A Guide to Testing Products and Management Practices

Part II: The Nuts and Bolts of Experimental Design

by Wendy Gelernter, Ph.D. and Larry J. Stowell, Ph.D.

Bottom line: Properly designed tests can be invaluable in identifying the products and practices that are most useful at your specific location. But badly designed tests will yield poor information, will waste your time, and can lead to erroneous decisions. When designing a testing plan, simplicity should be your goal. Carefully choosing the treatments you want to test, understanding how to replicate and randomize for best results, and making sure that plots are placed and sized correctly are the keys to a clean, simple and effective experimental design.

In Part 1 of this series ("Getting Started", in the October, 2000 issue of *PACE Insights*), we discussed how a testing program can help you improve the technical decision making process on the golf course. In this installment of the four part series "How to Test Products and Practices", we will address some of the nuts and bolts of experimental design - how to determine which treatments to test, how many treatments to test, and how to lay out the test on your golf course. If you find that you enjoy conducting your own on-site testing program, and that it has been valuable, there are a few additional sources of information, listed in the "References" section at the end of this article, that also provide some useful information. While most of these references do not address turfgrass systems specifically, they do provide some basic information on testing strategies that will help round your library out.

The experiment pictured below was designed to evaluate 3 different rates of the growth regulator Proxy on poa greens. Each 5 foot X 10 foot rectangular area represents one test plot. Each rate of Proxy was replicated in three different plots. The plots are arranged in three tiers, with each tier representing one replicate.

Perform a background check

Spending quality time thinking and doing some background reading as you plan your experimental project will pay off in the end. Get hold of information on your topic of interest from trade journal articles, scientific publications, technical product literature, product labels, the internet, or data presented at meetings. This allows you to find out which questions have already been addressed, and which questions remain unanswered. You'll also learn which testing methods have worked (or haven't worked) and which pitfalls to avoid. At a minimum, this type of background check will save you time by helping you to avoid repeating the mistakes of others. And at its best, you'll get some new ideas on how to approach your own testing program.

Control the urge

A frequent urge when starting out your own testing program is to test all of your ideas at once. However, this strategy frequently leads to more questions than answers - something a good experimental design can help you to avoid. Remember that some of the most effective experiments are also some of the simplest, where only two treatments are tested -- one new product or cultural practice compared against your current practice, for example.

To be successful in answering your questions with technically sound answers that can be effectively put into practice on your golf course, limit the number of treatments to a manageable number (in test plot lingo, a "treatment" is anything you want to test the effects of - from a new turf variety, to a product rate or formulation, to a new aerification method).

A good rule of thumb is to limit the number of treatments to five or fewer, and not more than 10. There may be times when you will need to exceed these numbers, but be assured you will be more confident in your results when fewer treatments are evaluated in an experiment. Time spent culling out unnecessary treatments before you start your test will be repaid many times over.

Break it down

If you have penciled out more than 10 treatments, consider breaking the experiment into its logical components. For example, if you are interested in determining the best timing and rates of application for fungicides labeled to control summer patch, you might at first choose to look at three different fungicides, each at the low and high labeled rates, and at two different application dates - preventative (before disease symptoms appear) and curative (after symptoms appear). That sounds like a fairly simple experiment, but in fact you would end up with 13 treatments! That would be:

3 (fungicides) X 2 (rates) X 2 (times of application)
= 12 treatments +1 non-treated control
= 13 treatments.

As you can see, adding extra factors can cause an experiment to blossom into a design that will be difficult to execute and will produce results that are hard to analyze. To simplify the experiment and to make the results easier to interpret, consider splitting the test up into two or more experiments.

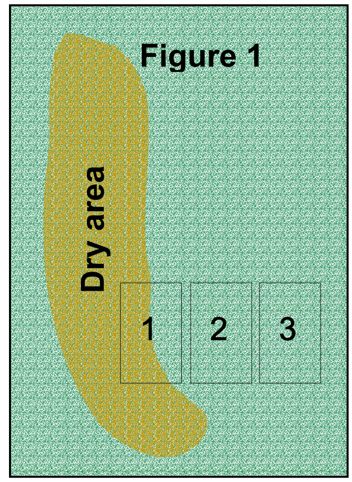
In the example above, the two main factors under investigation are rate of application and timing of application. Why not look at rates first, by keeping application timing the same for all treatments, but varying the rates? The three fungicides could then be tested at the low and high labeled rates, but the timing of application would be the same for all fungicides and all rates; this would be a seven treatment trial including the non-treated control. A second study could look at the effect of different application timings; this time, the rate would be kept the same for all treatments. By breaking the trial down, the execution of the trial and evaluation of the data at the end of the experiment are easier to handle.

Over and over: the beauty of replication

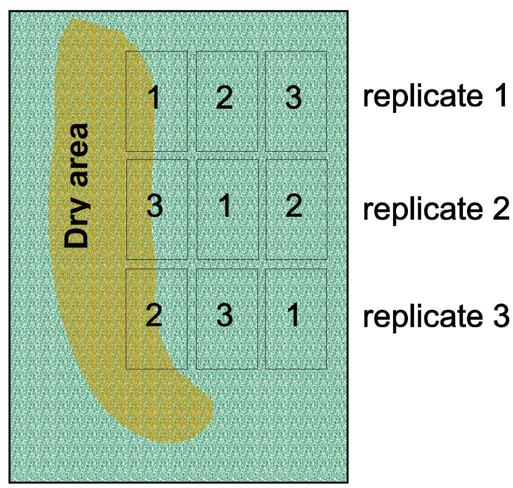
Despite our best efforts, the turf on a green or fairway is usually not homogeneous. There are differences in microclimate, moisture, turf quality and a host of other factors that result in variability that is beyond the control of the researcher. Without **replication** (repeating a treatment in two or more locations), this type of variability can lead us to draw the wrong conclusions from a research trial.

In the example illustrated in Figure 1, two fertilizers (labeled "1"and "2") are being compared against an untreated control (treatment 3). However, a pre-existing dry spot is located on one side of the green, causing poorer quality turf in the dry area.

This lack of homogeneity in turf quality, which is common, causes an unfair bias against treatment 1. Because the trial in Figure 1 is unreplicated (each treatment is repeated only once), you run the risk of a serious error – wrongly concluding that fertilizer 1 didn't work at all, and in fact made your turf look worse! Given the experimental design of only one replicate, it's impossible to tell if the results were due to a bad fertilizer, or due to some other factor, such as a pre-existing dry spot.



However, if each treatment is replicated in additional locations as in the illustration below, the chance of unfair bias against any one treatment is greatly reduced.



In most cases, three replications should be sufficient to separate out the good from the lousy treatments. If you read some of the scientific literature, you'll see that researchers frequently also use three replicates, but may use up to six replicates (and sometimes more) if they are trying to tell the difference between some very closely related treatments. However, for your purposes, where you are looking for treatments that are different enough to have a strong impact on turf quality, three replicates (as in the photograph on page 1) is plenty.

Size it right

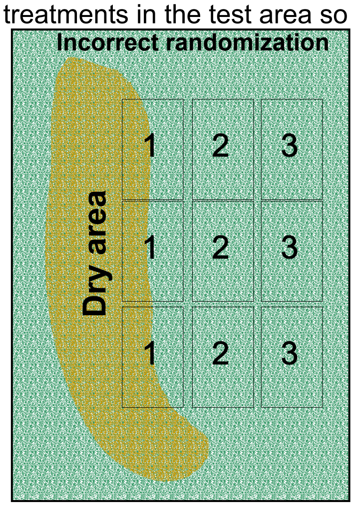
Most research trials are conducted using small plots, usually all placed within one green, one tee or one fairway. Despite the use of the term **small plots**, the last thing you will want to hear when conducting a field experiment is: "your plots are too small." We have found that the larger you make your test plots, the less likely it is that the whole plot will be destroyed by a mishap. For example, a hydraulic leak might damage half of the plot so that it is no longer usable. With larger plots however, the experiment can continue with the non-hydraulic-fluid-damaged areas of all plots being rated. Larger plots also ensure that a disease, insect or weed will be found in the test area. The smallest plots that we recommend for on-site testing are 4 ft X 4 ft (16 sq ft), but our usual small plot size is 5 ft X 10 ft (50 sq ft). For most small plot work, this is a convenient size for a sprayer that applies a 5 ft swath width.

Split greens or macho plots

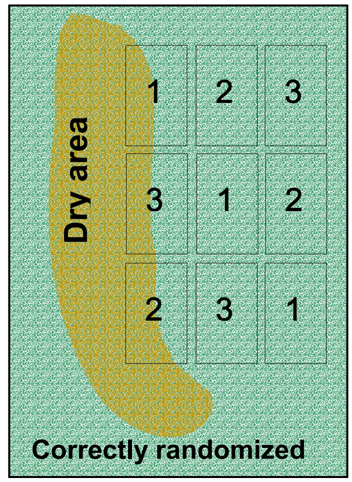
If you are not adapted to small plot work or just don't want to bother with specialized research equipment (see below), treat one-half or some other portion of a green or fairway using your standard equipment for applications. This is the best way to test a system prior to full adoption of a new cultural practice or new product. In this case, replication will probably have to take place on three separate greens or fairways due to size of the test area. A typical test would entail splitting the green in half and applying a new procedure to one half of a green and your standard treatment to the other half of the green.

Randomization, or rolling the dice

Replication (see above) helped us solve some of the problems presented by natural variability on turf. However, we need yet another tool - **randomization** - to attack this problem and to avoid bias in our design and results. The use of a randomized design helps us to properly arrange the treatments in the test area so that variability is minimized. Using the example of the fertilizer experiment illustrated in Figure 1 once again, there are several choices available to us in how we arrange the different treatments in each replicate. Problems can arise if the treatments are arranged in the same order in each replicate, as in the illustration showing incorrect randomization.



Hopefully, you have noticed that treatment 1 is receiving an unfair amount of pressure, because the dry spot is concentrated in the treatment 1 plots. Using this design, would you be able to tell whether the poor performance of treatment 1 be due to the negative effects of fertilizer 1, or is it due to the our having placed treatment 1 plots where soil conditions are dry? You have no way to find this out, using a non-randomized design.



In contrast, treatments can be arranged randomly as in the illustration to the left. In this case, the randomization has been done correctly, and the negative effect of the dry area is more evenly spread over all of the treatments, giving you a fairer look at the performance of each treatment. This illustration represents just one way in which these plots could be randomized: can you think of any others?

A simple method for randomizing treatments within each replicate is to use a deck of cards. Remove the numbered cards that correspond to each of your treatment numbers (another reason to use only 10 treatments!). Shuffle the cards and lay them down to give you a random arrangement of treatment numbers -- it's that simple.

Avoid “the nursery effect”

An interesting phenomenon that occurs at most golf courses is the "nursery effect." As a result of lack of traffic, nurseries frequently survive without disease and stress damage when most of the greens in play are struggling to survive. Under the low-stress conditions seen on most nurseries, products and practices usually perform differently than on the rest of the course. For this reason, nursery greens should be avoided for experiments - unless you are investigating some really wild idea that is too risky to try on greens in play.

Measurements

There are many factors that will influence the outcome of an experiment and only a few of which you will be able to control. The accuracy and precision with which products and practices are applied to the turf is one area where you do have some control, and where cutting corners may result in wasted time. The more care taken in precisely making measurements and calibrations, the more likely the results will be meaningful and repeatable.

You will have to accurately measure time using a stop watch, distance using a tape measure, volumes using graduated cylinders, syringes, pipettes, or precision flow meters, and weights using balances (scales) that can measure within 1 - 5% of the desired unit.

Try to measure all components with an accuracy of 1%. This means that to measure 100 grams of a product, the balance will have to have accuracy of 1 g. A standard triple beam balance will provide this level of accuracy for about $150. Volume measurements can be carried out using a variety of instruments including disposable pipettes with accuracy down to 0.01 ml for small volumes.

Equipment costs and your time

There are a variety of sources for testing equipment. Your existing equipment is the first place to start. However, if you are interested in small plot applications, Table 1 provides a list of suppliers and recommended items to assist in your efforts. Don't be fooled by the relatively low cost of the equipment needed to conduct testing programs. The investment of your time during experiment design, execution, observation and summary are far more costly than any equipment you might purchase. For that reason, a carefully designed experiment is one that will provide the greatest benefit at the least cost. Your golf course turf quality will benefit and your budget may drop, but be sure that you can afford the time needed to complete an experiment before you get started.

And, as a rule of thumb, if you think it will take half an hour to calibrate your sprayer, allot twice that time. For some strange and perverted reason, experiments always take at least twice as long as you think they will when you are sitting at your desk drafting up the objectives and materials and methods.

It's the law

Remember, it is illegal to use any pesticide that is not properly labeled, stored, and handled according its label. This extends to the use of labeled products on pests or application to sites that are not explicitly listed on the product label. Check with your County Agricultural Commissioner's office to find out what local and regional regulations must be complied with if you are testing a non-labeled product, or a non-labeled use of that product. Obtain the proper permits and certificates before conducting trials with products outside the constraints of the product label - it's the law. If you are uncertain, it's best to stick to experiments with labeled products.

References

Camper, N.D. ed., 1986. Research methods in weed science. Southern Weed Science Society, Champaign. 486 pp.

Hickey, K.D., 1986. Methods for evaluating pesticides for control of plant pathogens. APS Press, St. Paul. 312 pp.

Little, T.M., Hills, F.J., 1978. Agricultural experimentation. John Wiley and Sons, NY. 350 pp.

Table 1. Commonly used equipment in product testing. Prices are ballpark estimates for each item to provide a rough idea of the relatively low cost of equipment needed to test products on-site.

Source	Description	Cat. No.	Quantity	Price
A.M. Leonard 800-433-0633	36"Gandy spreader	36H12	1	252.00
Cole Parmer 800-323-4340	150 g Scale	H-11300-16	1	125.00
"	5000 g scale	H-1100-3-20	1	111.00
"	Container for 5000 g scale	H-11003-60	1	10.00
"	Graduated cylinder 500 ml	H-6137-90	2	15.00
"	Pipette pump	H-06221-03	1	21.00
"	Serological pipettes 1.0 ml	H-13000-06	1,000	146.00
"	Serological pipettes 10.0 ml	H-13000-36	500	170.00