

SENIOR CONSULTANT

The Voice of the Investment Management Consultant

Identifying Investment Manager Skill

Ron Surz, President, PPCA, Inc.

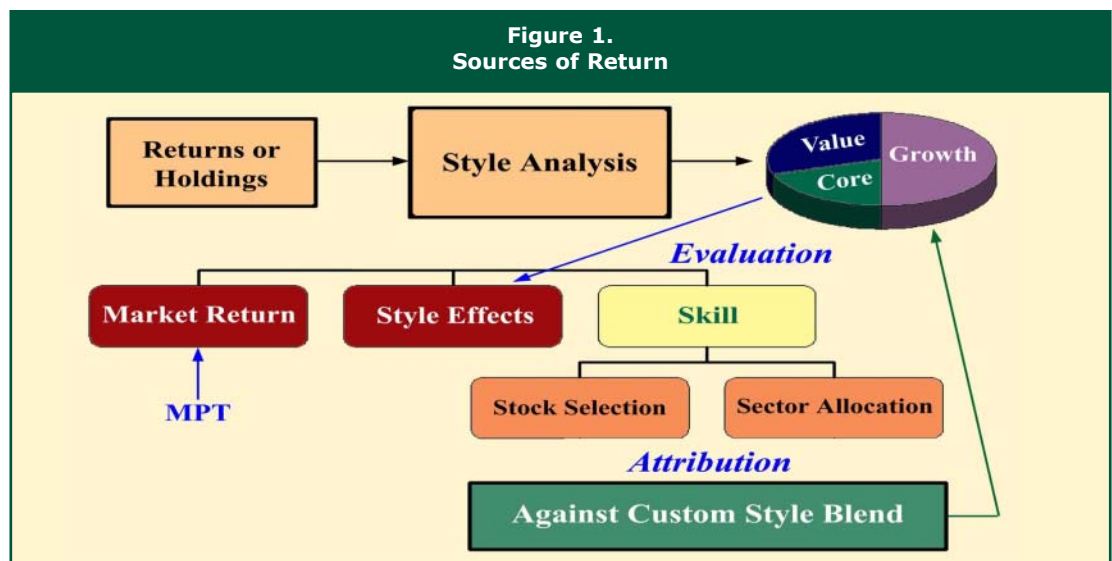
Editor's Note: This article is a representative chapter from the book, *The World of Money Management*, edited by Lyn Fisher and Sydney LeBlanc, to be released February 2004.

Do investment managers have skill? Can these managers be identified? Are my consultants personally invested with these managers? These very important questions could not be answered with confidence just a few short years ago. The motto of professional investment performance evaluators has long been “*evaluate skill, not luck.*” For years everybody attempted to make this distinction by using Modern Portfolio Theory (MPT), but this doesn't work, primarily because the focus of MPT is solely on broad market effects. However, about five years ago, researchers discovered the significance of investment style in identifying skill. A list of some of these studies is provided in the Appendix to this chapter. That is, they learned that skill could be properly identified only if we first lift the thick clouds of style that routinely distort our perspective, in addition to accounting for the broad market effects dictated by MPT. It was more complicated than MPT would suggest, but not that much more. Put simply, good growth equity managers tend to continue to be good growth equity managers and ditto for value, but if you're only looking at broad market effects, you'll never see this persistence. In the past, the problem with identifying skill has been that skill has routinely been confused with style. Witness the numerous firings of value managers that occurred as the growth stock bubble of the late 1990s inflated. Accordingly, the motto for 21st century evaluators is

gradually becoming “*evaluate skill, not style.*” We say *gradually* because the evolution beyond MPT has been slow, so not all would agree with what follows next.

In the remainder of this article, we take the perspective of the professional evaluator. Professional evaluators have an advantage over academic researchers who have discovered style-adjusted persistence in performance because these evaluators understand the other three Ps: People, Process, and Philosophy. Accordingly, they can use style-adjusted alpha as a first cut in their search for skill. This determination of positive risk and style-adjusted value added is called *performance evaluation*. Professional evaluators can then determine the reasons for the alpha and verify that these reasons substantiate the other three Ps. The examination of the reasons for performance is called **attribution analysis**. The reasons revealed by attribution analysis are stock selection and sector allocation. Importantly, to make sound decisions, evaluators look for persistence over time in these sources of added value. Furthermore, they confirm that the value added is coming from a source consistent with the management process. If the management process is predominantly top-down, one would expect alpha to derive primarily from sector allocation. Similarly, a bottom-up manager should excel in stock selection. This total performance evaluation and attribution picture is shown in Figure 1.

Figure 1.
Sources of Return





Note that while alpha, or skill, can be estimated using either holdings or returns, holdings are required to complete the picture with the components of skill, or attribution analysis. Also note that it is important that style be taken into account in both performance evaluation and performance attribution.

We will first discuss performance evaluation and then move on to performance attribution. As a practical matter, the search for skill should begin at the macro level with managers whose performances are good. Then due diligence can proceed with an understanding of the people, process and philosophy that produced the good performance. And then last, but not least, performance attribution confirms that the sources of this good performance are consistent with the people, process and philosophy. Throughout this process, we keep in mind that the resultant decisions are all about the future, even though we use the past as a guide.

Performance Evaluation

The answer to the question “is performance good or bad?” rests on yet another question: “Relative to what?” The investment industry has two answers to this second question: passive alternatives and peer groups. Most evaluators use both of these benchmarks. Passive alternatives are indexes, or combinations of indexes, that could have been purchased for a low fee in lieu of the selected active manager. Peer groups are collections of other managers who could have been hired, instead of the selected active manager. Both of these benchmarks have problems, but there are solutions to these problems.

Peer Groups

The investment industry has been using peer groups for so long that no one thinks to question them, except for the occasional discussion about survivor bias, which seems to be the only bias that people know and understand. Survivor bias raises the hurdle by including only those portfolios that have remained in business for the entire evaluation period, which is generally five or more years. In actuality, peer groups suffer from a collection of biases – only one of which is survivor bias – and each peer group has its own unique set of idiosyn-

cratic distortions. As a result, the exact same performance number will rank differently against different peer groups, even when all of the peer groups are for the same management mandate, such as large cap growth.

Even if a bias-free peer group existed, it still would not qualify as a valid investment performance benchmark as defined by the criteria established by Richards & Tierney, a Chicago-based investment consulting firm specializing in custom benchmarks and supported by the Association for Investment Management and Research (AIMR). These criteria are summarized below:

Criteria for Good Benchmarks

1. **Unambiguous:** Names and weights of securities are clearly stated.

TRADITIONAL PEER GROUPS MEET NONE OF THESE CRITERIA, EXCEPT POSSIBLY NO. 3 [MEASURABLE: THE BENCHMARK'S RETURN CAN BE CALCULATED ON A REASONABLY REGULAR BASIS]

2. **Investable:** Investors have the option to forego active management in lieu of a passive alternative.
3. **Measurable:** The benchmark's return can be calculated on a reasonably regular basis.
4. **Appropriate:** The benchmark accurately represents the manager's approach.
5. **Reflective of Current Investment Opinions:** The manager has current knowledge of the securities constituting the benchmark.
6. **Specified in Advance:** The composition of the benchmark is agreed upon and constructed prior to the start of all evaluation periods.

Traditional peer groups meet none of these criteria, except possibly No. 3.

Performance evaluation is all about making a judgment as to whether performance is good or bad. This judgment ought to be made relative to a passive alternative. *It doesn't matter*

how other managers in a particular peer group have fared, since they too should each be evaluated against their respective passive alternatives. Put another way, if you don't like your ranking in one peer group, choose another peer group.

What does matter is the degree of success or failure you experience relative to your benchmark. This can only be captured through a relatively new technique that creates all of the portfolios that could have been held by the manager, selecting from stocks in a benchmark that meet the criteria above. This new approach combines the better characteristics of peer groups with those of passive alternatives, while eliminating the problems of each. A manager's ranking in this scientific universe is an indication of the degree of his success or failure.

Passive Alternatives (Indexes)

Scientific universes bridge the gap between peer groups and passive alternatives. Passive alternatives have two problems: defining them and waiting. The industry's experience in manager performance evaluation has led it to conclude that most managers use a blend of styles. This conclusion has been reached with the help of a relatively new technology called

returns-based style analysis (RBSA). RBSA solves for the blend of style indexes that has behaved most like the manager and identifies this blend as the manager's effective style mix. Note the similarity of this approach to that of a “normal portfolio.” Normal portfolios are custom benchmarks designed to capture the essence of an individual manager's process and philosophy. Sometimes called *information-neutral portfolios*, these designer benchmarks were intended to reflect the portfolio to which the manager would retreat if one day he had no insights or ideas to implement. Almost everyone agreed at first that normals were a great idea, but it turned out that only a few consulting firms could construct them properly. Due to the cost and effort involved, normal portfolios never really caught on. By contrast to normal portfolios, which are composites of individual securities that are extremely difficult to construct, effective mix portfolios are blends of styles that are easy to create. The old idea of



designer benchmarks is back, and this time it is actually do-able.

The collection of style indexes used in RBSA is called a *style palette*. It is important that this palette be the best possible, so you can rely on the analysis. It's like finding the best color master for creating custom-blended paint. RBSA uses return history and optimization techniques to determine the blend of styles that most closely emulates the behavior of the investment portfolio. As with any statistical process, data problems in RBSA may go undetected, leading to faulty inferences. One such problem is multicollinearity, which occurs when the style indexes used in the regression overlap in membership. Multicollinearity invalidates the regression and usually produces spurious results. While most users of RBSA focus on optimal fit, it is important to note that good style palettes should also possess certain characteristics, as described below. The first two characteristics have been put forth as requirements by the developer and creator of RBSA. The last two characteristics are ours and represent our opinion of good characteristics.

Characteristics of Good Style Palettes

- **Mutually Exclusive:** No stock is categorized into more than one style. Accordingly, multicollinearity is minimized.
- **Exhaustive:** All stocks are classified. Some index vendors throw out data, e.g., stocks with negative earnings or small companies. Finding a good fit is impossible if any of the portfolio's stocks have been eliminated.
- **Inclusion of Core:** This continues to be a novel idea. It's a way to deal with stocks in that gray area between value and growth. Most index providers deal with this problem by either throwing out these stocks, violating the exhaustive rule, or by classifying them into multiple styles, violating the mutually exclusive rule. Interestingly, core doesn't always perform between value and growth. Sometimes it is better than both, and sometimes it is worse.
- **Quarterly Rebalancing:** Things change rapidly. Calling a cheap high-tech stock "growth" because it had a high price/earnings ratio a year ago doesn't make sense. Of

course, more frequent rebalancing makes an index harder to track because its composition is changing.

Be aware that the popular style palettes meet none of these criteria. Popularity is not synonymous with quality when it comes to RBSA. Regardless of the style palette, the resultant style blend meets the criteria for a good benchmark and importantly, represents a portfolio that the investor could have purchased for a low fee instead of hiring an active investment manager, so the problem of defining the benchmark has a contemporary solution.

The next problem is somewhat more difficult to solve. It takes many decades to develop statistical confidence in the manager's ability to beat his benchmark. For example, if your

**SO WITH THE OLD
TECHNOLOGY WE CAN PEEL
THE APPLE LIKE AN ORANGE
OR SLICE THE ORANGE LIKE
AN APPLE, BUT ALL WE'VE
GOT TO SHOW FOR IT IS A
ROTTEN FRUIT SALAD**

manager's return is 12% per year and the custom style mix has returned 10%, you'll have low confidence that this 2% annual spread is skill if it occurred over a 5-year period and might still have insufficient confidence over a 55-year period. This is because it takes a long time to develop confidence in the parameters that define the underlying statistical process.

But this waiting time problem can be solved by operating in the cross-section, rather than across time. By forming all of the portfolios that could have been formed from the passive style blend, you can determine statistical significance in a very short period of time. It's classic statistics. I'm testing the hypothesis that the manager has succeeded, so I let the monkeys form portfolios at random. If my manager's performance is in the top 10% of this random distribution, I accept the hypothesis with 90% confidence, and I do so even if the measurement period is as short as a few months or weeks. Let's put this another way. If a

manager delivers a 50% return in a month and his custom index is flat, it's a good chance that this is a significantly good return, and you shouldn't have to wait decades to come to this obvious conclusion. Cross-sectional performance evaluation produces the statistical backdrop necessary to make this determination, while traditional cross-temporal regression approaches (aka "alpha") do not. So performance is best evaluated against a scientific universe formed from all of the portfolios that the manager could have held, selecting stocks from his custom style blend. The next question is "Why is performance good or bad?"

Performance Attribution

Like performance evaluation, there has been an evolution in performance attribution. The performance measurement industry is fairly young, having started in the 1970s. Much of the attribution analysis that had been used until recently was developed in the 1980s, when we were only beginning to understand that there was more to life than MPT. We knew back then that characteristics like capitalization, price/earnings ratio and dividend-yield mattered, but hadn't figured out how to best integrate these factors into attribution analysis.

Consequently, we wrote "slicers and dicers" that segmented the portfolio and the benchmark by whatever characteristic we liked. Want to see how the segment of your portfolio with high P/Es fared against the comparable segment of stocks in the S&P500? No problem. Just draw the P/E line wherever you want and voilà. The problems with these old approaches are standardization and benchmark inflexibility. If you draw the P/E line at 15 and I draw it at 20, we'll each get different insights. Also, as described above, we'd like to use a custom style blend as the benchmark, but we can't do so with the 1980s technology because it doesn't provide the ability to customize the benchmark as a blend of indexes. So with the old technology we can peel the apple like an orange or slice the orange like an apple, but all we've got to show for it is a rotten fruit salad.

By contrast, contemporary technologies encourage the use of custom style-blended benchmarks and standardized style definitions so there is comparability across managers. In



this way a manager's stock selection and sector allocation skills are not confused with his style. Figures 2 and 3 summarize this evolution.

In the search for skill, we look for persistence in the reason(s) for good performance and for confirmation of the people, process and philosophy. The following Figure 4 shows a real-life manager who has consistently added value through stock selection, although the amount of value added has slowed somewhat in the recent past. This particular manager is a bottom-up stock picker, and the attribution analysis confirms his skill in this endeavor. Sector allocation has also added some value, which is consistent with bottom-up stock picking. Only trading activity has had a modest negative effect on performance. Trading activity measures the intra-period effects on performance of transactions executed during the period. If this manager were looking for ways to improve performance, a place to start would be the trading desk.

These relatively new tools give professional evaluators the insights needed to determine whether good performance is likely to continue into the future. The discussion thus far applies to all types of traditional portfolios – U.S. stocks, foreign stocks, bonds, etc. With some extensions, it can also be applied to hedge funds.

Hedge Funds

The growing popularity of hedge funds has engendered a tendency to confuse form with substance. The hedge fund form is excellent for skillful managers, since they can implement efficiently on most of their insights. As a result, some equate hedge funds with skill, even calling them "skill-based strategies." This skill part is not necessarily true, so the substance may, in reality, be missing from some hedge funds. Because of the fees involved and the degree of latitude afforded to the hedge fund manager, it is even more important that skill is present in hedge funds than in traditional investing. This determination must consider all of the moving parts in a hedge fund, as depicted in Figure 5:

In this context, style and attribution analyses begin by looking independently at the short portfolio and the long portfolio, just as if they were each long-only portfolios. Attribution analysis then blends the two portfolios, adding

Figure 2.
1980s Slicer and Dicer



Figure 3.
Today's Graft

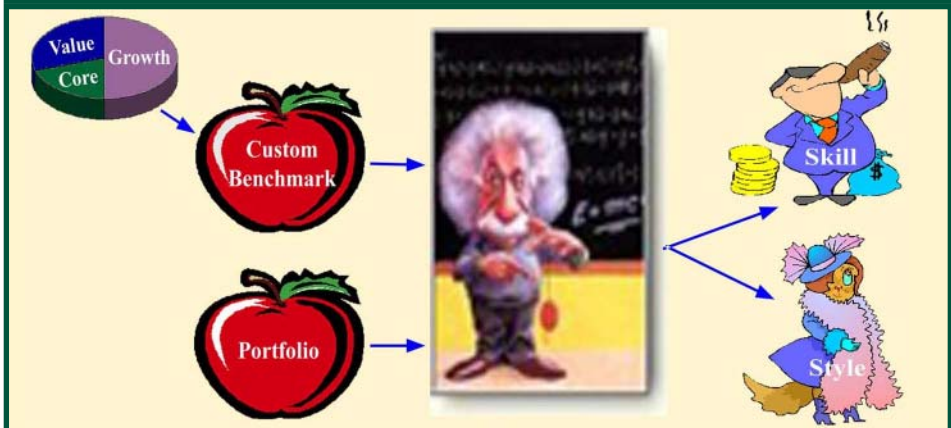


Figure 4.
Sources of Value Added: Persistence

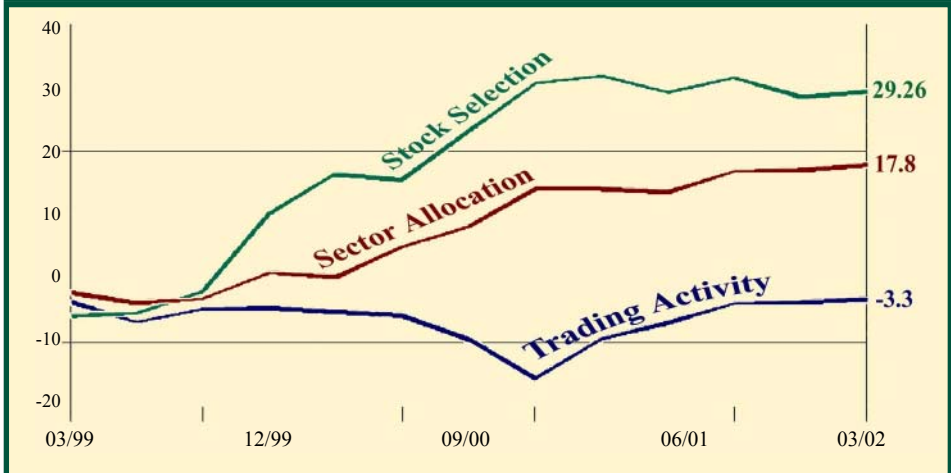
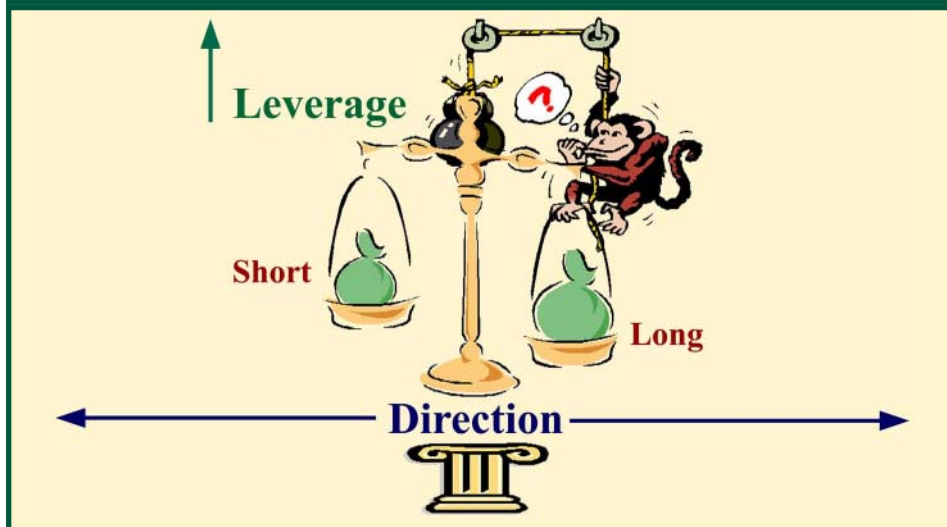


Figure 5.
Jones Model Hedge Funds



in the effects of directional bets, which are amounts – long or short – away from the target long and short exposures for the fund. The final level of attribution is leverage. Here we measure the effects of the targeted leverage as well as any deviations from this target. As in traditional analyses, the end game is persistence in one or more of these sources of value added and persistence that confirms the skill of the people, process and philosophy of the manager.

Conclusion

The search for investment manager talent puts a lot of emphasis on recent past performance. Unfortunately, in evaluating past performance, style is routinely confused with skill. The retirement industry is particularly notable for making this mistake as a group. Perhaps it's because the test of fiduciary prudence rests partially with what other fiduciaries have done.

After general market effects, the most important determinant of performance is style, followed by a distant third residual that we use to find manager skill. Detecting skill is tough for this reason. Although it's easy to confuse style with skill, it's hard to make good decisions once this mistake has been made. ■

Appendix: Studies Finding Evidence of Persistent Style-Adjusted Excess Return

- Coggin, Daniel T. and Charles A. Trzcinka, "A Panel Study of U.S. Equity Pension Fund Manager Style Performance," *Journal of Investing*, Vol. 9 (Summer 2000), pp 6-12.
- Gruber, Martin J. "Another Puzzle: The Growth in Actively Managed Mutual Funds," *Journal of Finance*, Vol. 51, No. 3, 1996, pp 783-810.

Ibbotson, Roger and Amita Patel. "Do Winners Repeat With Style," Ibbotson Associates research paper, November 2001.

Kahn, Ronald N. and Andrew Rudd. "The Persistence of Equity Style Performance: Evidence from Mutual Fund Data," *The Handbook of Equity Style Management*, 2d ed., (New Hope, Pa: Frank J. Fabozzi Associates, 1997).

Stewart, Scott D. "Is Consistency of Performance a Good Measure of Manager Skill?" *Journal of Portfolio Management*, Vol. 24, No. 3, 1998, pp 22-32.

About the Author

Ronald J. Surz is president of [PPCA, Inc.](http://www.ppcainc.com), a California-based computer software firm specializing in innovative analytic tools for sophisticated investors and their consultants. PPCA is the developer of StokTrib, a holdings-based style analysis and performance attribution system, applicable to both U.S. and non-U.S. portfolios. PPCA is also the creator of Portfolio Opportunity Distributions (PODs), which offers superior scientific benchmarks for evaluating investment performance. Use of these PODs to measure performance against the popular indexes ([PIPODs](http://www.ppcainc.com/PIP/PIPODs)) is a recommended *Senior Consultant Toolbox Solution* (<http://www.srconsultant.com/Toolbox/toolbox.html>).

Ron is a member of the Investment Performance Council of the Association for Investment Management and Research (AIMR) and is also on the board of the Investment Management Consultants Association (IMCA). He has earned an MS in Applied Mathematics and an MBA in Finance, and holds the Certified Investment Management Analyst (CIMA) designation. Ron can be reached at 949-488-8339 or Ron@PPCA-Inc.com

SENIOR CONSULTANT

THE VOICE OF THE INVESTMENT
MANAGEMENT CONSULTANT

JAMES P. OWEN
Co-Founder

STEPHEN C. WINKS
Co-Founder, Publisher & Editor-in-Chief

SYDNEY LEBLANC
Consulting Editor

MAMIE WOO MCNEAL
Production Editor

EDDIE BRYANT
Marketing Consultant



Advisory Board

JERRY BOTT
Bott Anderson

JOHN BROCK
Brock-Hazzard/Wachovia Securities

DICK CHARLTON
New England Pension Consultants

BOB CLUCK
Canterbury Capital

HAROLD EVENSKY
Evensky Brown & Katz

JEFF FRUM
Wells Fargo

RICH GLEASON
Salomon Smith Barney

KATHLEEN E. HEGENBART
Salomon Smith Barney

BRIAN HUNTER
Prudential Securities

GREG HUNTER
Merrill Lynch

BILL JOHNSON
CapTrust

JOHN KELSEY
Salomon Smith Barney

KEITH PHILLIPS
Morgan Stanley Dean Witter

BOB ROWE
Morgan Stanley Dean Witter

DICK SMITH
Capital Advisory Group

JIM YANNI
Yanni Partners

SENIOR CONSULTANT

1457 Crystal Springs Lane
Richmond, Virginia 23231

Ph 804-643-1075 ■ Fax 804-643-1544

[WWW.SRCONSULTANT.COM](http://www.SRCONSULTANT.COM)