



The Behavioral Alpha Benchmark

Assessing active portfolio managers based on skill,
not past performance.

Chris Woodcock
Snežana Pejić, Ph.D.
Isaac Kelleher-Unger



Research Authors



Chris Woodcock

Chief Product Officer and Head of Research

chris.woodcock@essentia-analytics.com



Snežana Pejić, Ph.D.

VP of Data Science

snezana.pejic@essentia-analytics.com



Isaac Kelleher-Unger

Data Scientist

issac@essentia-analytics.com

About Essentia

Essentia Analytics is a leading provider of behavioral data analytics and consulting for professional investors. Led by a team of experts in investment management, technology, and behavioral science, Essentia combines next-generation data analytics technology with human coaching to help active fund managers capture performance that was previously being lost to biases or other common decision-making deficiencies.

Contact us to find out more:

✉ info@essentia-analytics.com

🌐 www.essentia-analytics.com



Foreword

*It's time to consider the **quality** of a portfolio manager's decisions — not just near-term performance.*

In the 2011 film *Moneyball* — adapted from the Michael Lewis book of the same name — Oakland A's General Manager Billy Beane, played by Brad Pitt, sits with his scouts as they try to figure out their roster for the upcoming season. Beane hears a lot of idle chatter, including how good players look in a pair of jeans. He decides that there must be a better way.



He's assessing the factors that matter, not chasing past performance.

The investment management industry has a similar measurement challenge. On the one hand, the footnotes to every investment ad refer to the fact that historical performance is no guarantee of future returns. On the other, assets have very predictably chased performance since there has been investing and **continue to do so today**.

Performance is not a measure of investment skill. It is a measure of outcome. It is a *function* of skill (at least where professional investors are involved) — but it's also a function of luck.

In allocating capital to asset managers, investors should be paying for skill, not luck. So why do we end up basing manager selection decisions on one- to five-year performance, which is almost always affected by — if not dominated by — luck? It's because, similar to Beane's scouting reports, that's all that has been available — until now.

Essentia Analytics has been helping active equity portfolio managers make measurably better decisions for nearly a decade. Using high-granularity investment activity data rather than fund performance figures, we've honed in on a methodology for skill assessment and visualization that has been used by hundreds of equity portfolio managers around the world: the **Essentia Behavioral Alpha® Benchmark**.

You may be thinking, “Great, so is this measure of past skill actually predictive of future performance?” Of course you are: we humans just can’t help but look for something to be predictive of future performance!

Think of it this way: a “good” decision is one that produces a more positive outcome — on average — than would have been achieved by chance. If an investor consistently makes more good decisions than bad, and their good decisions consistently add more value than their bad decisions destroy, then over time this investor should perform better than someone for whom this is not the case. Whatever the influence of luck on their historical returns, this manager has shown evidence of possessing true skill.

A skilled investor can have a bad year (**or even ten!**) performance-wise. By using a skill lens, a portfolio manager, his or her firm, or indeed, the portfolio’s end investors, can gain insight into whether poor (or strong) performance is due to luck or skill.

With this knowledge, the manager unlocks the means to continuous improvement. But the first step is measurement: the creation of a baseline.

After nearly a decade of sharing these insights with our clients individually, I’m thrilled to be introducing the Behavioral Alpha Benchmark as a skills-based manager assessment and comparison tool. With it, the investment industry can finally begin its own *Moneyball* era.



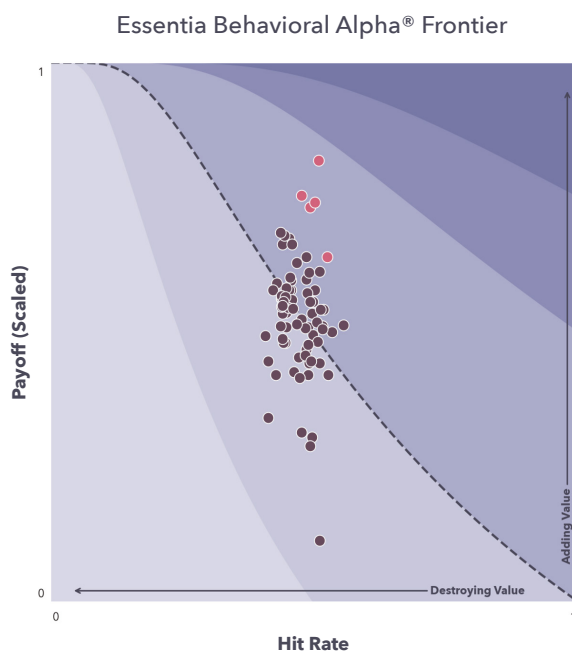
Clare Flynn Levy
CEO, Essentia Analytics



The Behavioral Alpha Benchmark

Summary

- Using our proprietary database of manager behavior in 195+ portfolios over 18 years, we have constructed a new evaluation system to assess and compare a manager's skill — **while mitigating the confounding effects of luck** on the manager's past performance. We call this system the Behavioral Alpha® Benchmark.
- We measured the demonstrated skill of 76 managers in **seven key decision types** over the past 36 months: stock picking, entry timing, sizing, scaling in, size adjusting, scaling out, and exit timing. We isolated each decision type and measured its impact on the portfolio using methods we have developed in conjunction with our clients over the past eight years.
- Managers were scored on an Essentia Behavioral Alpha Frontier (EBAF) that considers the hit rate (percentage of decisions that added value) and payoff (net amount of value added) for each decision type.



The five highest-scoring managers, represented by the magenta dots, are those towards the upper-right on this frontier diagram — furthest away from 0,0.

- Among our findings:
 - **Most managers got their decisions wrong most of the time.** When we look across all decision types, only 18% of the managers in our study had a hit rate above 50% — and the “most-right” manager got it right just 55% of the time.
 - These **low hit rates were largely — but not entirely — compensated for by higher payoffs:** when they got a decision right, most managers (68%) added more value than they lost when they got it wrong. As a result, overall, the **managers added value almost half the time** (43.4%) with the decisions they made.
 - **Most managers demonstrated strong skill in selecting good stocks** for their portfolios: 58% added value through their stock picking decisions. In contrast, managers typically destroyed value through sizing — only 38% added value. Amongst the other skills, managers were notably poor at scaling out of positions.
 - **The five managers who added the most value** due to their demonstrated skill in these seven key decision types over the last three years have agreed to be identified¹ and are:

Inaugural Behavioral Alpha® Award Winners



Rank	Manager(s)	Portfolio
1	Vishal Gupta	Morgan Stanley Investment Management - Emerging Markets Leaders
2	Mark Denham	Carmignac - Portfolio Grande Europe
3	Jonathan Good	Baird - Small/Mid Cap Growth
4	Xavier Hovasse Haiyan Li-Labbé	Carmignac - Emergents
5	Martin Walker	Invesco - UK Opportunities Fund (UK)

These five managers had the top scores in our Essentia Behavioral Alpha Benchmark as of March 31, 2022.

¹ Essentia will never reveal the details of its clients to anyone without their express prior permission.



Introduction

Luck and Skill

The ability to prove whether a manager is skilled, lucky, both, or neither has been a sort of holy grail for allocators since the earliest days of investment management.

That's because there is a dearth of tools that are able to focus solely on decision-making — i.e., factors within a manager's control. In the same vein, academic research covering luck and skill in investment decision-making uses ex-post portfolio returns rather than focusing on the set of decisions a manager chooses to make.

Managers are typically assessed based on the relative performance of their fund against a benchmark. And yet, where managers are shown to outperform, even over long periods, research repeatedly shows that performance tends to revert to the mean [Jensen (1968), Malkiel (1995), Fama & French (2010)]. This leads to the popular refrain that "past performance is not indicative of future results," and investors are left with a framework that will never support them in their task of separating the skilled managers from the lucky ones.

Moreover, performance is not a measure of manager skill, but the product of skill and luck, with no way to determine the influence of each. Again, the research clearly shows that even known skilled managers would be expected to underperform for extended periods [Kaplan & Kowara (2019)].

Meanwhile, index benchmarks, while understandable, are only one yardstick; measuring performance relative to one offers a narrow and uncertain perspective on the actual skill of the manager, and can produce perverse incentives detrimental to the real returns to clients [SCM Direct (2015), ESMA (2016)].

With the Essentia Behavioral Alpha Benchmark, we seek to disaggregate skill from luck and present an evaluative framework based on the quality of decisions made — not on overall portfolio performance. Using our proprietary data set consisting of the reconciled trades and holdings from 65,000 equity investment episodes across 195 portfolios, we look at the elements of a manager's role over which they have control (the actual decisions they make) and identify patterns within them.

We operate on the basis that, a priori, if a manager can make better decisions, they will improve as a professional investor, however that is measured. We incorporate the standard approach of benchmark-relative returns, but also seek to look beyond that into the opportunity space of the manager and ask, "How well did this manager make decisions within the set of all decisions they could have made?"

Even known skilled managers would be expected to underperform for extended periods.

This is an opportunity space for the manager, and tuning these decisions can lead to better outcomes.

Decisions, Decisions, Decisions

Investment managers are called upon to make a number of decisions with regard to the positions in their portfolios: what to buy, how much to buy, when to buy it, how quickly, how to adjust the amount as new information becomes available, when to sell, and how quickly to do so. Our analysis is based on these seven decision types, which we categorize as: **stock picking, entry timing, sizing, scaling in, size adjusting, scaling out, and exit timing.**

Two of these, when also judged in relative terms, are correlated with benchmark-relative performance: stock picking and (average) sizing. We call these “performance generators,” and they tend to align with traditional measures of manager evaluation.

However, the other five decision types are also important components in the lifecycle of an investment. Each can have a clear impact on the outcome, and they are all in the power of the manager to influence. Yet the manager’s skill at these decisions may not be captured in standard benchmark-relative performance: at least not in a way that is easy to decipher. This is an opportunity space for the manager, and tuning these decisions can lead to better outcomes.

Moreover, for anyone trying to determine the elusive measure of manager skill, the Behavioral Alpha Benchmark offers an alternative framework for discussion between the investor and the manager. Importantly, it offers the manager information *that they can do something about* rather than chasing performance over which they ultimately exert little control. The asset owner or allocator can then monitor how a manager is doing on these metrics and whether that manager is improving over time.

One question remains: how do we calibrate the skill levels of managers? What is “good,” and in comparison to what? It is not realistic for a manager to be highly skilled in all types of decisions, and not all fund management styles require the same set of skills to successfully invest.

Which brings us to one of the most compelling aspects of this research: **peer comparison.** By plotting many managers on an Essentia Behavioral Alpha Frontier (EBAF), the relative skill levels of each manager can be quickly assessed versus one another in a simple, clear illustration. We believe the potential value of this technique as a skills-based manager evaluation model is enormous.

We have identified the top five managers among those we evaluated across decision types and in the aggregate; these are discussed in detail in the Results section below.



Methodology

Decision Types

Investment managers are called upon to make a number of different types of decisions on an ongoing basis. In this case, we identify seven decision types: stock picking, entry timing, sizing, scaling in, size adjusting, scaling out, and exit timing.

Each of these decisions has an associated measure, which is referred to as the “value added” metric. For each type of decision, the value added is a metric that measures the efficacy of that type of decision using methods we have developed in conjunction with our clients over the past nine years. It involves comparing the actual decision with a reasonable, baseline alternative.

For each decision, we split the target metric into two further components: hit rate and payoff. The definitions of hit rate and payoff are described below. The use of hit rate and payoff gives the managers who are subject to these analyses a more practical understanding of why they are seeing the results they are seeing — *and how they can improve them*.

Hit Rate & Payoff

Hit Rate refers to the proportion of investment *episodes* (that is, the period comprising the first purchase to last sale of a given holding) that are winners. Winners are episodes with a positive value added, y , over the time period.

$$\textit{Hit Rate} = \frac{\textit{count}(y|y \geq 0)}{\textit{count}(y)}$$

Payoff refers to the average value added, y , of winning episodes divided by the average value added of losing episodes.

$$\textit{Payoff} = \frac{\textit{mean}(y|y \geq 0)}{\textit{abs}(\textit{mean}(y|y < 0))}$$

For hit rate, 50% is considered neutral or equivalent to chance (getting the same number of decisions right as wrong); while for payoff, 100% is considered neutral or chance (the magnitude of outcomes of good decisions is equal to the magnitude of outcomes of bad decisions, on average).

The Essentia Behavioral Alpha® Frontier

For each decision type, as well as an aggregate of all decision types, the results are presented using the Essentia Behavioral Alpha Frontier. The EBAF is a helpful visual tool for displaying the hit rate and payoff of a collection of decisions and for comparing data points.

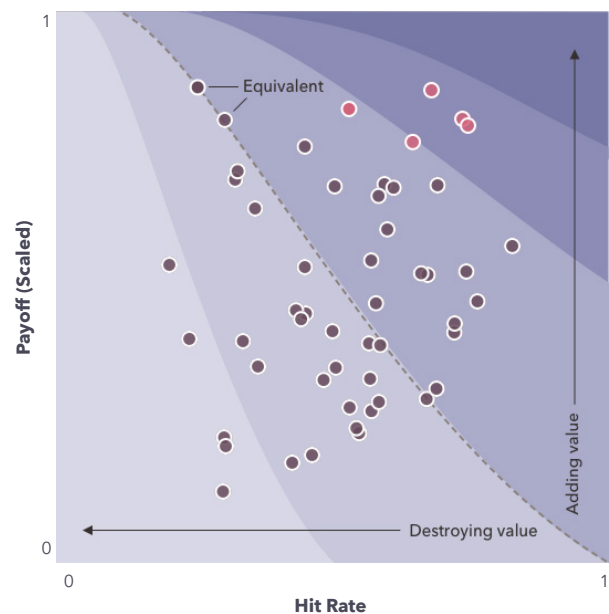
Hit rate varies between 0 and 1 and is plotted as the x axis on the frontier. Payoff varies between 0 and $+\infty$, centered on 1. We can smoothly transform the payoff measure onto the (0,1) domain, centered on 0.5, using an appropriate logarithmic function. These transformed values, labeled Payoff (Scaled), are then plotted as the y axis on the frontier.²

The EBAF is designed to visually equate the varying combinations of hit rate and payoff that occur between managers. Each mark on the EBAF represents the decisions of one portfolio. The further up and to the right (i.e., the higher the payoff and hit rate), the better.

The frontier curves on the visual — the borders between different colored regions — are indifference curves on the relationship between hit rate and payoff. That is to say, within this analysis, you would be indifferent between any two managers that both sat on the same curve (see illustration).

In other words, any mark plotted on the frontier has the same overall value added, or “edge,” as all other points lying on that same curve.

It should be noted that the particular curves drawn on the frontier chart are intended to be illustrative: they are drawn to emphasize the approach described above and do not necessarily have any significance in and of themselves. There are, of course, infinitely many curves that could have been chosen to highlight. The exception is the thicker dotted curve between the second and third colored regions from the left, which is drawn to emphasise the “neutral” — zero value added — frontier. It encompasses the specific point (0.5,0.5) on the frontier.



² The transformation is monotone increasing, meaning that between two points, higher payoff will always result in a higher payoff (scaled) value.



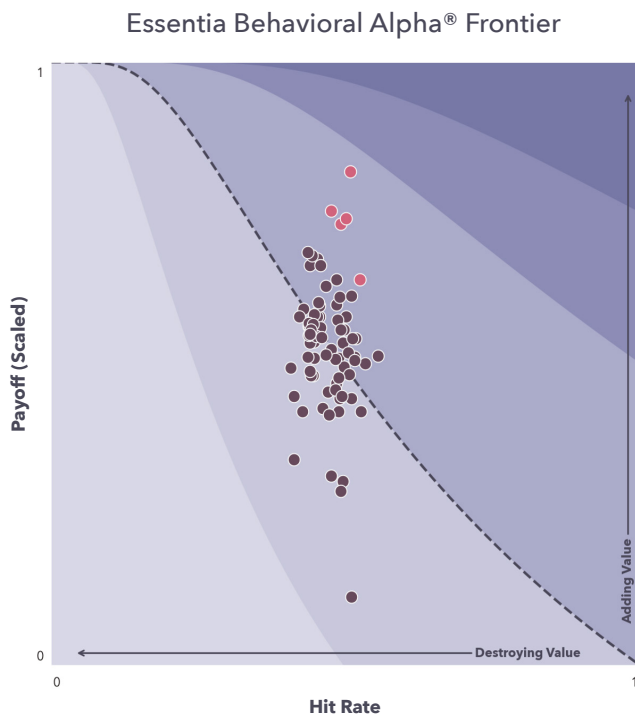
Portfolios

Essentia's proprietary analytical dataset comprises 50,000 investment episodes, 276,000 active decisions, and 166 million rows of data derived from the reconciled trades and holdings of over 195 long only portfolios. The study spanned the 36-month period ending in the first calendar quarter of 2022 (i.e., from Q2 2019 to Q1 2022 inclusive). To be included in the final results, a portfolio had to have data spanning at least two-thirds of the period and no more than the initial or final two quarters absent, which left 76 portfolios and 65,000 decisions.

All data on all managers was anonymized and any identifying information was (and always is) kept strictly confidential. We do not know the identities of the managers in the results tables; we have only identified the five overall winners (in magenta below), and obtained the permission of each before naming them.

Results

All Decisions



When considering all decisions together, we observe that 43.4% of portfolio managers in the analysis added value overall compared with chance (a 50% hit rate and a 100% payoff). Just 18% of portfolio managers had an overall hit rate greater than 50%, while 69% of managers had payoffs above 100%.

The figure demonstrates that managers' decision hit rates fall within a narrow band, whereas payoffs have a wider distribution. Indeed, hit rate varied from 41% to 56%, while payoff varied between 17% and 248%.

The top five managers, colored magenta in the figure, are listed below.

Rank	Manager(s)	Portfolio
1	Vishal Gupta	Morgan Stanley Investment Management - Emerging Markets Leaders
2	Mark Denham	Carmignac - Portfolio Grande Europe
3	Jonathan Good	Baird - Small/Mid Cap Growth
4	Xavier Hovasse Haiyan Li-Labbé	Carmignac - Emergents
5	Martin Walker	Invesco - UK Opportunities Fund (UK)

Summary Table

	Percent of portfolios adding value	Percent of portfolios hit rate positive	Min. Hit rate	Max. Hit rate	Percent of portfolios payoff positive	Min. Payoff	Max. Payoff
	43.4%	18.4%	40.7%	55.6%	68.4%	17.4%	247.5%
entry	56.6%	65.8%	34.9%	78.3%	46.1%	19.8%	300.3%
exit	50.0%	28.9%	18.2%	66.7%	63.2%	47.4%	320.3%
scaling in	44.7%	47.4%	25.0%	100.0%	46.1%	32.7%	258.0%
scaling out	26.3%	25.0%	0.0%	75.0%	22.4%	0.5%	409.7%
size adjusting	51.3%	47.4%	27.8%	78.6%	48.7%	31.3%	557.8%
sizing	38.2%	21.1%	25.5%	63.6%	60.5%	41.2%	187.5%
stock picking	57.9%	13.2%	18.2%	58.1%	81.6%	30.8%	258.6%



Individual Skills & Decisions

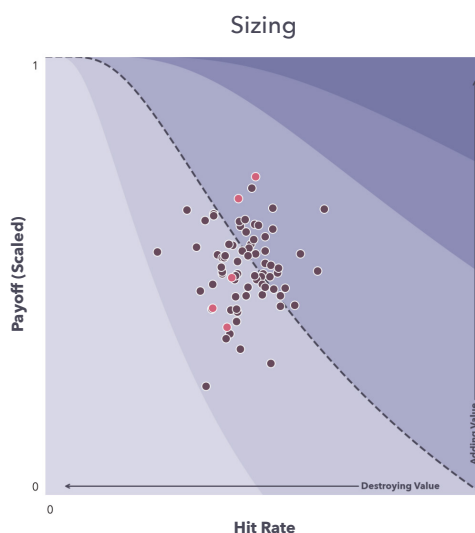
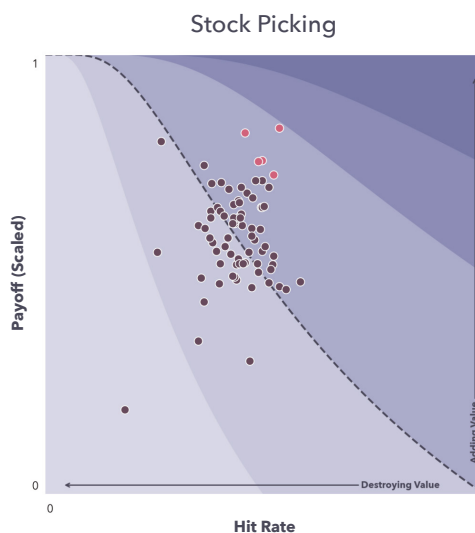
In all the following figures, the same top five managers overall are again colored magenta.

Performance Generators: Stock Picking and Sizing

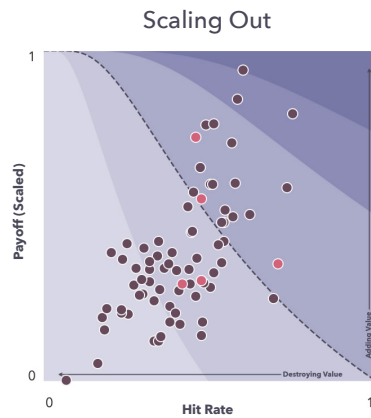
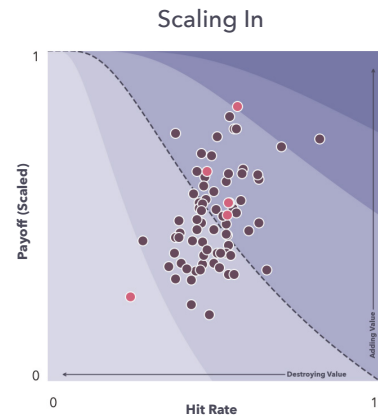
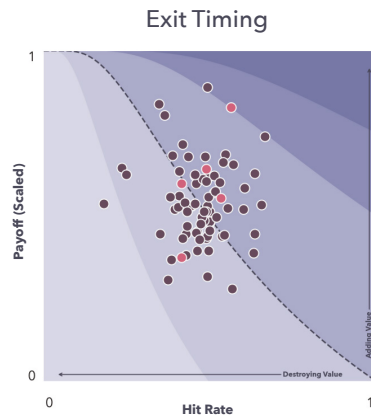
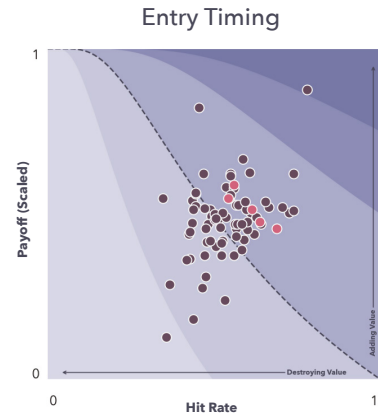
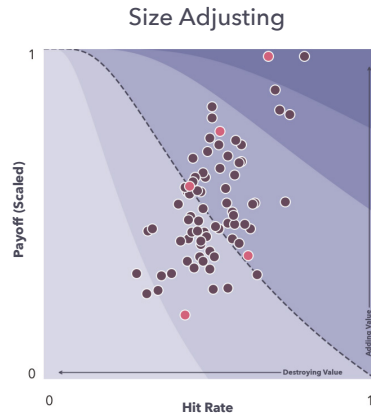
In general, managers proved adept at stock picking. 58% of portfolio managers in the analysis added value overall through their picking decisions.

Just 13% of portfolio managers got their picking decisions right more often than wrong (i.e., had a hit rate > 50%), but 82% of managers had positive payoffs >100% (i.e., their good picks outperformed their bad ones, all else being equal).

When it comes to sizing, only 38% of managers added value overall. 21% had a positive hit rate and 60% had a positive payoff. In other words, consistent with other literature, the majority of portfolio managers are shown here to be poor at sizing decisions. Even those that add value tend to get most sizing decisions wrong (hit rate < 50%) and make up for it by getting their big sizing decisions more right than wrong (high payoffs).



Opportunity Space: Size Adjusting, Entry Timing, Exit Timing, Scaling In, and Scaling Out



Size Adjusting looks at adding and trimming decisions. A small majority of portfolio managers added value through size adjusting: 51%. 47% had a positive hit rate, while 49% had a positive payoff.

57% of managers added value through their entry timing, whereas 50% of managers added value through their exit timing. Managers appear to be more adept at scaling in (45% adding value) than scaling out (26% adding value). Getting out of positions is a sore spot for most fund managers.



Discussion

If developing a reliable framework for assessing manager skill was a straightforward matter, it would have been done successfully already. While we think the lens of decision quality is a meaningful addition to the subject, the associated metrics we have developed are by no means the final word. By releasing this paper into the public domain, we look forward to stimulating healthy debate about how to best to measure decision efficacy. We will continue to evolve the way we measure decisions based on the feedback we receive.

In order to compare “apples to apples,” we have had to make some informed, yet arbitrary, decisions with regard to some of the parameters in the decision metrics, based on the averages across our database. For example, we consider a window after the entry point to determine entry timing efficacy. The size of that window has a fixed value of one month from the entry point because, on average, that is the most relevant value across the universe. Nevertheless, it may not be the optimum parameter for assessment for a given manager. These parameters are fully flexible when it comes to deeper individual manager analysis.

The universe at our disposal consists of clients and former clients of Essentia Analytics, hence the sample reported here was self-selecting — Essentia’s service appeals to managers who are continuous improvers by nature. Over time, as this approach gains traction, we hope to broaden the universe beyond those that have been early adopters of behavioral analytics.

Manager styles, such as growth and value, are important when considering a diversified set of allocations. In addition, not all fund management styles require the same set of skills to successfully invest. It is our intention to segment future analysis by such peer groupings once the dataset has grown sufficiently large. Peer groupings will also lead to more homogeneous risk profiles within groups, mitigating the heterogeneity that exists in the current data set.

Market regimes undoubtedly influence the measurement of decisions, although to what extent has not yet been investigated. It is one reason to use a peer comparison, which in some sense normalizes for market regime — especially by considering managers in the context of others of a similar style, and relative to the appropriate benchmark. In addition, it is not obvious that market regime would in fact influence, for example, an assessment of sizing decisions relative to an equally-weighted portfolio.

Finally, style drift is one of the key risks that allocators must monitor in order to maintain the balance of their portfolio. It may be possible to use the Essentia Behavioral Alpha Frontier diagram to monitor style drift: is the decision efficacy of a particular manager so far removed from their peers that it looks like they are playing a different game? In the future, we look to incorporate this aspect into our analysis more formally by grouping managers not on their own self-identification, but by the characteristics of their decision making. This could offer allocators a tool for monitoring style drift: does your favorite growth manager make decisions in a way that looks a lot like a bunch of value managers?

Is the decision efficacy of a particular manager so far removed from their peers that it looks like they are playing a different game?

The Behavioral Alpha Benchmark has the potential to “raise all boats” for investors and managers alike.

Conclusion

It is our sincere wish that this new skills-based manager assessment model — the Essentia Behavioral Alpha Benchmark — will usher in a new era of manager assessment in the investment management industry: one that recognizes demonstrated proficiency in a range of investment skills and eliminates the confounding effect of luck on a manager’s reported performance.

We hope that portfolio managers will use this approach — as our current clients do — to understand their strengths and weaknesses and track progress over time in a data-driven cycle of continuous improvement. And we hope that investors and allocators will use it in evaluating and selecting managers for the assets they steward.

This would represent a sea-change in the industry. It has the potential to “raise all boats” for investors and managers alike, allow highly-skilled active managers to demonstrate their value over passive investing, and empower those managers who have not yet achieved a high skill level to do so.

To the five top-scoring managers in our peer comparison, we offer our congratulations. Their strong showing across the skills we measure is an impressive achievement, which we believe demonstrates real value-add as active managers of their investors’ assets. We look forward to their continued success, and we are eager to see more managers in our universe move steadily toward the upper-right quadrant of the EBAF in the quarters to come.

References

- Jensen, M.C. (1968), THE PERFORMANCE OF MUTUAL FUNDS IN THE PERIOD 1945–1964. *The Journal of Finance*, 23(2), 389-416.
- Malkiel, B.G. (1995), Returns from Investing in Equity Mutual Funds 1971 to 1991. *The Journal of Finance*, 50(2), 549-572.
- Fama, E.F., French, K.R. (2010), Luck versus Skill in the Cross-Section of Mutual Fund Returns. *The Journal of Finance*, 65(5), 1915-1947.
- Kaplan, P., Kowara, M. (2019). Are Relative Performance Measures Useless? *The Journal of Investing*, 28(4), 83-93.
- Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: a query theory of value construction. *Journal of experimental psychology: Learning, memory, and cognition*, 33(3), 461.
- ESMA (2016), Public Statement on Potential Closet Indexing
- SCMDirect (2015), Closet Indexation: The UK Epidemic Continues



DELIVERING BEHAVIORAL ALPHA®

Essentia Analytics is an award-winning fintech company that provides behavioral analytics services to professional investors.

Our proprietary research directly fuels our work with clients, delivering critical insights that can guide them toward their best practices, helping them turn good intentions into good habits. Essentia’s skills-based manager evaluation system — the Behavioral Alpha Benchmark — is a case in point.

The Benchmark methodology looks beyond historical returns — and the effects of luck — to instead **measure a client’s demonstrated skill across a range of investment decision types**. Active fund managers then have clarity on where to improve and how to demonstrate their value over passive index funds.

*Learn how Essentia can help unlock the **behavioral alpha** - the excess return that results from mitigating one’s biases - that’s hidden in your investment decision-making process.*

WWW.ESENTIA-ANALYTICS.COM
INFO@ESENTIA-ANALYTICS.COM



Know thyself.

Socrates

