CEAB's Graduate Attribute 3.1.3: Investigations – Is it really possible for engineering graduates to possess this attribute?

Leonard Lye, PEng, PhD, FCSCE, FEC
Associate Dean (Graduate Studies)
Faculty of Engineering and Applied Science
Memorial University of Newfoundland
llye@mun.ca

Abstract

Starting 2014, engineering programs in Canada will be evaluated by CEAB based on twelve Graduate Attributes and institutions must demonstrate that their graduates possess these attributes at the time of graduation. One of these attributes is "Investigations" which is defined by CEAB as "an ability to conduct investigations of complex problems by methods that include appropriate experiments, analysis and interpretation of data, and synthesis of information in order to reach valid conclusions." This is similar to one of ABET's student outcomes which states that students attain "an ability to design and conduct experiments, as well as to analyze and interpret data". In this paper, it will be argued that with the current curriculum of most, if not all, engineering schools in Canada, it is almost impossible for graduates to possess this attribute unless a compulsory course is introduced to specifically teach the subject. Proper design, conduct, and analysis of experiments of complex problems cannot be learned by osmosis or by doing standard labs where the procedure and analysis methods are given. Engineering educators and graduates thinking that somehow that the skill to design, conduct and analyze experiments will be learned in an engineering program do not fully appreciate the myriad of issues that are involved with experimentation to study a complex problem. Examples of these issues include: a large number of variables, multiple responses with conflicting objectives, linear or nonlinear responses, interaction among variables, etc. In this paper, these issues and many others will be discussed. How they can be addressed will be discussed and a course that will help graduates achieve competence in "Investigations" is also proposed.

1 Introduction

In a short two years from now, engineering programs in Canada will be evaluated by CEAB using an outcome based approach [9]. Institutions must demonstrate that their graduates possess 12 specific graduate attributes at the time of graduation. These attributes are very similar to the list of 11 student outcomes as specified in [1]. One of the graduate attributes (3.1.3) is "Investigations" which is defined as "an ability to conduct investigations of complex problems by methods that include appropriate experiments, analysis and interpretation of data, and synthesis of information in order to reach valid conclusions." This is similar to one of ABET's student outcomes which states that students attain "an ability to design and conduct experiments, as well as to analyze and interpret data". In this paper, it will be argued that with the current curriculum of most if not all engineering schools in Canada, it is almost impossible for graduates to possess that stated attribute unless a compulsory course with appropriate labs and projects is introduced to specifically teach the proper methodologies for the design and analysis of experiments. Proper design, conduct, and analysis of experiments of complex problems cannot be learned by doing standard lab exercises where the objective, procedure and analysis methods have been given by the instructor. All the student has to do is to follow instructions. Engineering educators and graduates thinking that somehow that the skill to design, conduct and analyze experiments will somehow be learned in an engineering program, probably do not appreciate fully the myriad of issues that are involved with experimentation to study a complex problem. It is also likely that they do not know that there exists a whole body of knowledge on the science of experimentation that has been in existence for about 100 years. In this paper, some of the myriad of issues that arise with experimentation of a typical complex problem will be discussed. How they are properly handled will also be briefly discussed. It will be seen that these issues and techniques will not be obvious to those untrained in proper experimental design techniques. According to R.A. Fisher (about data obtained from bad experimental design) - "nothing much can be gained from statistical analysis; about all you can do is to

carry a postmortem and decide what such an experiment died of" [8].

In the next section, a case will be made about why CEAB's criteria 3.1.3 cannot be met based on the current engineering curriculum. This will be followed in Section 3 by a discussion of the some of the myriad of issues that arise in experimentation of a typical complex problem. How these issues are resolved with proper techniques will also be briefly discussed. In Section 4, a course outline that will help graduates achieve competence in the attribute of "Investigations" is proposed. Concluding remarks are given in Section 5 followed by the acknowledgement and references.

2. Current engineering curriculum

In a typical undergraduate engineering program, students normally take courses in chemistry, physics, electric circuits, materials, mechanics, fluids, and others in which there is a laboratory component that involves hands-on experimental work. The students are usually given a lab manual with quite detailed instructions on exactly what to do. The objectives are usually well-defined, what to measure and how to analyse and graph the results are also usually well established. As part of the lab component, students may be instructed on the basic concepts of the various types of experimental errors and basic uncertainty analysis such as propagation of errors [10]. Basic dimensional analysis is also sometimes taught and if students have already taken a basic statistics course, students may also be asked to calculate means and standard deviations, calculate confidence intervals, and perform simple regression analysis of their experimental data. It would be quite rare to expect students to perform multiple regression or ANOVA. Students will however learn how to use various tools and techniques in the lab, such as use of an oscilloscope, load cells, strain gauges, and various digital data acquisition systems. More details are given in [21].

However, CEAB criteria 3.1.3 "Investigations" states that a graduate must possess an ability to conduct investigations of complex problems by methods that include appropriate experiments, analysis, and interpretation of data, and synthesis of information in order to reach valid conclusions. Surely what is done in the standard undergraduate labs and basic statistics course do not address the CEAB requirement. The labs only deal with simple well-defined problems with usually one variable (X) to change or adjust and one response (Y) to measure. The lab problems can hardly be classified as a complex problem. Hence the data analysis techniques are also very elementary and

definitely insufficient to tackle problems of a more complex nature. There is also no opportunity for students to actually design an experiment on their own and they definitely do not know what issues are involved when designing and analysing an experiment with more complexity.

3. Some issues with complex problems

In this section, some characteristics of complex problems that require experimentation and possible issues to be dealt with will be discussed. Below is a sample list of issues that may be encountered in practice when dealing with typical engineering problems with some complexity.

1. Budget and time are always limited. Each experimental run may be expensive and/or time consuming.
2. Large number of factors or variables, some may be discrete, continuous or categorical. Some of the factors may or may not be controllable by the experimenter.
3. Some factors may interact with other factors although they may not be important on their own.
4. Several responses that may or may not be related. Responses may be discrete, continuous, time dependent, or even categorical.
5. Multiple objectives involved. Some responses may be more important than others, and some may be conflicting. One response may need to be maximized while another may need to be minimized or may need to be between certain limits.
6. Responses may be linear or nonlinear functions of the various factors.
7. Functional relationships between the responses and the factors may be desired for prediction and for design optimization purposes.
8. Some factors may be hard to change and hence minimizing such changes would be more cost effective.
9. Some combinations of factors may not be feasible or even dangerous.
10. Measurement of responses may require high precision and/or a complex set up.
11. Firm grasp of theory may be important or there is no known theory to guide the experiment.
12. Responses may be dependent on the proportions of individual components that sum to a given value.
13. Extraneous factors may affect responses and lead to misleading results if not taken into account.
14. Experiments may involve people or animals.
15. There may be missing or bad data from the experiments.

2

From the above list, it is clear that the numerous labs that engineering graduates have done while going through their program do not address any of the issues listed. It is more likely that they are not even aware of the issues involved. This is also true of the engineering educators. It is likely that engineering educators and students are suffering from the so-called Kruger-Dunning effect. This is an interesting human behavior where those who are unskilled in a domain are unaware that they are actually unskilled in that domain. They also have the tendency to overestimate their ability and performance relative to objective criteria [14]. The use of design of experiment techniques is also a key component of the Six-Sigma quality improvement process that is widely used by major corporations and institutions worldwide. It has been almost a 100 years since the invention of proper experimental design by Sir Ronald A. Fisher but it has yet to permeate the Canadian undergraduate engineering curriculum.

3.1 Design and analysis of experiments

An experiment has been formally defined in [17] as: A test or a series of tests in which purposeful changes are made to the input variables or factors of a system so that we may observe and identify the reasons for changes in the output response(s). Hence every experiment involves a sequence of activities. It starts with a conjecture or a hypothesis that motivates the experiment. Then an experiment is designed and conducted to investigate the conjecture. The data from the experiment are then analyzed. Finally, the conclusions reached as a result of the experiment. What has been learned about the original conjecture from the experiment? Do we need any follow up experiments? A properly executed experiment requires careful planning so that appropriate data can be analyzed by statistical methods that results in valid, objective, and meaningful conclusions. Hence it involves two closely linked aspects: design of the experiment, and statistical analysis. Experiments are conducted in practically of all fields of engineering. It is used to compare basic design configurations, evaluate different materials and methods, select design parameters so that the design will work well under a wide variety of field conditions, and to determine the key design parameters that impact performance. Experiments may be carried out in a laboratory, in the field or on a computer using a sophisticated computer model. Statistically designed and analysis of experiments or DOE is a well-established field and has been in existence for almost a 100 years. Currently, it is widely used in the agricultural, chemical and process, manufacturing, business, and more currently as a key component of the Six-Sigma quality

improvement process. All experiments should be designed experiments. Unfortunately, most experiments are poorly designed wasting valuable resources and resulting in inconclusive or even wrong conclusions. Statistically designed experiments permit efficiency and economy, and the use of statistical methods in examining the data result in scientific objectivity when drawing conclusions. In general, by using DOE, experimenters can learn about the process being investigated, screen for important variables, build a mathematical model that predicts how input variables interact to create responses, and optimize the responses, if needed. Statistical significance is tested using Analysis of Variance (ANOVA) and the prediction equation is obtained using multiple regression analysis. Hence DOE is a methodology for systematically applying statistics to experimentation. At a basic level, the use of statistics is not always necessary. In fact, the basic techniques of DOE are easy enough to be taught to junior high students. All that is needed is basic arithmetic skills such as addition, subtraction, division, and multiplication.

Consider a rather seemingly simple (but actually quite complex) household experiment that can be conducted by anyone – determining the best tasting micro-waved popcorn with the least wastage. Taste and wastage may depend on the type of microwave oven and the wattage. Hence the only way to resolve the issue is to do an experiment. There are three factors to consider: two brands of popcorn (cheap vs. expensive), two possible power levels (75% and 100%), and two possible time (3 and 5 minutes). There are two responses of interest: taste, and amount of unpopped corns. This experiment looks simple at first glance because it is an experiment that is quite easy to conduct. Basically, put a bag of popcorn in the microwave oven at a certain power level and for a certain length of time, then check the taste and maybe count the number of unpopped corns. Try a different brand, power level or time, and check taste again. Do this several times and eventually, an optimal combination of power and time may emerge. This strategy of guessing and checking or trial and error is one obvious approach that one can use. A more systematic approach is to adjust one variable or one factor at a time, the so-called OFAT approach. This approach would pick a brand, fix the power level say at 75%, and then check the taste and unpopped corns at 3 minutes. Then, the brand and power level are kept at the previous levels, the time is now changed to 5 minutes. The brand and time are now held constant and the power level is now changed. This continues systematically changing one factor at a time. The OFAT method of experimentation is very prevalent and it has been outdated since the 1920s with R. A.

Fisher's invention of modern methods of experimental design. The method has been shown to be inefficient, and can be even disastrous. It cannot determine interaction effects and does not provide a means to test for statistical significance and obtain prediction equations. More details on the dangers of OFAT are given in [12], [3], and [17]. The correct, modern, and most efficient approach to conducting a multi-factor experiment such as the popcorn experiment is by a factorial approach. In this approach, the factors are varied together in a systematic way. The number of experimental runs is known in advanced, the effect of each factor and their interaction can be studied, and prediction equations can be obtained. For example, for the popcorn experiment, since there are 2 levels for each factor and there are 3 factors, hence $2^3 = 8$ experiments are needed. With 8 experiments, the effect of brand type, power level, time in oven, and their interaction effects can be determined. The 8 experiments should also be run in random order. In fact, 7 effects can be determined with the 8 runs. As a comparison, using OFAT, one would require 16 runs for the same efficiency as the factorial design and only 3 main effects can be estimated. Interaction or joint effects cannot be detected by the OFAT method of experimentation. Interactions occur when the effect of one factor depends on the level of one or more other factors. Sometimes these are far more important than the individual effects. The choice of factors and number of levels in this popcorn experiment is rather straightforward. In a more complex experiment this choice can be critical to the success or failure of the experiment and must be carefully considered [13].

What happens if we have another almost identical microwave oven available so that the experiment can be run in half the time? How would one divide up the 8 runs for the two microwave ovens? The effect of the two types of microwave oven is not of interest but the conclusions should not be affected by them. An obvious way to divide up the runs is to do it randomly. Yet, random assignment could be the worst strategy if there is a systematic bias between the two microwave ovens. To do this properly so that any difference between the ovens cancels out requires the use of a DOE technique known as blocking. This involves confounding an effect with the block. This technique would not be obvious to the uninitiated. Dividing a large number of experimental runs into smaller units or blocks requires careful consideration using proper DOE techniques.

The popcorn experiment could easily include a few more factors to consider such to elevate or not to elevate the bag, pre-heat or not to pre-heat the oven, etc. With say 5 factors each at two levels would require a total of $2^5=32$ runs which may be time consuming and expensive. But if one is familiar with DOE techniques, the 5 factors can be studied with very little loss of information with half the number of runs. This technique is called fractional factorial design or 2^{k-p} design. How to choose the 16 runs is also quite technical and again not obvious to the uninitiated.

In some experiments, there may be factors that are hard to change. For example, if instead of a microwave oven, the experiments are carried out in a conventional oven and temperature setting is one of the factors. Clearly it would be very time consuming and inefficient to run the experiment in random order because it would take a long time for the temperature in the oven to stabilize to the desired value. Hence it would be better if fewer changes are made to the temperature. That is, randomization has been restricted in this case. Some factors are changed more than others. If the analysis is carried out assuming a randomized design, the results may be misleading or wrong. When there are hard to change factors and easy to change factors in an experiment, the proper way to analyse the data as a split-plot design with two error structure. Again, this is not trivial and it would be practically impossible for students without training to know this.

By using a two-level factorial design, it is assumed that there is a linear response as the factor changes from one level to the next. To model non-linearity would require the addition of at least another factor level by augmenting the two-level design or a new design. To fit a second-order model would need at least three levels. The fitting of such models are in the realm of response surface methodology or RSM. Well-known designs include the central composite and Box-Benhken designs (CCD and BBD) and there are entire text books such as [18] dealing with this topic. So it would not be something that students can just pick up during their spare time or after they graduate.

There are still many more DOE techniques available but in the interest of space will not be discussed further here. One example is a technique to find the combination of factors to maximize taste but minimize wastage.

The above examples and some of the issues involved are just scratching the surface of the science of experimentation. It is not something that students can learn by osmosis or by doing some standard labs. It needs to be formally taught in a course with hands-on activities. A proposed course outline given in the next

section would hopefully help to bridge the current knowledge gap in DOE methodology.

4. Proposed Course Outline

Proper design and analysis of experiments cannot be done without some basic statistical knowledge. Hence a DOE course is best taught after the first course in statistics. A good time to have the DOE course would be in the third year. In fact the DOE course would cover 7 of the 12 CEAB Graduate Attributes: 1) knowledge base for engineering, 2) problem analysis, 3) investigation, 4) design, 5) use of engineering tools, 6) individual and teamwork, and 7) communication skills. The techniques learned can then be used in senior level labs, design courses, and capstone projects. The course outline is as follows:

1. Introduction to experimental design principles: strategies for experimentation; guide to good experimental design; concepts of randomization, replication, and blocking. [3 lectures]
2. Simple comparative experiments: review of basic statistics for DOE; review of the t-test; confidence intervals; F-test and single factor ANOVA; checking assumptions of ANOVA; dealing with more than 1 factor. [3 lectures]
3. Two-level factorial (2^k) design: basic concepts; interaction of factors; factorial vs. OFAT designs; developing prediction equations. [3 lectures]
4. Statistical details of 2^k designs: Use of software for DOE; ANOVA for k factors; multiple regression, assumptions and diagnostics; response transformation; checking for curvature. [3 lectures]
5. Experiments in smaller blocks: concept of blocking; confounding and partial confounding; replicated and unreplicated cases. [3 lectures]
6. Fractional factorial designs: 2-level fractional factorial (2^{k-p}) designs; defining contrasts and concept of aliasing; design resolution. [3 lectures]
7. Minimum-run designs: Augmenting a design; foldover and semi-foldover; computer generated minimum-run designs. [3 lectures]
8. Hard to change factors: recognizing and analyzing a split-plot design; few and many hard to change factors; use of software for analyzing split plot designs. [3 lectures]
9. Response surface methodology (RSM): adding a 3rd level for 2nd order modeling; central composite design (CCD) and Box-Benhken design (BBD); concept of rotatability; multiple response optimization. [3 lectures]
10. Course review and brief overview of more advanced topics (optional): general factorials; design and analysis of computer experiments; space-filling designs; optimal designs; mixture designs. [3 lectures]

In addition to the lectures, a set of five assignments should be sufficient to cover all the topics covered. It is also crucial that students actually design, conduct, analyze, and report on several hands-on multi-factor experiments. These would be done as 3-hour labs. These labs should be done as a team of four or five. There are many well documented experiments that can be used for these labs and 20 such experiments have been compiled by Anderson [2]. The course should also culminate in a final "capstone" type project where the team has to come up with a multi-factor experiment that they have to design, conduct, analyze, and report on. This project could be technical or non-technical in nature. Some useful toys and software for teaching DOE are given in [15].

As to the choice of text books, several are given in the list of references. Perhaps the least expensive and most basic are the set by Anderson and Whitcomb [3] and [4] who are principals and developer of the Design-Expert software, one of the best DOE software in existence today. Another good practical book is [5]. Good choices for more advanced study and reference are [6] – [8], [11], [16] - [20].

5. Conclusion

The author has taught a DOE course at the graduate level to engineering students from all disciplines for the last 15 years. For the graduate level course, the optional topics listed above and some additional topics are covered in more detail. The students also have to do two journal paper reviews and do an individual project written in a journal paper format. It is a core engineering graduate course at Memorial University. Many of the graduate students who have taken the course do not even have a previous course in statistics and still managed to attain the target competence thresholds. Many of their final projects have resulted in journal or conference publications. Like typical engineering graduates they started the course not knowing anything of proper design of experiments principles. Afterwards they apply it in their thesis and design work, and have helped changed the mindset of their supervisors, employers, and other engineers.

For the undergraduate engineering program, the situation is more critical. It is one of the 12 attributes that graduates should possess by 2014 and the time available to address this issue is limited. It may be necessary to have this course replace an existing course in the engineering curriculum. Deciding which course to replace is obviously going to be an issue in

an already overloaded curriculum. However, as pointed out by Box [8], "Take almost anything out. Surely an engineer who doesn't know how to run an efficient experiment is not an engineer".

6. Acknowledgement

7. References

[1] ABET (2011): Accreditation Criteria for Engineering Programs, Engineering Accreditation Commission. www.abet.org. Accessed on December 8, 2011.

[2] Anderson, M. (2011): DOE It Yourself, available at www.statease.com/pubs/doe-self.pdf.

[3] Anderson, M. and Whitcomb, P. J. (2007): DOE Simplified, 2nd Edition, Productivity Press.

[4] Anderson, M. and Whitcomb, P. J. (2005): RSM Simplified, Productivity Press.

[5] Anthony, J. (2006): Design of Experiments for Engineers and Scientists, Butterworth-Heinemann.

[6] Berger, P. D. and Maurer, R. E. (2002): Experimental Design, Duxbury, Wadsworth Group.

[7] Box, G., Hunter, J. S. and Hunter, W. G. (2005): Statistics for Experimenters, 2nd Edition, John Wiley and Sons.

[8] Box, G. and Friends (2006): Improving Almost Anything- Ideas and Essays, Revised Edition, Wiley.

[9] CEAB (2010): Accreditation Criteria and Procedures. Canadian Engineering Accreditation Board, Engineers Canada.

[10] Coleman, H. W. and Steele, W. G. (1999): Experimentation and Uncertainty Analysis for Engineers, 2nd Edition, John Wiley and Sons.

[11] Cornell, J. (2002): Experiments with Mixtures, John Wiley and Sons.

[12] Czitrom, V. (1999): One-Factor-At-A-Time Versus Designed Experiments, The American Statistician, Vol. 53, No. 2, 126-131.

[13] Czitrom, V (2003): Guidelines for Selecting Factors and Factor Levels for an Industrial Designed Experiment, Handbook of Statistics, Vol. 22, R. Khattree and C.R. Rao (Editors), Elsevier Science B.V.

[14] Kruger, J., and Dunning, D. (1999): Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. Journal of Personality and Social Psychology, Vol. 77, No. 6, 1121-1134.

[15] Lye, L. M. (2009): Teaching Design of Experiment Methodology Using Toys and Software Tools, Proceedings CDEN/C^{2}E^2 Conference, Hamilton, Canada.

[16] Mason, R. L., Gunst R. F., and Hess, J. L. (2003): Statistical Design and Analysis of Experiments, John Wiley and Sons.

[17] Montgomery, D. G. (2008): Design and Analysis of Experiments, 7th Edition, John Wiley and Sons.

[18] Myers, R. H., Montgomery, D. G. and Anderson-Cook, C. M. (2009): Response Surface Methodology, John Wiley and Sons.

[19] Ryan, T. P. (2007): Modern Experimental Design, John Wiley and Sons.

[20] Schmidt, S. R. and Launsby, R. G. (2005): Understanding Industrial Designed Experiments, 4th Edition, Air Academy Press.

[21] Wheeler, A. J. and Ganji, A. R. (2010): Introduction to Engineering Experimentation, 3rd Edition, Prentice-Hall.