

Right Sizing Quality Assurance

Walt Lipke
Oklahoma City Air Logistics Center

Generally, quality assurance (QA) functions are sized at the direction of management and are rarely sized commensurately with their need. Over the years, influenced strongly by in-vogue attitudes and real-world circumstances, the size of the QA function has exhibited extremes: (1) inordinately large after an embarrassing product failure, or an executive's overreaction from attending a W.E. Deming seminar, or (2) completely eradicated when perceived to be unneeded or too expensive. This article introduces quality efficiency indicators that facilitate right sizing the quality assurance function, i.e., sizing QA to the customer's need, or the producer organization's own quality goals. The interpretation and application of the indicators is explained, and a simple example is provided demonstrating the calculation for sizing the QA function.

After a decade of performing process improvement, rework for our organization's software development projects was dramatically reduced from approximately 75 percent of the total effort to a very low value of 3 percent. When the percentage was high, rework was easily identified; for a small amount of quality assurance (QA) effort, a large quantity of rework was generated. As our production process improved, it became increasingly more difficult to identify defects. With rework now at 3 percent, we began to examine the economics of further improvement and the possibility of reducing the QA effort. Economically, the concept arises of right sizing the QA function with respect to the needs of the customer(s) or the quality goals of the producer organization.

Background

Generally speaking, companies are concerned with the quality of their products. Consequently, an organizational entity exists that is devoted to performing reviews, inspections, and testing for conformity to the product requirements, i.e., the QA function. However, the QA function is a cost affecting the price of a company's products. There is a cost for quality; it is not free. Thus, the QA function is connected to economic benefit.

At a minimum, QA functions should be sized sufficiently to satisfy the customer's requirement for product quality. In conflict, several pressures influence the size of the QA function. The customer wants the product at a low price with no flaws. The producer wants to *make money*, be competitive, and increase business – QA is a cost to be trimmed. Clearly, it is impossible to simultaneously satisfy these parties.

There are conflicting dynamics within the producer's organization, too. In competitive areas (multiple producers of the same product), the marketplace decides the product price. In turn, this places a constraint on the amount of rework and quality assurance the product can have and still be competitively priced. Regardless, the QA function

has the desire to achieve zero defects for the entire production process and believes it is in the best interest of the company to support this goal. However, a defect-free product most likely will not be affordable. Without some balance to the interests of the QA function, it can become too large. These are the influences of the classic market-share dilemma.

From the producer's perspective, QA needs to be efficient and rework minimized. Minimizing the cost of QA and rework makes the product more competitively priced and maximizes profit. A good production process will satisfy nearly all of the customer's requirements without QA, i.e., quality is built in, not inspected in. Likewise, a good QA process will identify most, if not all, of the nonconformance.

The customer, reasonably, cannot expect a perfect product. However, customers can mitigate their risk of purchasing poor products by testing performance and inspecting physical details during the production process and prior to accepting delivery. By performing product acceptance, the customer increases his cost of acquiring the product. His investment in product testing and inspection is an expense, and a portion of the product price is attributable to the customer-generated rework.

Defects not identified by the producer are subject to detection by the customer during his product testing and inspection. The customer's perception of product quality is created largely from the defects he identifies. To gain repeat business or good references for new business, the producer strives to minimize the defects that propagate, or leak, through his production and QA processes.

Quality Process Indicators

Minimizing the expenditure for QA yet meeting the customer's quality requirement is not a simple matter. To accomplish the task, management must have indicators for improving the processes and achieving the needed level of quality. In the following,

three measures of quality efficiency are proposed for determining the effectiveness and stability of the production and quality processes.

To better understand the subsequent discussion, the intended meaning of defects and rework is provided. The product requirements are the potential defects. A defect is nonconformance to a requirement, created as a function of the production process and its employees. Defects may be identified at any time during the production process up to customer acceptance. Rework results from the defects identified. Therefore, rework is a function of the QA process, QA employees, and customer testing and inspections. In mathematical form, defects and rework are expressed as follows:

Defects = f(production process, production employees)

Rework = f(QA process, QA employees, customer verification)

For an adequate understanding, a producer must have knowledge of the effectiveness of production and QA processes. Also, the producer needs to have information concerning the quality efficiency (QE) of the QA process itself. By having this information, the processes can be improved and the amount of improvement can be quantified.

Three measures are proposed to satisfy the information needed by the producer. These measures provide the capability for determining the goodness of the production and QA processes. The definitions of the measures are described below:

$$QE_1 = R(\text{process}) / R \quad 1)$$

where,

R = total rework costs
R = R(process) + R(customer)
R(process) = rework from the production process

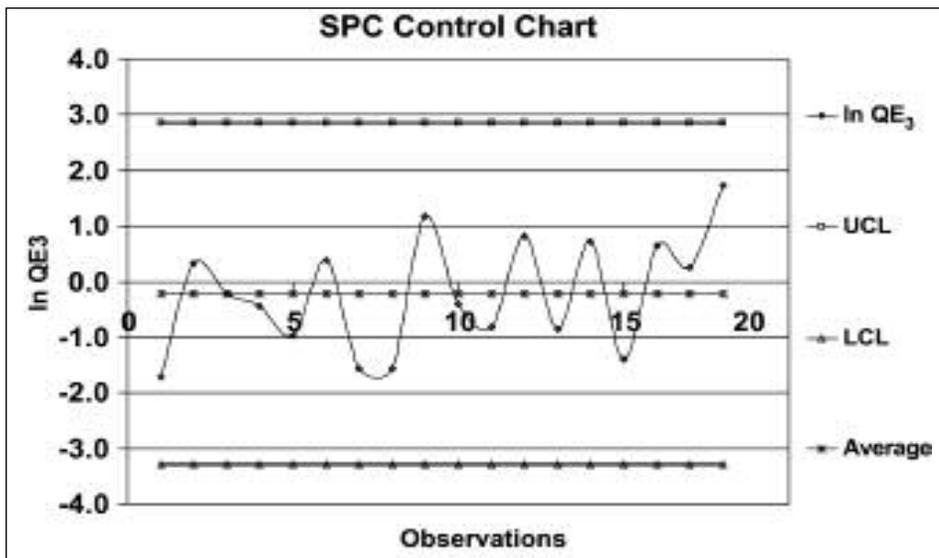


Figure 1: Statistical Process Control Chart

R(customer) = rework from the product inspections and testing conducted by the customer

The indicator is a measure of the QE of the quality process. When QE₁ indicates the customer identifies an excessive number of defects, improvement is needed from the QA process and its employees. Rework can come from non-requirements when good requirements management is not practiced. However, only rework from nonconformances to requirements is used in the calculation of the indicator.

$$QE_2 = P / T \quad 2)$$

where,

P = production costs
T = P + R + Q = total effort
Q = quality assurance costs

The indicator is a measure of efficiency of the production process. When QE₂ indicates excessive defects, the performance of the production process and its employees requires improvement.

$$QE_3 = R(\text{process}) / Q \quad 3)$$

The indicator is a measure of efficiency of the production and QA processes. When QE₃ is much greater than 1.0, the production process is examined for improvement. Conversely, when QE₃ is much less than 1.0, the QA process requires review and improvement.

Analysis

Satisfactory QA is indicated when all three indicators approach the value 1.0. As seen from examining the equations, it is possible for QE₁ and QE₃ to be equal to 1.0. However, it is not possible for QE₂ to have a

value of 1.0 when R and Q are not zero. The only condition for which QE₂ can equal 1.0 is when R=0.0 and Q=0.0, i.e., perfect process quality. It has been written that the minimum value of QA needed to maintain a high achieving quality process is 2.5 percent of the total effort [1]¹. Thus, the maximum value expected for QE₂ is 0.975.

The indicator QE₁ has the most influence on the customer's perception of product quality. Of the three indicators, it is the only one for which perfection (QE₁=1.0) can be consistently achieved. Thus, R(customer) = 0.0 (i.e., zero defects are identified by the customer) can (and should) be an expected outcome of the production and QA processes².

Under normal conditions, the value of QE₃ will approach 1.0, when the QA process is effective. However, as QE₁ and QE₂ approach the value of 1.0, QE₃ will approach zero. Using the equation for QE₃, this circumstance is more clearly understood. As the production process improves and approaches zero defects, the numerator, R(process), approaches 0.0. Concurrently, the denominator, Q, approaches its minimum value (2.5 percent of total effort), and thus, QE₃ approaches 0.0.

Indicators QE₁ and QE₂ may be used as evidence of defect prevention. The concept of defect prevention is that the QA process minimizes or eliminates the propagation of defects to the customer, and the production process has been optimized such that rework and QA are minimized [2]. QE₁ provides information concerning the amount of defect leakage from the QA process to the customer. Simultaneously, QE₂ provides information concerning the optimization of the production process. Taken together, these indicators show how well defect prevention is being achieved. When QE₁

approaches 1.0 and QE₂, simultaneously, nears 0.975, the production and QA processes are performing defect prevention at a level nearing perfection.

The indicators, QE₁, QE₂, and QE₃, are to be observed as both cumulative³ and periodic values. The cumulative number provides information as to the status of the process over a span of time. The periodic values yield trend information and help to answer the question, "Is the process improving, or is it getting worse?"

Quality Function Sizing

When the indicators QE₁, QE₂, and QE₃ are satisfactory with respect to the customer's needs or the organization's quality goals, and QE₃ is in statistical control, the QA function can be reliably sized. Likewise, the QA function can be sized for a new project using the data from a historical project, as long as the production and quality processes are unchanged. A statistical process control (SPC) control chart of the periodic observations of QE₃ is used to determine if the quality and rework processes are in control [3]⁴. The control charts may also be used as a *Rum* chart [3] for detecting the process reaction to improvements implemented.

As an example, Figure 1 is a SPC control chart created from real project data, shown in Table 1. As clearly seen in Figure 1, all observed values are within the upper and lower control limits shown as upper confidence limit (UCL) and lower confidence limit (LCL), respectively. Thus, the processes governing QE₃ are statistically stable.

Upon achieving statistical control, the QA function is sized from the periodic observations of Q/T, i.e., the quality investment as a fraction of total effort. From the average of these observations and their statistical variation, a 95 percent confidence value can be calculated for Q/T. At 95 percent confidence, we are 95 percent certain the actual QA requirement will be less than the size of the function created. Sizing QA at 95 percent confidence mitigates the risk of not sizing the QA function adequately.

The 95 percent confidence we are seeking is the UCL of the 90 percent confidence interval; 10 percent of the normal distribution is outside of the confidence interval, 5 percent below the LCL, and 5 percent above the upper limit.⁵ Having a QA requirement less than the lower confidence limit is not a concern; therefore, only the upper limit is used.

The 95 percent confidence limit, (Q/T)_u, is used in a linear relationship between the total effort cost (T) and the size of the QA function, i.e.,

$$Q = (Q/T)_u \times T$$

where,

Q is the expected cost for QA

This relationship is to be used with the project plan, specifically the monthly expenditures for total effort, to *right size* the application of QA resources. Performing the computations for the monthly values of Q will yield a funding profile for the QA function. In turn, this profile may be converted and used as the staffing profile.

To compute the 95 percent confidence limit, the periodic observations of Q/T are used as logarithms to make the statistical calculations⁴. The standard deviation σ is estimated for $\ln(Q/T)$, while the logarithm of the cumulative value, $(Q/T)_c$, is the estimate for the average value. Therefore, the confidence limit is first computed as a logarithm. Thus, the equation for the calculation of the 95 percent confidence limit follows:

$$(Q/T)_u = \text{antilog} [\ln(Q/T)_c + 90\% \text{ confidence interval}] \text{ (see Note 6)}$$

The antilog value, $(Q/T)_u$, is the appropriate number for the sizing computation.

Using the project data from Table 1, the value of $\ln(Q/T)_c$ is computed to equal -2.7662, with a standard deviation, $\sigma=0.5048$. From the values for z (=1.645), σ , and n (=18), the 90 percent confidence interval is calculated to be 0.1957. Adding $\ln(Q/T)_c$ and the 90 percent confidence interval yields the value -2.5705. The value of $(Q/T)_u$ is then computed from the antilog of the sum, and is determined to be 0.0765. For this project, the right size for the QA function is computed to be 7.65 percent of the total effort.

Summary

To economically apply QA requires three indicators of quality efficiency converge and approach 1.0. Two indicators are measures of defect leakage to the customer and from the production process, and the third measures the efficiency of identifying defects. The indicators are useful for improving the production and QA processes. Ultimately, upon achieving *in control* processes, the quality assurance function can be sized commensurately with the customer need, or the producer's quality goals⁷. ♦

References

1. Crosby, Philip B. Quality Is Free. New York: McGraw-Hill, 1979.
2. Paulk, M., et al. Capability Maturity Model for Software, Version 1.1. Software Engineering Institute, CMU/SEI-93-TR-24. Feb. 1993.
3. Pitt, H. SPC for the Rest of Us. Reading, MA: Addison-Wesley, 1995.

4. Crow, E. L., et.al. Statistics Manual. New York: Dover, 1960.

Notes

1. *High achieving* means nearly all of the producer's effort is in production. Extremely small efforts are performed for QA and rework to achieve the product requirements. In the author's opinion, very good quality for software producers would be $QE_1 \geq 0.98$, $QE_2 > 0.8$, and QE_3 between 0.6 and 1.2. World-class quality would be characterized by $QE_1 = 1.0$, $QE_2 > 0.9$, and QE_3 between 0.8 and 1.1.
2. The customer is still at risk of product defects, even when $R(\text{customer}) = 0.0$. Defects may be missed by the customer's inspection and testing.
3. Cumulative values for the three quality efficiency indicators are computed using the total values of the two parameters involved. For example, the cumulative for QE_2 would use total values for P and T.
4. When applying statistics, it is recommended to use the logarithm values of the periodic observations of QE_3 and Q/T. These parameters have been statistically tested as logarithms, and appear to be normally distributed. The results of statistics applications such as SPC and Confidence Interval are improved when the representation of the observations approximates a normal distribution.
5. The Confidence Interval is the region surrounding the computed average value within which the true value lies with a specified level of confidence. The end points of the interval are the Confidence Limits. The equation for the Confidence Limits is :

$$\langle x \rangle \pm z (\sigma/\sqrt{n})$$

where,

$\langle x \rangle$ is the average value of x , while z is from the standard unit normal distribution and corresponds to the area selected (for this application, $z = 1.645$ at 95 percent of the distribution area), σ is the standard deviation of the observations of x , and n is the number of observations [4].

6. The calculation is easily performed using the capability within personal computer spreadsheet applications, such as Microsoft's Excel.
7. Sizing the QA function using the method presented in this article assumes there is a semi-smooth flow of effort, and the requirement for QA is not sporadic.

Month	Rp	Q	T
1	15	83	1784
2	234	170	2808
3	124	154	3445
4	106	165	3051
5	39	103	2303
6	546	373	6178
7	30	143	2371
8	32	154	3374
9	247	77	2020
10	53	79	2321
11	75	169	3638
12	82	36	1473
13	221	518	4294
14	227	111	1111
15	191	768	4669
16	159	84	3571
17	144	111	3218
18	449	80	2059

Table 1: *Real Project Data*

About the Author



Walt Lipke is the deputy chief of the Software Division at the Oklahoma City Air Logistics Center. The division employs approximately

600 people, primarily, electronics engineers. He has 30 years of experience in the development, maintenance, and management of software for automated testing of avionics. In 1993 with his guidance, the Test Program Set and Industrial Automation (TPS and IA) functions of the division became the first Air Force activity to achieve Level 2 of the Software Engineering Institute Capability Maturity Model[®] for Software (SW-CMM[®]). In 1996, these functions became the first software activity in federal service to achieve SW-CMM Level 4 distinction. The TPS and IA functions, under his direction, became ISO 9001/TickIT[®] registered in 1998. These same functions were honored in 1999 with the Institute of Electrical and Electronics Engineers Computer Society Award for Software Process Achievement. Lipke is a professional engineer with a master's degree in physics.

**OC-ALC/MAS
Software Division
Tinker AFB, OK 73145-9144
Phone: (405) 736-3341
Fax: (405) 736-3345
E-mail: walter.lipke@tinker.af.mil**