



Facts and data have always been the lifeblood of Continuous Improvement.

One might expect that with such easy access to more facts, figures and computing power, today's improvement projects should be yielding astonishing, almost unbelievable results!

But the presence of such overwhelming amounts of data doesn't always translate into breakthroughs. Are you taking advantage of "big data"? Has it given you better insight and helped you make better decisions? Or are you missing the mark?

Mary Jane Conway-King

Old Challenges, New Possibilities

By: Sheila Julien, Senior Associate

"In God we trust; everyone else must bring data."

It has been over 50 years since Dr. Deming began pounding the data-drumbeat. In the ensuing decades, with help from Bill Conway and other talented leaders we have seen companies the world over make astonishing progress by using data to study and reduce variation, reduce costs, reduce defects and all while increasing the use of the talents of the people closest to the work.

But oh, how the world has changed! Access to data has been transformed beyond what was imaginable even a decade ago. The sheer volume of data is staggering. For example, Wal-Mart collects more than 2.5 petabytes of data every hour. A petabyte is one quadrillion bytes. Speed of access has increased so that a lot of data that was once only available monthly is now available throughout the day. And we now have an astonishing variety of data forms – texts, tweets, blog posts as well as data from sensors, scanners, even cell-phone locations. For example, the MIT Media Lab used location data showing the number of cell phones in Macy's parking lot on Black Friday to infer the sales on Black Friday even before Macy's had recorded the sales.¹ What's more, the cost of storing large amounts of new data is rapidly declining.

To what uses would Dr. Deming or Bill Conway put all the new possibilities we now have if they were starting their work today? What new tools do we have to help us in process improvement?

This article explores the amazing potential of some of the new technologies available today as well as some hazards we face and the new capabilities we will have to develop as we try to harness 'Big Data' to make better, faster, management decisions. It will also discuss how some of the new technologies can help us cultivate new capabilities in our organizations and to unleash a new wave of internal collaboration and innovation.

New Possibilities for Old Challenges

Better Customer and Market Information

The classic analytical methods of understanding our how we can increase value for our customers include continually analyzing customer complaint logs and lost order reports and using that data to target improvement opportunities. We may also gather episodic information from a focus group, surveys or sampling and observing some of our customers using our products or services.



Today we can do all that and much more!

Measuring and Analyzing Customers' Satisfaction and Perception

Everyone who is focused on improving quality tracks customer complaints so that they are able to identify patterns and prioritize ways to improve services or products. But for every customer who takes the time to call you, many others do not. In the old days, these people would pass on their opinions to friends and family, but today some of them may tweet their views or post a review on-line for the world to see.

The changing technological landscape now gives businesses unprecedented access to customers' opinions. We can learn what people are tweeting about us with Twitter searches, and we can read what they post on Yelp or Google reviews. We can study patterns for early warning about an emerging performance or a perception problem.

We can also study the same information about our competitors to identify their weak spots or any emerging threats.

Deeper Customer Relationships

Google Alerts can also help us stay close to our most important relationships or prospects, potentially identifying developing business needs for our products or services. For example, a Google Alert that a local company has just had a banner year with rapid revenue growth may signal an emerging credit need that their banker could address proactively. News that one of his clients has announced a merger may signal a supplier of network services to proactively visit to discuss an emerging need for additional networking.

Matching Offerings and Advertising to Customers' Needs and Interests

Traditionally we use surveys and/or focus groups to help us choose what ads or offerings would create value. At best, our offerings target specific audiences on a very rough scale – such as customized ads in trade journals or mailing to purchased lists. But every marketing communication that misses its target is waste. Every need we could have satisfied but did not is a wasted opportunity. The better able we are to predict needs and interests for a segment or an individual, the more waste we can eliminate from mismatches between our advertising information or products and the customers' needs and interests at the time.

Today companies can provide a website where customers or potential customers can seek out specific value adding information. A bank's website might offer a link to home buying advice; a computer consultant's website may have links to problem solving tips or to an article about how to choose a network provider; a canoe manufacturer may have links to articles about different types of paddling: white-water, flat-water, sea-kayaking. While this provides valuable information to potential customers, tracking the clicks can provide information about who needs what types of products and services. If a client clicks 'How to Save For College,' we now know something about her current needs and interests that allow us to share the most relevant marketing messages and offerings.

We can also study patterns in order to better predict the interests of similar customers. Amazon.com pioneered this approach of reporting to the customers what similar shoppers looked at or purchased, to simultaneously add value for their customers while boosting their own sales.

Expanded Public Data to Address Problems and Opportunities

The rapidly expanding sources of public information also create new opportunities for research and problem solving. For example, we now have data available on adoption rates of certain technologies, housing, demographic changes, weather patterns, and countless other types of information that a company can use to learn more about the problems and opportunities in their industry, if they know what to look for. For example, it has recently been estimated that if we were able to make effective use of the growing volume of health information data now available, we could improve efficiency and quality at a rate of \$300 billion a year in the United States alone.



Increased Agility from Speed of Info

If data can be collected and analyzed from digital interactions with customers, then organizations can learn from, and quickly respond to customers' interactions. For example, a change in the pattern of customer clicks could tell a business how much their customers were influenced by specific promotions, reviews, and/or page layouts. In the past, we would look at monthly sales figures to evaluate if an impact was made, but results could easily be influenced or offset by other factors. A statistical analysis could much more quickly identify if an approach has been successful. With the existence of today's data about customer actions and interactions, an organization can design experiments with the website, promotions, layouts, etc. to continuously improve their forms, offerings and messaging.

New Challenges

But with all the potential for using data for improvement, we also face quite a lot of potential for 'big data' to lead to costly mistakes. Here are a few of the hazards we must be careful of.

Confirming Biases

Ample data can lead to selective data mining and actually increase confidence in erroneous conclusions. In his book, [Thinking Fast and Slow](#), the psychologist and Nobel Lauriat in Economics, Daniel Kahneman, demonstrates how when human beings sift through large amounts of information, our pre-existing opinions significantly influence our perception of the data. Data that supports our pre-existing opinions are far more likely to be remembered or found credible, while contrary data are far more likely to be forgotten or judged less credible. So the more data you have, the easier it is to find in it whatever supports your previously held position – right or wrong.

MIT researchers McAfee and Brynjolfsson report exactly the same phenomenon in the field. They report observing managers using 'big data' to support opinions arrived at independent of the data: "Too often we saw executives who spiced up their reports with lots of data that supported decisions they had already made using the traditional 'Highest-Paid-Person's Opinion' approach. Only afterward were underlings dispatched to find the numbers that would justify the decision." As Nate Silver points out in the introduction to his new book, *The Signal and the Noise, Why So Many Predictions Fail but Some Don't*, when access to data explodes, our biases tend to amplify the noise faster than we can objectively discern the signal.

The Law of Small Numbers and Other Cognitive Illusions

Even absent any bias from a previously held position, the human habit of inferring causation from patterns exposes us to serious mistakes when dealing with truly random occurrences. This is especially true when dealing with small numbers. We know that if we flipped a coin 1,000 times we would feel pretty confident that it is no more likely to be heads than tails. If we flipped a coin 10 times it would hardly be surprising to get a 60/40 split or one even more uneven. If we did not know the underlying probabilities, it is quite likely we would draw an incorrect conclusion about the pattern.

When we analyze data, we do not have the benefit of knowing in advance which data points are the signal and which the noise – or random variation. 'Big data' cuts both ways. The quantities of data better inform our panoramic view, but we also wish to make use of more granular insights. The smaller the sample size, the more likely a purely random distribution will indicate a pattern – and sometimes a convincing pattern. Daniel Kahneman has a short and illuminating article in [The Week](#) with several powerful examples of how our minds can fool us into completely erroneous conclusions about random events, especially with small numbers of observations. He concludes: "The law of small numbers is part of a larger story about the workings of the mind: Statistics produce many observations that appear to beg for causal explanations but do not lend themselves to such explanations. Many facts of the world are due to chance, including accidents of sampling. And causal explanations of chance events are inevitably wrong."



Our minds are not accustomed to statistical thinking, and we are too quick to identify a pattern in a series of random events. Furthermore, once we have discerned a pattern, our preconception bias causes us to consistently strengthen that bias by noticing and remembering the data or instances that confirm our erroneous conclusion and cause us to systematically overlook or underweight the contrary data. A study reported in [Freakonomics podcast](#) suggests that the more numerically sophisticated a person is, the more he or she is prone to this systematic bias. If we are to *benefit from* rather than *fool ourselves with* big data, we will need to develop the methods, skills, and discipline to correctly interpret the data we have.

More New Challenges: New Capabilities Needed

It is hard to overstate the importance of aligning the organization's 'big data' objectives with its capabilities, processes, and culture. An organization can waste a lot of time, money, and organizational energy jumping on the 'big data' bandwagon before their internal capabilities, processes, and needs are aligned with the data collection and analysis they are attempting. The organization's execution wing must be aligned with the new information so that they are able and willing to act on the new information before the potential benefits will accrue. As has been well documented in the business press, many companies lost a lot of time and money on CRM systems because their organization's processes, practices, and personnel were not aligned to benefit from the valuable new customer information and technology.

To avoid investing in data and information that does not produce smarter decisions and execution, keep things as simple as possible. Focus on answering specific and important business questions, and start by building the capability to answer those. In his inimitable way, Dr. Deming once observed to a client that they would benefit more from \$2k worth of brains than from the \$2MM computer they were considering acquiring. Identify the few vital questions, the information required to answer them, and how you will apply the answers to create value for your customers, grow the business, improve quality or cut costs.

Competitive advantage will arise not so much from the answers you have but from the questions you ask. McAfee and Brynjolfsson echo Deming and Conway in advising executives to challenge the data. They advise that executives should "start with two simple techniques. First they can get in the habit of asking 'What do the data say?' when faced with an important decision and following up with more-specific questions such as 'Where did the data come from?' 'What kinds of analyses were conducted?' (People will get the message quickly if executives develop this discipline.) Second, they can allow themselves to be overruled by the data; few things are more powerful for changing a decision making culture than seeing a senior executive concede when data have disproved a hunch."

In addition to asking good questions, organizations will need to develop much deeper capability with data analysis, especially the study of variation and the ability to distinguish special causes from random variation in order to arrive at sound answers. This includes the obvious need for people with a facility for the nuts and bolts of statistics, but further, a need for statistical understanding and thinking at the managerial level. Many organizations have people at the managerial level who are not comfortable with statistical analysis and do not know how to evaluate its soundness or how to use it effectively.

Nate Silver describes three fundamental new skills we must develop to make better use of data as a predictive tool. The first and most important skill is to think 'probabilistically.' Rather than looking at a problem to determine the 'right' answer, probabilistic thinking means attempting to understand what different possibilities are in the 'feasible set' and what the associated probabilities are. This could mean understanding what the probability curve would look like and whether 'most likely' carries a probability of 40% or of 80%. Understanding the range of uncertainty will lead to far better managerial decisions than arriving at a single answer.

The principles Silver describes as second and third both advise the incorporation of additional information as



it is available. His second principle is to keep our mind open and revise our conclusions if we have new or different information. John Maynard Keynes is quoted as saying, “When my information changes, I alter my conclusions. What do you do, sir?” Silver also observes that incorporating multiple perspectives, the third principle of better forecasting, also improves accuracy. “Quite a lot of evidence suggests that aggregate or group forecasts are more accurate than individual ones, often somewhere between 15% and 20% more accurate depending on the discipline.” (He adds, “That doesn’t necessarily mean the group forecasts are good.”)

Not only do we need better skills with statistical analysis and managerial skills, we also need people with skills with the new technologies. Much of the newly available information comes in forms that traditional IT groups are unaccustomed to and that traditional databases are not well-designed to handle – such as tweets, posts, clicks, etc. Developing new capabilities with a variety of data will be essential to using this data to improve our understanding of the customer.

New Possibilities to Address New Challenges

It is fine to say we must increase our capabilities with data, statistical analysis, and new technologies, but how? People with these skills are hard to find, training budgets are often tight, and evening classes are often impossible for employees because of distance or other responsibilities.

How fortuitous that we have recently seen an explosion of new ways that people and organizations can build technological and quantitative analytical skills for free!

Quantitative, Technological, and Managerial Skills

Top quality and free on-line educational opportunities have recently taken a giant step forward providing interactive education with tests and even certificates of completion. Alison.com offers hundreds of excellent tutorials for free in technology, accounting, finance, human resources, production management, etc. KhanAcademy.org has hundreds of outstanding tutorials and practice problems from pre-algebra through differential equations and linear algebra. They also offer tutorials in finance and economics, computer science, and humanities.

Coursera.org offers free classes from some of the best teachers at Stanford, Princeton, Columbia, Duke, and many other excellent universities around the world – and with a new twist to on-line education. Unlike its predecessors, Coursera.org classes follow a calendar timeline with homework, due dates, quizzes and a final exam. Some, but not all, will present a ‘certificate of completion’ or ‘completion with distinction’ to a student who has satisfactorily completed the homework and exams. Never has there been so much opportunity with so few obstacles to increasing organizational capabilities.

New and expanded skills will be sorely needed as we move into the era of ‘big data,’ and these tremendous tools for on-line learning could not have come at a better time.

Innovation and Collaboration – Idea Blogs, Crowd Sourcing, Wikis

The development of technical skills is just one of the new tools we have to accelerate process improvement. Many organizations struggle to surface ideas and insights that employees have about improving the work. Some companies address this with suggestion boxes, but these have some drawbacks: the ideas require management attention (often spread too thin already) to read, evaluate, and to respond to each idea; many ideas are focused on low impact issues; and a suggestion box lacks interactivity, so a good idea that is not clearly stated is likely to be lost.

All that is transformed with today’s technology.



Idea Blogs

Today, organizations can use internal communication technology to create idea blogs to overcome the drawbacks to the traditional suggestion box. Questions can be raised by readers so that the submitter can clarify or others can chime in. Barriers to implementing the idea can be raised and the idea can be refined and improved to address the barriers.

Crowd Sourcing

Perhaps most valuable of all, the organization's talents and creativity can be focused on the organization's most pressing challenges by applying an internal version of 'crowd sourcing.' With crowd sourcing, problems or challenges are broadcast to a large group with a call for solutions. People in 'the crowd' submit solutions. The contributor of the winning solution is, in some cases, paid for the work or awarded a prize and sometimes the compensation is simply recognition. For example, a grocer may seek ideas for keeping displays full and fresh during the hours when traffic is light. A manufacturer may look for ideas to reduce scratches in final assembly. An operation may look for ideas to reduce the backup at the bottleneck. Of course, these problems should first be posed in production meetings with the people closest to the work, but where they remain unsolved, posting the challenge to 'the crowd' with a video of the operation to analyze – may surface new ideas.

Wikis

A 'wiki', sometimes defined as 'the simplest online database that could possibly work', is a repository for information, documents and communication. Organizations can create internal wikis where project teams or special interest groups can collaborate, store and share information, keep project and product development moving quickly with quick and consistent communication among the entire team. Confluence is an internal wiki one of our clients uses to help their process improvement teams collaborate more effectively and there are others readily available as well.

Once again we find ourselves at a new frontier. As we seek to apply the new capabilities to our timeless charge of continuously studying and improving the work, forever, we will find some breakthroughs and some setbacks. If you have additional examples of ways newly available technology can further problem solving and process improvement, email us. In the meantime, below are some links to some tools and tips that might be helpful.

[Dave Nelson's Ten Do's and Don'ts of Social Media](#)

<http://www.polleverywhere.com/>

<http://www.educrations.com/>

[1] Harvard Business Review, "Big Data: The Management Revolution", McAfee, Brynjolfsson October 2012