

On the Measurable

Roger Cohen, July 2012

Executive Summary

Some musings on **what gets measured gets made**. Humans are fascinated with metrics and measurement. The mere act of measuring something is enough to influence behaviour, sometimes in ways unexpected. Examples from sport, technology, finance are used to illustrate this.

Heisenberg's Uncertainty Principle states that both the position and momentum of a particle cannot be simultaneously measured with infinitely high precision. The mere act of measuring will interfere with the position and/or the momentum of the particle at the time of measurement.¹ This principle applies not only on the atomic scale, but is readily evident when we observe human fascination with measuring, cataloguing and comparing.

That we are fascinated with measuring, cataloguing and comparing is nothing new. It is evident across all ages, cultures and timescales. Children spend inordinate amounts of time counting and comparing their toys, boasting about what is bigger, stronger or faster. Games like [Top trumps](#), where a collection – real or fictitious – has various enumerated attributes. Players compete by selecting the winning attribute for the card they hold. The skill lies in picking the attribute which is relatively stronger, when an absolute does not exist. Of course, luck plays a part as the cards are initially randomly distributed. [Winning Moves](#), the company behind Top Trumps sells millions of packs annually, and operates internationally. In adulthood, we are a nation of sports enthusiasts, market watchers and lottery participants. We are fascinated, not only with outright scores and outcomes, but with all sorts of statistics. Batting averages, hit rates, error rates, second serve aces, there is a statistic for almost anything that can be measured or recorded. Furthermore, we are not shy to bet on sports and many other outcomes. Betting has evolved from simply picking winners, into the ability to wager on almost anything observable.

Individuals, and also governments and organisations are huge producers and consumers of statistics. From polls to studies, referendums and surveys; all types of data are collected, collated and disseminated. As well as being fundamental to the way our country and corporations are run, competition, performance measurement and many other outcomes are based on statistics. Our fascination with the measurable and the comparable, as well as providing an understanding of outcomes, has produced many unintended, and sometimes catastrophic consequences.

What gets measured gets made. The old adage is universally applicable. True when incentivising sales people, setting performance metrics and guiding behaviour patterns, it has also produced a fair share of unintended outcomes. The wise salesman who is measured by the number of calls made will maximise this number, irrespective of whether these calls lead to actual sales, or not. A friend and I realised in our early high school days, that our relative academic esteem was raised by the frequency and choice of library books we borrowed. My borrowings far exceeded my capacity and interest in reading. To this day, I still haven't read War and Peace, or the Lord of the Rings Trilogy. I did however

¹ <http://hyperphysics.phy-astr.gsu.edu/hbase/uncer.html>

have a very good relationship with the librarians, and was highly regarded as a prolific reader. With the internet, we have become fascinated with hit rates, page views and other web traffic statistics. Because of the ease of measuring these, extensive data is collected. Not all of it translates directly to desired outcomes. Search engine optimisation; strategies to actively drive traffic to websites, and to particular pages within those sites; and measuring the habits and patterns of users of the web are now disciplines in themselves. Voting by expressing a like/dislike of a post, following a writer, blogger or Tweeter is another means of measuring the consumption of information. Whether it is accurate is highly debatable, but it is widely used and accepted. These measures, in turn become motivators of strategy and design of internet related activity, and a badge of success or failure. Often, this becomes self propagating. Once the audience passes through a threshold, the mere volume of hits generates further interest. This positive feedback generates more hits. The target “goes viral”. This is no different to the epidemiological behaviour that diseases exhibit. Below a critical threshold, a piece of information will essentially die (or just exist in isolation). Above this threshold, it is guaranteed to attract an audience. Information aggregators aid this process by presenting material based on how much it has been viewed. This has no bearing on the quality of the information, or on its value if measured in ways other than by mere audience size and frequency.

Away from the internet, the measurable and the made are also evident. Another topical focus is on the way front office staff in banks and financial institutions are rewarded. The **call option** effect, where large revenue based bonuses become possible, has been a driver of risky and short term behaviour. It has also been blamed for observed excesses, blowups and even negligent or criminal behaviour. As a consequence, considerable effort is now going into designing metrics to prevent such risk taking behaviour. Fair remuneration models which encourage the “right” behaviour on trading desks are very topical. The same focus is being applied to CEO and senior executive remuneration. Although this has led to increased transparency, and to modifications of so called best practice, it is as yet not clear what the effect will be, or whether this will indeed be successful. Human behaviour is unlikely to change – especially when the rewards are high.

Despite much focus on aligning metrics with results, there are numerous situations where this is not happening. In some of these situations, the metrics are still not aligned with the type of behaviour that is being promoted, in other cases, it is too difficult to measure with immediacy the effect that metrics and benchmarks have on behaviour.

Recently, there has been a focus on “big data”. With technological advances, we collect masses of data whether it be spending patterns, observations or measurements of systems. Data is logged for later use. Loyalty card schemes are designed as a means of monitoring individuals spending habits. The rewards or loyalty bonuses offered are just hooks to obtain this information. Much information is just stored for later use. It is not yet known what will be derived from it. Indeed, as more data is collected, and as computing power increases, more sophisticated analysis will be performed. Who knows what patterns will emerge.

The fascination with measuring and comparing means that we continually seek things to measure. As a child, I remember comparing cars on the basis of brake horse power (in fact this harks back to one of the first Top Trumps packs I possessed). I thought that the more of this horsepower stuff a car had, the better (faster, more powerful) it was. Later I learned that power to weight ratio was a better measure. It was a much less spectacular measure, in that it was the combination of two things – power and

weight. As such, it didn't directly follow the "bigger is better" principle that applied to brake horse power. With the proliferation of computers, many new measures of power have proliferated. In the early days of personal computers, processor speed was the all important number. The more megahertz one had, the better was their computer. Benchmarking programs were created to directly measure this. Over time, it was realised that memory capacity was just as important, as it was vital for moving data quickly. Not enough memory would throttle any machine even if it had the best megahertz count. So RAM (Random Access Memory) became a new measure. Now, superseded, when it was realised that overall storage capacity, parallelism and many other factors, influence the speed of a computer. Today, connectivity is important, so network bandwidth is one of the new metrics. And it won't end here.

Another digital metric that is misused is the megapixels we apply to cameras. We are still at the stage where more megapixels are better. Any camera or device with a camera associates will display a megapixel rating. A good article by Ken Rockwell [The Megapixel Myth](#) explains the misuse of this metric better than I can. The conclusion is that more is not necessarily better. More information is needed to make a comparison if image quality is the desired result.

Besides being obsessed with the ability to measure and compare, we must not ignore that comparison is still a real, and most important, means of assessing similarity or difference. Obfuscation is thus purposely introduced to many arenas, to subvert efforts at making comparisons. Have you ever tried to compare mobile phone contracts. Even offerings from a single provider are almost impossible to fathom. Each plan includes different combinations or bundles of services. More minutes, free texts, after hours usage, calls to same network, voice and data together, with or without a handset, and so on. So many combinations of so many different parameters make simple direct comparisons impossible. Further, comparing plans across providers is an order of magnitude more complex than comparison of the offerings of just one provider. Supermarkets have (usually by law) normalised prices for comparison purposes. Just looking at various offerings of a product (such as shampoo or tea bags or fruit juice to name a few) it is not possible to directly compare prices. Size, weight, packaging and presentation all differ. Normalisation, by displaying a cost per unit, whether it is weight, volume or per standard piece goes a long way towards making comparison on a like for like basis possible.

Spurious metrics abound. The trade off between what is easily measurable, and what the measurement is intended to provide can and does influence behaviour. Several more examples follow:

Various forms of property taxes. The number of windows in a dwelling was from the 1600's, used as a proxy for household wealth in parts of Europe. This practice was widespread, and continued into the 1800's. Some detail can be found in regencyredingote.wordpress.com, and www.projectbook.co.uk with a lot more on the web. In parts of England, windows were bricked up to keep their number below the taxable threshold. In Amsterdam, a tax based on the frontage of a house saw the building of some almost impractically narrow houses – now tourist attractions. Derived as a simple means to raise revenue in proportion to the scale of a property, (which was assumed to mirror the scale of its inhabitants), these forms of taxation had an effect on architecture. The term "daylight robbery" may have been an outcome of the window tax. Even in retrospect, it is hard to measure whether such taxes were effective, especially if the detrimental effect of reduced natural light, frontage or other aspects is taken into account. These sorts of tax proliferated as they were based on easily measurable and

auditable characteristics. Enforcement was easy, and perceived to be not open to manipulation. Despite this, the outcome was skewed by the architectural modifications that took place.

Online share trading has spawned a thriving CFD (contract for difference) market. In the UK, CFDs are the outcome of stamp duty levied on trading in UK listed shares. Originally, the stamp duty was introduced to raise revenue by taxing share transactions. Although CFD's deliver the same economic exposure as trading in physical shares, they do not carry any stamp duty as they are not share transactions in themselves. CFD providers are generally exempt from stamp duty for transactions related to hedging CFDs. Trading thus moves to exempt participants, and the revenue potential of taxing all share transactions is diluted. (This is especially evident when looking at the way global hedge funds trade in different markets. Predominantly, their UK activity is CFD based, while in other markets, they may use CFDs or physical transactions via a prime broker). Whether this tax is effective or not is difficult to determine. It certainly has allowed a host of new products to be developed, and influenced the behaviour of market participants who just want the economic outcome of trading shares, but do not necessarily want to physically own them. See for example www.bullbearings.co.uk.

Value at Risk or VAR is another example of an easily measurable quantity whose interpretation diverges from its intended purpose. The idea of using VAR was to provide simple numbers which measure the risk inherent in an instrument, a trading book, a portfolio or even a whole organisation. Because VAR can be systematically calculated, and aggregated it has become a widely used measure of risk. Banks used VAR internally to measure risk in various business units. Aggregated, it was used by regulators and managers, as a measure of risk in an organisation. VAR has many shortcomings. It is based on simplistic assumptions – including that the returns of a portfolio are distributed in a well understood way. This does not mean that the return of any given position can be predicted. It just assumes that on a probabilistic basis, the distribution of returns can be described. What has resulted, is that the distributions in VAR models were not representative of what has actually occurred. So called “[black swan](#)” events or fat tails mean that events which statistically are seen as highly unlikely actually do occur somewhat regularly. This means that VAR is not a good measure of real risk under real conditions. Perhaps the simplicity and theoretical elegance of the methodology overshadowed its shortcomings.

Many more instances abound. Without going into detail, here are a few: In financial markets, fund products are described in terms of management fees or expense ratios. The perception that these must be minimised has led to obfuscation and misinformation. Products are developed with the aim of showing a low headline fee, which may or may not be a true reflection of the costs embedded in the product; benchmarking by using market indices influences the behaviour of markets rather than measures it in some cases; Social media relies on crowd generated statistics as a measure of popularity. I find it hard to believe that Lady Gaga and Justin Bieber, each have in excess of 25 million followers on Twitter, while President Barack Obama has a mere 18 million; Commission free foreign exchange is widely touted by name brand providers at airports and popular tourist locations. How many users of these services realise that the rate they are charged is usually several percentage points away from the fx rates quoted and used by market participants. This list is not exhaustive.

We know that if it can be measured, it will be made. A corollary: **If it can be measured, we will compete**, is just as true. We human beings tend to compete in almost everything we do. Some competition is necessary for survival. For example, protection against direct threats, or competition

for food where there is not enough for the whole population. In the present day, living in a well developed society, almost all competition is by choice. Where we play a sport, a game, or even use a random outcome have become competitive. Reward can vary from recognition to substantial financial gain. It seems however not to matter what the reward is, we, as a species still seem to want to compete. Derived competition, where we become stakeholders in an outcome which we can't influence or control is also prevalent. Most visible is betting, or wagering on sporting events and other observable outcomes. Whether we are competitors or observers, we still thrive on measuring and observing, and taking a stake in our desired outcome.

In summary, we seem to need to measure, make comparisons, and become stakeholders in outcomes, whether or not we have input into them. It seems part of our nature to actively seek out things to measure, compare and use for purposes of both competition and observation. Whether we are protagonists in measurable activities, or just observers does not seem to influence our passion for observation and measurement. Not always does this make sense. Understand that just by introducing a measurement we can alter behaviour in unintended ways. Not only does what get measured get made, but what gets made is bound to get measured as well.