

University of Chicago Law School

Chicago Unbound

Journal Articles

Faculty Scholarship

2004

Market Efficiency and Rationality: The Peculiar Case of Baseball 2004 Survey of Books Relating to the Law

Cass R. Sunstein

Richard H. Thaler

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles



Part of the [Law Commons](#)

Recommended Citation

Cass R. Sunstein & Richard H. Thaler, "Market Efficiency and Rationality: The Peculiar Case of Baseball 2004 Survey of Books Relating to the Law," 102 Michigan Law Review 1390 (2004).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

MARKET EFFICIENCY AND RATIONALITY: THE PECULIAR CASE OF BASEBALL[†]

Richard H. Thaler and Cass R. Sunstein***

MONEYBALL: THE ART OF WINNING AN UNFAIR GAME. By *Michael Lewis*. New York: W. W. Norton, 2003. Pp. ix, 288. \$24.95.

In this lively book, Michael Lewis¹ explores a topic that would seem of interest only to sports fans: how Billy Beane, the charismatic general manager of the Oakland Athletics, turned his baseball team around using, of all things, statistics. What next — an inspirational tale about superior database management? But there are some general lessons in Lewis's book that make it worth the attention of people who do not know the difference between a slider and a screwball (a group that, unfortunately, includes many lawyers and law professors). Those lessons have to do, above all, with the limits of human rationality and the efficiency of labor markets. If Lewis is right about the blunders and the confusions of those who run baseball teams, then his tale has a lot to tell us about blunders and confusions in many other domains. In that sense, the tale bears directly on continuing debates about behavior, cognition, and law.²

I. MEASURING PERFORMANCE

Lewis focuses on the extraordinary success of Beane, who has produced a terrific baseball team despite one of the lower payrolls in baseball. Since 1999, when Beane took over, the Athletics have

† This review is a significantly revised and expanded version of *Who's On First*, originally published in *The New Republic*; the relevant parts are reprinted by permission.

* Robert P. Gwinn Professor of Economics and Behavioral Science, University of Chicago Graduate School of Business. B.A. 1967, Case Western Reserve; Ph.D. 1974, University of Rochester. — Ed.

** Karl N. Llewellyn Distinguished Service Professor of Jurisprudence, University of Chicago Law School and Department of Political Science. A.B. 1975, J.D. 1978, Harvard. — Ed.

1. Journalist and popular author, whose other works include the national best-seller *Liar's Poker* and articles in numerous national periodicals.

2. See *BEHAVIORAL LAW AND ECONOMICS* (Cass R. Sunstein et al. eds., 2000). For critical discussion, see RICHARD A. EPSTEIN, *SKEPTICISM AND FREEDOM* (William M. Landes & J. Mark Ramseyer eds., 2003), and RICHARD A. POSNER, *FRONTIERS OF LEGAL THEORY* (2001).

compiled an amazing record. In 1999, the Athletics ranked eleventh (out of fourteen teams) in the American League in payroll and fifth in wins. In 2000, the Athletics ranked twelfth in payroll and second in wins, a feat that they duplicated in 2001. In 2002, the last year covered in the book, they ranked twelfth in payroll again — and first in wins. This past season, 2003, was yet another division-winning season with a cheap payroll.

How did Beane pull this off? He did it largely by ignoring or defying baseball's conventional wisdom, otherwise known in baseball lingo as *The Book*. (As in, "The Book says that you should bunt in this situation.") It turns out that many chapters of *The Book* are simply wrong. Sacrifice bunts are rarely a good strategy, and steals are vastly overrated. (Unless a base stealer succeeds at least two-thirds of the time, his running efforts reduce runs scored rather than increase them.³) The portion of *The Book* that was most in need of revision, and the most important edge that Beane was able to exploit, involved player evaluation. Here he tried to figure out, statistically, how much a player was likely to contribute to his team's chances of winning. If he couldn't measure a factor or its impact, he dismissed it as subjective.

Beane found that, as a statistical regularity, players drafted out of high school are much less likely to succeed than players drafted out of college (p. 16). And so he drafted no high school players, regardless of how highly they were touted. He hired a young assistant named Paul DePodesta, a Harvard economics graduate, who relied on his computer to project players' performances, without so much as ever seeing a player swing a bat. Much of the tension, and the comedy, of Lewis's book comes from the conflict between Beane's and DePodesta's statistical methods of evaluation and the well-established strategies of experts who have scouted, played, and breathed baseball for decades. The verdict? Statistical methods outperform experts.⁴ It's not even close.

As Lewis tells the tale, Beane's particular approach has intensely personal foundations. Beane himself was a top high-school prospect, one of the most sought after in the nation. He was fast; he was tall; he was strong; he could hit the ball a mile. The baseball scouts loved him. As one of them admitted, "I never looked at a single statistic of Billy's. It couldn't have crossed my mind . . . He had it all" (p. 9). According to those who watched him, "The boy had a body you could dream on. Ramrod-straight and lean but not so lean you couldn't imagine him filling out. And that face!" (p. 7). Beane was selected in the first round

3. See JOHN THORN & PETE PALMER, *THE HIDDEN GAME OF BASEBALL* (1985).

4. This is a general theme in social-science research. See, e.g., Robin Dawes et al., *Clinical Versus Actuarial Judgment*, in *HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT* 730 (Thomas Gilovich et al. eds., 2002); William Meadow & Cass R. Sunstein, *Statistics, Not Experts*, 51 *DUKE L.J.* 629 (2001).

of the draft, with the highest of expectations. He was destined to be a star.

There was just one problem: Beane did not play professional baseball very well. He thought too much. He was too emotional. His failures notwithstanding, baseball people saw his body, and his face, and his raw talent, and concluded that he was bound to succeed. "Teammates would look at Billy and see the future of the New York Mets. Scouts would look at him and see what they had always seen . . . The body. The Good Face" (p. 47). He certainly had talent, and once in a while he would do something truly sensational. But after several years in major league baseball, his performance was woefully bad. With only 301 at-bats, he hit .219; more embarrassingly, he had 80 strikeouts and only 11 walks. Abruptly, he quit the game. While playing for Oakland, he told the team's general manager that he no longer wanted to be a player, and would prefer the job of advance scout, an employee who travels ahead of the team to analyze future opponents. The team's general manager was stunned: "Nobody does that. Nobody says, I quit as a player. I want to be an advance scout" (p. 55).

Beane was a much better baseball analyst than baseball player, and he quickly moved up the Oakland club's hierarchy. He became interested in a simple question: What is the most efficient way to spend money on baseball players? The origins of Beane's iconoclastic answers can be found in the writings of Bill James, a once-obscure but now-legendary baseball writer/statistician.⁵ While working as a night watchman for a pork-and-beans factory, James decided that he wanted to write about baseball in a way that would illuminate what really happened and why. In his view, conventional statistics were insufficiently helpful and sometimes downright misleading. Consider the area of defensive play. When a player mishandles a ball or makes a bad throw, he can be assigned an "error." A player who accumulates a lot of errors seems like a bad fielder, whereas one with few errors seems really good. The problem is that a player may accumulate errors in part because he is unusually good at getting to the ball. If you do not get to the ball, you are most unlikely to get an error (according to the chapter on scoring in *The Book*). So errors are a crude measure of fielding ability.

Or consider walks. Since the late nineteenth century, walks have been treated, in official statistics, as neutral — neither good nor bad. According to a nineteenth-century expert whose advice is followed to the present day, "There is but one true criterion of skill at the bat, and that is the number of times bases are made on clean hits" (p. 70). Of

5. For a good introduction to Bill James, readers might try BILL JAMES, *THE NEW BILL JAMES HISTORICAL BASEBALL ABSTRACT* (2001).

course, many people realized that a walk is a positive event for the hitting team and a negative event for the team in the field, but this commonsense notion was not incorporated into baseball's most common measure of batting skill, the batting average, which leaves walks out. James found this preposterous, and he pushed for the use of the "on-base percentage" as an improvement.

James also criticized the use of "runs batted in" as a standard measure of a hitter's value. James pointed out that some players are in a position to bat in a lot of runs because they either are lucky or play on good teams. Other players bat in fewer runs, but only because they do not have the opportunities of their apparent superiors: "There is a huge element of luck in even having the opportunity, and what wasn't luck was, partly, the achievement of others" (p. 71).

Eventually James punctured countless myths about what was important to winning in baseball. And he had a positive agenda, too. He devised a formula to measure "runs created" — a formula that predicted, from just a few aspects of a player's performance, how many runs he would produce for an average team (p. 77). James's formula had explosive implications. It suggested that professional baseball experts, those who ran the teams, were placing far too much emphasis on batting averages and stolen bases, and far too little on walks and extra base hits. Statistical measures were moving actual decisions in a way that resulted in inferior team performance. After a slow start, James was widely read; his books became bestsellers, and he became a kind of cult figure among certain baseball fans. But for many years, baseball's experts and executives treated James's work as irrelevant. He had no effect on what they did. And with a few exceptions, the tried-but-not-so-true baseball statistics such as batting average and RBIs remained the only ones reported.

So Billy Beane, the "can't-miss" prospect who missed, became an avid Bill James reader. As Lewis writes,

James had something to say specifically to Billy: you were on the receiving end of a false idea of what makes a successful baseball player. James also had something general to say to Billy, or any other general manager of a baseball team who had the guts, or the need, to listen: if you challenge the conventional wisdom, you will find ways to do things much better than they are currently done. (pp. 97-98)

Wanting to ensure that statistical analyses were done, and done right, Beane hired DePodesta to study player performances with the aid of a computer. Some of Lewis's most hilarious passages illustrate the debate between old baseball wisdom and statistical knowledge:

"The guy's an athlete, Billy," the old scout says. "There's a lot of upside there."

"He can't hit," says Billy.

"He's not that bad a hitter," says the old scout.

“Yeah, what happens when he doesn’t know a fastball is coming?” says Billy.

“He’s a tools guy,” says the old scout . . .

“But can he hit?” asks Billy.

“He can hit,” says the old scout, unconvincingly.

Paul reads the player’s college batting statistics. They contain a conspicuous lack of extra base hits and walks.

“My only question is,” says Billy, “if he’s that good a hitter why doesn’t he hit better?” . . .

Over and over the old scouts will say, “The guy has a great body,” or, “This guy may be the best body in the draft.” And every time they do, Billy will say, “We’re not selling jeans here,” and deposit yet another highly touted player, beloved by the scouts, onto his shit list. (p. 30)

Beane ends up seeking, and getting, young players that other teams simply do not want. These were people largely ignored by the professional scouts, typically because they had something wrong with them — they did not match up with the scouts’ mental prototype of a successful ballplayer. Consider, for example, the star pitcher Barry Zito, passed over by other teams because of his slow fastball, but chosen by Beane on the basis of statistics: “Billy made us take Zito,” the oldest of the scouts confessed (p. 39). While scouts on other teams were still searching for young players who looked like Beane did in high school, DePodesta was busy surfing the Internet. “The evaluation of young baseball players,” Lewis writes, “had been taken out of the hands of old baseball men and placed in the hands of people who had what Billy valued most (and what Billy didn’t have), a degree in something other than baseball” (p. 41).

The statistical method was the only way for Beane to solve a serious problem: obtaining first-rate talent without a lot of money. After all, the New York Yankees had three times the budget of the Oakland Athletics. And if Beane did find good players, and they performed well, they would be bid away by richer teams. Owing to his low payroll, he would be forced to replace his own greatest successes. In 2001, Oakland won 102 games in the regular season, the second-highest total in baseball. They lost three players widely regarded as their best, and they were expected by many to have a catastrophic fall. Instead they used statistical methods to try to replace the lost players with new ones who would provide statistical equivalents — and they ended up winning 103 games, the most in baseball that year. Their payroll for the year was \$34 million, less than half that of their division rivals the Seattle Mariners. In Lewis’s account, Beane was able to succeed because “the market for baseball players was so inefficient,

and the general grasp of sound baseball strategy so weak, that superior management could still run circles around taller piles of cash” (p. 122).

II. RATIONALITY AND COMPETITION

Lewis has a wonderful story to tell, and he tells it wonderfully. His account of Beane’s success is punctuated by descriptions of numerous colorful characters, among them a promising fat catcher dumbfounded by Beane’s interest in him, an excellent pitcher whose fastball is extremely slow, and of course Beane himself. Lewis also raises some serious puzzles that he does not resolve, and his account has some large and perhaps profound implications that he does not much explore. Baseball is an extremely competitive market, and teams have a great deal to gain from winning more games. Beane’s success, and the blunders of many of his colleagues, suggests the persistence of boundedly rational behavior in a domain in which market pressures might well have been expected to eliminate them.

A. *Why Do Executives Make Mistakes?*

Why do professional baseball executives, many of whom have spent their lives in the game, make so many colossal mistakes? They are paid well, and they are specialists. They have every incentive to evaluate talent correctly. So why do they blunder? In an intriguing passage, Lewis offers three clues. First, those who played the game seem to overgeneralize from personal experience: “People always thought their own experience was typical when it wasn’t.” Second, the professionals were unduly affected by how a player had performed most recently, even though recent performance is not always a good guide. Third, people were biased by what they saw, or thought they saw, with their own eyes. This is a real problem, because the human mind plays tricks, and because there is “a lot you couldn’t see when you watched a baseball game” (p. 18).

Several findings in cognitive psychology help explain Lewis’s clues. In making judgments, people tend to use the “availability heuristic.”⁶ As Daniel Kahneman and Amos Tversky have shown, people often assess the probability of an event by asking whether relevant examples are cognitively “available” — the sense that they come readily to mind.⁷ Thus, people are likely to think that more words, on a random page, end with the letters “ing” than have “n” as their next to last letter — even though a moment’s reflection will show that this could

6. See Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, in *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 11-14 (Daniel Kahneman et al. eds., 1982).

7. See *id.*

not possibly be the case.⁸ It follows, from the availability heuristic, that recent events will have more salience than less recent ones — one of Beane's explicit claims. Now, it is not dumb to use the availability heuristic. Heuristics are used because they are, on average, helpful. Yet reliable statistical evidence will outperform the availability heuristic every time. In using data rather than professional intuitions, Beane confirmed this point.

Consider next Lewis's suggestion that people believed, falsely, that their own experience "was typical." This is also a prominent finding in studies of human judgment: People systematically exaggerate the extent to which others are like they are.⁹ To illustrate this in class one year, one of us asked students whether they owned a cell phone, and also to guess what percentage of their classmates owned cell phones. As expected, cell-phone owners guessed a higher percentage of cell-phone ownership than nonowners. This bias in judgment is undoubtedly connected with the availability heuristic, because one's own behavior is always available and salient (and, like the availability heuristic, often useful). But the belief in the typicality of one's own behavior is not simply about availability; it is a distinctive kind of bias.

Above all, perhaps, Lewis's tale demonstrates the use (and vices) of the representativeness heuristic.¹⁰ Baseball scouts use that heuristic by relying on a prototypical picture of a promising prospect — one that has a certain physique, a set of "tools," even the Good Face — in a way that leads to systematic errors. Perhaps the most famous example of the representativeness heuristic involves the likely career of a hypothetical woman named Linda, described as follows:¹¹ "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and also participated in antinuclear demonstrations." Researchers asked people to rank, in order of probability, eight possible futures for Linda. Six of these futures were fillers (e.g., psychiatric social worker, elementary school teacher); the two crucial ones were "bank teller" and "bank teller and active in the feminist movement." Most people said that Linda was less likely to be a bank teller than to be a bank teller and active in the feminist movement. This is an obvious logical mistake, a conjunction error, in

8. See Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, in *HEURISTICS AND BIASES*, *supra* note 4, at 19, 21.

9. See Lee Ross et al., *The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes*, *J. PERSONALITY & SOC. PSYCHOL.* 279-301 (1977).

10. See Daniel Kahneman & Shane Frederick, *Representativeness Revisited*, in *CHOICES, VALUES, AND FRAMES* (2003).

11. See *id.*; Barbara Mellers et al., *Do Frequency Representations Eliminate Conjunction Effects?*, *12 PSYCHOL. SCI.* 269 (2001).

which characteristics A and B are thought to be more likely than characteristic A alone. The error stems from the representativeness heuristic: Linda's description seems to match "bank teller and active in the feminist movement" far better than "bank teller." In an illuminating reflection on the example, one that bears on baseball and many personnel decisions, Stephen Jay Gould observed that "I know [the right answer], yet a little homunculus in my head continues to jump up and down, shouting at me — 'but she can't just be a bank teller; read the description.'"¹²

In Lewis's account, baseball scouts rely on a prototype of a good player, and they ask, too often, whether a prospect fits the prototype, instead of asking whether he is likely to produce runs (pp. 24-34). Beane himself was a prototype of the prototype: "Great body, plus wheels, plus arm, good instincts, and the ability to hit the ball over light towers" (p. 41). The problem was that Beane, like many who fit the mold, couldn't succeed at the major-league level. Instead of emphasizing what players look like, Beane endorsed "performance scouting" (p. 38), based not on "what you can see" the player "doing in your mind's eye" (p. 38), but on the player's statistics.

Beane's preference for performance scouting, as opposed to prototype-based scouting, fits quite neatly with a recent discussion of the representativeness heuristic by Kahneman and Frederick, who emphasize the relevance of dual-process theories of cognition.¹³ Those theories distinguish between two families of cognitive operations, sometimes labeled System I and System II. System I is intuitive; it is rapid, automatic, and effortless (and it features Gould's homunculus). System II, by contrast, is reflective; it is slower, self-aware, calculative, and deductive. Kahneman and Frederick suggest that System I proposes quick answers to problems of judgment, and that System II operates as a monitor, confirming or overriding those judgments. Old-fashioned scouts, stressing the "mind's eye," use System I. Beane insists on a System II override, based on one thing: "Paul's computer" (p. 37).

B. *Why Not Earlier Beanes?*

But the evident usefulness of Paul's computer raises an even larger puzzle. Why didn't someone like Beane come along sooner? Why didn't baseball executives start using statistics a decade, or two decades, or three decades, earlier? Why have falsehoods and mistakes persisted? The economic stakes are extremely high, after all, and if

12. STEPHEN JAY GOULD, *BULLY FOR BRONTOSAURUS: REFLECTIONS IN NATURAL HISTORY* 469 (1991).

13. See Kahneman & Frederick, *supra* note 10. For an overview, see *DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY* (Shelly Chaiken & Yaacov Trope eds., 1999).

Lewis is correct, the management of most baseball teams could have saved many millions of dollars simply by making more rational personnel decisions. Here, then, is a situation in which boundedly rational behavior persisted even though the economic stakes were significant. Market pressures did not drive out that behavior. Nor was the important information hard to find. James's arguments have been around for nearly two decades. In a market as competitive as major-league baseball, surely the information should have been used, and fast. What went wrong? The question has general importance: Will markets usually ensure rational behavior, or can a form of quasi-rationality¹⁴ last in some situations even with strong economic incentives?

The initial problem is not that baseball professionals are stupid; it is that they are human. Like most people, including experts, they tend to rely on simple rules of thumb, on traditions, on habits, on what other experts seem to believe. Even when the stakes are high, rational behavior does not always emerge. Researchers who have studied expert decisionmaking have often replicated the same biases that psychologists had earlier illustrated with laboratory studies of college students. For example, physicians have been shown to display overconfidence, and security analysts issue biased forecasts.¹⁵ It takes time and effort to switch from simple intuitions to careful assessments of evidence. This point helps to explain why baseball owners have been slow to copy Beane's approach. But at least they are starting. The Toronto Blue Jays and the Boston Red Sox have hired general managers who follow Beane's general approach. And, in one of the longest-overdue moves in baseball, the general manager of the Red Sox, appointed when just twenty-eight years old, hired James as a consultant.

C. *Lesson 1: The Problem of Bad Statistics*

Many of the lessons in this book apply in domains far from baseball. One involves the harmful repercussions of using bad statistics or performance measures. We have already seen that batting average and RBIs can lead teams in unfortunate directions. Consider also the case of the "save" statistic. A save is awarded to a relief pitcher who comes in near the end of a close game with his team ahead and "saves" the win for his team. Most thoughtful observers realized long ago that this is a really dumb statistic. Why should

14. See RICHARD H. THALER, *QUASI-RATIONAL ECONOMICS* (1993).

15. See Werner DeBondt & Richard Thaler, *Do Analysts Overreact*, in *HEURISTICS AND BIASES*, *supra* note 4, at 678; Derek J. Koehler et al., *The Calibration of Expert Judgment*, in *HEURISTICS AND BIASES*, *supra* note 4, at 686.

pitching the last inning of a game have any special significance? A pitcher who comes in to pitch the sixth inning of a tied game and pitches three scoreless innings has done something much more important than one who just pitches the ninth inning, protecting a three-run lead. Now, a dumb statistic could be harmless, but in this case, as is often true, the very fact that the number is collected and tabulated ends up influencing behavior. The existence of the “save” statistic (and the muddled thinking that goes along with it, namely, the idea that runs scored at the end of the game count more) seems to have altered the way teams use their relief pitchers. In the past twenty years, most teams have settled on using their best relief pitcher in a specific role that has acquired a name: the closer. The closer typically comes in just to pitch the ninth inning in close games with his team ahead. Though this strategy is nearly universally used, it is clearly stupid. A team that is ahead by three runs going into the ninth inning has a 97 percent chance of winning the game (a percentage that has not changed since the advent of the closer).¹⁶ It makes more sense to use your best relief pitcher in more crucial situations, such as a tied game. Saving a three-run lead is much easier than protecting a tie, since you can give up two runs and still win the game.

What should a team do if it figures this out? One strategy would be to take your best reliever and use him more strategically, sometimes in the seventh inning of a close game with the best players of the other team due to bat, other times with a one-run lead in the ninth. This strategy would win more games, but it would not create many “saves” for the ace. Beane did this, but he also did something deviously clever: he created closers in order to sell them. As Lewis puts it:

Established closers were systematically overpriced, in large part because of the statistic by which closers were judged in the marketplace: ‘saves.’ The very word made the guy who achieved them seem vitally important. . . . You could take a slightly above average pitcher and drop him into the closer’s role, let him accumulate some gaudy number of saves, and then sell him off. You could, in essence, buy a stock, pump it up with false publicity, and sell it off for much more than you paid for it. (p. 125)

It is interesting to speculate about the possibility that similar phenomena occur in other domains. Consider the role of citation counts in the academic labor market. While the hard copy edition of the Social Science Citation Index has existed for years, citations have become more salient in recent years because of the advent of web-based citation search engines, including LEXIS and WESTLAW. We speculate, partly on the basis of the experience of one of us (Sunstein), that legal scholars are now reacting to this rise in the prominence of

16. Thanks to Bill James for providing these facts.

citation counts. In law, another example is placements in top law reviews. This is a flawed proxy for quality, to say the least, but lateral appointments are often affected by them.

D. *The Difficulty of Organizational Change*

Beane's use of relief pitchers, and his success in general, bear on a more general problem. To what extent are the top managers in an organization — here the owners and the general managers — able to push a rational but radical change down through an organization? Beane has an owner who is sympathetic to his philosophy, but if he wants to try something new, such as using the relief-ace flexibly, he has to convince the field manager to implement his strategy. He also has to avoid a rebellion by the players. To get his manager to use the player Beane thinks is his most effective pitcher in tight situations, Beane tells the manager to think of him as the closer before the ninth inning. A relief ace would likely complain about being used in the optimal manner, because he would accumulate fewer saves and thus would be worth less on the open market. Similarly, suppose a player takes more pitches in an attempt to draw more walks and as a result increases his on-base percentage at the cost of lowering his batting average. His team might like this trade-off, but if it lowered his value to other teams, then the player might suffer in the free-agent market.

Finally, there is the impact of the media and the fans. When James was hired by the Red Sox last winter, there was great anticipation about how the team would deal with relief pitchers. The rational strategy of using pitchers to maximize the chance of winning the game was quickly dubbed "bullpen by committee" by Boston sportswriters, who knew (from *The Book*, naturally) that this was a terrible idea. When Red Sox relievers lost the opening-day game to the woeful Tampa Bay Devil Rays and suffered through an awful opening month, James was viewed as the villain. Of course, James does not advocate bad pitching, and, presumably with his help, the team has acquired three new relief pitchers. But, interestingly, they designated one as their closer, perhaps deciding to let this particular battle wait for another day.

The difficulty of achieving sensible change in organizations is hardly special to baseball. If a new CEO comes in and wants to change how things are done, a selling job must be performed all the way down the organizational ladder. Every institution has organizational norms, ways of doing things (the in-house version of *The Book*) that are hard to overcome. The new guy is told, well, we don't do things that way here. If the CEO forces his views down the organizational ladder and his methods are unconventional, then lower-level workers may face the same dilemma as the player trying to get more walks. Do they maximize their value to the current CEO or to the outside world?

E. *Bad Equilibrium*

As we have emphasized, Lewis's account bears more generally on the performance of markets. He often refers to the inefficiency of the market for player talent, as the example of the closer illustrates. If it were not for this inefficiency, Beane would not be able to procure a winning team while spending one-third as much as the Yankees and one-half as much as his division rivals. How can this market be so inefficient? Simply put, the baseball owners seem to have evolved into a "bad equilibrium." Cascade effects might well be involved: People rely on the apparent judgments of others, believing that they carry information, even if those other judgments are themselves a product, much of the time, of the apparent judgments of others.¹⁷ Teams have thought the same way for years; *The Book* has never been revised; and so there are massive inefficiencies that have been relatively easy for Beane to exploit. You do not need a Harvard economics graduate to realize that "a walk is as good as a hit," an expression that was around well before James started writing. Beane's key insight was that this market was inefficient and could therefore be exploited. Now that Beane has succeeded, it is likely that the market for baseball-player talent will get more competitive. Still, it is embarrassing that it took so long for this to happen, especially for those who think that competitive markets always lead to a rational allocation of resources.

What does this tell us about other markets? Lewis poses this question: "If professional baseball players could be over- or undervalued, who couldn't? Bad as they may have been, the statistics used to evaluate baseball players were probably far more accurate than anything used to measure the value of people who didn't play baseball for a living" (p. 72). Right! On the basis of first principles, the market for baseball players should be one of the most efficient labor markets on earth. It is hard to think of any high-paid profession in which performance is measured so precisely — and is publicly available to every other potential employer.

Compare the market for baseball players with the market for corporate executives. A company looking for a new director of human-resource management would be hard-pressed to get any objective data on the past performance of job candidates. Instead, such a company would be forced to make choices based on interviews with the candidates — a process that is even less accurate than the one the old scouts use to size up a high-school player. Interviews are

17. For a helpful overview, see Sushil Bikchandani et al., *Learning From the Behavior of Others: Conformity, Fads, and Informational Cascades*, 12 J. ECON. PERSP. 151 (1998).

notoriously imperfect predictors of future job performance.¹⁸ In many contexts their predictive value is essentially zero.

The biases caused by labor markets using subjective evaluations instead of objective measures of output are potentially huge. We can glimpse the scope of the problem by studying one subjective variable that most people would say should be irrelevant to accurate job evaluations in most positions: physical beauty. Away from places such as the back lots of Hollywood and the runways of Milan, most of us who do not look like Ben Affleck or Jennifer Lopez would like to think that successful people get where they do because of their accomplishments, not their attractiveness. Alas, it ain't so. To take one simple example, height matters: in the United States, the taller presidential candidate has usually won. And beauty matters in domains where we might not expect it to matter: law and business.¹⁹ In both fields, and for both sexes, career success and earnings are correlated with good looks.

And consider another problem. We have said that interviews, especially unstructured interviews, are not very useful predictors of job performance — except that once you have conducted an interview, it is almost impossible to avoid the conclusion that you have learned a lot. But the facts are clear: unstructured interviews are nearly useless at predicting anything except whether the interviewer will subsequently like the interviewee.²⁰ So a rational Beane-like strategy would be to interview only employees who will work directly with you and otherwise make all decisions based on objective criteria and statistical models. Rather than conduct interviews, firms should give tests and pay more attention to grades in school (a good predictor of both intelligence and diligence, admirable traits in nearly any job). Numerous law firms and law faculties place great weight on relatively unstructured interviews, thinking that as a result of such interviews, they learn a great deal about likely performance. But this is probably a

18. Some authors find that interviews are nearly useless, see, for example, John E. Hunter & Ronda F. Hunter, *Validity and Utility of Alternative Predictors of Job Performance*, 96 PSYCHOL. BULL. 72, 90 (1984). Others find that if the interview is “structured” it can provide some limited value. See Michael McDaniel et al., *The Validity of Employment Interviews: A Comprehensive Review and Meta-Analysis*, 79 J. APPLIED PSYCHOL. 599 (1994).

19. Daniel Hamermesh has produced a series of papers on this topic. See Daniel Hamermesh, *Beauty Papers*, <http://www.eco.utexas.edu/faculty/Hamermesh/Beautystuff.htm> (last visited June 1, 2004). For published versions, see Daniel Hamermesh et al., *Dress for Success: Does Primping Pay?*, 9 LAB. ECON. 361 (2002); Daniel Hamermesh & Jeff E. Biddle, *Beauty, Productivity and Discrimination: Lawyers' Looks and Lucre*, 16 J. LAB. ECON. 172 (1998); Daniel Hamermesh & Jeff E. Biddle, *Beauty and the Labor Market*, 84 AM. ECON. REV. 1174 (1994).

20. See McDaniel et al., *supra* note 18.

blunder. In fact the representativeness heuristic undoubtedly affects evaluations, leading firms and law schools to rely, much of the time, on a prototype of the successful lawyer or teacher. Their own homunculus boldly announces that the interviewee is or is not “right” for the job. Most of the time, they would do better to ignore the homunculus and to place greater weight on actual performance measures.

Emboldened by Lewis and Beane, both of us have tried to encourage our respective divisions at the University of Chicago to reduce the time and weight devoted to the interview process and to rely instead on other measures. To date, the effort has been repudiated by numerous highly articulate homunculi (otherwise known as colleagues and admissions officers). But in many contexts, more objective procedures would not only improve accuracy; they would also reduce discrimination, inadvertent or otherwise, based on factors such as race, gender, or beauty.²¹ Symphony orchestras have found that if they conduct auditions with the candidates hidden behind a screen, more women are hired.²² And when employees are hired, firms should seek objective measures of performance, which distinguish luck from skill.

CONCLUSION

Some enlightened organizations, like the Boston Red Sox and the Toronto Blue Jays, are undoubtedly taking steps along these lines. Many more should be doing so. One of the noteworthy features of Lewis’s book is his clear demonstration of the pervasiveness of boundedly rational behavior even when the stakes are high and even when market pressures might be expected to supply a corrective. We suspect that countless areas of enterprise, both private and governmental, would benefit from their own Billy Beanes and Paul DePodestas, challenging widespread intuitions, or what “everyone knows,” with statistical information about what works and what does not, and with performance measures that more accurately reflect the true contribution to organizational success. Baseball is not the only realm for which *The Book* is in need of revision.

21. Cf. Marianne Bertrand & Sendhil Mullainathan, *Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*, AM. ECON. REV. (forthcoming).

22. Claudia Goldin & Cecilia Rouse, *Orchestrating Impartiality: The Impact of “Blind” Auditions on Female Musicians*, 90 AM. ECON. REV. 715 (2000).