NEW YORKER

WHAT BASEBALL TEACHES US ABOUT MEASURING TALENT

The clash between data and intuition opens onto a larger debate.



By Louis Menand April 1, 2019



You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe now.</u>

Expand



Scouts judge and "scorers" measure. Which method should we trust? Illustration by David Plunkert

0:00 / 23:07

Audio: Listen to this article. To hear more, download Audm for iPhone or Android.

The subject of Christopher Phillips's "Scouting and Scoring: How We Know What We Know About Baseball" (Princeton) is baseball, but it's worth reading for more than just the baseball. The book is an effort to help us understand one of the oldest problems in modern societies, which is how to evaluate human beings. Do we scout or do we score?

The "scouting" in Phillips's title refers to the traditional baseball scout. He's the guy who sizes up the young prospect playing high-school or college ball, gets to know him away from the diamond, and draws on many years of experience hanging out with professional ballplayers to decide what the chances are that this one will make it to the bigs—and therefore what his price point should be for the club that signs him.

The "scorer" is what's known in baseball as a sabermetrician. (And they don't call it scoring; they call it "data capture.") He's the guy who punches numbers into a laptop to calculate a player's score in multivariable categories like war (wins above replacement), fip (fielding independent pitching), whip (walks plus hits per inning pitched), woba (weighted on-base average), and O.P.S. (on-base percentage plus slugging). Quantifying a player's production in this way allows him to be compared numerically with other available players and assigned a dollar value.

The scout thinks that you have to see a player to know if he has what it takes; the scorer thinks that observation is a distraction, that all you need are the stats. The scout judges: he wants to know what a person is like. The scorer measures: he adds up what a person has done. Both methods, scouting and scoring, propose themselves as a sound basis for making a bet, which is what major-league baseball clubs are doing when they sign a prospect. Which method is more trustworthy?

The question is worth contemplating, because we're confronted with it fairly regularly in life. Which applicant do we admit to our college? Which comrade do we invite to join our revolutionary cell? Whom do we hire to clean up our yard or do our taxes? Do we go with our intuition ("He just *looks* like an

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe</u>

Expand

Expand

Many readers will already be familiar with the scout-versus-scorer dilemma in baseball from Michael Lewis's best-selling "Moneyball," which was published in 2003 and made into a movie, starring Brad Pitt. "Moneyball" is the story of how a baseball team that did not have a lot of money to spend on players, the Oakland A's, deployed a new way of evaluating talent and proceeded, for several years, to compete with teams that had much bigger stars and much higher payrolls, like the New York Yankees. It was a way for small-market teams to keep up with their richer big-city rivals.

Lewis colorized his story a bit by casting the scouts as a bunch of Don Zimmer-y old-timers who spit tobacco juice and say things like "I can see this guy in somebody's pen throwing aspirin tablets someday" (meaning he throws hard) and "This kid wears a large pair of underwear" (meaning his body is wrong for baseball). The scouts put a lot of stock in whether a player has "the good face"—a time-honored term of the scouting art. Lewis writes, "The old scouts are like a Greek chorus; it is their job to underscore the eternal themes of baseball."

Lewis's "scorers" are geeky Harvard grads who speak stats-talk and who actively disidentify with the culture of the game. Their whole approach is based on disdaining the wisdom of the scouts. They don't need to see prospects, they don't even need to see games, because, for them, a player is not a body; he's a row of numbers. As it must be for all industry disrupters, the scorers' advice *has* to be the opposite of the scouts': if it was identical, their services would not be needed. As Lewis puts it, "The new outsider's view of baseball was all about exposing the illusions created by the insiders on the field."

Between the scouts and the scorers in "Moneyball" is the general manager of the A's, Billy Beane, a former hot prospect who fizzled out in the major leagues, and therefore knows the world of the scouts, but who, from a kind of manic desperation, puts his faith in the geeks and is rewarded for it by the sporting gods. Lewis is a journalist; he's trying to tell a story. But his sympathies are with the scorers. Phillips is an academic. His field is the history of science, and he is not telling a story. He is making an argument based on scholarship. And "Scouting and Scoring" is basically presented as an answer to "Moneyball."

People-measuring arose with the modern nation-state—the word "statistics" comes from the word for "state." In the beginning, the data that states collected were demographic: population size, birth and death rates, marriage, disease. In the early nineteenth century, statistical methods started being applied to human beings, and were used to determine, for example, the chest size of the average Scot.

Statistics were also used to make predictions. How many suicides would there be in France next year? How many homicides? How many homicides involving the use of poison? It turned out that there were regularities in all these categories. There was a kind of natural law governing the rate of murder by poison.

Debate ensued about whether the state could reduce the average annual number of these murders by, say, restricting access to poison. The statistician's—the scorer's—answer was that poison doesn't kill people; people kill people. Homicides are going to occur at a certain rate per unit of population no matter what the laws are. We're having the same debate today.

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe</u> now.

people found this sort of thing deterministic and upsetting. Why? Again, it was a case of scouts versus scorers. Scout types refused to believe that people decide to marry in accordance with some statistical law. Marriage is about feelings, not the price of goods. Scorer types argued that, whatever unique sentiments might motivate the partners to marry in any specific case, the numbers do not lie. The fact is that when corn is cheap more people will get married. It's a prediction you can take to the bank.

It is during this period of statistics mania, the mid-nineteenth century, that baseball, basketball, tennis, football, rugby, soccer, and other forms of organized athletic competition come into prominence, and we start to get leagues and championships and standardized rules of play. It is also when we start to get widespread scoring in sports. As both Phillips and Lewis point out, baseball scoring begins almost as soon as there is baseball. The first known box scores, breaking down the stats of games, appear in 1845; by the eighteen-sixties, Phillips says, "nearly everyone in baseball was counting in one form or another."

The key figure in making these numbers official and insuring their accuracy was a man named Henry Chadwick, a British émigré who ended up working as a sports reporter in Brooklyn. Chadwick was obsessed with scorekeeping not just because it was a way to memorialize games but because he thought that data science could explain why games are won or lost. "In time," he said in 1868, "the game will be brought down almost to a mathematical calculation of results from given causes." Chilling words for fans who think that today's game is overquantified.

Welcome words, however, for gamblers. Keeping score is a natural pastime. Some fans bring their own scorecards to games, and most fans check the box scores and other stats when they read the sports page. Numbers are part of the sports experience—especially for fans of baseball, which seems to have a statistic for everything. Since at least the eighteenth century, though, a lot of people have followed sports mainly for the purpose of placing bets, and there is little question that the creation of reliable statistical summaries of players' performances had something to do with the gambling market.

Baseball bettors don't bet only on outcomes. They bet on how many runs the teams will score, how many bases will be stolen, how many bunts will be laid down, how many pitches will be thrown. They could not do this if they did not have a lot of data on the players involved. A team's manager wants to be able to predict how certain players will perform in certain situations. So does the gambler.

In fact, the stats revolution that Lewis writes about in "Moneyball" took off at the same time as a gambling-like offshoot of the game, fantasy baseball—or, as it used to be known, after the Manhattan restaurant in which it was conceived, in 1980, Rotisserie Baseball. Fantasy baseball is played for money (usually small amounts). Participants select a lineup of active players, and winners and losers are determined by how those players end up performing on a given day. Obviously, the better the information about the players who are available for selecting, the better the chances of winning.

In the beginning, as Lewis points out, fantasy leagues used conventional measures of performance, because that's all there was—things like earned-run average and runs batted in, categories that the scorers in "Moneyball" would view with derision. But fantasy baseball gave participants an incentive to find or to create

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe now.</u>

Expand

Expand

baseball, Bill James, has said that he started rethinking how games are won because he wanted to win at a tabletop game that was a precursor to fantasy baseball.

In 1982, two years after Rotisserie Baseball got started, James's book of new statistical measures, "Baseball Abstract," became a best-seller. Two years after that, "The Hidden Game of Baseball: A Revolutionary Approach to Baseball and Its Statistics," by John Thorn and Pete Palmer, came out, and sabermetrics was born. The success of those books, as Phillips says, showed that baseball data could be monetized. Today, virtually everything in baseball that can be measured is measured: exit velocity, launch angle, spin rate. Even sportscasters load the play-by-play with statistics. Geeks rule.

Phillips says complimentary things about "Moneyball," and his own book is more a correction than a refutation. He brings considerable historical knowledge to the task of establishing a point that does not actually seem all that controversial, which is that scouting involves measuring and scoring involves judging. "Facts don't just appear," as Phillips puts it. "They must always be made."

Scouts and scorers may be looking at different things, but knowing which things to look at involves a judgment. Do we care about how far a ball was hit? What the defensive alignment was? The quality of the umpiring? Turning knowledge of the things that are judged relevant into a prediction requires a quantification, even if it's only of the "on a scale of one to ten" variety. That's still an act of reducing information to a number.

"What I discovered was that historically the ways scorers and scouts produced knowledge and established facts were not all that different," Phillips writes. "Any claimed division between scouting's judgment-based subjectivities and scoring's data-based objectivities doesn't have a strong purchase." The scout's job is the same as the scorer's: he puts a number on a player, and that number represents the player's potential value.

MORE FROM THIS ISSUE **APRIL 8, 2019 AMERICAN CHRONICLES CUPERTINO POSTCARD** SHOUTS & MURMURS **FICTION** The Challenge of Going Tim Cook's Big Apple It's a Flat Earth "Lulu" Off Psychiatric Drugs Circus By Te-Ping Che By Gary Richardson By Rachel Aviv By Nathan Heller

To show how intertwined judging and measuring are, Phillips spends a lot of time on a concept that was crucial to player evaluation: the error. The error was an object of fascination for Henry Chadwick, too. If a

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe now.</u>

really an error? Here, the scout has the advantage, because the error is a judgment call. You have to have seen a lot of games to know whether the fielder should have made the play. There is no substitute for experience. You are not going to learn how to score an error in a Harvard math class. (But the math department is working on it. It has become possible, based on the hang time of a batted ball and its distance from a defensive player, to calculate the probability that it will be caught.)

Errors are a subset of a much more cosmic category: luck. It's an adage of professional sports that lucky breaks, like bad calls, even out in the course of a season. (Actually, given a normal distribution of lucky breaks, there must be outliers. There is a luckiest person in the world. Probability theory requires it.) A major aspect of scoring, therefore, is figuring out how to take luck out of the equation.

Largely, this means discounting for fielding. Once a ball is put into play, many things can happen to it. A spectacular catch of a ball headed for the seats is recorded as an out; a dribbler in front of the plate that no defender can get to in time is a hit. And the evidence shows that these occurrences don't necessarily even out. One of the findings of sabermetricians (as reported by Lewis) is that the number of hits surrendered by a pitcher can be quite inconsistent from year to year. Batted balls that are caught one year just happen to "drop in" in another.

But outcomes that a pitcher *can* control—mainly walks, strikeouts, and home runs—are predictable, which argues for making those, rather than hits and runs allowed, the proper basis for evaluating pitchers. Over all, the luck factor makes the scorer's job harder, since scorers need reliable numbers to work with. But it hardly affects the scout's job at all, since scouts are looking at intangibles like discipline, athleticism, and the will to win.

T alent has to be evaluated in every professional sport. Why all the attention to baseball? One answer is that the gap in the level of play between high-school or college ball and major-league ball is very great. In basketball and football, many college stars can play in the N.B.A. or the N.F.L. right away. That's not the case in baseball.

More than a thousand baseball players are drafted every year, and less than ten per cent of them ever play in the majors. Everyone else toils away somewhere in baseball's enormous farm-team archipelago: two hundred and forty teams, more than seven thousand players. (There are seven hundred and fifty active major leaguers.) Virtually everyone in baseball spends time in that system, which means that scouts and scorers are often trying to guess what a seventeen- or twenty-year-old is going to play like when he's twenty-five and finally arrives in the big leagues.

What makes this process even more difficult is that the game changes, and precisely according to what teams value in their players. The "shift" is a much discussed recent example—moving most defenders to one side of the field against hitters who are predicted to pull the ball. With a lower chance of getting base hits, those batters now try to hit home runs. (This is why "launch angle" has become so important.) So do you value future home-run hitters? Or hitters who can put the ball in play anywhere on the field? You can't know which type of player will be needed five or eight years in advance. The ideal type is a moving target

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe now.</u>

Expand

body type does. For the scorers, though, players who know how to get bases on balls in college will get bases on balls in the majors. It doesn't matter what sort of physique they have. In either case, evaluation is not simple, and since there is more and more money involved, there is a high premium on getting it right.

Lewis's book has a lot of examples where scouts got it wrong but scorers got it right, so it's regrettable that Phillips doesn't provide much in the way of examples where the reverse is true. That may be because he wants to make a more philosophical point about the nature of data—that they're always a hybrid of objectivity and instinct, analytics and intuition. He concludes that scouting has as much claim to being scientific as scoring does: "it serves as a well-developed and well-crafted set of heuristics for arriving at stable, generalizable, and reliable facts about the natural world." He argues that since teams continue to use scouts, their experience must be irreplaceable.

Phillips does appear to sidestep what may have been for many readers the revelatory takeaway of "Moneyball," which is that, for decades, baseball was sunk in the sports equivalent of primitive theology. Baseball minds genuflected before idols—the stolen base, the sacrifice bunt, the hit-and-run—that turned out to have little to do with winning games in the real world of professional sports. Baseball players are notoriously superstitious, and this trait seems to have infected the culture of scouting a little.

That's what Billy Beane—who didn't do the math himself—learned from his scorers. A lot of the action in "Moneyball" is not about draft prospects but about players who are already big leaguers and who either become available for trade or are on the market as free agents. This is an area where Beane's approach paid off. He was able to snatch up players who were undervalued according to the old rubrics, and to acquire players no one else saw value in, because they didn't know the new rubrics.

One of Beane's favorite strategies relied on the overvaluation of closers. The closer is a relief pitcher who typically enters the game in the ninth inning with his team ahead by three runs or less and, normally, with no runners on base. If he preserves the lead, he is credited with a save. If you think about it, this is the easiest job on the pitching staff. The closer comes into the game with a fresh arm to face batters who are seeing him for the first time, and he starts out already ahead. All he has to do is get three outs before the other team can tie the score.

But, starting in the nineteen-seventies, a mystique grew up around the closer, partly because pitchers like Sparky Lyle, Goose Gossage, and Al (the Mad Hungarian) Hrabosky—and, later on, Jonathan Papelbon and Brian Wilson—developed outsized mound personalities. Major facial hair, hulking frames, Frankenstein's-monster windups, and demonic intensity became the attributes of the closer persona. The closer was a berserker, a danger artist, a Lord High Executioner—even though all he was doing was mopping up games his teammates had already won.

The monetary value of the closer got inflated accordingly. So Beane took so-so minor leaguers and starting pitchers who had begun to fade with age, and he turned them into closers. This drove up their stats and hence their market value, and they became trade bait for other teams. Beane could swap those pitchers, repurposed as lights-out relievers, for younger and cheaper players

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe now.</u>

Expand

realms—for example, college admissions. The debate today over fairness in college admissions is oversimplified to the point of absurdity. There is no single standard for admissions at select colleges, because there are many different buckets to fill. When one applicant displaces an arguably more meritorious applicant, she is almost always displacing an applicant within her own bucket.

Still, you can see the scout-versus-scorer opposition in the way people talk about admissions. Some critics of the current system deplore the reliance on standardized-test scores, on the ground that privileged students are prepped for the tests. And some critics—often they are the same ones—identify certain accepted applicants (legacies, varsity athletes, Jared Kushner) as undeserving because their test scores are below average.

In other words, people want college admissions to reflect SAT scores and G.P.A.s (scoring), but people also want them to consider an applicant's background, motivation, and personal qualities, like having overcome disadvantages (scouting). One standard is frequently going to be at the expense of the other.

Are there other lessons to be learned from these stories about baseball scouts and scorers? Lewis and Phillips both seem to think so. For Lewis, the important lesson was about business. What the scorers working for the Oakland A's showed was that markets have inefficiencies, and that if you can find those inefficiencies and exploit them you can beat your competition.

This was a lesson appropriate to the era in which "Moneyball" was written, shortly after the first dot-com boom, in the days of a startup-, hedge-fund-, derivatives-fuelled gold rush. To put it in scout-versus-scorer terms, that was a time when hands-on experience, bred-in-the-bone wisdom, and seat-of-the-pants intuition—the "human element" in decision-making—were perceived as obstacles to less expensive and more efficient ways of doing things. Crunching the numbers was how you got ahead.

Today, this seems a very 2000 way of thinking. Reading "Moneyball" back then, you found yourself rooting for Billy Beane and his geeks against the bloated Yankees and their bullying owner, George Steinbrenner. Now you realize that what you were actually rooting for was a bunch of guys who were trying to figure out a way to underpay their players. In exchange for the chance to be in the big leagues, Oakland's players settled for less compensation than they were worth, simply because no one else knew how to value them. There are foreshadowings of the gig economy in the "Moneyball" story.

Phillips's book is appropriate to our more chastened post-recession moment, when social confidence in Big Tech is going through a rough patch. His book is a reminder that algorithms and machine intelligence are only extensions of the men and women who create them, and that there is no substitute for human judgments based on experience with actual people. Scorer types aren't interested in history; Phillips tries to show us that knowing the past can help us grasp what's at stake in the choices we make in the present.

Phillips has another lesson to draw, though. This one is about the death of expertise. "Though it is a leap from *Moneyball* and data science to the rechristening of falsehood as 'alternative facts' in 2017 by a special advisor to the president of the United States," he writes, "self-styled outsiders, armed with a bit of data, have

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe now.</u>

Expand

expert."

Whoa. It *does* seem a bit of a leap from "Moneyball" to climate-change deniers and post-truth politicians —O.K., lying politicians. Is our political mess really a scouts-and-scorers situation? The deniers and the liars in public life don't use "bits of data." They just assert, on the theory that many people prefer to believe what they already want to believe. The comparison also seems out of proportion. We need to trust our public institutions, and that crisis deepens every day. But we don't need to trust our team's front office. They are free to make up the facts. The peril is on them.

Phillips seems to be ignoring the lesson of his own book, which is that history shows that the tension between scouting and scoring is always with us. It's never going to go away. If we're tipping toward a scoring phase right now, somewhere down the road we'll tip back the other way.

And in the end a fresh consensus, a new conventional wisdom, does emerge. Now that everyone in baseball has figured out how to use sabermetrics, we're back to where we started, with the rich clubs lording it over the rest. Bill James now works as a consultant for the Boston Red Sox, who won the most recent World Series and have the highest payroll in baseball. They are planning to get through next season without hiring a closer. •

This article appears in the print edition of the April 8, 2019, issue, with the headline "Twist and Scout."



Louis Menand has been a staff writer at The New Yorker since 2001. He teaches at Harvard University. Read more »

Read something that means something.

Try *The New Yorker* and get a free tote. Cancel anytime. Subscribe now. »



CONDÉ NAST

© 2019 Condé Nast. All rights reserved. Use of and/or registration on any portion of this site constitutes acceptance of our User Agreement (updated 5/25/18) and Privacy Policy and Cookie Statement (updated 5/25/18). Your California Privacy Rights. The material on this site may not be reproduced, distributed, transmitted, cached or otherwise used, except with the prior written

You are reading your last free article this month. Try *The New Yorker*, plus get a free tote. <u>Subscribe now.</u>

Expand