I_i \text{ quantified}} Q_i, \quad R_g = \frac{\sum_{i \in g, I_i \text{ quantified}} Q_i}{\sum_{i: I_i \text{ quantified}} Q_i} \quad (4)$$

The mean priority score compares the average event score among quantified events across categories. The priority share measures each category's contribution to the total score across all quantified events.

For each event category, proxy coverage was calculated as the number of quantified events divided by the total number of events in that category. Category rankings were reported when at least 30 events were quantified and proxy coverage was at least 25%. The sensitivity of the rankings

to these reporting thresholds was assessed by varying the minimum quantified count over 20, 30, 50 and 100, and the minimum coverage over 25%, 50% and 75%. In the coverage-expansion sensitivity analysis, the 4,570 unresolved loose-ball fouls were assigned P_s .

3. Results

3.1. Event-category variation in INC assessments

Across all 42,114 non-call events, the observed INC proportion was 6.78% (95% CI: 6.53%-7.02%). Observed INC proportions varied substantially across event categories (Table 2). Defensive three seconds had the highest proportion (39.60%), followed by travel/dribble violations (34.52%) and timing/throw-in/lane violations (31.82%). In contrast, shooting fouls had a lower proportion of 5.50%. However, because shooting fouls were the most common category, with 12,183 events, they accounted for the largest share of observed INCs (23.48%). This shows that categories with higher observed INC proportions were not necessarily those contributing the largest share of observed INCs.

Table 2. Observed INC proportions, shares of observed INCs, and adjusted odds ratios by event category

Event category	Non-call events, n	INC n, observed INC proportion, % (95% CI)	Share of observed INCs, %	Adjusted OR (95% CI)
Defensive three seconds	702	278; 39.60 [35.77, 43.53]	9.74	14.76 [12.20, 17.87]
Travel/dribble violation	1,231	425; 34.52 [31.75, 37.21]	14.89	11.98 [10.30, 13.92]
Timing/throw-in/lane violation	773	246; 31.82 [28.59, 35.18]	8.62	10.94 [9.10, 13.15]
Out of bounds	459	65; 14.16 [10.92, 17.52]	2.28	3.87 [2.91, 5.15]
Kicked ball/backcourt violation	178	24; 13.48 [8.62, 18.67]	0.84	3.53 [2.27, 5.49]
Administrative/technical/other	225	30; 13.33 [8.94, 18.10]	1.05	3.90 [2.61, 5.82]
Goaltending/basket interference	249	27; 10.84 [7.17, 14.94]	0.95	2.84 [1.86, 4.33]
Loose-ball foul	4,577	280; 6.12 [5.41, 6.82]	9.81	1.45 [1.25, 1.69]
Shooting foul	12,183	670; 5.50 [5.08, 5.92]	23.48	1.35 [1.20, 1.53]
Away-from-play foul	368	20; 5.43 [3.21, 7.91]	0.70	1.20 [0.76, 1.91]
Personal foul	11,266	480; 4.26 [3.91, 4.64]	16.82	1.00 [reference]
Offensive foul	9,903	309; 3.12 [2.79, 3.46]	10.83	0.72 [0.62, 0.83]

Notes: INC denotes Incorrect Non-call and OR denotes odds ratio. The 95% CIs for observed INC proportions are percentile intervals from 5,000 season-stratified bootstrap samples in which complete L2M reports were resampled within season. Adjusted ORs use Personal foul as the reference category and adjust for remaining-time group and season. Their 95% CIs are based on report-level cluster-robust standard errors. Shares of observed INCs use all 2,854 observed INCs as the denominator. Categories are ordered by observed INC proportion.

After adjustment for the remaining-time group and season, the three categories with the highest observed INC proportions also had the three largest adjusted ORs, in the same order. Compared with personal fouls, defensive three seconds had the highest adjusted OR (14.76), followed by travel/dribble violations (11.98) and timing/throw-in/lane violations (10.94). In contrast, shooting

fouls had a much lower adjusted OR of 1.35, although they accounted for the largest share of observed INCs.

Observed INC proportions were similar across the three remaining-time groups: 6.89% for 0-30 seconds, 7.16% for 31-60 seconds and 6.49% for 61 or more seconds. After adjustment for event category and season, the adjusted ORs (95% CIs), relative to the 61-or-more-second group, were 1.02 (0.92-1.12) for 0-30 seconds and 1.09 (0.99-1.21) for 31-60 seconds. Across seasons, the observed INC proportion ranged from 5.17% in 2024-25 season to 8.64% in 2021-22 season. After the 15 records outside the nominal 120-second boundary were excluded, the largest absolute relative change in an event-category adjusted OR was 4.08%. Therefore, the remaining-time differences were small and excluding these 15 records had little effect on the event-category adjusted ORs.

3.2. Model validation and stability

Report-grouped cross-validation produced an AUC of 0.712, a Brier score of 0.0575 and a calibration slope of 0.990. The AUC indicated moderate discrimination between INC and CNC events. The calibration slope was close to 1. The mean leave-one-season-out AUC was 0.699, indicating similar discrimination in omitted seasons.

3.3. Primary impact proxy and review priority

Under the rule-constrained impact mapping, 25,786 events were assigned a primary impact proxy, giving overall proxy coverage of 61.23%. Season-specific free-throw values ranged from 0.7729 to 0.7843 points and possession values from 1.1009 to 1.1476 points. Across quantified events, the primary impact proxy had a mean of 1.3229 points and a median of 1.1476 points. The remaining 16,328 events had no primary impact proxy, including 11,261 personal-foul records with unresolved team-penalty status, four special personal-foul (two Personal Take and two Double Personal) with no uniquely identifiable rule consequence, and 4,570 loose-ball fouls with unresolved penalty context. These events had no event score and they were not included in the priority rankings.

Among categories meeting the reporting thresholds, travel/dribble violations ranked first by mean priority score (0.3893), followed by timing/throw-in/lane violations (0.3600), defensive three seconds (0.3085), and goaltending/basket interference (0.1899, Table 3, Panel A). This ordering reflected the joint contribution of fitted INC probability and the primary impact proxy. Although defensive-three-second events had the highest mean fitted INC probability (39.60%), their mean primary impact proxy was only 0.779, the second lowest among the ranked categories. As a result, defensive three seconds ranked third by mean priority score.

Travel/dribble and timing/throw-in/lane violations had fitted probabilities of 34.52% and 31.85% and mean primary impact proxies of 1.129 and 1.132 points, respectively. They had the two highest mean priority scores. Goaltending/basket-interference events had a lower fitted probability of 10.84% but the largest mean primary impact proxy among the ranked categories at 1.749 points. They ranked fourth by mean priority score.

Table 3. Mean priority scores and priority shares by event category

Panel A. Categories ranked by mean priority score					
Rank	Event category	Quantified n (proxy coverage)	Mean fitted INC probability	Mean primary impact proxy	Mean priority score
1	Travel/dribble violation	1,231 (100.00%)	34.52%	1.129	0.3893
2	Timing/throw-in/lane violation	711 (91.98%)	31.85%	1.132	0.3600
3	Defensive three seconds	702 (100.00%)	39.60%	0.779	0.3085
4	Goaltending/basket interference	249 (100.00%)	10.84%	1.749	0.1899
5	Out of bounds	331 (72.11%)	14.04%	1.132	0.1586
6	Kicked ball/backcourt violation	91 (51.12%)	13.23%	1.132	0.1495
7	Shooting foul	12,183 (100.00%)	5.50%	1.559	0.0857
8	Away-from-play foul	368 (100.00%)	5.43%	0.778	0.0422
9	Offensive foul	9,903 (100.00%)	3.12%	1.127	0.0351
Panel B. Categories ranked by priority share					
Rank	Event category	Quantified n	Mean priority score	Priority share	
1	Shooting foul	12,183	0.0857	42.18%	
2	Travel/dribble violation	1,231	0.3893	19.37%	
3	Offensive foul	9,903	0.0351	14.05%	
4	Timing/throw-in/lane violation	711	0.3600	10.35%	
5	Defensive three seconds	702	0.3085	8.75%	
6	Out of bounds	331	0.1586	2.12%	
7	Goaltending/basket interference	249	0.1899	1.91%	
8	Away-from-play foul	368	0.0422	0.63%	

Table 3. (continued)

9	Kicked ball/backcourt violation	91	0.1495	0.55%
---	---------------------------------	----	--------	-------

Notes: Panel A ranks categories by mean priority score and Panel B by priority share. A quantified event has a non-missing primary impact proxy. Rankings are reported only for categories with at least 30 quantified events and at least 25% proxy coverage. For categories with incomplete coverage, reported means and score sums use quantified events only. Priority shares use the event-score sum across all 25,786 quantified events as denominator, including quantified events from categories that did not meet the reporting thresholds. Personal foul, loose-ball foul and Administrative/technical/other did not meet the thresholds and are not ranked.

The priority share produced a different ranking because it incorporates both the mean priority score and quantified-event count. Shooting fouls had the largest priority share at 42.18%, followed by travel/dribble violations at 19.37%, offensive fouls at 14.05%, and timing/throw-in/lane violations at 10.35%. Although shooting fouls had a mean priority score of only 0.0857, their 12,183 quantified events offset this relatively low value. Similarly, offensive fouls combined a mean priority score of 0.0351 with 9,903 quantified events and ranked third by priority share.

The priority rankings were stable in the proxy-value sensitivity analysis. The five highest-ranked categories were unchanged for both mean priority score and priority share, and Spearman rank correlations with the primary rankings ranged from 0.983 to 1.000. The priority-share top five also remained unchanged across all reporting thresholds. The top five by mean priority score remained unchanged at coverage thresholds of 25% and 50%. At the 75% threshold, out-of-bounds events were excluded because their proxy coverage was 72.11%, and shooting fouls entered the reported top five. After assigning P_s to the unresolved loose-ball fouls, the top five by mean priority score remained unchanged. However, the priority-share ranking changed: loose-ball fouls moved into fourth place and displaced the defensive three seconds from the top five.

4. Discussion

This study supports review decisions at two levels. For individual-event screening, the fitted INC probability ranks events by modeled risk, while the event score Q_i combines that probability with the primary impact proxy to rank quantified events by review priority. At the category level, the mean priority score identifies the categories with higher average scores for quantified events, while the priority share, which takes the number of quantified events into account, shows in which categories the total scores are concentrated. Therefore, these measures support different tasks: event-level triage and category-level planning. However, the actual resource allocation would also require information on review time and costs.

Several limitations should be considered. First, the findings are restricted to the close-game intervals covered by NBA L2M reports and to the league's postgame CNC/INC classification. Generalization to other game contexts requires separate validation. Second, the risk model is based on event category, remaining time and season, so it does not capture finer differences among events within the same category. Third, the league-average primary impact proxy does not capture event-specific bonus status, possession value, player context, or subsequent game development. Therefore, the resulting scores provide approximate ranking measures based on the information available for each event. Finally, missing primary impact proxies were concentrated in several categories, so differences in proxy coverage affected category-level priority shares. The sensitivity analyses showed that priority share was more sensitive to which events could be quantified than to the choice among plausible proxy values. These limitations suggest two directions for future work. Possession-

level or tracking data could support more event-specific primary impact proxies [16]. The rankings could also be applied in later seasons to examine whether they can improve review efficiency when review time and cost information is available.

5. Conclusion

This study examined how NBA-assessed Incorrect Non-Call (INC) risk varied across event categories and how these differences could be combined with rule-constrained impact information to organize review priorities. Across 42,114 non-call events, both observed INC proportions and adjusted ORs differed substantially across categories. Defensive three seconds, travel/dribble violations, and timing/throw-in/lane violations showed the strongest adjusted associations with INC classification. However, shooting fouls accounted for the largest share of observed INCs because they occurred much more often. Together, these results show that within-category INC risk and total contribution to observed INCs describe different aspects of officiating performance.

For the 25,786 quantified events, the event score combined the fitted INC probability with the primary impact proxy. This extended the category-risk comparison by adding the main scoring or possession consequence of each quantified non-call event to the review-priority analysis. Travel/dribble violations ranked highest by mean priority score, while shooting fouls ranked highest by priority share. This difference is useful in practice: mean priority score identifies categories whose quantified events tend to score higher, whereas priority share shows where the total score is concentrated when event frequency is considered. The results show how INC risk, primary impact proxy and event frequency shape different aspects of review priority. In L2M close-game settings, these measures can help compare quantified non-call events, identify individual events with higher review priority, and show which categories may need more attention in overall review planning. This can help reviewers use limited review time to focus on events with greater potential impact on the game.

References

- [1] National Basketball Association. (2026, July 15). NBA last two minute reports—frequently asked questions. *NBA Official*. <https://official.nba.com/nba-last-two-minute-reports-frequently-asked-questions/>
- [2] Hancock, D. J., & Pizzera, A. (2025). The psychology of sport officiating. *Psychology of Sport and Exercise*, 80, 102899. <https://doi.org/10.1016/j.psychsport.2025.102899>
- [3] Deutscher, C. (2015). No referee bias in the NBA: New evidence with leagues' assessment data. *Journal of Sports Analytics*, 1(2), 91–96. <https://doi.org/10.3233/JSA-150012>
- [4] Belasen, A. R., Belasen, A. T., & Olbrecht, A. M. (2025). With the game on the (betting) line: NBA referee performance in the last two minutes. *Journal of Sports Economics*, 26(8), 913–935. <https://doi.org/10.1177/15270025251369447>
- [5] Pelechrinis, K. (2023). Quantifying implicit biases in refereeing using NBA referees as a testbed. *Scientific Reports*, 13, 4664. <https://doi.org/10.1038/s41598-023-31799-y>
- [6] Shu, F., Zhang, L., Xu, X., & Xu, R. (2026). Situational influences on officiating performance: An analysis of foul calls in the Chinese Basketball Association league (CBA). *PLOS ONE*, 21(2), e0342151. <https://doi.org/10.1371/journal.pone.0342151>
- [7] National Basketball Association. (n.d.). NBA officiating last two minute reports archive. *NBA Official*. Accessed July 15, 2026, from <https://official.nba.com/nba-officiating-last-two-minute-reports-archive/>
- [8] National Basketball Association. (2026). 2025–26 NBA officiating last two minute reports. *NBA Official*. Accessed July 15, 2026, from <https://official.nba.com/2025-26-nba-officiating-last-two-minute-reports/>
- [9] National Basketball Association. (2026). NBA stats: Teams traditional. *NBA Stats*. Accessed July 15, 2026, from <https://www.nba.com/stats/teams/traditional/>
- [10] National Basketball Association. (2026). NBA stats: Teams advanced. *NBA Stats*. Accessed July 15, 2026, from <https://www.nba.com/stats/teams/advanced/>

- [11] Zeileis, A., Köll, S., & Graham, N. (2020). Various versatile variances: An object-oriented implementation of clustered covariances in R. *Journal of Statistical Software*, 95(1), 1–36. <https://doi.org/10.18637/jss.v095.i01>
- [12] Roberts, D. R., Bahn, V., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- [13] Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., & Steyerberg, E. W. (2019). Calibration: The Achilles heel of predictive analytics. *BMC Medicine*, 17, 230. <https://doi.org/10.1186/s12916-019-1466-7>.
- [14] Kubatko, J., Oliver, D., Pelton, K., & Rosenbaum, D. T. (2007). A starting point for analyzing basketball statistics. *Journal of Quantitative Analysis in Sports*, 3(3), 1-24. <https://doi.org/10.2202/1559-0410.1070>.
- [15] National Basketball Association. (2026). NBA Rulebook. *NBA*. <https://official.nba.com/rulebook/>.
- [16] Jutamulia, I., & Hosoi, A. E. (2025). Leveraging expected action value to measure on-court decision-making skills in basketball. *Journal of Sports Analytics*, 11, 1–15. <https://doi.org/10.1177/22150218251325018>.