Are There "Pitchers" Umpires and "Hitters" Umpires?

Mark Pankin
SABR 47
June 30, 2017
New York, NY

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 insights
- Retrosheet: Game logs and other data
- SABR for all it does
- Bill James: Early abstracts (1976-78 seasons) had strikeout, walk rates by umpire crew

David Vincent has studied umpires and provided helpful insights and information. He passed away on 7/2/2017 and will be greatly missed.

Bill James did it by crew since he was not sure he could tell who was the HP ump for each game. Retrosheet game logs have that information

Overview

- Most favorable home plate umps for pitchers, hitters in 1960-2016
- Former players as umpires
- Has pitch calling improved by being more consistent after 2000?

Wanted at least 50 years of data and Retrosheet game logs are by decade. Using a different time period likely would give different results.

Example of Computations

- Eric Gregg
- Maybe the most "pitcher friendly" game:
  - 1997 NLCS, Game 5: 2-1 Marlins over Braves
  - 25 Ks, 3 BBs, 8 hits in 8.5 innings
  - 15 Ks by Livan Hernandez
  - Extremely wide strike zone

In the Koufax perfect game, there was only one hit. Koufax fanned 14, but Hendley had only 3 (and 1 walk), so not an indication of an extreme strike zone, just outstanding pitching.

Even having the same last name did not help as Tommy Gregg (no relation) fanned as a pinch hitter.

Strike zone was so wide that it distorted the game and motivated baseball to try to make pitch calling more uniform.
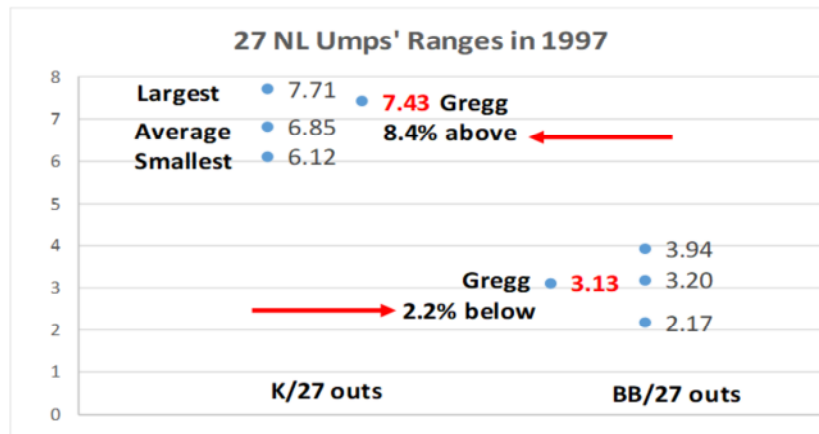
Eric Gregg: 1997 data

- 28 games at HP in NL parks, (3 in AL parks excluded)
- 1494 outs (= 498 half innings)
- 411 Ks, 173 non-intentional BBs
- 7.43 K/27 outs, 3.13 BB/27 outs

Due to DH rule, not appropriate to combine games in the two leagues for individuals.

After 2000, discussed later, computed K/27 outs and non-IW BB/27 outs in each league for an umpire and then averaged them since overall among all same number of games in each league. Individuals may vary quite a bit such as 12 in one league and 20 in the other.

Gregg compared to NL HP umps

(27 had at least 25 games behind plate)

27 NL Umps' Ranges in 1997

Largest • 7.71 • **7.43** Gregg
Average • 6.85 **8.4% above** ←
Smallest • 6.12

• 3.94
Gregg • **3.13** • 3.20
**2.2% below**
• 2.17

K/27 outs   BB/27 outs

Nothing magic about requiring a minimum of 25. To some extent needed to get enough umps in the data. Higher would have been better for statistical confidence. However, these days with vacation weeks during the season and rotation through the review crews in NY, don't get much more than 30 behind the plate for individuals.

The 8.4% above the K rate and 2.2% below the BB rate averages for the NL in 1997 are the data used in the evaluations.
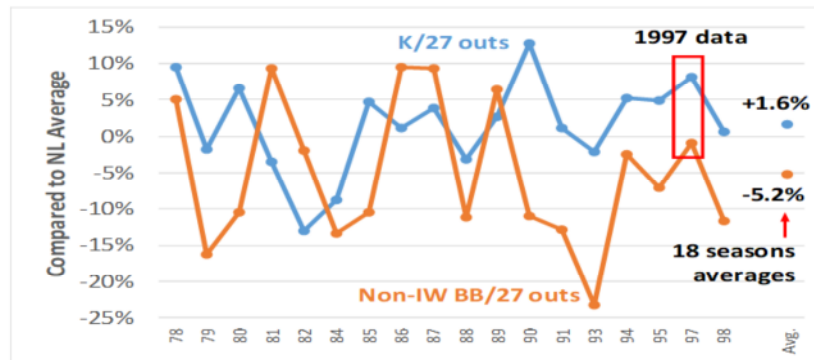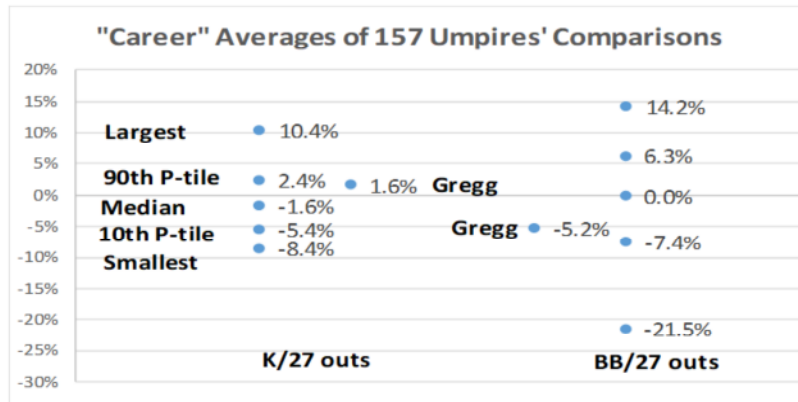
Gregg was somewhat inconsistent (suspect that is true for many) but generally more favorable to pitchers.

The averages for his 18 seasons with at least 25 games behind the plate (+1.6% for K/27 outs, -5.2% for BB/27 outs) will determine if he was one of the most pitcher friendly umpires in the study.

Umpires' ranges 1960-2016

At least 9 seasons of 25 or more games at HP

"Career" Averages of 157 Umpires' Comparisons

Largest — 10.4% — 14.2%
90th P-tile — 2.4% — 1.6% Gregg — 6.3%
Median — -1.6% — 0.0%
10th P-tile — -5.4% — Gregg -5.2%
Smallest — -8.4% — -7.4%
— -21.5%

K/27 outs   BB/27 outs

Nothing special about requiring 9 seasons of 25 or more games. Chosen to some extent to get more of the current umps. Ratings will be based on at least 225 games behind the plate, and for most of them, considerably more.

K/27 outs and BB/27 outs show the largest, smallest, median, and the 10% away from the extremes levels.
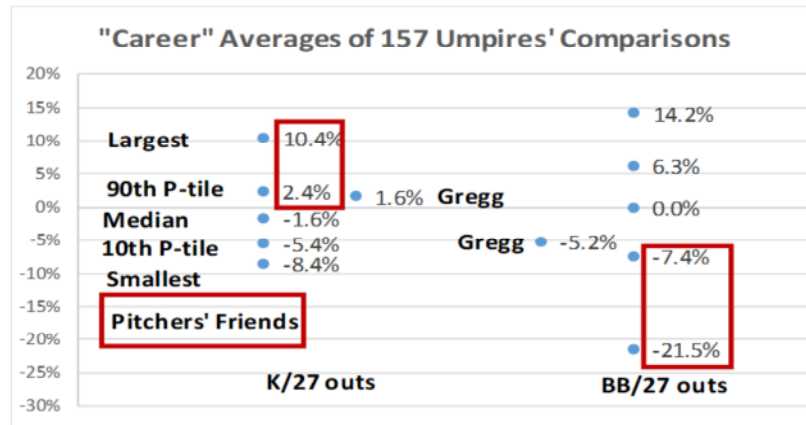
Not surprising that walks have more variation because there are "unintentional, intentional" ones in the data. Also "Umpire Analytics" article by Brian Mills in the downloadable SABR Book on Umpires and Umpiring shows "true" balls tend to be called correctly more than true strikes.

We see Gregg is not in the top 10% of K rates and not in the bottom 10% of BB rates.

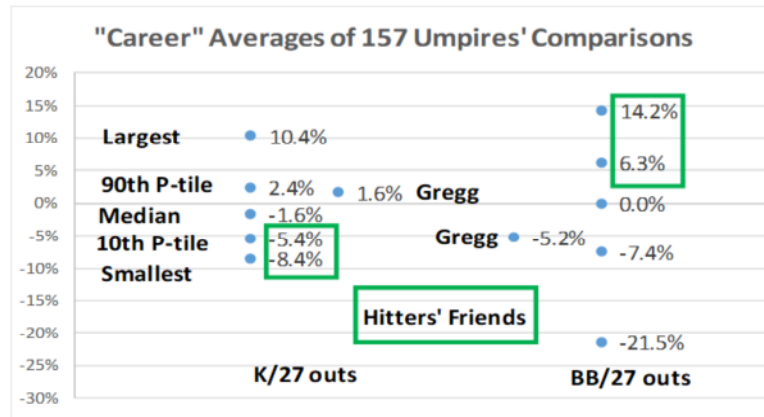My definition of "friendliness" is being in the extreme 10% groups of both categories. Pitcher-friendly: with Ks well above, BBs well below.

Hitter's friends are the opposite: lowest K rates, highest BB rates.

Friendliest Umpires

	Compared to group		Years	
	Ks/27 outs	BB/27 outs	First	Last
PITCHERS' FRIENDS				
Ed Runge	10.4%	-14.3%	1960	1970
Dave Pallone	5.0%	-9.1%	1979	1988
Lee Weyer	5.0%	-8.2%	1963	1987
Greg Bonin	4.6%	-8.9%	1988	1998
Ken Burkhart	3.6%	-9.9%	1960	1973
Mark Hirschbeck	3.2%	-9.2%	1989	2002
Larry McCoy	3.2%	-10.8%	1971	1998
HITTERS' FRIENDS				
Lance Barksdale	-8.4%	4.1%	2001	2016
Paul Schrieber	-7.9%	12.3%	1998	2015
Alfonso Marquez	-7.7%	7.7%	2000	2016
Randy Marsh	-6.8%	11.5%	1982	2009
Greg Gibson	-6.7%	7.4%	2000	2016
John McSherry	-6.3%	8.6%	1971	1995

Sorted by K/27 outs, largest for pitcher, smallest for hitters.

"As They See 'Em" by Bruce Weber cites Jim Kaat as saying the pitchers loved it when they saw Ed Runge would be behind the plate. Runge supposedly told hitters he wanted to see them swing the bat, not walk. (p. 186 in paperback, copyright 2009). He started in 1954, so only part of his career is shown. (Kaat mentions Ed Hurley as being friendly to hitters, but his career ended in 1965, so not enough years to be included in the analysis.)

Barksdale is 3rd most accurate in Mills article, but his correct ball % (92.56%) is much higher than correct strike (88.25%). Two others are more extreme, but they have not had nine years as MLB umps. He is the only overlap between the list here and the ten most and least accurate (combined balls and strikes).

List would be different for other periods, say 2000-16.

Former Players as Umpires

- Do former pitchers favor pitchers?
- Do former position players favor hitters?
- Looked back to 1920
  - At least 7 seasons with at least 25 games behind the plate
  - Usually many more due to smaller crews
  - 8 former pitchers, 8 former position players

Started with 1920 to avoid the dead ball era. Seven seasons chosen to get data on more of them. Since most worked with 2 or 3 man crews, they were at HP much more often during a season than current ones.

Former Pitchers

	Career averages (seasons >=25 at HP)				
	Compared to league		Lg	Years	
	Ks/27 outs	BB/27 outs		First	Last
Ken Burkhart	3.5%	-9.6%	NL	1957	1973
Bill Dinneen	2.7%	11.1%	AL	1920*	1937
Tom Gorman	1.0%	-4.0%	NL	1952	1976
Bill Kunkel	1.5%	-1.9%	AL	1968	1984
Charlie Moran	-2.5%	-8.7%	NL	1920*	1939
Hank O'Day	-4.0%	-10.9%	NL	1920*	1926
George Pipgras	3.2%	11.3%	AL	1939	1945
Eddie Rommel	-1.4%	11.3%	AL	1938	1959
Average	0.5%	-0.2%			

1920* indicates umpired before 1920

Listed in alphabetical order by last name.

Names in red were favorable to pitchers: had career above league averages for K/27 outs and below averages for non-IW BB/27 outs. (In some seasons IW were not an official record). Green name was favorable to hitters with opposite comparisons. Others were in the same direction for both comparisons.

Burkhart is highlighted because he was on the pitcher friendly list earlier. Enough data for Gorman and Kunkel, but not large enough to make most pitcher-friendly list.

Average of the eight is pretty close to league averages.

Former Position Players

	Career averages (seasons >=25 at HP)				
	Compared to league		Lg	Years	
	Ks/27 outs	BB/27 outs		First	Last
Charlie Berry [C]	3.4%	-0.4%	AL	1957	1973
Jocko Conlan [OF]	-1.2%	-8.8%	NL	1941	1964
George Hildebrand [OF]	2.0%	6.3%	AL	1920*	1932
Barry McCormick [IF]	0.6%	-10.5%	NL	1920*	1929
George Moriarty [IF]	-1.6%	18.1%	AL	1920*	1940
Babe Pinelli [IF]	2.3%	-12.8%	NL	1935	1956
Frank Secory [OF]	0.0%	-2.0%	NL	1952	1970
Vinnie Smith [C]	1.9%	3.7%	NL	1957	1965
Average	0.9%	-0.8%			

1920* indicates umpired before 1920

▌ Slightly more favorable to pitchers

Similar layout to former pitchers page. Berry had enough years after 1960 to be considered for the most pitcher friendly, but his comparisons to the AL averages are not nearly large enough.

As with former pitchers, wide variation among this group. Its averages are also close to the leagues', but slightly more favorable to pitchers than the former pitchers.

Better Pitch Calling Now?

- Before 2000, each league had it own umpires and supervisors
- Since 2000, unified under MLB
- QuesTec, PITCHf/x started in 2001, in all parks since 2008
- Is the variability among HP umps less now than before 2000?

Eric Gregg NLCS game shown earlier may have been one of the motivations to combine the leagues and have only office supervising them.

Umpire's union strike led by Richie Phillips in 1999 provided an opportunity to combine the leagues' supervision.

Better Pitch Calling Now?

- Will look at annual standard deviations of K/27 outs, BB/27 outs to league averages
- Umps with >=25 games at HP
- Graphs for 1983-2016: 17 seasons before, after change in 2000
- Separate for each league
  - DH may have an effect
 - Kunkel in 1984 was last to use balloon

Might have used a different measure of consistency. Standard deviation has the advantage of taking all umpires into account and giving more weight to those at the extremes than near the averages.

Use of the balloon ump in the AL and standing directly over the catcher rather than over the shoulder on the inside part of the plate will have at most a trivial effect on the first two years shown for the AL.

Standard Deviations of K/27 outs
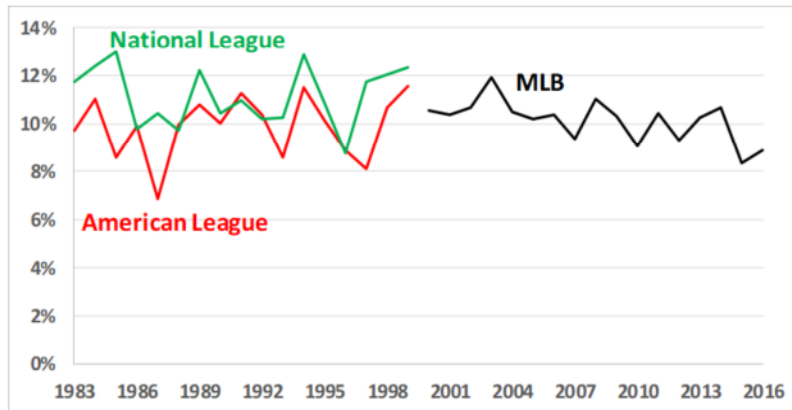
Individuals' comparison to league averages
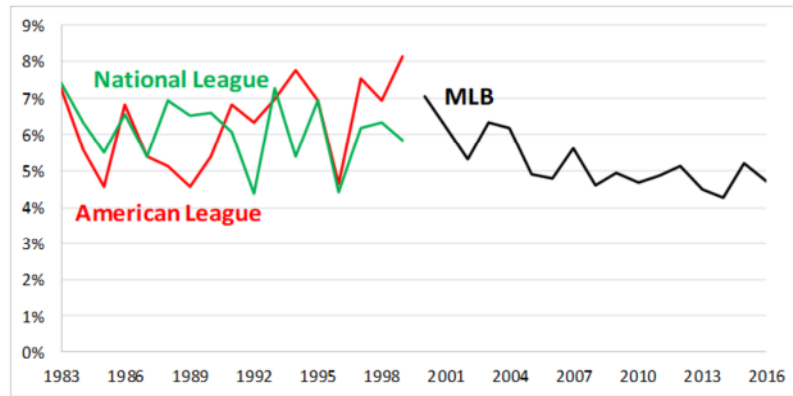
Lower is "better"; note range of about 4-8%.

Considerable variation between the leagues, both relatively and absolutely with no obvious trend.

Once under MLB, trend has been toward lower values, so more consistent strike calling among the umpires.

Standard Deviations: BB/27 outs

Individuals' comparison to league averages

Range is 7-13%, both higher and wider than Ks, which is not surprising since some walks are "unintentional intentional" or result from "pitching around" or very carefully to some hitters. Also, possibly due to more accurate calling of true balls than strikes (at least recently according to Mills article)

Although not as well defined with Ks, trend seems to be to lower values.

Mills article says pitch calling has become more uniform in recent years. Moreover, newer umpires tend to be more accurate (better training, evaluation, feedback, need to keep job?) so as older ones retire overall accuracy improves.

Summary

- Can identify HP umps who seem to favor pitchers, and others who favor hitters
- Hitter-friendly ones more recent
- Former players: no favoritism as a group
- Variation in K&BB/27 outs has gone down since 2000 MLB takeover of supervision
- PITCHf/x: can evaluate by pitch calls

Group averages for former players are close to the league averages. In each group (former pitchers, position players), 3 more favorable to pitchers (although not a whole lot for some), 1 more favorable to hitters.

"Umpire Analytics" article by Brian Mills in the downloadable SABR Book on Umpires and Umpiring digs into data on pitch calling and strike zones. That should be a better evaluation of pitch calling "quality" than the method here. However, what I have done is applicable historically.

Web sites, e-mail

www.retrosheet.org

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.