son: hitting might have improved just as much, if not more. Why should hitting be uniquely exempt from a general betterment in all other aspects of play? Isn't it more reasonable to assume that batting has improved in concert with other factors of baseball? I shall show that general improvement in hitting has not only kept pace with betterment in other aspects of play, but that baseball has constantly fiddled with its rules to assure that major factors remain in balance. The extinction of 0.400 hitting must therefore arise from other causes.

·8·

A Plausibility Argument for General Improvement

However tempted we may be to indulge in fanciful reveries about dedication during "the good old days," the accepted notion that decline in batting skills caused the extinction of 0.400 hitting just doesn't make sense when we consider general patterns of social and sports history during the twentieth century. This context, on the contrary, almost guarantees that hitting has improved along with almost anything else we can measure at the apogee of human achievement. Consider just three of many arguments that virtually cement the case, even before we examine a single baseball statistic.

1. LARGER POOLS AND BETTER TRAINING. In 1900, 76 million people inhabited the United States, and only white men could play major league baseball. Our population has since ballooned to 249 million people (1990

census), and men of all colors and nations are welcome. Training and coaching were absent to slapdash in the past, but represent a massive industry today. Players follow rigorous and carefully calculated programs for working out (even, if not especially, during the off season, when their predecessors mostly drank beer and gained weight); they no longer risk careers and records by playing hurt. (Joe DiMaggio once told me that he was batting 0.413 with two weeks to go in the 1939 season. He caught a serious cold, which clouded his left [leading] eye, and he could not adequately see incoming pitches. The Yanks had already clinched the pennant. Any modern counterpart would sit on the bench and preserve his record; DiMaggio played to the last game and fell to 0.381, his highest seasonal average, but below the grand plateau.) No one—neither the players nor the owners—can afford to take risks and fool around today, not with star salaries in the multiple millions for careers that last but a few years at peak value. What possible argument could convince us that a smaller and more restricted pool of indifferently trained men might supply better hitters than our modern massive industry with its maximal monetary rewards? I'll bet on the larger pool, recruitment of men of all races, and better, more careful training any day.

2. SIZE. I don't want to fall into the silly mythology of "bigger is better" (okay for a few things, like brains in the evolution of most mammalian lineages, but irrelevant for many items, like penises and automobiles). Still, *ceteris paribus* as the Romans said (all other things being equal), larger people tend to be stronger (and I say this as a short man who loved to watch Phil Rizzuto and Fred Patek). If height and weight of ballplayers have augmented through time, then (however roughly) bodily prowess should be increasing.

Pete Palmer, sabermetrician extraordinaire and editor, along with John Thorn, of *Total Baseball,* the best (and fattest) general reference book of baseball stats, sent me his chart (reproduced here as Table 1) of mean heights and weights for pitchers and batters averaged by decades. Note the remarkably steady increase through time. I cannot believe that the larger players of today are worse than their smaller counterparts of decades past.

TABLE 1

DECADAL AVERAGES FOR HEIGHTS AND WEIGHTS OF MAJOR LEAGUE BASEBALL PLAYERS

	BATTERS		PITCHERS	
	Height (inches)	Weight (pounds)	Height	Weight
1870s	69.1	163.7	69.1	161.1
1880s	69.6	171.6	70.2	172.7
1890s	69.8	172.1	70.6	174.1
1900s	69.9	172.6	71.5	180.7
1910s	70.3	170.5	72.1	180.7
1920s	70.4	171.2	72.0	179.8
1930s	71.1	176.8	72.6	184.8
1940s	71.4	180.3	73.0	186.5
1950s	72.0	183.0	73.1	186.1
1960s	72.2	182.7	73.6	189.3
1970s	72.3	182.3	74.1	191.0
1980s	72.5	182.9	74.5	192.2

3. RECORDS IN OTHER SPORTS. All major baseball records are relative—that is, they assess performance against other players in an adversarial role—not absolute as measured by personal achievement, and counted, weighed, or timed by a stopwatch. A 0.400 batting average records degree of relative success against pitchers, whereas a four-minute mile, a nineteen-foot pole vault, or a 250-pound lift is unvarnished *you* against an unchanging outer world.

Improvements in relative records are ambiguous in permitting several possible (and some diametrically opposed) interpretations: rising batting averages might mean that hitting has improved, but the same increase might also signify that batting has gotten worse while pitching has deteriorated even more sharply (leading to relative advantages for hitters as their absolute skills eroded).

Absolute records, however, have clearer meaning. If leading sprinters are running quicker and vaulters jumping higher . . . well, then they are performing their art better. What else can we say? The breaking of records doesn't tell us why modern athletes are doing better—and a range of diverse reasons might be cited, from better training, better understanding of human physiology, new techniques (the Fosbury flop), to new equipment (fiberglass poles and the immediate, dramatic rise in record heights for pole vaulting)—but I don't think that we can deny the fact of improvement.

Therefore, since the relative records of baseball must be ambiguous in their causes, we should study the absolute records of related sports. If most absolute records have been improving, then shouldn't we assume that athletic prowess has risen in baseball as well? Wouldn't we be denying a general pattern and creating an implausible, ad hoc theory if we attributed the extinction of 0.400 hitting to a decline in batting skills? Shouldn't we be searching for a theory that can interpret the death of 0.400 hitting as a consequence of generally superior athleticism—thus making this most interesting and widely discussed trend in the history of baseball statistics consistent with the pattern and history of almost every other sport?

I don't want to worry a well-understood subject to death, and I don't want to bore you with endless documentation of well-known phenomena. Surely all sports fans recognize the pervasive pattern of improvement in absolute records through time. The first modern Olympic marathon champ, Spiridon Loues, took almost three hours in 1896; more recent winners are nearly down to two. The allure of the four-minute mile challenged runners for decades, while Paavo Nurmi's enticing 4.01 held from 1941 until Roger Bannister's great moment on May 6, 1954. Now, most of the best runners routinely break four minutes nearly every time. By 1972, for the 100-meter freestyle, and by 1964 for the 400 meters, the best women swimmers eclipsed the Olympic records of the 1920s and 1930s, set by the two great Tarzans (both played the role in movies) Buster Crabbe and Johnny Weismuller. I will let one chart stand for the generality, based on

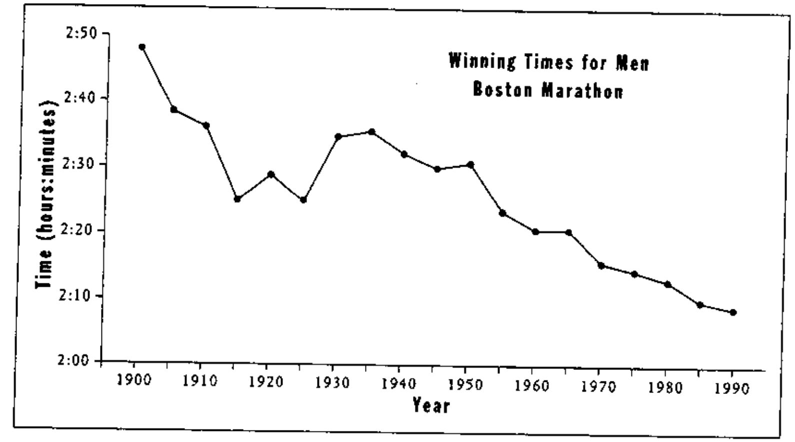

FIGURE 12
The steadily decreasing record time for men in the Boston Marathon. Dots are five-year averages (my calculation).

data closest to hand for the most famous of local races at my workplace—the Boston Marathon (see Figure 12). The general pattern is clear, and the few anomalies record changes in distance (the "standard" 26 miles 385 yards has prevailed in most years, but early winners, from 1897 to 1923, ran only 24 miles 1,232 yards for their longer times; with a rise to 26 miles 209 yards from 1924 to 1926; the standard distance from 1927 to 1952; and a shortened 25 miles 958 yards from 1953 to 1956, until reestablishment of the standard distance in 1957).

For almost every sport, the improvement in absolute records follows a definite pattern with presumed causes central to my developing argument about 0.400 hitting. Improvement does not follow a linear path of constant rate. Rather, times and records fall more rapidly early in the sequence and then slow markedly, sometimes reaching a plateau of no further advance (or of minutest measurable increments from old records). In other words, athletes eventually encounter some kind of barrier to future progress, and records stabilize (or at least slow markedly in their frequency and amount of improvement). Statisticians call such a barrier an asymptote; vernacular language might speak of a limit. In the terminology of this book, athletes reach a "right wall" that stymies future improvement.

Since we are considering the world's best performers in these calculations, the probable reasons for such limits or walls should be readily apparent. After all, bodies are physical devices, subject to constraints upon performance set by size, physiology, and the mechanics of muscles and joints. No one will argue that curves of improvement can be extrapolated forever—or else runners would eventually complete the mile in nothing flat (and, finally, in negative time), and pole vaulters would truly match a gentleman of legend and leap tall buildings in a single bound.

We can best test the proposition that physical limits (or right walls) cause the slowing and plateauing of improvements by comparing curves for athletes operating near the extremes in human capacity with performers who probably retain much room for further advance. What conditions might place people far from the right wall, and therefore endow them with great scope for improvement? Consider some potential examples: new sports where athletes have not yet figured out optimal procedures; new categories of people recently admitted to old sports; records for amateur play. As an example, the Boston Marathon was opened to women only in 1972. Note how much more rapidly women have improved than men from their beginning to the present (Figure 13).

We may generalize this principle by setting up a hierarchy of decreasing improvement (also a ranking of increasing worth in the value systems of some rather old-fashioned and well-heeled folks): women, men, and horses. Winning times for major horse races have improved, but ever so slightly over long intervals. For example, between 1840 and 1980, thoroughbred horses in the three great English races of St. Leger, Oaks, and the Derby have shaved twelve, twenty, and eighteen seconds off record times, for a minuscule gain of 0.4 to 0.8 percent per generation (Eckhardt et al., 1988). These gains are tiny even when compared with the other great arena of breeding in domesticated animals: improvement of livestock, where gains of 1 to 3 percent per year are often achieved for features of economic importance.

This limited improvement makes perfect and predictable sense. Thoroughbreds have been rigorously raised from a limited stock for more than two hundred years. Stakes could not be higher, as the slightest improvement may be worth millions. More effort has gone into betterment of this breed than into almost any other biological endeavor of economic impor-

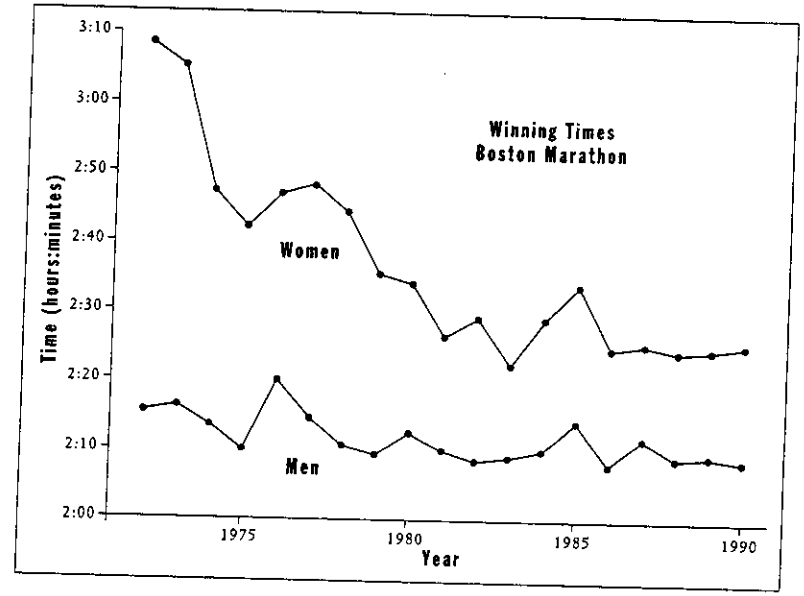

FIGURE 13
Women's records decline precipitously while men's records remain relatively stable between 1970 and 1980 in the Boston Marathon. Since 1980 both records have changed little. Dots are actual winning times for each year.

tance to humans. We might therefore suppose that the best thoroughbreds have long resided at the genetic right wall for the breed, and that future improvements will be negligible to slow. But since (thank God) we have not yet reached brave new world, we do not breed humans for optimized physical performance, and records for people should therefore show more flexibility—for we have no purposeful purebreds at our right walls.

In most popular and established men's events, we note the pattern of rapid initial improvement followed by flattening of the curve.[4] Exceptions may be found in such events as the marathon, where length and complexity provide great "play" for experimenting with new strategies, and where re-

4. All bets are off when fundamentally new equipment or procedures enter the field, as in the fiberglass pole, or (God forbid) the aluminum bat, which (we may hope and pray) will never darken the doorstep of major league dugouts. Such innovations will produce sudden blips in curves of improvement. In fact, such innovations are usually better treated statistically as the beginning points of new curves.

cent surges in popularity have brought large increases in prestige and participation. (Note that the curve of improvement for the Boston Marathon has remained virtually linear for men, and did not slow before 1990—though the pattern may now be shifting into the usual mode as the world's best runners now compete and improvements begin to abate.)

Many commentators have noted that most women's records are both falling faster than men's for the same event, and are not yet flattening, but maintaining a linear pace of improvement. Interestingly (see Whipp and Ward, 1992), most men's running events (200 to 10,000 meters) have improved in the same range of rates regardless of the event's total distance—5.69 to 7.57 meters per minute improvement per decade. (Improvement in the marathon has been greater, at 9.18 meters per minute per decade, thus supporting my claim that this event remains "immature" and still in the category of potentially linear improvement—that is, not near the right wall.) But for women in the same events, rates of improvement run from 14.04 to 17.86 minutes per meter per decade (with a whopping 37.75 meters for the marathon).

These findings have led to all manner of speculation, some rather silly. For example, Whipp and Ward (1992) just extrapolate their curves and then defend the conclusion that women will eventually outrun men in most events, and rather soon for some. (The extrapolated curves for the marathon, for example, cross in 1998 when women should beat men by this argument.)

But extrapolation is a dangerous, generally invalid, and often foolish game. After all, as I said before, extrapolate the linear curve far enough and all distances will be run in zero and then in negative time. (False extrapolation also produces the irresponsible figures often cited for growth of human populations—in a few centuries, for example, humans will form a solid mass equal to the volume of the earth and no escape into outer space will be possible because the rate of increase will cause the diameter of this human sphere to grow at greater than the speed of light, which, as Einstein taught us, sets an upper bound upon rapidity of motion.) Clearly we will never run in negative time, nor will our sphere of solid humanity expand at light speed. Limits or right walls will be reached, and rates of increase will first slow and eventually stop.

Women may outclass men in certain events like ultra-long-distance

swimming, where buoyancy and fat distribution favor women's physiques and endurances over men's (women already hold the absolute record for the English Channel and Catalina Island swims). The marathon may also be a possibility. But I doubt that women will ever capture either the 100-meter dash or the heavyweight lifting records. (Many women will always beat most men in any particular event—most women can beat me in virtually anything physical. But remember that we are talking of world records among the very best performers—and here the biomechanics of different construction will play a determining role.)

The basic reason for more rapid gains (and less curve flattening) in women's events seems clear. Sexism is the culprit, and happy reversals of these older injustices the reward. Most of these events have been opened to women only recently. Women have been brought into the world of professionalism, intense training, and stiff competition only in the last few years. Women, not so long ago (and still now for so many), were socialized to regard athletic performance as debarred to their gender—and many of the great women performers of the past, Babe Didrikson in particular, suffered the onus of wide dismissal as overly masculine. In other words, most women's curves are now near the beginning of the sequence—in the early stages of rapid and linear improvement. These curves will flatten as women reach their own right walls—and only then will we know true equality of opportunity. Until then, the steep and linear improvement curves of women's sports stand as a testimony to our past and present inequities.

·9·

0.400 Hitting Dies as the Right Tail Shrinks

Granting the foregoing argument that hitting must be improving in some absolute sense as the best athletes first rush, and then creep, toward the right wall of biomechanical limits on human performance, only one traditional explanation remains unrefuted for viewing the extinction of 0.400 hitting as the deterioration of something at bat—the possibility that, while hitting has improved, other opposing activities (pitching and fielding) have gotten better, even faster, leading to a *relative* decline in batting performance.

This last holdout of traditionalism fails the simplest and most obvious test of possible validity. If pitching and fielding have slowly won an upper hand over hitting, we should be able to measure this effect as a general decline in batting averages through the twentieth-century history of baseball. If mean batting averages have fallen with time, as pitching and

fielding assert increasing domination, then the best hitters (the 0.400 men of yore) get dragged down along with the masses—that is, if the mean batting average were once 0.280, then a best of over 0.400 makes sense as an upper bound, but if the mean has now fallen to, say, 0.230, then 0.400 might stand too far from this declining mean for even the best to reach.

This entirely sensible explanation fails because, in fact, the mean batting average for everyday players has been rock-stable throughout our century (with interesting exceptions, discussed later, that prove the rule). Table 2 (page 102) presents decadal mean batting averages for all regular players in both leagues during the twentieth century. (I included only those players who averaged more than two at-bats per game for the entire season, thus eliminating weak-hitting pitchers and second-stringers hired for their skills in fielding or running.)[5] The mean batting average began at about 0.260, and has remained there throughout our century. (The sustained, though temporary, rise in the 1920s and 1930s presents a single and sensible exception, for reasons soon to come, but cannot provide an explanation for the subsequent decline of 0.400 hitting for two reasons: first, the greatest age of 0.400 hitting occurred before then, while averages stood at their usual level; second, not a soul hit over 0.400 throughout the 1930s, despite the high league means—I include Bill Terry's 0.401 of 1930 itself in the 1920s calculation.) Thus our paradox only deepens: 0.400 hitting disappeared in the face of preserved constancy in average performance. Why should the best be trimmed, while ordinary Joes continue to perform as ever before? We must conclude that the extinction of 0.400 hitting does not reflect a general decline in batting prowess, either absolute or relative.

When issues reach impasses of this sort, we usually need to find an exit by reformulating the question—and reentering the field by another door. In this case, and following the general theme of my book, I suggest that

5. The recent disparity between the two leagues records, in large part, the introduction of the "designated hitter" to the American League alone—a permanent "pinch hitter" for the pitcher. His substitution for the pitcher doesn't affect the decadal average per se, because I don't include pitchers in this calculation. But the designated hitter still provokes a small general rise in the American League mean by introducing another good bat into the lineup, whereas the National League retains more relatively poor hitters in the bottom part of the order. Nonetheless, I remain an adamant opponent of the DH rule—the one vital subject in our culture that permits no middle ground. You gotta either love it or hate it!

we have been committing the deepest of all errors from the start of our long-standing debate about the decline of 0.400 hitting. We have erred—unconsciously to be sure, for we never considered an alternative—by treating "0.400 hitting" as a discrete and definable "thing," as an entity whose disappearance requires a special explanation. But 0.400 hitting is not an item like "Joe DiMaggio's favorite bat," or even a separately definable class of objects like "improved fielders' gloves of the 1990s." We should take a hint from the guiding theme of this book: the variation of a "full house" or complete system should be treated as the most compelling "basic" reality; averages and extreme values (as abstractions and unrepresentative instances respectively) often provide only partial, if not downright misleading, views of a totality's behavior.

Hitting 0.400 is not an item or entity, a thing in itself. Each regular player compiles a personal batting average, and the totality of these averages may be depicted as a conventional frequency distribution, or bell curve. This distribution includes two tails for worst and best performances—and the tails are intrinsic parts of the full house, not detachable items with their own individuality. (Even if you could rip a tail off, where would you make the break? The tails grade insensibly into the larger center of the distribution.) In this appropriately enlarged perspective, 0.400 hitting is the right tail of the full distribution of batting averages for all players, not in any sense a definable or detachable "thing unto itself." In fact, our propensity for recognizing such a category at all only arises as a psychological outcome of our quirky propensity for dividing smooth continua at numbers that sound "even" or "euphonious"—witness our excitement about the coming millennial transition, though the year 2000 promises no astronomical or cosmic difference from 1999 (see Gould, 1996, essay 2).

When we view 0.400 hitting properly as the right tail of a bell curve for all batting averages, then an entirely new form of explanation becomes possible for the first time. Bell curves can expand or contract as amounts of variation wax or wane. Suppose that a frequency distribution maintains the same mean value, but that variation diminishes symmetrically, with more individual measures near the mean and fewer at both the right and left tails. In that case, 0.400 hitting might then disappear entirely, while the mean batting average remained stable—but the cause would then reside in whatever set of reasons produced the shrinkage of variation around a constant mean. This different geometrical picture for the disappearance of 0.400 hitting does not specify a reason, but the new model forces us to reconsider the entire issue—for I can't think of a reason why a general shrinkage of variation should record the worsening of anything. In fact, the opposite might be true: perhaps a general shrinkage of variation reflects improvement in the state of baseball. At the very least, this reformulation weans us from traditional, locked-in, and unproductive modes of explanation—in this case the "certainty" that extinction of 0.400 hitting must be recording a trend in the degeneration of batting skills. We are now free to consider new explanations: Why should variation be shrinking? Does shrinking record improvement or degeneration (or neither)—and, if so, of what?

Does this alternate explanation work? I have already documented the first part of the claim—preservation of relatively constant mean batting averages through time (see Table 2). But what about the second component? Has variation been shrinking symmetrically about this mean value during the history of twentieth-century baseball? Let me first demonstrate that mean batting averages have been stabilized by an active effort of rulemakers—for natural shrinkage about a purposely fixed point presents an appealing picture that, in my view, establishes our best argument for viewing 0.400 hitting as a predictable and inevitable consequence of general improvement in play.

Figure 14 presents mean batting averages for all regular players in both leagues year by year (the National League began in 1876, the American League in 1901). Note the numerous excursions in both directions, but invariable returns to the general 0.260 level. This average level has been actively maintained by judicious modification of the rules whenever hitting or pitching gained a temporary upper hand and threatened to disrupt the saintly stability of our national pastime. Consider all the major fluctuations:

After beginning at the "proper" balance, averages began to drift down, reaching the 0.240s during the late 1880s and early 1890s. In response, and in the last major change ever introduced in the fundamental structure of baseball (number 1 on Figure 14), the pitching mound retreated to its current distance of sixty feet six inches from the plate during the 1893 season. (The mound had begun at forty-five feet from the plate, with pitchers de-

TABLE 2

League Averages for the Twentieth Century, by Decades

	AMERICAN LEAGUE	NATIONAL LEAGUE
1901–1910	.251	.253
1911–1920	.259	.257
1921–1930	.286	.288
1931–1940	.279	.272
1941–1950	.260	.260
1951–1960	.257	.260
1961–1970	.245	.253
1971–1980	.258	.256
1981–1990	.262	.254

livering the ball underhand, and had migrated steadily back during baseball's early days—the reason for limited utility of nineteenth-century statistics in these calculations.) Unsurprisingly, hitters responded with their best year ever. The mean batting average soared to 0.307 in 1894, and remained high until 1901 (number 2 on Figure 14), when adoption of the foul-strike rule forced a rapid decline to propriety (foul balls had not previously been counted for strikes one and two). Averages remained anomalously low until introduction of the cork-centered ball prompted an abrupt rise in 1911 (number 3 in Figure 14). Pitchers quickly accommodated, and averages returned to their proper 0.260 level as the decade advanced.

The long excursion (number 4 on Figure 14), nearly twenty years of high hitting during the 1920s and 1930s, represents the one extended exception to a pattern of long stability interrupted by quick blips—and the fascinating circumstances and putative reasons have long been debated by all serious fans. In 1919, Babe Ruth hit a wildly unprecedented twenty-

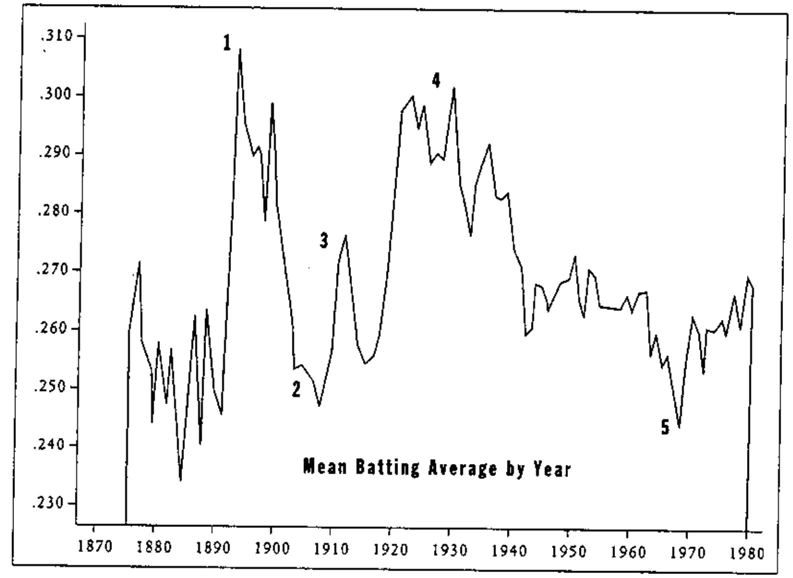

FIGURE 14

The mean batting average for regular major league players has remained quite steady at about 0.260 during the entire history of major league baseball. The few exceptions can be explained and were "corrected" by judicious changes in the rules. Averages rose after the pitching mound was moved back *(1)*; declined after adoption of the foul-strike rule *(2)*; rose again after the introduction of the cork-centered ball *(3)*, and then again during the 1920s and '30s *(4)*. The decline in the 1960s *(5)* was reversed in 1969 by lowering the pitching mound and decreasing the strike zone.

nine homers, more than most entire teams had garnered in full seasons before; then, in 1920, he nearly doubled the total, to fifty-four. At all other times, the moguls of baseball would have reacted strongly to this unseemly change and would, no doubt, have reined in these Ruthian tendencies by some judicious change of rules. But 1920 represented the crux of a unique threat in the history of baseball. Several members of the 1919 Chicago White Sox (the contingent later known as the Black Sox), including the great 0.400 hitter Shoeless Joe Jackson, had accepted money from a gambling ring to throw the World Series of 1919. The resulting revelations almost destroyed professional baseball, and attendance declined precipitously during the 1920 season. The owners (whose pervasive stinginess had set the context that encouraged such admittedly dishonest and indefensible behavior) turned to Ruth as a *deus ex machina.* His new style of play packed in the crowds, and owners, for once, went with the flow and allowed the game to change radically. Scrappy, one-run-at-a-time, any-way-possible, savvy-baserunning, pitcher's baseball became a style of the past (much to Ty Cobb's permanent disgust); big offense and swinging for the fences became *de rigueur.* Mean batting averages rose abruptly and remained high for twenty years, even breaking 0.300 for the second (and only other) time in 1930.

But why were Ruth and other hitters able to perform so differently when circumstances encouraged such a change? Traditional wisdom—it is ever so, as we search for the "technological fix"—attributes this long plateau of exalted batting averages to introduction of a "lively ball." But Bill James, baseball's greatest sabermetrician, argues (in his *Historical Baseball Abstract,* Villard Books, 1986) that no major fiddling with baseballs in 1920 can be proven. James suspects that balls did not change substantially, and that rising batting averages can be attributed to alterations in rules and attitudes that imposed multiple and simultaneous impediments upon pitching, thus upsetting the traditional balance for twenty years. All changes in practice favored hitters. Trick pitches were banned, and hurlers who had previously scuffed, shined, and spat on balls with abandon now had to hide their antics. Umpires began to supply shiny new balls whenever the slightest scuff or spot appeared. Soft, scratched, and darkened balls had previously remained in play as long as possible—fans even threw back foul balls (!), as they do today in Japan, except for home runs. James argues that the immediate replacement of soft and discolored by hard and shiny balls would do as much for improved hitting as any supposedly new construction of a more tightly wound, livelier ball.

In any case, averages returned to their conventional level in the 1940s as the war years siphoned off the best in all categories. Since then, only one interesting excursion has occurred (number 5 in Figure 14)—another fine illustration of the general principle, and recent enough to be well remembered by millions of fans. For reasons never determined, batting averages declined steadily throughout the 1960s, reaching a nadir in the great pitchers' year of 1968, when Carl Yastrzemski won the American League batting title with a minimal 0.301, and Bob Gibson set his astonishing, off-scale record of a 1.12 earned run average (see page 127 for more

on Gibson). So what did the moguls do? They changed the rules, of course—this time by lowering the pitching mound and decreasing the strike zone. In 1969, mean batting averages returned to their usual level—and have remained there ever since.

I do not believe that rulemakers sit down with pencil and paper, trying to divine a change that will bring mean batting averages back to an ideal. Rather, the prevailing powers have a sense of what constitutes proper balance between hitting and pitching, and they jiggle minor factors accordingly (height of mound, size of strike zone, permissible and impermissible alterations of the bat, including pine tar and corking)—in order to assure stability within a system that has not experienced a single change of fundamental rules and standards for more than a century.

But the rulemakers do not (and probably cannot) control amounts of variation around their roughly stabilized mean. I therefore set out to test my hypothesis—based on the alternate construction of reality as the full house of "variation in a system" rather than "a thing moving somewhere"—that 0.400 hitting (as the right tail in a system of variation rather than a separable thing-in-itself) might have disappeared as a consequence of shrinking variation around this stable mean.

I did my first study "on the cheap" when I was recovering from serious illness in the early 1980s (see chapter 4). I propped myself up in bed with the only book in common use that is thicker than the Manhattan telephone directory—*The Baseball Encyclopedia* (New York, Macmillan). I decided to treat the mean batting average for the five best and five worst players in each year as an acceptable measure of achievement at the right and left tails of the bell curve for batting averages. I then calculated the difference between these five highest and the league average (and also between the five lowest and the league average) for each year since the beginning of major league baseball, in 1876. If the difference between best and average (and worst and average) declines through time, then we will have a rough measurement for the shrinkage of variation.

The five best are easily identified, for the *Encyclopedia* lists them in yearly tables of highest achievement. But nobody bothers to memorialize the five worst, so I had to go through the rosters, man by man, looking for the five lowest averages among regular players with at least two at-bats per game over a full season. I present the results in Figure 15—a clear confirmation of my hypothesis, as variation shrinks systematically and symmetrically, bringing both right and left tails ever closer to the stable mean through time. Thus, the disappearance of 0.400 hitting occurred because the bell curve for batting averages has become skinnier over the years, as extreme values at both right and left tails of the distribution get trimmed and shaved. To understand the extinction of 0.400 hitting, we must ask why variation declined in this particular pattern.

Several years later I redid the study by a better, albeit far more laborious, method of calculating the conventional measure of total variation—the standard deviation—for all regular players in each year (three weeks at the computer for my research assistant—and did he ever relish the break from measuring snails!—rather than several enjoyable personal hours propped up in bed with the *Baseball Encyclopedia*).

The standard deviation is a statistician's basic measure of variation. The calculated value for each year records the spread of the entire bell curve, measured (roughly) as the average departure of players from the mean—thus giving us, in a single number, our best assessment of the full range of variation. To compute the standard deviation, you take (in this case) each individual batting average and subtract from it the league average for that year. You then square each value (multiply it by itself) in order to eliminate negative numbers for batting averages below the mean (for a negative times a negative yields a positive number). You then add up all these values and divide them by the total number of players, giving an average squared deviation of individual players from the mean. Finally, you take the square root of this number to obtain the average, or standard, deviation itself. The higher the value of the standard deviation, the more extensive, or spread out, the variation.[6]

Calculation by standard deviation gives a more detailed account of the shrinkage of variation in batting averages through time—see Figure 16,

6. I referred to my first method as working "on the cheap" because five-highest and five-lowest represents a quicker and dirtier calculation than the full standard deviation of all players. But I knew that this shortcut would provide a good surrogate for the more accurate standard deviation because standard deviations are particularly sensitive to values farthest from the mean—a consequence of squaring the deviation of each player from the mean at one point in the calculation. Since my quick-and-dirty method relied entirely on values farthest from the mean, I knew that it would correlate closely with the standard deviation.

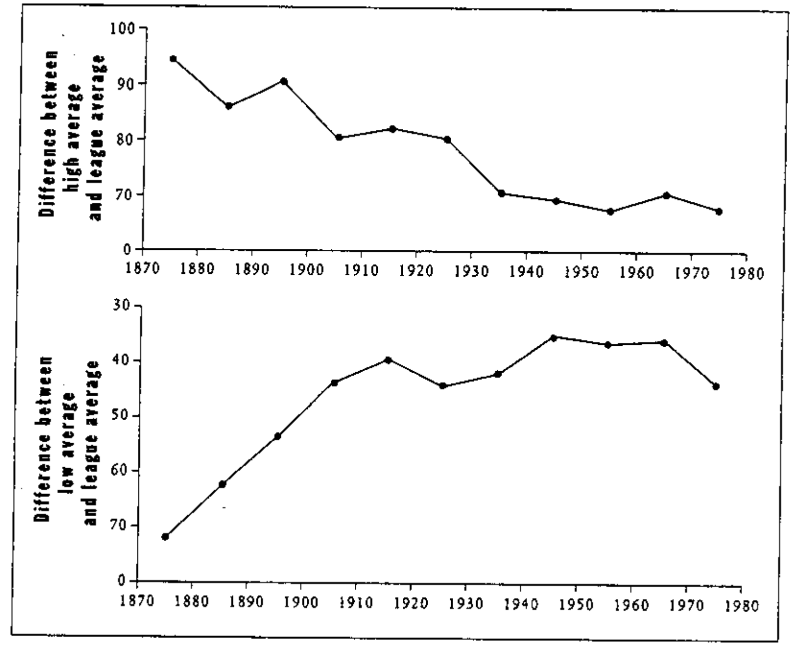

FIGURE 15
Declining differences between highest and lowest batting averages and league means throughout the history of baseball.

which plots the changes in standard deviation year by year, with no averaging over decades or other intervals. My general hypothesis is confirmed again: variation decreases steadily through time, leading to the disappearance of 0.400 hitting as a consequence of shrinkage at the right tail of the distribution. But, using this preferable, and more powerful, method of standard deviations, we can discern some confirming subtleties in the pattern of decrease that our earlier analysis missed. We note in particular that, while standard deviations have been dropping steadily and irreversibly, the decline itself has decelerated over the years as baseball has stabilized—rapidly during the nineteenth century, more slowly during the twentieth, and reaching a plateau by about 1940.

Please pardon a bit of crowing, but I was stunned and delighted (beyond all measure) by the elegance and clarity of this result. I knew from my previous analysis what the general pattern would show, but I never dreamed that the decline of variation would be so regular, so devoid of exception or anomaly for even a single year, so unvarying that we could even pick out such subtleties as a deceleration in decline. I have spent my entire professional career studying such statistical distributions, and I know how rarely one obtains such clean results in better-behaved data of controlled experiments or natural growth in simple systems. We usually encounter some glitch, some anomaly, some funny years. But the decline of standard deviations for batting averages is so regular that the pattern of Figure 16 looks like a plot for a law of nature.

I find the regularity all the more remarkable because the graph of mean batting averages themselves through time (Figure 14) shows all the noise and fluctuation expected in natural systems. These mean batting averages have frequently been manipulated by the rulemakers of baseball to maintain a general constancy, while no one has tried to monkey with the standard deviations. Nonetheless, while mean batting averages go up and down to follow the whims of history and the vagaries of invention, the standard deviation has marched steadily down at a decreasing pace, apparently disturbed by nothing of note, apparently following some interesting rule or general principle in the behavior of systems—a principle that should provide a solution to the classic dilemma of why 0.400 hitting has disappeared.

The details of Figure 16 are impressive in their exceptionless regularity. All four beginning years of the 1870s feature high values of standard deviations greater than 0.050, while the last reading in excess of 0.050 occurs in 1886. Values between 0.04 and 0.05 characterize the remainder of the nineteenth century, with three years just below at 0.038 to 0.040. But the last reading in excess of 0.040 occurs in 1911. Subsequently, decline within the 0.03-to-0.04 range shows the same precision of detail in unreversed decrease over many years. The last reading as high as 0.037 occurs in 1937, and of 0.035 in 1941. Only two years have exceeded 0.034 since 1957. Between 1942 and 1980, values remained entirely within the restricted range of 0.0285 to 0.0343. I had thought that at least one unusual year would upset the pattern, that at least one nineteenth-century value would reach a late-twentieth-century low, or one recent year soar to a nineteenth-century high—but nothing of the sort occurs. All yearly mea-

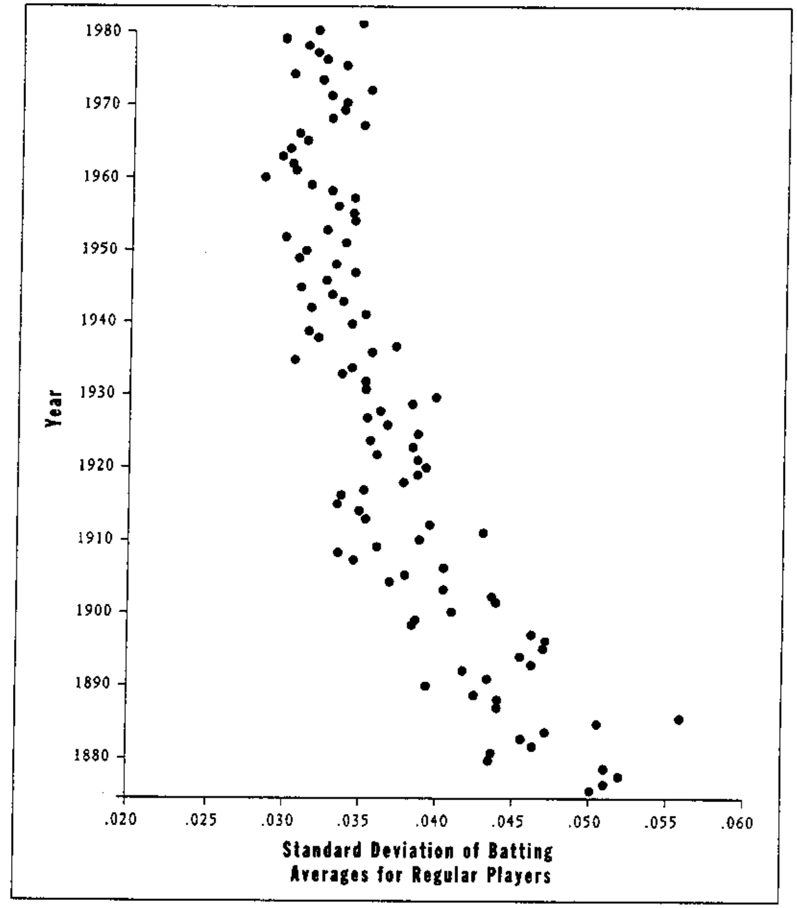

FIGURE 16
Standard deviation of batting averages for all full-time players by year for the first 100 years of professional baseball. Note the regular decline.

sures from 1906 back to the beginning of major league baseball are higher than every reading from 1938 to 1980. We find no overlap at all. Speaking as an old statistical trouper, I can assure you that this pattern represents regularity with a vengeance. This analysis has uncovered something general, something beyond the peculiarity of an idiosyncratic system, some rule or principle that should help us to understand why 0.400 hitting has become extinct in baseball.

·10·

Why the Death of 0.400 Hitting Records Improvement of Play

So far I have only demonstrated a pattern based on unconventional concepts and pictures. I have not yet proposed an explanation. I have proposed that 0.400 hitting be reconceptualized as an inextricable segment in a full house of variation—as the right tail of the bell curve of batting averages—and not as a self-contained entity whose disappearance must record the degeneration of batting in some form or other.

In this different model and picture, 0.400 hitting disappears as a consequence of shrinking variation around a stable mean batting average. The shrinkage is so exceptionless, so apparently lawlike in its regularity, that we must be discerning something general about the behavior of systems through time.

Why should such a shrinkage of variation record the worsening of anything? The final and explanatory step in my argument must proceed beyond the statistical analysis of batting averages. We must consider both the nature of baseball as a system, and some general properties of systems that enjoy long persistence with no major changes in procedures and behaviors. I therefore devote this section to reasons for celebrating the loss of 0.400 hitting as a mark of better baseball.

Two arguments, and supporting data, convince me that shrinkage of variation (with consequent disappearance of 0.400 hitting) must be measuring a general improvement of play. The two formulations sound quite dissimilar at first, but really represent different facets of a single argument.

1. *Complex systems improve when the best performers play by the same rules over extended periods of time. As systems improve, they equilibrate and variation decreases.* No other major American sport permits such an analysis, for all others have changed their fundamental rules too often and too recently. As a teenager, I played basketball without the twenty-four-second rule. My father played with a center jump after each basket. His father (had he been either inclined or acculturated) would have brought the ball downcourt with a two-handed dribble. And Mr. Naismith's boys threw the ball into a peach basket. While the peach basket still hung in the 1890s, baseball made its last major change in procedure (as discussed in the last chapter) by moving the pitcher's mound back to the current distance of sixty feet six inches.

But constant rules don't imply unchanging practices. (In the last chapter I discussed the numerous fiddlings and jigglings imposed by rulemakers to keep pitching and hitting in balance.) Dedicated performers are constantly watching, thinking, and struggling for ways to twiddle or manipulate the system in order to gain a legitimate edge (new techniques for hitting a curve, for gobbling up a ground ball, for gyrating in a windup to fool the batter). Word spreads, and these minor discoveries begin to pervade the system. The net result through time must inevitably encourage an ever-closer approach to optimal performance in all aspects of play—combined with ever-decreasing variation in modes of procedure.

Baseball was feeling its juvenile way during the early days of major league play. The basic rules of the 1890s are still our rules, but scores of subtleties hadn't yet been invented or developed. Rough edges careered out in all directions from a stable center. To cite just a few examples (taken

from Bill James's *Historical Baseball Abstract*): pitchers only began to cover first base in the 1890s. During the same decade, Brooklyn developed the cutoff play, while the Boston Beaneaters invented the hit-and-run, and signals from runner to batter. Gloves were a joke in these early days—just a bit of leather over the hand, not today's baskets for trapping balls. As a fine symbol of broader tolerance and variation, the 1896 Philadelphia Phillies actually experimented for seventy-three games with a lefty shortstop. Unsurprisingly, traditional wisdom applied. He stank—turning in the worst fielding average with the fewest assists among all regular shortstops in the league.

In baseball's youth, styles of play had not become sufficiently regular and optimized to foil the accomplishments of the very best. Wee Willie Keeler could "hit 'em where they ain't" (his motto), and compile a 0.432 batting average in 1897, because fielders didn't yet know where they should be. Slowly, by long distillation of experience, players moved toward optimal methods of positioning, fielding, pitching, and batting—and variation inevitably declined. The best now meet an opposition too finely honed to its own perfection to permit the extremes of accomplishment that characterized a more casual and experimental age. We cannot explain the disappearance of 0.400 hitting simply by saying (however true) that managers invented relief pitching, while pitchers invented the slider—for such traditional explanations abstract 0.400 hitting as an independent phenomenon and view its extinction as the chief sign of a trend to deterioration in batting. Rather, hitting has improved along with all other aspects of play as the entire game sharpened its standards, narrowed its ranges of tolerance, and therefore limited variation in performance as all parts of the game climbed a broader-based hill toward a much narrower pinnacle.

Consider the predicament of a modern Wade Boggs, Tony Gwynn, Rod Carew, or George Brett. Can anyone truly believe that these great hitters are worse than Wee Willie Keeler (at five feet four and a half inches and 140 pounds), Ty Cobb, or Rogers Hornsby? Every pitch is now charted, every hit mapped to the nearest square inch. Fielding and relaying have improved dramatically. Fresh and rested pitching arms must be faced in the late innings; fielders scoop up grounders in gloves as big as a brontosaurus's footprint. Relative to the right wall of human limitation, Tony Gwynn and Wee Willie Keeler must stand in the same place—just a few inches from theoretical perfection (the best that human muscles and bones can do). But average play has so crept up upon Gwynn that he lacks the space for taking advantage of suboptimality in others. All these general improvements must rob great batters of ten to twenty hits a year—a bonus that would be more than enough to convert any of the great modern batters into 0.400 hitters.

I have formulated the argument parochially in the terms and personnel of baseball. But I feel confident that I am describing a general property of systems composed of individual units competing with one another under stable rules and for prizes of victory. Individual players struggle to find means for improvement—up to limits imposed by balances of competition and mechanical properties of materials—and their discoveries accumulate within the system, leading to general gains toward an optimum. As the system nears this narrow pinnacle, variation must decrease—for only the very best can now enter, while their predecessors have slowly, by trial and error, discovered better procedures that now cannot be substantially improved. When someone discovers a truly superior way, everyone else copies and variation diminishes.

Thus I suspect that similar reasons (along with a good dollop of historical happenstance) govern the uniformity of automotive settling upon internal combustion engines from a wider set of initial possibilities including steam and electric power; the standardization of business practices; the reduction of life's initial multicellular animal diversity to just a handful of major phyla (see Gould, 1989); and the disappearance of 0.400 hitting in baseball as variation shrinks symmetrically around a stable mean batting average.

In the good old days of greater variation and poorer play, you could get a job for "good field, no hit"—but no longer as the game improved and the pool of applicants widened. So the left tail shriveled up and moved toward the mean. In those same legendary days, the very best hitters could take advantage of a sloppier system that had not yet discovered optimalities of opposing activities in fielding and pitching. Our modern best hitters are just as good and probably better, but average pitching and fielding have so improved that the truly superb cannot soar so far above the ordinary. Therefore the right tail shriveled up and also moved toward the mean.

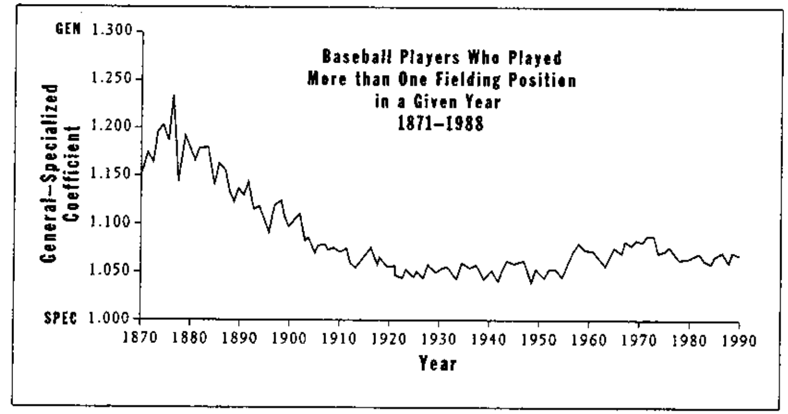

FIGURE 17
Increasing specialization as shown by decline in the number of players who fielded more than one position in a given year.

I first published these ideas in the initial issue of the revived *Vanity Fair* in March 1983. To my gratification, several fellow sabermetricians became intrigued and took up the challenge to test my ideas with other sources of baseball data. The results have been most gratifying. In particular, my colleagues have provided good examples of the two most important predictions made by models for general improvement marked by decreasing variation.

Specialization and division of labor. Ever since Adam Smith began *The Wealth of Nations* with his famous example of pinmaking, specialization and division of labor have been viewed as the major criteria of increasing efficiency and approach to optimality. In their paper "On the tendency toward increasing specialization following the inception of a complex system—professional baseball 1871–1988," John Fellows, Pete Palmer, and Steve Mann plotted the number of major leaguers who played more than one fielding position in a single season. Note (see Figure 17) the steady decrease and subsequent stabilization, a pattern much like the decelerating decline of standard deviations in Figure 16—though in this case measuring the increase of specialization through baseball's history (I do not know why values rose slightly in the 1960s, though to nowhere near the high levels of baseball's early history).

Decreasing variation. My colleagues Sangit Chatterjee and Mustafa Yilmaz of the College of Business Administration at Northeastern University (baseball does provide some wonderful cohesion amid our diversity) wrote an article on "Parity in baseball: stability of evolving systems." In searching for an example even more general than shrinking variation in batting averages, Chatterjee and Yilmaz reasoned that if general play has improved, with less variation among a group of consistently better players, then disparity among teams should also decrease—that is, the difference between the best and worst clubs should decline because all teams can now fill their rosters with enough good players, leading to greater equalization through time. The authors therefore plotted the standard deviation in seasonal winning percentage from the beginning of major league baseball to the present. Figure 18 shows a steady fall in standard deviation, indicating a decreasing difference between the best and worst teams through the history of play.[7]

2. *As play improves and bell curves march toward the right wall, variation must shrink at the right tail.* I discussed the notion of "walls" in chapter 4—upper and lower limits to variation imposed by laws of nature, structure of materials, etc. (There I illustrated a minimal left wall in the story of my medical history—an obvious and logical lower bound of zero time between diagnosis and death from the same disease. Part Four will focus upon a left wall of minimal complexity for life—for nothing much simpler than a bacterial cell could be preserved in the fossil record.) We would all, I think, accept the notion that a "right wall" must exist for human achievement. We cannot, after all, perform beyond the limits of what human bone and muscle can accomplish; no man will ever outpace a cheetah or a finch. We would also, I assume, acknowledge that some extraordinary people,

7. These statistics can also be broken down to yield finer patterns that validate the hypothesis. The National League began in 1876, the American in 1901. Since the hypothesis holds that systems equilibrate through time by decelerating decrease in variation, we might predict that, from 1901 to 1930, when the American League was new but the National already in middle age, variation in American League records should decrease more rapidly than comparable measures in the National League. This pattern does indeed emerge, both for standard deviations of batting averages in my calculations, and for the history of differences in best versus worst teams in the data of Chatterjee and Yilmaz.

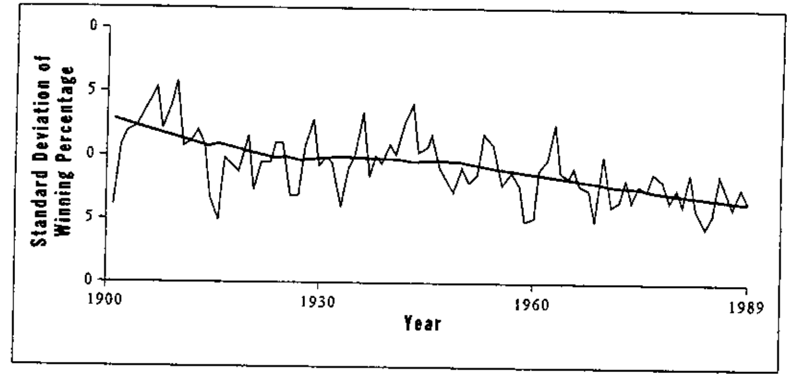

FIGURE 18
Decline in the standard deviation of winning percentage for all teams in the National League through time. The trend shows greater equalization of teams in the history of baseball, a consequence of increasing general excellence of play.

by combination of genetic gift, maniacal dedication, and rigorous training, push their bodies to perform as close to the right wall as human achievement will allow.

Earlier I discussed the major phenomenon in sports that must be signaling approach to the right wall—a flattening out of improvement (measured by record breaking) as sports mature, promise ever greater rewards, become accessible to all, and optimize methods of training (see pp. 92–97). This flattening out must represent the approach of the best to the right wall. The longer a sport has endured with stable rules and maximal access, the closer the best should stand to the right wall, and the less we should therefore expect any sudden and massive breaking of records. When George Plimpton, several years ago, wrote about a great pitching prospect who could throw 140 miles per hour, all serious fans recognized this essay in "straight" reporting as a spoof, though many less knowledgeable folks were fooled. From Walter Johnson in the 1920s to Nolan Ryan today, the best fastball pitchers have tried to throw at maximal speed, and no one has consistently broken 100 mph. In fact, Johnson was probably as fast as Ryan. Thus, we can assume that these men stand near the right wall of what a human arm can do. Barring some unexpected invention in technique, no one is going to descend from some baseball Valhalla and start throwing 40 percent again as fast—not after a century of trying among the very best.

These approaches to the right wall can easily be discerned in sports that keep absolute records measured as times and distances. As previously discussed, record times for the marathon, or virtually any other timed event with stable rules and no major innovations, drop steadily—in the decelerating pattern of initial rapidity, followed by later plateauing as the best draw near to the right wall. But this pattern is masked in baseball, because most records measure one activity relative to another, and not against an absolute standard of time or distance. Batting records mark what a hitter does against pitchers. A mean league batting average of 0.260 is not an absolute measure of anything, but a general rate of success for hitters versus pitchers. Therefore a fall or rise in mean batting average does not imply that hitters are becoming absolutely worse or better, but only that their performance relative to pitchers has changed.

Thus we have been fooled in reading baseball records. We note that the mean batting average has never strayed much from 0.260, and we therefore wrongly assume that batting skills have remained in a century-long rut. We note that 0.400 hitting has disappeared, and we falsely assume that great hitting has gone belly-up. But when we recognize these averages as relative records, and acknowledge that baseball professionals, like all other premier athletes, must be improving with time, a different (and almost surely correct) picture emerges (see Figure 19)—one that acknowledges batting averages as components in a full house of variation with a bell-curve distribution and that, as an incidental consequence of no mean importance, allows us finally to visualize why the extinction of 0.400 hitting must be measuring improvement of play as marked by shrinking variation.

Early in the history of baseball (top part of Figure 19), average play stood far from the right wall of human limits. Both hitters and pitchers performed considerably below modern standards, but the balance between them did not differ from today's—and we measure this unchanging balance as 0.260 hitting. Thus, in these early days, the mean batting average of 0.260 fell well below the right wall, and variation spread out widely on both sides—at the lower end, because the looser and less accomplished system did provide jobs to good fielders who couldn't hit; and

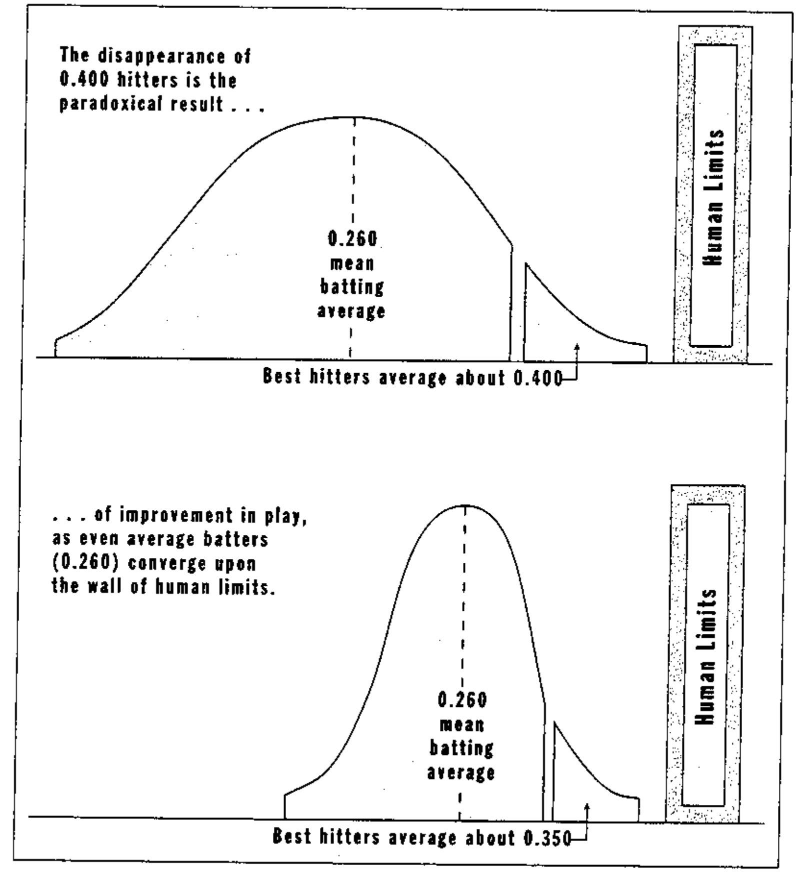

FIGURE 19
Four hundred hitting disappears as play improves and the entire bell curve moves closer to the right wall of human limits while variation declines. Upper chart: early twentieth-century baseball. Lower chart: current baseball.

at the upper end, because so much space existed between the average and the right wall.

A few men of extraordinary talent and dedication always push their skills to the very limit of human accomplishment and reside near the right wall. In baseball's early days, these men stood so far above the mean that we measured their superior performance as 0.400 batting.

Consider what has happened to modern baseball (lower part of Figure 19). General play has improved significantly in all aspects of the game. But the balance between hitting and pitching has not altered. (I showed on pp. 101–105 that the standardbearers of baseball have frequently fiddled with the rules in order to maintain this balance.) The mean batting average has therefore remained constant, but this stable number represents markedly superior performance today (in both hitting *and* pitching). Therefore, this unchanged average must now reside much closer to the right wall. Meanwhile, and inevitably, variation in the entire system has shriveled symmetrically on both sides—at the lower end, because improvement of play now debars employment to men who field well but cannot hit; and at the upper end, for the simple reason that much less room now exists between the upwardly mobile mean and the unchanging right wall. The top hitters, trapped at the upper bound of the right wall, must now lie closer to the mean than did their counterparts of yore.

The best hitters of today can't be worse than 0.400 hitters of the past. In fact, the modern stars may have improved slightly and may now stand an inch or two closer to the right wall. But the average player has moved several feet closer to the right wall—and the distance between ordinary (maintained at 0.260) and best has decreased, thereby erasing batting averages as high as 0.400. Ironically, therefore, the disappearance of 0.400 hitting marks the general improvement of play, not a decline in anything.

Our confidence in this explanation will increase if supporting data can be provided with statistics for other aspects of play through time. I have compiled similar records for the other two major facets of baseball—fielding and pitching. Both support the key predictions of a model that posits increasing excellence of play with decreasing variation when the best can no longer take such numerical advantage of the poorer quality in average performance.

Most batting and pitching records are relative, but the primary measure of good fielding is absolute (or at least effectively so). A fielding average is you against the ball, and I don't think that grounders or fly balls have improved through time (though the hitters have). I suspect that modern fielders are trying to accomplish the same tasks, at about the same level of difficulty, as their older counterparts. Fielding averages (the percent of errorless chances) should therefore provide an absolute measure of changing excellence in play. If baseball has improved, we should note a decelerating rise in fielding averages through time. (I do recognize that some improvement might be attributed to changing conditions, rather than absolutely improving play, just as some running records may fall because modern tracks are better raked and pitched. Older infields were, apparently, lumpier and bumpier than the productions of good ground crews today—so some of the poorer fielding of early days may have resulted from lousy fields rather than lousy fielders. I also recognize that rising averages must be tied in large part to great improvement in the design of gloves—but better equipment represents a major theme of history, and one of the legitimate reasons underlying my claim for general improvement in play.)

Following the procedure of my first compilation on batting averages, I computed both the league fielding average for all regular players and the mean score of the five best for each year since the beginning of major league play in 1876. Figure 20, showing decadal averages for the National League through time, confirms the predictions in a striking manner. Not only does improvement decelerate strongly with time, but the decrease is continuous and entirely unreversed, even for the tiny increments of the last few decades, as averages reach a plateau so near the right wall.

For the first half of baseball (the fifty-five years from 1876 to 1930),

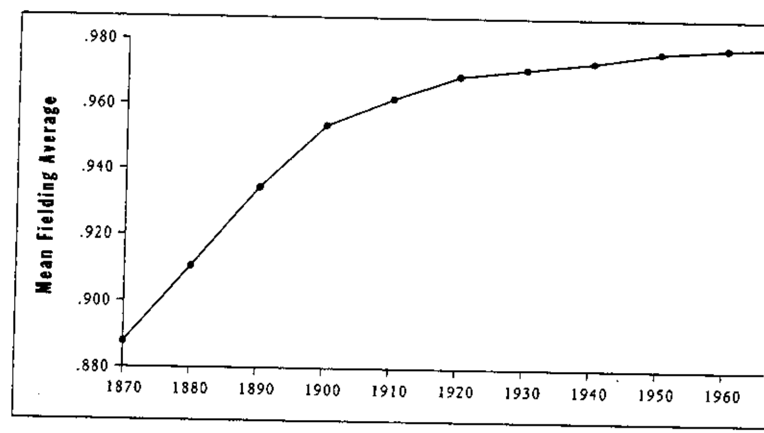

FIGURE 20
Unreversed, but constantly slowing, improvement in mean fielding average through the history of baseball.

decadal fielding averages rose from 0.9622 to 0.9925 for the best players, for a total gain of 0.0303; and from 0.8872 to 0.9685 for average performers, for a total gain of 0.0813. (For a good sense of total improvement, note that the average player of the 1920s did a tiny bit better than the very best fielders of the 1870s.) For baseball's second half (the fifty years from 1931 to 1980), the increase slowed substantially, but never stopped. Decadal averages for the best players rose from 0.9940 to 0.9968, for a small total gain of 0.0028—or less than 10 percent of the recorded rise of baseball's first half. Over the same fifty years, values for league averages rose from 0.971 for the 1930s to 0.9774 for the 1970s, a total gain of 0.0064—again less than 10 percent of the improvement recorded during the same number of years during baseball's first half.

These data continue to excite me. As stated before, I have spent a professional lifetime compiling statistical data of this sort for the growth of organisms and the evolution of lineages. I have a sense of the patterns expected from such data, and have learned to pay special attention to noise and inevitable departures from expectations. I am just not used to the exceptionless data produced over and over again by the history of baseball. I would have thought that any human institution must be more sensitive than natural systems to the vagaries of accident and history, and that baseball would therefore yield more exceptions and a fuzzier signal (if any at all). And yet, here again—as with the decline of standard deviations in batting averages (see page 106)—I find absolute regularity of change, even when the total accumulation is so small that one would expect some exceptions just from the inevitable statistical errors of life and computation. Again, I get the eerie feeling that I must be calculating something quite general about the nature of systems, and not just compiling the individualized numbers of a particular and idiosyncratic institution (yes, I know, it's just a feeling, not a proof). Baseball is a truly remarkable system for statisticians, manifesting two properties devoutly to be wished, but not often encountered, in actual data: an institution that has worked by the same rules for a century, and has compiled complete data (nothing major missing) on all measurable aspects of its history.

For example, as decadal averages for the five best reach their plateau in baseball's second half, improvement slows markedly, but never reverses—the total rise of only 0.0028 occurs in a steady climb by tiny increments: 0.9940, 0.9953, 0.9958, and 0.9968. Lest one consider these gains too small to be anything but accidental, the first achievements of individual yearly values also show the same pattern. Who would have thought that the rise from 0.990 to 0.991 to 0.992, and so on, could mean anything at all? An increment of one in the third decimal place can't possibly be measuring anything significant about actual play. And yet 0.990 is first reached in 1907, 0.991 in 1909, 0.992 in 1914, 0.993 in 1915, 0.994 in 1922, 0.995 in 1930. Then, thank goodness, I find one tiny break in pattern (for I was beginning to think that baseball's God had decided to mock me; the natural world is supposed to contain exceptions). The first value of 0.996 occurs in 1948, but the sharp fielders of 1946 got to 0.997 first! Then we are back on track and do not reach 0.998 until 1972.

This remarkable regularity can occur only because, as my hypothesis requires in its major contention, variation declines so powerfully through time and becomes so restricted in later years. (With such limited variation from year to year, any general signal, however weak, should be more easily detected.) For example, yearly values during the 1930s range only from 0.992 to 0.995 for best scores, and from 0.968 to 0.973 for average scores. By contrast, during baseball's first full decade of the 1880s, the yearly best ranged from 0.966 to 0.981, and the average from 0.891 to 0.927.

This regularity may be affirmed with parallel data for the American League (shown with the National League in Table 3). Again, we find unreversed decline, though this time with one exception as American League values fall slightly during the 1970s—and I have no idea why (if one can properly even ask such a question for such a minuscule effect). Note the remarkable similarity between the leagues in rates of improvement across decades. We are not, of course, observing two independent systems, for styles of play do alter roughly in parallel as both leagues form a single institution (with some minor exceptions, as the National League's blessed refusal to adopt the designated hitter rule indicates in our times). But nearly identical behavior in two cases does show that we are probably picking up a true signal and not a statistical accident.

Data on fielding averages are particularly well suited to illustrate the focal concept of right walls—the key notion behind my second explanation for viewing the disappearance of 0.400 hitting as a sign of general improvement in play. Fielding averages have an absolute, natural, and logical

TABLE 3

Decadal Fielding Averages for Five Best Players and for All Players in Major League Baseball

	National League		American League	
	All Players	Five Best	All Players	Five Best
1870s	.8872	.9622		
1880s	.9103	.9740		
1890s	.9347	.9852		
1900s	.9540	.9874	.9543	.9868
1910s	.9626	.9912	.9606	.9899
1920s	.9685	.9925	.9681	.9940
1930s	.9711	.9940	.9704	.9946
1940s	.9736	.9953	.9740	.9946
1950s	.9763	.9955	.9772	.9960
1960s	.9765	.9958	.9781	.9968
1970s	.9774	.9968	.9776	.9967

right wall of 1.000—for 1.000 represents errorless play, and you cannot make a negative number of errors! Today's best fielders are standing with toes already grazing the right wall—0.998 is about an error per year, and nobody can be absolutely perfect. (Outfielders, pitchers, and catchers occasionally turn in seasons of 1.000 fielding, but only one infielder has ever done so for a full season's regular play—Steve Garvey at first base in 1984.)

If you doubted my explanation for shrinking variation at the upper end of the bell curve for batting averages—that as the mean moves toward the right wall, variation scrunches up into an ever smaller available space, and must therefore decrease—you will surely grant me the argument for

fielding averages so close to an absolute wall. Even the 1870s didn't provide much space, but fielders had a bit of breathing room for improvement between their first decadal best of 0.962 and the wall. And improve they did, and steadily. But now, with the five best averaging 0.9968, there just isn't any more space, barring the construction of truly errorless robotic fielding machines.

As the mean moves toward the wall, variation must decrease. For absolute measures of fielding, high numbers persist and low values get axed. But for relative measures of hitting, the wall itself bears no number. The advancing mean retains the same value (as a balance between hitting and pitching), while both hitting and pitching move in lockstep toward their right walls of human limitation. Thus, 0.400 hitting disappears as the league mean of 0.260 marches steadily toward the wall. But the 0.400 hitters of yore are alive and well, probably more numerous than ever, and standing where they always have resided—just inches from the right wall. But their current best does not measure 0.400 anymore, because everyone else has improved so much, raising average play to a level where an unchanged (or even slightly improved) best can no longer soar so far above the norm.

The best hitters of early baseball could compile 0.400 averages by taking advantage of a standard in average play much lower than today's premier batters encounter. Wade Boggs would hit 0.400 every year against the pitching and fielding of the 1890s, while Wee Willie Keeler would be lucky to crack 0.320 today. Since pitching and batting both feature relative records, and presumably exist in effective balance throughout the history of baseball, we should be able to detect similar phenomena in the statistics of pitching through time. The best pitchers of the past, legendary figures like Christy Mathewson, Cy Young, Walter Johnson, Three Finger Brown, and Grover Cleveland Alexander, should be no better than their modern counterparts Sandy Koufax, Bob Gibson, Tom Seaver, and Nolan Ryan. But the old pitchers, standing next to their own right wall and facing much poorer average batting, should have racked up numbers that modern hurlers just can't equal.

The fascinating and well-known history of minimal earned run averages provides our best illustration of symmetry between batting and pitching—another indication that these statistics record the general behavior of systems, not just a peculiarity of batting in baseball. As the best batters sacrificed their 0.400 averages because variation declined while average play improved, the best pitchers lost their earned run averages below 1.50 because ordinary hitters became too good.

The list of the hundred best seasonal ERAs shows a remarkable imbalance. More than 90 percent of the entries were achieved before 1920. Since then, only nine pitchers have obtained an earned run average in the top one hundred (and remember that the number of pitchers, hence the number of opportunities, has expanded dramatically, first with the introduction of the American League and later with expansion from an original eight to our current roster of fourteen teams per league). Moreover, of these nine modern values, seven rank in the lower half. If we consider the modern achievements, from the bottom up, we get a good sense of the obstacles that must face our superb contemporary pitchers.

Tied at number 100 are Sandy Koufax (1.74 in 1964) and Ron Guidry (1.74 in 1978). Koufax was, well, Koufax—by general agreement the greatest of modern pitchers, perhaps of all pitchers anywhere, anytime (he also holds the ninety-seventh spot at 1.73 for 1966). Guidry, a wonderful Yankee pitcher for a few years, compiled a stellar season in 1978 (with an unmatched combination of total victories and winning percentage of 25–3, for 0.893), and then threw his arm out. Nolan Ryan occupies eighty-seventh place at 1.69 for 1981. And Ryan was, well, Ryan. Nothing else need be said. Carl Hubbell, perhaps the premier pitcher of the 1930s (Lefty Grove was no slouch, either) turned in 1.66 in 1933 for seventy-sixth place and the only entry for his high-hitting decade. Dean Chance, a strictly okay pitcher of the last generation, posted an anomalous 1.65 for seventy-first place in 1964—and I can't figure this one at all. Spud Chandler holds sixty-sixth place at 1.64 for 1943—a fine (if not fabulous) pitcher during the war years, when all decent hitters were blasting away at Germany or Japan instead. Luis Tiant, a damned fine pitcher but not among the greatest, holds sixtieth place at 1.60 for 1968—and I'll return to him in a moment. Dwight Gooden had a fabulous sophomore season in 1985, with a 1.53 ERA that puts him in forty-second place as one of only two modern pitchers in the first half-hundred. He then fell victim to what the newspapers politely call "substance abuse."

We then come to what may be the finest record in modern sports—

Bob Gibson's truly incredible 1.12 ERA of 1968, for fourth place, surrounded by forty old-timers before we meet Doc Gooden at number forty-two. Gibson's only superiors are Tim Keefe with 0.86 in 1880, Dutch Leonard at 0.96 for 1914, and Three Finger Brown at 1.04 for 1906. How could Gibson compile such a record—the only post-1920 value below 1.50, and way, way below at that—in our modern era of greatly improved average hitting?

I don't want to take a thing away from Bob Gibson, who absolutely terrified me in the 1967 World Series, when he almost single-handedly beat the Red Sox by winning three games and casting a pall of inevitability over the whole proceedings. But, in slight mitigation, 1968 was a really funny year, as mentioned previously (see page 104). For some set of reasons that no one understands, pitching took a dramatic upper hand that year, capping a trend of several years' duration. (As explained before, the rulemakers then restored the usual order by lowering the pitching mound and decreasing the strike zone; batting averages and ERAs rose appropriately in the 1969 season and have remained in balance ever since.) The 1968 season didn't just belong to Gibson; in that year, low ERAs sprouted like dandelions in my garden. In most years of modern baseball, no pitcher in either league has posted an ERA lower than 2.00. Uniquely in 1968, all five leading American League pitchers bettered this mark, as Yastrzemski won the batting title with a paltry 0.301 (Tiant at 1.60, McDowell at 1.81, McNally at 1.95, McLain at 1.96—a banner year for Scotland—and John at 1.98. As I said, Tiant was a terrific pitcher and great fun to watch, but not one of the game's greatest. If he could post 1.60 for 1968, baseball was really out of whack that year.) So Gibson certainly took maximal advantage of a weird year, but let's not take anything away from him. No one, no matter how good, had any statistical right to post a value so much better than anything achieved for sixty years, especially when general improvement in play should have made such low ERAs effectively unobtainable. Gibson had one helluva year!

In quick summary of a long and detailed argument, symmetrically shrinking variation in batting averages must record general improvement of play (including hitting, of course) for two reasons—the first (expressed in terms of the history of institutions) because systems manned by best performers in competition, and working under the same rules through time, slowly discover optimal procedures and reduce their variation as all personnel learn and master the best ways; the second (expressed in terms of performers and human limits) because the mean moves toward the right wall, thus leaving less space for the spread of variation. Hitting 0.400 is not a *thing,* but the right tail of the full house for variation in batting averages. As variation shrinks because general play improves, 0.400 hitting disappears as a consequence of increasing excellence in play.

·11·

A Philosophical Conclusion

Some people regard this explanation as a sad story. One can scarcely decry a general improvement in play, but the increasing standardization thus engendered does seem to remove much of the fun and drama from sports. The "play" in play diminishes as activities become ever more "scientific" in the pejorative sense of operating like optimized clockwork. Perhaps no giants inhabited the earth during baseball's early days, but the best then soared so far above the norm that their numbers seemed truly heroic and otherworldly, while our current champions cannot rise nearly so far above the vastly improved average.

But I suggest that we should rejoice in the shrinkage of variation and consequent elimination of 0.400 hitting. Yes, excellence in play does imply increasing precision and standardization, but what complaint can we lodge against repeated maximal beauty? I have now been a fan for fifty years. I have seen hundreds of perfectly executed double plays and brilliant pegs from outfield to home (that may or may not have beaten the runner charging from third)—the kind of beautifully orchestrated precision that probably occurred only rarely in baseball's early years. I do not thrill any less with each repetition. The pinnacle of excellence is so rare, its productions so exquisite. Did we ever get bored with Caruso or Pavarotti in their prime? I would much rather have my expectation of excellence affirmed when I go to the ballpark or the opera house than to take potluck and hope for a rare glimpse of glory in a sea of mediocrity.

Moreover, the rise in general excellence and consequent shrinkage of variation does not remove the possibility of transcendence. In fact, I would argue that transcendence becomes all the more intriguing and exciting for the smaller space now allocated to such a possibility, and for the consequently greater struggle that must attend the achievement. When the norm stood miles from the right wall, records could be broken with relative ease. But when the average player can almost touch the wall, then transcendence of the mean marks a true outer limit for conceivable human achievement. (Again, I would make an analogy to musical performance. Do we not rejoice when every string in a symphony orchestra plays with exquisite beauty and consummate professionalism? And do we not thrill all the more when, in this context of superb general performance, a great soloist does something so special that only angels in heaven could have contemplated the possibility?) I would carry the argument even further and point out that a norm near the right wall pushes the very best to seek levels of greater accomplishment that otherwise might never have been conceptualized. I will speak in the final chapter about the heroic efforts, often with attendant accident and loss of life, that such "pushing of the envelope" imposes in the almost holy mania that infects the greatest performers in the circus arts and other dangerous activities. Call it foolish (and swear up and down that you would never so act yourself), but acknowledge that human greatness often forms a strange partnership with human obsession, and that the mix sometimes spells glory—or death.

The possibility of transcendence can never die, because this pinnacle of admiration in sports can be reached by several attainable routes. First of all, a kind of democracy infests individual games. When we go to the ballpark, we never know what we will witness. At any time, even the worst team may execute a thrilling play with awesome perfection. The event may

occur only once a year (or much less often) on average, but the day of your attendance may feature a triple play, a steal of home, a rip-roaring, bench-clearing brawl (yes, as *Homo ludens* and *Homo stupidus,* we also root for this sort of rare nonsense from the underside of our complete lives), or an inside-the-park homer, with the runner just slipping under the catcher's tag. You never know.

The enormous variability of individual performance guarantees that even a mediocre player can, for one day of glory, accomplish something never done before, or even dreamed of in baseball's philosophy. Harvey Haddix was a fine pitcher, but not the greatest. Yet one day he hurled twelve innings of perfect ball—and then lost the game in the thirteenth (as the opposing pitcher had shut out Haddix's side for the first twelve innings). Bobby Thomson was a better-than-average outfielder for the New York Giants, but one day in 1951 he hit a home run, perfectly ordinary by the physics of distance, but meaningful beyond measure in baseball's enclosed system, because this single blow won the pennant for the Giants against their archrivals, the Brooklyn Dodgers, in the last inning of the last game of a play-off series culminating the greatest comeback in the history of baseball (the Giants had trailed the Dodgers by thirteen and a half games in August, and had entered this last inning with an apparently insurmountable three-run deficit). I was a ten-year-old Giants fan watching the game on our family's first television set, and I have never been so thrilled in all my life (except for one other time).

Don Larsen was a truly mediocre pitcher for the Yankees, but he achieved baseball's definition of perfection when it mattered most: twenty-seven Dodgers up, twenty-seven Bums down on October 8, 1956, for a perfect game in the World Series (no one before or since had ever thrown a no-hitter of any kind in a World Series game). I was a fifteen-year-old Yankees fan (many New Yorkers rooted for two teams, one local club in each league), trying to persuade my French teacher to let us listen to the game's end on the radio. I have never been so thrilled in all my life (except for one other time).

When we move to the statistics of seasonal or lifetime performance, this kind of democracy vanishes, and only the truly great can achieve transcendence. But some humans can push themselves, by an alchemy of inborn skill, happy fortuity, and maniacal dedication, to performances that just shouldn't happen—and we revel when such a man reaches farther and actually touches the right wall. Bob Gibson had no business compiling an ERA of 1.12 in 1968. And I can show you with copious statistics that Joe DiMaggio should never have hit in fifty-six straight games in 1941 (see Gould, 1988). I delayed writing the last paragraph of this chapter for several days because I couldn't bear not to share vicariously in a great moment of transcendence. So I am sitting at this old typewriter on September 6, 1995, as Cal Ripken plays his 2131st consecutive game, eclipsing the "unbreakable" record of the Iron Horse, Lou Gehrig.

No records lie beyond fracture (unless rules or practices have changed to make an old achievement unattainable in modern performance). Perhaps I have exaggerated by discussing the "extinction" of 0.400 hitting in this section. (I am a paleontologist and hate to avoid one of the favorite words in my trade.) But I meant extinction in the literal sense of snuffing out a candle that might be lit again, not in the evolutionary and ecological meaning of species death where, by an accurate motto of our times, extinction is truly forever.

I am not arguing that no one will ever hit 0.400 again. I do say that such a mark has become a consummate rarity, achieved perhaps once in a century like a hundred-year flood, and not the common pinnacle of baseball's early years. The fifty-year drought since Ted Williams supports this view, and I think that this part has identified the reason by reconceptualizing 0.400 hitting as the right tail in a shrinking bell curve of batting averages with a stable mean—all as a necessary and predictable consequence of general improvement in play. But someday, someone will hit 0.400 again—though this time the achievement will be so much more difficult than ever before and therefore so much more worthy of honor. When the idiots on both sides in the great pissing contest of 1994 (otherwise known as a labor dispute) aborted the season and canceled the World Series, Tony Gwynn was batting 0.392 and on the rise. I believe that he would have succeeded had the season unfolded as history and propriety demanded. Someday, someone will join Ted Williams and touch the right wall against higher odds than ever before. Every season brings this possibility. Every season features the promise of transcendence.

Part Four

The Modal Bacter: Why Progress Does Not Rule the History of Life