Analytics vs The Eye Test: Red Sox Playoff Mistakes

There I was at 161st, Yankee Stadium, section 305, watching Chad Tracy pace out to the mound to pull his starting pitcher, Payton Tolle, with two outs in the bottom of the fifth. Well below his per-game average of 91, Tolle had thrown 83 pitches, only giving up one run on five hits with four strikeouts. Everyone was confused about the managerial decision, especially the Boston native sitting in front of me, who loudly shared his true feelings on Tracy. Three pitches later, the Yankees led 2-0.

After the game, while scrolling on X, I came across a fan talking about the “stupid analytically backed move”. Personally, I don’t like the words “stupid” and “analytically” put together, so my goal in this article is to break down the decision and find out which analytics, if any, the move was inspired by. Ultimately, I want to answer the modern-day question: Do analytics overrule the eye test?

My first step was to look at the matchup against Ben Rice. So far this game, Rice was 1-2 against Tolle with a double and a strikeout. Across their career matchups, Rice was just 1-3 with a strikeout. Simply put, there isn’t a large enough sample size to definitively state if Tolle was pulled because of their prior history. While it is plausible that the earlier double was the largest factor in the call to the bullpen, I don’t believe it was the sole reason, especially after Tolle retired 8 of the last 10 batters.

Another important factor in the decision is handedness. Either way, the at-bat was going to be a lefty-on-lefty matchup, whether Tolle stayed in the game or manager Chad Tracy went to Erik Miller.

One defining feature of Peyton Tolle is his reverse splits. Left-handed hitters hold a .704 OPS against him with a 22.5% strikeout rate, compared to righties, who have just

a .624 OPS and a 30.7% strikeout rate. Clearly, lefties have an easier time facing Tolle than righties do. To me, this is the exact analytical backing behind the call to the bullpen and bringing in Miller specifically, who holds lefties to a stifling .536 OPS, only reinforces that logic.

Miller had also proven to be a trustworthy relief arm across his 2026 stints with the Giants and Red Sox, logging 51.1 innings with a 1.38 WHIP, while Rice entered the game 0-for-1 with a strikeout against him. Overall, from a pure matchup standpoint, going to the bullpen was the analytically sound move, but there are deeper variables worth unpacking.

To get a better idea of the bullpen call, we have to look at pitch mix and the pitch selection during the Miller-Rice at-bat.Tolle's arsenal leans heavily on his fastball - 45% four-seamers, 24% sinkers, and 19% cutters - along with low-usage curveballs and changeups. Miller, by contrast, is a two-pitch pitcher, throwing 49% sinkers and 41% sliders.

Looking at Rice's performance against specific pitch types, he leads the league in run value against four-seamers at +24 (.331 BA, .843 SLG, .520 wOBA). Against sinkers and sliders, however, he is far more vulnerable, posting a -2 run value (.368 wOBA) and a +6 run value (.386 wOBA), respectively. Once again, the analytics show that Miller's arsenal is much better suited to neutralize Rice.

In the actual at-bat, Miller started with a slider for a called strike, missed outside with a second slider, and then left a sinker middle-up that Rice drove out of the yard. Even so, no matter how deceptive Tolle's four-seamer might be, keeping Rice off his favorite pitch was an analytically sound move.

The final factor to investigate is one of baseball's classic rules of thumb: a pitcher gets noticeably worse the third time through the order. For Tolle, however, the numbers oppose the trend. Opponents actually post a lower batting average and OPS the third time they face him compared to the first and second. Naturally, his walk rate ticks up as fatigue sets in, though it's hard to imagine Chad Tracy was overly worried about putting the Yankees' best hitter on base. Looking strictly at his late-game effectiveness, the analytics suggest the Red Sox could have trusted their ace.

Without a doubt, Tracy's decision was based on analytics and the favorable matchup that Erik Miller presented. I can't say that the decision was incorrect because the result of another Tolle-Rice matchup is unknown; proving to be a moment for the baseball "What If?" collection. To revert back to the original question, yes, in this case analytics overruled the eye test. There is an opportunity to dive deeper into the analysis and begin to look at swing path and pitch movement but the statistics and advanced stats themselves tell us a lot about Tracy's thinking. We could also start looking at home run probability models to get a statistical prediction on what would've happened if Tolle stayed in. I would never categorize the move as a "stupid analytical" one, but I think Tracy should've trusted his ace in the biggest moment of the season. I'm not one to criticize a manager, but after pulling Sonny Gray early in Game 2, Boston fans have a legitimate reason to be upset.

Statmuse
ESPN
MLB.com
Savant