

The Implications of Style Analysis on Mutual Fund Performance Evaluation

Starring: TIC-TAC-TOE

Also Featuring: Pascal, God, Chess, and War Games

K

Keynote Speech by

John C. Bogle, Founder and Chairman, The Vanguard Group

before the Morningstar Investment Conference

Chicago, Illinois

June 13, 1997

Consider the child's game tic-tac-toe. There is simply no way to win, even if a genius is playing against an opponent of only moderate intelligence. Each player, in turn, simply blocks the other player's previous move. (Of course, if one player is dull-witted or bereft of the power of concentration, a loss is easily accomplished.) In short, as a game that cannot be won, only lost, tic-tac-toe is the ultimate loser's game.

Exhibit I: Tic-Tac-Toe

X	O	X
X	O	O
O	X	X

Curiously enough, the new Morningstar Category Rating System is played on a field with a pattern identical to tic-tac-toe. Because of this similarity, the nine-box system for analyzing fund investment styles raises, perhaps inadvertently, the question: Does the search for fund performance resemble the search for three Xs (or Os) in a row in a child's game? Put another way, if no one can win consistently when nearly all participants have at least average skill, is not fund selection, too, a loser's game?

This analogy quickly brings to mind one of the truly seminal articles about the challenges of investment management in increasingly efficient financial markets. Written by Charles D. Ellis, founder of Greenwich Research Associates, and published in the July/August 1975 issue of *The Financial Analysts Journal*, it was called, of all things, "The Loser's Game." In his article, Mr. Ellis observed:

"The investment management business is built upon a simple and basic premise: professional managers can beat the market. That premise appears to be false. The ultimate outcome (of the game) is determined by who can lose the fewest points, not who can win the most. Money management has been transformed from a Winner's Game to a Loser's Game."

When the article was written—now more than two decades ago—the Standard & Poor's 500 Index was virtually the only standard used by institutions to measure market returns. (Even it wasn't used very often!) And in those ancient days, the portfolios of most institutional managers—and most mutual funds—were dominated by a blended list of the large cap stocks in the Index. In this modern day and age, however,

other styles have developed, some with extreme emphasis on value or growth, or on medium or small cap stocks. Given variations in investment performance among these styles (at least over interim periods) and in volatility risk (in all periods), it seems only good judgment to compare “like with like.”

To date, most mutual fund performance evaluations have been fairly simplistic: how has a fund performed relative to “the market”? The Standard & Poor’s 500 Stock Index is usually used as a proxy for the market, despite the fact that it accounts for only about 70% of the capitalization of the U.S. stock market and is dominated by corporations with gigantic market capitalizations. (Its largest 25 stocks account, on average, for 1% of the entire market; the 6500 “non-500” stocks in the market have an average weight of 4/1000 of 1%.) But today, many funds resemble “the market” only tangentially.

So, under the concept of style analysis, a mutual fund is compared not with “the market,” but with its peers following a similar investment style. For many years, this analysis was used by institutional investors *via* a box with a vertical axis running from large to small market capitalization, and a horizontal axis running from value to growth (usually based on ratios of market-to-book value or price-to-earnings). Each account got an “X” somewhere along each axis. It wasn’t very complicated, but neither did it make it very simple to evaluate comparative performance.

Exhibit II: Institutional Style Box

	Value	Growth
Large		X
Small		

Enter Morningstar. Its contribution—and it is, as advertised, “a more intelligent way to select and monitor mutual funds”—was to divide the simple box punctuated with dots into a nine-box matrix—just like tic-tac-toe—where each fund is, in effect, forced into one of nine boxes: large, medium, or small capitalization on the vertical axis; value, blend (mixed), or growth on the horizontal axis. The beauty of this system is that it immediately becomes possible to quantify the vital statistics of each fund’s performance relative to its peers. Large cap growth funds are compared with other large cap growth funds; small cap value funds are compared with other small cap value funds; and so on. And, under the Morningstar system, each fund then gets a Category Rating, ranging from “one” (lowest 10%) to “five” (highest 10%)—both, therefore, are very tough leagues to break into. Here is the current mix of the 741 equity funds with five-year records followed by Morningstar, which makes their detailed records remarkably accessible through its incredible Principia data base. This is the first of nine tic-tac-toe boxes I’ll present today:

Exhibit III: Number of Funds (741)

	Value	Blend	Growth
Large	100	211	58
Medium	54	84	90
Small	52	32	60

While this analysis is important, I cannot emphasize sufficiently the importance of achieving superior total returns in the long run, irrespective of style or category. If a given style group, say, small cap value, fails to outpace “the market” over twenty years, for example, it would seem counterintuitive to give much credit to a manager who created such a “product” (as they say) for clients, even if it outpaces other small cap value funds. By the same token, if a large cap blended fund (in effect, paralleling the Standard & Poor’s 500 Index) outpaced the market for twenty years, it ought to get some credit, even if it fell slightly short of its peers. (I’ll leave aside for the moment the critical issues of: (a) why the small cap fund didn’t beat the Index; and (b) whether the past two decades has any relevance for the next two decades.)

But the Morningstar Category Rating System does accurately reflect general differences or similarities in return among the various categories. In the past five calendar years, interestingly, similarities were in the driver’s seat. Only large cap growth funds (annual returns averaging about +12%) strayed from the +13% to +15% returns of all the other groups. Returns for each of the nine categories are shown in Exhibit IV.

Exhibit IV: 5-Year Return (%)

	Value	Blend	Growth
Large	13.8	13.2	11.9
Medium	14.2	14.0	13.3
Small	15.1	15.1	15.0

Differences in risk, however, are much more sharply defined among the nine categories. Using standard deviation as a proxy for risk—although it is really something slightly different: a measure of volatility—the variability of returns has ranged from a low of 9.8% (large cap value) to a high of nearly double that figure, 18.7% (small cap growth). Curiously, then, despite the identity in *return* among the three small cap categories, the differences in *risk* were extreme (11.6% for value and 18.7% for growth). Exhibit V shows these sharp differences in risk:

Exhibit V: Standard Deviation (%)

	Value	Blend	Growth
Large	9.8	9.9	12.0
Medium	9.9	11.3	15.8
Small	11.6	13.9	18.7

These differences in risk in the face of the similarity of returns give rise, of course, to large differences in risk-adjusted returns. Here, I’ll use the Sharpe Ratio, developed by Nobel Laureate William F. Sharpe, which in effect calculates fund rates of return in excess of the risk-free rate per unit of risk-- more accurately volatility, as measured by standard deviation. (Morningstar publishes the *three-year* Sharpe Ratios, but on a relative basis these numbers correspond broadly with the *five-year* risk-adjusted return ratings.)

As Exhibit VI shows, the differences in *risk-adjusted return ratings* are also extremely wide—in fact, exactly 100%, from 120 for large blend funds to 60 for small growth funds. To make the point clear, if two funds had an equal volatility of 10%, a fund with a 120 risk-return ratio would return 16%, while a fund with a risk-return ratio of 60 would return 10% (assuming a risk-free rate of 4%). *This is hardly a trivial difference.*

The risk-adjusted return ratings among the nine boxes vary widely, largely reflecting the differences in the risks of the nine market segments during the period.

Exhibit VI: Risk-Adjusted Return Ratings

	Value	Blend	Growth
Large	117	120	101
Medium	105	98	65
Small	91	86	60

Given these variations, it seems to me, it makes consummate good sense to evaluate each fund’s returns on a category basis, if we are looking to appraise a manager’s abilities to use the tools he or she has chosen to use. In effect, what results is a peer group that, while by no means perfect, is as good as is available today.

Now let’s take a look at what happens when we begin to evaluate equity funds on the basis of their investment styles, as measured by their Morningstar categories. I’m going to use returns and standard deviations of return for the past five calendar years for this analysis in performance appraisal, and I’ll try to answer the questions of what conclusions flow from style analysis. My first example is the Large Capitalization Blend Group—mutual funds investing in giant companies with both value and growth characteristics. This category is composed of more than twice as many funds as any other group (211 of 741 funds analyzed over the five-year period), and some 40% of the assets of all domestic equity funds (\$450 billion of \$1.2 trillion of equity assets in the Principia data base), so it provides a solid platform on which to begin the analysis. Here is how the performance looks, ranking funds into four quartiles based on total returns for the period:

Exhibit VII: Large Capitalization Blend Funds

	Ranked by Return		
Return	5-Year	5-Year	Risk-Adjusted
<u>Quartile</u>	<u>Total Return</u>	<u>Risk</u>	<u>Rating</u>
First (highest)	15.9%	10.1%	141
Second	14.1	9.8	128
Third	12.6	9.7	114
Fourth (lowest)	10.2	10.0	95
Average	13.2%	9.9%	120

We can see that even though returns rise, risk in this category remains virtually unchanged, with standard deviation remaining remarkably constant over the quartiles. Obvious result: the risk-adjusted return ratio increases by the same magnitude as the total return, from a ratio of 95 to 141—fully 46 points from the lowest to the highest. This 50% difference, dare I say, is “statistically significant.” As it happens, this

outcome for this large cap, blend (middle-of-the-road) fund category is typical. Seven of the nine categories (the exceptions are small cap value and medium cap growth) have fairly steady risk scores, whether returns are high or low. Hence, the top *risk-adjusted* ratings are consistently earned by the funds with the highest total returns.

The previous table, of course, is simply a recounting of the past. But as I looked at the data, I wondered whether there was an element that might have been used to determine *in advance* which large cap blended funds might most likely fall into the various quartiles. Of course, my first thought (this will hardly astonish you!) was whether relative fund operating expenses would not give an investor some forecasting ability. So, I divided the funds into cost quartiles, with funds with the lowest expense ratios comprising the first quartile, and the funds with the highest ratios comprising the fourth quartile. I don't think it will surprise anyone who has seriously studied investment returns—either from a theoretical academic basis or from pragmatic industry experience—that *costs matter*.

In fact, the funds in the group with the lowest expense ratios had the highest net returns. At the same time, they assumed an identical level of risk (volatility), and therefore provided distinctly higher risk-adjusted returns. Here are the same data that I presented earlier, but arrayed by expense quartiles.

Exhibit VIII: Large Capitalization Blend Funds

	Ranked by Cost		
Cost Quartile	5-Year Total Return	5-Year Risk	Risk-Adjusted Rating
First (lowest)	14.2%	9.8%	136
Second	13.8	9.9	125
Third	12.5	9.9	113
Fourth (highest)	12.3	9.9	105
Average	13.2%	9.9%	120

Now, we seem to be on to something important. *With risk astonishingly constant, high returns are directly associated with low costs.* The risk-adjusted ratings provided by the lowest expense funds, in the large cap blend group, at 136 were more than 13% above the average of 120; the returns provided by the highest expense funds were 13% below average--a 26% spread. Clearly, expenses are a compelling factor.

Given this finding, I decided to add the expense ratios to the *net* returns to see how similar the *gross* returns would have been. Again, perhaps unsurprisingly, the *gross* returns in each quartile were substantially the same.

Exhibit IX: Large Capitalization Blend Funds

Net Returns vs. Gross Returns(%)

Cost Quartile	5-Year Net Return	Expense Ratio	5-Year Gross Return
First (lowest)	14.2	0.50	14.7
Second	13.8	0.90	14.7
Third	12.5	1.10	13.6
Fourth (highest)	12.3	1.70	14.0
Average	13.2	1.00	14.3

This example clearly confirms the theory that cost is a key determinant of relative total return.

Now the question is: do these relationships between return and risk prevail across the style boxes? The answer: they do, and they do so remarkably well. This next table shows the percentage difference between the risk-adjusted returns of the first quartile (lowest expense) funds and the fourth quartile (highest expense) funds, using the average risk-adjusted rating *for that style box* as the standard. For example, in the large cap blend category, low expense funds had risk-adjusted returns 13% greater (113) and high expense funds 13% lower (87) than the average.

Exhibit X: Relative Risk-Adjusted Return Ratings

<u>Low-Expense Funds</u>				<u>High-Expense Funds</u>			
	Value	Blend	Growth		Value	Blend	Growth
Large	+16	+16	+8	Large	-36	-15	-9
Medium	+23	+7	+19	Medium	-31	-13	-13
Small	+5	0	+7	Small	-12	+6	-9

The consistency of the relative risk-adjusted ratings in each box is striking—even shocking. In eight of the nine boxes, the spreads were strongly positive for the low-cost funds, strongly negative for the high-cost funds. The only exception was in the small cap blend category, where the lower expense funds provided an average performance, while the higher expense group provided a positive result. But this is a bit of a statistical anomaly. There were only 32 funds in this group—eight in each quartile. If we simply divide the group in half, the 16 lower-cost funds would have shown +8, the 16 higher expense funds -8, quite consistent with the other eight categories.

The strong implication—if not the virtual certainty—of these figures is that, in each of the nine style boxes, an investor who doesn’t seriously consider limiting selections to funds in the low expense group, and eschewing funds in the high expense group is someone who should take off the blinders—perhaps even a bit of a fool.

The mutual fund world is one in which forecasting relative (to say nothing of absolute) returns based on past performance is indeed a fool’s game—in general, a zero sum relative game. And past performance is all we have . . . almost. But we do have cost data, for those willing to look at it. And we now know—I would argue, as a certainty—that costs matter. It matters for equity funds in the aggregate; far more, to be sure, for bond funds; and infinitely more—indeed cost is virtually everything—for money market funds. (But those groups are beyond the scope of my talk today.) And we know that it matters—indeed it is a prime differentiator—in the nine-box equity style analysis whose pattern parallels the surface of a tic-tac-toe game.

So, why not take the position that investors should act on the full implications of the thesis that costs matter. Because the lowest cost funds out there in the marketplace are index funds, why not just buy index funds in each of the nine style boxes? It is hardly a specious argument.

This matrix shows both the return and risk of a low-cost index fund in each of the nine boxes, compared with the average return of the equity funds managed in that style. The index funds are operating index funds in the three large cap groups (Standard & Poor’s Indexes), and hypothetical index funds based

on publicly produced indexes (with returns reduced by estimated fund costs of 0.3%) in the medium and small cap groups (Frank Russell Indexes). This pair of tables reflect the spreads of risk and return in each category:

Exhibit XI: Indexes vs. Funds: Returns and Risks (Percentage Points)

<u>Added Index Return</u>				<u>Reduced Index Risk</u>			
	Value	Blend	Growth		Value	Blend	Growth
Large	+2.8	+1.8	+1.5	Large	-0.9	-1.3	-2.5
Medium	+2.9	+1.5	-0.4	Medium	-1.3	-2.0	-4.4
Small	+3.1	+0.6	-2.8	Small	-2.8	-3.3	-4.9

Let me summarize the outcome: The average *return* for the all funds in the index group in total was 1.4 percentage points above the average return for the equity group—+15.1% vs. +13.7%. In six of the nine boxes, the passively managed market index outpaced the average return of the actively managed equity funds; in two cases the results were about even; and in just one case (this time among the 60 small cap growth funds), the regular funds did better. But the average risk assumed by the indexes—and this is a truly remarkable finding—was far lower—by 2.2 percentage points: 9.7% for the indexes, 11.9% for the equity group. (Small and medium cap growth funds took particularly large extra risks).

The net result is that the *risk-adjusted* ratings averaged 124 for the index group and 99 for the regular funds—an average premium of fully 25% in risk-adjusted return. It is a strikingly consistent premium, one that is remarkably parallel across the matrix. The relative risk-adjusted ratings are so dramatically in favor of the low-cost index approach as to defy even the most optimistic (or, for active managers, pessimistic) expectations. Here, there are *no* exceptions whatsoever to the pattern. Indeed, its magnitudes are so devastating to the concept of high-cost active management that I, for one, could barely believe the figures. But we’ve checked them “eight ways to Sunday,” and correct they are:

Exhibit XII: Risk-Adjusted Ratings of Indexes vs. Equity Funds

	Value	Blend	Growth
Large	+17	+26	+62
Medium	+18	+10	+22
Small	+31	+12	+10

Lest this difference seem unimportant, at the 25% level the annual return in two funds with the same 10% standard deviation would be +16.5% for the nine passively managed index funds versus +14.0% for the actively managed traditional funds—a truly remarkable enhancement of 2.5 percentage points per year. Much of this spread, of course, is accounted for by the lower expense ratios and lower portfolio transaction costs for index funds. (Further, of course, the index funds would also come hand in hand with substantial tax advantages.)

To be sure, we should only go so far with three-year numbers in a bull market. But we also reviewed the past five years, which includes two poor market years and three good ones—hardly unrepresentative of the market’s long-term pattern. In fact, results during this period reflect an identical pattern, but with a substantially higher advantage for the indexes. Suffice it to say that our analysis deserves testing in other periods, and it is (as you might imagine) high on our agenda. But what we have today is surely a striking pattern, and one that appears to give the lie to the often expressed—now even trite—notion that “indexing works only in large cap markets.” *Given this data, that notion no longer has the ring of truth.* The Morningstar Category Ratings, indeed, may ultimately prove to be the biggest boon to indexing since the first S&P 500 Index mutual fund was founded in 1975.

That’s enough numbers for today. (A table showing the detailed data is presented as a separate appendix, attached.) Now, let me close with a few reflections on Pascal, on God, on index funds, and on a movie by the name of “War Games,” which will bring us back full circle to the game of tic-tac-toe. As Peter Bernstein tells the story in his marvelous book, *Against the Gods*, Blaise Pascal, the father of probabilities, cast the question of the existence of God into a game of chance: “A coin is tossed. Which way would you bet: on heads (God is) or tails (God is not)?” Paraphrasing Pascal, consider the chances of being on the losing side of the bet. If you bet God is . . . you will live a holy life and give up a few enjoyable temptations, but that’s all you lose. If you bet God is not and you are wrong, by leading an evil life you will be forever damned. *Consequences must outweigh probabilities.*

Turning to the stock market, Bernstein continues, if you believe it is efficient (and you are right). . . the best strategy is to buy an index fund. If you believe it is efficient (and you are wrong) . . . you will earn the market’s return but a few actively managed funds will beat you. But if you bet that the market is not efficient, the probability of underperforming is high. The risk, in short, is much greater if you bet on inefficiency rather than on efficiency.

And that is ultimately the conclusion of equity style analysis in the mutual fund industry: no matter what fund style you seek, emphasize the low-cost funds, eschew the high-cost funds. And, if you want the best bet of all, you should consider indexing in the Category Rating box in which you seek style representation in your portfolio. (A simpler course, to be sure, is to index your *entire* portfolio, with the S&P 500 Index or—probably a more conservative wager—with a *total* stock market index.)

If, because of high costs, investing in mutual funds is a (relative) loser’s game (though almost surely a winner’s game in absolute terms over the long run), is it not just another game between battling global armies? I won’t try to answer the question, but I will close by using a profound conclusion from the 1983 movie *War Games*. We are in the NORAD war room where our generals are trying to ward off an incipient global nuclear war, precipitated by a young computer nerd who has cracked the U.S. security system. The boy says he can solve the problem he has created, and, all other hope lost, the generals agree to let him try. He programs the U.S. air defense computer . . . with a game of tic-tac-toe. Calculating at a furious pace, the computer realizes that neither opponent can win the game—or the nuclear war—and the screen goes blank. The action ceases. Peace reigns. Then these words appear on the computer screen: “A strange game. The only winning move is not to play . . . How about a nice game of chess?”

Have investment management games, like global warfare games, become loser’s games, just like tic-tac-toe? Think about it.

I am deeply indebted to Walter H. Lenhard, Assistant to the Chairman, for his invaluable statistical and research support. I also want to thank James M. Norris, my former Assistant, for suggesting the inclusion of the “War Games” anecdote. Note: The opinions expressed in this speech do not necessarily represent the views of Vanguard’s present management. © Copyright 1997 by John C. Bogle