Measure, attribute, appraise

Performance Measurement

1 A story: the check-up

- A nurse measures: weight, pulse, blood pressure
- The doctor explains: which habits caused the readings
- Then the doctor judges: healthy, and will it last?
- A portfolio review runs in the same three steps

IN THE LESSON

- Think about a visit to the doctor.
- First the nurse measures: weight, pulse, blood pressure.
- Those readings tell you what is happening, and nothing more.
- Next the doctor explains: which habits produced those readings.
- Last the doctor judges: is this patient really healthy, and is it likely to last?
- A portfolio review works the same way.
- Measurement says what the portfolio returned and how risky the ride was.
- Attribution says why, by tying the result to decisions the manager took.
- Appraisal says whether the result was skill that is likely to repeat.
- Dana Whitfield uses all three for the foundation, and she always works in that order, because each step leans on the one before.

2 Three questions, in order

Each step feeds the next; the findings feed the process

Measure: what happened →
Attribute: why it happened →
Appraise: was it skill

IN THE LESSON

- Here are the three stages as boxes, read from left to right.
- The first box is measurement.
- It works out the portfolio's return over the period, usually time weighted, which strips out the effect of when clients paid money in or took it out.
- It also measures risk.
- The second box is attribution.
- It takes the gap between the portfolio and its benchmark and pins it on particular decisions, such as the sector weights or the choice of securities.
- The third box is appraisal.
- It adjusts for the risk taken and asks whether the pattern is strong and steady enough to expect again.
- And the findings do not stop there.
- They go back into how the portfolio is run.

3 Excess return and the information ratio

Excess return = Rp − Rb
Information ratio = excess return ÷ tracking error

Portfolio return	Time-weighted, for the period
Benchmark return	Return on the mandate benchmark
Tracking error	Volatility of the gap
Information ratio	Return per unit of active risk

IN THE LESSON

- Two short formulas carry measurement and appraisal.
- The excess return is the portfolio return minus the benchmark return, both over the same period.
- It is the raw material for attribution.
- The tracking error is the standard deviation of that gap, and it measures how much active risk the manager ran.
- You met it at the earlier levels.
- The information ratio divides the excess return by the tracking error.
- So it tells you how much extra return the manager earned for each unit of active risk.
- A higher ratio is better, and because it uses risk in the bottom of the fraction, a big return earned by taking a big risk does not look as good as it first seems.

4 Example 1: reviewing Alder Ridge

Alder Ridge, 3-year average; Brindle's global equity manager

Portfolio return	time-weighted, a year	**11.0%**
Benchmark return		**8.0%**
Excess return	11.0 − 8.0	**3.0%**
Tracking error		**6.0%**
Information ratio	3.0 ÷ 6.0	**0.5**

IN THE LESSON

- Now the first case.
- Dana Whitfield reviews Alder Ridge Investors, the foundation's global equity manager, which calls itself a bottom up stock picker.
- Over three years its time weighted return averaged eleven per cent a year, against eight per cent for its benchmark.
- That is measurement.
- Eleven minus eight gives an excess return of three per cent a year.
- The tracking error, the volatility of that gap, was six per cent.
- Appraisal divides the three by the six, so the information ratio is point five.
- In plain words, the manager earned half a point of extra return for every point of active risk it took.
- Those are illustrative figures.

5 Attribution: where the 3.0 came from

Points of excess return a year

Allocation	**0.5**
Selection	**2.2**
Interaction	**0.3**
Total excess	**3.0**

0.5 + 2.2 + 0.3 = 3.0

IN THE LESSON

- Attribution takes that three per cent and asks where it came from.
- The vertical axis is points of excess return a year.
- The first bar is allocation, the gain from holding sectors in different weights from the benchmark.
- It is only half a point.
- The second bar is selection, the gain from picking better stocks inside each sector.
- It is two point two points, by far the biggest.
- The third bar is interaction, the joint effect of the weights and the stock choices together, three tenths of a point.
- The last bar is the total.
- The three effects add back to three.
- Notice how well this fits what Alder Ridge says it does: most of the gain is stock selection, the edge it claims.

6 Example 2: a wrong benchmark travels downstream

Same manager; a different benchmark; tracking error held at 6.0%

Information ratio, right benchmark	3.0 ÷ 6.0	**0.50**
Wrong benchmark return		**9.5%**
Excess return, wrong benchmark	11.0 − 9.5	**1.5%**
Tracking error, held		**6.0%**
Information ratio, wrong	1.5 ÷ 6.0	**0.25**

IN THE LESSON

- Each stage leans on the one before, so a mistake early on spreads.
- Suppose the committee had, by mistake, measured Alder Ridge against a different index, one that returned nine and a half per cent over the same three years.
- The manager's return is still eleven.
- But the excess return falls from three to one and a half.
- To keep the comparison simple, hold the tracking error at six.
- Then the information ratio falls from point five to point two five.
- Half the apparent skill vanished, without the manager changing anything.
- And the attribution would also have shifted.
- So an error in measuring returns or choosing the benchmark contaminates both the explanation and the verdict.

7 Example 3: is three years enough?

Alder Ridge: information ratio 0.5 over 3 years

Information ratio	3.0 ÷ 6.0	**0.5**
Years of record		**3**
t-statistic	0.5 × √3	**0.87**
t-statistic usually wanted	about	**2**
Years needed	$(2 \div 0.5)^2$	**16**

IN THE LESSON

- Now the trap that exam cases love.
- A good return is only a measurement result.
- It is not yet proof of skill.
- How long a record do we need?
- A statistic called the t statistic tells us whether a result stands out from luck.
- For an annual information ratio, with returns that are unrelated from one year to the next, it is about the ratio times the square root of the number of years.
- Half times the square root of three is point eight seven.
- A common rule of thumb asks for about two.
- Work backwards: two divided by half, squared, is sixteen years.
- So Alder Ridge looks promising, but the committee cannot yet call it skilled.
- A short, strong record is a reason to keep watching, not a verdict.

8 A story: the neat map of the wrong city

- A neat map of the wrong city still misleads
- A report can add up and still tell you nothing
- Effective means complete and faithful to the process
- It also needs sound data and a fair benchmark

IN THE LESSON

- Imagine a satnav with a beautifully drawn map, every street labelled.
- But it is the map of a different city, or last year's roads.
- It looks perfect, and it will still send you the wrong way.
- An attribution report can be like that.
- The numbers may add up neatly, yet the report may not help the committee at all.
- It may leave a big slice of the result unexplained.
- It may credit the manager with decisions it never took.
- Or it may rest on bad data.
- Dana Whitfield therefore checks four features before she trusts any attribution.
- Effective attribution is not about neat arithmetic.
- It is about being complete, being faithful to the process, and being built on sound inputs.

9 Four features of effective attribution

What Dana Whitfield checks first

Feature	The test
Complete	Residual small
Mirrors the process	Same decisions
Counts each active decision	Fair benchmark
Sound inputs	Clean data

IN THE LESSON

- Here are the four features, and each is something you can check.
- First, it is complete.
- It accounts for the whole of the excess return, or the whole of the active risk, with no large unexplained residual.
- A big residual usually signals missing trades, poor data or the wrong model.
- Second, it mirrors how the manager actually decides.
- Third, it quantifies each active decision the manager took, against a benchmark that fits the mandate, so the output shows where value was added or lost.
- Fourth, it is built on sound inputs: accurate holdings, prices and transactions, and returns calculated in a consistent way.
- Exam cases often ask which of these a described process breaks.

10 Example 4: the report that breaks two features

Alder Ridge, year four: excess return 2.0%; sector report

Allocation + selection	0.8 + 0.4	**1.2%**
Excess return to explain		**2.0%**
Residual	2.0 − 1.2	**0.8%**
Share left unexplained	0.8 ÷ 2.0	**40%**
Allocation credit, a stock picker		**0.8%**

IN THE LESSON

- Here is an exam style case.
- In year four, Alder Ridge beat its benchmark by two per cent.
- The custodian's report is a sector breakdown built from holdings at the start of each month.
- It shows allocation of point eight and selection of point four.
- Add them and the report explains one point two of the two.
- That leaves a residual of point eight, which is forty per cent of the excess return unexplained.
- That is too large, so completeness fails, and the likely cause is the trades made during each month, which month start holdings cannot see.
- There is a second failure.
- The point eight of allocation credits sector bets to a manager that never makes sector bets.

11 Match the breakdown to the manager

The breakdown must follow how the manager decides

Manager's process	Fitting view
Top-down sector rotator	By sector
Bottom-up stock picker	By security
Factor-based process	By factor

IN THE LESSON

- The second feature, mirroring the process, deserves its own look.
- The breakdown must follow the order in which the manager really takes decisions.
- A top down sector rotator first decides which sectors to overweight, then picks stocks.
- For that manager an allocation and selection breakdown by sector is the right tool, and the allocation effect is a real, deliberate choice.
- A bottom up stock picker chooses stocks first and lets the sector weights fall out as a by product.
- Analysed by sector, that manager shows allocation effects nobody chose.
- It needs a security level view.
- A factor based process needs a factor view, with the return from each exposure plus a specific part.
- Pennant Capital, the foundation's sector rotator, fits the first row.

12 The same 2.0, explained twice

Points of excess return

Allocation	**0.8**
Selection	**0.4**
Residual	**0.8**
Selection	**1.7**
Trading	**0.2**
Residual	**0.1**

Red: sector model, month-start holdings. Green: security model, all trades

IN THE LESSON

- Here is the same year for Alder Ridge, explained in two ways.
- The vertical axis is points of excess return.
- The first three bars, in red, come from the sector model on month start holdings.
- It credits point eight to allocation and point four to selection, and leaves point eight in the residual.
- The last three bars, in green, come from a security level model that uses every trade.
- Almost everything lands in selection, one point seven, where the manager says its edge lies.
- There is a small trading effect of point two, and the residual is only point one.
- Both reports add to two.
- But only the second fits the manager and explains the result, and that is the one the committee can act on.

13 A story: where the time was won

- Where was the time won? That is return
- Where was the energy burned? That is risk
- The director sees the whole team: macro
- One rider sees his own race: micro

IN THE LESSON

- Picture a cycling team.
- At the end of a race, the director looks at where the team won time: perhaps on the flat stage.
- That is the story of the result, told after the event.
- But the director also wants to know where the riders burned their energy.
- Maybe most of it went on the mountain stage, where they gained nothing.
- That is the story of the risk.
- There is a second choice of view.
- The director sees the whole team and its tactics.
- Each rider sees only his own race.
- A foundation looks at its portfolio in just these ways: return or risk, and the whole fund or one manager.

14 Four kinds of attribution

Two splits: return or risk, and macro or micro

Kind	Answers
Return attribution	What was earned
Risk attribution	Where risk sits
Macro: the whole fund	Owner's choices
Micro: one manager	Manager's choices

IN THE LESSON

- Let us pin the words down.
- Return attribution explains the sources of excess return, after the event.
- Risk attribution explains where the portfolio's risk comes from.
- For a benchmark relative mandate that risk is the active risk, the tracking error.
- For an absolute return mandate it is total volatility.
- And it can be run before the event, on expected risk, as well as after.
- The other split is the level.
- Macro attribution works at the level of the whole fund and the asset owner's decisions: the strategic weights, the tactical deviations from them, and the choice of managers.
- Micro attribution works inside one manager's portfolio, explaining that manager's allocation and selection decisions against the mandate benchmark.

15 Macro above, micro below

Macro: the foundation's decisions

Policy weights → Tactical tilts → Choice of managers

Micro: inside one manager

Sector weights → Stock selection

IN THE LESSON

- This picture shows the two levels.
- Start with the top row, the macro level.
- It belongs to the asset owner, here the foundation's investment committee.
- The first box is the policy weights, the long term mix it chose.
- The second is the tactical tilts, where it leaned away from that mix.
- The third is the choice of managers, and the benchmark each is given.
- Now the bottom row, the micro level, which sits inside a single manager's portfolio.
- The first box is the manager's sector weights against the mandate benchmark.
- The second is its stock selection within the sectors.
- Each manager answers only for its own row, against its own mandate.
- The committee answers for the top row.

16 Example 5: macro attribution for the fund

Brindle fund, last year; peeled in layers

Fund return		**8.0%**
Policy benchmark return		**7.0%**
Managers' benchmarks, actual mix		**7.4%**
Owner's decisions	7.4 − 7.0	**0.4%**
Managers' value added	8.0 − 7.4	**0.6%**

IN THE LESSON

- Now a macro example.
- Last year the whole Brindle fund returned eight per cent.
- Its policy benchmark, the return on the long term mix, was seven per cent.
- The managers' own benchmarks, blended at the weights the fund actually held, returned seven point four.
- We peel the eight in layers.
- The step from seven to seven point four is point four.
- That comes from the committee's own decisions, its deviations from the policy weights and its choice of benchmarks.
- The rest, eight minus seven point four, is point six, the managers' combined value added.
- The trap is blaming or rewarding the managers for a result that was driven by the policy mix, such as a weak year for equities.

17 Example 6: micro return against micro risk

Pennant Capital: excess return 1.5%, tracking error 4.0%

Technology	return 1.2, risk 0.6	**80% and 15%**
Energy	return 0.0, risk 2.4	**0% and 60%**
Everything else	return 0.3, risk 1.0	**20% and 25%**
Totals	excess 1.5, tracking error 4.0	**1.5 and 4.0**

IN THE LESSON

- Now a micro example, for Pennant Capital, the sector rotator.
- Its excess return was one and a half per cent, and its tracking error four per cent.
- The return effects add to the excess return, and the risk contributions add to the tracking error.
- Technology earned one point two of the return, which is eighty per cent of it, and used point six of the risk, fifteen per cent.
- Energy earned nothing, yet used two point four of the risk, which is sixty per cent.
- Everything else earned point three, twenty per cent of the return, and used one point zero, a quarter of the risk.
- Return attribution alone would praise the technology calls.
- Risk attribution shows energy eating the risk budget.

18 Share of the return against share of the risk

Share of total (%)

Tech: return	**80%**
Tech: risk	**15%**
Energy: return	**0%**
Energy: risk	**60%**

Pennant Capital: same year, two views

IN THE LESSON

- Here are those shares as a picture.
- The vertical axis is the share of the total, in per cent.
- The first pair of bars is technology.
- The first bar is its share of the return, eighty per cent.
- The second is its share of the risk, only fifteen per cent.
- A good deal.
- Now energy.
- Its share of the return is zero, nothing at all.
- But its share of the risk is sixty per cent.
- Perhaps the position was a deliberate hedge, or a view that has not paid off yet, or an accident of stock choice.
- The committee needs to know which.
- A manager can earn its return in one place while spending its risk budget in another, and only the second pair of bars shows it.

19 Mistakes to avoid so far

- A high return taken as proof of skill
- A poor benchmark distorting every later step
- Ignoring a big residual in the attribution
- Praising allocation skill in a bottom-up stock picker
- Reading return attribution without risk attribution

IN THE LESSON

- Let us collect the common mistakes from the first three topics.
- One: treating a high return as proof of skill, when it is only a measurement result.
- Two: letting a poor benchmark distort the excess return, the attribution and the information ratio together.
- Three: shrugging at a big residual, when it signals missing trades, bad data or the wrong model.
- Four: crediting a bottom up stock picker with sector allocation skill that was never a deliberate choice.
- Five: reading return attribution without risk attribution, and so missing a manager who spends its risk budget where it earns nothing.
- Keep these five in mind, because the exam cases are built around them.

20 A story: three ways to rebuild a match

- Final score only: returns-based
- Team sheet at kick-off: holdings-based
- Full commentary, every change: transactions-based
- More detail needs more data, and costs more

IN THE LESSON

- Suppose you missed a football match and want to know what happened.
- The cheapest way is to look up the final score.
- That tells you who won, but not how.
- A better way is to read the team sheet from kick off.
- Now you know who started, but you still do not know about substitutions or tactical changes.
- The best way is a full commentary that records every substitution and every change.
- It tells you everything, but someone has to write it all down.
- The three methods of attribution are exactly these.
- Returns based uses the final score.
- Holdings based uses the team sheet.
- Transactions based uses the full commentary.

21 More data, more precision

- Returns-based
- Holdings-based
- Transactions-based

IN THE LESSON

- Let us place the three methods on one line.
- The left end is little data and a rough answer.
- The right end is all the data and a precise answer.
- Returns based attribution sits at the left.
- It needs only a return series, so it is quick and cheap, and works where holdings are hidden, as with many hedge funds.
- Holdings based attribution sits in the middle.
- It uses the positions at the start of each period.
- It is more accurate, especially with frequent snapshots and low turnover.
- Transactions based attribution sits at the right.
- It uses the holdings and every trade, so it reconciles fully to the actual return.
- The price of moving right is data, effort and cost.

22 Returns-based attribution is a regression

$R_{p,t} = \alpha + \beta_1 \cdot F_{1,t} + \beta_2 \cdot F_{2,t} + \varepsilon_t$

Fund return	The return series, nothing more
Exposure	How much behaves like each factor
Factor return	A style or factor index
Unexplained	Alpha and error, left over

IN THE LESSON

- Returns based attribution is a regression.
- Each period's fund return is explained by the returns on a few factor or style indexes.
- The exposures on the right are the regression coefficients, and they say how much of the fund behaves like each factor.
- The residual is what the factors do not explain.
- It includes the intercept, the alpha, and the error term.
- Notice what the regression never sees.
- It sees only a series of returns.
- It cannot name a stock, a sector or a trade.
- And if the manager changes its exposures during the sample, the regression averages over the change and can mislead.
- That is the price of needing so little data.

23 Example 7: Quill Partners, a hidden hedge fund

Quill Partners returned 9.0%; growth index 10.0%, value index 6.0%

Growth share	70% × 10.0%	**7.0%**
Value share	30% × 6.0%	**1.8%**
Explained by style mix	7.0 + 1.8	**8.8%**
Fund return		**9.0%**
Left unexplained	9.0 − 8.8	**0.2%**

IN THE LESSON

- Quill Partners is a hedge fund in the Brindle portfolio.
- It will not disclose its positions, so the only tool is returns based attribution.
- Dana Whitfield regresses its returns on a growth index and a value index.
- The result says that seventy per cent of the fund behaves like the growth index, which returned ten per cent.
- That contributes seven.
- Thirty per cent behaves like the value index, which returned six.
- That contributes one point eight.
- Together the style mix explains eight point eight.
- The fund actually returned nine, so point two is left over.
- That is useful.
- We know the style tilt.
- But we cannot say which stocks, or which trades, earned it.

24 Example 8: the mid-month sale

100m, 50 in A and 50 in B; A +6% then sold, would then fall 4%; B +1.0%

Actual return	(53.0 + 50.5) ÷ 100 − 1	**3.50%**
Holdings-based return	(50.88 + 50.5) ÷ 100 − 1	**1.38%**
Unexplained by holdings	3.50 − 1.38	**2.12%**
Transactions-based credit	to the mid-month sale	**2.12%**

IN THE LESSON

- Now why holdings based attribution can miss things.
- In one month Alder Ridge starts with one hundred million, fifty in stock A and fifty in stock B.
- A rises six per cent by mid month, and the manager sells all of it, worth fifty three.
- Had it been kept, A would then have fallen four per cent, ending the month at fifty point eight eight.
- B finished at fifty point five.
- The actual return is fifty three plus fifty point five, over one hundred, minus one: three point five per cent.
- The holdings based figure assumes the start of month positions were kept: fifty point eight eight plus fifty point five.
- That gives one point three eight.
- The gap is two point one two, the value of the sale, lost in the residual.

25 What each method sees

Return for the month (%)

Actual	**3.50%**
Holdings-based	**1.38%**
Left in residual	**2.12%**

Alder Ridge, one month

IN THE LESSON

- Here is that month as a picture.
- The vertical axis is the return for the month, in per cent.
- The first bar is the actual return, three point five.
- The second bar is the holdings based return, one point three eight.
- It is right about what it measures.
- If the manager had kept its start of month positions, that is what it would have earned.
- The third bar is what is left in the residual, two point one two.
- That is the value of the sale, the manager's best decision of the month, and it looks like noise or a data error.
- A transactions based method records the trade, so it reconciles to the first bar and credits the sale.
- This is why holdings based works best with low turnover.

26 Strengths and weaknesses

Three methods, one trade-off

Method	Verdict
Returns-based	Cheap, but rough
Holdings-based	Exact, misses trades
Transactions-based	Complete, costly

IN THE LESSON

- Let us put the three methods side by side.
- Returns based is quick and cheap, and usable where holdings are hidden.
- But it is the least precise, it cannot identify individual decisions, and it can mislead when exposures change.
- Holdings based is more accurate.
- But it ignores trades within the period, so trading effects and transaction costs leak into a residual.
- Frequent snapshots and low turnover help.
- Transactions based uses holdings plus every trade.
- It reconciles fully to the actual return and captures trading decisions.
- It is the most accurate, and also the most data hungry and costly.
- So the choice is a trade off between precision and the data you can actually get.

27 Example 9: choosing a method for each manager

Match the method to the data and the turnover

Quill Partners	returns only, holdings hidden	**Returns-based**
Pennant Capital	turnover 25%, monthly holdings	**Holdings-based**
Alder Ridge	turnover 160%, all trades	**Transactions**

IN THE LESSON

- To finish, Dana Whitfield picks a method for each of the three managers.
- Quill Partners shows only returns and keeps its holdings hidden, so returns based is the only option.
- Pennant Capital turns over about a quarter of its portfolio a year and sends monthly holdings, so holdings based is accurate enough, because little happens between snapshots.
- Alder Ridge turns over one hundred and sixty per cent a year and can send every trade.
- Heavy trading is exactly where holdings based leaks into a residual, so transactions based is worth its cost.
- The exam rule is simple: match the method to the data you have and the turnover the manager runs.

28 Recap: the lesson in five lines

- Measure, then attribute, then appraise
- Effective: complete, faithful, fair benchmark, sound data
- Return versus risk; macro versus micro
- Returns, holdings, transactions: more data, more precision
- A high return alone never proves skill

IN THE LESSON

- Let us pull the lesson together.
- First, measurement says what happened, attribution says why, and appraisal asks whether it was skill, and each depends on the one before.
- Second, an effective attribution is complete, mirrors the manager's process, counts each active decision against a fair benchmark, and uses sound data.
- Third, return attribution and risk attribution answer different questions, and macro looks at the fund while micro looks inside one manager.
- Fourth, returns based is cheap and rough, holdings based is accurate but misses trades, and transactions based is complete but costly.
- Fifth, a high return alone never proves skill.

29 Try three quick questions

1. Excess 2.4%; allocation 1.2%, selection 0.6%. Residual?
   Answer: 2.4 − 1.8 = 0.6, which is 25% unexplained: too large.
2. Excess 2.0%, tracking error 5.0%. Years for t = 2?
   Answer: IR = 0.4; $(2 \div 0.4)^2$ = 25 years.
3. A heavy mid-month trader: which method sees the trades?
   Answer: Transactions-based; holdings-based leaves them in the residual.

IN THE LESSON

- Now try three quick questions yourself.
- Pause after each one.
- Question one.
- A manager beat its benchmark by two point four per cent.
- The attribution shows allocation of one point two and selection of point six.
- What is the residual, and is it a problem?
- Two point four minus one point eight is point six, which is a quarter of the excess return, too large to ignore.
- Question two.
- A manager has an excess return of two per cent and a tracking error of five per cent.
- How many years until the t statistic reaches two?
- The information ratio is point four, and two over point four, squared, is twenty five years.
- Question three.
- A manager trades heavily in the middle of each month.
- Which method sees those trades?
- Only transactions based attribution does, while holdings based leaves them in the residual.