
The Case for Systematic Decision-Making

by Wesley R. Gray Ph.D.

“If you do fundamental trading one morning you feel like a genius, the next day you feel like an idiot....by 1998 I decided we would go 100% models...we slavishly follow the model. You do whatever it [the model] says no matter how smart or dumb you think it is. And that turned out to be a wonderful business.”

This quote, from Jim Simons, founder of the world’s most successful hedge fund, Renaissance Technologies, demonstrated the utility of systematic decision-making in an MIT video.

The urge to use our judgment throughout the investing process is strong. I argue that, while investors need human experts to design models, they should let computers be in charge of applying those models and fight the urge to use their judgment in the implementation process. “Gut-based,” or discretionary, stock pickers certainly have a compelling story: Invest countless hours in research, identify investment opportunities and profit from the hard work. Stock pickers, however, rely on the false premise that “countless hours of being busy” adds value in the context of investment management. The empirical evidence on the subject of systematic versus discretionary decision-making is abundantly clear: Models beat experts. In fact, the late Paul Meehl, one of the great minds in the field of psychology, describes the body of evidence on the “models versus experts” debate as the only controversy in social science with “such a large body of qualitatively diverse studies coming out so uniformly in the same direction.”

Econs and Humans

University of Chicago professors Dick Thaler and Cass Sunstein in their bestseller book “Nudge: Improving Decisions About Health, Wealth, and Happiness” (Yale University Press, 2008) describe two types of people that can be found in the world: econs and humans. Econs are fully rational, continuously calculating and have both unlimited attention and mental resources. Humans are a decidedly less rational and more emotionally driven bunch. This view is based on an understanding of two ways of thinking that are innate to humans. As described in Daniel Kahneman’s great work “Thinking, Fast and Slow” (Farrar, Straus and Giroux, 2011), humans are driven by two modes of thinking: System 1 and System 2. System 1 decisions are instinctual and automated by the brain; System 2 processes are rational and analytical.

System 1, while imperfect, is highly efficient. For example, if Joe is facing the threat of a large tiger charging him at full speed, System 1 will trigger Joe to turn around and sprint for the nearest tree, and ask questions later. As an alternative, Joe's System 2 will calculate the speed of the tiger's approach and assess his situation. Joe will examine his options and realize that he has a loaded revolver that can take the tiger down in an instant.

On average, if Joe immediately sprints to the tree he may get lucky and outrun the tiger. If, on the other hand, Joe pauses and calculates his best option, which is shooting at the tiger with his revolver, his tactical pause may end with Joe trying to remove a 500-pound meat-eating monster from his jugular vein.

Joe's tiger situation highlights why evolution has created System 1: On average, running for the tree is a life-saving decision when faced with a high-stress situation where survival is on the line. The issue with System 1 is that heuristic-based mechanisms often lead to systematic bias: Joe will almost always run, even when sometimes he should shoot. System 1 certainly served its purpose when humans were faced with life and death situations in the jungles, but in modern day life, where decisions in chaos have limited consequence, the benefits of immediate decisions rarely outweigh the costs of flawed decision-making. The necessity of avoiding System 1 and relying on System 2 in the context of financial markets is of utmost importance.

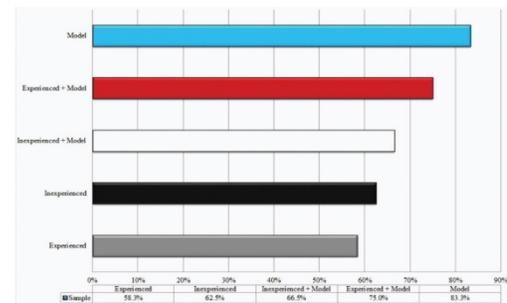
Perception Is Not Reality

Ted Adelson, a vision scientist at MIT, has developed an illusion that highlights the fallibility of the human brain. This illusion is shown as [Figure 1](#).

Stare at cells A and B in [Figure 1](#). Do the colors of the squares look different? How confident are you that A is a different color than B? What odds would you accept in a bet? 5-1? 20-1? If you are a human, you should be confident that A and B are different. However, if you are an econ, your computer-like brain will identify a pixel in cell A and B, compare the red-green-blue values and identify that each is 120-120-120, a perfect match. Stare a little longer, but this time cut pieces of paper to create a small box around cells A and B. Now it should be clear: A and B are the same. The lesson here, and its applicability to decision making, is best described by Mark Twain, "It ain't what you don't know that gets you into trouble, it's what you know for sure that simply ain't so." As investors, we need to be most wary of situations where "we know" something is bound to happen.

factors. The model utilizes a systematic approach based on a statistical model of prior data; the humans utilize their experience and intuition. The results from the study are striking. First, the simple quantitative model has a classification accuracy ratio of 83.3%. The most experienced clinicians only have a hit rate of 58.3%. Interestingly, the less experienced clinicians were slightly better at 62.5%. The model clearly beat the experts as [Figure 2](#) shows.

Figure 2. Hit Rate for Classification Accuracy (%)



Source: Research by Dano Leli and Susan Filskov.

The researchers then took their analysis one step further. They wanted to explore what would happen when the experts were armed with a powerful prediction model. A natural hypothesis is that experts combined with models can outperform the stand-alone model. In other words, models represent a floor on performance—to which experts can add incremental value—and not a ceiling. In follow-on tests, the researchers gave the clinicians the output of the model and disclosed that the model has “previously demonstrated high predictive validity in identifying the presence or absence of intellectual deterioration associated with brain damage.” Experienced clinicians significantly improved their accuracy ratio from 58.3% to 75% and the inexperienced clinicians moved from 62.5% to 66.5%. Nonetheless, the experts were still unable to outperform the stand-alone model, which had an 83.3% accuracy rate.

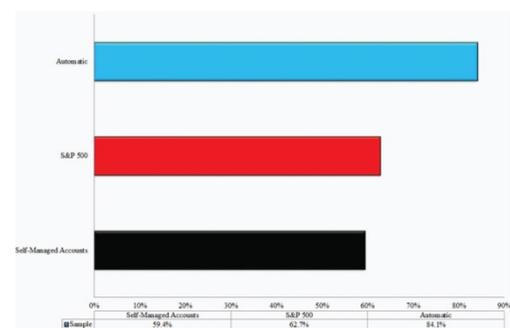
This study suggests that models represent a ceiling on performance, not a floor. Why? Models are built by humans when they are in the System 2 rational mode of thinking. The models are then implemented in a systematic way, devoid of System 1 bias. In contrast, human experts develop an internal model and then implement their thesis in a discretionary way. Unfortunately, discretionary decision-makers are unable to deflect bias from System 1, which detracts from their ability to beat a systematic process.

But Discretionary Investors Beat Simple Models, Right?

One might argue that the clinicians in the Leli and Filskov (1984) study were subpar and perhaps the study design was flawed. Expert stock pickers have access to much better quantitative tools and can develop soft or qualitative information edges. Stock pickers can't possibly be beaten by simple models, can they?

Joel Greenblatt, famous for his bestselling books "[You Can Be a Stock Market Genius Even If You're Not Too Smart](#)" (Simon and Schuster, 1997) and "[The Little Book that Beats the Market](#)" (John Wiley & Sons, 2006) stumbled into a natural experiment he discussed on Morningstar. Joel's firm, Formula Investing, utilizes a simple algorithm that buys firms that rank highly on an average of their cheapness and their quality. A quantitative Warren Buffett, if you will. The firm offers investors separately managed accounts (SMAs) and investors have a choice: They can simply follow the model and purchase all the names suggested by the model, or they get a list of the model's outputs, but have the ability to use their discretion in making individual stock picks. Joel collected data on all their separately managed accounts from May 2009 through April 2011 and tabulated the results. I've presented the results in [Figure 3](#).

Figure 3. Cumulative Return May 2009 to April 2011 (%)



Source: Research by Joel Greenblatt.

The automatic accounts earned a total return of 84.1%, besting the S&P 500 index's 62.7% mark by over 20 percentage points. The self-managed accounts, in which clients were given the model's outputs, but were allowed to pick and choose stocks at their discretion, earned a respectable 59.4%. However, the 59.4% figure was worse than the passive benchmark,

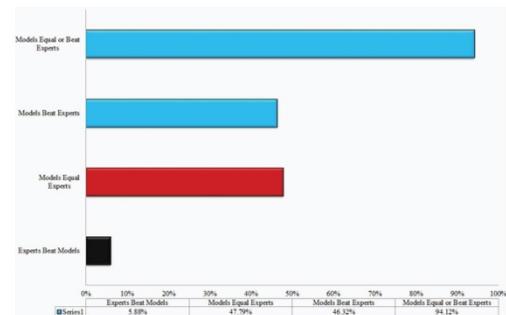
and much worse than the account performance for those that simply “followed the model.” This evidence is similar to the study on brain impairment accuracy: models represent a ceiling on performance, not a floor.

Further Evidence That Systematic Beats Discretionary

Thus far, I’ve presented a formal study published in 1984 and a somewhat ad hoc study of investor behavior. In order to make a more convincing case that models beat experts, we require more analysis. Luckily, one doesn’t have to look that far. There is a sophisticated body of academic literature that has studied the performance of systematic and discretionary decision-making for over 50 years. The breadth and depth of studies are overwhelming, but fortunately, professors William Grove, David Zald, Boyd Lebow, Beth Snitz, and Chad Nelson have performed a meta-analysis (a study of studies) on 136 studies that analyze the accuracy of “actuarial” (i.e., computers/models) vs. “clinical” (i.e., human experts) judgment.

The studies examined by Grove et al in their 2000 Psychological Assessment article included forecast accuracy estimates for just about every category one can imagine. A few examples include college academic performance, magazine advertising sales, success in military training, diagnosis of appendicitis, business failure, suicide attempts, and so forth. Figure 4 summarizes the compiled results of Grove et al’s meta-analysis.

Figure 4. Models vs. Experts Across 136 Academic Studies (Win %)



Source: Research by William Grove, David Zald, Boyd Lebow, Beth Snitz and Chad Nelson.

The study's results are stunning: Models equal or beat experts 94% of the time; experts beat models 6% of the time. The empirical evidence that systematic decision processes meet or exceed discretionary decision-making is overwhelming. An extract from the paper states it best:

“Superiority for mechanical-prediction techniques was consistent, regardless of the judgment task, type of judges, judges’ amounts of experience, or the types of data being combined. Clinical predictions performed relatively less well when predictors included clinical interview data. These data indicate that mechanical predictions of human behaviors are equal or superior to clinical prediction methods for a wide range of circumstances.”

Why Systematic Decision-Making Outperforms

The empirical evidence on the horse race between model-driven decisions and discretionary decision-making is clear, but the implications are unsettling. How is it possible that simple models can consistently beat expert opinion? Experts often have decades of experience, access to qualitative information (e.g., interviews, emotional cues, etc.), and work extremely hard to develop their forecasts. The answer to this conundrum lies with cognitive bias.

I highlight below five key reasons why human experts underperform the forecasts provided by simple models.

1. Same facts; different decisions

Humans, unlike models, can take the same set of facts and come to different conclusions. This can happen for a variety of reasons, but a lack of human consistency is often attributed to anchoring bias, availability bias, representative bias, or something as simple as hunger and fatigue. A computer suffers from none of these ailments—same input, same output.

2. Story-based decisions, not empirical-based decisions

Humans suffer from a tendency to believe in stories, or explanations that fit a fact pattern, but don't bother to examine the empirical evidence. For example, consider the following statement:

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

Is it more likely that Linda is a bank teller or that Linda is a bank teller and is active in the feminist movement? Our gut instinct is to think that it is more likely that Linda is a feminist bank teller, but this line of reasoning is incorrect. It is more likely that Linda is a bank teller because the subset of bank tellers that are also feminist is much smaller than the population of bank tellers as a group. And yet, our brain's love for a consistent story forces us to make poor probability decisions. An empirical-based decision would consider the fact that the bank teller population is larger than the feminist bank teller population and immediately understand that it is more likely that Linda is a bank teller.

3. Overconfidence

Humans are consistently overconfident, and much of their overconfidence is driven by what Kahneman describes as an "illusion of understanding." Overconfidence can be driven by cognitive errors such as hindsight bias—believing events were more predictable than they actually were—and self-attribution bias—attributing good outcomes to skill and poor outcomes to bad luck. Systematic decisions limit these problems. Models don't get emotionally involved and don't have an ego. Therefore, they are unable to get overconfident or overoptimistic—they simply compute.

4. Incorrect modifications to the model outnumber correct modifications

Humans can create modifications to a model that can create value. A popular concept in psychology is the "broken leg theory." Say a human expert develops a model to predict when people will go to the movie theater. The human expert identifies that someone has a broken leg and is able to update the quantitative model and outperform the model. The issue is that humans are unable to limit their "tinkering" with the model. The evidence from academic research suggests that the number of incorrect modifications experts impose on a model outnumbers the number of correct modifications. In summary, human modifications to quantitative models are similar to Lay's potato chips—you can't eat just one.

5. Experts need to feel their efforts are worthwhile

Humans need to fulfill what psychologist Abraham Maslow—famous for developing the human hierarchy of needs—calls our innate need for esteem and self-actualization. Bowing down to the fact that simple models outperform experts directly challenges our ability to achieve goals, gain confidence, and feel a sense of achievement. We want to feel that our efforts are worthwhile, but often put little effort into understanding if our activity actually adds value. Consider the act of banging one's head against the wall for 10 hours a day, seven days a week. Banging your head against the wall involves a lot of activity, but because the outcome of this activity is clearly “bad,” it is easy to know that our concerted efforts are a waste of time. However, what if we are spending 10 hours a day contacting CEOs about the prospects of their companies? Is this intense activity valuable? A lot of investors assume it is, but have they ever systematically reviewed this assumption? Unlikely. And the evidence from the previous research noted suggests that a lot of so-called “value-add” activity performed by experts is equivalent to banging one's head against the wall: The activity is detracting from value, not contributing to value.

Experts are Worthless?

Unfortunately, experts are humans and humans operate with a System 1 emotional brain and a System 2 analytical brain. We are flawed decision-makers that consistently underperform when pitted against systematic decision-makers. As humans, we naturally want our flesh and blood brethren to outperform the cold, calculating computer. Who didn't want Garry Kasparov to beat IBM's super computer Deep Blue in their epic chess matchup? We empathize because we understand how difficult decision-making can be.

And while this article may suggest that experts are worthless, nothing could be further from the truth. Experts are undoubtedly valuable to society. In fact, experts are critical. Experts are in charge of developing the algorithms and systematic models we need to use in our lives to ensure we make accurate and reliable decisions that are unaffected by System 1 thinking. The experts who have devised the algorithms we now use in our daily lives are priceless. These algorithms have saved countless lives in medical settings, enhanced economic wealth and even made the parole process more fair and reliable.

The conclusion one should take away from this body of research is that the world needs experts to design decision-making models, but computers need to be in charge of implementing the models.

→ Wesley R. Gray Ph.D., Ph.D., is the founder and executive managing member of Empiritrage LLC, an SEC-registered investment adviser, and Turnkey Analyst LLC, a firm dedicated to educating and sharing quantitative investment techniques with the general public. He is also an assistant professor of finance at Drexel University's LeBow College of Business.