



**BASIC QUALITY**

# Design of Experiments For Dummies

by **Willy Vandenbrande**

**D**espite all the efforts by specialists in quality and statistics, design of experiments (DoE) is still not applied as widely as it could and should be. When I talked to people who had some kind of introduction to DoE, I noticed they were reluctant or even

scared to apply it. To them, DoE deals with complex tests with many factors at various levels and requires in-depth knowledge of at least two statistics textbooks. In addition, the continuous fight between Genichi Taguchi adepts, Dorian Shainin enthusiasts and classicists does not make it easier for the interested newcomer.<sup>1</sup>

More interesting to me are people's comments about DoE's complexity. I dare say 90% of all tests performed in industry are one-factor two-level experiments, also known as 2<sup>1</sup> designs. We usually just want to know how the system, product or process will react if one factor is changed from one level to another level. In most cases, the first level is the original, currently used level of the factor, so we already have some knowledge about the behavior of the system at this setting.

Simple, you say? Do it all the time? Undoubtedly, but how often do you do it correctly? I know from experience an industrial experiment does not need to be complex to get messed up. There are two ways experiments get messed up: when they don't lead to any conclusions and when they lead to incorrect conclusions.

## In 50 Words Or Less

- Design of experiments is not as difficult as you may think.
- In most cases, you just want to know how the system, product or process will react if one factor is changed.
- Follow nine rules to ensure your conclusions are correct.



## Where Do We Go Wrong?

If you divide the experimentation process into four phases—set up the experiment, execute the tests, analyze the results and draw conclusions—you’ll see things can go wrong in all phases, though the most damage is often done before the tests start.

An incorrectly set up experiment cannot be saved, not even by the most advanced statistical software programs. You need to think before you start an experiment and use the basic rules of DoE to avoid problems. These experiments are so easy, the calculations and analyses will be effortless. You just need to be sure the results contain the answers to your questions.

To illustrate the many ways to mess up an experiment, I will use a simple example called the hunt for the red tomato starring Sam, a friendly, amateur gardener, who is a keen tomato grower. His local supplier just told him about a great new fertilizer called “red tomato” that is 20% more expensive than Sam’s current “tomato lover” but is guaranteed to give him a lot more tomatoes by the end of the season. The offer is tempting, but Sam wisely decides to test the product before changing over his entire crop. He knows a few things about experimentation and plans to follow nine basic rules of DoE.

### Rule One

**Write down the question(s) you want the experiment to answer.**

In Six Sigma terminology: Define the problem, remembering one aspect of problem definition is quantification. If you want to compare results, you’ll need to know on what basis the results will be compared, so you’ll need to quantify the results. In this case, Sam wants to know, does red tomato fertilizer increase my tomato harvest by at least 20% in weight?

### Rule Two

**Characteristics that are not part of the study also need to fulfill requirements.**

It’s not beneficial to solve one problem or improve one aspect and, at the same time, create three new problems or deteriorate several other characteristics. Sam would not be happy if, as a result of changing fertilizer, he

had 20% more tomatoes, but they all tasted bad.

Besides evaluating the taste, he might also look at the size of the tomatoes, their size distribution or their color. Although he wouldn’t perform the experiment to improve these characteristics, he would need to ensure they stay at an acceptable level, so he would measure and evaluate them at the end of the experiment.

### Rule Three

**Make sure you have a reliable measurement system.**

Be aware of the importance of the variation introduced by the measurement system and keep it to a minimum. Weighing systems are generally pretty good, but if Sam wants to be sure, he should analyze his measurement system to ensure it’s adequate.<sup>2</sup>

**TABLE 1** Two Types of Errors

| Decision taken based on the experiment   | Real (but unknown) situation  |  |
|--|---|--|
|  | There is a difference   | There is no difference   |
| There is a difference between the means  | OK  | Type I error ( $\alpha$ - risk)<br>Significance level = $1 - \alpha$ |
| There is no difference between the means | Type II error ( $\beta$ - risk)<br>Significance level = $1 - \beta$ | OK   |

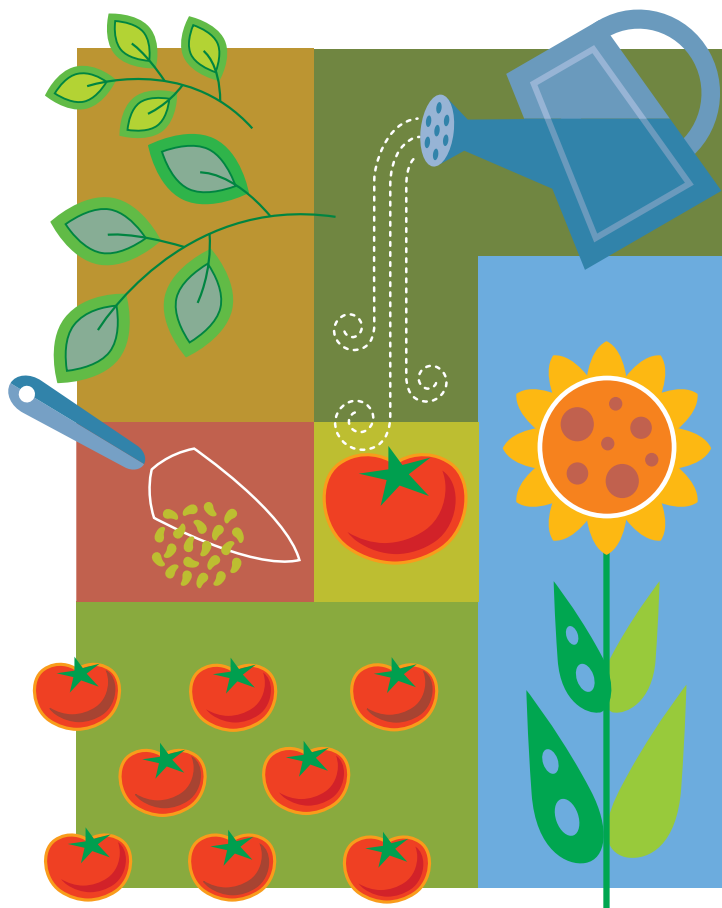
**TABLE 2** Sample Size Selection

**Assumptions:**  $\alpha$  - risk = 0.05 (significance level = 95%).  
 $\beta$  - risk = maximum 0.10 (power = minimum 90%).  
 Based on t-test for two means with equal variance.

| Minimum difference (In standard deviation units) | One-sided comparison (Is the new setting better?) | Two-sided comparison (Is there a difference between the old and new setting?) |
|--|---|---|
| 0.5  | 70  | 86  |
| 1.0  | 18  | 23  |
| 1.5  | 9   | 11  |
| 2.0  | 6   | 7   |
| 2.5  | 4   | 5   |
| 3.0  | 3   | 4   |

**TABLE 3** Results of Last Year's Harvest

| Tomato plant              | Yield (kg) |
|---------------------------|------------|
| 1                         | 4.94       |
| 2                         | 4.11       |
| 3                         | 4.64       |
| 4                         | 4.43       |
| 5                         | 4.86       |
| 6                         | 5.02       |
| 7                         | 4.68       |
| 8                         | 4.85       |
| 9                         | 4.51       |
| 10                        | 5.14       |
| 11                        | 4.80       |
| 12                        | 4.65       |
| 13                        | 4.57       |
| 14                        | 4.88       |
| 15                        | 4.26       |
| 16                        | 4.47       |
| 17                        | 4.71       |
| 18                        | 4.84       |
| 19                        | 5.51       |
| 20                        | 5.09       |
| 21                        | 4.33       |
| 22                        | 4.85       |
| 23                        | 4.63       |
| 24                        | 4.53       |
| 25                        | 5.52       |
| 26                        | 4.85       |
| 27                        | 4.82       |
| 28                        | 3.78       |
| 29                        | 5.37       |
| 30                        | 5.58       |
| <b>Average</b>            | 4.77       |
| <b>Standard deviation</b> | 0.41       |



### Rule Four

**Use statistics and statistical principles up front.**

Before starting a test, you must determine the appropriate sample size because you will use the sample result to draw conclusions for the behavior of the population. Table 1 shows the two types of incorrect conclusions and the associated risks you may encounter when comparing two means.

To determine the appropriate sample size, you need to determine the  $\alpha$  – and  $\beta$  – risk [or their counterparts: significance level ( $1 - \alpha$ ) and power ( $1 - \beta$ ) of the test] you are willing to take, the distribution and standard deviation of the population and the minimum difference you want to be able to detect between the two populations. This will require you to make several assumptions.

Table 2 can be used as a guideline for determining sample sizes if you want to compare the means of normally distributed populations with equal variance. You can use the historical variance of the current process as an approximation. The table is based on a two-sample t-test and contains the values for one-sided and two-sided comparisons. Despite its limiting assumptions, it is useful in many cases. If you want to detect small differences, your sample



sizes will increase drastically. For other cases, such as different significance and power levels or comparing other characteristics, refer to *Implementing Six Sigma*<sup>3</sup> and *How To Choose the Proper Sample Size*.<sup>4</sup>

The data from last year's harvest are shown in Table 3. Sam wants to switch to red tomato fertilizer only if it gives him an average yield of at least 20% more (one-sided comparison). This amounts to 0.95 kg or 2.3 standard deviations. Looking at Table 2, he decides to treat six plants with tomato lover and six plants with red tomato. This should put him on the safe side when drawing his conclusions.

### Rule Five

#### Beware of known enemies.

Figure 1 gives you an idea of Sam's garden and the location he has reserved for his tomato plants. One thing is immediately obvious: Some of the plants will receive a lot more sunshine than others. Now suppose Sam treats six plants in the shade with tomato lover and six plants in the sun with red tomato, and at the end of the season he sees the red tomato plants clearly yield more tomatoes. What has he proven now? That the fertilizer is superior or that tomato plants in the sun yield more tomatoes? No one knows, because the effects of the two factors have been mixed.

Sam has three options. He can place all test

plants in the sun, he can place all test plants in the shade, or he can place half in the sun and half in the shade. He chooses the third option because the conclusions he can make from the test will have a stronger validity. It allows him to compare the effect of the fertilizer over various conditions of sunshine hours. In DoE terms this is called "blocking." For every known enemy, you have to develop a strategy by determining whether you will keep it constant for the test or use it as a block factor.

As a result of this rule, Sam marked 12 test spots—six in the shady area and six in the sunny area (see Figure 2). Now he has to select which three plants in each area he will treat with red tomato.

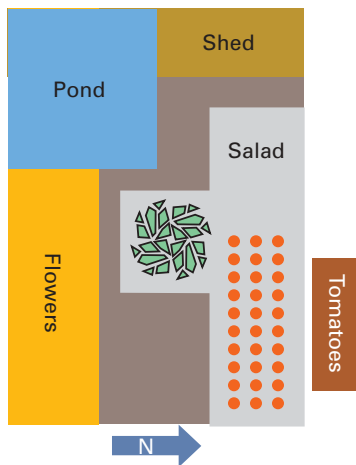
### Rule Six

#### Beware of unknown enemies.

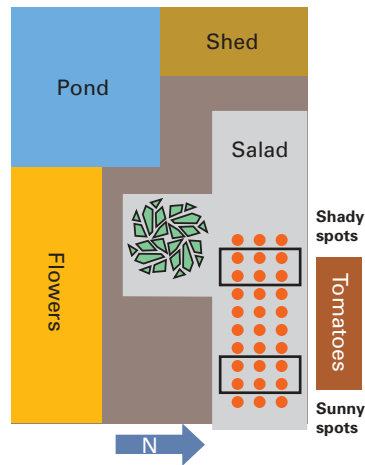
Gardens are mysterious places. They hold all sorts of differences we are not aware of, including small changes in soil composition, the effect of the wind and groundwater levels. All these factors may or may not influence the result of Sam's test. The only way he can defend himself against these enemies is to set up his experiment so these factors are distributed randomly, by chance, over the experiment.

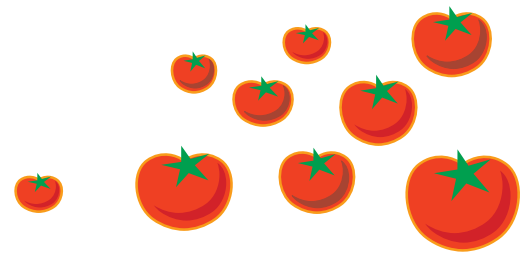
Because of the blocking, Sam must randomize within each block. This will be simple because there are only two levels to consider. So Sam takes

**FIGURE 1** Schematic Representation Of Sam's Garden



**FIGURE 2** Known Enemy: Amount of Sunshine



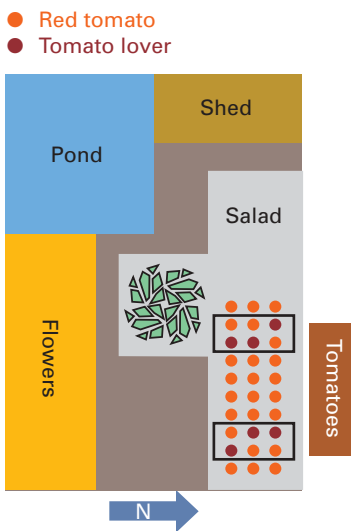


three black and three red playing cards, shuffles them and, at each test location within the block, has his daughter pick one card. If it is a black card, he will treat the plant with tomato lover; if it is a

red card, he will treat the plant with red tomato (see Figure 3).

This example illustrates a randomization in location. In many industrial tests, a randomization in time is what's needed, meaning the sequence of executing the tests has to be decided by chance within each block.

**FIGURE 3** Select Test Plants With Playing Cards



### Rule Seven

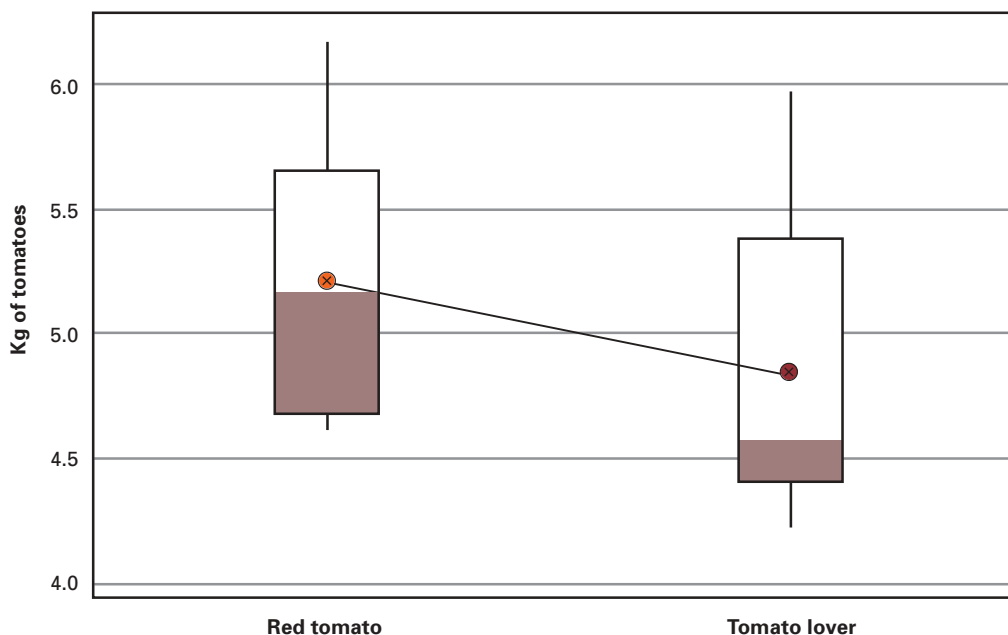
**Beware of what goes on during testing.**

Sam will not have too many problems following this rule because he's doing everything himself and will have full control of what will occur during testing. The only thing he may want to do is instruct his daughter not to touch anything in the tomato garden and not to become overenthusiastic and start watering some of the plants.

It isn't always so easy in industrial experiments. There is no end to what can go wrong during testing. In many cases, the people performing the test were not part of the team that designed it, so they have no idea what it's about or why it's being done. Keep these two golden rules in mind:

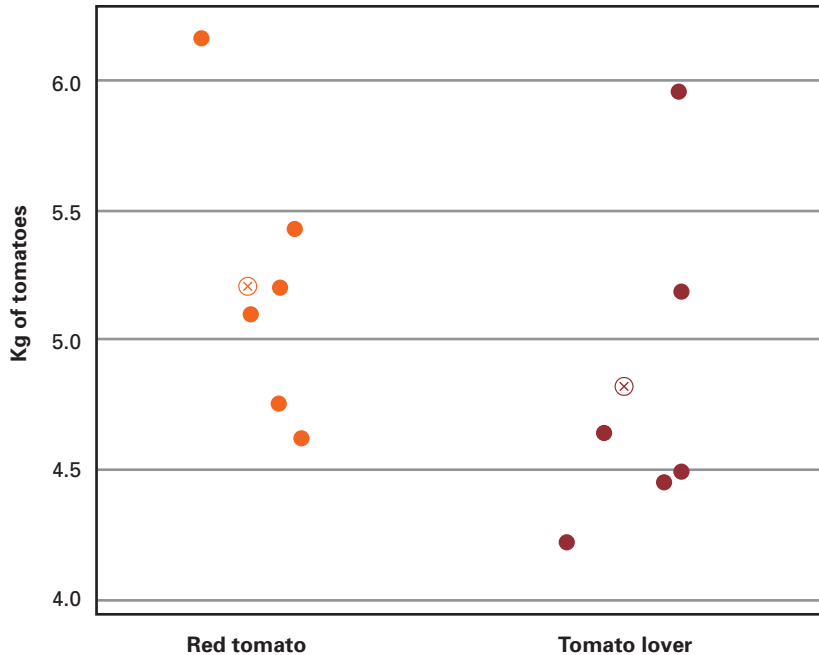
1. He who communicates is king.
2. Be where it happens when it happens.

**FIGURE 4** Box Plot of Tomato Yield Per Fertilizer





**FIGURE 5** Individual Value Plot of Tomato Yield Per Fertilizer



**TABLE 4** Test Results

| Test number | Tomato lover yield in kg | Red tomato yield in kg |
|-------------|--------------------------|------------------------|
| 1           | 4.46                     | 6.17                   |
| 2           | 4.65                     | 5.11                   |
| 3           | 5.19                     | 4.76                   |
| 4           | 5.97                     | 5.21                   |
| 5           | 4.23                     | 4.62                   |
| 6           | 4.49                     | 5.43                   |
| Average     | 4.83                     | 5.22                   |

### Rule Eight

Analyze the results statistically.

Once the tomato season came to an end, Sam weighed all the tomatoes and compiled the test results in Table 4. In rule four you learned he only wants to change if the new fertilizer brings him at least 20% or 0.95 kg more tomatoes. Sam’s observed difference is only 0.39 kg.

Statistically, you’ll use a t-test<sup>5</sup> to test the null hypothesis that the difference between the means is 0.95 kg vs. the alternative that the difference is larger than 0.95 kg. If you use the assumption of unknown but equal variances, you should use the following formulas:

$$t_0 = \frac{(\text{difference}) - 0.95}{s \sqrt{\frac{1}{n_1} - \frac{1}{n_2}}}$$

$$\text{with } s = \sqrt{\frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}}$$

and  $n_1 + n_2 - 2$  degrees of freedom.

Using the data of Table 4, you’ll find  $t_0 = -1.63$ , with 10 degrees of freedom.

The critical value at the 0.05 level is 1.812. The only way you could say with 95% confidence that the new fertilizer gives more than a 20% increase in yield would be if the calculated t-statistic were  $> 1.812$ . This is clearly not the case. In fact, that data do not show a statistically significant difference in mean at all.

If the result were positive, Sam would still have to analyze all the other characteristics to see if minimum requirements were fulfilled.

### Rule Nine

Present the results graphically.

Those who understand statistics will see from a t-test there is not a significant difference between the two fertilizers. However, not all people involved in the experiment will be knowledgeable in statistics. That’s why graphical presentation of results is important in communicating the results of the test.

Figures 4 (p. 63) and 5 show there is definitely not a 20% higher tomato harvest with red tomato

fertilizer. Figure 5 also illustrates there is a lot of overlap on the individual test results, indicating the observed difference in mean could be due to chance and has no great significance. In most cases, the graphical representation will tell the whole story.

There is no such thing as a simple experiment. No matter how simple it may look, you need to follow these nine rules if you want to be able to draw correct conclusions from your tests. Don't forget, it is just as expensive to run a bad experiment as it is to run a good experiment. The only difference is the good experiment has a return on investment.

#### REFERENCES

1. Willy Vandenbrande, "Make Love, Not War: Combining DoE and Taguchi," *Proceedings of the 54th Annual Quality Congress*, May 2000, pp. 450-456.
2. For more information on measurement systems analysis, see *Measurement Systems Analysis*, third edition (DaimlerChrysler, Ford Motor Co. and General Motors Corp., March 2002).
3. Forrest W. Breyfogle III, *Implementing Six Sigma*, John Wiley & Sons, 1999, chapters 16-20.
4. Gary G. Brush, *How To Choose the Proper Sample Size*, ASQ Quality Press, 1988.
5. Breyfogle, *Implementing Six Sigma*, p. 322, see reference 3.

#### NOTE

All statistical analyses and graphs were created using Minitab 14.

**WILLY VANDENBRANDE** is a consultant with QS Consult in Brugge, Belgium, where he focuses on production process improvement. He holds a master's degree in metallurgical engineering from the University of Ghent in Belgium. Vandenbrande is a senior member of ASQ, a certified Six Sigma Black Belt and ASQ's country councilor for Belgium. He has spoken at several Annual Quality Congresses and is a keen gardener, though no tomato expert.

### Please comment

If you would like to comment on this article, please post your remarks on the *Quality Progress* Discussion Board at [www.asq.org](http://www.asq.org), or e-mail them to [editor@asq.org](mailto:editor@asq.org).

#### SOFTWARE REVIEW

### WINSPC SOFTWARE



Statistical Process Control software that's easy to configure and use. WINSPC software provides administrators with a color-coded dashboard that automatically displays the status of plant-wide data collection operations on a single screen in real time and/or historically (by week, shift, etc).

Administrators receive real-time alarming and can drill-down on any cell to view data, statistics and charts at any collection station.

#### DataNet Quality Systems

Phone: (248) 357-2200 • Fax: (248) 357-4933

Email: [info@winspc.com](mailto:info@winspc.com) • [www.winspc.com](http://www.winspc.com)

#### ASQ MARKETPLACE



### <http://www.asq.org/shop/marketplace> Need Solutions? Look to ASQ

You've been assigned a project. You have been searching for solutions. Look no further. ASQ Quality Marketplace is your premier resource for quality products, services, and information. We connect you with solutions from the manufacturing, health care, service, and education industry.

ASQ Sales

800-248-1946

#### ASQ MAGAZINE



### Free QP Live!

If you'd like to preview the next issue of *Quality Progress*, subscribe to our FREE electronic newsletter, *QP Live*, a summary of the next issue's contents. Visit [www.asq.org/keepintouch.html](http://www.asq.org/keepintouch.html) for instructions on how to subscribe, or contact ASQ customer care at [help@asq.org](mailto:help@asq.org).

Stay on top of the issues affecting the quality profession with *QP Live!*

ASQ

800-248-1946