Cognadev Technical Report Series

7

17th June, 2022

The ROI of the Gallup Q12:

Assessing the true value of high-cost HR interventions

Cognadev UK Ltd., 2 Jardine House, Harrovian Business Village, Harrow, Middlesex, HA1 3EX. Registered UK # 05562662

Preamble

1. An intervention is defined as any HR initiative which seeks to intervene in the day-to-day functioning of employees and their work.

2. If money, time, or effort is expended on any activity within a business, then, unless that business is a charity, there should be a firm expectation that the expenditure can be assessed for its eventual return (profit) – expressed monetarily as a Return on Investment (ROI).

3. ROI modelling is defined to be the entire and complete costing of an HR intervention, taken together with an explicit detailing of the process of "how and when" the benefits of such an intervention are to be realised by a company.

Why model at all?

1. Because HR interventions can be high-cost. If these costs produce no measurable financial return, then those costs should be avoided.

2. Because most HR interventions cannot be costed in any straightforward manner. Thus, they will need to be costed and modelled as evolved "systems" operating over time in an organisation. By forcing HR intervention "champions" to make explicit both the evidence and assumptions that underpin the foundation for organisational implementation, computational models can be programmed that embody certain parameters from this evidence and assumptions. These models can then be evolved (run) over a specified time period in order to determine whether the expected returns can be demonstrated, and the conditions which may or may not affect an expected outcome.

3. Because some HR interventions involve "soft" effects that cannot ordinarily be measured by any simple process, test, or criterion-oriented variable. Thus, explicit assumptions will need to be made and modelled about how such "soft" effects might operate and diffuse throughout an organisation, such that expected or desired outcomes can be assessed for plausibility.

4. Because if HR cannot demonstrate that what it does has a measurable effect on a company's balance-sheet, then it becomes a cost-burden to that company. But, if it can demonstrate it is having an effect on the bottom-line, then it can potentially continue to improve and accelerate this effect in order that even greater rewards might be earned for the company.

A small example

I discuss some aspects of the Gallup Workplace Audit as an exemplar of some of the thinking and detailed calculations than can take place prior to the model construction process. The Gallup example is unique in that a strong body of evidence is brought forward to support its implementation. However, it is demonstrated that probabilistic evidence requires detailed understanding in order to fully appraise the likelihood and timing of the expected outcomes.

- ROI modelling is "value free".
- The modelling examines whether the ROI is generated via a "holding the line" effect against market adversity v "an increase in profitability" effect.
- The purpose of such modelling is not to criticise, attack, or defend HR interventions, company policy, or individuals. Rather, its sole purpose is to provide decision-makers with the means to determine whether expected outcomes for expensive interventions actually do occur over time – and to explicitly quantify the ROI of all such interventions.
- Such modelling will also enable optimisation and cost-minimisation strategies to be envisaged – given models can be evolved that incorporate such "tweaks".
- This is HR metrics in practice – although perhaps somewhat more sophisticated than most.

Some Initial Definitions

Intervention

It is defined as any HR initiative which seeks to intervene in the day-to-day functioning of employees and their work. That is, employees are employed by an employer to carry out a set of tasks and requirements. Employees will generally seek to do this as efficiently, productively, and easily as possible. Some may balance all three, some may favour the latter one over the other two. An HR intervention involves the external application of a process or system into this relatively private work domain of employees. I say *private* here because the contract of employment is between an employee and employer, with the requirements and expectations for a specific job to be done for that employer being presented by a named manager, as the designated agent of the employer. When an HR representative or department, external to this manager-subordinate employee relationship, decides that a group of employees should undertake some special process or take up a new system of activities at work, this should be viewed as an intervention into an existing work-oriented process.

Return on Investment (ROI)

If money, time, or effort is expended on any activity within a business, then, unless that business is a charity, there should be a firm expectation that the expenditure can be assessed for its eventual return (profit). This return need not be directly financial though, although it must be possible to show that there will be a financial return eventually, however indirect the path from the investment to the return. More generally, the ROI is expressed in percentage terms, as the economic gain divided by the total costs associated with an intervention or HR program (Cascio, 2002).

However, this disguises actual loss. If you spend $1million on an intervention, and see no financial gain, then the ROI by this calculation is 0.0. However, what this does not express is that in actuality, you have lost $1million. If you had kept it in a bank account, it might have earned 5% income over that year. If you had just put it in a suitcase under the bed – and income and profitability for the company remained the same at the end of the year – then that $1million might not have earned you any income, but neither did you waste it. So, I am defining ROI here as:

$$ROI\% = \left[\frac{(\$Gain - \$Cost)}{\$Cost} \right] \text{ x } 100$$

You can observe an ROI of -100% if you spend $1million but can detect no financial gain (say any improvement in profitability year on year). Alternatively, you can perhaps observe an ROI of +300% if the gain is 3 times the computed costs. Of course, the critical issue here is how to estimate the value for "gain". For example, when all other organizations' sales are plummeting due to market conditions, but yours plummet less – perhaps arguably due to the HR intervention, then do you still conclude "no gain" or perhaps compute that the "gain" is there, but now acting as a "buffer" against the more severe trends experience by other companies? This effect is quantifiable and can enter into such a calculation as that above.

Medium to High Cost HR

These constitute HR interventions that may cost up to $¼ million or more to implement across the corporate workforce (or some component of it). At a corporate level, this amount of money is no longer insignificant, although it may still be considered fairly trivial. Some HR projects though can cost several

millions to "roll out", especially leadership and performance management products taken across an entire international corporate. Whilst there is no fixed bound as to what amount might be considered a high cost for a corporate, it is always sobering to compute how much money needs to be earned by a company in order to utilise the profit as a cost-item elsewhere.

ROI Modelling

ROI modelling is defined to be the entire and complete costing of an HR intervention, taken together with an explicit detailing of the process of "how and when" the benefits of such an intervention are to be realised by a company. Such a model may be programmed into a computer, and evolved over time in the manner of an econometric model, or less deterministically, similar to that utilised within an artificial life model. It is only worth contemplating when the cost of the modelling, evaluation, and costing process itself is far less than the expenditure of the HR intervention. The aim of this modelling is to potentially save $millions of dollars over a short period of years, as well as augment corporate profitability by permitting re-investment or redirection of funds that might otherwise have been dissipated in a non-productive HR intervention.

Why "Model" at all instead of just calculating standard costs and benefits?

Well, unlike the rather direct approach to costing taken by Cascio (2000), and Fitz-Enz and Davison (2002) of the Saratoga Institute in the US, the kind of large-scale, high cost HR interventions that require modelling are those that cannot be costed in any simple "cookbook" fashion. Worse, their outcomes are barely defined in many cases, except as a series of non-quantitative statements in which there is no obvious direct link between "input" and "output", and little explicit information as to when their effects might be detected. However, all promise the kind of desirable high economic gain outcome that makes the initial high cost input seem a worthwhile risk. The problem of course is that rarely, if ever, is the outcome evaluated against those attractive expectations. I do not think that in every case this is a purposeful decision on the part of the HR professional, but rather, reflects the very real difficulty of trying to capture and measure what look to be almost intangible outcomes, and certainly the kinds of outcomes that are confounded with perhaps so many other variables that the attribution of causality (input A leads to outputs B, C, and D) appears to be non-discoverable.

However, the apparent complexity of the problem should not be taken as a signal that it is insoluble, or unworthy of the effort to force it to yield to a solution. If we approach the problem constructively, the initial phase of precise intervention costing (consisting of preliminary, setup, maintenance, and future implied costs) is relatively trivial although necessarily highly detailed. Cost is quantifiable as time, resources, personnel, and monetary expenditure. The "expected outcomes" are where we invariably face what looks to be an insoluble problem. This is because the expectations are left largely undefined, or defined so loosely that no quantitative outcome can be easily inferred, or are defined in such a way that the causal calculus between the inputs and outputs of a system is non-existent for any simple set of deterministic calculations. Further, such outputs may be said to evolve over time. That is, the expected outputs do not appear "instantly" at any point in time, but are expected to be observed gradually over time, spread across many areas of output. This nebulous evolutionary process eventually becomes the reasoned justification for not examining the ROI of the intervention.

Thus, instead of imparting to the CEO of a corporate with "it should work, but I can't be very specific about how or when it will be seen to work", the promise of ROI modelling should permit the CEO to view a rational model of the intervention, its total costs, how and when its expected benefits will be

detected within the company, and especially the expected ROI trend over a specific duration of time. Further, such a model can be evolved over time. That is, the model parameters and expectancy functions may be evaluated from a starting and finishing date, using an initial set of values to begin the process. In this way, a decision to "go, not go" might well be possible prior to the actual implementation of the suggested HR intervention. Alternatively, the intervention may be implemented, but then modelled later in order to develop baselines for the detection of expected positive outcomes (and the subsequent "flagging" of potential negative trends that may suggest the increasing likelihood of intervention failure).

ROI computational modelling is politically neutral and judgement-free

All it does is expose the reasoning, attempt to quantify features, and model the hypothesised processes of HR interventions such that the proposed benefits of an intervention may be assessed, either in advance of implementation, or post-implementation. It is designed to permit the recognition of explicit success or failure of HR interventions, such that decision makers can be provided with an evidence-base for making more rational decisions about the proposed implementation, continuance, or otherwise for a high-cost intervention.

Although no-one has yet attempted HR ROI computational modelling in the manner indicated above, the computational modelling of some specific HR-oriented organizational behaviours such as absenteeism, counter-productive behaviours, and some team dynamics has been modelled with some degree of success (Hulin, Miner, and Seitz, 2002; Ilgen and Hulin, 2000). However, this is leading-edge research that is barely out of the starting-block. Further, it is quite possible some large organisations may be undertaking such research, but are not permitting the academic publication of any results.

Some examples of HR interventions as worthwhile targets of ROI modelling.

1. The Gallup Survey
2. A Performance Management System
3. A 360-degree process
4. An Incentive-Bonus Scheme for Senior Management

Let me outline briefly the kind of approach that might be taken with the Gallup Survey.

The Gallup Q12 Survey

Over the course of the past 30 years, researchers with the Gallup Organization have conducted thousands of qualitative focus groups across a wide variety of industries. The approach underlying this research has been founded on what might be called "positive psychology" (e.g., Seligman & Csikszentmihalyi, 2000)— specifically, the study of the characteristics of successful employees and managers and productive work groups. In developing measures of employee perceptions, Gallup researchers have focused on the consistently important human resource issues on which managers can develop specific action plans. An instrument was developed from studies of work satisfaction, work motivation, supervisory practices, and work-group effectiveness. The instrument, the Gallup Workplace Audit (**GWA**; The Gallup Organization, 1992–1999), is composed of an overall satisfaction item plus 12 items that measure employee perceptions of work characteristics. These 13 items were developed to measure employee perceptions of the quality of people-related management practices in business units. The criteria for selection of these questions came from focus groups, research, and management and scientific studies of the aspects of employee satisfaction and engagement that are important and able to be influenced by the manager at the business-unit or work-group level. An article (Harter, Schmidt, and Hayes, 2002) presents a meta-analysis of studies conducted by The Gallup Organization to calibrate the instrument's relatedness to business-unit outcomes, generalizability across organizations, and usefulness in differentiating more effective work groups from less effective ones in relation to a variety of desirable business outcomes. The GWA was designed to reflect two broad categories of employee survey items: those measuring attitudinal outcomes (satisfaction, loyalty, pride, customer service intent, and intent to stay with the company) and those measuring or identifying issues within a manager's control that are antecedents to attitudinal outcomes.

In a nutshell, the aim of the organizational survey is to permit item responses to be "benchmarked" against a database of such responses from many organizations. Given a positive relation has been previously empirically determined between the level of response and a positive business outcome, the idea is to determine where your organization stands in relation to others, and then seek to develop intervention strategies that will subsequently lead to an increase in GWA ratings next year, with the implied benefit that the causes of such ratings changes will indeed lead to more positive business outcomes such as increased profitability.

The Gallup Q12 Evidence Base

An impressive body of data (Harter et al, 2002) is presented in support of this argument, with near 8000 business "units", within 36 companies, and almost 200,000 employees being part of the "benchmark" sample. Outcome measures were variables such as profitability (% of revenue), productivity, employee turnover, and composite performance. In every case there is a positive ***statistical*** advantage to obtaining a higher score on overall job satisfaction and overall employee engagement. That is, the higher the score, the higher the profitability, the lower the turnover, the higher the productivity and the higher the customer-satisfaction/loyalty ratings.

But, that word "statistical" needs to be completely understood. What this indicates is that the positive advantage being claimed has been observed as a probability statement, and not as a statement of certainty. For example, increasing a Gallup score on employee engagement is associated with a *probability* of observing higher profitability. If this probability is high, then one can have increasing confidence that the effect may well be observed in one's company in which the Gallup procedures are followed.

Let us for the moment stay with how Gallup report their results – in terms of "effect sizes". Quite simply, an "effect size index" in our context is a measure of how much impact a certain intervention has on an outcome. So, if .0 represents *no effect*, and 1.0 represents *absolute certainty of an effect*, then what do we conclude if the effect size is just .10?

This in fact is the observed effect size (Pearson correlation) between profitability and the mean score of the GWA items. We can correct this value for the unreliability associated with the averaging across companies and profitability estimates. The correlation increases to .11. We can also correct it independently for the unreliability associated with averaging across business units and profitability estimates. The correlation now increases to .15. Finally, we can make a joint correction for all such sources of unreliability, and the correlation increases to **.17**.

But, these corrected values are somewhat problematic, as LeBreton, Scherer, & James (2014) have recently noted. For what we are saying when we make them is "**if there were no measurement error**, what would be the true value of a relationship between two variables". We impose a condition that cannot be met in reality in order to demonstrate the absolute theoretical maximum for a variable relationship. This is perfectly acceptable when used with experimental test psychometrics, or for theoretical information about the nature of observed relationships. But here, we are trying to assure a CEO that the relationship is almost twice the size of what actually has been observed, even though the relationship can never be demonstrated in practice (without using perfectly reliable measurement).

A demonstration of a specific kind of modelled ROI

Let me demonstrate graphically what these correlations look like. I generated 100,000 data values, sampled from a bivariate normal distribution composed of pairs of Gallup GWA mean scores (12-60) and profitability values (0-100). The population bivariate relationship is fixed as per Gallup at **.17** *(the theoretical population correlation)*. The complete *a priori* parameterisation/constraints and generated parameters are as follows:

Correlation and Sample Size

	Specified	Actual
Required Correlation	**0.170**	0.169532
Sample Size	100000	100000
Number of Plot Points	10000	10000
Number of Iterations	200	7

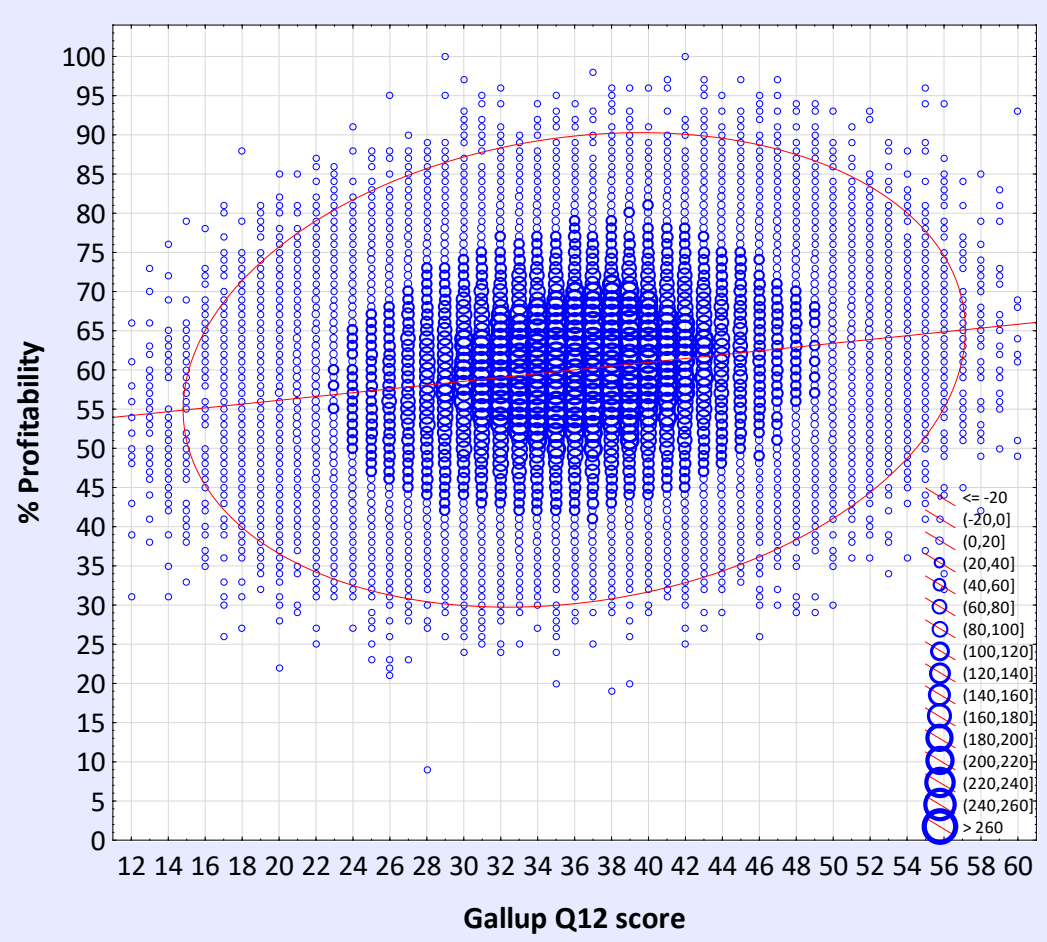
X-axis Setup — Gallup Q12 Score

	Specified	Actual
Mean	36	35.982740
Std. Dev.	7	6.977047
Minimum	12	12.000000
Maximum	60	60.000000
Diagnostic Axis Title	☑	
Integer Values Only	☑	

Y-axis Setup — % Profitability

	Specified	Actual
Mean	60	60.011380
Std. Dev.	10	9.975276
Minimum	0	9.000000
Maximum	100	100.000000
Diagnostic Axis Title	☑	
Integer Values Only	☑	

A *frequency* scatterplot of all 100,000 sample points, with a 99% confidence ellipse, is presented below:



What we see from the above is that a Gallup mean score of 40 may be associated with profitability that ranges from 29% through to 96%. Likewise, a very low Gallup mean score of 28 is associated with a profitability range of 24% through to 97%. The red ellipse indicates the boundary within which we would be 99% sure of finding all sampled values, **if** our sampling was of a perfectly normal bivariate distribution (which has been assumed in all Gallup calculations involving effect sizes). Further, note that this is a good as it gets for these particular pairs of variables as we are already working with a solely theoretical maximum effect size.

It is possible that your company might be in the top half of the ellipse in terms a relationship between increased profitability and increased Gallup Scores. The concern is that as your company demonstrates an *increase* in Gallup scores year on year, it is nevertheless related to a *decrease* in profitability. Using our 100,000 sampled pairs of observations, if we managed to obtain an increase in the mean Gallup score from 36 to 42 *(or somewhere in between)* in the space of one year, we can calculate the probability of observing an *increase* in **profitability** above the mean value -vs- observing the mean value or less given a correlation of .17 between the two variables. This calculation tells us:

Probability of an increase in profitability = .52
Probability of a decrease in profitability = .48

We have only a slightly better than chance probability (**.50**) of increasing our profitability by increasing our mean Gallup score year on year, given two assumptions:

1. the relationship between profitability and Gallup Mean score remains at 0.17 on occasion 2.
2. we assume that the interventions applied to companies with many individual scores of 36 at Occasion_1, only change in line with the global population effect.

Neither assumption may be reasonable, but this is a question that can be answered using empirical data. What it does show is that nothing can be taken for granted here - as the capacity for a "lack of effect" may be probabilistically quite high. In short, it seems essential to continually monitor the Gallup "effects" for their expected vs actual outcomes. If, over 2 years in your company, the increases in Gallup scores do not seem to be reflected in any hard performance indicator (productivity, sales, profitability), then you might have spent over $1million for no apparent reason or result. The ROI will be negative, severely so.

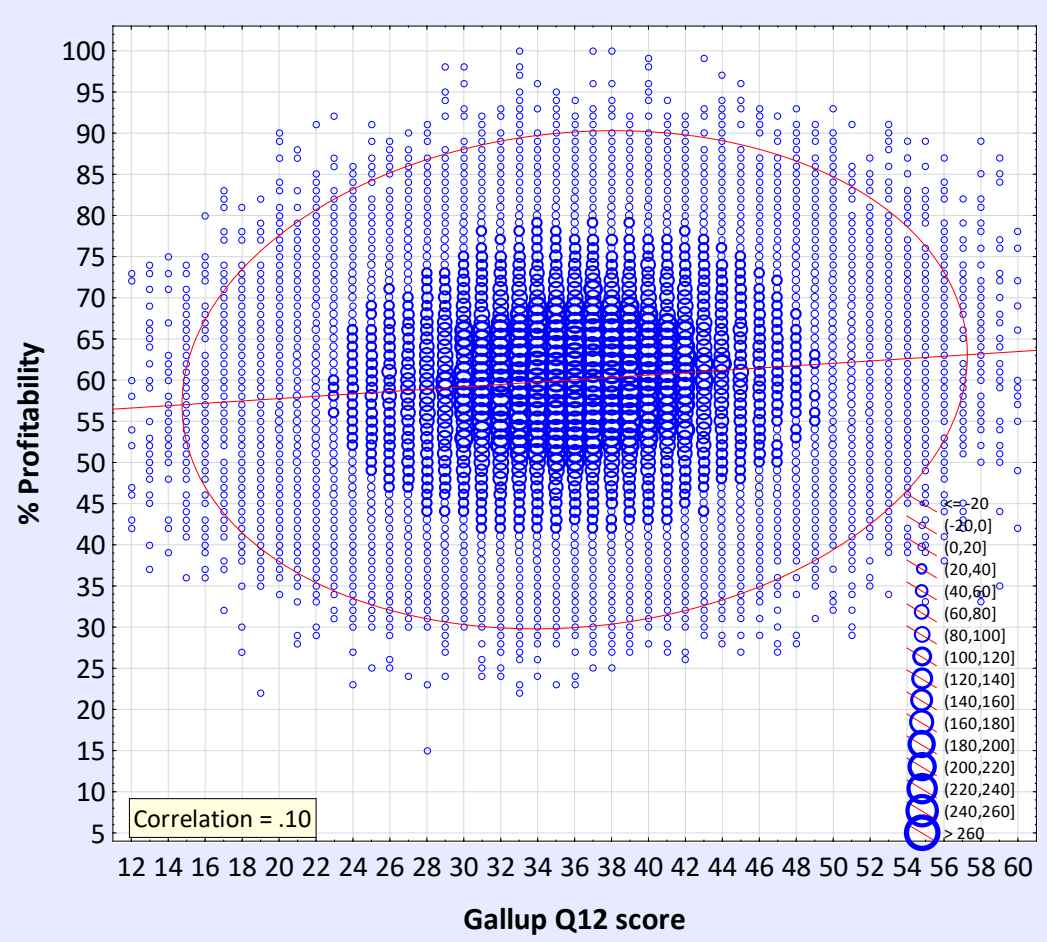
If we focus on a strategy of attempting to increase the Gallup Q12 engagement scores, year on year, assuming an improvement from the baseline mean score, we see the following:

	Estimated correlation of **.17**		Observed correlation of **.10**	
New annual Gallup Q12 score above the mean of 36	**Probability of equalling the mean score of 60% or less profitability**	**Probability of an increase in profitability above the mean score of 60%**	**Probability of equalling the mean score of 60% or less profitability**	**Probability of an increase in profitability above the mean score of 60%**
37	.50	.50	.506	.494
38	.50	.50	.51	.49
39	.49	.51	.51	.49
40	.48	.52	.49	.51
41	.47	.53	.48	.52
42	.45	.55	.486	.514

I included the probabilities for the expected observed correlation magnitude of just **.10**, with all generating parameters the same as for the correlation of .17; as this is the relationship one would expect in practice rather than in the abstract. With this magnitude of relationship, it is not until we can improve the Q12 scores from a baseline mean score of **36** to **52** that the odds of improving profitability perhaps begin to look worth the financial risk:

Probability of an increase in profitability = .60
Probability of a decrease in profitability = .40

The frequency scatterplot of this .10 relationship is:



What is actually required here is some computational modelling of the relationship between Gallup score and outcomes, within and between business units within your company, and their longitudinal relationship over time (test occasions). It might be that the Gallup "effects" are only detectable after several years, or that a threshold effect is present in the data (below a certain threshold the relationship is weak, above this and almost certain benefits ensue). Further, the relationship between the cost of implementing HR interventions targeted to produce increased Gallup ratings, and the "dose-response" relationship of these to the Gallup score changes and the subsequent outcome benefits should be computed and monitored. It may be that the cost of raising a score exceeds the potential benefits to be attained. By modelling the expectations and known error-rates from available Gallup meta-analytic data, it should be possible to indicate the gradient point whereby the cost of interventions can be seen to exceed the financial benefit (direct or indirect) of any outcomes.

Also, by examining the evidence-base for Gallup in so much detail, it might be possible to identify areas of expected outcome for which no empirical data actually exists. This could indicate that seeking to improve Gallup scores is currently proceeding in your company with little specific *a priori* evidence that increases year on year in these scores have any effect whatsoever on the criteria of interest. This is not necessarily a bad thing to do, but by making explicit the "research" basis of the expectations, you then know it is working in "research mode" rather than from a strong evidence base – *with all that implies for implementation and risk of failure*. So, by forcing a modelled link, however tenuous and convoluted between an input and output, we can discover which expectations are based upon empirical evidence and which are based upon inference alone. If too many are based upon inference, then this again begs the question of what exactly might or should be the ROI of the Gallup survey, given various scenario analyses.

Does the above address whether or not the Gallup survey is currently working within any specific company?

No. Rather, it shows the kinds of areas that need to be investigated and modelled in order to demonstrate the likelihood or actuality of a substantive ROI for the use of Gallup. If there is no return – or in fact a negative return – then why continue with the investment at all? Over 5 years, the full cost to a large organization might be in the order of $7-10million dollars – taking into account the full and complete cost of implementing the Gallup survey and subsequent HR interventions implemented throughout each following year in an effort to raise 'engagement' among employees.

By judicious use of modelling, we can perhaps state that the utilization of these funds is producing a financial return, or terminate the intervention altogether if no return appears detectable, and is unable to be detected in a longitudinal evolved expectancy model of Gallup effects. The same kinds of thinking and analysis can also be focused on the three other HR strategic interventions.

For HR management to attain the strategic status it has desired for so long in the boardroom, this is one area in which it can engage professionally as a leading member in the "strategic business conversation" with others around that boardroom table.

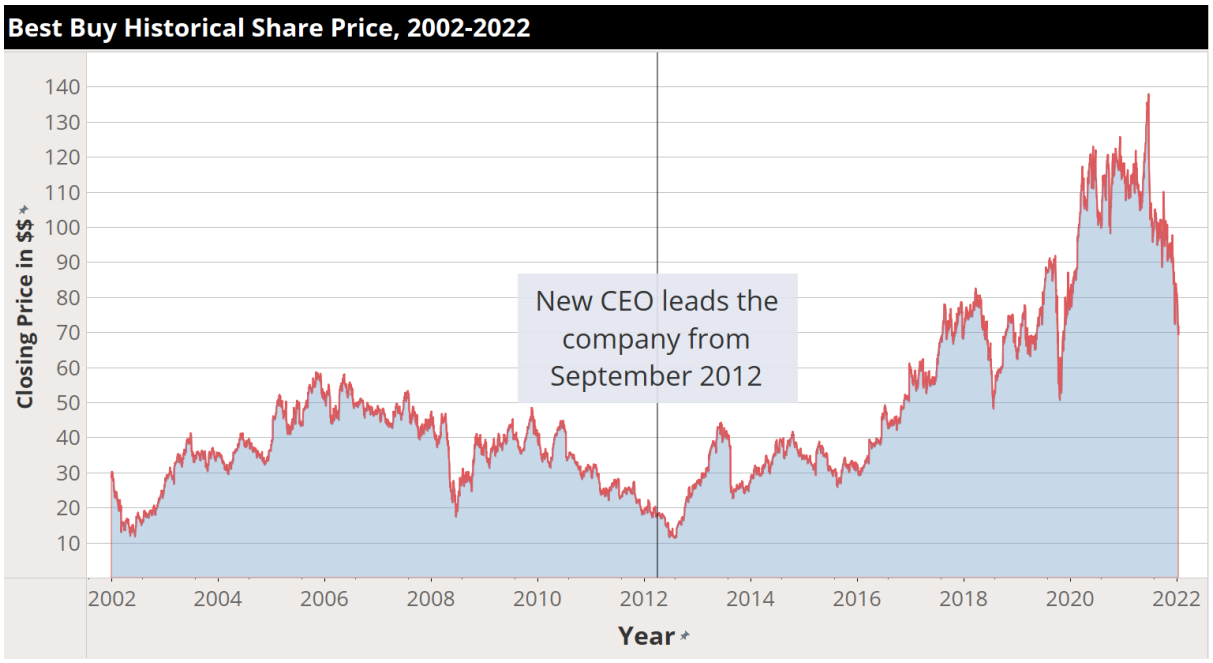
The lesson from Best Buy Inc.

The message from Gallup was that Best Buy was a poster-company for its engagement assessment, scores within the top 10% of the benchmarked norms with profitability to match. But I noted this good news was from around 2003 – 2010. In a 2011 story on the financial benefits of employee engagement in the magazine Chief Executive:

> "… Best Buy found just that. Higher employee engagement scores did lead to better store performance. The company found that for every 10th of a point it boosted employee engagement, its store saw a $100,000 increase in operating income annually. (CFO Magazine, "Measuring Up," June 26, 2007).

I took a look at the share price, corporate accounts, and company trading matters of Best Buy Inc over historical and current years. If engagement scores were this high so consistently, then the company should be doing really well if engagement was to be seen as causal for organizational performance in a non-trivial way. Using data and graphics supplied by Yahoo Finance, the following share price data for Best Buy Inc., from 2002 to 2012 was:

The company was in trouble come 2012, announcing the closure of 50 stores in the US and the laying off of 400 employees. Given the size of the company these figures are not so significant as they seem, but the ominous 'writing on the wall' for Best Buy was in this simple sentence from an analyst's report at the time: "industry watchers contend that Best Buy stores increasingly serve as physical showrooms for online retailers." People were going into the stores to examine products physically and check them out for functionality, then bought them online. Amazon Inc was killing Best Buy. All the employee engagement in the world could not stop the decrease in profitability caused by senior executives not realizing (until too late) how Amazon would change consumer buying habits. However, the organization had recovered somewhat by 2016 and the share price continued its overall upward trajectory until recently.



Was that dramatic improvement the slow-release diffusion of 'engagement' throughout the organization – or as Harvard Business Review reported in a 2016 article entitled: "Reinventing Best Buy", this change of fortunes was solely due to a new CEO leading the company from 2012?

References

Barrett, P.T. (2002). *Evidence-Based HR: Can it be done?* NZ Psychological Society Annual Conference, Christchurch, NZ.

Cascio, W.F. (2000). *Costing Human Resources 4th Edition*. Ohio: South Western College Publishing.

Cascio, W.F. (2002). *Managing Human Resources 6th Edition*. New York: McGraw-Hill

Fitz-Enz, J. and Davison, B. (2002). *How to measure Human Resources Management 3rd Edition*. New York: McGraw-Hill

Harter, J.K., Schmidt, F.L., and Hayes, T.L. (2002). Business-Unit-Level Relationship between employee satisfaction, employee engagement, and business outcomes: a meta-analysis. *Journal of Applied Psychology*, 87, 2, 268-279.

Hulin, C., Miner, A.G., and Seitz, S.T. (2002). *Computational Modeling in Organizational Sciences: Contributions of a Third Research Discipline*. In F. Drasgow and N. Schmitt (Eds), Measuring and Analyzing Behavior in Organizations. San Francisco: Jossey-Bass

Ilgen, D.R. and Hulin, C.L. (2000). *Computational Modeling of Behavior in Organizations*. Washington: American Psychological Association

LeBreton, J.M., Scherer, K.T., & James, L.R. (2014). Corrections for criterion reliability in validity generalization: A false prophet in a land of suspended judgment. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 7, 4, 478-500.

Seligman, M.E.P. and Csikszentmihalyi, M. (eds.). (2000). Special Issue on happiness, excellence, and optimal human functioning. *American Psychologist*, 55, 1.

The Gallup Organization (1992-1999). *Gallup Workplace Audit* (Copyright Registration Certificate TX-5 080 066) Washington DC: US Copyright Office.

Technical Note

The simulations were conducted using my free-to-install/use Corviz correlation data simulation software, coupled with further analysis and graphics using Statistica 14.