A GUIDE TO TESTING PRODUCTS AND MANAGEMENT PRACTICES FOR GOLF COURSE SUPERINTENDENTS: PART I GETTING STARTED

Larry J. Stowell and Wendy Gelernter
Pace Consulting and Pace Turfgrass Research Institute
San Diego, California

Within the past two years, at least four new fungicide active ingredients, a new herbicide and a new insecticide were introduced for use in pest management at golf courses in California. Combined with the currently registered pesticides, fertilizers, amendments, and various new pieces of cultivation equipment and turfgrass management approaches, the task of selecting the correct blend of products and practices can sometimes be daunting. The only way to gain more confidence that a management system is the best one for your site is to start a testing program. This article is the first in a series of three articles that will describe how to set up a successful testing program at a golf course. Part I, Getting Started will describe the basic elements of testing program, attitude, components of an experiment, record keeping; Part II, Experiment Design will provide expanded details on how to set up experiments that will be simple and efficient, plot design, replication and randomization techniques will be discussed; Part III, Interpreting Results will provide a background on how to evaluate your findings and how to decipher scientific publications with an introduction to basic statistics. Combined, these three articles will provide sufficient information to develop an effective testing program.

You've got to have an ATTITUDE!

One of the most important factors to consider when setting up a testing program is your attitude. If you have a feeling that you want to prove a product works or doesn't work, your attitude will get in the way – your biased from the start and your bias may influence the way you perceive the results. Before venturing into the realm of testing products and processes, spend a minute of soul searching before you start a test – if you really want to see a product fail, it possibly will because your perception is biased toward failure. There is no reason to run a test under these conditions. If you are open minded, however, your level of bias will be lower and your results will be more valuable to your. With your attitude adjusted, you will run tests to COMPARE, EVALUATE, and DEMONSTRATE, never to prove.

The Experiment

A testing program is really a series of experiments that have several well-defined components that must be considered before and during the execution of each experiment. 1) Clearly state the OBJECTIVE of the test. Why is the test being conducted in the first place? 2) List the MATERIALS that were used and the METHODS of application or how the equipment was used. This should include the sprayer or spreader configuration and calibration information. 3) Once the experiment has been started, begin recording OBSERVATIONS for each product or process being tested. Observations can be descriptions of visual characteristics, numerical ratings (objective measurements, weights of clippings, soil electrical conductivity readings) or relative ratings (subjective performance estimates of quality 1 – 9 scales) of turf performance. 4) At the end of the experiment, reread all of the notes and write a discussion of that summarizes the findings and potential future tests. These four components, Objectives, Materials and Methods, Observations, and Discussion are the essential components of any testing program. Omit any one of these components and you will find that it is difficult to determine what actually happened during the experiment and you may never be able to reproduce the results.

An Example

Testing programs can take on many flavors. In this example, the superintendent suspects that fall aeration when soil salinity is elevated may have resulted in water channeling down the sand-filled aeration holes instead of uniformly flowing through the entire soil profile. The symptom is green polka dots throughout the low areas of the green surrounded by chlorotic plants. How would we test the hypothesis that there is no difference between the salt content of the sand extracted from aeration holes and the soil between the aeration holes? The experiment might look something like this:

Objective: Compare the soil electrical conductivity (EC) under chlorotic turf surrounding green

polka dots of healthy turf to the soil EC under the green polka dots.

Materials and Methods: Three cup cutter samples will be taken from green 15 in the area where the chlorotic (yellow) turf with regularly spaced green polka dots are located (front left where traffic enters the green). A knife will be used to dig out the sand under the green polka dots and the sand will be placed into a coffee cup labeled "holes." Similar samples will be collected from between the aeration holes for comparison and placed into coffee cups labeled "between." Water will be added to each sand sample until they are saturated. A Cole Parmer TDS-4 meter will be used to measure soil salinity and values will be converted to saturated paste equivalent values.

Observations: The green polka dots of healthy turf occurred over aeration holes. The roots were white and more than an inch long in the aeration holes. The roots were short and off-white colored under the chlorotic turf. The EC for the sand in the aeration holes under the green turf was 2.2, 2.8, and 2.5 dS/m and the EC of the soil between the holes was 5.1, 4.6, and 4.8 dS/m.

Discussion: The salinity of the new sand under the green polka dots is lower than the surrounding older sand. This may indicate that water is channeling down the aeration holes rather than percolating uniformly through the soil or there is some other problem with the soil between the aeration holes (e.g. compaction, low soil air movement, pathogens may be present). Because poa only tolerates a soil EC of 3.0 dS/m, the front of green 15 needs to be leached to drop soil salinity levels between the aeration holes. Soil salinity should be monitored following leaching to be sure that the soil salts have been reduced to below 3.0 dS/m and the chlorosis disappears. Recovery should also be monitored. Traffic and compaction are probably making the problem worse so golfers should be rerouted using a rope barrier on Monday, Wednesday and Fridays to move golfers to the right-front side of the green.

The above experiment is an example of how golf course testing programs can be used to provide a wide variety of benefits. This simple test described can be conducted in a matter of about one hour and the results of the test will help improve management practices by identifying the need for increased leaching and modified traffic patterns. A golf course testing program should not be limited to evaluation of products.

The Power of Nothing

An introduction to experimentation would not be complete without stressing the value nothing. For readers familiar with experimentation, nothing will interpret into a non-treated control or check area. This is an important concept and it is not as easy to define as it sounds. The non-treated area is an area within the green, tee, fairway, etc. that is managed just like the area where the treatment (fertilizer application, pesticide application, cultivation) is going to take place with the exception that nothing will be applied or nothing will be changed in your typical management program. The non-treated check area will be the yardstick to measure product or process improvements in, or damage to turfgrass quality. The location of the non-treated control area is important because must fairly represent the entire area being treated for comparison. If the non-treated control area is not fairly placed, the test will be biased – the results may be confusing – your time may be wasted.

Where Plywood is King

A piece of plywood, used properly, can save thousands of dollars in unneeded fungicide applications. How is this possible? Combined with our understanding of the power of nothing described above, plywood is one of the most effective means of providing non-treated control plots. Simply place a piece of plywood (4'x8' or smaller but usually not less than 4'x4') on an area of turf that you wish to evaluate the performance of a fungicide for control of a conspicuous disease prior to application of a test fungicide. The area under the plywood will be the non-treated control.

For example, some superintendents treat greens whenever light green or yellow rings form on greens during the fall and spring. These rings are frequently caused by *Rhizoctonia cerealis*, a relative of *Rhizoctonia zeae* but without the punch of the more difficult to control *R. zeae*. *R. cerealis* produces the unsightly rings that frequently disappear without fungicide treatment in dry weather and reappear during heavy dew or rainfall (without fungicide treatment, *R. zeae* won't disappear but the turf will). If a few irregular rings caused by R. cerealis are not a serious aesthetic problem, why apply a fungicide? But what if the problem is *R. zeae*? The safe bet is to apply a fungicide but there are many reasons why a fungicide

should not be applied to control *R. cerealis*.

These fungi produce sufficiently different symptoms that a superintendent can tell the two apart – but not without some testing first. When the irregular yellow rings show up, spray as you would normally (for Rhizoctonia, Prostar is a good choice). But this time, place a piece of plywood over some of the irregular yellow circles before spraying so that nothing is applied to these areas – the non-treated control plot. Mark the location of the corners of the plywood using turf paint so the non-treated areas can be easily identified for three to four days. Observe the plots daily for several days and record your observations. In this case, what would the objective of the study be? Can you list the materials and methods for the study proposed? What might the outcomes be? How much money might you save if the irregular yellow rings disappear in the non-treated control plots – and why?

Record Keeping

The foundation of any testing program is record keeping. Before starting any testing program buy several composition notebooks at a stationery store. These notebooks are inexpensive and they will provide a reliable place to keep a chronological log of your testing programs. Leave a few pages blank at the beginning of the book to use as an index. Use only ballpoint pens that do not have water-soluble ink. Tape a business card into the inside cover in hopes that someone will return the record book if it is misplaced. Write freely in the book any observations that you have regarding the performance of a product, the reaction of golfers to a management practice, or difficulty handling or applying a material. Excess information is better than insufficient information. This notebook will be the object of discussion in resolving disputes about which practice or technique is best and where or how a product was applied. Date each entry and take notes carefully so that your efforts are not wasted.

Stay tuned for the next installment of A Guide to Testing Products and Management Practices for Golf Course Superintendents: Part II Experiment Design. You will find out that it is easy to set up a meaningful experiment and even easier to design an experiment that leaves you with more questions than you had to start with. Keep It Simple Superintendent!

Part 2, Experimental Design

In Part 1, Getting Started, we left off with record keeping. This installment of How to Test Products and Practices will pick up from that point and provide some guidelines in experimental design - how to select treatments and how many treatments makes a good trial. The information provided here is only a launching point for to get you started in an testing program. If you enjoy the process and find on-site testing has been valuable, there are a few additional sources of information that will help you develop additional testing programs (Camper 1986, Hickey, K.D. 1986, Little and Jackson 1978). Unfortunately, most of these resources do not address turfgrass systems but they do provide a variety of ideas on testing practices.

Thinking Pays Off

Spending quality time thinking about an experimental project will pay off in the end. Gather up your ideas and find some supporting research before you develop your plan. The research can be something provided by a salesman, an article in a trade journal, or a scientific publication. If you can't find any supporting research for an idea that you want to test, you haven't looked hard enough. This doesn't mean that you shouldn't test something entirely new - it means that there are few ideas that are new and you should take advantage of other people's efforts so that you can contribute information builds on the information already available. In most cases your efforts will be directed at improving or modifying an existing process or use of a product. The payoff is in better performance, cost savings, and improved turf quality.

For example, what if you have heard that nitrate nitrogen fertilization may aid in root development and reduced shot growth compared to urea and ammonium nitrogen sources. A review of the literature finds some supporting evidence (Glinski et.al., 1990) so you want to see if nitrates improve rooting in a golf green. The publication or report should detail methods that can be adapted to a test under real golf course operational conditions. In this case, the best root to shoot ratios were obtained when the plants were supplied with three nitrate:ammonium in a ratio of 3:1. Although other ratios were evaluated, why not try the best case first compared to current nitrogen fertilization and fine tune the system later.

Control the Irge

A frequent urge when starting out testing your own ideas is to test all of the ideas at once. This strategy frequently leads to more questions than answers. To be successful in answering more questions than you started with, limit the number of questions or treatments to some number that is manageable. A good starting point is to limit the number of treatments to five or fewer if possible and not more than 10 except in special cases. There will 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 will be repaid many times over.

Break it Down

Because some questions are complex and experiments may become difficult to execute, break these projects into smaller components that are easier to manage. If you have penciled out more than 10 treatments, break the experiment into its components. For example, if you are interested in the timing and rates of application for fungicides labeled to control summer patch disease of poa greens, the top three fungicides might be evaluated at the low and high labeled rates at two times of application - preventative (before disease symptoms appear) and curative (after symptoms appear). 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. In stead of this more complicated design, split the experiment into its component parts.

In the example above, the two main factors under investigation are rate of application and timing of application. Why not use two different greens and test the rates of application using the standard preventative application program and compare the three products at low and high-labeled rates. This study will compare the effectiveness of the three products under normal preventative application conditions. It is a seven-treatment trial including the non-treated control. The second study would only look at the high labeled rate of each product used as a curative treatment. This study is a four-treatment study including the non-treated control. 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

In order to be sure that the differences observed during an experiment are the result of a treatment and not simply variation in the quality of the turf across the test area, each treatment should be repeated, or replicated. The use of replicated treatments in a small experiment allows the experimenter to evaluate the variation that naturally occurs across a test area compared to the performance of the product in the test area. In most cases, three replications should be sufficient to separate out the good from the lousy treatments. If the differences are very small between a treated and the non-treated control plots - then the test material or process probably isn't much better than the non-treated control. Three replicate plots will be sufficient to pick out the superior and inferior treatments. More than three will improve your ability to separate differences that are close together but these differences may not be large enough to improve turf quality under normal conditions.

Size it Right

The last thing you will want to hear when conducting a field experiment is "you've got small plots." We have found that the larger you can make your test plots, the less likely that the whole plot will be destroyed by a mishap. Larger plots also make sure that a disease, insect or weed will be found in the test area. The smallest plots that we recommend for on-site testing is 4 ft X 4 ft (16 sq ft). Our usual small plot size is 5 ft X 10 ft (50 sq ft). When we are more concerned about variability in the test area, performance of a product or distribution of a pest, we increase plot size to 7 ft X 10 ft (70 sq ft). For most small plot work, a 5 ft X 10 ft plot is a convenient size for a sprayer that applies a 5 ft swath width. Unexpected events undoubtedly will occur during your experiments, for example a hydraulic leak that damages half of the plot so that it is no longer usable. With large plots, the experiment can continue with the non-hydraulic-fluid-damaged areas of all plots being rated.

Split Greens or Macho Plots

If you are not adapted to small plot work or just don't want to bother with the experimental equipment, treat half or portions of greens or fairways using your standard equipment for applications. This is the best way to test a system prior to full adoption of a cultural process or product change. In this case, replication will probably have to take place on separate greens or fairways due to size of the test area. A typical test would entail splitting greens and applying a procedure to half of a green and the standard treatment to the other half of the green. As mentioned in part 1 of this series, plywood can also perform a valuable role by providing a non-treated area when large area is being treated.

Role of the Dice

In addition to replication that helps remove the effects of variation in the test area, randomization is needed to remove our bias in locating test plots within the test area. Randomization applies to the location of each treatment area or plot within the entire area. The role of randomization is to make sure that no plots occur in the same order in each block. Randomization helps take into account any systematic impacts on the trial. For example if someone drives through the first block, the extra traffic will impact all three treatments. In the perfect world, each treatment would be replicated in each block. For our purposes, three replicates should serve the needs of identifying strong improvements in turf quality and strong detriments to turf quality.

A simple method of randomizing treatments within each replicate block is to use a deck of cards. Remove the numbered cards that correspond to each of the treatment numbers (another reason to use only 10 treatments?). Shuffle the treatments and lay the cards down in order that the treatments will be located in each block. In many cases, the first replicate will be placed in order corresponding to a treatment number in a protocol to make viewing the first replicate easier. If you need more treatment numbers, simply add numbers to several other playing cards in the deck.

The Nursery Effect

An interesting phenomenon that occurs at most golf courses is the "nursery effect." Nurseries somehow survive without disease and stress damage when most of the greens in play are struggling to survive. The lack of traffic on a nursery green makes this area a poor candidate for experiments - except really wild ideas that are too risky to try on greens in play. Do not use nurseries for experiments - results probably will not interpret into useful management decisions for the rest of the course 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 cutting corners may result in wasted time. The more care taken in carefully making measurements and calibration the more likely the results will be repeatable.

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

As a rule of thumb, try to measure all components with an accuracy of 1%. That means to measure 1 gram of a product the balance will have to have accuracy of 0.01 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 from disposable pipets with accuracy down to 0.01 ml for small volumes

Equipment Costs and Your Time

There are a variety of sources for equipment to help apply products. 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 in your time during experiment design, execution, observation and summary are far more costly than any equipment that you might purchase. For that reason, a carefully designed experiment is one that will provide the greatest benefit at the least cost.

Cost being your time and efforts. 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 take 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 the 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. A research authorization (RA) must be obtained from the California Department of Pesticide (DPR) Regulation and the County Agricultural Commissioner must be notified prior to application of any product to a site or for control of a pest that is not listed on the product label. In addition, research conducted under a research authorization will require that a qualified applicator certificate holder be certified for demonstration and research applications. Stick to experiments with labeled products or obtain the proper permits and certificates before conducting trials with products outside the constraints of the products label - its the law.

References:

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

Glinski, D.S., Mills, H.A., Karnok, K.J., and Carrow, R.N. 1990. Nitrogen form influences root growth of sodded creeping bentgrass. HortScience 25:932-933.

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 for use in product testing. Products can include fertilizers, fungicides, herbicides, insecticides, nematicides, biostimulants, adjuvants, wetting agents, etc. 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/Quantity
A.M. Leonard				
800-433-0633	36"Gandy	36H12	1	252.62
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
	Pipet pump	H-06221-03	1	21.00
	Serological pipets 1.0 ml	H-13000-06	1,000	146.00
	Serological pipets 10.0 ml	H-13000-36	500	170.00
	Electrical conductivity meter, TDSTestr 4	H-19800-30	1	51.95
	Calibration solution for EC meter (500 ml)	H-01482-70	1	14.95
R&D Sprayers				
318-942-1001	Plot sprayer for liquids	Model AS	1	604.50

Part III, Interpreting Results: An Introduction to Basic Statistics

Luck. An uneducated gambler in Las Vegas depends upon it, usually to their surprise and dismay, while the more experienced gambler carefully calculates the probability that they can win. Similarly, as a professional and as a superintendent, you prefer not to rely on lucky guesses when making management decisions on the golf course. Instead, you strive to make sure that your key decisions are based on factual information that allows you to accurately predict how new products and management practices will perform on the golf course.

When designed properly, a good testing program helps to support you in this effort. While you can never eliminate the possibility of unexpected results, you can surely reduce the possibility that you will be unpleasantly surprised by basing your decisions on data from a sound testing program.

Statistics: Managing the Game of Chance

When a "fair" coin is tossed into the air, the likelihood that the coin lands with the heads facing up is 1/2, or 50%. This probability represents the number of heads on the coin (1) divided by the total number of sides on the coin (2, heads and tails). Probability theory tells us that there is a 50% chance that you will win a bet every time the coin is tossed regardless of whether you select heads or tails. Even if five tosses of the coin come up tails, the chance that the next toss of the coin will be heads is still 50% — no more and no less than for any other toss of a fair coin.

When a field test is conducted, the odds are not so easily calculated as they are for a coin toss. Why is this? The answer is that the number of **variables**, or factors than can contribute to the outcome, are much higher for a field test than for a coin toss. In a field test, the turfgrass variety, turfgrass stress, soil type, traffic patterns, weather etc. can have a big effect on the performance of products and practices. In contrast, the number of variables contributing to the outcome of a coin toss are limited.

Because we cannot use guesswork, probability theory or any other system to predict how a product or practice will perform, field tests are conducted to give us information that can be used to make the best possible predictions and decisions.

Statistics is the tool that allows you, as objectively as possible, to analyze the information collected from field tests, and to predict, with as much confidence as possible, which products or practices will give you the best results. In the first two installments of this series (Stowell and Gelernter, 1997; Stowell and Gelernter, 1998), we described how to set up field tests, and how to collect the results. In this 3rd and final installment, we will describe the final steps — how to analyze the results statistically, by calculating the **mean**, the **standard deviation**, and the **confidence interval.** In addition, we'll review methods that will allow you to clearly represent the results in the form of **line graphs, bar charts** and **data tables,** for use in your own records and for presentations to greens committees, general managers and others.

A Real Life Example

To give our discussion of statistics some grounding in reality, we will use results from a field test conducted by the PACE Turfgrass Research Institute in 1997. This test was conducted with the assistance of Bill Gallegos, CGCS at Los Coyotes Country Club, and with financial support from Valent Corporation The objective of the test was to look at the performance of 3 different rates of an experimental fungicide (procymidone, Valent Corporation) and to compare it to a standard fungicide, iprodione (Chipco 26019, Rhone-Poulenc) for control of dollar spot, *Sclerotinia homeocarpa,* on a creeping bentgrass nursery. The five different treatments tested are listed in Table 1. Each treatment was replicated three times, and treatments were randomized. Results were collected two weeks after the fungicides were applied by making a visual estimate of percent turf damage due to dollar spot.

How do we use this data to make a decision on the best product and best rate to use for controlling dollar spot? By calculating the mean, the standard deviation and the confidence interval, as described below.

Calculating the Mean

Usually, the first statistic that is calculated is the **mean**, or average rating for each treatment. The mean is calculated by summing the values for each replicate of a given treatment, and then dividing by the number of replicates. For example, in our dollar spot experiment, the mean for percent dollar spot in plots treated with procymidone at 0.5 oz active ingredient/1000 sq ft is 4.667 (rounded to 4.7 in Table 1):

sum of values (5 + 7 + 2) ÷ number of replicates (3) = mean (4.667)

This process is repeated for each treatment, as illustrated in Table 1. The mean can be easily calculated with pencil and paper. If you are using a calculator, the mean may be represented by the symbol "**X**" with a horizontal line drawn above it.

Although the mean is a powerful statistic, when used by itself, it can be misleading and can push you towards poor decisions. This is because the calculation of the mean doesn't take into account the **variability** of the results.

Variability: a Complicating Factor

There are many factors beyond our immediate control at work on a golf course, such as microclimate, moisture, turf quality, pest pressure, etc. As discussed in Part 2 of this series (Stowell and Gelernter, 1998), these factors exert a powerful force on the way a product or practice performs, and how consistently it performs. The use of replication (repeating a treatment in two or more locations) in designing your field test helps to minimize the effects of variability, but it can't erase them. As a result, it is extremely rare for a given treatment to produce the same result each time it is applied. In the dollar spot experiment, for example, procymidone applied at 0.5 oz active ingredient/1000 square feet produced three different disease incidence levels in each of three identical plots — 5%, 7% and 2% (Table 1).

How does variability affect your interpretation of the results? Let's assume that in the dollar spot experiment above, variability was much lower. In that case, the percent dollar spot values for the procymidone 0.5 oz treatment would be much more similar, for example, 4.6%, 4.7% and 4.7%. The mean for these hypothetical values would be identical to the mean calculated above — 4.667, but the variability would be less.

Which set of data gives you a greater guarantee that the product will perform the same way the next time you apply it? Which data set gives you a greater sense of confidence? Statisticians tell us that the data set with the lowest variability gives us the best predictions for how products will perform. So, even when the means are the same for two data sets, we still want to know how variable the data was.

Measuring Variability: the Standard Deviation and the Confidence Interval

There are a variety of statistics used to measure variability, but the most commonly used measure is the **standard deviation,** frequently represented by the symbol "**S**" on a hand calculator. A small standard deviation indicates that there is less variability associated with the mean — the data is more consistent — than the same mean with a large standard deviation. In the dollar spot example presented in Table 1, the highest standard deviation (6.3) occurs in the non-treated check treatment, and the lowest standard deviation (0.6) occurs in two of the procymidone treatments — the 1.5 oz and 2.5 oz rates.

Calculating the standard deviation is more complicated than calculating the mean, and we encourage you to purchase a calculator (most simple scientific calculators include standard deviation), or use a spread sheet program, such as Microsoft Excel, that performs the standard deviation function.

Looking at the means and standard deviations in Table 1, which treatment or treatments do you think gave the best dollar spot control? We still have one more calculation to perform before we can answer that question — **the confidence interval.**

The confidence interval is related to the standard deviation, and is an easy way to represent the interval, or range of values, or degree of variability associated with a mean. The lower end of the interval is calculated by subtracting the standard deviation from the mean, and the higher end of the interval is calculated by adding the standard deviation to the mean. Staying with the example of procymidone at 0.5 oz, the confidence interval for this treatment would range from 2.2 (4.7 - 2.5) to 7.2 (4.7 + 2.5). In other words, we have a high level of confidence that the mean value for this treatment falls between 2.2% and 7.2%; and our best estimate for that mean is 4.7%.

To find out which treatments performed statistically differently from another, look for the treatments where the range of values of the confidence intervals do not overlap. For example, the non-treated check, with confidence limits of 13.0 - 25.6, is statistically different from all of the other treatments, whose confidence intervals never get as high as 13.0. In contrast, procymidone at 1.5 oz and 2.5 oz have overlapping confidence intervals. This means that, based upon the data from this trial, the treatments did not perform differently

Once all of your calculations have been completed, make a **summary table** similar to that in Table 2. This table shows a letter following each mean value, something you will frequently encounter when reading scientific papers. These letters are a way of illustrating which confidence intervals overlap, and which don't. For example, values (such as 1.3%, 1.7% and 2.3%) followed by the letter "a" have overlapping confidence intervals and are therefore not statistically different from one another. In contrast, values that are followed by different letters ("b" or "c" in the case of Table 2), are statistically different from those followed by "a"s.

In fact, all of the information required to determine which treatments are best, which are worst, and which are the same, is contained in Table 2, but it's difficult for most of us to read tables. That's where graphs come in.

One Picture is Worth aThousand Words

One of the best approaches towards interpreting results is to graph the information. There are two types of graphs that are used to illustrate data collected from field trials such as the fungicide trial described above, **the bar chart** and **the line chart**.

For either type of chart, there are two axes, or lines, that define the chart — the horizontal axis, also called the "X" axis, and the vertical axis, also known as the "Y" axis. Figure 1 illustrates the results of the dollar spot fungicide experiment presented in a bar chart. The X axis has no numerical units, just treatment names. The Y axis represents the mean percent dollar spot values presented in Table 1. Thus, the bar for the non-treated check is the tallest bar, registering at 19.3% dollar spot. The vertical lines extending above and below the tops of each bar are called **error bars** and represent the confidence intervals for each treatment mean.

We suggest that you always try to graph your data. You can plot the results by hand using graph paper, or you can let a spreadsheet program on the computer do it for you, automatically.

Finale

This series of articles has described simple methods for designing a field testing program — from developing an experiment plan with clear objectives, to executing an experiment, to analyzing the results. Although this is a cursory look at the scientific process, we hope it encourages you to begin, or if you have already started, to continue testing new ideas. Remember to take care to record your objectives, materials

and methods, results, and conclusions. The next time someone asks you why you selected a particular practice or product, you may be able to pull a notebook from the shelf and point to a graph illustrating the advantages of your approach. Aside from personal pride, there is no better way to answer an agronomic practices question than to run a carefully designed, simple experiment.

References:

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

Stowell, L.J., and Gelernter, W. 1997. How to test products and practices: Part I, Getting started. Calif. Fairways, November/December, 1997. pp. 20-22.

Stowell, L.J., and Gelernter, W. 1998. How to test products and practices: Part II, Experiment design. Calif. Fairways, Jan/Feb 1998. pp. ???

Table 1. Results of a fungicide trial for control of dollar spot on creeping bentgrass. Rates of fungicides are represented as ounces of active ingredient per 1000 sq ft (oz ai/M). Check refers to the non-treated check plot.

	Procymidone oz ai/M		Iprodione Oz ai/M		Check
	0.5	1.5	2.5	2.0	
Percent dollar spot (replicate 1)	5	2	2	2	13
Percent dollar spot (replicate 2)	7	1	1	4	25
Percent dollar spot (replicate 3)	2	1	1	1	20
Total	14	5	4	7	58
Number of replicates	3	3	3	3	3
Mean (=total ÷replicates)	4.7	1.7	1.3	2.3	19.3
Standard Deviation (S)	2.5	0.6	0.6	1.5	6.3
Confidence Interval (mean - S) to (mean + S)	2.2 - 7.2	1.4 - 2.0	0.7 - 1.9	0.8 - 3.8	13.0 - 25.6

Table 2. Summary of dollar spot control results using tabular format. The numbers in the percent Dollar Spot column followed by the same letter are not significantly different using the standard deviation as the confidence interval. If the confidence intervals overlap, it is unlikely that the means are different.

Treatment (oz/M)	Mean % Dollar Spot	Confidence Interval
Procymidone 2.5	1.3 a	0.7 - 1.9
Procymidone 1.5	1.7 a	1.4 - 2.0
Iprodione 2.0	2.3 ab	0.8 - 3.8
Procymidone 0.5	4.7 b	2.2 - 7.2
Check 0.0	19.3 c	13.0 - 25.6