

Lean Six Sigma Quality Transformation Toolkit (LSSQTT)*
LSSQTT Tool #9 Courseware Content
“Attribute Data, The Obvious Starting Point For Lean, Six Sigma, Service”

1. Statistical principles in quality—more basic definitions
2. Normal curve, variation and standard deviation as foundations, six sigma
3. Attribute data discussion beginning
4. Additional perspective: attributes and check sheets as pre-control
5. P chart steps for attribute data
6. Attributes, checklists and charting, and more for lean, service
7. Quality characteristics, accurate data and variable charting introduced
8. Team based problem solving, six sigma variation reduction for lean

*Updated fall, 2007 by John W. Sinn.

Statistical Principles In Quality—More Basic Definitions

The use of statistical methods enables us to make quantitative statements about data that often could not be communicated as readily by other means. Statistics refers to data, the collection of numbers that represent raw materials, products or systems we work with. These may be test scores, reaction times, frequency of rejects, numbers of product produced, or other similar numerical measures or indicators of items or behaviors. But statistics also means the processes or methods, generally as reflected in formulas, which have been derived and developed to treat data in meaningful ways which are helpful, if not critical in technological organizations and technical situations.

The task of statistics is to reduce groups of data to meaningful and useful information. These meaningful values, or information, can then assist us in several ways. This includes planning, policy changes and general decision making, setting standards, and others. Based on numerical indicators and measures from production, we can take some of the guess work out of making decisions--and become increasingly competitive in the process. When we do not use data, simply stated, we are going to be less certain, less definitive, and more oriented to making "seat of the pants" types of decisions. Statistics can help provide a more solid and formal base from which to make improved decisions.

Based on statistical data generated, policy can be configured or altered. If existing modes of organizing and operating prove to be problematic in quality terms, changes must be made. Changes would be based on statistical indicators, and after changes, new data generated could be used for comparative

analysis. Over time and with training, personnel at all levels will be exposed to data which help make the case for decisions, changes, and managerial issues which are being confronted. This is fundamentally what data based decision making is about--generating and using data to assist in ongoing improvement and change.

Yet the broader problem solving environment is underpinning the entire discussion. This has been reinforced throughout the toolkit in earlier tools and through the model shown periodically. Part of the complexities being configured relate to the broader culture, as well as internal and external customers and suppliers, technological complexities and realities, and others. Beyond the data are documentation and technical leadership within that competitive culture of change. This is presented and discussed throughout since this is the complex environment where we must do data based functions.

Industry-wide terms and other key statistical indicators. Descriptive statistics seek to accurately describe a situation through numeric values/data. These deal with the present time and are limited to describing present conditions for the group of numbers actually studied. Meaningful values describe the results of a particular sample of behavior. The purpose of descriptive statistics is to communicate something about a particular group of observations, as in the distribution of weights on a football team or grade point averages. When we say that 90% of the product was good or 10% of the product was defective, it describes certain areas of concern which we may want to know more about. For example, we may wish to study type and number of defects contributing to defectives in the 10% earlier.

Inferential statistics deals with inferences, implications, and generalizations from the sample to a larger population. Inferential statistics are generally

concerned with the future or with broadening conclusions to include other groups beyond those actually studied, such as probability for occurrences of events. The attention shifts from describing a limited group of observations to making inferences about the population. The task of inferential statistics is to draw inferences or make predictions concerning the population on the basis of data from a smaller sample. An example could be a poll of one hundred (100) voters (the sample) to predict how the nation (population) voted. The sample of voters polled must be representative of all (or most) voters and citizens if we wish to make useful inferences from the sample to the broader population. Regarding product on the line if a sample does not gain truly representative data, we increase the likelihood that we can not accurately infer from the sample to the population.

Central tendency is three measures, including mean or average; median or positional value; and, mode or most frequent. Uses of central tendency occur in several ways. The average or mean is the one used most often. It is used to report average size, average yield, average percent defective, and so on. The median is usually used only in some special situations such as data that can be ranked but not easily measured including color, softness, or smoothness of a surface. The mode is the most frequent or most representative value. It is sometimes useful in eliminating the effects of extreme values in a distribution.

Range tells us the extremes in a given group of data. In the data group 10, 7, 3, 12, 21, 14, 15 and 12, the range of data is 3 to 21. The top of the range is 21 and the bottom of the range is 3. We could rank order the data within the range to look as follows:

21
15
14
12
12
10
7
3

When it is ranked from high to low, we can better understand and more readily begin to use the data. For example, we can now determine that the mode is the number 12, appearing twice, and the average is $84/8 = 10.5$. This was based on the sum of eight numbers being 84 divided by the total number of values in our data set.

Note that what started out as just a simple bunch of numbers has become a rather useful set of

descriptive data. If we were to then indicate that this was the temperature for the first eight days in the month of January for the year 1994, it would give even more meaning to the data and to those days during the month of January in 1994. This is essentially what descriptive data is about. Inferential data would reflect behavior over time in another way--still using our existing example.

For example, looking at the ten year average of the same set of eight days in January we find that each of the averages is as follows:

1985 = 11
1986 = 12
1987 = 12.5
1988 = 13
1989 = 11.5
1990 = 14
1991 = 15
1992 = 18
1993 = 20
1994 = 10.5

We could then sum these and provide what is called a grand mean. This would be 137.5, and when divided by 10, it would be 13.75. We would generally say that the mean, or \bar{X} , of all of the 10 means is a grand mean, or $\bar{\bar{X}}$. Thus 13.75 is the grand mean of the ten year average. The inference which may be acceptable to draw from this data is that the temperature for the first eight days of January in most years will be in the range of 10.5 to 20 degrees, and on average will be 13.75 degrees.

Translating this into product, as an example, it is increasingly important for us to be able to say that on average we can produce product with certain quality levels. It is also important for us to be able to retrieve data from our part histories and say with reasonable assurance to current, and perhaps more important, potential customers, that we can produce products of certain types at this or that level of tolerance. Tolerance in this case means a range of values--possibly length or width on a given product line. Note the use of the term, range, used in reference to tolerance.

If the customer says they will remain a customer only as long as dimensional accuracy falls within the tolerance with a mean or average value which is capable at the 1.50 or 1.60 level, relative to the predetermined tolerance, how will we know if we can achieve this demand? We must establish--and begin to understand--that this is in fact the case--we must produce within tolerance or specification as it is sometimes called at a level currently 1.30 plus

relative to the tolerance. Our ability to obtain and keep contracts in the future will be directly contingent upon our knowledge and understanding of range of tolerance, mean values within the tolerance, or capability, and other related statistical terms.

Following the concept of tolerance, variable and attribute data must be introduced. Variable data is that which we can measure with a graduated instrument such as a dial indicator, ruler, micrometer or other incrementally based tool, studied as variation from tolerance. Attribute data are judgement calls based on subjective observations of defect and defective. If we say color looks wrong, product is shaped funny, or it smells different, we are making attribute judgements. These are go and no-go, not based on graduated measures, but rather simple judgements of fit or acceptance.

It should be briefly pointed out that much of the basis for attribute data is the applications to service industry. Much of the variable data logic, upon which the statistical process control systems, such as six sigma, were developed, has been manufacturing. Service industries such as education, government, health care, retail and other consumer businesses, distribution, information systems, and so on are the new and emerging markets for attribute type systems being explored in the current tools. Checklists, cost related information and forms, standard procedure forms, time related collection devices for documentation, and other data and documentation systems can provide significant improvement opportunities in service activities.

Normal Curve, Variation And Standard Deviation As Foundations, Six Sigma

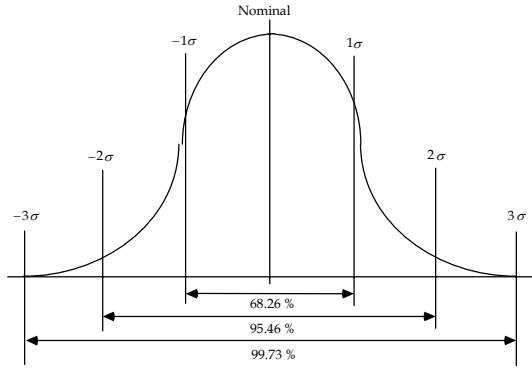
The context for six sigma is represented in the normal curve. The normal curve represents, statistically, all inputs from production within the statistical process control system. Six sigma is a project approach for improvement which parallels capability, gage repeatability and reproducibility, attribute and variable measures, and other elements of the toolkit approach to improvement, or kaizen. Six sigma is based on the assumption that by using data, documented over time in various ways shown throughout the toolkit, we can systematically improve in disciplined ways, measurable by the data and documentation. This is further explained based on the functions and applications of the normal curve.

Dispersion is represented by three measures including range, variance, and standard deviation. Use of measures of dispersion occur in several ways.

Range is the minimum value subtracted from the maximum value in a sample or set of numbers, and as a measure of variability of a set of values is useful when the number of values is small. The standard deviation as an indicator of total variation, is also reflective of what is called central tendency in and among data. Variance, in general, is simply the square of standard deviation, providing a broader look at the behavior around the mean of data. All of these values, as a function of descriptive data, fall under the broad rubric of the normal curve. The "normal curve" is illustrated nearby for further description and discussion throughout this section.

While we do not necessarily use the range, standard deviation or variance directly, we do use them indirectly in numerous ways. For example, just about all of the statistical process control calculations we will use rely upon the standard deviation in some capacity. This is true for attribute and variable SPC, capability, and gage repeatability and reproducibility (R & R) calculations. Our concern with understanding range, standard deviation, variance and other measures of central tendency is not only to use them directly, but rather to understand their functioning and role in the broader applications tools for ongoing improvement.

Standard deviation is a calculation of dispersion among the mean, or sample total variation around the mean. With the use of standard deviation the total population can be predicted as being within + or -1 standard deviation 68% of the sample will fall, within + or - 2 standard deviation 95.5% of the population will fall, and within + or -3 standard deviation 99.7% of the population will fall. The standard deviation is, by far, the most valuable and used measure of a frequency distribution. It expresses dispersion in a single number and a very important relationship existing between the "standard deviation" and normal curve. This is illustrated in the normal curve graphic nearby.



The Normal Curve, Distribution Of Variation In Populations.

If we translate this into useful information for improvement, we begin to understand that variation in product quality is the basic issue. When we take variable measures out on the line, and assuming we wish to place this information into a normal curve type representation of variation, this could be configured graphically. When the data was summed and a mean calculated, the value for the mean can form the basis for further analysis. Presented in a graphical form, it is commonly shown as a normal curve, such as we saw earlier in the short form. The normal curve represents all of our group of sampled data--all else is directly related back to the normal curve, and forms the basis of much of our broader SPC relationship for analysis and improvement.

To calculate the standard deviation, we can use a simple formula called the sum of squares. This provides the grouped value around the mean or average performance numerically. We can understand and use the basic relationships inherent in all of this since, when calculated, the standard deviation is the average value of all deviations, represented as a mean. This then means that the average value from the mean in the normal curve is the standard deviation or variation from the mean. Powerfully, all other basic SPC functions rely upon this foundational statistical concept--all within the larger group of data--and the normal curve. The standard deviation, for our purposes, is expressed as:

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}}$$

Where σ is sigma, also called the standard deviation, and X_i represents each individual X value or raw data.

Referring back to our previous examples and values, it should be clear that this is actually a rather simple statistical application--yet rather powerful and useful. By (1) summing all values (the X_i summed and averaged); (2) determining differences in each X_i value from X_i average; (3) squaring and summing these; and (4) dividing by one less than our original sample size (n), we are obtaining the grand deviation average from the total group of data. After determining this, we can then make certain assumptions based on the standard deviation from the grand mean value of the group--generally termed as sigma--and used in numerous ways by our customers and suppliers. Sigma, actually standard deviation, will be relied upon both directly and indirectly throughout discussions of basic and advanced SPC.

We would commonly say that + 3 sigma is three standard deviations on the high side of the mean or average. Again referring back to our graphics, three standard deviations on the low side of the mean is -3 sigma. Approximately 99.7% of all values will fall within + 3 sigma and -3 sigma, of the mean, if under statistical control (plus or minus three sigma). A statistical limit that is set three standard deviations on the high side of the mean, or +3 sigma, is termed upper control limit. The lower control limit is a statistical limit that is set three standard deviations on the low side of the mean, or -3 sigma. Similarly, approximately 95.5% of all values will commonly fall within the range of +2 sigma on the high side of the mean to -2 sigma on the low side. And the remaining +1 to -1 sigma on either side of the grand mean represent approximately 68% of all data from the larger sampled group in any given situation.

By understanding sigma, based on standard deviation, we can perform various other useful, and necessary, analytical functions for improvement in process and product. The terms six sigma should now begin to take on new and more relevant meanings. As we begin to see that when the customer says we must have a six sigma value of 2.00 capability, what is meant is that all values will be, on average, well within the upper and lower control specification limits. While we may begin to understand the concept of six sigma and capability, we must also understand that it will take strong commitment and teamwork to achieve 2.00.

Attribute Data Discussion Beginning

Attribute data, are those data which can be judged as either having or not having some characteristic which was previously determined to be

desirable. Attributes are characteristics which are of a human judgement nature, such as taste, appearance, or general size and shape. Attributes may or may not involve gaging and simple go no-go devices for assisting in making judgements. But it is critical to understand that an attribute is either acceptable or it is not acceptable from a quality standpoint.

Variable data is measurable with graduated scales, digital devices, or other numerical instruments and systems. This is an important difference, since while most products are measurable in variable instrumentation ways, it is not always practical and certainly not always cost effective to do so. Variable data is not necessarily always needed--and frequently it may not be easily done--or cost effective.

An attribute example could include appearance of some types of clothing. In many cases judging the appearance of the clothing is fine as an attribute, but it is also clear that specific sizes of clothing, as variable data are quite important. Similarly, when we eat food we generally know whether it is good or bad by taste, but it may also be very important under some circumstance to determine by careful measurement, the food's temperature in preparation, or to use related variable measures to determine other characteristics in quality. If bacteria were present in undesirable quantities, we would want to quantify this, beyond a simple attribute subjective "taste test".

Picking up on the clothing circumstance alluded to further, attributes can be further defined and explained. Attributes such as overall appearance, color, material texture, general fit, and others are characteristics which can be observed and tracked. Other characteristics such as length, diameter, cloth density, color density, and others are specific and measurable with some type of variable gage. Previous attribute characteristics are judgement calls that may or may not use a simple go or no-go gage. In each case, attribute or variable, further definition could and likely would, be given in characteristics. Variable length in sleeves, pants, and so on could be specified and general fit attributes could be defined as bulky, tight, and so on.

While variable data are generally the more desirable due to specificity and precision for mechanical fit and function, many quality situations and characteristics do not require this type of precision and detail. In many cases if we determine how to track some characteristics with variable data when a simple attribute judgement is sufficient, we simply will not be competitive. As production is increasingly less oriented to craftsmen, and more mass production and automated, we will increasingly see more variable data being generated, of necessity, particularly in continuous production. The start will

often be attribute and judgmental in nature, later to shift to more sophisticated variable measures.

The shift will be done in part to support requests from customers who will demand variable data-based, precision measurement oriented, statistical process control. But can the equivalent system be set up for situations where attributes are being used to track quality? Assuming yes, attributes can be charted, what would the calculations be, and how would charts be pursued and maintained?

Much of the same logic applies to attributes, and attribute charting, as is true in variable charting. Similar to variable \bar{X} and R charts, we must determine the purposes of our charts, and fully understand the application and environment we are pursuing. This relates to understanding the process and characteristics which are to be tracked over time, and why they are being studied or charted. This will often initially be driven by the customer, but we must also remember that any charting costs us all, and thus it becomes important for us to consistently be in communication with the customer and others about the possibility of re-evaluating characteristics over time. At some point we should be able to reduce the tracking of a given characteristic, as it becomes increasingly under control, allowing us to move to other characteristics and opportunities for improvement--both attribute and variable.

Defects, defectives and other non-conformances. Following the introduction of attribute data and characteristics in quality, it is important to distinguish between defects and defectives. While related, the two are also generally different as related to data, attributes and quality. This is because the defect may be one of many which contributes to what may or may not be a defective product. By tracking occurrences in production, we can begin to sort out types and significance of defects versus defectives.

Even with multiple defects, and different types of defects, the product could remain sufficient for use. This is also indicative of the need to be precise in our quality judgements, leading to the desire for variable data-based measures. Defects, when detected by some type of mechanical and/or human judgement, generally will lead to the need for determining what constitutes a defective product, and what to do about it.

A defect is defined as any characteristic which is not in conformance with standards or requirements as identified by the customer. A defective is a part which possesses more than the allowable number of non-conformities, usually the presence of one or more defects, again as identified by the customer. The defective would also generally constitute product which cannot be shipped, while the individual defect

may or may not qualify for non shipment. It is important to differentiate between individual defects and multiple defects in the form of defective product.

The broader application of defects and defectives being clarified also relates to pre-control and short runs. This can be done through a simple checksheet or checklist which provides the basis for identification of the nonconformity at the operator level in a disciplined manner. The checksheet can be used to begin the charting process, and also serves as the start of decision making at the proper levels.

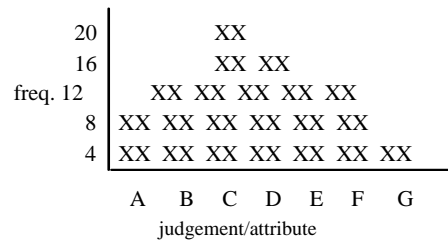
Histograms, bar, pie and Pareto charts.

Another, possibly less complex charting technique in the attribute judgement method relates to simple histograms, bar, pie and Pareto charts. These simple charts are bars, lines or other symbols which represent a corresponding number of values or observations tracked in production--attributes identified and logged--converted as summaries to identify areas needing improvement. Bars can be horizontal or vertical and are generally proportionally constructed relative to information they represent.

Histograms are constructed by first determining the number of classes (cells) to use. The number of cells in a histogram of data should be determined in part based on the number of values in the sample. Cells or blocks of information may also be influenced by other "common sense" information based on what we know about the situation. For example, it may be obvious that the data generally falls into groupings based on the nature of the process, characteristics or other factors. Steps for constructing histograms are:

1. Place all data/observations on a chart (form) to begin organizing, determining the class (cell) intervals, based on total group actual data.
2. Construct a tally chart with classes in rank order. Classes should provide a category for each type of observation or attribute.
3. Tally each data observation into appropriate classes and list total frequency for each cell.
4. Prepare vertical and horizontal axes. Vertical axis is frequencies of classes and horizontal will be classes of the frequency tally.
5. Fill bars to height of total frequency of class.
6. Based on graphic, analyze and interpret data to determine appropriate steps for improvement.

An example histogram, based on previous steps, is shown using frequencies ranging from 4 through 20, and various judgements of attributes ranging A to G.



One Type Histogram, Bar Chart.

The chart shows frequencies of attribute A = 8, B = 12, C = 20, D = 16, E = 12, F = 12, and G = 4 occurrences. This provides the start of a bell type normal distribution, based on natural occurrence of data. Some inaccuracies may exist with the approach taken based on groupings and cells resulting from "forcing" all occurrences into predetermined groups. This may be a valid criticism, yet it provides an efficient way to describe grouped data. This could be "modified" in later analyses based on additional measurements, and cells reduced or expanded in size.

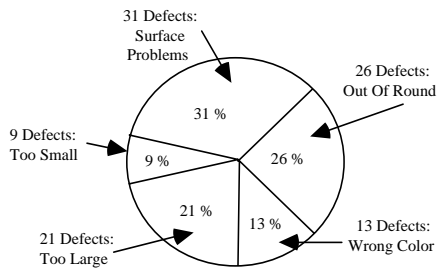
A variation on the histogram is the bar graph. Bar graphs can use horizontal groupings as shown below where TA, TB, and TC are frequencies of judgements given in more detail and sub-grouping in cells. Judgements shown reflect simple distributions of observation ranging from A to H, and could be color, surface quality, or other common attributes.

Attribute/ Judgement	TA	TB	TC	Tally A	Tally B	Tally C
A	1	2	1	*	**	*
B	2	3	2	**	***	**
C	3	4	3	***	****	***
D	5	5	4	*****	*****	****
E	6	6	5	*****	*****	*****
F	4	7	4	****	*****	****
G	3	4	3	***	****	***
H	2	3	2	**	***	**

Attributes Organized Into Bar Chart Form.

General considerations for these type bar graphs include their usefulness for viewing groups of information and for comparing and analyzing overlapping and similar groups of data. These groupings approximate normal curve shape, assumed since behaviors studied as attributes occur randomly and represent the broader population.

Related to other attribute analysis tools, another useful graphical chart is the pie chart shown nearby. This has 100 product defects by categories, sorted and organized in the chart--by count and %.



Pie Chart Showing 100 Product Defects, By Count And %.

This analytical tool is a simple "slice out of the pie" for each area represented. The power in this approach is "seen" clearly by comparing in rather straight forward ways, the larger versus smaller areas--and where we need to go to work to make improvements. This is particularly true as related to attribute data at the workplace for quick and easy analytical aids for operators. Defects and defectives, as in conformance and non-conformance, acceptance and rejection based on subjective judgements, are frequently shown in these ways with pie charts.

An additional bar type graph, related to histograms and gaining in popularity, is the Pareto chart. The reason for Pareto popularity, similar to many of these tools, is its quick graphical analytical focus for quality improvement. The Pareto orders data, based on occurrence, ranks the occurrences by percentage, and gives a graphical pictorial for quick reference and ease of use. An example Pareto application is shown on the next page and additional blank forms are provided in application forms. Procedures for constructing Pareto charts are:

1. Identify the problem or attribute to be studied.
2. Collect data of frequency on occurrence by attribute, studied categorically or individually.
3. Identify maximum level of occurrence in attributes, placing most from left to right, least.
4. Overall behavior of problem is 100%.
5. Construct graphic of data per occurrence.
6. Left vertical matrix represents frequency.
7. Right vertical matrix is up to 100% downward.
8. Horizontal matrix is category of occurrence.
9. Vertical bar is percentage of occurrence.
10. Connect vertical bars at various points to demonstrate behaviors in groups, relationships.
11. Interpret, provide improvements and repeat.

General communication and prioritization capability of the Pareto is underscored. An example nearby

shows a possible occurrence of external customer complaints. Occurrences are given numerical value with "x's" in the vertical bar columns. It may be helpful to place the greatest to least column (item) of occurrence from left to right on the chart. It should also be noted that the % value on the right is not necessarily always 100%--in fact it is unlikely that any one set of occurrences would ever equal 100%.

Additional Perspective: Attributes And Checksheets As Pre-control

The use of these tools may seem so simple that we may not even need to bother with their use. They are simple--but certainly far from insignificant. As with much about world class and ongoing quality, it is important to remember that improvement is generally incremental, and it must start somewhere.

The starting point for much of quality will be tracking non-conformance or defectives in some readily observable manner, perhaps bar charts or Paretos. This also relates to discussions concerned with converting from attribute tracking systems to the possibly more elaborate and rigorous variable charting system. Since the bar, histogram and Pareto charts are often associated with attribute and non-conformities, it may make sense to consider beginning with Pareto in many applications. But with time and experience, convert to a histogram or bar chart reflecting variable data.

World class and total quality is about discipline and systematic, repetitive behaviors in process, certainly supported by people. Particularly as we are trying to get this type thinking and behavior started, as well as continued and evolved, we will likely wish to start with the simpler Paretos, or Pie charts, and convert from there. We can expect to see increases in the disciplined approach if we require and support collection of some type of data in the systematic manner being described.

Similarly, we will be much better prepared and equipped to make data based decisions with the Pareto or other chart based information, relative to not having this--even at the minimum level. But it should not end here--this is only the beginning. When we consider strong relationships relative to evaluating characteristics, changes in specifications, and other engineering and manufacturing opportunities for cooperation and collaboration, we begin to understand the need to begin tracking. Perhaps most important, the basis for proving changes implemented lies in tracking system. We need attribute data, in the form of various charts, such

as p, Pareto, bars, and others, to know what is working and not working--after making changes for improvements. How will we know what the problem was/is, if we have not tracked attributes representing defects and defectives? Attribute data forms the fundamental basis for understanding problems which will lead to solutions and improvements.

Attribute charting systems. The need to track and chart various attributes, under different types of circumstances, has led to the development of several types of attribute control charting systems. These are further defined as p, np, c, and u charts, generalized as appropriate in this section. Part of the discussion focuses on general concerns which must be attended to regarding attribute charting systems. Some of this is applicable to variable charting--and some similarities may be noted. The primary emphasis, for our purposes, will be p charting since it provides a good general basis from which to initiate and pursue all charting, as a broader SPC system. P charting is the most used attribute chart among those being discussed, a basis for others. Based on the p chart, c, np and u charting will be discussed. Similarities are drawn, and differences noted, where appropriate to aid in the broader understanding of attributes tracking and analysis, leading to variable charting systems.

As with any charting effort, it should be clearly determined what the actual purpose of the chart is at the outset, prior to proceeding. This relates to determining where and how actual inspection will occur, and the specific characteristics to be addressed--what are the details? What size is the sample and when should it be taken--and is it to remain constant over time or can it change without affecting the system? What other information should be placed on the chart and the attribute study in general--and how should this be used over time?

General steps for attribute p charting include collecting data, calculating fraction defective, and upper and lower control limits. Based on calculations, plot results for analysis, over time, and complete appropriate analysis and corrective actions if needed--depending on findings. Similar to variable charting steps, makeup of the team will depend on complexities encountered, severity of non-conformities or defectives, general environmental issues, and so on.

P Chart Steps For Attribute Data

The p charting approach will be briefly defined and explained, calculations provided and examples shown. The p chart is used where it is necessary or advantageous to know if the numbers of rejected sub-components or products is acceptable or

unacceptable in a given situation. While the focus is primarily on the p, most of the steps and procedures are also generalizable to all attribute charting systems. Specific steps are:

Attribute Step 1: Collect Data. Data are collected as the number inspected (n), and divided into subgroups. This is generally per customer request, or based on some inspection SOP, consistent with the type product or process under study. The inspection strategy is identified by date or lots for future documentation and tracking needs. Subgroup size is consistent with type attribute being studied, but no less than 50 to facilitate generalizations to the larger population over time. Consistent subgroup sizes of 60 are shown in the example on the forms.

Attribute Step 2: Collect/Calculate Each Fraction Defective. The second major attribute step is to calculate the fraction defective or nonconforming for each subgroup. This information should be entered on the chart or some type of collection form for tracking and further analysis purposes. The fraction defective is determined by dividing the number of defectives (np) by the subgroup size (n), expressed as np/n mathematically.

Attribute Step 3: Calculate Average Fraction Defective. The average fraction defective is a function of summing pn and dividing by the total n for an average or mean value, commonly expressed as p bar. This converts individual subgroups to grouped data, similar to all SPC functions relying on standard deviation, reflecting normal curve behavior.

Attribute Step 4: Computing The Control Limits And Central Line Values. Upper and lower control limits are calculated based on the p bar value converted to a percentage, or times 100. The p bar converted by percentage is used as the central line in the chart, centrally located between upper and lower controls (UCL and LCL). UCL and LCL are calculated based on standard deviation, plus or minus 3 sigma, generally. It is also possible to use other levels of sigma (i.e., 1, 2 or more), but generally 3 sigma is used. Commonly used formulas are:

$$UCL / LCL = \bar{p} \pm 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

As As illustrated through examples and applications, the UCL is +3 sigma, the LCL is -3 sigma, p is a function of np divided by n, and the p bar value represents sums of the total sample population.

Attribute Step 5: Plotting The Points. After all calculations are performed, for each subgroup, the

information is plotted on a form similar to what is shown nearby. Note that "ups and downs" may be observed in the step 5 charting graph are a function of varied subgroup sizes, sometimes unavoidable. This also points to the need of using consistent subgroup (n) sample sizes. Note an application as a variable n, with control limit lines fluctuating.

Attribute Step 6: Analysis Of The Chart. Over time, as data is tracked and plotted, the chart will provide a quick and generally easy-to-use tool for determining if the attribute being studied is sufficiently conforming to the desires of all concerned. Actions appropriate to the circumstances should be taken to adjust and improve over time.

c, np and u charts for attributes. Based on the previous discussion related to p charting, three other attribute charting methods are presented. These are np, c and u, each briefly discussed and an example provided. Following a brief discussion of each, referenced to the p chart, steps for calculating will also be provided.

The np chart. The np chart, while similar to the p chart, is slightly different. While the p chart studies the fraction defective (p), np looks at the number of defectives. The np chart may prove more useful in some situations, or it may prove easier to do. Again, similar to the p chart, np requires a constant sample size from each tracking period to the next. Steps to follow in conducting the np application include:

1. Identify on the forms that this is a np chart. This is done for you in the forms included in the toolkit, but some organizations use the same form for all attributes.
2. Record the constant sample sizes on the chart, similar to how p data were recorded.
3. Record the number of defectives found in each sample. This is np on the chart.
4. Plot the number of defectives found on the chart.
5. Calculate the average for np, or np bar, based on the formula:

$$n\bar{p} = \frac{\sum np}{\sum N}$$

6. Using the np bar value, calculate upper and lower control limits based on:

$$UCL/ LCL = n\bar{p} \pm 3\sqrt{n\bar{p}\left(1 - \frac{n\bar{p}}{n}\right)}$$

Interpretation of the np chart is essentially the same as the p chart. The obvious difference is that the np is concerned with number of defectives as compared to the fraction defective, as is the case in p. Process capability is the average performance of np bar over time.

The c chart. The c chart is an attribute chart which focuses on the control in number of defects. While the p chart studies the fraction defective, and np looks at the number of defectives, the c chart is concerned with defects as individuals. We would be well advised to remember that a defect, while significant, does not necessarily constitute a defective product. Multiple defects may help constitute a defective, but the defect in and of itself is not necessarily a defective. The c chart centerline is identified as the average or c bar, found by summing total defects and dividing by number inspected. Steps for conducting the c chart application are:

1. Identify on the forms that this is a c chart.
2. Record the sample sizes on the chart, similar to how p and np data were recorded.
3. Record the number of defects found in each sample. This is c on the chart.
4. Plot the number of defects found on the chart.
5. Calculate average for c, or c bar, based on:

$$\bar{c} = \frac{\sum c}{n}$$

6. Using the c bar value, calculate the upper and lower control limits based on the formula:
- 7.

$$UCL/ LCL = \bar{c} \pm 3\sqrt{\bar{c}}$$

Interpretation of the c chart is essentially the same as the p and np charts. The obvious difference is that the np is concerned with number of defectives as compared to the individual defects, as is the case in c.

The u chart. The u chart is used when it is important to view data from the perspective of defects associated with a specific unit of production. While the p chart studies the fraction defective, the

np looks at number of defectives, a c chart is concerned with defects as individuals, u is concerned with one item of production and defects in the same. Once again, we would be well advised to remember that a defect, while significant, does not necessarily constitute a defective product. The u chart centerline is identified as the average or \bar{u} , found by summing the total number of defects and dividing by the number inspected. Steps to follow in conducting the u chart application include:

1. Identify on the forms that this is a u chart.
2. Collect the data from one or more units, being careful to identify defects per each unit.
3. Record the number of units inspected, generally identified as n, and the defects found in each sample. Note that, as in the past c charts, defects are c on the u chart.
4. Calculate the u for each subgroup using:

$$u = \frac{c}{n}$$

5. Plot each u value calculated, becoming the points around the centerline.
6. Calculate the average for u, or \bar{u} , based on:

$$\bar{u} = \frac{\sum c}{\sum n}$$

8. Using \bar{u} value, calculate UCL and LCL:
- 9.

$$UCL/ LCL = \bar{u} \pm 3 \frac{\sqrt{\bar{u}}}{\sqrt{n}}$$

Interpretation of the u chart is essentially the same as the p, np, and c charts. The obvious difference is that the u is concerned with number of defects found in a specific unit as compared to the total defects found.

Decisions and issues summarized. Several basic issues are summarized here specific to attribute charting, but they are also related to most other types of charting. These will be further generalized and applied in the next tool relating to variable charting. Decisions/issues in preparation to completing the control chart could include:

- what is the purpose of the chart?
- where should inspection occur?
- which characteristics/attributes should be charted?
- what should the size of the sample be?
- should the sample size be consistent/same over time?
- how frequently should samples be taken?
- what information should be included on the chart?
- how does the p chart relate to the Pareto and perhaps other tracking systems?

Issues regarding startup of control charts include:

- making good human judgements about defects and defectives
- recording data accurately--how, who, etc.?
- calculation of p, completing the forms
- plotting the points

Continuing control charts--issues overtime may be:

- calculation of the control limits
- plotting the control limits
- plotting the sampled data points
- interpretation of control/lack of control
- comparison of current charts to past behaviors/charts
- having discipline to follow through on problems with corrective action

Possible reports/actions based on p charting could be:

- bringing/keeping process in control at satisfactory level
- review of specifications and characteristics periodically
- providing information to various groups/persons
- ongoing improvements based on charts

Similar to any form of pre-control, it is sometimes desirable to simply arbitrarily determine the upper and lower controls, rather than calculate them. It is also true that over time, given constant improvement, we will come to know that incremental "tightening" of the control limits will be sufficient for gaining constant improvement. At a quick glance this may seem to be inconsistent with the use of calculated controls, but we must recognize that as our systems mature, we can know that we are under control (or out of control) and we will know that improvement is

incrementally being made. This is consistent with the "start-up" nature of attribute charting and data, often aimed broadly at longer term variable systems conversions.

Attributes, Checklists, Charting And More For Lean, Service

The broader application of defects and defectives being clarified also relates to pre-control and short runs. Short runs can be done through a simple checksheet or checklist which provides the basis for identification of the nonconformity at the operator level in a disciplined manner. Whether attribute or variable, the checksheet can be used to begin the charting process, and also serves as the start of decision making at the proper levels. Not only is the immediate judgement on defect and defective being provided, but the broader corrective action for improvement should (and can) be noted. This is graphically provided nearby, and example checklist, for starters, are provided throughout this section.

The need to track and chart various characteristics, under different types of circumstances, has led to the development of several types of control charting systems. These are being defined as variable and attribute, at a start-up level, and generalized as appropriate in this section, to be further defined in the next tool. Part of the discussion focuses on general concerns which must be attended to regarding all charting systems--and again some of this is applicable to other data and charting--and some similarities may be noted.

Several checklist examples are provided, primarily as related to documentation systems and Kaizen team building. Within the previous discussion relating to tracking and studying characteristics, the checklist can be a pivotal tool. The checklist can be pivotal for doing all things we have been trying to reinforce. The checklist:

1. Provides indicators of defects or defectives at the point of production.
2. Is operator friendly, a good transitional mechanism on data collection.
3. Frequently ties into team building and process improvement as a basic work instruction or standard operating procedure (SOP).
4. Can be a pivotal tool in Kaizen analysis, reinforcing the data and documentation relationship at a grass roots level.

Most of the above will be borne out as several example checklists are provided. These include

housekeeping and safety checklists, process flow and charting analysis checklists, and various applications possibilities for production scheduling and balancing, through various modified checklists. The first, and simplest of these are various operation checklists, and a simple work sampling technique, shown nearby in the applications section.

Through the use of relatively simple checklist tracking systems, we can apply traditional time and motion analysis tools as well as non-traditional Kaizen tools. One technique which can help is the determination of process capacity based on operations within the process. After individual capacities are determined, we can balance total capacity through improvements at individual operations and collectively through improved process. This is where checklists and other related documentation techniques are powerful allies to data collection systems for improvement.

This form, essentially a reworked version of the earlier time and motion or method analysis form, is now targeted to further analyze the specific operation as part of the previously determined and analyzed process. Again, there will be delays, transports, and other wastes in the operation, needing to be identified and improved upon as waste. But what is being pursued here is the need to identify and provide an optimum operating time for production within the broader process. And it must be remembered that all waste issues in Kaizen are essentially attribute or variable data, and certainly may eventually be identified as characteristics.

It should be recognized that the operation capacity study is based on the assumption that we can detail out our operation processes and functions, all which will be studied further for improvement later. It should also be apparent that the capacity determination tool, while quite useful as an independent technique for improvement in any isolated circumstance, is a necessity for determining work loads for balancing and improvement through synchronous techniques.

Other documentation tools useful to technologists for addressing productivity are plant layout, process flow charting and inventory related approaches. Process flow charting uses symbols placed in a condensed format. The advantage of process flow charting is that the entire enterprise system (or a sub-component) can be analyzed from a graphical schematic. Times and costs can be placed alongside each process on the chart, permitting further comparison and analysis.

We should observe, through actual and simulated production problems, how the layout and

flow diagrams can prove useful in personnel and machinery allocation, both on a day-to-day basis as well as for plant/capital expansion in the future. This provides linkages for checklists and beginnings of tracking and analyzing our work areas for improvement. This all relates to total time and costs to produce the product, much of what is driving profitability and competitiveness in the marketplace.

Quality Characteristics, Accurate Data And Variable Charting Introduced

Quality characteristics were introduced earlier as a key part of the overall quality system. The quality characteristics are important dimensions, features or parts of a component. This involves the classification of numerous important areas of a product, such as dimensions, speed, hardness, weight, finish, each according to their relative importance in contributing to the quality of the product. This enables the quality effort to be directed to the matters of greatest importance, through priorities. It also simplifies selection of sampling plans, increasing inspection accuracy on important characteristics, providing a clearer focus. The quality characteristic is at the core of the entire quality system.

Quality characteristics help assure minimum quality cost by decreasing chances of error and aiding in general quality of communication. Establishment of characteristics as priorities to be inspected may occur simply based on the key problems which have been identified. But it is also likely that quality characteristics' determination would occur based on input from the customer, engineering functions, and perhaps others. This also relates to collecting information, or data, about the characteristics, in production. And the need for reliable, trustworthy, data that we can all rely on--gathered by operators and workers at the job site--is obvious.

Attributes are quality characteristics which are either possessed or not possessed by an item being inspected. Piece attributes may be evaluated after being produced and could include (but not be limited to) under, to, or over size, or the number of clearly identified flaws in the piece. Quality attributes are generally identified as observations and judgement calls as opposed to measurable characteristics, and are either viewed as acceptable or unacceptable in conformance quality terms. The process charts generally associated with attribute data are P charts, Np charts, C charts, and U charts.

A quality variable is any characteristic of a product or piece to be measured on a continuous

basis or scale. Shaft diameters, strength in materials, density in photography, are all variables related to quality. Variables, then, are used as measures of quality, and are often the basis for statistical quality control and analysis. Variable data are commonly associated with X bar and R charts.

Whether attribute or variable, we would be well advised to revisit the need to gather and handle data carefully and accurately. During collection and manipulation, such as organizing into columns or in forms, it is vitally important that we be mindful of the need to use extreme care and precision in our work. As we will see, increasingly, the data will formulate the basis of virtually all decisions in the workplace and throughout the organization. If we do not exercise extreme care and caution when measuring and recording, and perhaps as we calculate, we may be helping put ourselves, and others, out of work.

The sample used as the basis upon which to draw all conclusions must be carefully accomplished for deriving accurate data. How frequently to measure, when and where, how to measure, size of subgroup, who should do it, and what to do with the data, are all the beginnings of the right questions to ask about sampling. This all must be done through careful interaction and team work with persons in quality, engineering, and throughout manufacturing--getting accurate sampled data is not easy--but it is vitally important since so much relies on this foundational data. It is also vitally important that we begin to take steps to integrate this into all that we do and all that we are about--from the point of the raw material being ordered to the finished product being shipped. Just as basic SPC forms the basis for our technical future, attitudes we must build and engender for all of our people are vitally important--all of the tools in the world cannot make any difference if I do not want to learn them or use them.

Each type of charting and inspection system has advantages and disadvantages. For example, the attribute system generally requires little if any sophisticated gaging apparatus, and is relatively easy to use, since it is essentially a "judgement call". But the tradeoffs in ease of use and lowered immediate cost due to little or no equipment are balanced with less than accurate data and variation among and between operators--judgement calls.

But in many cases the attribute system, with the emphasis on color, texture, smell, and so on, may be perfectly acceptable. This is particularly true where we are just beginning to track a characteristic or critical dimension. And if we actually obtain more accurate data, will it do us any good? Not in all circumstances, but in many situations it is critical for

us to strive to upgrade attribute systems to variable systems. The cost for obtaining the data is generally directly related to the accuracy required in measurement. We may also conduct a short run study using attribute tracking and logging charts, and take charting no further based on the data collected.

Variable data affords the opportunity to provide a tolerance of acceptability and function. While the attribute approach merely "appears" acceptable or unacceptable, we can come much closer to "knowing" with reasonable certainty and definiteness, that we are actually producing product within the acceptable limits. A degree of "seat of the pants" decision-making and management style can be eliminated or upgraded through conversion to variable from attribute data and systems.

It is also important to try to build the variable data collection system to be automatic as opposed to manual. That is, it may be acceptable to begin the variable system with a scale, caliper, or micrometer which is manually measured and recorded on the chart. But ultimately, it is important to extract the data via electronic gaging systems which are automatic or semi-automatic, and "on-line". On-line means that data is being automatically extracted, and electronically, or digitally, logged in for comparative analysis and enhanced control. In other words, the operator has the information virtually at their fingertips for enhanced decision-making.

Obviously, the intention of the broader SPC system, using all of the statistical tools and techniques, but particularly the various charting systems, is to aid in solving problems, controlling production and improving the overall situation over time. But this does not happen by chance. It is only with a disciplined and by design, overt effort, that employees can expect to improve quality and productivity, and reduce cost and improve customer relations in the process.

Part of the reasoning for using collected data is to calculate upper and lower control limits for control. This is known as process control, since the control limits are being calculated from the actual data means in production. This, contrasted to pre-control or trending, will show actual relationships among on-going production shift to shift, day to day and so on, over time. This is important because the actual measurements being taken represent the actual problems (and inherent opportunities for improvement) in production day to day. If we arbitrarily apply control limits based on customer or other demands, rather than based on actual calculated means in production, we simply do not get the full benefit of the data--for problem solving purposes.

It therefore should become the goal of all of us to shift from simple pre-control to process control over time, since we get a more useful reading on actual production as it is happening--depending on how readily we can react (real time) to the production derived data--as well as how quickly we can gather and chart it, and so on. But this concept--shifting from pre-control to process control with actual calculated versus arbitrarily determined control limits--is also consistent with moving from attribute to variable data and inspection systems. Both of these--attribute to variable conversion, and pre-control to process control--are basic approaches to on-going improvement, and represent very real and identifiable measures of success in our work areas.

We should eventually gain sufficient control over time to enable phasing out the charting system all together. Assuming we do improve over time, and that we learn, getting better and better in the process, theoretically at least, we should be able to phase out the entire chart at some point. This also relates to earlier statements that we may quit tracking selected characteristics eventually, based on the ongoing improvement in process.

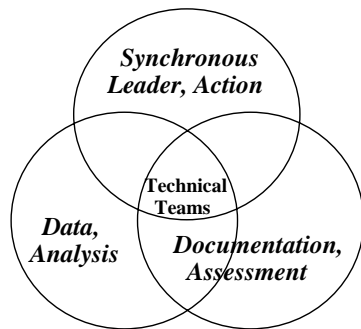
Integral to making the system work are teams--fully communicating and focused on appropriate technical problems and issues for improvement. This also assumes that teams have fully functioning members from the floor, using operator expertise along with quality, engineering, maintenance, supervisors and certainly others, as part of the managerial function--and perhaps other functions as needed. But pivotal to their success is the ability to spot problems, numerically, graphically, and in other very real day-to-day ways, in production and in other places.

Team Based Problem Solving, Six Sigma Variation Reduction For Lean

Focusing specifically on problem solving, and based on what was presented in the past section, this section provides a context within the broader infrastructure. This section is designed to give general guidance to a technical problem solving system which is designed particularly around the various toolkits. The broad basis for the system being discussed is the tool kit technological change model, using data and documentation with teams and leadership within the proper culture as depicted in the model earlier. As has been pointed out in various sections, this is essentially the context of the six sigma kaizen approach for lean.

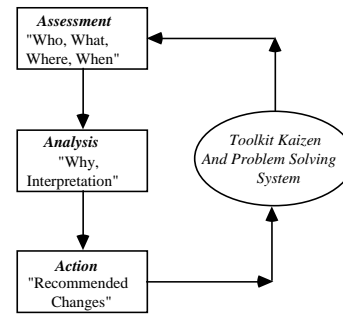
A graphic nearby depicts relationships between and among tool kit elements for team building, problem solving and improvement. Appreciating that the team requires data and documentation seems less difficult to understand on the surface, relative to how to conduct the broader problem solving act. It is the synergy inherent in the relationships built around and between the data, documentation and leadership, all synchronized toward the collective team effort which can and must provide the technical solutions' infrastructure as well as mechanism. This only happens if proper design consideration is given to infrastructure and organizational aspects technologically and regarding human resource issues.

The overall problem solution will be a function of three fundamental phases, each wrapped within the broader context of data, documentation and synchronized leadership as depicted throughout the tool kit. The phases are assessment, analysis and action, each to be further explored and defined within the remainder of this section. This would seem to be at the core of the concept of ongoing improvement.



Toolkit Relationships For Technical Team Building, Problem Solving And Improvement.

The graphic nearby provides the three tiered linear relationship between assessment, analysis and action, all related and ongoing based on feedback within the context of problem solving. While depicted as a straight line linear function, obviously the functions will not always be this discreet, straight forward and simplistic. Relationships embodied provide a useful strategy for the context of bringing forward technical solutions and improvements.



Toolkit Kaizen And Problem Solving System.

Assessment. During the assessment phase of problem solving, the team must document the current circumstances surrounding the problem or opportunity for improvement. This may involve demographic data such as persons and equipment involved, process flow charts of the macro process as well as the micro process. Regardless, much documentation will be involved to flush out the "who, what, where and when" type issues surrounding the way we currently do what we do. This could be a total line or production job site at the macro level and a micro work area within the broader system further analyzed. Both would likely require layouts, time and cost data, standard operating procedures, and flow charts on the current process and system.

Product design and specifications documentation would also be well advised as part of the assessment. Various tools for data analysis and documentation would begin to be formulated as a function of the nature of the product and process. It should also be clear that the data and documentation tools selected and used in the assessment phase will have a direct relationship to outcomes overall for the study in general, and subsequent phases in particular. Based on a thorough survey of all persons engaged in the work areas, it is quite likely that specific areas for further analysis will become apparent.

Analysis. While the major focus for assessment was to determine the current methods for processing product, the analysis phase builds on and around the assessment. Data and documentation begun in the assessment phase are fine tuned and multiple iterations may be required based on further analysis. Various experiments or trials may be run to determine optimum conditions or to further analyze what was flushed out from the original assessment. Pivotal in the analysis phase is the establishment of baseline data and documentation as performance

baselines upon which to base measures of improvement. As baselines are established, sources of variation are determined, focused on and causes flushed out for optimization. Stabilization in process must be achieved in reasonable ways, facilitating a clearer understanding of broader relationships in production process. As this occurs, factors and levels appropriate for further study will begin to surface. But this all assumes that under control conditions can facilitate a sufficiently "noise free" circumstance for focused improvements. This analytical environment can demonstrate optimum conditions in process.

Conflicting views or information may be found in the assessment, requiring various analytic tools and/or further clarification. Tools being required at this phase may consist of basic data such as attribute and variable charts, gage R & R, CPC, and so on, all organized within the ongoing process control plan (OPCP) and failure mode and effects analysis (FMEA) documentation tools. When these tools are used in the team mode, the overall complexity of the problem solving situation has shifted. Perhaps only one or two tools will be used, rather than all at the same time. But the array of tools available for analysis should not be under stated. The number of iterations with any one tool, to continue to interpret and understand the overall problem circumstance for improvement, will vary. Quality of problem solution will determine whether further iterations may occur.

Action. The final phase in the pursuit of a problem solution will be recommendations for action. Actions consist of new procedures to be followed uniformly in the process, new equipment based on conclusions that processes analyzed were not capable, or others. All of this drives establishment of new standards, training and additional studies. Assuming new equipment is evaluated as being appropriate for implementation as part of the solution and improvement, new studies and iterations will be required. It would be quite common to determine that additional training were required, or better gaging needed, or shifts from one characteristic to another to be studied identified. Costs of such actions will need to be detailed and presented with justifications for changes, and hopefully, improvements noted.

SOP as a pivotal part of the system. It is important that we all recognize the importance of the SOP as a documentation and communication device. Many persons, depending on function, may not have had a great deal of involvement in building and/or using the SOP. As teamwork and cooperation become increasingly prevalent for all persons,

particularly for supervisors and operators, it is believed this device will become more useful and necessary. The SOP is a basic communication and documentation device representing and summarizing many other inputs and sources of information.

The importance of the SOP as a documentation device cannot be over stated. This is true since:

1. When customers wish to determine how we are processing their product, the SOP provides a foundational part of the basis.
2. When, internally we wish to determine how something is being done, processing a product, the SOP provides one key part of the basis.
3. When suppliers wish to better understand what is happening to raw material or component in our procedures, SOP's are a key starting point.
4. When detailed analysis is done on any application, SOP's have critical information.
5. When being considered for new customer programs or projects, SOP's demonstrate capability in similar and applications.
6. SOP's document a necessary "paper trail" of information for certification and verification of process if questions arise in the future.
7. The SOP is an excellent system for evaluating ourselves over time, based on strong and accurate documentation of process at various points in the evolution of a given product.
8. SOP's provide a rather straight forward method for teams, individuals, departments and others to see and understand their involvement and role in product--and their responsibility and accountability can become increasingly apparent--and stronger.

Each SOP can be a substantial document, one that forms the basis for much that is important to the future. The SOP relies upon accurate and timely information--gathered and compiled by many persons and organizations. Obviously, the SOP can only be as good as the information upon which it is based--and the people which are providing the data. The role of the operator, supervisors, teams and others who support them, ought not be underestimated in the SOP--either in building or using them.

One key function of SOP's is training. Whether used with new or ongoing employees, if the SOP is correctly and carefully put together, it can form the basis of essential steps for the operator to follow. When "coached" by knowledgeable operators and others, such as a supervisor, based on the SOP documentation, the new operator can be efficiently

brought up to speed--and proficient in the correct methods and procedures--based on the SOP.

One additional connection remains related to variation reduction based on documentation, and used both as routine communication as well as for non-traditional training and education purposes. The issue is one of having the ability to get all persons on the same sheet of music, applying the documentation in creative ways to assure that all are running equipment in similar ways. When this occurs it represents a substantial step toward variation reduction since we can more readily attain consistency in production. This becomes even more essential as we connect the data with the documentation to perform increasingly robust analyses and studies for improvement.

