

Center for Quality and Productivity Improvement
UNIVERSITY OF WISCONSIN
610 Walnut Street
Madison, Wisconsin 53705
(608) 263-2520
(608) 263-1425 FAX

Report No. 123

Total Quality: Its Origins and Its Future

George Box

January 1995

The Center for Quality and Productivity Improvement cares about your reactions to our reports. Please direct comments (general or specific) to: Report Editor, Center for Quality and Productivity Improvement, 610 Walnut Street, Madison, WI 53705; (608) 263-2520. All comments will be forwarded to the author(s).

Total Quality: Its Origins and Its Future

George Box

*Center for Quality and
Productivity Improvement*

University of Wisconsin
Madison, Wisconsin

ABSTRACT

This article discusses how an efficient organization is characterized by its knowledge and learning capability. It examines the learning ability of the human animal, the logic of continuous, never-ending improvement, the catalysis of learning by scientific method, and Grosseteste's Inductive-Deductive iteration related to the Shewhart Cycle. Total Quality is seen as the democratization and comprehensive diffusion of Scientific Method and involves extrapolating knowledge from experiment to reality which is the essence of the idea of robustness. Finally, barriers to progress are discussed and the question of how these can be tackled is considered.

KEYWORDS: *Extrapolating knowledge; Knowledge and learning;
Robustness; Total quality.*

Total Quality: Its Origins and Its Future

George Box

This article discusses how an efficient organization is characterized by its knowledge and learning capability. It examines the learning ability of the human animal, the logic of continuous, never-ending improvement, the catalysis of learning by scientific method, and Grosseteste's Inductive-Deductive iteration related to the Shewhart Cycle. Total Quality is seen as the democratization and comprehensive diffusion of Scientific Method and involves extrapolating knowledge from experiment to reality which is the essence of the idea of robustness. Finally, barriers to progress are discussed and the question of how these can be tackled is considered.

1. AN EFFICIENT ORGANIZATION IS CHARACTERIZED BY ITS KNOWLEDGE AND CAPABILITY TO LEARN

The reason that, having started as a chemist, I became a statistician was that Statistics seemed to me of much greater importance. It was about the catalysis of scientific method itself.

Scientific method is concerned with efficient ways of generating knowledge. It is also the essence of the quality movement which, at the most fundamental level, is not so much concerned with quality but with knowledge. As we continuously get to know more about our products, our processes and our customers we can do a better job of producing a high quality product, of reducing development times, of lowering costs and, indeed, of achieving whatever objective we choose. Furthermore the process of generating knowledge itself provides motivation and improves the morale of the whole work force.

2. THE LEARNING ABILITY OF THE HUMAN ANIMAL AND THE LOGIC OF CONTINUOUS IMPROVEMENT

Human beings are distinguished from the rest of the animal kingdom by their extraordinary ability to learn. There seems, however, to be a great deal of misunderstanding about the way this process of learning works. For example, for people trained in mathematical optimization never-ending improvement may seem a foolish goal. The law of diminishing returns – they say – surely implies that once you have reached a point somewhere close to the optimum it is wasteful to continue seeking

improvement.

This argument is fallacious. It supposes that, in a relationship $y = f(x)$, the identity of the "output responses" y , the function f , and the "input variables" x , are all immutable and given. In the real situation (see for example Box, 1994) we learn about them as we go. New possibilities are revealed for the choices of x , of f , and of y as we learn more about what we are doing.

The process of never-ending improvement is before our eyes in the past history of science. If we look at the remarkable discoveries of such scientists as Faraday, Darwin and Mendel, we see that what they found were not "truths". They were major steps in the never ending (and diverging) process of finding approximations which help predict natural phenomena. Thus, when Watson and Crick first discovered the structure of DNA, their paper appeared in a few pages in *Nature*. A problem was solved, but inquiry did not cease. On the contrary, there are now tens of thousands of papers describing further discoveries made possible only because of that paper. Each of these new discoveries leads to more questions.

Thus, *continuous never-ending improvement* is nothing else but the logical description of scientific method itself.

3. THE CATALYSIS OF LEARNING BY SCIENTIFIC METHOD

History shows that the human animal has always learned but progress used to be very slow. This was because learning often depended on the chance coming together of a potentially informative event on the one hand and a perceptive observer on the other

(Box, 1989). Scientific method accelerated that process in at least three ways:

- i) By experience in the *deduction* of the logical consequences of the group of facts each of which was individually known but had not previously been brought together.
- ii) By the passive observation of systems already in operation and the analysis of data coming from such systems.
- iii) By experimentation – the deliberate staging of artificial experiences which often might ordinarily never occur.

A misconception is that discovery is a "one shot" affair. This idea dies hard. It is characterized by the famous but questionable story of Archimedes exclaiming "Eureka!" as he leapt from his bath. We also have the example of Samuel Pierpoint Langley, at that time regarded as one of the most distinguished scientists in the United States. He evidently believed that a full sized airplane could be built and flown largely from theory alone. This resulted in two successive disastrous plunges into the Potomac River, the second of which almost drowned his pilot.

This experience contrasts with that of two bicycle mechanics Orville and Wilbur Wright who designed, built and flew the first successful airplane. But they did this after hundreds of experiments extending over a number of years. They began first with a large kite, then with a glider and finally a powered aircraft. One of many essential discoveries they made along the way was that the vitally important formula for the lift of an aerofoil was wrong. This led to building their own wind tunnel which, among other things, resulted in a major correction to the formula.

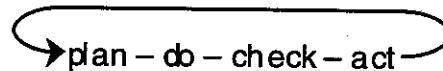
4. INDUCTIVE-DEDUCTIVE LEARNING ITERATION

That scientific inquiry is a deductive-inductive learning process seems to have been in the mind of Aristotle and was specifically stated by Grosseteste and later by Francis Bacon. Suppose we wish to solve a particular problem and that initial speculation about available data has produced some possibly relevant conjecture, idea, theory or model. On the basis of this tentative conjecture, we seek appropriate data which will further illuminate it. This data gathering could consist of (a) a trip to the library, (b) a brainstorming meeting between and/or within executives and workers, (c) careful passive observation of a process or (d) active experimentation. In any case the facts/data which are gathered sometimes confirm

what we believed already. More frequently, however they suggest that the initial idea is only partly right or, sometimes, that it is totally wrong. In the latter two cases, the *difference* between what we previously deduced and what the data actually seemed to say sparks off an inductive phase which seeks possible explanations for the discrepancy. These can point to a modified idea/model or even a totally different one which, if true, might explain what has been found. This new idea/model can again be compared with the data and with new data specifically gathered in order to test it. It is now well known that human kind have a two-sided brain specifically adapted to carry out such a continuing deductive-inductive conversation. This iterative process can lead to a solution of the problem. However, we should not expect the nature of the solution nor the route by which we arrive there, to be unique.

Each one of the general points made above relates to a particular aspect of quality improvement.

- i) We recognize in the deductive-inductive scientific iteration the Shewhart cycle



Within that cycle the various means of increasing knowledge are all represented.

- ii) Sharing already available knowledge to deduce previously unknown consequences is the objective of such devices as the "Seven Management Tools". For example, cross-functional deployment can bring together the combined expertise of many different departments and lead to an understanding of where the emphasis should be to achieve various customer requirements.
- iii) Shewhart control charts, flow charts and the SPC tools of Ishikawa are all devices for hypothesis generation from passive observation made on an operating system. This system might be student admission to a University or the manufacture of some commodity.
- iv) Experimentation in engineering and science can be made far more efficient by the use of statistical design aided by (largely graphical) analysis which makes direct appeal to the subject matter specialist.

5. DEMOCRATIZATION AND COMPREHENSIVE DIFFUSION OF SCIENTIFIC METHOD

A special contribution, for which the Total Quality Movement is responsible, is the democratization and comprehensive diffusion of simple scientific method. Many of the problems that beset us are, in essence, simple. Also, the essential information necessary to solve the problems is known to people who are in immediate contact with the process – the nurse with the patient – the process worker with the machine. Unhappily there is often no mechanism for this information to be used and no incentive to make it available. By providing the workforce with simple tools (such as those referred to in *iii*) above) and empowering their use, we ensure that thousands of brains are active not only in producing product but in producing information on how this product/process can be improved. When the resulting improvements are used fairly to benefit the workforce as one part of the organization, it can be remarkably effective not only directly, but indirectly, by producing a highly motivated workforce and by raising morale.

6. THE QUALITY MOVEMENT AS AN ANTIDOTE TO KNOWLEDGE FRACTIONATION

The situation with which human beings are presently confronted is illustrated in Figure 1. We are assured by human geneticists that human genetic structure has not materially changed over the last 10,000 years. If it were possible to take a baby born 10,000 years ago and bring it up in our present environment, it would turn out much the same as any modern day human being. Thus, the contribution of nature to the *nature plus nurture* equation (which might well be something like 50%) is essentially fixed. The process illustrated in the figure – of rather slow chance learning, gradually accelerating into scientific learning, further catalyzed nowadays by continuous improvement – has led to an ever widening gap between the technological environment in which we live and our natural genetic heritage. This has many consequences.

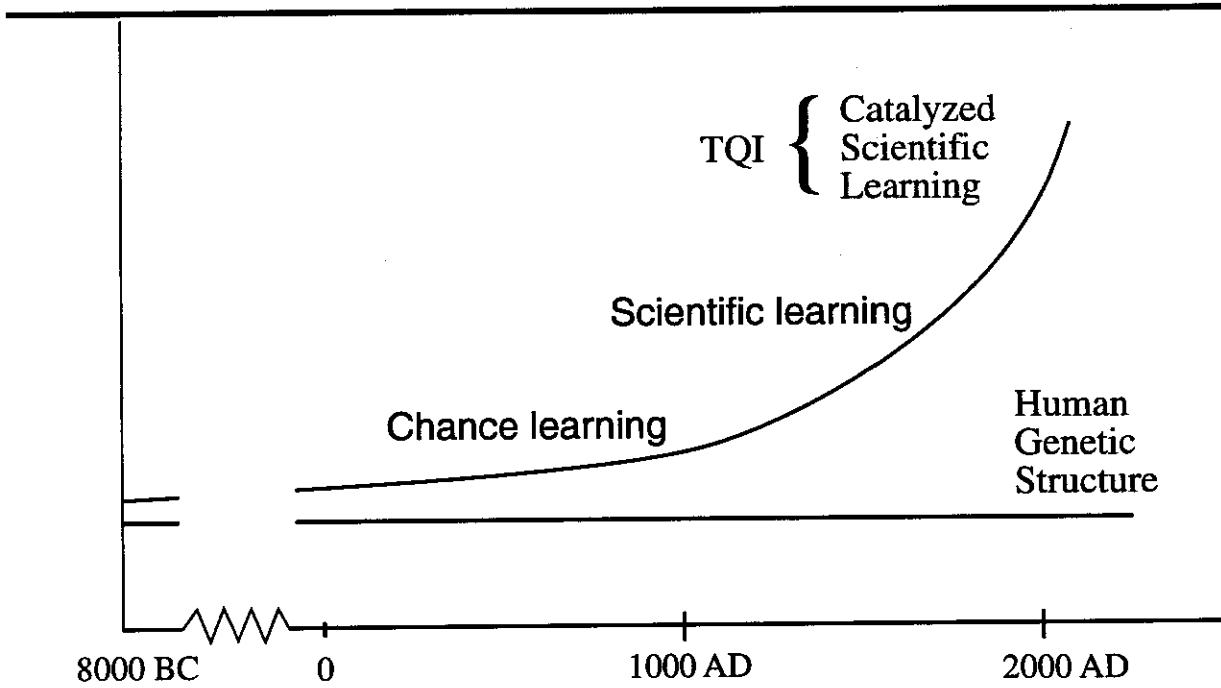


Figure 1. Three phases of increased human learning contrasted with nearly static human genetic structure.

6.1 NEED FOR IMPROVEMENT IN EDUCATION

We have attempted in the past to fill the gap between the two graphs by a process of formal education. In the developed world, this lasts from between $\frac{1}{6}$ and $\frac{1}{3}$ of a lifetime. But the educational process is presently stretched to breaking point by the ever widening rift. It is necessary now to apply the processes of quality improvement to education itself – to the teaching, administration and research of schools and universities. Acceleration of knowledge generation also emphasizes the need for *life long* education. The trained teacher, scientist or engineer can no longer regard what they have learned at the university as supplying their needs for the rest of their lives.

6.2 FRACTIONATION OF KNOWLEDGE AND MANAGEMENT TOOLS

We do not need to go back very far to find a time when available knowledge could be appreciated by the minds of single individuals. This happy state is evidenced, for example, in the early meetings, some 350 years ago, of what is now the Royal Society. Nowadays, *use of* the tools of every day experience – the airplane, the computer, the fax – and the knowledge of *how they work* becomes more and more remote. Furthermore, our mentors – engineers, scientists, business executives, and physicians – know more and more about less and less.

Whether they are employed in the manufacture of a car or the running of a university, quality management tools mentioned in 4(ii) above, can help to negate the effects of this knowledge fractionation.

6.3 HUMAN POSSIBILITIES AND LIMITATIONS

Computer technology has made available to us remarkable new ways of storing old knowledge. The skills of modern knowledge-retrieval are easily learned. Our schools and universities should, therefore, largely give up the hopeless task of teaching old knowledge. Instead they should spend much more time illustrating and practicing the deductive-inductive solution of novel unstructured problems – a skill which only the human brain possesses.

For the foreseeable future, at least, human beings must be an essential part of any operating system. But they cannot be treated as computers, robots, or machines. The brain possessed by every human being has inductive capabilities that are totally beyond our understanding. But what motivates and interests the

owners of these valuable brains, is what has motivated human beings throughout the ages, in particular curiosity and the opportunity to take pride in their accomplishment.

7. THE POTENTIAL OF SCIENTIFIC STATISTICS AS A CATALYST TO KNOWLEDGE ACQUISITION

Obviously the catalysis of scientific method is concerned with dealing intelligently with the generation and analysis of data. This is what the subject of statistics ought to be about. But its potential to be this has been hamstrung by the destructive idea that statistics is a branch of mathematics. Now it is true, of course, of any scientist – physicist, chemist or statistician – that the more mathematics s/he knows the better. However, the *raison d'être* for statistics and quality improvement is not mathematics but science. How much better off we would be if statisticians renamed their subject Scientific Statistics and prepared themselves to be statistical scientists. Any serious statistician must, of course, study the mathematics of statistics. But there is much more to the subject than that, and it is from science that the leadership and motivation must come.

The quality movement clearly shows the unfortunate consequences of past mathematical preoccupation. It also has the potential to show us how to get back on the right road.

I was told recently that, of the industrial statisticians that were members of the American Statistical Association, three quarters were employed by the pharmaceutical industry. The majority of these I suspect were hired to meet government testing regulations. Such facts help to throw light on the present crisis in statistics (Iman, 1994).

Consider the case of the development of a new drug to combat a specific disease (Box, 1994). There is, on the one hand, a process of *finding* the drug and a way to manufacture it – typically this takes many years of biological and chemical research and involves very many inductive-deductive iterations. On the other hand, there is the final stage of *testing* the drug for efficiency and safety. This is an activity which is, very properly, regulated by strict government control.

The process of *finding* the drug is like that of a detective finding the perpetrator of a crime using appropriate iterative inductive-deductive techniques. The process of *testing* the drug is like that of the trial of the suspect after he has been apprehended and

must be conducted in accordance with very formal procedures appropriate after the investigation is all over. The null hypothesis of innocence is to be tested against an alternative of guilt defined very specifically in the charge. Usually the products of our departments of mathematical statistics are reasonably well-trained to conduct this second function but have little or no knowledge of the first which is clearly of much greater importance and much more fun.

The production of statisticians who are testers rather than investigators arises, I believe, because in their training they were steeped in the one shot mathematical paradigm of

theorem → *proof*

not the iterative scientific paradigm of



The former requires one to state all one's assumptions A at the beginning and then to prove a theorem T with a statement of the kind: if A then (T|A). Since truth is never attainable, we substitute conditional truth. But we must remember that conditional truth includes such statements as "If the moon were made of green cheese it would be a wonderful place for mice."

Scientific statistics cannot take this way out. Again absolute truth is not attainable but we may be able to find how to approximate the world's behavior by an iterative process of discovery. This is a very different operation. In particular it is different because it cannot take place using mathematics alone. It requires at each iteration, the injection of *subject matter knowledge* by the engineer or other specialist. Thus the statistician cannot operate on his own but must be involved with the engineer or scientist – and not merely as a consultant, but as a colleague who is part of the investigational team. If we continue to train statisticians free from the contamination of practice, they will continue to be limited both in their studies and their usefulness to a rather small and uninteresting piece of scientific endeavor.

8. EXTRAPOLATING KNOWLEDGE FROM EXPERIMENT TO REALITY: ROBUST PRODUCTS AND PROCESSES

Applications of science are almost always* *extrapolations* from the original discovery (see for example Deming (1986)). For example, the reported results of an experiment performed in a chemical engineering laboratory in Dresden are of no interest unless that same phenomena can be repeated in other laboratories where, to some extent, the experimental environment is bound to be different. Furthermore, they may not be of much practical interest unless the principles involved apply, for example, to full scale manufacturing processes. From a statistical point of view, this means that only rarely can an application be credibly regarded as a member of the relevant reference set of the original investigation. Thus, we can't rely on such statistical reasoning to justify most applications of science. Suppose that an expensive industrial plant is to be built based on an entirely new principle discovered by a process of scientific investigation. The person responsible for the decision to go ahead with the project, perhaps a senior-engineer/manager, asks "What makes you think that this project is justified?" If you tell him, "Well we ran an experiment A in this test tube and an experiment B in this one and A was better than B" he is unlikely to be very happy. On the other hand, suppose you can tell him that in the investigation you ran carefully planned experiments in which a whole number of "environmental factors" were varied, to find out whether the process was sensitive or insensitive to the kinds of changes that would be met in its application. Also you checked out the most important of these on the small scale plant. Then he will feel much happier.

This idea of design for robustness is very evident in work at Guinness's early in this century. Varieties of barley were sought which would give reasonably good yields irrespective of the very different conditions in which they would be grown on different farms in Ireland. Later these ideas were formalized by Fisher (1935) in his discussion of a randomized block experiment by Immer in which six different varieties of barley were tested at six different locations in the State of Minnesota in two different years. In the same publication discussing the advantages of factorial experiments, he says "(extraneous factors) may be

* An exception is Evolutionary Operation whether the results of experiments on the actual process are applied to that process itself.

incorporated in (factorial) experiments designed primarily to test other points; with the real advantages that, if either general effects or interactions are detected, that will be so much knowledge gained at no expense to the other objects of the experiment; and that, in any case, there will be no reason for rejecting the experimental results on the ground that the test was made in conditions differing in one or other of these respects from those in which it is proposed to apply the results." These ideas were extended further by Youden (1961a) when working at the Bureau of Standards. He developed what he called "rugged" methods of chemical analysis. His experiments used *fractional* factorial designs, earlier introduced by Finney (1945) to further increase the number of extraneous factors that could be tested. Thus the concept of using statistics to design a product that would operate well in the conditions of the real world clearly has a long history going back at least to Gosset in the early part of the century.

An early *industrial* example is due to Michaels (1964). After explaining the usefulness in industry of *split plot* designs (often synonymous with Taguchi's (1987) inner and outer arrays, he illustrates with a problem of formulating a detergent. Michaels says, "Environmental factors, such as water hardness and washing techniques, are included in the experiment because we want to know if our products perform equally well *vis-a-vis* competition in *all environments*. In other words, we want to know if there are any Product \times Environment interactions. Main effects of environment factors, on the other hand, are not particularly important to us. These treatments are therefore applied to the Main Plots and are hence not estimated as precisely as the Sub-plot treatments and their interactions. The test products are, of course, applied to the Sub-plots." The quality revolution has once again emphasized the importance of ideas previously overlooked. We owe to Taguchi (1987) and Morrison (1957) demonstration of their wide industrial application.

While in some investigations the major concern is with producing a change in mean value of the response (increasing yield, for example), in others it is with the reduction of variation. In many instances desirable conditions can only be found by considering both mean and variance together.

Thus, studies in both the analysis and the synthesis of variation are important. The former uses experimental designs to determine variance components from different sources (a problem first considered by Henry Daniels, in the 1930's). The latter requires error transmission studies. This is a subject about which statisticians have traditionally

behaved somewhat irrationally. For example, on the one hand, a standard statistical model that might be used to formally justify the analysis of a factorial experiment supposes that all the error is in the response (output) variable y . On the other hand, in any discussion of replication we are told that to obtain a genuine estimate of experimental error a complete repetition of the whole *process of setting up each experimental run* must be made. The implication is clearly that an important part of the experimental error is transmitted from the inputs (factors, independent variables).

The application of error transmission studies to engineering design was first considered by Morrison (1957). The problem is to design an assembly such that error transmitted by deviations of the components in the assembly from nominal values will be minimized. Morrison's work emphasizes two points which were later missed in Taguchi's work, (a) that this is a problem in numerical minimization (not requiring inner and outer arrays, or indeed, an experimental design of any kind), (b) that the solution is extraordinarily sensitive to what is assumed about the variances of the components (see Box and Fung, 1994) and remarks by Box in the discussion edited by Nair, (1992). Once more, however, we owe to Taguchi's (1986) demonstration of the wide industrial importance of these rather different robust design ideas.

9. THE FUTURE

9.1 THE QUALITY MOVEMENT WILL NOT GO AWAY

Particularly in the area of management, we have experienced the coming and going of many initiatives – management by exception, management by objectives and so forth. There is a natural concern that the quality movement will experience the same cycle of rise and demise. However in so far as this movement concerns the application, diffusion and democratization of scientific method whether in the guise of an affinity diagram, a quality control chart, an experimental design or a study of group dynamics, it can scarcely fail – we are always better off with more knowledge about our system than with less.

9.2 BARRIERS TO PROGRESS

In a world of accelerating change some of our ideas must necessarily change. We are told by O'Shaughnessy (1844-1881), 'Each age is a dream

that is dying or one that is coming to be.' Thomas Kuhn (1962) takes up this theme, as it applies to science, in his book *The Structure of Scientific Revolutions* and St. Paul tells us to 'Test all things, hold fast to that which is good'. This process of change can be painful. Many of us try to hold on to a treasured belief too long, others seem to be deflected by every wind that blows. Salvation lies somewhere in between.

I often think the Quality Gurus do not help. They tend to teach a particular subset of quality ideas which comprise an immutable creed that they and their disciples believe must be followed to the letter. They rarely acknowledge (a) to what degree their particular nostrums overlap those of others or (b) that nothing is immutable except the fact of scientific change itself.

At the other extreme, we are faced with junior but powerful executives that are anxious to become senior executives and are poised to spring on or off every band wagon that passes. Closely related to this problem is the danger imposed by a myopic and exaggerated concern with short term results. 'Since water runs down hill,' any organization that is set up in such a way as to encourage such behavior will lose when competing with another organization with a longer term view. The history of industry in the United States and Japan between 1960 and 1990 surely demonstrates this.

Another barrier to change is mathematical statistics which, as I said, encourages a paradigm for scientific method which is appropriate only for testing and not for discovery. The important subject of Scientific Statistics will be taught in departments of statistics if they are far sighted enough to change. Otherwise it will be taught elsewhere and statistics departments may cease to be.

9.3 THE PROMISE OF THE QUALITY MOVEMENT

I have spoken of the divergence of scientific discovery. But the history of science tells us also of periods when disciplines, previously thought to be separate, were brought together. For example, it has been clear for a long time that, to solve certain problems, chemistry, physics and astronomy must be considered together; also only since Watson and Crick, the previously unsuspected relationship between genetics and coding theory has emerged. The Modern Quality Movement brings together ideas from, for example, Systems Analysis, Operations Research, Problem Solving, Statistics, Engineering, Group Dynamics, Management Science, Human

Genetics, and Organizational Development.

The Quality Movement will, from time to time, undergo healthy changes and may even be called by different names. I am sure, however, that it is here to stay and can be of great benefit to us all. In particular it (a) can help solve the many problems that face us (b) it can produce a far more efficient and happier work culture and (c) it can free us from many of the frustrations produced by poor quality of services and goods.

ACKNOWLEDGMENTS

This work was supported by a grant from the Alfred P. Sloan Foundation.

REFERENCES

- Box, G. (1989). "Quality Improvement: An Expanding Domain for Scientific Method," *Phil. Trans. R. Soc. London A*, Vol. 327. pp. 617-630.
- Box, G. (1994). "Statistics and Quality Improvement," *J.R. Statist. Soc. A*, Vol. 157, Part 2. pp. 209-229.
- Box, G. and Fung, C. (1994). "Is Your Robust Design Procedure Robust?," *Quality Engineering*, Vol. 6, No. 3. pp. 503-514.
- Deming, D.E. (1986). *Out of the Crisis*. MIT Press; Cambridge, MA.
- Finney, D.J. (1945). "Fractional Replication of Factorial Arrangements," *Annals of Eugenics*, Vol. 12. pp. 291-301.
- Fisher, R.A. (1935). *The Design of Experiments*. Oliver and Boyd; Edinburgh.
- Iman, R.L. (1994). "Statistics Departments Under Siege," *Amstat News*, Vol. 6, No. 212.
- Kuhn, T. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press; Chicago, IL.
- Michaels, S.E. (1964). "The Usefulness of Experimental Design," (with discussion), *Applied Statistics*, Vol. 13, No. 3. pp. 221-235.

- Morrison, S.J. (1957). "The Study of Variability in Engineering Design," *Applied Statistics*, Vol. 6, No. 2. pp. 133-138.
- Nair, V.N. (1992). "Taguchi's Parameter Design: A Panel Discussion," *Technometrics*, Vol. 33. pp. 127-161.
- O'Shaughnessy, A.W.E. (1844-1881). "The Music Makers," *The Oxford Book of English Verse*. Oxford.
- St. Paul, from the First Epistle to the Thessalonians.
- Taguchi, G. (1986). *Introduction to Quality Engineering*. UNIPUB/Kraus International; White Plains, NY.
- Taguchi, G. (1987). *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*. UNIPUB/Kraus International; White Plains, NY.
- Youden, W.J. (1961a). "Experimental Design and ASTM Committee," *Materials Research and Standards*, Vol. 1. pp. 862-867. Reprinted in *Precision Measurements and Calibration*, Vol. 1, Special Publication 300. National Bureau of Standards; Gaithersburg, MD. 1969, ed. H.H. Ku.
- Youden, W.J. (1961b). "Physical Measurement and Experimental Design," *Colloques Internationaux de Centre National de la Recherche Scientifique*, No. 110, le Plan d'Experiences. pp. 115-128. Reprinted in *Precision Measurement and Calibration*, Vol. 1, Special Publication 300. National Bureau of Standards; Gaithersburg, MD. 1969, ed. H.H. Ku.