Statistics in Sports: Baseball Overview

ZACHARY BINNEY, PHD MPH

OXFORD COLLEGE OF EMORY UNIVERSITY

FALL 2021

Roadmap

- 1. Sports Analytics History
  - “Culture Wars”
- 2. (Player Valuation) Analytics in Baseball:
  - Batting
  - Fielding
  - Pitching
- 3. What Makes a Good (Sports) Statistic?
  - Define. Your. Question!

Sports Analytics History

First Use of “Statistics” in Sports?

- Who can push us back the farthest?

STOP!

DON’T LOOK AT FUTURE SLIDES YET.

First Use of “Statistics” in Sports?

- Haven’t they always been there, in a sense?
- Doesn’t everyone – including **coaches** and **scouts** – use them?
  - Jocks vs. nerds?
  - Culture war vs. actual disagreement on using numbers in sports
    - Matter of degree (literally and figuratively)
    - Outsiders vs. insiders

Analytics Culture War

- It's not (just) dumb nerds resisting jocks who "know how to do their job better"
  - "Nerd" = "Outsider"
  - Have to recognize we're coming in to *their* house. Be humble, "there to help"
  - Try and integrate ourselves in existing culture/space while pressing for change. Tough balance!

Football Research's Closest Comparisons:		
Age	28-29	
Position	WRTE	
Injuries in Past 3 Years		
Injury #1	Knee_ACL	
Injury #2	Hamstring	
Injury #3		
Number of Matching Players	11	**Age and Position Averages**
% Playing in 3 Seasons (2000-2009 seasons only)	29%	46%
Average Games Missed Due to Injury Next Season	1.9	1.8
Average Games Played Next Season	10.4	11.4
Average Snaps Next Season	484.6	469.0

List of Players with Similar Histories Going in to X Season	Games missed due to Injury	Games played	Scout Grade - Listed Season	Scout Grade - Prior Season	Change
DANIELS, OWEN - 2011	0	15	6.2	5.9	Rise category
DANIELS, OWEN - 2012	1	15	6.2	6.2	Maintain category
WILLIAMS, ROLAND - 2003	0	1		5.8	Not graded
CALDWELL, RECHE - 2007	0	10	5.3	5.9	Fall category
GIVENS, DAVID - 2008	0	0		3.5	Did not Play
AVERY, DONNIE - 2012	0	16	6.2	5.8	Rise category
ENGRAM, BOBBY - 2002	1	15	5.8	5.6	Maintain category
BURLESON, NATE - 2010	2	14	5.2	5.9	Fall category
LEWIS, JERMAINE - 2004	7	9		3.5	Not graded
DYSON, KEVIN - 2003	10	3	5.6	5.8	Maintain category
ENGRAM, BOBBY - 2001	0	16	5.6		Unknown

Modern History of Statistics in Sports

- DISCLAIMER: In no way comprehensive. Weighted towards sports I know more about (football, baseball)

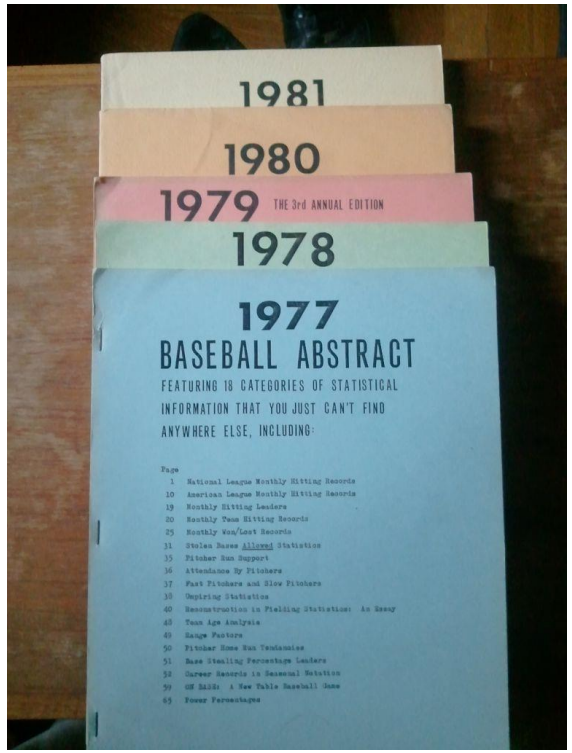
- Here's a baseball history

Modern History of Statistics in Sports

- Early Modern Era: 1970-2000ish
  - Statisticians, academics, and other “outsiders” argue they can improve sports with numbers

Modern History of Statistics in Sports

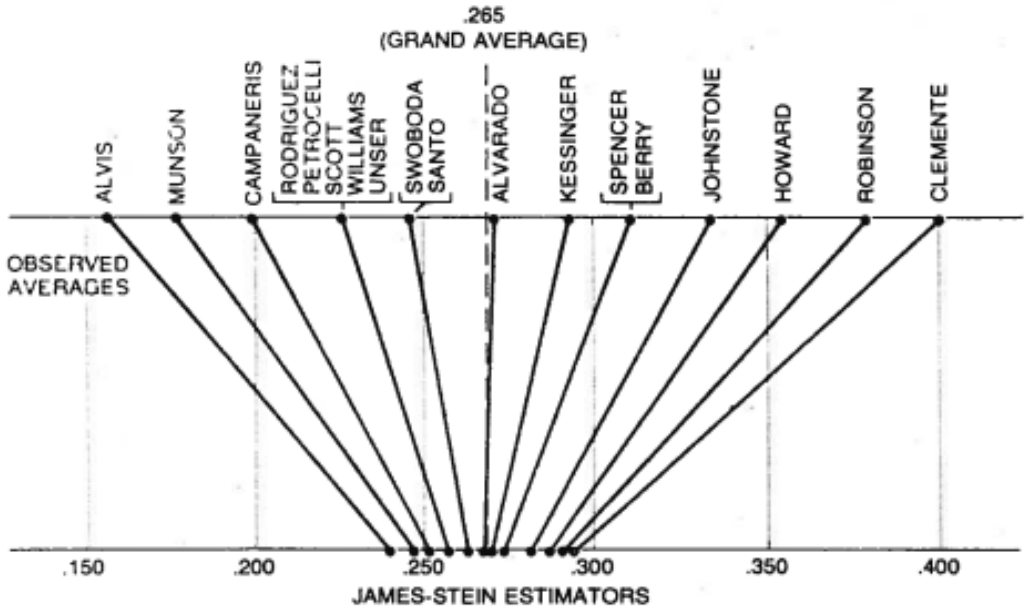
- Early Modern Era: 1970-2000ish
  - "Operations research in football" – Carter & Machol, 1971
  - "On the optimal time to pull the goalie" – Morrison, 1976
  - Bill James Baseball Abstract – 1977
  - Rotisserie Leagues/fantasy baseball (by mail)



Source: https://homemlb.wordpress.com/2014/02/05/1977-new-candidates/

Modern History of Statistics in Sports

- Early Modern Era: 1970-2000ish
  - "Stein's Paradox in Statistics" – Efron and Morris, 1977



  - *Hidden Game of Football* – Carroll, Thorn, and Palmer, 1988

Source: https://statweb.stanford.edu/~ckirby/brad/other/Article1977.pdf

Modern History of Statistics in Sports

- Turn of the Millennium: 2000-2015ish
  - Greater data availability
  - Internet, bloggers, public analyses → team hires
  - Public profile grows
    - *Moneyball*, 2003

Source:

Modern History of Statistics in Sports

- Turn of the Millennium: 2000-2015ish
  - BaseballProspectus.com (1997)
  - Baseball-Reference.com (2000)
  - Pro-Football-Reference.com (2003)
  - FootballOutsiders.com (Aaron Schatz; 2003)
  - *Basketball on Paper* (Dean Oliver; 2004)
  - Sports Reference (Sean Forman; 2004; now includes basketball, hockey, soccer, college basketball/football)
  - Fangraphs (2005)
  - Advanced Football Analytics (Brian Burke; 2006-16)

FOOTBALL OUTSIDERS™

BP

FANGRAPHS

SPORTS REFERENCE

Baseball | Basketball | NFL | Hockey | CFB | CBB | Soccer | Stathead | Blog | Full Site Menu Below ▾

Source: https://homemlb.wordpress.com/2014/02/05/1977-new-candidates/

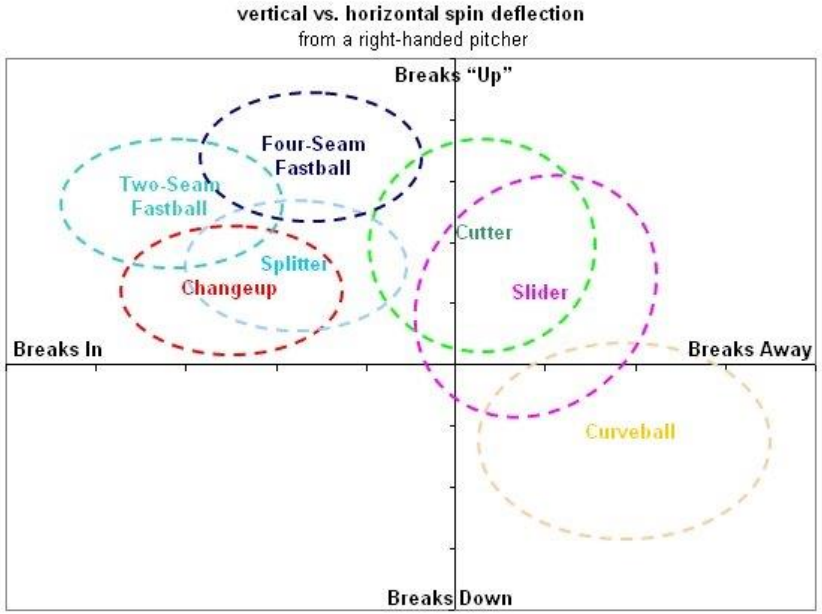
Modern History of Statistics in Sports

- Firehose/Player Tracking Era: 2015ish-Now
  - Proliferation of individual blogs, websites, independent sports analyses, sports analytics clubs
  - Massive increases in data for public but especially for private team work
  - **Tracking data** firehose; data engineers, architects
  - Team hiring accelerates (non-baseball 0 → some, baseball some → whole departments)

Source:

Modern History of Statistics in Sports

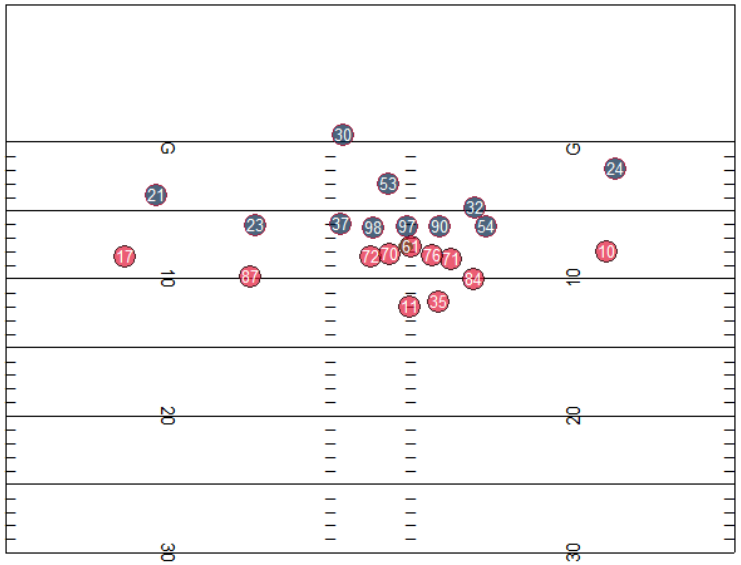
- Firehose/Player Tracking Era: 2015ish-Now
  - Baseball: Statcast/Hawkeye
    - Previously Trackman, PitchF/X

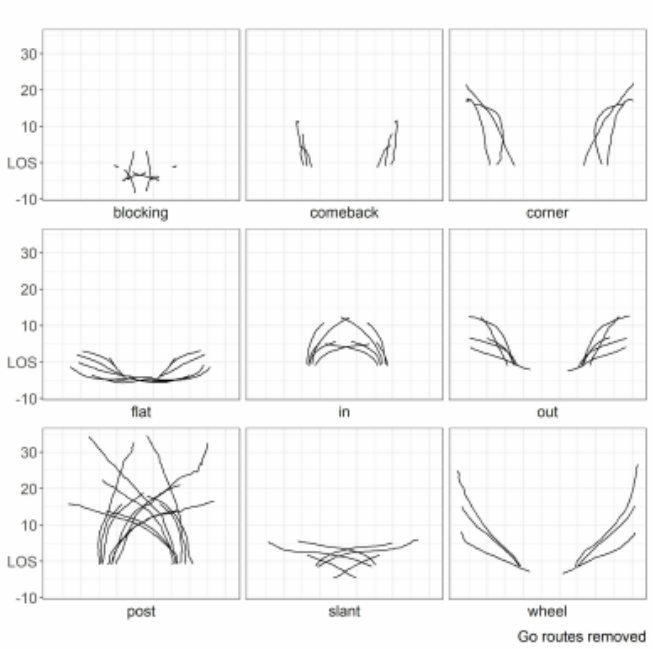


Source: https://www.etcentric.org/mlbs-statcast-tech-could-revolutionize-the-game-of-baseball/; https://library.fangraphs.com/pitch-type-abbreviations-classifications/

Modern History of Statistics in Sports

- Firehose/Player Tracking Era: 2015ish-Now
  - Football: Player Tracking data





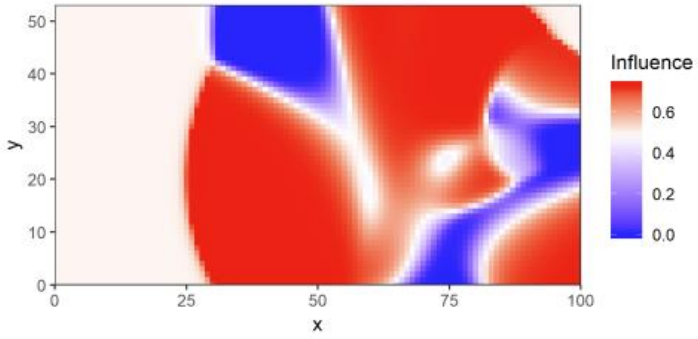
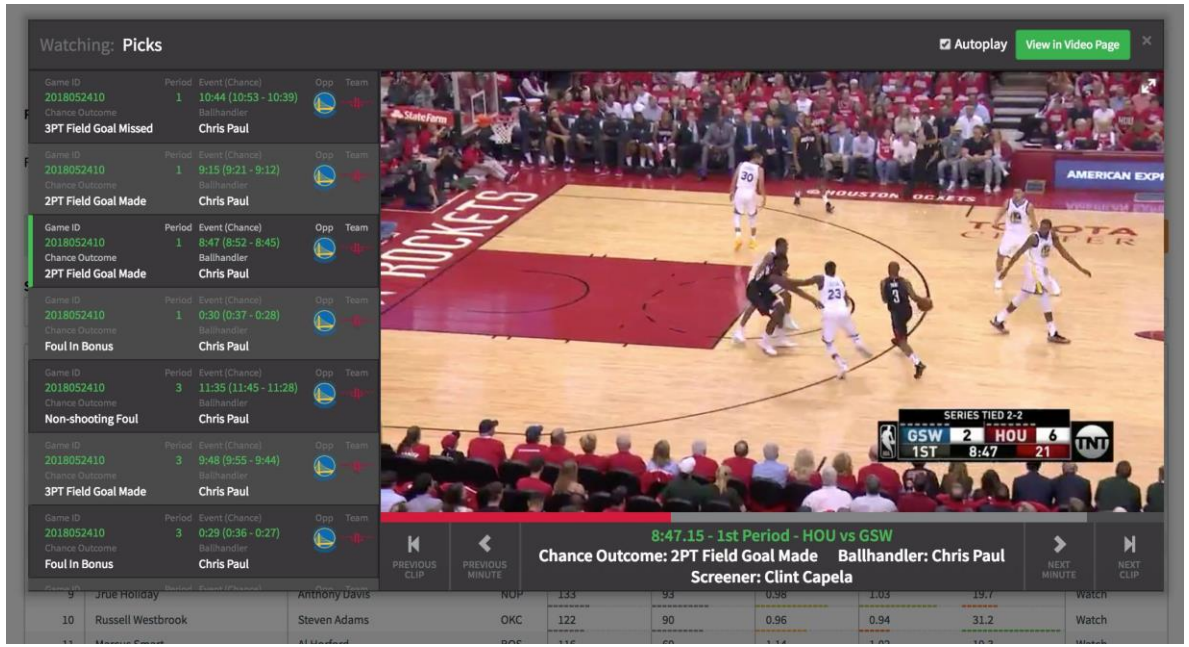


Figure 7: Offensive influence at the moment the ball is released by the QB to the receiver

Source: https://statsbylopez.netlify.app/post/r-for-nfl-analysis; https://operations.nfl.com/media/3670/big-data-bowl-sfu.pdf

Modern History of Statistics in Sports

- Firehose/Player Tracking Era: 2015ish-Now
  - Basketball: Second Spectrum (prev. SportVu)



Source: https://fansided.com/2018/06/28/second-spectrum-redesigning-nba/; https://www.secondspectrum.com/press/2016-05-13.html

(Player Valuation) Analytics in Baseball

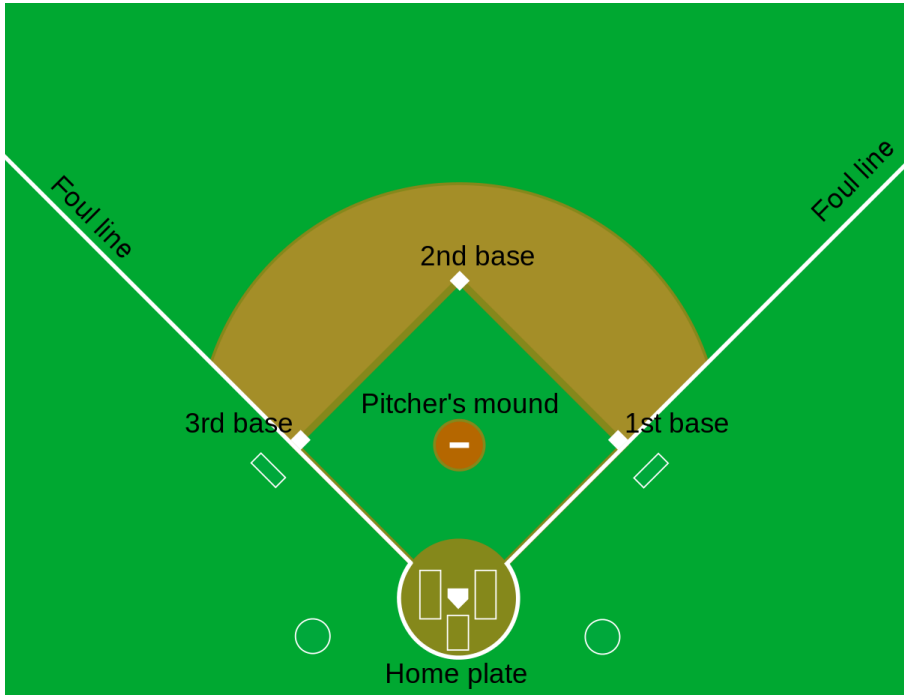
Player Valuation

- Focus on **player valuation**
  - Even ignoring some parts of that, like catcher defense/pitch framing
- Ignores other cool topics including:
  - Opponent scouting (shift)
  - Strategy questions (should you bunt?)
  - Team comparisons (rankings, GOAT team)
  - Preventing pitching injuries
  - Umpire analytics
  - Statcast data, spin rates

Source: https://memegenerator.net/instance/47078543/liam-neeson-theres-no-time

Baseball

- How do you play baseball? Let's take a look at a recent random Yankees-Royals game



Source: https://www.youtube.com/watch?v=CIYAC2PX93Y&t=19s; https://commons.wikimedia.org/wiki/File:Baseball_diamond_simplified.svg

Player Valuation: Defining the Question

- What is player **value**?

- Step back: what is the goal/purpose of a baseball team?

Batting

- DISCLAIMERS: In no way even a comprehensive list of baseball batting player valuation stats
  - Not using Statcast data (e.g. out-of-zone swing %s)

- Many different “flavors” of some stats with different formulas – we use Fangraphs

Batting

- Runs (R) and Runs Batted In (RBIs)
  - Problems/Limitations? **(For this and all subsequent slides, don't look forward until I say so.)**

R and RBI

Batting

- Batting Average (BA) = $\frac{Hits\ (H)}{At\text{–}Bats\ (AB)}$

  - **Rate stat** vs. **counting stat** – solves denominator problem
  - Also improves dependence on other players
  - Problems/limitations?

R and RBI → BA

Batting

- On-Base Percentage (OBP) = $\frac{H + Walks\ (BB) + Hit\ by\ Pitch\ (HBP)}{AB + BB + HBP + Sacrifice\ Flies\ (SF)}$

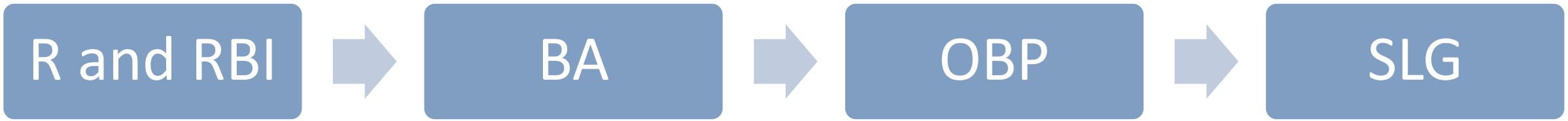
  - Solves undervaluing of walks, HBP
  - Problems/limitations?

≈ Plate Appearances (PA)

R and RBI → BA → OBP

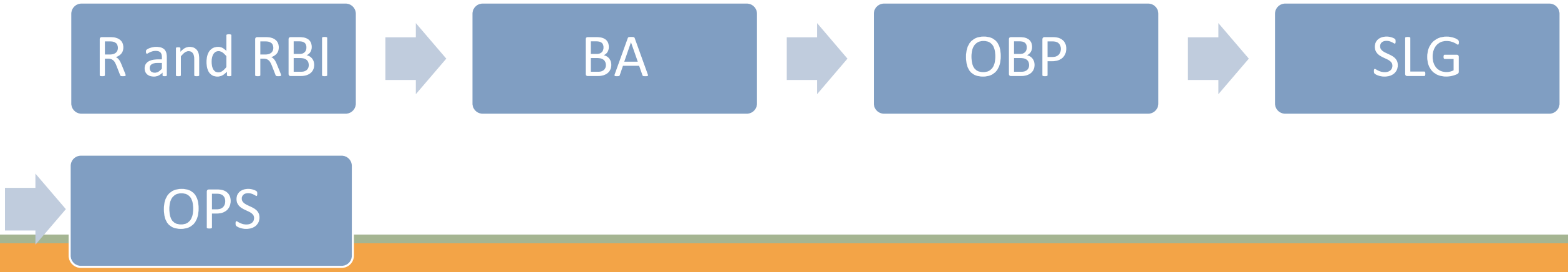
Batting

- Slugging Percentage (SLG) = $\frac{Total\ Bases\ (TB)}{AB}$
  - Accounts for power, XBH and homers
  - Problems/limitations?



Batting

- On Base Plus Slugging (OPS) = $OBP + SLG$
  - Closer to all-encompassing stat for batting skill
  - Problems/limitations?



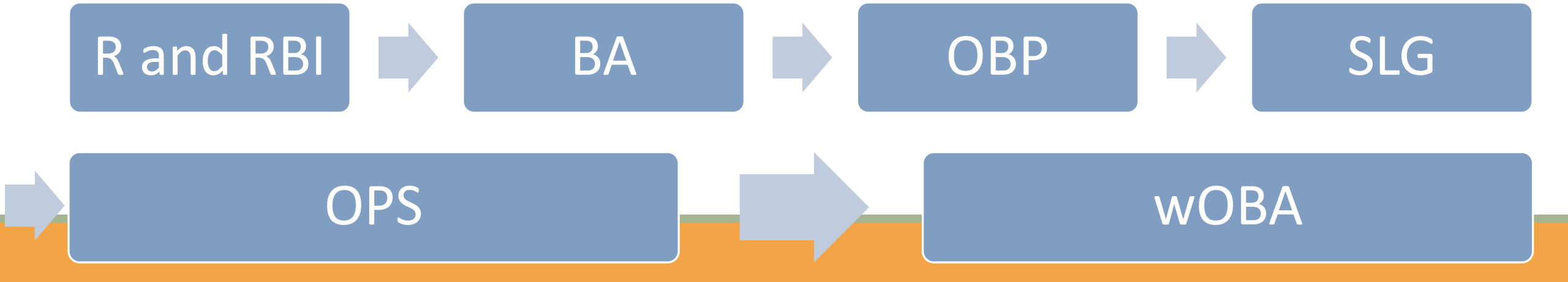
Batting

Weights vary slightly by season based on season OBP

- Weighted On Base Average (wOBA) =

$$\frac{0.69*uBB + 0.72*HBP + 0.89*Singles\ (1B) + 1.28*2B + 1.64*3B + 2.14*HR}{AB + BB - Intentional\ Walks\ (iBB) + SF + HBP}$$

  - More accurate weights than SLG based on **run expectancy**
  - Scaled to OBP, interpreted similarly
  - Problems/limitations?



Run Expectancy

- **Run Expectancy (RE):** every PA moves you from one state to another
- Ex.: leadoff single goes from 0.481 to 0.859 → worth 0.378 expected runs
- Average over all singles → RE for singles, which is (part of) the singles weight for wOBA

Run Expectancy Matrix 2010-2015			
Runners	0 outs	1 outs	2 outs
_ _ _	0.481	0.254	0.098
1B _ _	0.859	0.509	0.224
_ 2B _	1.100	0.664	0.319
1B 2B _	1.437	0.884	0.429
_ _ 3B	1.350	0.950	0.353
1B _ 3B	1.784	1.130	0.478
_ 2B 3B	1.964	1.376	0.580
1B 2B 3B	2.292	1.541	0.752
SOURCE: Tom Tango			

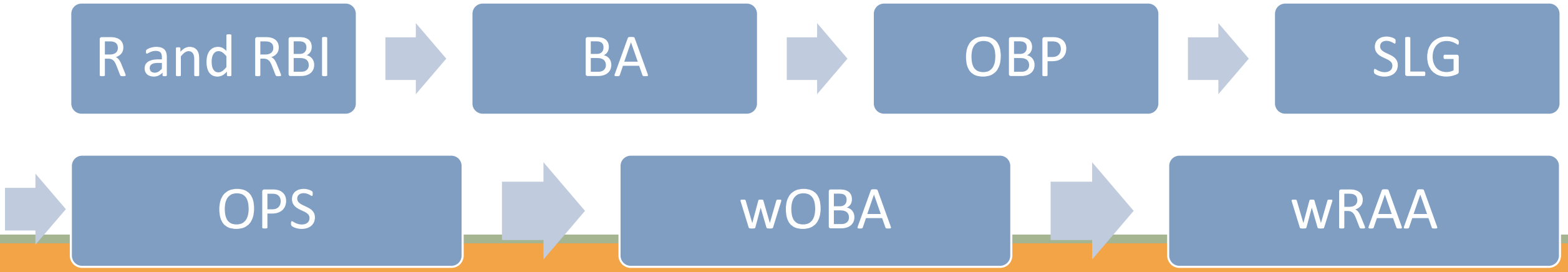
Source: https://library.fangraphs.com/the-beginners-guide-to-deriving-woba/

Batting

- Weighted Runs Above Avg (wRAA) = $\frac{wOBA - League\ avg.wOBA}{wOBA\ Scale} * PA$

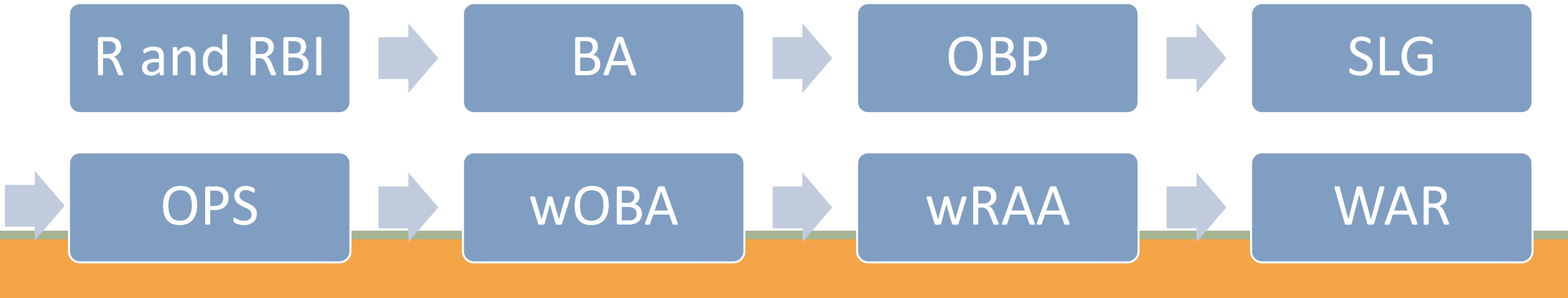
Undoes scaling so wOBA represents actual expected runs

  - Translates wOBA to actual runs above *average* player
  - Problems/limitations?



Batting

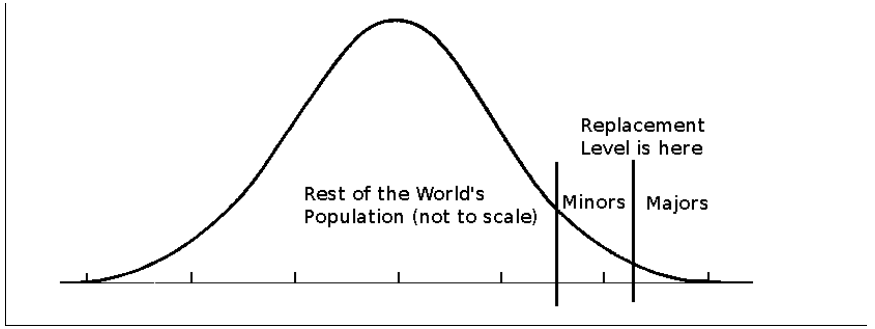
- (Batting) Wins Above Replacement (WAR) = *Complicated*
  - Adjusts for park, league, position; compares to **replacement player**
  - Then translates runs to wins (1 win ≈ 9-10 runs)
  - **Ultimate value. How much $ is a win worth x (total) WAR = contract value**



Replacement Level

- **Replacement Level:**

Replacement level is simply the level of production you could get from a player that would cost you nothing but the league minimum salary to acquire. Minor league free agents, quad-A players, you get the idea. The concept is pretty tidy. These are the



This definition exists because we want to be able to compare the number of wins a player is worth compared to the player a team would have to acquire to fill their shoes. If a great player is making $20 million and a replacement player is making $500,000, that great player is providing you X number of wins for $19.5 million because you would have to allocate that half million to the roster spot no matter what.

Source: https://library.fangraphs.com/misc/war/replacement-level/; https://www.baseball-reference.com/about/war_explained.shtml

“Let’s See Your Stats Capture This”

Source: https://gfycat.com/shallowscarcegermanspitz-los-angeles-trea-turner-baseball-dodgers; http://hyunyrun.blogspot.com/2020/05/kbo-bat-flip-collection-gif.html; https://giphy.com/explore/bat-flip

Fielding

- Errors
  - Problems/Limitations?

Errors

Fielding

- Fielding Percentage = $\frac{Putouts+Assists}{Putouts+Assists+Errors}$
  - **Rate stat** vs. **counting stat**
  - Problems/Limitations?

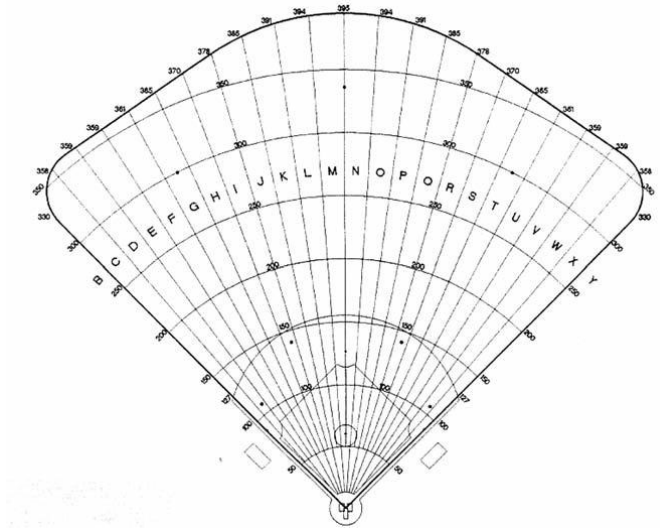
Errors → Fielding %age

Fielding

- Defensive Average (DA) or Zone Rating (ZR) =

$$\frac{Play\ Made}{Opportunities\ in\ Zone}$$

Turned ball into an out



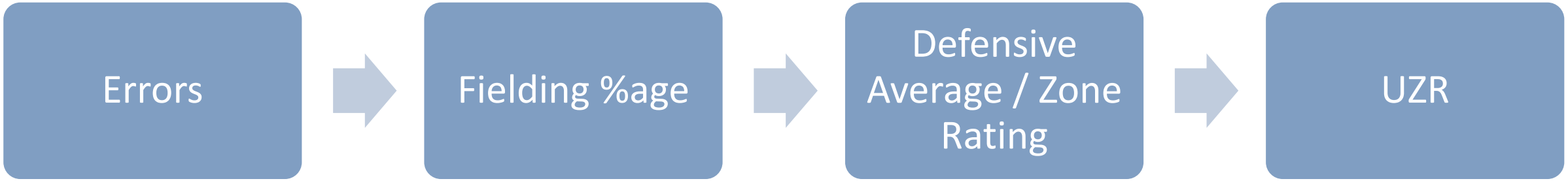
  - Counts balls a fielder misses
  - Problems/Limitations?

Errors → Fielding %age → Defensive Average / Zone Rating

Source: https://sabr.org/journal/article/measuring-defense-entering-the-zones-of-fielding-statistics/

Fielding

- Ultimate Zone Rating (UZR): extremely complicated
  - Accounts for park, skill turning double plays, outfield assists; just generally more comprehensive
  - Translates to runs saved
  - Problems/Limitations?



Fielding

- (Fielding) WAR = *Complicated, again*
  - Translates UZR runs to wins (1 win ≈ 9-10 runs); treats catchers separately (steals and passed balls prevented); adjusts for position
  - **Ultimate value. How much $ is a win worth x (total) WAR = contract value**

Pitching

- Wins (or Win-Loss Record, W-L)

Indians

In pitching, the only thing that really matters is wins: Paul Hoynes' Rant of the Week

Updated Jan 12, 2019; Posted Sep 12, 2010

- Problems/limitations?

W-L

Source: https://www.cleveland.com/tribe/2010/09/in_pitching_all_that_matters_a.html; https://nesn.com/2014/04/mets-pitcher-bartolo-colon-swings-hard-misses-pitch-loses-helmet-gif/

Pitching

- Earned Run Average (ERA) = $\frac{Runs\ Not\ the\ Result\ of\ an\ Error}{Innings\ Pitched\ (IP)} * 9$

  - Accounts for *errors* by defense, eliminates offensive performance
  - Problems/limitations?

W-L → ERA

Breaking Down Pitching Performance

- Fundamental problem:

$$Pitching\ Performance = Pitcher\ Skill + Defense\ Skill + Luck$$

- How do we isolate *just* pitcher skill?

- Enter **Defense-Independent Pitching Statistics (DIPS)** – despite name, try to deal with defensive skill *and* luck

Source: https://i.gifer.com/OOxX.gif

Breaking Down Pitching Performance

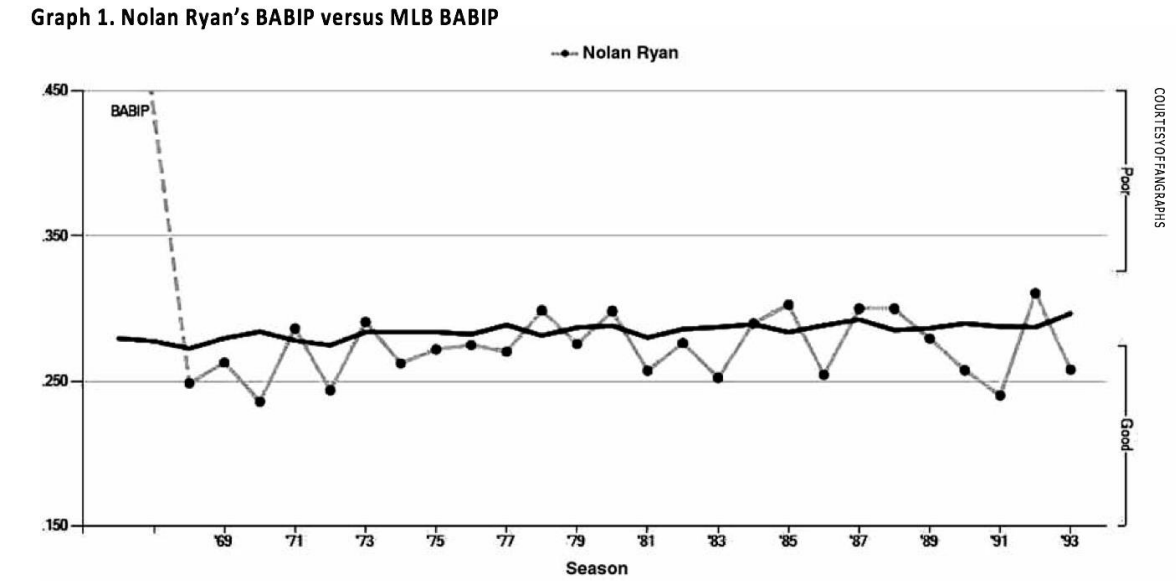
- $Pitching\ Performance = Pitcher\ Skill + Defense\ Skill + Luck$

- In any plate appearance, the ball can end up **in-play** (must be fielded) or **not-in-play** (strikeout (K), BB, HBP, HR)

  - Voros McCracken's key insights:
    - 1) We should split outcomes into these categories
    - 2) Pitchers have complete control over **not-in-play** outcomes, but are at the mercy of the defense for **in-play** balls

Batting Average on Balls in Play (BABIP)

- $Pitching\ Performance = Pitcher\ Skill + Defense\ Skill + Luck$
- Strong year-to-year correlations among pitchers for strikeouts and walks (≈0.7-0.8 or higher), pretty strong for HRs (≈0.4-0.5)
- Much weaker for BABIP (≈0.15-0.25)



Source: http://sabr.org/journal/article/the-many-flavors-of-dips-a-history-and-an-overview/

Pitching

To make FIP read like ERA

- Fielding Independent Pitching (FIP) = $\frac{13*HR+3*(BB+HBP)-2*K}{IP} + Constant$
  - Only measures things a pitcher has or appears to have control over ("defense-independent statistics" that are consistent year-to-year); excludes all BIP
  - Problems/limitations?

W-L → ERA → DIPS (FIP)

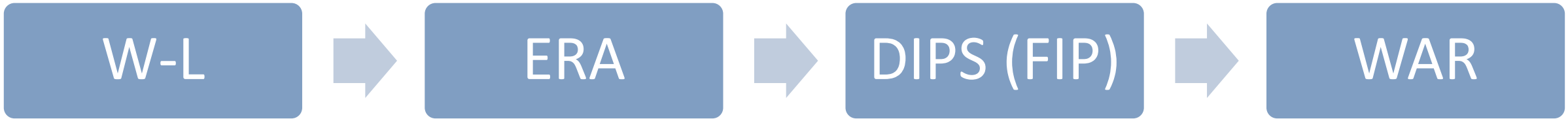
Ground and Fly Balls

- Of course it's not really that simple
- Pitchers don't control *overall* BABIP, but do control:
  - % ground balls (GB)
  - % fly balls (FB)
  - Ratio GB/FB
  - Etc.
- Stats like tRA, QERA, SIERA take this batted ball data into account
- In-season vs. next season projections

Source: https://medium.com/@mxoliverleigh/really-it-should-not-be-this-complicated-fd8a13056747

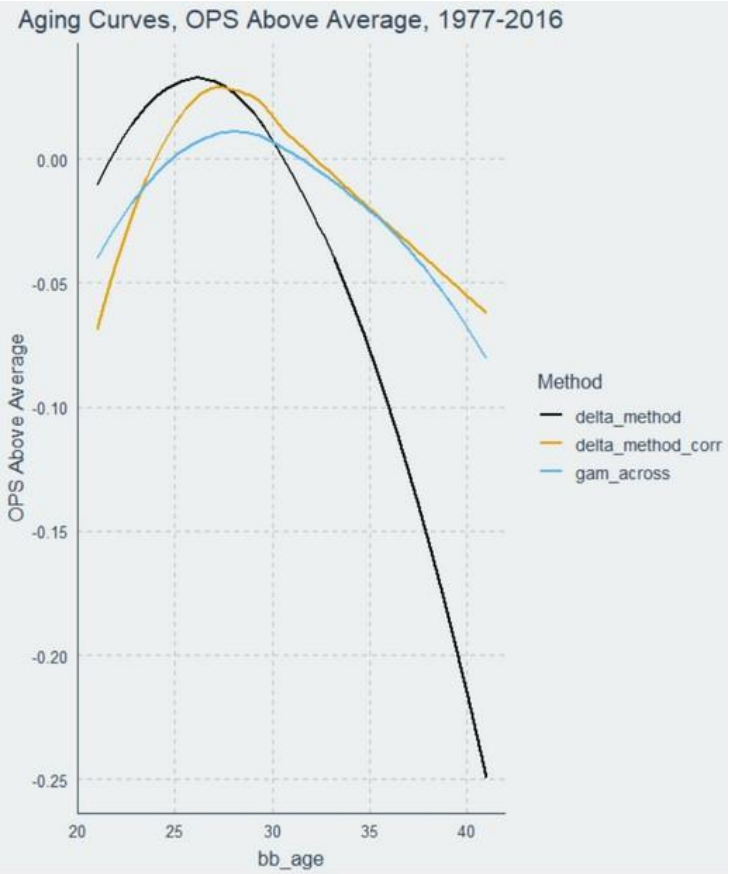
Pitching

- (Pitching) WAR = *Complicated, again*
  - Adds infield flies to FIP, translates to runs, adjusts for park and league, then translates to wins (but here every pitcher has a different runs per win metric), compares to replacement level
  - Remember this is Fangraphs WAR; other sources may use something other than FIP as basis
  - **Ultimate value. How much $ is a win worth x (total) WAR = contract value**



Predicting Future Value

- Different from **cross-sectional** stats describing what an athlete *is*, we need to know what an athlete *will be*
- Two challenges: 1) Try to identify a player's **true talent**, and then 2) project that forward
- **Repeatability:** look for stats that are either stable year-to-year, or at least predictive of future performance
  - Not luck-based
  - Adjust for things you know will happen (e.g. **aging curves**)



Source: https://www.baseballprospectus.com/news/article/59972/the-delta-method-revisited/, https://arxiv.org/abs/1802.00998

What Makes a Good (Sports) Statistic?

The Question

- A statistic starts with a QUESTION
- It should be designed to answer that QUESTION
- **A statistic may be good for one QUESTION but not another**
  - Be *transparent* and *explicit* about what your QUESTION is and what the statistic is designed to do – and, critically, *not* do!

Source: https://twitter.com/bernybelvedere/status/1171435503586926593

Match Your Stat to Your Question

- How good is this hitter? – wOBA
- How good is this pitcher? – FIP
- How good *will this pitcher be?* – SIERA

Match Your Stat to Your Question

- Teams: How much should I pay this player? – WAR, and make sure it has *predictive* not just *descriptive* value

- Fantasy Managers: Who should I draft or keep? – ZiPS Projections

Match Your Stat to Your Question

- Who's the GOAT? – Need era-, park-, league-adjusted statistic
  - Some aren't, because <u>they're meant for another question!!</u>

Source: https://www.tvovermind.com/is-the-new-match-game-doing-justice-to-the-original/

Three Types of Questions

- **1. Descriptive**
  - Describing the world as it *is*
- **2. Predictive**
  - Describing the world as it *will be*
- **3. Causal (Counterfactual Prediction)**
  - Describing the world as it *could be*

General Qualities of a Good Statistic

- Grabiner, Sabermetric Manifesto:
  - 1. Is the stat important for success in the sport?
  - 2. Does it capture/isolate an individual player's contribution?
    - Pre-supposes a player valuation QUESTION!
  - 3. Is there a better alternative to measure what you're trying to measure?
- We'll add:
  - 4. Is it repeatable/stable?
    - Is it measuring signal or noise? "True talent" or luck?

Source: http://www.seanlahman.com/baseball-archive/sabermetrics/sabermetric-manifesto/

Better Alternatives, and the Scientific Process

- 3. Is there a better alternative to measure what you're trying to measure?
  - **Science is built on accumulating small answers that lead to big insights – "get a little less wrong each day"**
  - **No single analysis answers everything – understand and accept that**
  - BUT sometimes an analysis is worse than no analysis at all; judgement call
    - CDC MMWR from Provincetown: vaccinated vs. unvaccinated viral loads with COVID delta variant

Assessing Our Baseball Statistics

- 1. Is the stat important for success in the sport?

Batting	Fielding	Pitching
R or RBIs	Errors	W-L
BA	Fielding Average	ERA
OBP	Defensive Average/Zone Rating	FIP
SLG	UZR	(Pitching) WAR
OPS	(Fielding) WAR	
wOBA		
wRAA		
(Batting) WAR		

Assessing Our Baseball Statistics

- 2. For player valuation questions: does it capture/isolate an individual player's contribution?
  - Which depend on player's: Teammates? Ballpark? Era?

Batting	Fielding	Pitching
R or RBIs	Errors	W-L
BA	Fielding Average	ERA
OBP	Defensive Average/Zone Rating	FIP
SLG	UZR	(Pitching) WAR
OPS	(Fielding) WAR	
wOBA		
wRAA		
(Batting) WAR		

Assessing Our Baseball Statistics

- 3. Is there a better alternative to measure what you're trying to measure?

Batting	Fielding	Pitching
R or RBIs	Errors	W-L
BA	Fielding Average	ERA
OBP	Defensive Average/Zone Rating	FIP
SLG	UZR	(Pitching) WAR
OPS	(Fielding) WAR	
wOBA		
wRAA		
(Batting) WAR		

Assessing Our Baseball Statistics

- 4. Is it repeatable/stable
  - Is it measuring signal or noise? “True talent” or luck?

Batting	Fielding	Pitching
R or RBIs	Errors	W-L
BA	Fielding Average	ERA
OBP	Defensive Average/Zone Rating	BABIP
SLG	UZR	FIP
OPS	(Fielding) WAR	(Pitching) WAR
wOBA		
wRAA		
(Batting) WAR		

Thanks!

- Questions? zbinney@emory.edu, @binney_z on Twitter