

# **A Replays Play**



**Mark Pankin**

**SABR 46**

**July 29, 2016**

**Miami, FL**

Notes provide additional information and were reminders during the presentation. They are not supposed to be anything close to a complete text of the presentation or thorough discussion of the subject.

Use Acrobat Reader's ability to enlarge what appears on the screen if you have trouble reading a graph or table.

## Special Thanks



- ⌘ David Vincent for data and analysis
- ⌘ Retrosheet: Replays (HR, expanded) can be downloaded
- ⌘ SABR for all it does

David Vincent has compiled the list of replays used for this talk and which appears on [Retrosheet.org](http://Retrosheet.org).

He also provided some of the data summaries presented and insight into how umpires work.

## **Play in Four Acts**



- ⌘ Act 1: Effects on homer replays
- ⌘ Act 2, scene 1: Expanded replay overview
- ⌘ Act 2, scene 2: Some details, breakouts
- ⌘ Act 3: Positioning – HP collisions, slides
- ⌘ Grand Finale: Possible first inning tactics

## Effects on Homer Replays

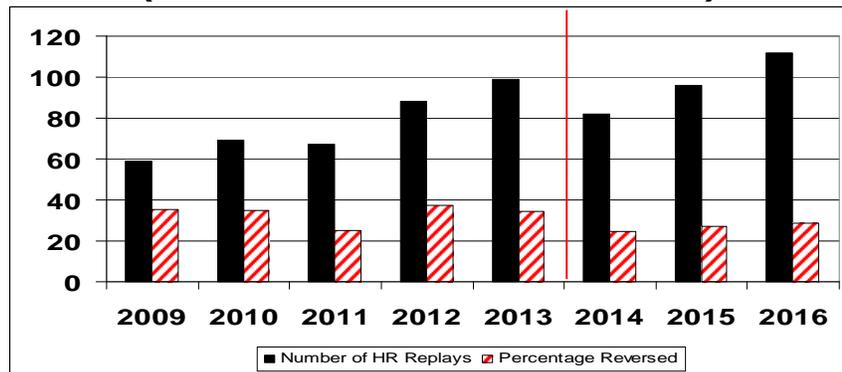
### ⌘ More or fewer?

- ☒ Fewer since teams can take a quick look before asking for a review
- ☒ More because umpires get NY do them and don't go under the stands, so quicker

### ⌘ Change in percent overturned?

## HR Replays, Percent Reversed

(2016 as of June 30 doubled)



Trend in replays may be due to more homers being hit this year.

Lower percentage reversed may be due to having reviews done in central location in NY rather than in each ballpark. Or it may indicate umpires are more willing to ask for reviews on ones that are not all that questionable.

## Expanded Replay Reviews

- ⌘ Began in 2014 (HR replays since 2008)
- ⌘ Team gets one challenge per game
  - ☑ Many, but not all plays are eligible
  - ☑ If successful, gets a another challenge
  - ☑ Umpire can grant more after 6<sup>th</sup> inning
- ⌘ Umpires can call for reviews
  - ☑ Mainly for homers, HP collisions (in 2014)
  - ☑ Others before 7<sup>th</sup> inning are infrequent

Having one challenge available provides another tactic that managers can use.

Team gets some time (supposedly less than 30 seconds, but can be more) to decide whether or not to issue a challenge.

## Annual Totals

	Number	Average per Game	Time taken (minutes)			
			Average	Median	Longest	Shortest
2014	1276	0.53	1.77	1.63	6.28	0.38
2015	1351	0.56	1.85	1.65	6.42	0.43
2016*	1478	0.64	1.69	1.52	6.02	0.27

\* 2016 through June 30, Number is doubled

⌘ Trend is for more per game

☒ Mainly due to managers issuing more:

☒ 2014: 1063 2015: 1186 2016: 1288

June 30 is not quite half the season. Some teams had not quite yet played 81 games. It is closer to half a season than the All-Star Game break when teams have played 90 or close to 90 games.

Another way of seeing the increase is the percentage of games that have at least one review:

2014: 41.5%

2015: 42.4%

2016 (through June 30): 46.3%

(Does not include postseason games)

## Types of Reviews

Type	Percent	% reversed	Avg. Time
Force play	38.7%	59.5%	1.61
Tag play	37.3%	46.0%	1.95
Home run	7.7%	27.5%	1.64
Hit by pitch	4.4%	52.0%	1.84
HP collision	4.3%	13.7%	2.00
Slide rule*	3.8%	14.3%	1.92
Fair/foul (outfield)	1.4%	38.8%	1.84
Catch/no catch	1.1%	50.0%	1.81
Others	1.3%	37.2%	1.97

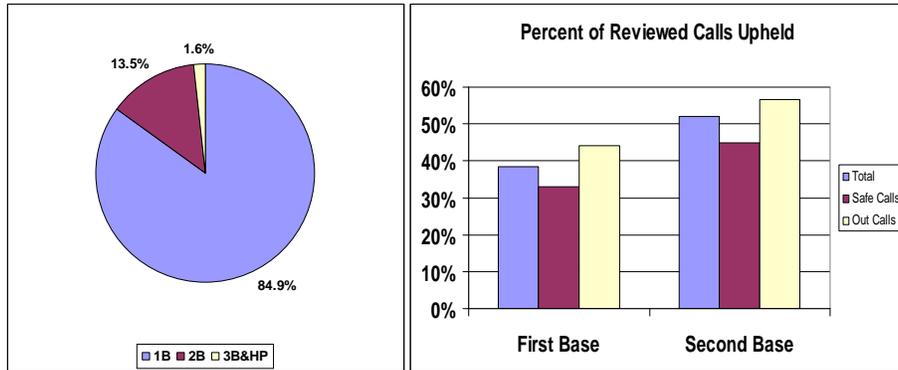
\* Slide rule data based on 2016 only

### **Others**

Boundary call  
 Grounds rule  
 Passing runners  
 Record Keeping  
 Rules Check  
 Tag-up  
 Timing play  
 Touching a base  
 Trap play (outfield)

Retrosheet web site and mlb.com provide more information about the nature of the rules and what they mean.

# Force Play Reviews

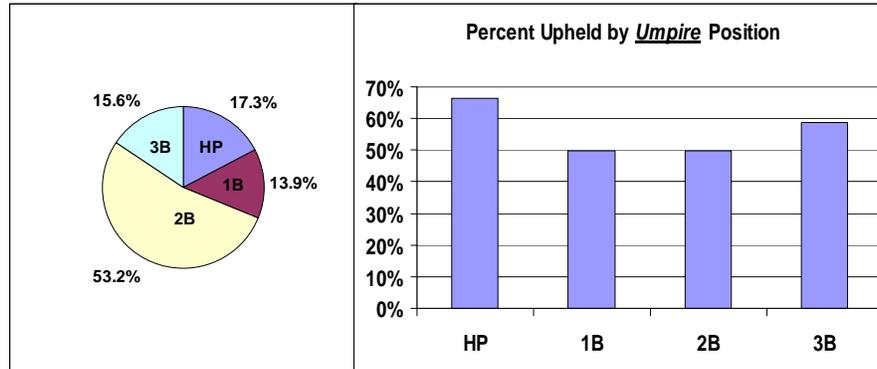


⌘ Forces include possible LDP, FDP (<1%)

Play on runner getting back to a base after a catch is considered a force play because he is required to get back and a tag is not required, just the fielder catching the ball and being on the base before the runner returns.

Force call is “easier” at second than first because on a close play runner will slide making it easier to tell when his foot or hand touches the base, most likely at the “front”, than when he is running to first and foot is coming down on any part of the base.

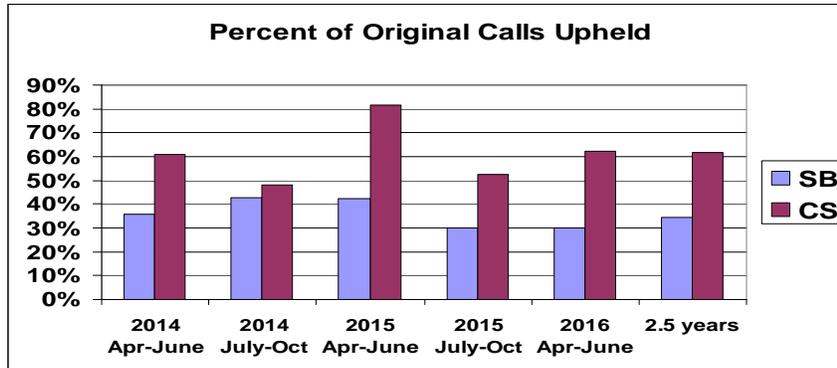
## Tag Play Reviews by Umpire Positions (Not Bases)



⌘ Umps make some calls at different bases

Due to rotation on balls hit into the outfield, umpires often move to cover the different base of the one who goes into the outfield. This is most likely for the 3B umpire and least likely for the home plate one because he needs to stay there if there is a possible play at the plate.

## Tag Play Reviews: SB2/CS2



⌘ 60% of 2<sup>nd</sup> base ump's; 32% of tag replays

It seemed that in early 2014 there were quite a few SB2 that were reversed because the runner lost contact with the base for a very brief time and the fielder maintained the tag. I thought runners would make an effort to maintain contact and fielders would try harder to maintain the tag. I thought the runners learning might show up in fewer SB2 being overturned.

Divided into half seasons to see if there was evidence of the learning effects. Fewer overturns in second half of 2014 than first half. That continued in the first half of 2015, but then percentage of SB2 upheld dropped back below the level of the first half of 2014.

No idea if there is anything other than randomness involved.

## Hit by Pitch Reviews

⌘ 148 in the 2.5 seasons

☒ All challenges by managers, umps  $\geq 7$ th

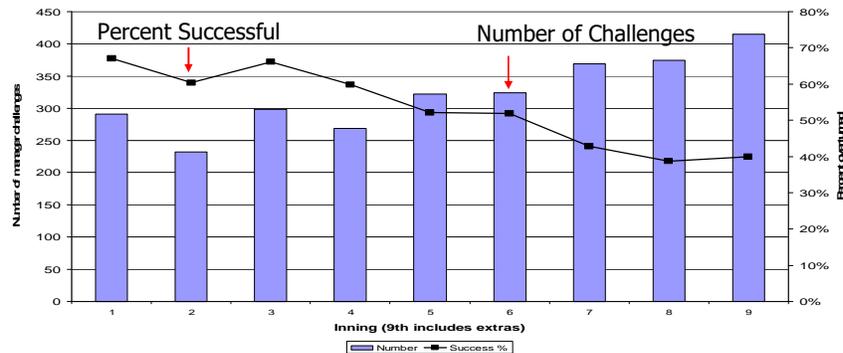
☒ 45% of original calls were hit by pitch

⌘ 48% of all upheld, so 52% overturned

☒ 62% of HBP calls upheld

☒ 37% of not hit by pitch upheld

## Manager Challenges by Inning



### ⌘ Managers use challenges “strategically”

- ⏏ Early innings: good chance to succeed
- ⏏ Later innings: might as well use it

Managers more likely to challenge after the fourth because they know umpires can give them additional challenges in the seventh or later if they don't have one left.

More in the first and third than the second and fourth may be due to those two innings being higher scoring on average so more likely to have runners on and more chances to challenge.

## Highest, Lowest Success %

Total challenges by team			Total against team		
Team	Number	Success %	Team	Number	Success %
NYA	82	64.6%	DET	116	56.9%
MIA	98	61.2%	NYA	78	56.4%
KCA	105	58.1%	WAS	101	55.4%
SFN	119	54.6%	MIL	134	54.5%
MIN	105	53.3%	NYN	93	53.8%
MLB Average (2014 - 6/30/2016): 48.0%					
WAS	117	41.9%	KCA	125	42.4%
SLN	116	41.4%	CLE	83	42.2%
TEX	131	39.7%	TBA	121	41.3%
TBA	151	38.4%	MIA	114	41.2%
TOR	157	33.8%	BAL	117	38.5%

⌘ Looks like fewer challenges → higher success %, but

☒ OAK: 87, 47.1%; MIL: 90, 50.0%; BAL: 91, 46.2%

☒ CHN: 148, 47.3%; BOS: 138, 42.0%; PIT: 133, 50.4%

☒ Modest statistical relationship: R-squared of 41%

Not sure if percentages against teams has any meaning. Are the Tigers disfavored by the reviewers or could it mean that ump's on the field tend to give them the benefit of the doubt on close plays. Likely neither, just randomness.

R-squared of 41% means that percentage in the variation of success percentages is explained by the number of challenges. So well over half is not due to the number.

Apparently there is a difference in philosophy between the Yankees and Blue Jays, or more generally teams that challenge more and those that do it less often. Not clear if one approach is better than the other. It would likely take quite a bit more data and analysis to evaluate this. Would need a way of estimating the value of a successful challenge (on the expected runs scored in the remainder of the inning or possibly the probability of winning) against the loss of value if the challenge failed. Quite dependent on the game situation and things like pitchers in the game and available, etc. so unlikely to come up with anything close to "firm" answers.

## Home Plate Collision Reviews

- ⌘ Umpires initiated in 2014, mainly managers since
- ⌘ Many more in 2014
- ⌘ "Learning curve" by umps, catchers, managers

	Reviews initiated by		Original out calls	
	Umpires	Managers	Number	% upheld
2014	95	0	2014	93 86.0%
2015	2	30	2015	31 87.1%
2016*	1	19	2016*	18 88.9%

\*2016 thru June 30

- ⌘ Only 5 original safe calls reviewed, 4 upheld

There was quite a bit of discussion of how catchers would adapt to the "Posey rule." In 2014 when it was up to the HP umpire to ask for a review (usually at the urging of a manager), they were more prevalent since umpire did not have to worry about losing his challenge. Seeing the high percentage upheld, managers learned to be careful about challenging.

Safe original calls on this type of review indicates that the catcher blocked the plate without the ball.

High percentage of out calls upheld indicates that catchers had worked on their positioning. Not clear if slightly higher percentage in in 2015-16 is meaningful (randomness, relatively fewer).

## Slide Rule Reviews

- ⌘ New in 2016
- ⌘ Is there a learning curve on these?
- ⌘ 28 thru 6/30, all by managers, umps > 6th
- ⌘ All but one original call: no interference
  - ☒ April: 11 no interference, 10 upheld
  - ☒ May: 11 no interference, all upheld
  - ☒ June: 5 no interference, 3 upheld
  - ☒ One interference in April was upheld

I wondered if it would take some time and rulings for runners to learn how they needed to slide into second in possible DPs. Most managers said they told their players to slide directly at the base. Apparently that is what is happening since there has been only one review of an interference call. Don't know how many interference calls have been so obvious that there was no point in a challenge.

Rule is sometimes called the "Utley rule," but to be consistent with Posey rule, it should be the (Rueben) Tejada rule.

It looks like managers have figured out that there is no point in challenging the no interference unless there is an excellent chance of winning.

I think the interpretation of the rule is still evolving. Saw a play in May when the runner clearly did not slide at the base, but he did not contact the fielder, so the ruling was no interference because his slide did not prevent not completing the DP.

## Odds and Ends

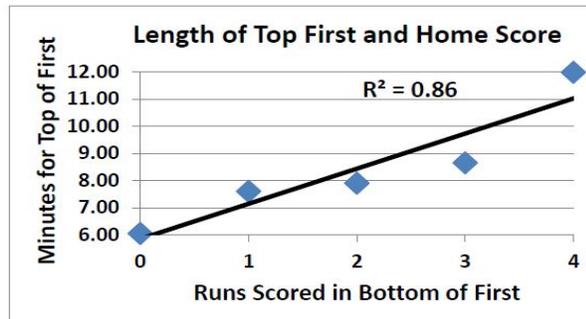
- ⌘ 19 games with 3 overturned
  - ☒ 5/22/14 all were on one ump's calls
- ⌘ Managers with 3 successes in a game
  - ☒ 8/8/15, Clint Hurdle (4 challenges)
  - ☒ 8/19/15, A.J. Hinch (3 challenges)
  - ☒ 4/9/16, Walt Weiss (3 challenges)
- ⌘ 4/20/16, COL at CIN, pitching change in the middle of a review

4/20/16: single scoring runner from second in top of 7<sup>th</sup> that would have tied the game. Runner missed third, which would have been the third out. Cinci manager did not see that or that one of his fielders was indicating the runner missed, so he went ahead with a pitching change. After the change, he know what had happened, so the new pitcher stepped off the mound and threw to third for the appeal, which was upheld. Rockies manager then challenged the call that the runner had missed the base, but the original call was upheld.

New pitcher got credit for 0.1 inning pitched due to the sequence. He batted in the bottom of the 7<sup>th</sup> (was not due up, but Reds had a longer, scoring inning), so chance to use pinch hitter was lost. In top of 8<sup>th</sup>, the pitcher faces two batters, both of whom reached. So he did not retire anyone and credit for two batters faced and 0.1 inning pitched.

## Top of First Inning Tactic?

- ⌘ Dave Smith's 2006 SABR presentation
- ⌘ Longer top of first → more runs in bottom



⌘ Additional minute in top, about 2/3 more on average in the bottom (based on the black line)

R-squared of 86% is high for baseball relationships. It means a large portion of the variation in run scoring in the bottom of the first can be explained by (note: does not necessarily mean is caused by) the length of the top of the first.

Should this affect whether or not a manager challenges a play in the top of the first?

## **Top of First Tactic? (2)**

- ⌘ Typical replay review: 1.5 to 2 minutes
- ⌘ Increases average home team scoring in bottom of first by about 1 run
- ⌘ Time to decide on challenge by manager
- ⌘ Home team: make close call challenge
- ⌘ Road team: want to be fairly certain
- ⌘ Complex analysis depends on many factors other than bottom of first scoring

Home team might want to challenge a play they are not all that confident will be reversed. The time it takes may lead to (more) runs scored in their half of the inning. If nothing else, they may want to delay making the decision to challenge or not as long as they can.

Road team should challenge only when they have a high degree of confidence. Earlier slide shows a success rate of 67% for first inning manager challenges, which likely should be continued for decisions in top of first by road team

This is simplistic since there are quite a few factors to consider. Perhaps the most important is the value, positive or negative, of having or not having a challenge to use later in the game before the seventh when umpires can grant challenges. Also the game situation and the effects of keeping or reversing the play that has just taken place. May be others such as supporting players.

## Summary



- ⌘ More than 3/4 are force, tag plays
- ⌘ About half of all result in reversals
- ⌘ Safe calls on force, tag plays more likely to be reversed than out calls
- ⌘ Managers are using challenge as a tactic
- ⌘ Possible top of first tactic
- ⌘ Learning about them is ongoing process

One of the major learning areas is how managers should use their challenges. Seems to be more of a “use it or lose it” attitude. Not clear if Yankees with fewer reviews and higher success rate and doing it better than the Blue Jays who issue more and get fewer reversed.

## **Web sites, e-mail**

[www.retrosheet.org](http://www.retrosheet.org)

look for link on home page to replays

E-mail: `sabr --ATsign-- pankin.com`

Plan to post slides, notes on Retrosheet site

These slides and some notes will eventually be posted on the Retrosheet Research page.