



STATISTICAL START-UP

by

Dr. Jeffrey T. Luftig

(Edited by Bob Pahlkotter)

A low/no cost approach to equipment qualification and optimization

IN 1985, WHILE WORKING WITH IBM, GENERAL MOTORS, AND FORD MOTOR COMPANY, it was discovered that a significant need existed in industry. The requirement was a need for a disciplined, low cost approach to the successful installation of new equipment. In a number of cases when new equipment, lines, and plants were “started up” it took months or sometimes years (and in a few cases, never) for the installed systems to reach a projected or acceptable level of performance related to operational cost, reliability, maintainability, internal and/or customer quality and delivery requirements. Clearly, what was needed was a method that could be utilized to avoid these conditions, while at the same time minimizing the additional time or cost required for the design and installation process.

Developing a discipline and methodology to accomplish these goals, in the late 1980s a number of projects were successfully piloted. First, in a series of automotive plants and then followed by a number of installations in the aluminum, steel and food industries. Since the discipline is heavily focused utilizing statistics and statistically-based sciences, the term was coined “*Statistical Start-Up*™”. To date the methodology has been applied and utilized both inside and outside the United States on such equipment as aluminum cold mills, rice mills, tin plating lines, coating lines, metal slitters, and various other production lines.

Historically, companies have tended to design and purchase equipment with a minimum of requirements, mostly around *functional testing* for the expected performance characteristics for the production process. Additionally, while some of the required product quality characteristics might be mentioned in purchase agreements, only *acceptance testing* would be conducted. Oftentimes, the results of such a minimal design and planning process would lead to serious and catastrophic results.

Functional testing, refers to that portion of the engineering and installation process where the equipment is installed, “turned on,” and tested to determine if it is operating properly as designed and specified (e.g., speeds, feeds, gauging, etc.).



The term *acceptance testing* refers to a level of analysis beyond functional testing that speaks to whether the machine or line is performing at a minimally acceptable level. Basically, *acceptance testing* as a disciplined analysis is utilized to;

1. Compare the product as manufactured to the required characteristics as previously specified; and
2. Compare the operational or performance characteristics as observed to the conditions specified by the contractual requirements, and/or the parameters upon which the equipment was originally justified.

Much of the acceptance testing that's done relates directly to the performance of the equipment, and often has nothing to do with the product. However, an appropriate assessment of acceptance testing involves a number of conditions.

The first category of conditions that are normally subjected to acceptance testing is what is referred to as *Product Quality Characteristics*. These characteristics, which usually have been defined with a target or nominal value with an upper and/or lower specification, describe the attributes associated with the product. Examples of these attributes would include thickness, length, width, I.D., O.D., coating weight, viscosity, taste, surface quality, and color. Usually, these quality characteristics are defined by the customer(s) for the product.

The identity of those quality characteristics that are related to the finished product, are often referred to as *End-of-Line (EOL)* quality characteristics. Quality characteristics that the customer has indicated, or that we have identified as being **very** important to form, fit, function, use, or safety considerations are often termed *critical* or high risk elements as is sometimes used in the automotive industry. Usually, most companies installing new equipment or lines will minimally assess the *critical EOL quality characteristics* as defined by the customer(s) for the product. When we observe products/units that fail to meet one or more specification(s), we refer to the unacceptable product or unit as *nonconforming*.

The second category of conditions that are usually assessed in acceptance testing are *product defects*, or *non-conformities*. Defects are those product conditions by which their presence in terms of frequency, severity, or both, render the product unacceptable, or *defective*. Common examples of defects include scratches, gouges, blemishes, voids, slivers, cracks, contaminants, dents, and blisters. Again, most companies installing new equipment or lines will minimally assess a defined number of products or units to assess whether these defects are present, and at what level.

The third category of conditions that are generally assessed in acceptance testing are *Equipment Performance Requirements*. This category includes the equipment performance measures or *Key Performance Indicators (KPI)*, which determine the productivity and therefore, the associated cost of a production process. Luftig & Warren International uses a comprehensive model that describes the system in relation to Total Asset Utilization (TAU), and includes elements that describe and quantify these performance metrics. The elements that are employed include *Availability, Efficiency, Duty Cycle, and Yield, or Recovery*. The calculations for each of these

components are fairly complex, however it should be pointed out that the Availability is broken down into a number of sub-components, including *Reliability* and *Maintainability*. Most professionals working in this field recognize that common measures related to these two sub-components are respectively, *Mean Time Between Failure (MTBF)* and *Mean Time To Repair (MTTR)*. Many companies installing new equipment will generally try to assess the condition of the newly installed equipment or line in terms of these two conditions, as well as efficiency and duty cycle(s). However, some companies—surprisingly—still don't measure all of these elements to verify that they received what they paid for.

It's the fourth category that historically does not get a lot of attention in acceptance testing for new equipment, and this relates to the analysis of *Product Performance Characteristics*. This category includes criteria or requirements associated with how the product performs in use—typically at the customer's location or while in the hands of the ultimate consumer. Examples of these types of measures include such conditions as formability, mobility, line speeds, in-process rejection rates, and stamping die wear rates; again, at the customer's site. Experience indicates that many companies simply haven't tried to include these characteristics into their acceptance tests for new equipment for production lines. Part of the explanation for this is that it requires a company to understand the interrelationship between their customer's performance requirements and their product's quality characteristics. Companies who haven't partnered or had early involvement with their customers in the design phase, or who haven't implemented a Quality Function Deployment (QFD) process, wouldn't have this data base, and would, as a result, have difficulty building it into their acceptance testing process. As early involvement efforts and cost reduction requirements become more and more important in the future, the ability to perform this analysis will become much more important to obtain new and retain existing business.

If acceptance testing is the *only* testing that is executed, the results can be catastrophic. In other words, there's nothing wrong with acceptance testing, *unless* it's all the testing that is going to be done. It is equally important to understand the *limitations* of acceptance testing. In other words, what inferences can be made about the process?

Assume a new plating line has started up and the goal was to start-up the line so that all functional tests were satisfied, and to pass acceptance tests according to Specifications Checklist. (*Refer to the Specifications Checklist to review the product and equipment requirements*).

Next, suppose that 10 coils were produced and inspected during the first 80 hours of production. Imagine that the acceptance tests yielded the results shown in Figure 1. The results indicate that all categories have passed inspection. However, now observe Figure 2. The figure shows the results of acceptance testing conducted three months later. Although the figures show different results, the line is exactly the same as it was when the acceptance testing was conducted the first time. The fact is that all of the acceptance test results are representative of what one would expect to see from a sample of 10 coils and 80 hours of scheduled production. ***One can't make any inferences about future states or conditions based on the results of an acceptance test.*** Regardless of how good the results are, comparing the observed single-point-in-time data to requirements or expectations simply can't indicate how the process will perform after, say, 3

days, much less 3 months. What is needed at this point is a testing process referred to as statistical qualification.

SPECIFICATIONS CHECKLIST	
<ul style="list-style-type: none"> Product Quality Characteristic: Average Thickness per Coil 	<p style="text-align: center;">Target = 0.1025 Upper Specification Limit (USL) = 0.1050 Lower Specification Limit (LSL) = 0.1000</p>
<ul style="list-style-type: none"> Product Defects: No more than 2 coating voids on any coil Product Performance Characteristic: No more than 2% units per coil (each coil produces 100 panels or units) rejected in Customer Stamping Plant #1 Equipment Performance Requirement: MTBF equal to or greater than 2 hours per 80 hours of scheduled run time. 	

Figure 1

RESULTS OF ACCEPTANCE TESTING ON PLATING LINE

Characteristic Category	Requirements Specifications	Observed Outcomes*	Decision/Conclusion
Product Quality	0.1025±0.0025	0.10211, 0.10492, 0.10371, 0.10416, 0.10156, 0.10370, 0.10400, 0.10272, 0.10244, 0.10325	PASS
Product Defects	No More Than 2 Voids per Coil	1, 0, 2, 1, 1, 2, 2, 2, 1, 2	PASS
Product Performance	2% Units per Coil Maximum	2, 0, 1, 0, 0, 2, 2, 1, 1, 2	PASS
Equipment Performance	MTBF ≥ 2 Hours	MTBF = 2.35 Hours	PASS

*Product observations derived from each coil tested; equipment performance data logged after 80 hours of scheduled run time.

Figure 2

RESULTS OF ACCEPTANCE TESTING ON PLATING LINE 3 MONTHS AFTER INITIATION OF FULL-SCALE PRODUCTION

Characteristic Category	Requirements Specifications	Observed Outcomes*	Decision/Conclusion
Product Quality	0.1025+/-0.0025	4.73 Defective	UNACCEPTABLE
Product Defects	No More Than 2 Voids per Coil	17% of Coils with >2 Voids	UNACCEPTABLE
Product Performance	2% Units per Coil Maximum	9.86% of All Coils >2% Defective Rate	UNACCEPTABLE
Equipment Performance	MTBF ≥ 2 Hours	MTBF = 1.75 Hours	UNACCEPTABLE

*Cumulative results after initial production period.

Statistical qualification is *an assessment of whether future production units or equipment performance conditions are likely to meet the internal or external customer requirements and expectations*, based on the observations gathered from an initial set of actual conditions. A statistical qualification allows one to predict what is going to happen, not just what has happened.

It should also be mentioned that it is possible to conduct a short-term qualification study called a potential study analysis. The term “potential” is used because the study is conducted in the short term, like an acceptance test, where stability, or statistical control has not been documented. In this case the results can’t be used to predict into the future, but the test provides more information than just comparing the results to the requirements. Since it can usually be done with the same data one gets from conducting the acceptance test, it doesn’t require much more time and effort, so it’s a good idea to use the technique. The time limitation still exists; however, so long term qualification tests are still required.

Most companies do not conduct this type of test for three reasons. The first is that it may take a bit more time and effort. If the acceptance tests are passed, sometimes company personnel think they have nothing to worry about, and the extra effort just isn’t required. That’s where the nasty surprises come from.

The second is that in a number of cases, which are getting to be smaller in frequency, the company is small and just doesn't have staff that is knowledgeable in how to conduct the qualification studies and finally, the third is political.

In a number of companies, the staff responsible for installing and functionally starting up the equipment or lines aren't the same people who are responsible for running production, assuring the quality of the output, or for the reliability or maintainability of the equipment. So if, for example, the engineering group is responsible for getting the equipment "properly" installed, operations is responsible for production, the maintenance group is responsible for TAU and downtime, and the quality group is responsible for quality and customer complaints, then the engineering group might take the position that once the line meets its functional and acceptance requirements, it is up to all those other groups to worry about "the rest."

Unfortunately, common sense is no match for a vertically integrated organization with departmental metrics. If qualification testing is done at all it is usually after engineering has "handed off" the equipment or line to operations, and then the surprises set in and by this time, the original start-up team is working on a new project.

Of course, the problem is that if the qualification procedure is started *after* production begins, scheduling and other customer commitments come into the picture and valuable engineering resources are off the project. Consequently, it takes a long time to understand and ultimately solve all the problems the line might have. *Much longer*, in fact, than if the tests were conducted *before* full-scale production was initiated.

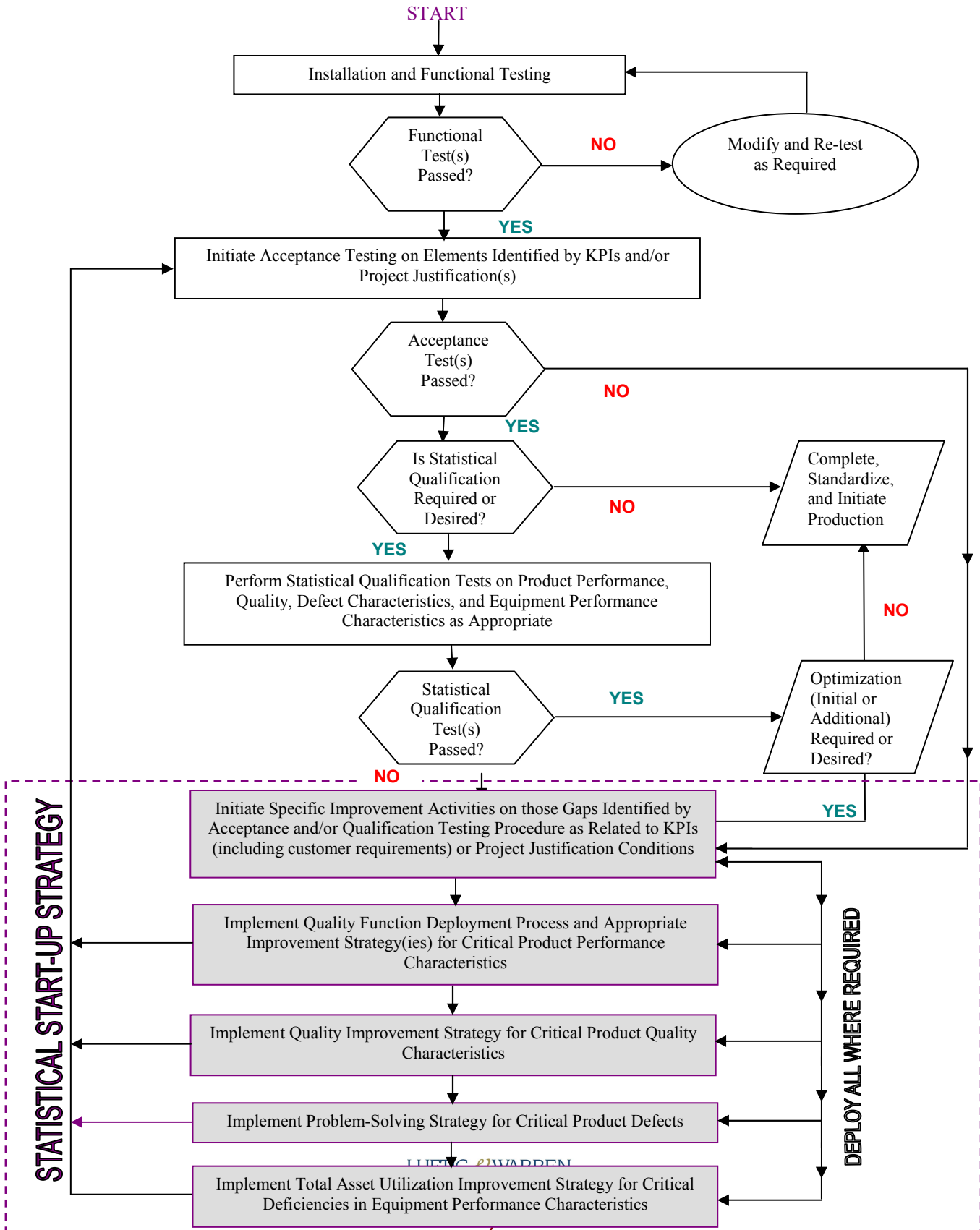
Assume that a new piece of equipment or line was being installed (Refer to Figure 3). Assume further that all functional and acceptance testing is conducted and everything passes. All of the test requirements, by the way, would have been specified before the equipment or line was ordered, minimally; and hopefully before the equipment was even designed.

Now, one of two conditions can occur. Suppose the line or process fails one or more of the acceptance tests. This, by definition, is also a failure on the qualification requirements, because they're more stringent. At this point, a Statistical Start-Up would be initiated on the elements where a gap exists, based on the acceptance tests. For the elements that passed the acceptance tests, in today's competitive environment in most industries, one would almost certainly want to conduct a qualification analysis. For any of the elements failing the qualification tests, these would be added to the first list, and subjected to the Statistical Start-Up procedure. Those that passed would be standardized and put into place using the Quality Operating System (QOS).

So, basically, at this point, one could now have a set of product performance characteristics, product quality characteristics, product defects, and equipment performance requirements that failed to pass at acceptance and/ or qualification, and it would be these conditions which would be the focus of the Statistical Start-Up.

Figure 3

OPTIONS IN THE PROCESS OF ESTABLISHING EQUIPMENT/LINE ACCEPTANCE, QUALIFICATION, AND OPTIMIZATION



The statistical startup methodology uses a large number of advanced strategies and tools in a highly disciplined, integrated fashion. There is a text that is used to guide personnel through the process. The major strategies and tools usually required in a standard statistical startup are shown in Figure 4 and Figure 5 shows how the data might appear in many industries and companies at the end of the qualification portion of the Statistical Startup process.

Figure 4

MAJOR STRATEGIES, PROCESSES, AND TOOLS COMMONLY UTILIZED IN STATISTICAL START-UP PROJECTS

Strategies	Processes	Tools
Quality Improvement Strategy	Supplier Quality Assurance Process	Statistical Process Control (SPC) Charts
Problem-Solving Strategy	Customer Quality Assurance Process	FMEAs, FMECAs
Total Asset Utilization Improvement Strategy	Quality Function Deployment	Design of Experiments (DOE) Technology
Measurement Control and Capability Improvement Strategy	Daily Management Process	Reliability Root Cause and Failure Tabulation Matrices
	Policy Deployment Process	Standardization Systems (SOPs)

Note that Figure 5 uses some statistical symbols and terms. For those readers who are proficient in the quality sciences, this should indicate that the process, for all relevant characteristics, is in a state of control and is minimally capable. Normally, at this point, if the qualification tests are passed and, based on the objectives, all requirements and expectations are met; the associated variables would be standardized, and implemented for full-scale production with continuous improvement activities employed through Kaizen. Or one could engage in the advanced methods component of a Statistical Start-Up and proceed to *Optimization*. However, if one or more of the product or equipment characteristics are shown to fall short of requirements or expectations in the qualification tests, full-scale production should not be initiated. Instead, the optimization component of the Statistical Start-Up process would be conducted.

Optimization, which is also referred to as *Statistical Optimization*, is the advanced component of the Statistical Start-Up process. It is defined as *the strategy by which product or equipment characteristics are brought into the highest level of control and capability available given the nature and condition of the existing process; at the minimum cost possible*. The term ‘existing’ refers to the notion that the optimization process gets the equipment or line running at its highest level of capability possible, given the equipment available as it has been designed and purchased. Only the process referred to as *Advanced Quality Planning*, properly executed, can guarantee, so to speak, that all of the requirements will end up performing in a state of control and capability.

Figure 5

RESULTS OF ACCEPTANCE TEST ON PLATING LINE 3 MONTHS AFTER INITIATION OF FULL-SCALE PRODUCTION AND FOLLOWING A STATISTICAL START-UP (Qualification Stage)

Characteristic Category	Requirements/ Specifications	Observed Outcomes*	Decision/Conclusion
Product Quality	0.1025+/-0.0025	99.994% In Specification with $\mu = 0.1025$	ACCEPTABLE
Product Defects	No More Than 2 Voids per Coil	99.865% of Coils with ≤ 2 Voids	ACCEPTABLE
Product Performance	2% Units per Coil Maximum	99.997% of All Coils <2% Defective Rate	ACCEPTABLE
Equipment Performance	MTBF ≥ 2 Hours	MTBF In Control at 2.25 Hours	ACCEPTABLE

*Projected cumulative results after initial production period.

The key is that the optimization process achieves a point where *the system is performing as well as it can, without making fundamental changes or additions to the equipment and spending additional capital*. Hence **Statistical Start-Up** may be defined as *a process by which advanced statistical methods are applied in a strategic fashion to critical product performance, quality, and defect requirements, as well as critical equipment performance requirements, in an attempt to achieve full statistical qualification or optimization prior to the initiation of full-scale production*.

The “hit rate” of success in utilizing this strategy is on the order of 95% or more for the product and equipment characteristics pursued, in those cases where the entire Statistical Start-Up discipline, including the Optimization component was utilized. Usually, the product and equipment design engineers have fundamentally done a good job. It’s the last piece of qualification and optimization that has historically been missing. Even where qualification or acceptance does not pass testing, the Statistical Start-Up strategy will at least indicate where the problem(s) is(are), and what one has to do to change the process to modify the outcome.

For example, Figure 6 illustrates results of what might be expected after qualification and after optimization from the same process with no additional costs invested. Clearly, this has implications for cost savings. For example, on one project, after using the process to achieve

qualification on a number of internal characteristics, on one *single product characteristic*, the effect realized in the reduction of internal rejection rates had an annual impact to the company translating into bottom line savings of approximately *1.6 million dollars*.

Given these observations one would think that more companies would utilize this process. However, as was mentioned earlier, there were at least three reasons why this is not so. Perhaps a fourth reason could also be what we might call an inappropriate decision related to priorities.

Usually, when talking to management teams about conducting a Statistical Start-Up, the politically, vertically-driven folks in the room who just see this as more work get uncomfortable. Inevitably, the question is asked, “How much more time and effort will *it* take?”

An appropriate response to this question is to address the word “it”. What’s “it”? If “it” is the goal of getting a product off the line that kind of looks like what we’re trying to make, or to be able to make one or two units that meet spec, and then dump the problem equipment or line on the operations folks, wash our hands and walk away, then yes, this is going to take extra time and effort.

But if “it” means starting up the equipment and line in a state of control and maximum level of capability possible, then, it will *always* take *less* time to achieve that goal during a Statistical Start-Up, than after full-scale production is initiated—and that’s a fact. Of course, some Start-Ups require additional calendar time, and the amount of that time is dependent on the size of the project, but it’s like the old saying, you can pay now, or pay (big) later.

This discussion took place once, with a V.P. of Operations, in front of the entire management team. After he challenged the supposedly inordinate amount of time, effort, and money required to perform a Statistical Start-Up, he was asked the question “What’s your oldest plant?” The response was “12 years”. He was then asked how many of the critical product and process conditions were operating in a state of control and capability. The response was “none”. In essence, the plant was still being started up. The message was well received by him and the rest of the management team. It is interesting to note that this particular facility utilized the Start-Up methodology on a plant which ended up coming in at \$80,000 a month under budget and a higher than projected level of production; all in a plant with less than 100 hourly and salaried workers.

Other benefits, which are included in the deliverables of a Statistical Start-Up, include SOPs, Reaction Plans, Audit Requirements, FMEAs, and training guides all, which are required to continue to operate the line or plant in a state of control. What this means is that if the company is seeking ISO certification, QS-9000 certification, or some other type of certification status, virtually no additional work needs to be conducted to achieve that goal. The output from this methodology plugs directly into this framework.

Of course, this same methodology can be applied to existing equipment, which may be performing very poorly and where management wants a breakthrough improvement in costs, product quality, or equipment performance. This discipline, in fact, can turn a losing line or plant into a profitable one. Further, it is not always necessary to stop the production lines. Often the requirement is partial production or a slight reduction in output for a short while.

Currently, this methodology has been successfully applied to more than a billion dollars of equipment, lines and plants for both domestic firms as well as firm's abroad.