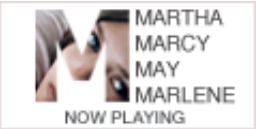
The New York Times Reprints

This copy is for your personal, noncommercial use only. You can order presentation-ready copies for distribution to your colleagues, clients or customers here or use the "Reprints" tool that appears next to any article. Visit www.nytreprints.com for samples and additional information. Order a reprint of this article now.



October 19, 2011

Don't Blink! The Hazards of Confidence

By DANIEL KAHNEMAN

Many decades ago I spent what seemed like a great deal of time under a scorching sun, watching groups of sweaty soldiers as they solved a problem. I was doing my national service in the Israeli Army at the time. I had completed an undergraduate degree in psychology, and after a year as an infantry officer, I was assigned to the army's Psychology Branch, where one of my occasional duties was to help evaluate candidates for officer training. We used methods that were developed by the British Army in World War II.

One test, called the leaderless group challenge, was conducted on an obstacle field. Eight candidates, strangers to one another, with all insignia of rank removed and only numbered tags to identify them, were instructed to lift a long log from the ground and haul it to a wall about six feet high. There, they were told that the entire group had to get to the other side of the wall without the log touching either the ground or the wall, and without anyone touching the wall. If any of these things happened, they were to acknowledge it and start again.

A common solution was for several men to reach the other side by crawling along the log as the other men held it up at an angle, like a giant fishing rod. Then one man would climb onto another's shoulder and tip the log to the far side. The last two men would then have to jump up at the log, now suspended from the other side by those who had made it over, shinny their way along its length and then leap down safely once they crossed the wall. Failure was common at this point, which required starting over.

As a colleague and I monitored the exercise, we made note of who took charge, who tried to lead but was rebuffed, how much each soldier contributed to the group effort. We saw who seemed to be stubborn, submissive, arrogant, patient, hot-tempered, persistent or a quitter. We sometimes saw competitive spite when someone whose idea had been rejected by the group no longer worked very hard. And we saw reactions to crisis: who berated a comrade whose mistake caused the whole group to fail, who stepped forward to lead when the exhausted team had to start over. Under the stress of the event, we felt, each man's true nature revealed itself in sharp relief.

After watching the candidates go through several such tests, we had to summarize our impressions of the soldiers' leadership abilities with a grade and determine who would be eligible for officer training. We spent some time discussing each case and reviewing our impressions. The task was not difficult, because we had already seen each of these soldiers' leadership skills. Some of the men looked like strong leaders, others seemed like wimps or arrogant fools, others mediocre but not hopeless. Quite a few appeared to be so weak that we ruled them out as officer candidates. When our multiple observations of each candidate converged on a coherent picture, we were completely confident in our evaluations and believed that what we saw pointed directly to the future. The soldier who took over when the group was in trouble and led the team over the wall was a leader at that moment. The obvious best guess about how he would do in training, or in combat, was that he would be as effective as he had been at the wall. Any other prediction seemed inconsistent with what we saw.

Because our impressions of how well each soldier performed were generally coherent and clear, our formal predictions were just as definite. We rarely experienced doubt or conflicting impressions. We were quite willing to declare: "This one will never make it," "That fellow is rather mediocre, but should do O.K." or "He will be a star." We felt no need to question our forecasts, moderate them or equivocate. If challenged, however, we were fully prepared to admit, "But of course anything could happen."

We were willing to make that admission because, as it turned out, despite our certainty about the potential of individual candidates, our forecasts were largely useless. The evidence was overwhelming. Every few months we had a feedback session in which we could compare our evaluations of future cadets with the judgments of their commanders at the officer-training school. The story was always the same: our ability to predict performance at the school was negligible. Our forecasts were better than blind guesses, but not by much.

We were downcast for a while after receiving the discouraging news. But this was the army. Useful or not, there was a routine to be followed, and there were orders to be obeyed. Another batch of candidates would arrive the next day. We took them to the obstacle field, we faced them with the wall, they lifted the log and within a few minutes we saw their true natures revealed, as clearly as ever. The dismal truth about the quality of our predictions had no effect whatsoever on how we evaluated new candidates and very little effect on the confidence we had in our judgments and predictions.

I thought that what was happening to us was remarkable. The statistical evidence of our failure should have shaken our confidence in our judgments of particular candidates, but it did not. It

should also have caused us to moderate our predictions, but it did not. We knew as a general fact that our predictions were little better than random guesses, but we continued to feel and act as if each particular prediction was valid. I was reminded of visual illusions, which remain compelling even when you know that what you see is false. I was so struck by the analogy that I coined a term for our experience: the illusion of validity.

I had discovered my first cognitive fallacy.

Decades later, I can see many of the central themes of my thinking about judgment in that old experience. One of these themes is that people who face a difficult question often answer an easier one instead, without realizing it. We were required to predict a soldier's performance in officer training and in combat, but we did so by evaluating his behavior over one hour in an artificial situation. This was a perfect instance of a general rule that I call WYSIATI, "What you see is all there is." We had made up a story from the little we knew but had no way to allow for what we did not know about the individual's future, which was almost everything that would actually matter. When you know as little as we did, you should not make extreme predictions like "He will be a star." The stars we saw on the obstacle field were most likely accidental flickers, in which a coincidence of random events — like who was near the wall — largely determined who became a leader. Other events — some of them also random — would determine later success in training and combat.

You may be surprised by our failure: it is natural to expect the same leadership ability to manifest itself in various situations. But the exaggerated expectation of consistency is a common error. We are prone to think that the world is more regular and predictable than it really is, because our memory automatically and continuously maintains a story about what is going on, and because the rules of memory tend to make that story as coherent as possible and to suppress alternatives. Fast thinking is not prone to doubt.

The confidence we experience as we make a judgment is not a reasoned evaluation of the probability that it is right. Confidence is a feeling, one determined mostly by the coherence of the story and by the ease with which it comes to mind, even when the evidence for the story is sparse and unreliable. The bias toward coherence favors overconfidence. An individual who expresses high confidence probably has a good story, which may or may not be true.

I coined the term "illusion of validity" because the confidence we had in judgments about individual soldiers was not affected by a statistical fact we knew to be true — that our predictions were unrelated to the truth. This is not an isolated observation. When a compelling impression of a

particular event clashes with general knowledge, the impression commonly prevails. And this goes for you, too. The confidence you will experience in your future judgments will not be diminished by what you just read, even if you believe every word.

I first visited a Wall Street firm in 1984. I was there with my longtime collaborator Amos Tversky, who died in 1996, and our friend Richard Thaler, now a guru of behavioral economics. Our host, a senior investment manager, had invited us to discuss the role of judgment biases in investing. I knew so little about finance at the time that I had no idea what to ask him, but I remember one exchange. "When you sell a stock," I asked him, "who buys it?" He answered with a wave in the vague direction of the window, indicating that he expected the buyer to be someone else very much like him. That was odd: because most buyers and sellers know that they have the same information as one another, what made one person buy and the other sell? Buyers think the price is too low and likely to rise; sellers think the price is high and likely to drop. The puzzle is why buyers and sellers alike think that the current price is wrong.

Most people in the investment business have read Burton Malkiel's wonderful book "A Random Walk Down Wall Street." Malkiel's central idea is that a stock's price incorporates all the available knowledge about the value of the company and the best predictions about the future of the stock. If some people believe that the price of a stock will be higher tomorrow, they will buy more of it today. This, in turn, will cause its price to rise. If all assets in a market are correctly priced, no one can expect either to gain or to lose by trading.

We now know, however, that the theory is not quite right. Many individual investors lose consistently by trading, an achievement that a dart-throwing chimp could not match. The first demonstration of this startling conclusion was put forward by Terry Odean, a former student of mine who is now a finance professor at the University of California, Berkeley.

Odean analyzed the trading records of 10,000 brokerage accounts of individual investors over a seven-year period, allowing him to identify all instances in which an investor sold one stock and soon afterward bought another stock. By these actions the investor revealed that he (most of the investors were men) had a definite idea about the future of two stocks: he expected the stock that he bought to do better than the one he sold.

To determine whether those appraisals were well founded, Odean compared the returns of the two stocks over the following year. The results were unequivocally bad. On average, the shares investors sold did better than those they bought, by a very substantial margin: 3.3 percentage points per year, in addition to the significant costs of executing the trades. Some individuals did

much better, others did much worse, but the large majority of individual investors would have done better by taking a nap rather than by acting on their ideas. In a paper titled "Trading Is Hazardous to Your Wealth," Odean and his colleague Brad Barber showed that, on average, the most active traders had the poorest results, while those who traded the least earned the highest returns. In another paper, "Boys Will Be Boys," they reported that men act on their useless ideas significantly more often than women do, and that as a result women achieve better investment results than men.

Of course, there is always someone on the other side of a transaction; in general, it's a financial institution or professional investor, ready to take advantage of the mistakes that individual traders make. Further research by Barber and Odean has shed light on these mistakes. Individual investors like to lock in their gains; they sell "winners," stocks whose prices have gone up, and they hang on to their losers. Unfortunately for them, in the short run going forward recent winners tend to do better than recent losers, so individuals sell the wrong stocks. They also buy the wrong stocks. Individual investors predictably flock to stocks in companies that are in the news. Professional investors are more selective in responding to news. These findings provide some justification for the label of "smart money" that finance professionals apply to themselves.

Although professionals are able to extract a considerable amount of wealth from amateurs, few stock pickers, if any, have the skill needed to beat the market consistently, year after year. The diagnostic for the existence of any skill is the consistency of individual differences in achievement. The logic is simple: if individual differences in any one year are due entirely to luck, the ranking of investors and funds will vary erratically and the year-to-year correlation will be zero. Where there is skill, however, the rankings will be more stable. The persistence of individual differences is the measure by which we confirm the existence of skill among golfers, orthodontists or speedy toll collectors on the turnpike.

Mutual funds are run by highly experienced and hard-working professionals who buy and sell stocks to achieve the best possible results for their clients. Nevertheless, the evidence from more than 50 years of research is conclusive: for a large majority of fund managers, the selection of stocks is more like rolling dice than like playing poker. At least two out of every three mutual funds underperform the overall market in any given year.

More important, the year-to-year correlation among the outcomes of mutual funds is very small, barely different from zero. The funds that were successful in any given year were mostly lucky; they had a good roll of the dice. There is general agreement among researchers that this is true for nearly all stock pickers, whether they know it or not — and most do not. The subjective experience

of traders is that they are making sensible, educated guesses in a situation of great uncertainty. In highly efficient markets, however, educated guesses are not more accurate than blind guesses.

Some years after my introduction to the world of finance, I had an unusual opportunity to examine the illusion of skill up close. I was invited to speak to a group of investment advisers in a firm that provided financial advice and other services to very wealthy clients. I asked for some data to prepare my presentation and was granted a small treasure: a spreadsheet summarizing the investment outcomes of some 25 anonymous wealth advisers, for eight consecutive years. The advisers' scores for each year were the main determinant of their year-end bonuses. It was a simple matter to rank the advisers by their performance and to answer a question: Did the same advisers consistently achieve better returns for their clients year after year? Did some advisers consistently display more skill than others?

To find the answer, I computed the correlations between the rankings of advisers in different years, comparing Year 1 with Year 2, Year 1 with Year 3 and so on up through Year 7 with Year 8. That yielded 28 correlations, one for each pair of years. While I was prepared to find little year-to-year consistency, I was still surprised to find that the average of the 28 correlations was .01. In other words, zero. The stability that would indicate differences in skill was not to be found. The results resembled what you would expect from a dice-rolling contest, not a game of skill.

No one in the firm seemed to be aware of the nature of the game that its stock pickers were playing. The advisers themselves felt they were competent professionals performing a task that was difficult but not impossible, and their superiors agreed. On the evening before the seminar, Richard Thaler and I had dinner with some of the top executives of the firm, the people who decide on the size of bonuses. We asked them to guess the year-to-year correlation in the rankings of individual advisers. They thought they knew what was coming and smiled as they said, "not very high" or "performance certainly fluctuates." It quickly became clear, however, that no one expected the average correlation to be zero.

What we told the directors of the firm was that, at least when it came to building portfolios, the firm was rewarding luck as if it were skill. This should have been shocking news to them, but it was not. There was no sign that they disbelieved us. How could they? After all, we had analyzed their own results, and they were certainly sophisticated enough to appreciate their implications, which we politely refrained from spelling out. We all went on calmly with our dinner, and I am quite sure that both our findings and their implications were quickly swept under the rug and that life in the firm went on just as before. The illusion of skill is not only an individual aberration; it is deeply ingrained in the culture of the industry. Facts that challenge such basic assumptions — and thereby

threaten people's livelihood and self-esteem — are simply not absorbed. The mind does not digest them. This is particularly true of statistical studies of performance, which provide general facts that people will ignore if they conflict with their personal experience.

The next morning, we reported the findings to the advisers, and their response was equally bland. Their personal experience of exercising careful professional judgment on complex problems was far more compelling to them than an obscure statistical result. When we were done, one executive I dined with the previous evening drove me to the airport. He told me, with a trace of defensiveness, "I have done very well for the firm, and no one can take that away from me." I smiled and said nothing. But I thought, privately: Well, I took it away from you this morning. If your success was due mostly to chance, how much credit are you entitled to take for it?

We often interact with professionals who exercise their judgment with evident confidence, sometimes priding themselves on the power of their intuition. In a world rife with illusions of validity and skill, can we trust them? How do we distinguish the justified confidence of experts from the sincere overconfidence of professionals who do not know they are out of their depth? We can believe an expert who admits uncertainty but cannot take expressions of high confidence at face value. As I first learned on the obstacle field, people come up with coherent stories and confident predictions even when they know little or nothing. Overconfidence arises because people are often blind to their own blindness.

True intuitive expertise is learned from prolonged experience with good feedback on mistakes. You are probably an expert in guessing your spouse's mood from one word on the telephone; chess players find a strong move in a single glance at a complex position; and true legends of instant diagnoses are common among physicians. To know whether you can trust a particular intuitive judgment, there are two questions you should ask: Is the environment in which the judgment is made sufficiently regular to enable predictions from the available evidence? The answer is yes for diagnosticians, no for stock pickers. Do the professionals have an adequate opportunity to learn the cues and the regularities? The answer here depends on the professionals' experience and on the quality and speed with which they discover their mistakes. Anesthesiologists have a better chance to develop intuitions than radiologists do. Many of the professionals we encounter easily pass both tests, and their off-the-cuff judgments deserve to be taken seriously. In general, however, you should not take assertive and confident people at their own evaluation unless you have independent reason to believe that they know what they are talking about. Unfortunately, this advice is difficult to follow: overconfident professionals sincerely believe they have expertise, act as experts and look like experts. You will have to struggle to remind yourself that they may be in the

grip of an illusion.

Daniel Kahneman is emeritus professor of psychology and of public affairs at Princeton University and a winner of the 2002 Noble Prize in Economics. This article is adapted from his book "Thinking, Fast and Slow," out this month from Farrar, Straus & Giroux.

Editor: Dean Robinson

OPEN

MORE IN MAGAZINE (4 OF 21 ARTICLES)

Bad Guys vs. Worse Guys in Afghanistan

Read More »