

Measuring The Marigolds

By Michael Bland

*‘As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.’
(Albert Einstein)*

‘This removes the last excuse for not evaluating PR,’ shrieked a CIPR newsletter of a new report on evaluation (it removed nothing of the sort)...

‘...the Luddites who say you can’t measure PR are a dwindling minority,’ proclaimed PR Week...

‘...of course, you’re anti-evaluation,’ a leading industry figure said to me...

I tell you, life is hard for a realist these days. Far from being anti-evaluation, I was proclaiming its virtues in my books when most of the current evaluation brigade were counting their teddy bears in the toy box and have always regarded assessment and evaluation as an invaluable element of communications strategy.

The revolution has eaten its children

As with any management discipline, most projects benefit from various types of research at different stages to assess factors such as audience attitudes, aim, types of message, effectiveness of programme etc. It can also throw up otherwise unidentified opportunities and audiences – as in the case of the London Olympic bid, where some very good research spotlighted a crucially important targeting gap.

For a long time, PR largely lacked this important discipline (and still often does) but along with the deployment of more research and evaluation has come the typical consequence of the revolution eating its children in the form of three flawed nostrums:

- that all PR *must* be evaluated. An extreme example of this was a (mercifully now dead) campaign by PR Week for 10% of all PR programme budgets to be spent on evaluation.
- that only certain types of ‘scientific’ evaluation tools are acceptable.
- that PR outcomes can be accurately measured (PR *inputs* can be measured and are not discussed here).

While it is true that a degree of evaluation is desirable for most projects, it doesn’t always have to be the case, especially in small, simple or low budget programmes. For example, when I was involved in promoting post-war Iraqi Kurdistan as a viable inward investment opportunity to Western businesses the usual initial benchmark study of knowledge of – and attitudes towards – the ‘product’ was deemed unnecessary as we already knew the answers: a) None and b) Not a lot.

When the subject of evaluation is studied in any depth a number of concerns and weaknesses in the arguments soon come to light:

Why PR can’t be measured precisely

The first is that it is almost always impossible to measure the effectiveness of PR precisely. In 30 years I have only encountered one campaign where all other factors could be isolated and the success of the PR program measured in isolation. Normally there are several other variables, such as:

- Advertising
- Marketing activity
- State of the economy
- Spending patterns
- Competitor activity
- Perceptions affected by actual product improvements
- Changing audience tastes.

To assess the relative impact of all the different variables calls for a high degree of subjective judgement and guesswork, which makes a nonsense of the concept of ‘precise’ evaluation.

Like the proverbial man who asked directions to Cork and was told he shouldn't be starting from here, there is little point measuring something accurately if the starting parameters are wrong in the first place.

I recently attended a conference of business professionals who were concerned at the low level of women in senior positions – estimated at 15%. They instigated a programme to raise this figure to 25% within three years. The words 'if you can measure it you can manage it' were actually used and applauded. This is another example of how arbitrary metrics can set a programme in stone and potentially lead everyone off down the wrong track. How was 'senior position' determined? Why is 25% the right level to achieve? Why three years? It might be that the target level and timescale were unrealistic and that an absolutely brilliant programme could only achieve a lower level – yet it would be deemed a failure because of the unrealistic targets. Conversely, it might be that the targets were easily achievable and an inadequate programme would be deemed a success.

Another variable which throws a spanner in the works is time span. A major function of PR is to reduce the time span between reality and perception, but it is almost impossible to know exactly how long it will take from, say, a product improvement to full awareness of that improvement being achieved. The choice of a three-year time span to achieve the aim in the above example was entirely arbitrary and the actual optimum time would vary according to a number of virtually unmeasurable factors.

Science is an inexact science

The various methods of evaluation involve increasingly sophisticated techniques – but it must not be assumed this makes them either truly scientific or particularly reliable. There is a naive belief among sections of the PR community in the infallibility of science. While the pure units of measurement of time, speed etc are extremely accurate (though even staples like the kilogram and astronomical time are not entirely constant), problems of inaccuracy and changing interpretations set in as soon as these metrics are deployed to measure complex and variable systems.

If measuring PR effectiveness were as simple as measuring a piece of string there wouldn't be a problem. But a better analogy for the scientific challenge of trying to measure something as complex and intangible as a PR programme is the analysis of a fractal such as a coastline. At first, a question like 'How long is the coastline of Britain?' seems easy – and no doubt the Ordnance Survey has an 'answer.' But where does the 'coastline' begin and end when you get to, say, the Thames Estuary? Do you hop across from Southend to Sheerness – or do you go round to Tilbury?

Or Maidenhead? Or right up to the source of the Thames? On the way round the coast, do you take in all the inlets and creeks? If you do, you will increase the measured length by a factor of several hundred percent. In short, there is no way of measuring the actual coastline, merely a variety of metrics to choose from – all of which will give different results.

Also, scientists themselves often admit to what is called the ‘experimenter effect,’ whereby researchers, no matter how objective they believe their experimentation to be, get the results they were looking for in the first place. This is one reason scientists are always arguing with each other!

The concept of science somehow being a hard measuring device which will tell us exactly how we are performing is reassuring but, sadly, a myth.

‘...many working scientists acknowledge that subjectivity plays a big role in their day-to-day thinking.’ (Robert Matthews, New Scientist, 13 March 2004).

Often, seemingly sound communications research gives ‘scientific’ results which experience and common sense tell us are nonsense. Examples include the well known ‘findings’ that only seven per cent of the impact of a message is down to the content (it turned out that the researchers didn’t actually experiment with different types of content in the first place) and that only 27 per cent of people trust the media as a source of information (most people *say* they don’t believe the media, but experience indicates that the media is actually the main source of much that they believe to be true). And much of Coca Cola’s misery over public outrage that its Dasani bottled water came from the Sidcup mains was caused by their belief in their research:

‘Our research told us that there was a very low level of interest in where the water actually came from...’ (Coca Cola UK, 2004).

Much of the faith which is placed in the accuracy of PR measurement techniques is therefore misplaced. Though full of charts, computer readouts and assorted mumbo jumbo they are often no more scientifically reliable than the ritual reading of a chicken’s entrails.

More to it than the media

Another concern is the over emphasis placed by the PR trade on media evaluation only. The terms ‘evaluation’ and ‘media evaluation’ are often synonymous.

An analysis of media coverage can be a good indicator in cases where the media are the strongest influencer but in a majority of cases, from motor cars to supermarkets, our

opinions are also formed by a range of other factors such as customer service, word of mouth and other non-media influences.

Indeed, many professional PR campaigns seek to generate a WOM ('word-of-mouth') effect via third parties **other** than the media. So the tendency for our trade to see media evaluation as the only benchmark can be detrimental to our ability to grasp the whole picture. Our propensity to buy a particular make of car, for example, is influenced more by non-media factors such as:

- Previous experience with that make of car
- Recommendations by others
- Advertising
- Sales talk
- Budget
- Available specification meeting customer requirements.

Good or bad media coverage about that car, and the volume of that coverage, may play some part – but the impact will vary from person to person and from how much that person has been exposed to other, non-media, influences.

The actual role of media coverage in influencing decision-making for the car buying universe is therefore impossible to measure and assess. Using media analysis as a (and often **the**) leading benchmark in evaluating the effectiveness of a car PR campaign could therefore be dangerously misleading.

Still on cars, even non-media evaluation methods can mislead. In the mid 1980s I was involved in a major £450,000 a year PR campaign to improve attitudes to the UK motor industry. Sure enough, over the ensuing three years the MORI favourability/familiarity analysis showed a significant improvement and everyone was happy. But in reality the industry had also made dramatic improvements in the preceding couple of years to product quality and service – and people were starting to notice. Actual experience – and word of mouth resulting from the physical improvements – will certainly have played a more significant part in the attitude shift than the PR campaign, which got all the credit.

Value of negative coverage

Media evaluation analysts also make the flawed assumption that negative coverage is automatically 'bad.'

In fact, negative coverage can be extremely ‘good.’ My book on crisis management cites a number of cases where negative publicity has done wonders for a product or organisation. Back on cars again, Jeremy Clarkson claims that every car he has panned has been a huge success! Indeed, some people even stir up negative controversy to increase awareness (e.g. Benetton and some celebrities).

It is impossible to assess the degree to which negative coverage has a favourable or unfavourable impact. In some cases it can actually enhance awareness of an organisation or product without harming propensity to buy or do business (e.g., Stena Line ferry grounding; Intel Pentium launch glitch) – but in others a sudden burst (e.g. Ratner) or steady stream (e.g. Wang) of bad coverage can be fatal. Evaluators cannot assess the real positive or negative impact of unfavourable publicity because there are just too many variables and imponderables to make any serious, scientific study of the subject possible. The mathematics involved are closer to chaos theory than bean counting.

The weakness of surveys

Much of our non-media research is survey based and surveys, too, can be misleading tools. Many are conducted on statistically invalid principles (e.g. too small, too localised or only eliciting response from a certain category of person). Back in my motor industry days, as the Chairman’s right hand person, any surveys sent to the Chairman were passed on to me by his assistant for completion. I suspect a similar ritual is carried out in most head offices, possibly all.

Even the well conducted surveys do not necessarily measure what people really think and how they really act.

‘Over the 17 years I have been working in research... I have looked at hundreds of datasets. These purport to tell us what customers want, but in fact all they tell us is what customers have said. These are two very different scenarios, but I and many like me thought they were the same.’ (Alison Bond, ABA Research Ltd, writing in Customer Management magazine, 2004).

A vivid example of this phenomenon was the screamer headlines that told us that one youth in 10 carried a gun. It is extremely unlikely that if all British youths were searched 10% of them would be found to be tooled up. What the research really told us was that one youth in 10 *said* he carried a gun, which is a very different matter.

Babies and bathwater

In the drive for ‘scientific’ evaluation some very useful babies have been thrown out with the bathwater, one of which is the almost obsessive ditching of the ‘advertising value equivalent’ (AVE) yardstick. This has been witch-hunted by the purists as an unreliable method which does not measure the degree to which the hearts and minds of the audience have been captured.

As described above, the more ‘sophisticated’ methods are not very reliable in this regard either, but AVE is in any case not all that unscientific. The advertising world carefully researches the value of the different volumes and positions of coverage, and editorial is generally more read – and more believed – than advertising copy. It is therefore fair to assume that editorial coverage in the same space is worth at least as much as the advertising equivalent.

More importantly, though, AVE is still the single best way to impress the great majority of Britain’s directors, most of whom lack the sophistication to know what PR actually **is**, let alone the niceties of how it should be measured. And it doesn’t cost anything, which impresses directors even more! It is still a useful – and free – sales tool for PR, and more valuable as a benchmark than it is given credit for.

Where there *is* a problem with AVEs is in the multiplier effect. Commonly, evaluators claim three times the value of the advertising cost on the grounds that editorial is more effective. I even judged some CIPR awards where an agency claimed a fatuous multiplier of 18 times! Clearly this is a nonsense and AVEs would have a more viable currency if they used a low multiplier or none at all.

Another useful method of evaluation – vital even – is gut feel. The trouble with the new ‘scientific’ approach is that it excludes the right hemisphere of the brain, which can often tell us more about what is really going on than all the statistics in the world. In business and finance, the likes of Lord Hanson and George Soros did not make their fortunes at their computer screens.

‘As the stakes and complexity of decisions rise, managers are becoming increasingly addicted to using highly formalized decision-making processes to arrive at a conclusion. It’s especially true of larger organisations. But by doing this they are squeezing out judgment. That really concerns me.’ (Malcolm Gladwell, author of *The Tipping Point* and *Blink*)

As many top scientists recognise, a blind reliance on scientific methodology can actually detract us from the true picture. Indeed, Niels Bohr turned to Eastern mysticism for the answers to atomic theory and Roger Penrose has said that he can only reconcile

the paradoxical conflict between relativity and quantum mechanics with a subjective approach involving human consciousness. And:

*'I believe in intuitions and inspirations...I sometimes **feel** that I am right, I do not **know** that I am.'* (Albert Einstein)

Just when real science is starting to soften up, PR quasi-science is committing itself to the 400 year-old Newtonian/Cartesian approach.

Using the term 'gut feel' in PR circles these days can elicit waves of scorn and derision. In fact, gut feel can be a highly valuable management tool, but this era of process driven, left brain belief in our ability to measure and control things is pushing the use of body-mind into the background. Our brain can out-compute anything on earth. At any one moment, the number of possible permutations of thought and analysis available to us through the synaptic connections of the brain is greater than the number of atoms in the known universe.

It is unlikely that gut feel alone will do the job for us but the right combination of metrics and intuition is far more effective than either on their own. As one director of a major PR company said: 'show me the cuttings Darling, and I'll tell you if it's working!' I would place more faith in the combination of her years of successful experience and a pile of press cuttings than I would in a plethora of PowerPoint charts, graphs and assorted gobbledygook.

'If management expects to be able to predict and measure quantitatively in public relations, public relations people will be inclined to confine themselves to functions that are predictable and measurable. The best thinking and creativity will be lost.'
(Alfred Geduldig, PRSA Conference, 1993).

Sophistication v cost

As with the science that comes up with a different conflicting finding on the same subject every five minutes (remember dioxins in Scottish farmed salmon? Three simultaneous studies by three reputed institutes gave three totally different findings – when will we learn?) the method of measurement often determines the fact, and not vice versa. What you measure is what you get.

One way to minimise these variations and arrive at something nearer the true picture is through a process of triangulation, whereby you apply a range of different metrics and can reasonably assume that in the areas where they all say the same thing you are looking at something like the true picture.

Good stuff; getting better; but meanwhile, back in the real world, I dream of finding a client who will pay for it. A fundamental problem with evaluation is that the more sophisticated the analysis, the more it costs – and it is hard enough to obtain budgets for a press cuttings service, let alone a triangulation study.

In the 1990s Shell, its image severely dented by episodes such as Brent Spar and being accused of supporting a Nigerian regime that executed its opponents, undertook a long and thorough CSR programme. The effectiveness of the programme was ‘measured’ by frequent, extensive international opinion research studies. And sure enough, over a period of x years the company’s favourability rating improved by y per cent.

On the surface, this is a role model of evaluation and ‘proof’ of return on investment (ROI). But if you dig a little deeper the exercise, as in almost all such cases, actually proves little. The questions to ask are:

1. Over time our angry feelings towards a person or organisation tend to soften in the absence of further aggravation. How do you measure this tendency and what mathematical formula would you apply to factor it out of the survey findings?
2. If evaluating the contribution of PR to the programme, how do you measure how much of the improvement was because of PR as opposed to personal experience of Shell and its service?
3. To what extent did the act of going out and surveying people about their attitudes make them feel better about the company? How do you measure this well known effect and factor it out?
4. Was the percentage improvement satisfactory? How do you know what you could/should have achieved?

The costs of this particular research were more than anyone but a giant multinational could dream of, let alone afford. Yet even at this level of sophistication the ‘measurement’ becomes very fuzzy when scrutinised.

A final thought

Back in the 1980s, the inspired PR director of Nestlé UK, Alan Allbeury, spotted Margaret Thatcher drinking bottled Perrier at a televised news conference. He offered to provide her with a free supply of Nestlé’s Ashbourne water to avoid having a xenophobic prime minister drinking French water. Not only was the water British (though ultimately owned by a Swiss company, but we’ll draw a veil over that) but it had the British flag on the label, which she would turn towards the cameras at every

opportunity for the rest of her long tenure. I smile when I think of the evaluators trying to calculate the opportunities-to-see (OTS) rating!

Let us by all means strive for ever higher standards and seek to evaluate our performance. But please don't let's kid ourselves that this evaluation can be done scientifically. It can't. And please let's remember that the more we impose standards and norms on any aspect of our trade the more we could lose sight of what we are about in the first place.

'The intuitive mind is a sacred gift and the rational mind is a faithful servant. We have created a society that honours the servant and has forgotten the gift.' (Albert Einstein)