

What is Statistical Engineering?
(Chapter 1, Section 1, Statistical Engineering Handbook)
Draft 2, March 2019
Roger Hoerl

1.1.1 Objectives

The purpose of this section is to explain what statistical engineering is; that is, how it is defined, how it works, why it is needed, as well as the basics of its underlying theory.

1.1.2 Outline

We begin with an elucidation of the definition of statistical engineering. Next, we explain why it is needed as a discipline, and then present the current state of the art in terms of its underlying theory.

1.1.3 Definition and Elaboration

The discipline of statistical engineering is: *the study of the systematic integration of statistical concepts, methods, and tools, often with other relevant disciplines, to solve important problems sustainably.*

Several words in this definition warrant explanation. First of all, statistical engineering is defined as a *discipline*, the *study* of something, not as a set of tools or techniques. Secondly, as an *engineering* discipline it does not focus on advancing the fundamental knowledge of the physical world, i.e., it is not a science. Rather, as with other engineering disciplines, it utilizes existing *concepts, methods, and tools* in novel ways to achieve novel results. In this sense it is complementary to statistical science, just as chemical engineering is complementary to chemistry.

Concepts, methods, and tools are each important, and need to be *integrated*. That is, formal statistical *methods*, such as time series or regression analysis, and individual *tools*, such as residual plots, need to be integrated with *concepts*, such as the advantages of randomization, and the need to understand the quality (“pedigree”) of observational data prior to developing models (Hoerl and Snee 2018). When addressing straightforward issues, a single statistical tool may suffice. However, as noted by Hardin et al. (2015), when solving the challenging problems often faced by practitioners, obtaining a viable solution typically requires integration of multiple methods into an overall strategy and sequential approach.

Such integration should be done in a *systematic*, rather than ad hoc manner. Throughout the history of statistics, good statisticians have generally figured out how to integrate concepts, methods, and tools to solve problems. One classic example would be Box and Wilson’s (1951) integration of experimental design and regression into an overall sequential strategy for the empirical optimization of processes, which we know today as response surface methodology.

It would appear clear, however, that despite many historical examples of successful integration, there is little existing theory in the literature on how to best accomplish such integration in general, that is, with a new problem. Due to a lack of theory, new integration problems are often attacked with a trial and error approach. However, the theory of statistical engineering, discussed below, provides guidance for a *systematic* approach, which is likely to be much more effective. In addition, such theory can be formally studied, taught, and advanced over time.

By the word *theory*, we do not refer to mathematical statistics. Rather, we refer to development of an overall methodology, based on the scientific method, by which one might approach integration in a methodical (systematic) rather than ad hoc manner. Note that theory may be defined as: “A coherent group of general propositions used to explain a phenomenon” (Hoerl and Snee 2017). Note that neither this nor other common definitions of theory contain explicit requirements for mathematics, although mathematics is often important.

In addition, for many of the important problems facing practitioners, such integration must include *other disciplines* beyond statistics. For example, almost by definition, information technology (IT) is required to address “Big Data” problems (see the ASA statement on Data Science at (<http://www.amstat.org/misc/datasciencestatement.pdf>). In fact, the authors of this handbook have found that IT is needed to some degree to solve most important real problems. Kendall and Fulenwider (2000) explain how critical IT is to successful Six Sigma projects, and we feel that the same is true of statistical engineering. Challenging problems, such as developing personalized medicine protocols through genomics, for example, are virtually impossible to resolve without effective and innovative use of IT.

Other disciplines may be needed as well, including natural sciences, other engineering disciplines, and also social sciences, such as organizational effectiveness, psychology, or social networking theory, depending on the specific problem being addressed. As one example, the improvement methodology Lean Six Sigma (Antony et al. 2017) is essentially the integration of diverse statistical methods, including control charts, experimental design, and regression, with various quality concepts and methods, including Pareto charts, mistake proofing, and quality function deployment (QFD), in addition to the efficiency concepts and methods from Lean manufacturing. These efficiency concepts and methods could be considered under the umbrella of the discipline of industrial engineering.

As an engineering discipline, the ultimate goal of statistical engineering is to *solve important problems*. While this may seem obvious, an emphasis on solving important problems gives statistical engineering perhaps its most important attribute, being *tool-agnostic*. That is, statistical engineering is neither Bayesian nor frequentist, neither parametric nor non-parametric (or semi-parametric), and does not promote either classical or computer-aided designs, per se. Rather, as an engineering discipline its “loyalty” is to solving the problem and generating results,

not to a predetermined set of methods. Tools are of course important, but within a statistical engineering paradigm they are chosen based on the unique nature of the problem to provide the best possible solution, rather than predetermined based on personal preferences. Various philosophies and tool sets may be employed and integrated.

Further, statistical engineering seeks solutions that are *sustainable*. We argue that many solutions, including those published in professional journals, provide technical solutions, but all too frequently these solutions are not actually sustainable over time. Of course, virtually no solution will be permanent, but statistical engineering seeks solutions that are *sustainable* beyond the immediately time frame, and hopefully last until the problem itself changes, or until new technology becomes available, enabling an even better solution.

In practice, purely technical solutions often overlook organizational, political, or psychological constraints. To be sustainable, the solution must eventually be embedded into standard work procedures and best practices, typically via IT. An interesting example from the related discipline of data science is the classic Netflix competition, in which Netflix paid \$1,000,000 to the team that developed the “best” model to predict customer ratings of movies.

As noted by Donoho (2017), however, the winning solution was never actually implemented by Netflix, because it found that the time and expense involved in maintaining the 107 individual models utilized within the overall ensemble (see Fung 2013) was not worth the small improvement in accuracy. So a team won the competition and the \$1,000,000 award, but it did not actually solve Netflix’s business problem. Clearly, the technical solution is only a part of solving important problems *sustainably*.

1.1.4 Why Statistical Engineering?

It is certainly logical to ask why a new discipline is actually needed, and even allowing that one is, why it should be statistical engineering. As noted previously, good statisticians have integrated multiple statistical methods, and tools from other disciplines, for a long time. In this sense, we could say that statistical engineering itself is old. However, as also noted above, such applications have typically been presented as isolated case studies utilizing ingenuity and creativity to provide novel solutions to complex problems. What has been missing is a concise presentation of an underlying theory as to how the researchers actually developed their solutions. A body of research is needed to fill in this gap, to develop an underlying theory as to how such problems should be addressed in general, and why. In this sense, we say that statistical engineering is a new discipline, even though statistical engineering itself is old.

The main reason statistical engineering was needed in these case studies was to solve problems that were not straightforward “textbook” problems. Textbook problems are typically well structured, have a clear objective, and a single, correct

answer; generally, one that can be looked up in somewhere in the textbook. For example, a data set might be presented with paired data, such as “before and after” weights from a diet evaluation study. Clearly, with paired data a standard independent samples t test would not be appropriate. Rather, a paired t test is likely to provide the “correct” analysis. We can look this up in the textbook in order to verify that it is the appropriate analysis, making reasonable assumptions.

However, real problems faced by practitioners are not usually so well structured. The specific problem to be solved may not be clear. Appropriate data for solving the problem may not yet exist. For example, suppose an international corporation’s reputation was damaged by the discovery that a supplier was - unknown to the corporation - using child or slave labor in a developing country. The corporation needs to address this issue immediately, so as not to support such human rights violations, and then begin perhaps a much longer process of rebuilding its reputation. But what exactly does “rebuilding its reputation” mean? How would this be measured and verified? How should the company go about acquiring data to set a baseline on its reputation? The answers to these questions are not obvious, and there is certainly no “correct” answer to look up in a textbook.

Further, it is unlikely that one statistical method would suffice to solve this problem. Some type of survey, or perhaps web scrapping of social media could be involved, followed by analysis of the data, perhaps with multiple tools. Additional data gathering and analysis steps might follow. In other words, there would be a need to first think through an overall strategy of how to attack the problem, then acquire data, then analyze the data using a mix of graphical and analytical tools. That is, there would be a need to link and integrate multiple tools in a sequential fashion, based on a strategy.

Very few statistical textbooks provide guidance on how to link and integrate multiple tools, especially through sequential cycles of data gathering and analysis. Rather, most textbooks provide details on individual methods, one method at a time; descriptive statistics, probability, confidence intervals, hypothesis testing, regression, and so on. Further, a theoretical foundation is needed to provide guidance on how to accomplish this integration, including the underlying theory of statistical engineering, which we present shortly.

Several other authors have noted this gap in the current body of research on tool integration to solve complex problems. For example, Meng (2009) pointed to the same issue, and subsequently added a new course in the Harvard statistics department curriculum, Stat 399, which “...emphasizes deep, broad, and creative statistical thinking, instead of technical problems that correspond to a recognizable textbook chapter”. Complex problems rarely correspond to a recognizable textbook chapter!

Shortly after the publication of Meng’s paper, Susan Hockfield, then President of MIT and a member of the General Electric (GE) Board of Directors, gave an interesting

perspective on the relationship between science and engineering, which has obvious ramifications for statistical engineering and statistical science (Hockfield 2010). She noted that around the dawn of the 20th century, physicists discovered the basic building blocks of the universe (i.e., the periodic table), which could be considered a “parts list.” However, it was engineers who figured out how this parts list could be put to best use, subsequently driving the electronics and computer revolutions. Similarly, Hockfield noted that biologists had recently discovered the basic building blocks of life (the human genome), another “parts list,” and now engineers are finding creative ways to use this parts list, such as in personalized medicine.

A key point Hockfield made was that there has been for some time a consistent “separation of labor” between science and engineering across diverse disciplines, although it is important that they collaborate. To be more precise in terminology, common definitions of the word “science” are similar to: “the study and advancement of the fundamental knowledge of the physical or natural world” (e.g., <https://www.merriam-webster.com/dictionary/science>). Various definitions of engineering are also available (e.g., <https://www.merriam-webster.com/dictionary/engineering>), but accepted definitions generally emphasize “utilization of existing science and mathematics in novel ways to benefit humankind”. An old saying in the engineering community is: “An engineer is someone who can accomplish for \$1 what any fool can accomplish for \$2.” While science emphasizes development of new fundamental knowledge, engineering finds creative ways to use this knowledge for the benefit of society.

We argue that this distinction between science and engineering applies to statistics quite well. Statisticians have been developing an excellent toolkit for over a century, which could also be considered a “parts list,” using Hockfield’s terminology. This is what the vast majority of statistics textbooks emphasize, as noted above. However, we argue that insufficient thought has gone into the engineering problem of how to best integrate multiple tools in creative ways to solve complex problems. At least, insufficient thought has gone into documenting the underlying theory of how to approach this engineering problem in general.

Acknowledging this problem, the American Statistical Association (ASA) published guidelines for the design of undergraduate statistics programs, noting (ASA 2014, p. 6):

Undergraduates need practice using all steps of the scientific method to tackle real research questions. All too often, undergraduate statistics majors are handed a “canned” dataset and told to analyze it using the methods currently being studied. This approach may leave them unable to solve more complex problems out of context, especially those involving large, unstructured data.... Students need practice developing a unified approach to statistical analysis and integrating multiple methods in an iterative manner.

Unfortunately, the ASA report did not suggest a specific method to provide a “unified approach to statistical analysis and integrating multiple methods in an interactive manner”. This is, in fact, the gap statistical engineering is intended to fill.

So there appears to be a clear consensus that a deep theoretical foundation in individual methods, while certainly valuable and needed, is not sufficient. In addition to sound statistical science, the profession also needs a well-developed theory and practice of statistical engineering, to ensure that society benefits from the many advancements that have been made in statistical science.

At first glance, some may feel that what we are calling statistical engineering is nothing more than a rebranding of applied statistics. However, this would be analogous to saying that chemical engineering is nothing more than a rebranding of applied chemistry. As a simplistic example, consider parents who buy their children a chemistry set for a birthday or holiday. If the children mix vinegar and baking soda, they might create a toy “volcano”, due to the subsequent chemical reaction. This is certainly applied chemistry! However, not many would consider this to be chemical engineering. The children did not use the laws of chemistry to engineering a solution to a real problem.

Similarly, whenever someone applies a statistical method to real data, this constitutes applied statistics. In many applications, particularly with relative straightforward problems, one method found in a textbook will suffice. The problem has now been successfully solved through applied statistics. However, with more complex problems, a single method will rarely suffice. More likely, a novel solution will have to be engineered, using the “parts list” of statistical science tools, perhaps integrated with tools from other disciplines.

Another important distinction between statistical engineering and applied statistics is that statistical engineering has an underlying theory, while “the theory of applied statistics” would be an oxymoron; it is *applied* statistics, not *theoretical* statistics. As noted by Nair (2008), there is a clear and well-established delineation between theoretical statistics and applied statistics, although hopefully these are intertwined.

1.1.5 The Underlying Theory of Statistical Engineering

1.1.5.1 What is Theory?

As we present the theory of statistical engineering, we should acknowledge that it is in its early stages of development. Of course, the theories of all known disciplines are in essence, “works in progress”, in that research in each continues to this day. For example, mathematics is one of the oldest known disciplines, and has been formally studied and researched for millennia. And yet, rigorous research in mathematics continues at universities and colleges around the world, with no evidence of slowing down. Having noted the ongoing development of the theories of all disciplines, the current state of statistical engineering theory is admittedly basic and relatively crude compared with more established disciplines, including traditional engineering disciplines. We anticipate that future research will add to

the current body of knowledge, eventually producing a rich literature documenting the theory of statistical engineering to a degree of rigor on par with other engineering disciplines.

The underlying theory of statistical engineering is quite different from the underlying theory of statistical science, which is based on mathematical statistics. That is, most of the theory of statistical science can be proven or derived using formal mathematics; calculus, real analysis, linear algebra, and so on. The theory of statistical engineering is not mathematical in nature, however. In other words, it is not based on a “theorem-proof” model. Rather, it is based more on empirical research, which demonstrates what does and does not tend to work to solve important problems sustainably, and why. Of course, there may be proofs that certain tools work better than others under specific assumptions.

While some statisticians might not consider such theory to be a true theory, it is important to keep in mind that the fundamental theory of most disciplines cannot be proven mathematically. For example, no one to date has mathematically *proven* that the Keynesian theory of economics is “correct” or even “better” than its main alternative, New Classical Economics (<https://www.econlib.org/library/Enc/KeynesianEconomics.html>). Of course, no one has *proven* that it isn’t correct either. Psychology, sociology, management science, and geology are a brief list of disciplines that have extensive bodies of research and underlying theory, but which rarely publish “theory-proof” articles in their journals.

Madigan and Stuetzle, in their discussion of Lindsay et al. (2004 , p. 409), made essentially this same point: “The issues we raise above have nothing to do with the old distinction between applied statistics and theoretical statistics. The traditional viewpoint equates statistical theory with mathematics and thence with intellectual depth and rigor, but this misrepresents the notion of theory. We agree with the viewpoint that David Cox expressed at the 2002 NSF Workshop on the Future of Statistics that ‘theory is primarily conceptual,’ rather than mathematical.”

The word “theory” itself must be properly understood in order to understand the points above, as well as the theory of statistical engineering. As with engineering and science, many definitions of the word theory are possible (e.g., <http://www.merriam-webster.com/dictionary/theory>). However, reasonable and accepted definitions typically state something similar to: “a coherent group of general propositions used to explain a phenomenon.” Obviously, there is no explicit requirement in such definitions for mathematics to be involved, although it often is. The underlying theory of physics, for example, involves considerable mathematics, but of course not all of the theory of physics is mathematical. If it were, physics would be considered a subfield of mathematics.

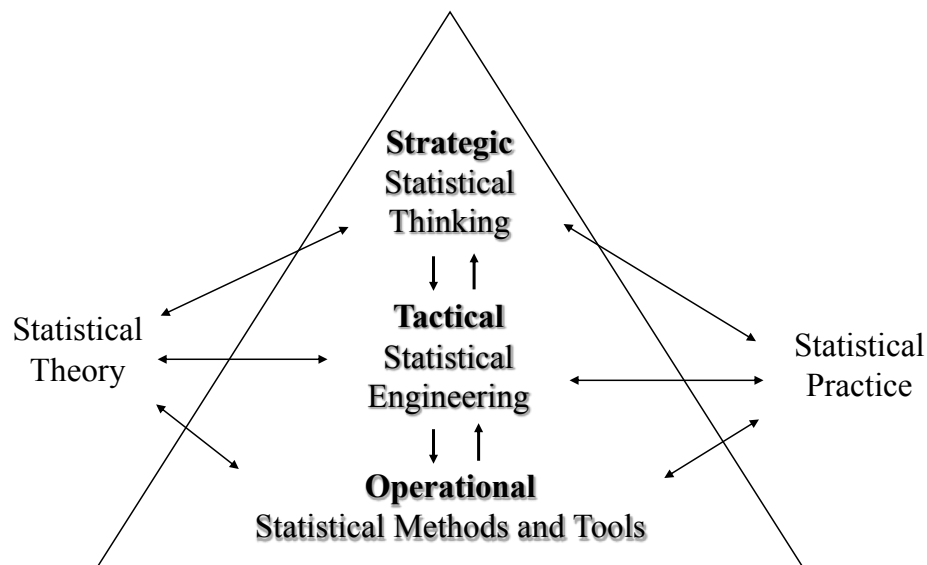
There is now, in fact, “a coherent group of general propositions used to explain” statistical engineering. These propositions are presented and explained below.

There are two other aspects of the underlying theory that we feel are also important: a conceptual model of the relationship between statistical engineering and the statistical methods, which we present first, and an overall model to guide application of statistical engineering to large, complex, unstructured problems, which we present third.

1.1.5.2 How Does Statistical Engineering Fit?

Figure 1 (Snee and Hoerl 2017) depicts the statistics discipline as a system, with strategic, tactical, and operational levels, each of which has both a theoretical and an applied aspect. The strategic-tactical-operational model is one that has been used in the military, business, government, and other organizations for a long time, perhaps millennia. The strategic level is where high-level decisions are made about the organization’s fundamental purpose, what it views as success, and how it will win in a competitive environment. This is where such things as vision, mission, values, and so on are determined. Per Meng (2009), statistical thinking is at the strategic level for the statistics discipline; that is, how we think about statistics itself, and its relationship with other disciplines. This includes how to interpret the world from a stochastic versus deterministic viewpoint, how we think about data and its relationship to subject matter theory in problem solving, and so on.

Figure 1.1 The Statistics Discipline as a System



The operational aspect of this type of model is where the “rubber hits the road,” that is, where the actual work of the organization is accomplished. In the military, it

would be the “front lines”, in manufacturing it would be the production floor, and in a hospital, where patients are seen and treated, to give just a few examples. In Figure 1, the methodologies of statistics, such as time series models, experimental design, statistical process control, and so on, would be at the operational level. In fact, when most people, both statisticians and non-statisticians, think about the statistics discipline, it is likely that they primarily think of this operational level – the tools themselves. Research on the tools over the decades has produced a rich and deep understanding of how and why these tools work, as well as invention of newer and more effective tools. When this theory is integrated with learning from actual applications, we refer to this combined body of knowledge as statistical science.

Note that Figure 1 illustrates a theoretical and applied component at all three levels. For example, at the operational level we both perform research on the theory of the individual methods, and also actually apply them to real problems. Similarly, we can debate the theory of statistical thinking; what should be the fundamental principles of the discipline? A stochastic view of the world would seem obvious as a core principle, but what about the proper relationship between statistics, data science, computer science, industrial engineering or operations research? There could no doubt be serious debates as to how the statistics profession should view its boundaries and proper relationships with these other disciplines.

Of course, these concepts are hopefully actually applied in practice, such as the Food and Drug Administration (FDA) insisting the clinical trials be based on randomized experiments, rather than solely on observational data. Fortunately, the FDA understands the qualitative distinction between observational data and data from randomized experiments.

The tactical level of the organization exists to develop tactics to carry out the strategy. In the business world, senior executives set strategy – where to place the “big bets” in new product development, which businesses or markets to get out of, which to get in to, and so on. However, the employees on the “front lines” in manufacturing, sales, logistics, and so on, are far removed, both physically and conceptually, from the executive office. Middle management therefore exists to take the strategic direction and figure out specific tactics within each function to ensure that the strategy actually succeeds. For success to occur, of course, the “front lines” need to take actions that are supportive of the overall strategy. In some ways, this tactical level of middle management has the toughest job, which is one reason that “middle management” has a negative connotation in many circles.

In the statistics discipline, we have found a serious gap between the higher-level principles of statistical thinking and utilization of the individual tools. That is, distinguished statisticians may opine on the proper way to think about the discipline and how it can succeed in expository articles, but such opining is far removed from the tools research being done in academia, or from the routine applications of practitioners. In essence, there is no “middle management” in the

statistics profession. In our view, the critical question of how researchers or practitioners should research and use statistical methods in such a way as to be consistent with the principles of statistical thinking has gone largely unanswered.

Wild and Pfannkuch (1999) identified this issue two decades ago, and provided some suggestions as to how to address it. We propose that statistical engineering can further fill this gap and serve as the tactical element of the discipline, linking the individual methods with the fundamental principles of statistical thinking. That is, statistical engineering, as we discuss below, is based on fundamental statistical thinking principles. It applies these principles to guide the linking and integration of individual tools to solve a real problem, typically one that is large, complex, and unstructured. Therefore, it is providing guidance on how to take the individual tools and utilize them in a manner consistent with the strategy. As one example, statistical engineering provides a specific “unified approach to statistical analysis and integrating multiple methods in an iterative manner”, one of the strategic principles mentioned in the ASA guidelines for undergraduate statistical education, discussed previously.

Again, while it is important to develop a theory of how to do this, it is equally important to actually apply this theory to real problems. Such application provides a feedback loop to the theory, noting what does and does not actually work in practice, when addressing real problems versus textbook problems.

1.1.5.3 A Coherent Group of General Propositions

The statistics profession has certainly learned and documented important principles over the decades concerning solution of large, complex, and unstructured problems. However, we do not feel that they have been effectively integrated into a formal framework. If integrated, however, they are in some sense a “theory,” that is, “a coherent group of general propositions used to explain a phenomenon.”

Most experienced practitioners learn these principles and pitfalls “on the job,” often through making their own mistakes. At this point, they might be considered principles of statistical practice, or applied statistics. However, we argue that such principles can be studied, documented, debated, and enhanced over time, as well as formally taught to students. Under these circumstances, they would be considered a theory. The logical expectation in most disciplines is that theory and practice should gradually converge over time; we believe that the same should be true of statistics.

In our view, the most critical propositions, or principles of statistical engineering applied to large, complex, unstructured problems can be loosely grouped into the five major categories listed in Table 1 (Hoerl and Snee 2017). The first principle emphasizes the need for developing an understanding of the problem context. With straightforward problems, little time needs to be invested in studying the background or context. If someone asks you what time it is, you don’t need to study the history of watchmaking to answer the question – just look at your watch or cell phone!

Table 1.1
Fundamental Principles of Statistical Engineering

1. Understanding of the problem **context**
2. Development of a problem solving **strategy**
3. Consideration of the data **pedigree**
4. Integration of sound **subject matter theory** (domain knowledge)
5. Utilization of **sequential approaches**

However, suppose a city wishes to address gang violence. One could no doubt come up with some “obvious” solutions, such as providing more police to patrol the streets, trying to infiltrate the gangs with informants, or even modifying the criminal justice system. However, with large, complex, unstructured problems such as these, “obvious” solutions rarely work well. Rather, to have a serious impact on gang violence the city would likely need to develop a deep understanding of the gangs themselves; why people join them in the first place, how they recruit and operate, their specific criminal activities, how the gangs relate to one another, their internal codes of conduct, and so on. An effective response is only likely to be identified after developing a deep understanding of these contextual issues. This same principle generally holds for large, complex, unstructured problems in business, engineering, and healthcare.

The second principle highlights the fact that serious thought needs to go into development of a problem-solving strategy, once the context is understood. With straightforward problems, the correct solution can often be found in a textbook; no overall strategy is needed. However, with complex problems, especially those that are unstructured, the plan of attack will often not be clear. In fact, there is rarely a single “correct” approach. Therefore, significant time and planning need to go into developing the overall approach to solve the problem. Hoerl et al. (2014) discussed these points in greater detail within the context of Big Data problems. Unfortunately, the word “strategy” rarely appears in the indices of statistics textbooks, providing another illustration of the difference between statistical engineering and statistical science.

Understanding of the data pedigree, point 3, is important in any data analysis. Too often analysts assume that the data are “perfect”, representing a random sample from the population of interest. Almost without exception, they don’t represent a random sample from the population of interest. Data always have some limitations, whether they involve biased or limited sampling, outliers, missing data, missing variables, the wrong timeframe of data collection, or just outright blunders, such as recording a “34” when the actual number was “43”. Murphy’s Law, which says that anything that can go wrong, will go wrong, certainly applies to data collection.

The pedigree documents how the data were collected, what specifically they represent, how samples were obtain and measured, and what, if any changes or deletions were made to the data over time (the “chain of custody”). Hoerl and Snee

(2018) provide more detail on the concept and use of data pedigree, and an elaboration of this topic is also given in the Data Acquisition chapter of this handbook.

The next two points emphasize that statistical engineering views statistical and other tools from the perspective of the scientific method. That is, statistical methods are viewed as enablers of the scientific method, not substitutes for it. While this point may seem obvious, we note that few statistics textbooks formally discuss the scientific method, or how statistics fits within it. In particular, few discuss the critical importance of subject matter knowledge in acquiring data, analyzing it statistically, and interpreting the analyses; this is the emphasis of the fourth principle.

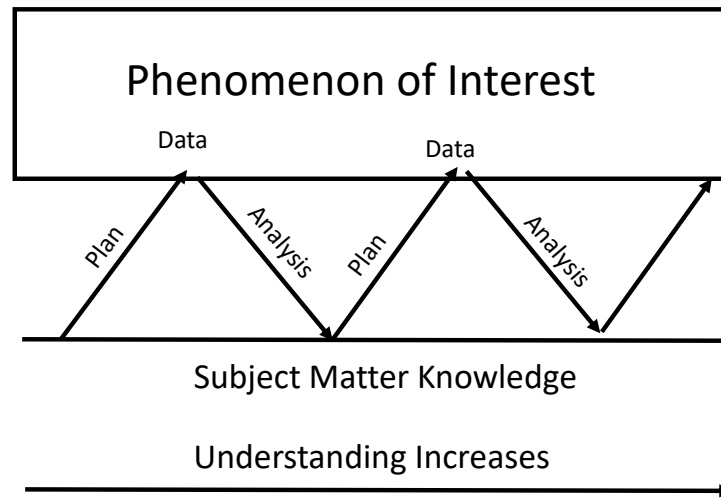
Subject matter (domain) knowledge is everything we know about the phenomenon under study, either from relevant theory, such as physics, epidemiology, or economics, or from previous data analyses. Such knowledge is needed from the very beginning of applications of statistical engineering, even in identifying the true problem, that is, the root cause, rather than just the symptoms. If scientists had all possible knowledge, they actually wouldn't need statistics or statistical engineering. Statistics is only needed because scientific knowledge is not complete, and empirical approaches – based on data collection and analysis – are often needed to “fill in the gaps” in our scientific knowledge. Eventually, after the data analyses are confirmed, they augment our previous scientific knowledge, enhancing our understanding. This process continuous through sequential cycles of the scientific method, eventually producing a mature discipline, such as physics or chemistry.

Sequential approaches are also core to the scientific method, and are the emphasis of the fifth principle in Table 1. Most applications in statistics textbooks tend to be “one shot studies”, where a data set is given, and the “correct” statistical method is applied, allowing the authors to move on to the next data set or next problem. The same is true of homework problems; for example, “what is the correct method to apply to this data”? Of course, real problems, particularly large, complex, unstructured problems, are not so simple. There is no single “correct” method, and in most cases multiple statistical methods and perhaps multiple disciplines are needed. In other words, a sequential approach is needed.

Each time practitioners determine the specific data needed, they do so based on their current understanding, that is, their current subject matter knowledge. They often have specific questions they need answered to “fill in the gaps”. Once they obtain the data and begin to analyze it, typically with multiple tools, they may answer some questions, but others may arise unexpectedly. For example, why is every fourth data point high? Therefore, additional rounds of data gathering and analysis are typically needed. Fortunately, with each round, they become a little more knowledgeable, and can ask better and more specific questions. That is, their understanding gradually increases through these sequential cycles of the scientific method, producing greater and greater understanding. This is illustrated in Figure 2,

based off a similar graph in Hoerl and Snee 2012, which is itself based off an earlier version from Box, Hunter, and Hunter (1978).

Figure 1.2 The Sequential Nature of Statistical Engineering



1.1.5.4 A Framework for Statistical Engineering Projects

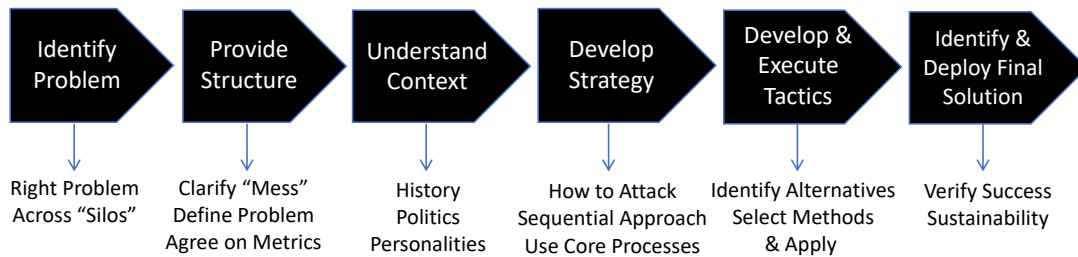
As previously noted, there is no “correct” solution to large, complex, unstructured problems. Similarly, there is no “cookbook” that will lead practitioners step by step through successful completion of all such projects. However, there is a framework to give some guidance as to how to think about approaching such problems. Figure 3, based on Hoerl and Snee (2017), shows the phases that statistical engineering projects typically go through. As an analogy, each child grows up to maturity along a different path; no two children, even “identical” twins, grow up exactly the same. However, the discipline of child development has documented the fact that virtually all children go through the same set of phases growing up, although uniquely. For example, “terrible twos”, “fantastic fours”, pre-teens, teenagers, and so on, are layman’s terms for these child development phases commonly used by parents.

So it is important to keep in mind that Figure 3 provides a general framework, not a “cookbook”. As previously noted in the discussion of fundamental principles, practitioners will generally need to develop a unique strategy for each problem, based its unique context. Therefore, while statistical engineering projects will generally go through each of these phases, they will do so in unique ways, just as children go through child development in unique ways.

It should also be noted that Figure 3 is similar in nature to other problem-solving frameworks, such as DiBenedetto et al. (2014), the Job Task Analysis (JTA) framework and “domains” from the Certified Analytics Professional (CAP) program (INFORMS 2018), and the Data Analytics Lifecycle (EMC Education Services 2015).

While each of these frameworks has distinctive aspects, with Figure 3 focusing on large, complex, unstructured problems, there is enough overlap to provide confidence that that each provides a reasonable approach.

Figure 1.3 The Phases of Statistical Engineering



The first phase in Figure 3 is to identify the problem. This might sound easy, and in some cases it is. However, as noted by D. K. J. Lin (2014, personal communication), "Finding a good problem is harder than finding a good solution." Also, large and complex problems typically cross organizational boundaries or "silos". Because it is usually easier to work "within" a silo than "across" silos, teams will often focus their problem-solving efforts on their silo, working on the symptoms of the larger problem that appear within their silo. If multiple teams work on the same large problem, but each focuses on the symptoms within their silos, the net result is often teams working at cross purposes, each trying to push the problem from their silo to another silo. The real problem, crossing multiple silos, may not even be recognized, much less addressed.

A classic example of this phenomenon occurs when businesses attempt to effectively manage their overall order fulfillment system, from sales to production planning to warehousing and inventory to logistics, ultimately delivering the product to customers in a timely fashion. Obviously, this overall order fulfillment system is a large, complex system. In most businesses, it is broken up into individual silos, representing each functional area involved, such as a sales team, a production planning team, a warehousing and inventory team, logistics or product delivery, and customer management, which focuses on "keeping the customer happy". Periodically, there will be a business drive to reduce inventory costs and working capital, putting pressure on the warehousing and inventory team to reduce the inventory levels as low as possible. At the same time, a team from logistics or customer management may be working on a project to provide more timely deliveries to customers, with no product outages.

It should be obvious that both teams are working on the same fundamental problem – order fulfillment, but both are only working within their own respective silos,

focusing on the symptoms they see; high inventory costs and late or incomplete customer deliveries, in this case. The net result is that one team is figuring out how to lower inventories, while the other is working on how to increase them. No one has identified the real problem they should all be working on; the large, complex, unstructured problem of optimizing the order fulfillment system, whatever that might mean when properly structured. The overall system could no doubt be improved, but this would require cooperation, and everyone having the same understanding of what the real problem is, and what success would look like. In other words, they would need to identify the *right* problem.

Once the right problem has been identified, it usually needs to be properly structured. As noted by X. Tort (2018, personal communication), what we typically see initially, unfortunately, is a “mess”. It is virtually impossible to solve a mess. Rather, we first need to convert the mess into a formal problem. Once we have a formal problem, we can move forward to solve it. The process of converting a mess into a problem is what we call providing structure.

In our order fulfillment example, we may initially see a mess in which we have too much finished product inventory (perhaps some is expiring before we can deliver it to customers), too much work-in-progress inventory, upset customers who don't know where their product is or why it was late or incomplete, manufacturing disruptions, dysfunctional teams that don't like each other and won't work across silos, recurring quality issues resulting in more work-in-progress and late shipments, and perhaps pressure from senior management demanding that the situation be “fixed” ASAP, but not providing any methodology to fix it.

As a next step, the organization would have to define the right problem properly, considering the overall order fulfillment system. Since it is typically impossible to minimize inventory while at the same time minimizing late customer deliveries, what exactly would success look like? How would it be measured? It should be clear that there is no obvious problem statement or a single, quantitative objective to be maximized. Considerable work may be required to convert this mess into a formal problem that can be attacked, and to obtain organization alignment across silos and with senior leadership.

The next phase is to understand the context of the problem, which is one of the fundamental principles of statistical engineering, and was discussed in the previous section. As noted there, large and complex problems have resisted solution for a reason; “obvious” solutions don't generally work. Only once the problem's root causes, the history of previous efforts – including why they failed, and the technical, political and social background of the problem are properly understood, can the team develop a viable approach to solution. This usually requires a lot of hard work, but is absolutely necessary for big problems.

Once the right problem has been properly identified, the mess has been converted to a formal problem statement, and the context of the problem is properly understood,

the team is in position to develop a strategy to address it. As noted above, a strategy is needed because multiple methods and perhaps multiple disciplines will be required, and all of these need to be integrated into an overall approach or game plan. In sports, one aspect of a head coach's responsibilities is to prepare a "game plan" for each opponent. The individual game plans may be quite different from each other, depending on the strengths and weaknesses of the opponents the team faces. This game plan is in essence a strategy that the coaching staff believes will maximize the chances of success. However, if the players are not all on the same page, and some are not following the game plan, failure is likely.

Similarly, a key role of project leadership is to develop a game plan, or strategy, to solve the problem, and then ensure that everyone on the team, even people from different silos, are all on the same page. This is easier said than done, because people from different silos and with different skills sets may have their own ideas about how the project should proceed. That is, they may not agree with the strategy, and start to go off on their own "closet projects". Such a splintering of the team rarely works well, just as it doesn't in sports. The statistical engineering strategy will typically involve application of a series of statistical and other methods, linked and integrated in a logical manner. Note that the strategy for solving a particular problem is obviously at a much lower level than the overall strategy for a business, university, or other organization, which we discussed previously. Both are examples of strategy, however.

Once a strategy has been developed and everyone is on the same page, the team needs to develop and employ tactics to actually carry out the project. A strategy, while critically important, is just a plan. To win on the sports field the team needs to actually block, tackle, pass, catch, and so on, in order to implement the strategy. Tactics are more detailed elements of the overall strategy that provide specific direction at the operational level. For example, suppose our strategy for order fulfillment includes a decision that for now we will prioritize customer fulfillment (minimizing late deliveries) over inventory reduction. We still need specific methods for fulfilling orders; a high-level plan is not sufficient. In the tactics phase, we figure out specific methods to fulfill orders more consistently, and then actually deploy these in operations to see how well they worked. The tactics will generally involve selection of individual statistical and other methods within each of the core processes discussed above.

Once the strategy and tactics are in place, the team can "take the field" and begin implementing them, i.e., solving the problem. For statistical engineering problems this will result in several statistical and non-statistical tools utilized in a sequential strategy. The results of the first analysis may change the ensuing tactics, just as when sports teams find themselves way behind at halftime they may "ditch the game plan" and start over, or perhaps make less dramatic halftime adjustments.

In the course of applying these methods in a systematic fashion, the team should begin to learn and identify specific actions they could take to address the problem.

In most cases, these actions will need to be piloted to verify that they actually work, and don't cause unforeseen issues. Gradually, a final solution is identified and deployed. If it doesn't work as well as anticipated, the team may need to reloop back to the strategy or tactics phases. Once a satisfactory solution is obtained, the team still needs to worry about sustainability. Therefore, a "control plan" is typically needed to embed the solution into standard work processes, as well as to identify how the system should be monitored over time, and what steps employees should take when backsliding is detected.

Even in a best-case scenario, there will be opportunity for further learning and improvement. Therefore, a new improvement initiative or project may make sense, to follow up on the first team's results. The cycle of improvement from the scientific method continues.

1.1.5.5 The Core Processes of Statistical Engineering

The methods needed within the statistical engineering strategy are often selected from five major categories, or "core processes", which represent the major "what's" of statistical science. That is, the core processes are not individual methods or tools, such as regression analysis or control charts, which could be considered "how's". They are called "processes" because they represent the major high-level activities performed in applications of statistics. Virtually all individual statistical methods fit conceptually into one of these processes. Of course, other non-statistical tools and competencies will be needed in the other phases of statistical engineering projects, as we explain shortly. In the typical order in which they are applied, the core processes are:

- Data Acquisition – proactively obtaining the highest quality data possible for the problem at hand, and documenting the data pedigree
- Data Exploration – understanding the data, observing patterns and trends, and beginning to develop or refine hypotheses, based on graphical and numerical methods
- Model Building – developing different types of formal models, depending on the data and problem being addressed
- Drawing Inferences (Learning) – considering what broader conclusions can be drawn about the phenomenon of interest beyond this particular data set
- Solution Identification and Deployment – determining the best course of action to take based on what has been learned from the previous processes, deploying it, and ensuring sustainability

Note that each of these high-level processes begins with a verb – they represent some action, rather than a specific tool. Of course, there are many tools to be considered for use within each process. The mix of tools will typically vary for each problem. There is also a set of overarching competencies that is generally needed to achieve success. These competencies are needed not only in the strategy and tactics phases, but rather across all phases of statistical engineering applications. These

overarching competencies include project management, teamwork, communication, and other competencies discussed in the chapter on overarching competencies.

1.1.6 Summary of Key Points

Key points that we would like to emphasize from this introductory section include:

- Statistical engineering is not a “buzzword”. It has been carefully defined to represent the engineering of solutions to statistically oriented problems.
- Large, complex, unstructured problems are particularly amenable to a statistical engineering approach.
- Statistical engineering emphasizes *integration*; that is, integration of methods and integration of disciplines.
- There is an underlying theory to statistical engineering that is admittedly a work in progress.
- Part of this theory is a set of generic phases through which most applications of statistical engineering progress. This framework provides general guidance to those applying statistical engineering.
- The overall strategy and tactics utilized in applications will typically involve a series of methods selected from statistical *core processes*, linked with other methods. The specific methods selected will depend on the unique aspects of the problem at hand.

References:

American Statistical Association Undergraduate Guidelines Workgroup (2014) “2014 Curriculum Guidelines for Undergraduate Programs in Statistical Science,” Alexandria, VA: American Statistical Association. Available at <http://www.amstat.org/education/curriculumguidelines.cfm>

Antony, J., Hoerl, R.W., and Snee, R.D. (2017) “Lean Six Sigma: Yesterday, Today, and Tomorrow”, The International Journal of Quality & Reliability Management, 34,7, 1073-1093.

Box, G. E. P., and Wilson, K. B. (1951) “On the Experimental Attainment of Optimal Conditions,” Journal of the Royal Statistical Society, Series B, 13, 1–45.

Box, G.E.P., Hunter, W.G., and Hunter, J.S. (1978) Statistics for Experimenters, John Wiley & Sons, Hoboken, NJ.

Di Benedetto, A., Hoerl, R.W., and Snee, R.D. (2014) “Solving Jigsaw Puzzles: Addressing Large, Complex, Unstructured Problems”, Quality Progress, June, 50-53.

Donoho, D. (2017) “50 Years of Data Science”, Journal of Computational and Graphical Statistics, 26:4, 745-766.

EMC Educational Services (2015) Data Science and Big Data Analytics, John Wiley & Sons, Hoboken, NJ.

Fung, K. (2013) "The Pending Marriage of Big Data and Statistics", Significance, 22-25.

Hardin, J., Hoerl, R., Horton, N.J., Nolan, D., Baumer, B., Hall-Holt, O., Murrell, P., Peng, R., Roback, D., Temple Land, D., and Ward, M.D. (2015) "Data Science in Statistics Curricula: Preparing Students to 'Think With Data'", The American Statistician, 69, 4, 343-353.

Hockfield, S. J. (2010) "Technical Challenges of the 21st Century," Niskayuna, NY: Presentation at GE Global Research.

Hoerl, R.W., and Snee, R.D. (2012), Statistical Thinking: Improving Business Performance, 2nd ed., John Wiley & Sons, Hoboken, NJ.

Hoerl, R.W., and Snee, R.D. (2017) "Statistical Engineering: An Idea Whose Time Has Come?", The American Statistician, 71, 3, 209-219.

Hoerl, R.W., and Snee, R.D. (2018) "Show Me the Pedigree!", accepted for publication in Quality Progress.

Hoerl, R. W., Snee, R. D., and De Veaux, R. D. (2014) "Applying Statistical Thinking to 'Big Data' Problems," Wiley Interdisciplinary Reviews: Computational Statistics, July/August, 221-232.

INFORMS (2018) Certified Analytics Professional Program and Examination: Candidate Handbook, 6th ed., Catonsville, MD.

Kendall, J., and Fulenwider, D.O. (2000) "Six Sigma, E-Commerce Pose New Challenges", Quality Progress, July, 31-37.

Lindsay, B. G., Kettenring, J., and Siegmund, D.O. (2004) "A Report on the Future of Statistics" (with discussion), Statistical Science, 19, 387-413.

Meng, X. (2009) "Desired and Feared-What Do We Do Now and Over the Next 50 Years?," The American Statistician, 63, 202-210.

Nair, V. (2008), "Industrial Statistics: The Gap Between Research and Practice," Youden Memorial Address, ASQ Statistics Division Newsletter, 27, 5-7.

Wild, C., and Pfannkuch, M. (1999) "Statistical Thinking in Empirical Enquiry," International Statistical Review, 67, 223-248.