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

<table>
<thead>
<tr>
<th></th>
<th>Compared to group</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ks/27 outs</td>
<td>BB/27 outs</td>
</tr>
<tr>
<td><strong>PITCHERS’ FRIENDS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ed Runge</td>
<td>10.4%</td>
<td>-14.3%</td>
</tr>
<tr>
<td>Dave Pallone</td>
<td>5.0%</td>
<td>-9.1%</td>
</tr>
<tr>
<td>Lee Weyer</td>
<td>5.0%</td>
<td>-8.2%</td>
</tr>
<tr>
<td>Greg Bonin</td>
<td>4.6%</td>
<td>-8.9%</td>
</tr>
<tr>
<td>Ken Burkhart</td>
<td>3.6%</td>
<td>-9.9%</td>
</tr>
<tr>
<td>Mark Hirschbeck</td>
<td>3.2%</td>
<td>-9.2%</td>
</tr>
<tr>
<td>Larry McCoy</td>
<td>3.2%</td>
<td>-10.8%</td>
</tr>
</tbody>
</table>

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

Web sites, e-mail

www.retrosheet.org
E-mail: sabr --ATsign-- pankin.com
Plan to post slides, notes on Retrosheet site