# TESTING VERSUS TESTIMONIALS: HOW DO YOU TELL THE DIFFERENCE? HOW AND WHY DO WE DO RESEARCH?

Joe Lauer[1]

Research is a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalized knowledge. So, research can lead to knowledge, but only if it's done well. Done "well" means using accepted scientific methods, which often include statistics. If we "just don't like" research outcomes, that does not mean it isn't science!

Applied agronomic research is a set of **planned comparisons** carried out over an **adequate** number of fields and years (sets of weather), with results accumulated, and analyzed to allow us to **predict** the response from tested inputs or practices when we use them in the future. Usually includes economics. We might think of this as a "branch" of science, in which probabilities of certain outcomes suggest whether to use certain inputs.

**Planned comparisons** are careful choices and placement of treatments to establish two or more (crop) inputs under the same (neutral) conditions, with results (yield, quality) from each input carefully measured. An "unplanned" comparison can become a planned one, but only if it meets the "same conditions" test.

An **adequate** number of fields and years depends on the expected variability of response, from none to some; depends on the frequency and cost of yield loss from using an input; depends on the cost of the product. Before we do the research, this obviously requires some guesswork.

A **prediction** is a statement of likelihood of expected results from use of a particular input: Usually includes an "average" expected result: "Product X increases yield by 3.4 bushels on average." It should include an economic assessment: "The average return to using Product X is $3.50 per acre, after subtracting its cost of $2.20 per acre." It should include some measure of uncertainty: "Product X is expected to provide a positive return 70 percent of the time, and net return is expected to range from -$2.20 (product cost, with no effect on yield) to +$10.50 per acre." If appropriate, "condition" statements should be included: "There is little return to use of this product under poor drainage conditions."

Knowing *what we want to say* when the work has been completed should be a critical precondition for undertaking on-farm, applied research. If you can't even guess what such a statement might look like, you might want to hold off. If you are "required" to reach a certain conclusion, is there any real point to the whole exercise?

*But still*, a certain percentage of all on-farm research projects are wasted effort or even harmful. The research may suggest the use of harmful inputs; it may suggest the use of useless inputs, or it may fail to predict (accurately) the benefit of a useful input. We want to reduce the frequency that this happens.

Farmers today have an increasing number of tools for managing crops. New developments in precision farming technologies, biotechnology, and advancements in pesticides, equipment, and other ag inputs are converging and arriving at the farm-gate at an unprecedented rate. Sifting

---

[1] Corn Agronomist, University of Wisconsin, Department of Agronomy, 1575 Linden Drive, Madison, WI 53706.

through the overwhelming milieu of technologies to find the tools that really work is a challenge for farmers and the consultants and agronomists that serve and support production agriculture.

Often farmers use technologies with little or no evaluation prior to use. Industry heavily invests in technology research and development, thus, “ramp-up” is fast and products are often marketed and distributed quickly in an attempt to recover investments in the early phases of technology adoption. Often farmers, usually at great expense, must learn and re-learn management of these technologies as new and improved versions are released.

In general, there are two major categories of on-farm research trials. The first is **replicated trials** that try to account for field variability with repeated comparisons. Examples include trials conducted by universities and by public and private plant breeders. The other type is **non-replicated demonstrations** such as yield contests, on-farm yield claims, demonstration trials and farmer observation and experience.

The advent of effective tools for collecting data related to crop production such as weigh wagons, on farm scales and yield monitors have removed many of the traditional barriers of on-farm trials. The next phase in the development of agriculture is necessary coordination of multi-site trials that will require collaborative specialists for data collection and analysis.

**Randomization:** Randomization prevents bias of any one treatment in any way (intended or unintended). To randomize a trial, randomly assign replicated check plots and treatments (Fig. 1) by drawing numbers out of a hat or flipping a coin as you assign treatments to plots.

**Replication:** Replication is used to determine whether the difference between plots is due to chance variation or treatment variation. Chance variation is caused by differences in weather, soil and other factors. These factors change significantly in space (field to field) and time (year to year). Through replication in both space and time, average treatment effect values can be obtained. Replication in space means that several plots of each treatment are grown in the field (field replication) or that single plots of each treatment are placed in several fields across the farm (farm replication). Replication in time is repeating the trial over several years. Comparisons between average values are more accurate than those between single plots. Replicating your check and treatment plots at least three or more times will give you much greater confidence in your results and final conclusions about a new practice if it is made only after being evaluated over several years and/or at several locations.

| Replicate 1 | Replicate 2 | Replicate 3 |
|---|---|---|
| Treatment 1 | Check | Treatment 2 |
| Check | Treatment 3 | Treatment 3 |
| Treatment 2 | Treatment 1 | Check |
| Treatment 3 | Treatment 2 | Treatment 1 |

Low ----------------------------------- High
Soil Fertility Gradient (or yield potential, organic matter, pH, etc.)

**Fig. 1. An example of a plot design, with a check plot and treated plots arranged in three replications. The entire trial area should be kept within a uniform soil condition. Other plot arrangements are possible.**

**Check (Control) Plots:** Your current practice is represented in the check or control plot. It does not receive the new technology being tested; rather it represents your current management style where your tillage practice, applied fertilizer, variety and/or applied fungicide is used in the usual manner. The check plot and the treated plot differ only in the specific treatment comparison being made. Aside from this treatment, plots are managed exactly the same to avoid biasing results.

In some trials, the new technology incorporates several practices. Avoid these if possible. For example, consider a trial that compares a farmer's current planting operation with another planting operation using different tillage, fertilization and row spacing systems. A fair comparison can only be made between the two complete systems, not any given part of either system. This kind of trial is difficult to interpret because of all the confounding interactions that may occur among the parts.

**Selection of treatments to be tested:** Keep treatments simple. Limit the number of treatments to no more than four, including one well-known treatment as a check (control) plot. As the number of treatments increase so does the complexity of the on-farm trial. Choose treatments that you expect to be significantly different. With experience you will gain confidence in your on-farm testing abilities and you can move on to testing treatments involving minor impact, difficult to test production practices. It is very important that production inputs remain constant, except for the tested treatments.

**Plot size:** Optimum plot size for on-farm tests may vary greatly with the size of uniform land area, number of treatments and size of equipment. Adjust plot lengths so that each treatment is within a reasonably uniform area or so that each uniformly covers the field variation as discussed above. Plots should be similar in size. Avoid using field edges in plots. Field edges should be left as borders. Plot width is determined by the width of equipment used to apply treatments (e.g., planter, sprayer, etc.) and/or harvest plots. The width of the established treatment should be larger than the harvest width. This way there will be a uniform harvest width and errors in harvesting will not affect side by side treatments. Typical treatment plots are between 1/10 and 1/2 acre.

**Management:** Each plot should be managed exactly the same and as close as possible to the conditions which normally occur on your farm.

**Measurements:** Depending on the test, take stand counts of the crop, ratings of weed control, disease or insects. Weigh the yield of each plot, take a moisture sample, and adjust yields to the same moisture content. Yield estimates are needed to make production and economic comparisons between treatments. Measure the size of the harvest area using a measuring tape or before or immediately after you harvest each plot. These distances are then multiplied by the width of the combine head to arrive at the harvested area and yield per acre.

Harvest the middle area of each treatment plot so that border effects do not bias the results. Yields can be measured with a local truck scale, a weigh wagon, or a properly calibrated yield monitor. Harvest equipment must be completely empty and clean before each treatment is harvested. Save a sample from each treatment to determine moisture content at harvest and any other quality factors that may be important such as test weight and protein content. If moisture contents differ between the treatments, you must be corrected to constant moisture.

**Data analysis and Statistics**

Data analysis largely depends on how the project was designed and conducted. Simple statistical software packages are available. Microsoft Excel has a very good analysis of variance procedure.

*Economics***:** Use a partial budget analysis where

Grower return = (Yield*Price) - [Yield * (Handling+ Hauling+ Storage+ Drying+ Trucking)]

- Price = Weighted Price per Bushel = 50% November 15 Average Cash price + 25% March CBOT Futures price ($0.15 basis) + 25% July CBOT Futures price ($0.10 basis). November 15 Average Cash price derived from Wisconsin Ag Statistics; CBOT Futures prices derived from closing price on first business day in December.
- Handling costs = $0.02 per bushel
- Hauling costs = $0.04 per bushel
- Storage costs = $0.02 per bushel for 30 days
- Drying costs = $0.02 per bushel per point of moisture
- Trucking costs = $0.11 per bushel for 100 miles

**Analysis:** Use averages over replicates to compare treatments. A well-conducted test will have small differences among plots of the same treatment and some large difference between treatment averages. Consider all-important traits and not just yield.

Variations in yield and other measurements because of variations in soil and other growing conditions lower the precision of the results. Statistical analysis makes it possible to determine, with known probabilities of error, whether a difference is real or whether it might have occurred by chance.

Means are often separated using a number labeled “LSD” which stands for least significant difference. LSD’s at an appropriate level of confidence (usually 10%) are used. Where the difference between two selected treatments within a column is equal to or greater than the LSD value at the bottom of the column, you can be sure that in nine out of ten chances that there is a real difference between the two treatment averages. If the difference is less than the LSD value, the difference may still be real, but the experiment has produced no evidence of real differences.

Statistics are only a tool to help prevent us from deceiving ourselves and others. Growing conditions in any particular year can have large effects on certain practices. Two years of replicated data are a minimum for supporting most practices. Statistics, as commonly used, often describe better than they predict. But, stats used over a lot of site-years can provide a measure of the usefulness of a prediction based on data. And yet, statistical statements always involve probability, and this is not always easy to “apply” when it comes to inputs. Statistics do NOT substitute for the large amount of data (site-years) that good on-farm testing always requires.

**Time:** Data from one field in one year may be misleading. About two to three years of your own tests in conjunction with other reliable information should be adequate to select treatments to be practiced on larger acreage. One year’s data may be adequate to discard poor treatments from the test. Replace discarded treatments with new treatments in any tests conducted next year.

Adjustments in the number site-years should be considered if expected variability due to soil type is high (then you need more soil types), if expected variability due to years (weather) is high (then you need more years), and/or if variability is expected to be high over both soils and years, (then you will need a lot of sites and years). If variability is expected to be high over varieties/hybrids, you have a problem.