BY BRANDON THEISS

NUMB3RS ARE NOT

ENOUGH

Improved manufacturing comes from using quality data to make the right connections and conclusions

The era of social networking has enabled a person sitting in a café in New York City to obtain up-to-the-minute "status" updates about a friend or colleague in Shanghai. This is in stark contrast to the industrial paradigm of considering information to be timely when it is reported days or weeks removed from its collection. Clearly the tangible, actionable value of immediate real-time data about the organization's processes is much higher than the disposable "tweets" of acquaintances regarding their morning coffee. Yet in the latter case information flows faster, across greater distances and to a wider audience. This paradox suggests that despite the ubiquity of technologies enabling the instantaneous flow of quality data, many organizations continue to restrict the flow of information, consequently operating at suboptimal levels.

An organization can be defined as the summation of a large number of different inter-related processes at different levels of abstraction. "Process" in this context is not limited to the scope of manufacturing engineering, but it can be applied more broadly to any verb-noun combination. The beauty and elegance of lean Six Sigma and other continuous improvement initiatives is that the same set of tools, techniques and methods can be used across the full range of processes from drinking a cup of coffee to the entire operation of a multinational corporation. The shared commonality is that a sequence of steps is performed to produce a desired output. However, the tools of lean Six Sigma are limited and constrained by the quality of the measured data.

For data to be considered "quality," it needs to be accurate, timely and complete. "Quality" is being used in the context of Joseph Juran's definition "fitness for intended use" and W. Edwards Deming's "meeting or exceeding customer expectations." To collect quality data, an organization must maintain a critical level of process intelligence at its most fundamental levels (the manufacturing floor, restaurant dining room, brokerage sales desk, etc). This intelligence can be manifested in the human capital of the operators or the physical capital tools and systems of production. The critical yet often sub-optimally implemented — or completely omitted — tool of production is the data collection system.

Getting the numbers

Data collection does not need to be particularly elegant or expensive. Sometimes, an operator with a pencil and clipboard will do. Other situations necessitate a comprehensive RFID-based supervisory control and data acquisition (SCADA) system. A cost-benefit analysis must assess the synergies and trade-offs between investments in hardware and/or human capital. Regardless of the allocation model, personnel must receive the necessary tools and an effective, timely feedback loop to facilitate and encourage proper behavior. To that extent, the type of data collection and the attendant cost considerations must facilitate that success while continuing to add value to the enterprise.

For operators working on a manufacturing floor, instruments need to be available at the right time, in the right location and be capable of producing the desired work product to the requisite specification. The operators need to be trained sufficiently and empowered to take ownership of the process so they can maximize the hardware's efficiencies to produce a quality product. At higher levels within the organization, the need for having the proper tools, training and ownership of the process remains paramount. In management, the process can be defined as decision making. The output of the process is the overall success of a project, team, division or enterprise. The inputs to this process are data. A conflict happens when there is a direct and quantifiable connection from the output (decision) to the bottom line, but the value of inputs cannot be directly inferred.

This contradiction arises from the fact that data in the abstract have little value. Only after connections and well-reasoned conclusions are drawn from the information can "numbers on a screen" be translated directly to the accountant's profit and loss statement. In the "new normal" of the current macro-economic downturn, accounting and finance teams are unwilling to allocate funds without documented hard savings and a short-term return on investment. No longer can projects be justified by soft savings, the hard to measure intelligible or long-term benefits. This presents a

MANUFACTURING STARTS AND STUTTERS

U.S. manufacturing continues to recover from the Great Recession, although supply chain disruptions following the July earthquake in Japan hampered second quarter results, according to the U.S. Federal Reserve.

Manufacturing output was unchanged from May to June. For the second quarter, the manufacturing index edged up at an annual rate of 0.2 percent, a smaller gain than in any quarter since the recession ended in the second quarter of 2009. Capacity utilization for manufacturing was unchanged in June at 74.4 percent. That's still about 10 percentage points above the June 2009 level but 4.6 percentage points below its average from 1972 to 2010.

Source: Bloomberg

case of the often cited "chicken or the egg" paradox. The question becomes, "How much should be invested in measuring A when its impact on B might not generate enough additional benefit to cover the cost of measuring A?"

Albert Einstein's quotation, "Our theories determine what we measure" provides insight into determining the answer. In science, the goal is to confirm or refute a theoretical model. In industry, the costs, purposes and uses of collected data must be tied directly to the enterprise's profitability. Many such models for industrial systems exist. The strength of the model is only established through the course of collecting data and testing the predictive validity of the theory. In the industrial setting, the data can be collected only if capital has been allocated; capital can be allocated only if the project's value has been pre-determined. This is impossible, of course, and represents an example of the closed-loop feedback system of the causality dilemma. The solution to this dilemma is to a priori declare the value. The inherently stochastic nature of an industrial enterprise makes this determination more of an exercise in creative writing than hard science.

This inherited random nature of industrial systems encourages using the statistical underpinnings of Six Sigma. The methodology must be understood as a set of statistical tools that do not yield absolute truth. The warning that "correlation does not imply causation" can be extended to all of the results of the often-computationally intensive analyses. The prototypical example is to attempt to correlate the rainfall in a certain part of the world with a major stock market average. Given the proper data set, a very high correlation value can be determined. Yet there clearly is not an underlying theoretical or physical connection between the variables.

As the mathematics become more involved, the number of variables increases and the interaction of the variables becomes harder to model and understand. The often erroneous conclusion is that since the statistical methods are rigorous, there must be a fundamental theoretical physical connection between the variables. The statistical tools are most powerful when they are used to focus efforts in a direction to explain the fundamental physical relationship or confirm understanding, but their value always will remain constrained by the quality of the input data.

Culture can kill

Data accuracy is not purely a function of the precision of the gauge used to take measurements. Lean Six Sigma introduces the concept of a Gage R&R (repeatability and reproducibility) to attempt to address this fact. However, this statistical tool only measures the amount of variation in the measurement system arising from the measurement device and the people taking the measurement. It does not address the culture of the organization collecting the data.

In one industrial example, a company manually mixed bulk materials to make a slurry. The batching process consisted of an operator using a Bobcat skid loader to dump bulk material into a large industrial mixer. At the beginning of the shift the operator was given a pre-printed form that prescribed the exact amount for each of the bulk ingredients. Prior to beginning the batching process, the operator moved the Bobcat onto a floor scale and recorded the empty tare weight of the loader. Then the operator navigated the Bobcat to the pile of bulk material, scooped up a load, navigated back to the scale to record the weight on the pre-printed form, and then finally to the mixer to load the material.

After completing an ingredient, the operator would use a pocket calculator to add up the respective weights, recording them on a form. This cycle continued for all of the ingredients in the recipe. Once the recipe was completed the operator turned in the form to his supervisor, who in turn submitted the form to the quality lab. The material being batched was used as an input to another manufacturing process, which was quite tolerant to batching errors. Process parameters could be adjusted easily to accommodate upstream errors if they were reported. Yet this manufacturing process had constant quality problems. The supervisors in manufacturing blamed the batching area for delivering an inferior product. The batching supervisors presented the batching records indicating that the batches always were made to specification. When a batch was made improperly, the operator was reprimanded severely and in most cases fired.

To the trained practitioner and the casual observer alike, this situation is riddled with potential error points. The operator manually writing data, using a hand-held calculator, improperly identifying the materials and handing off the physical form multiple times are opportunities for the data to be corrupted, compromised or lost. Other forms of "muda" include the wasted motion of the operator having to drive to the floor scale prior to the delivery of each load.

However, the true root cause of the problem did not lie in any of these process-related defects. Instead, the problem was the management response to reported defects. By virtue of using the Bobcat's coarse dispensing system, exact amounts cannot be dispensed. But the pre-printed form the operator used stated exact recipes out to a fraction of a pound on a scale that counted by 2-pound increments. The operators knew that reporting the actual amounts dispensed would jeopardize their job. Consequently, only the "correct" values were reported. The root cause of the quality problems in the downstream processes was the absence of data that could have been used to change process parameters and avoid the quality defects.

Liberating data

Unfortunately, the prior example is all too common. Instead of information flowing freely from ground levels of an organization to the upper levels of management, arbitrary barriers are erected. These barriers inhibit the necessary dialog between colleagues, managers, groups and divisions, further reinforcing a culture of suspicion and distrust. This culture spreads like cancer such that when data flows properly, suspicions about its quality discredit decisions based upon the information.

Data latency is an unfortunate characteristic of the human condition. For instance, it can take up to 20 minutes after you stop eating for your body to indicate that it is full. This delay has led many to overeat. To exacerbate the situation, it can take up to eight hours to digest food and convert it into usable energy. Likewise, the value of last week's newspaper primarily is limited to lining bird cages and wrapping fish. Even though the facts may be valid and the stories compelling, the passage of time has depreciated the information's value. Likewise, the value of industrial information declines exponentially as time passes. This does not stop many companies from using data collected weeks earlier to make decisions that affect the present.

An industrial example of data latency was discovered in the final assembly/packaging operation at a multinational manufacturing firm. Upper management had noted through a benchmarking exercise that one plant's productivity was dramatically less than similar plants in the region. An investigation compared the standard operating procedures, line layouts, tools, training and staffing of the lines at the respective plants. Minor differences were observed and changed to adopt the best practices of the higher performing plant. This resulted in minimal, incremental improvements, and the plant continued to underperform.

Management next dispatched a team of interns to observe the line and record the "up time." After collecting data around the clock for a week, the team determined that the line only met its target cycle time 65 percent of the time. Management was aghast. A group of Six Sigma practitioners was dispatched. Their process map revealed that the last step before packaging was to scan the parts' unique barcode, which for accounting purposes reported the product as a finished good.

The team members smugly thought they had identified the problem's root cause. They would use the tamp stamp information, which recorded when the piece was scanned, and then record the amount of time until the next part was scanned. The team created a simple software application to extract this elapsed time and create an X control chart. Area supervisors only would be required to explain the out-of-control conditions as shown directly in the chart. After recording why the line stopped, separate projects would fix each of the identified problems.

The practitioners proudly presented their solution to management. Management, partially blinded by the jargon and graphs, mandated that the charting software be used by the area supervisor at the completion of every shift. No one was to be released from the line until every out of control point was explained. The data collection operated for two weeks, and the practitioners gathered to look at the results. Much to their surprise, the data was incomplete and inconsistent.

The team had failed to consider several important factors. First, at the end of a 10-hour shift the operators and their supervisors would be anxious to go home and took shortcuts in completing the paperwork. Second, the team did not provide a standard lexicon for classifying the problems. Each supervisor classified the same stoppage differently. The most important factor the team failed to consider was that the operators simply could not remember why the line had stopped nine hours prior.

The team regrouped and developed a rather novel solution to invest in improving the machine intelligence. The team modified the system that scanned the barcodes on the line. If the operator failed to scan a barcode in the allocated cycle time, the system would lock and display a pop-up message. The message would display manning, materials, machines or other. The operator then would select the classification for line stoppage. For example, if the line was stopped because of stock-out of a box, the operator would select materials and then from the next menu select packaging.

The software was released into production for two weeks. Oddly, while the system was in testing, management observed that productivity increased. This uptick in productivity was disregarded as a Hawthorne effect induced anomaly. When the practitioners gathered to review the collected data, they were pleasantly surprised.

The data conclusively showed that the line's productivity had increased steadily each day. When graphed as a Pareto chart, the stoppage causes showed a clear trend. But why had the productivity increased? The system only recorded data; it did not fix any problems. The team was confused, so its members went to the production floors and asked the operators.

The operators explained that they were annoyed when the pop-up message appeared and never had a sense of how much time was being wasted. As a result, the team on the line worked proactively to fix potential stoppages. This rapid closed-feedback loop allowed for the system to self-correct. Even with this self-correction, the plant's performance still lagged. A formal analysis into the Pareto data discovered that the primary cause of line stoppages was that the operators did not know what material was coming next. Instead, they simply reacted to the material that was delivered by the upstream process.

This was simply a data flow problem with an equally simple solution. Large-format displays were installed in the packaging/final assembly area to give operators the estimated time of arrival of material. The setup was similar to flight monitor displays that let people in airports know arrival and departure times. This enabled the operators to know exactly what products would be packaged up to four hours before entering the line. This provided ample time to ensure that all components were available in sufficient quantities. After implementing the system and raising the system intelligence, the plant became the leader in productivity for the entire company.

The problem of data latency is not a new concept. Jay Wright and his colleagues at the MIT Sloan School of Management captured the need for data to be collected and communicated in their Beer Distribution Game, a simulation created in the 1960s to demonstrate key supply chain management principles. This study in the bullwhip effect depicts how a system will break down when information does not flow properly from the fundamental level (retailer) to consumers of the data (distributors) and finally on to upper management (manufacturer). Omnipresent technologies that facilitate this flow of information are ubiquitous in our common life, yet this problem remains intractable for industry.

Uncorrupted and complete

Even if the data is accurate and flowing in a timely fashion, the tools of lean Six Sigma will fail if the data set is incomplete. Take, for example, the Indian parable of the blind men and the elephant. A group of blind men touch an elephant to learn what it is like. Each one feels a different part, but only one part, such as the side or the tusk. The men then report their findings and learn that they are in complete disagreement, and no one has described an elephant. Although all of the men reported accurate and timely information, the information led to erroneous conclusions and decisions.

This situation was epitomized in the case of a high-volume manufacturer that had made a large capital investment and implemented a comprehensive ERP/MRP system. The system literally drove the enterprise by automatically translating customer orders directly to the production schedule and manufactured pieces. The system enabled management to have a complete picture of the enterprise from a single interface and extract granular information of the flow of material from the raw material arriving through manufacturing, supply chain and ultimate sales to the end user.

However, the Achilles heel of the system was the quality of the input data, a consequence of the convoluted process of collecting data. The operator would record by hand the number of parts produced. At the conclusion of the shift, these notes were turned in to the area supervisor, who would either personally key the information into the ERP system or delegate the task to a data entry person. The paper notes were lost frequently; there were accidental and deliberate transcription errors that caused the ERP to respond erroneously to changing the production schedule. If a sheet was lost, the ERP would drive more production of the part, though in physical reality the parts had been produced. Conversely, if a penmanship error over-reported production, the ERP would not schedule production.

This created an error-amplifying feedback loop that cost the company several million dollars prior to being corrected. This example shows that in the absence of data or in the presence of incomplete or corrupted data, the application of lean Six Sigma methods will result in erroneous decisions and the misallocation of resources.

Garbage data, garbage results

Data enables an individual, company or society to rise from the bondage of myths and half-truths. The tools, technique and methods of lean Six Sigma help to reveal these truths, yet they are limited by the input quality of the data. Inserting garbage data into the most elegant of calculations will result in nonsensical analysis and decisions. Getting enough of the right data, at the right time, to the right people is the constantly evolving challenge that must be met. Technology now facilitates the real-time flow of this information. The technology can be implemented and process improvements achieved only if an enterprise is willing to make the investment.

Brandon Theiss is a master black belt at American Standard Brands, where he is responsible for the continuous improvement efforts for the four chinaware facilities in North America. Prior to joining American Standard, Theiss worked as a systems engineer at Johnson Scale Co., where he was responsible for the design and implementation of process control and data collection systems. He has been published in Quality Progress *magazine. Theiss earned a B.S. in electrical and computer engineering and a B.A. in physics from Rutgers University, where he is pursuing an M.S. in industrial engineering.*