Testimonials versus Testing: How do you tell the difference?

Joe Lauer

University of Wisconsin-Madison

Wisconsin Crop Management Conference

Madison, WI

January 10-12, 2012

http://corn.agronomy.wisc.edu

Decision Making
The chances of getting it right

The sale
Odds= ?

Flipping a coin
50% chance

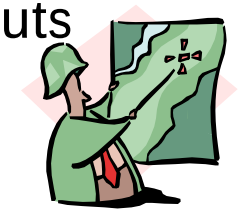
No method predicts with 100% certainty.

Rock/Paper/ Scissors
33% chance

The Scientific Method
Known probabilities
90-99% chance

For Today …

- **We will assume that everyone is interested in "neutral inquiry" without a requirement to produce a certain outcome (or meet a sales target)**
- **We will assume honesty in approach, meaning that we won't "bias by design" or cherry-pick results**
- **A certain percentage of all research projects are wasted effort or even harmful:**
  - ✓ Suggest the use of harmful inputs
  - ✓ Suggest the use of useless inputs
  - ✓ Fail to predict (accurately) the benefit of a useful input
- **Knowing what we want to say when the work has been completed should be a critical precondition for undertaking 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?**
- **The goal of research is <u>to predict future responses</u>.**



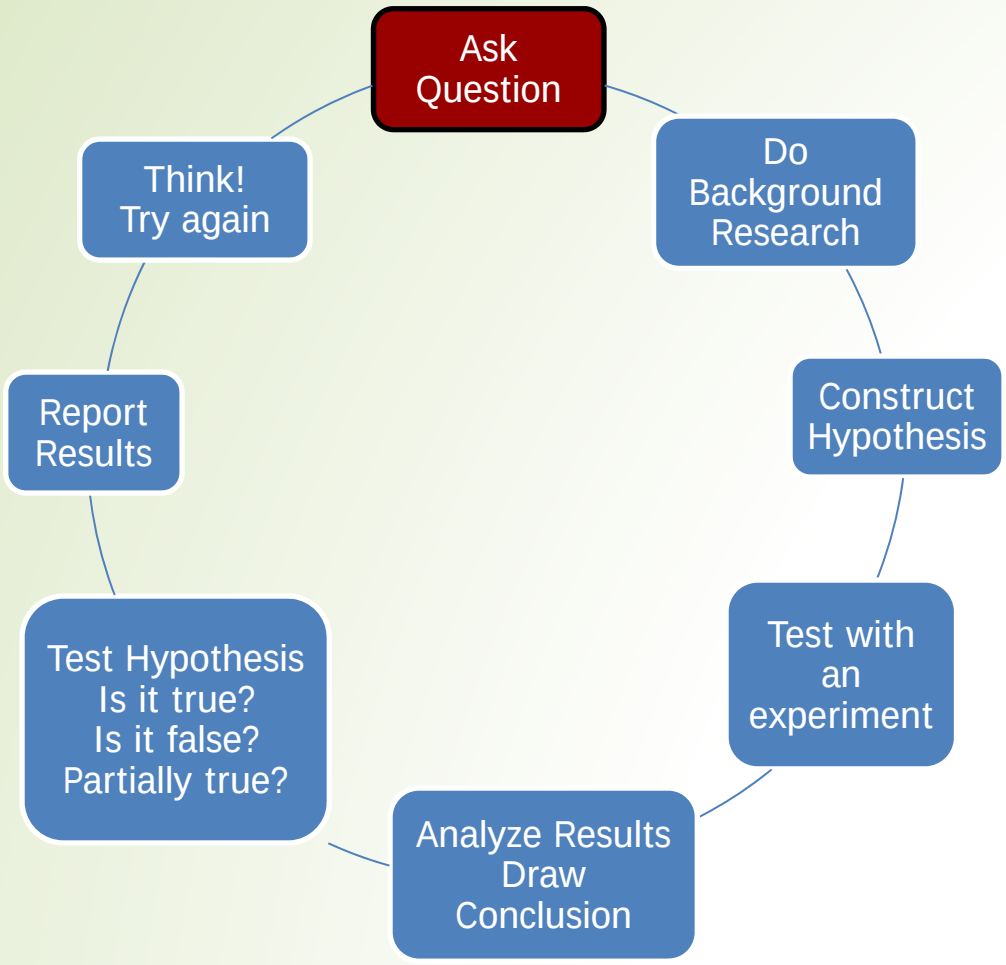
What is a good test?

- **"Planned comparison:" careful choice 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 planned, but only if it meets the "same conditions" test.
- **"Adequate" number of fields and years (Replication and Randomization)**
  - ✓ 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
- **"Prediction:" a statement of likelihood of expected results from use of a particular input**
  - ✓ Usually includes "average" expected result: "Product X increases yield by 3.4 bushels on average"
  - ✓ 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"
  - ✓ 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."

The Scientific Method

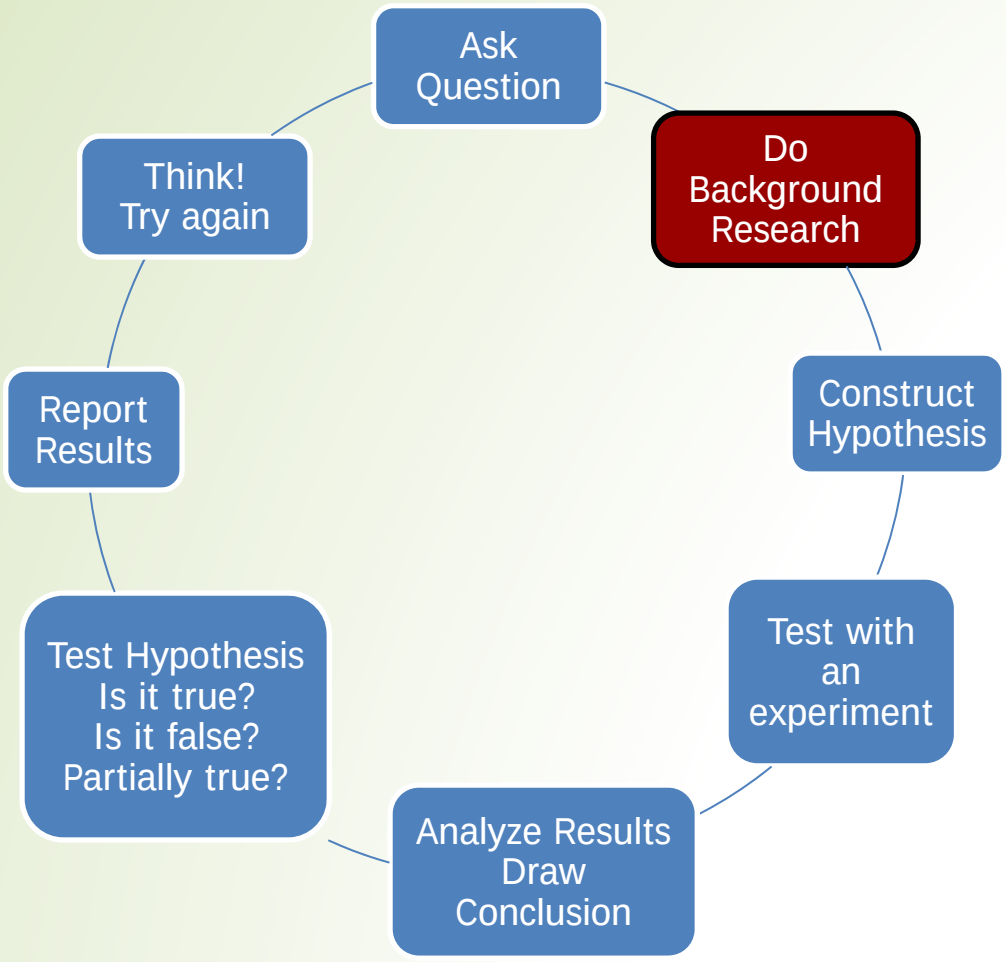
http://corn.agronomy.wisc.edu

The Scientific Method



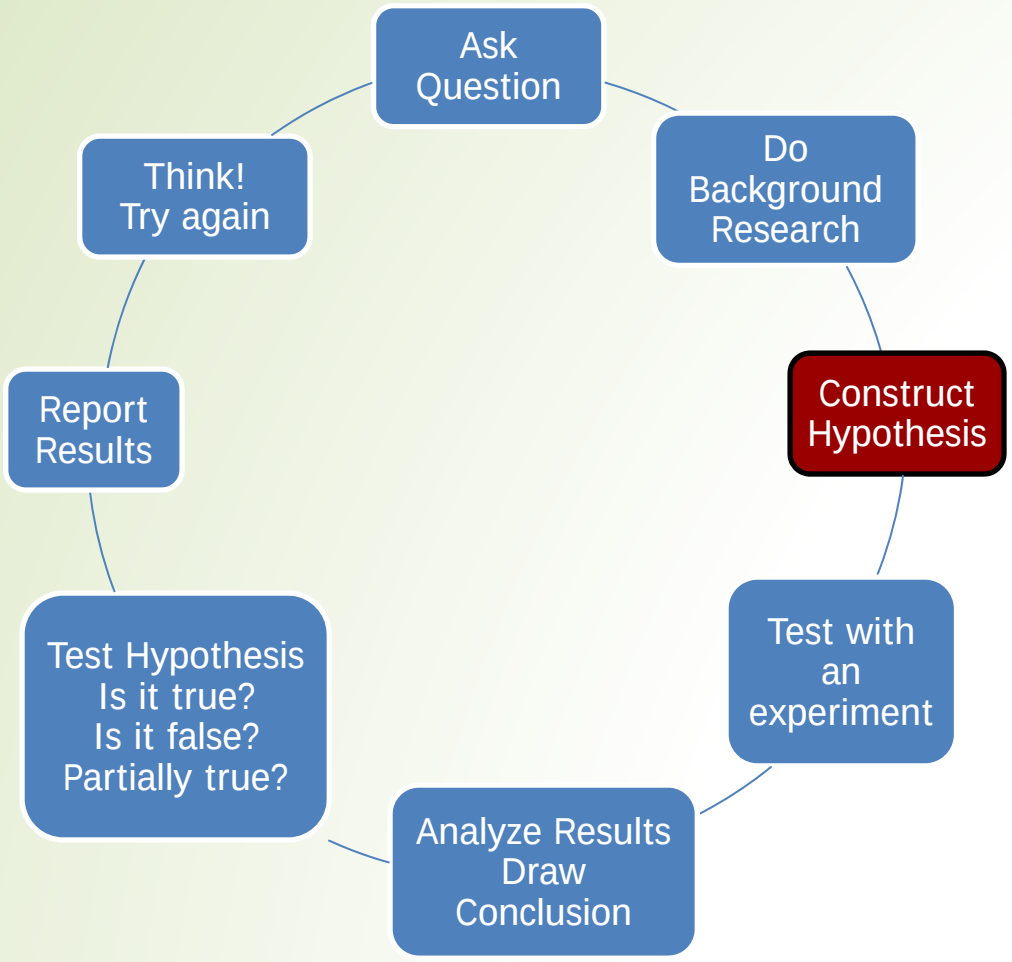
- **Ask a Question:**
  - ✓ The scientific method starts when you ask a question about something that you observe: How, What, When, Who, Which, Why, or Where?
  - ✓ And, in order for the scientific method to answer the question it must be about something that you can measure, preferably with a number.
- **When developing questions …**
  - ✓ Keep it simple, simple, simple!
    - ❑ Trials require time, energy & money.
    - ❑ Complex trials involve more of each.
  - ✓ Best questions involve a yes/no answer.
    - ❑ Herbicide ‘A’ versus herbicide ‘B’
    - ❑ Treated soybean versus non-treated

The Scientific Method



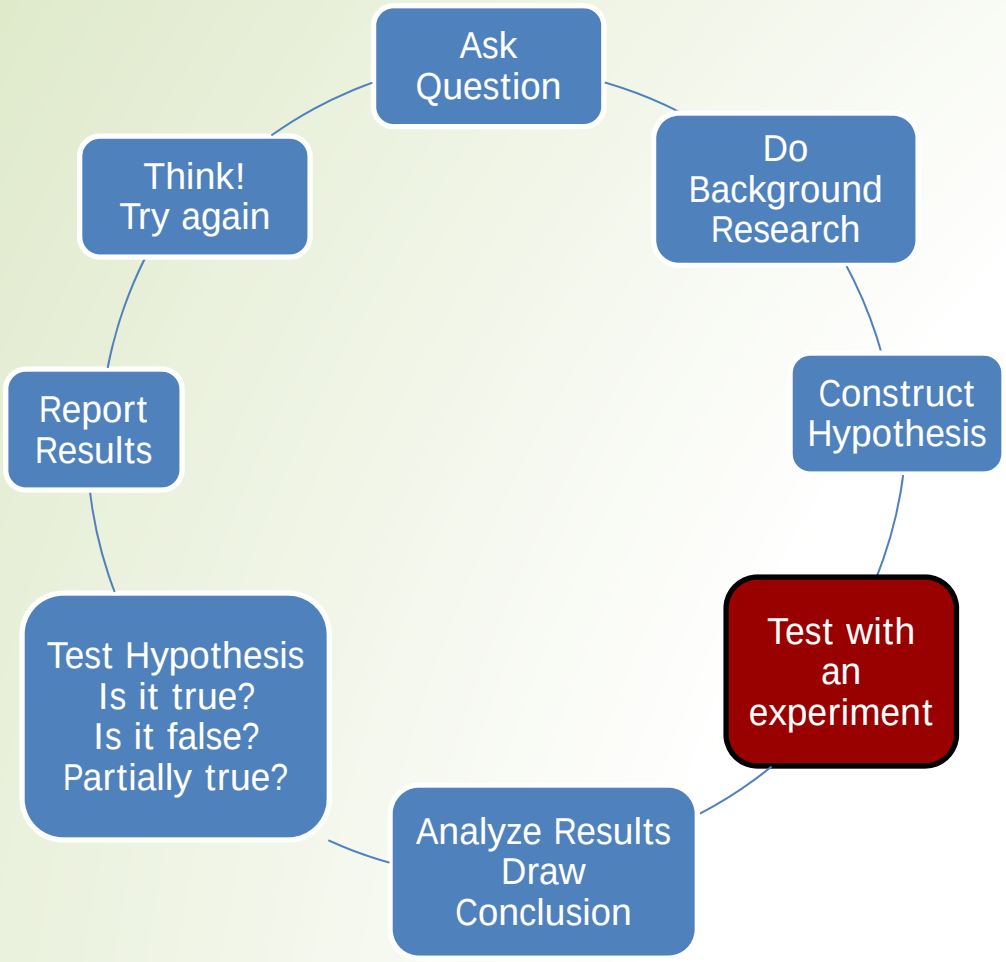
- **Do Background Research:**
  - ✓ Rather than starting from scratch in putting together a plan for answering your question, you want to use the library to help you find the best way to do things and insure that you don't repeat mistakes from the past.
- **Request help: A poorly planned experiment has a high risk of failure.**
  - ✓ Treatment choice, field selection, treatment replication, treatment randomization, plot layout & size, etc.
  - ✓ If research is not your profession, then ask for help from those who conduct research for a living.
    - ❑ University researchers & extension specialists
    - ❑ Industry researchers & agronomists
    - ❑ County Agents and Crop consultants

The Scientific Method



- **Construct a Hypothesis:**
  - ✓ A hypothesis is an educated guess about how things work: "If _____[I do this] _____, then _____[this]_____ will happen."
  - ✓ You must state your hypothesis in a way that you can easily measure, and of course, your hypothesis should be constructed in a way to help you answer your original question.

The Scientific Method



- **Test Your Hypothesis by Doing an Experiment:**
  - ✓ Your experiment tests whether your hypothesis is true or false. It is important for your experiment to be a fair test. Plan to objectively (without bias) test the question.
  - ✓ You conduct a fair test by making sure that you change only one factor at a time while keeping all other conditions the same.
  - ✓ Decide what treatments to compare
  - ✓ Design the field layout of the experiment
  - ✓ Determine what you will measure
  - ✓ Repeat your experiments several times to make sure that initial results weren't just an accident.

Experimental Unit

- **The "Experimental Unit" is the experimental material to which a treatment is applied.**
- **Sometimes it can be difficult to identify**
  - ✓ For field experiments, the Experimental Unit is usually some land area

Selecting treatments for research trials

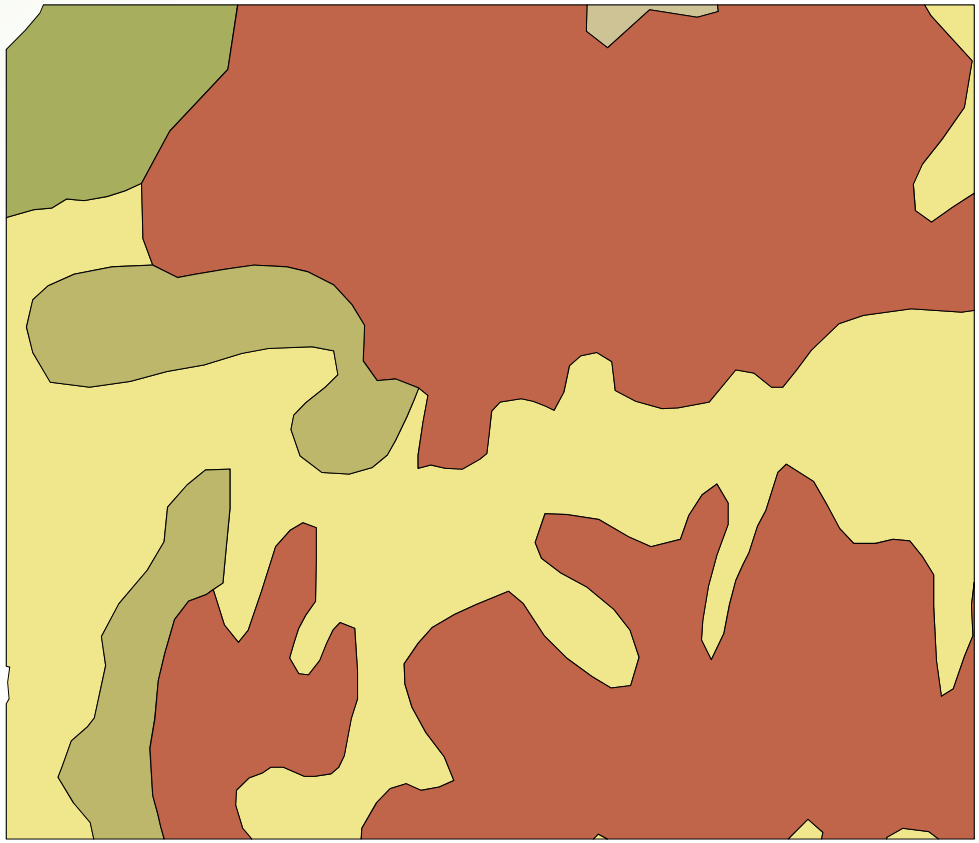
- **Choose treatments to meet objectives**
- **Keep the trial manageable**
  - ✓ Usually two treatments
  - ✓ No more than 12 plots (JL)
- **Include a control or check treatment.**
  - ✓ A good control may be the grower's standard practice.
- **Experimental designs**
  - ✓ Paired comparisons (two treatments)
  - ✓ Randomized complete blocks (three or more treatments)
- **Include a range of treatment levels if variable inputs are tested**
  - ✓ e.g., corn plant densities of 25K, 30K, 35K, and 40K seeds per acre
- **To improve efficiency of managing the experiment, you need to consider grower equipment for experiment layout**
  - ✓ Planter
  - ✓ Combine
  - ✓ Sprayer

Randomization and Replication

- **Randomization – Eliminates bias when assigning treatments to experimental units**
- **Replication – Improves the estimate of the treatment effect and provides estimate of error**
  - ✓ Number of replications for trials
    - ❑ Minimum is two
    - ❑ Recommend three or four within a field
    - ❑ Can replicate by fields and farms
- **These two factors separate research experiments from demonstration plots**
  - ✓ Grower can make valid conclusions and ultimately wise business decisions

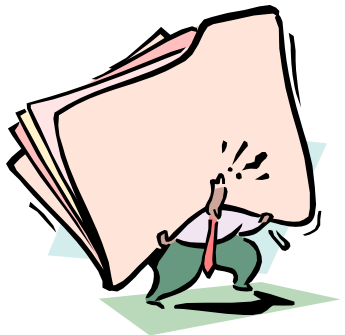
Location of research trials

- **Choose a uniform field**
- **Soil type**
- **Slope**
- **Fertility**
- **Tillage**
- **Crop history**
- **Incorporate field and plot border/buffer areas to ensure that treatments do not influence each other**
  - ✓ Do not use compacted end rows, fence line grasses, or field roads

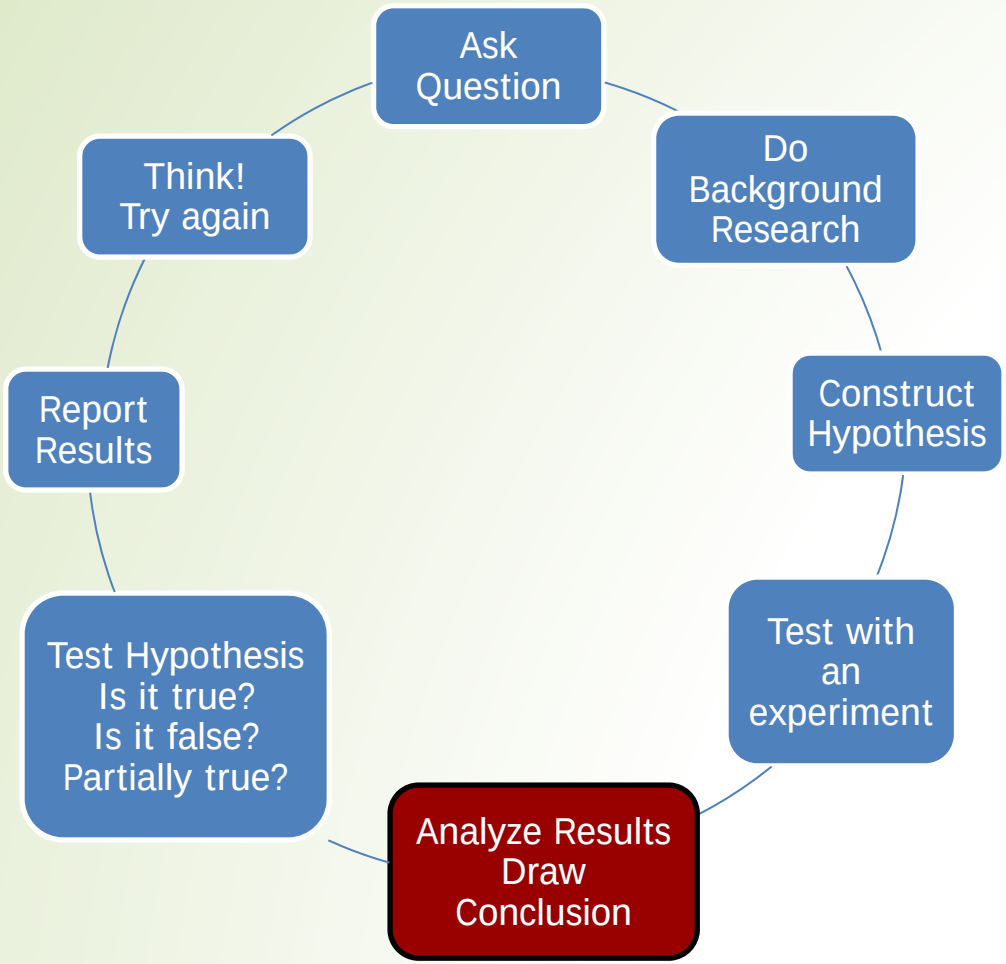


Records for research trials

- **Keep good records (diary of crop and weather conditions, photos)**
  - ✓ Identifying and labeling strips in the field
    - ❑ Mark strips in the field with stags or pvc flags (and GPS record)
    - ❑ Code the identification until all data are recorded
    - ❑ Uncode when results are summarized
  - ✓ Write it down (or electronic files)
  - ✓ Draw a sketch (map) of the experiment layout for later reference
- **Data collection for each experimental unit**
  - ✓ Use properly calibrated weigh wagon or yield monitor for harvest weights
  - ✓ Collect moisture and test weight
  - ✓ Record all information that may affect results (notes, soil fertility, plant height, insect thresholds, weed densities, planting and harvest populations, protein)



The Scientific Method



- **Analyze Your Data and Draw a Conclusion:**
  - ✓ Collect your measurements, analyze them to test if your hypothesis is true or false.
  - ✓ Focus on economics
    - ❑ What is the cost:benefit ratio?
    - ❑ Consider non-tangible benefits (i.e. soil quality, environment, etc.)
  - ✓ Often the hypothesis is false, and in such cases a new hypothesis is constructed starting the process over again.
  - ✓ Even if the hypothesis was true, it may tested again in a new way.
  - ✓ Statistically analyze data (determines probabilities that the differences were caused by treatments versus chance, random variation, or error)
  - ✓ Draw conclusions from more than one location and/or year

Experimental Error

- **Any measurement on an experimental unit has a certain amount of error associated with it.**
  - ✓ Error consists of variability among plots due to other, uncontrolled, trait influencing factors.
- **Statistics allow us to quantify and assess error.**
- **Error cannot be assessed with one measurement.**
  - ✓ Need multiple measurements (replications).

Error can be …

- **Human error in conducting the trial.**
  - ✓ Mistake in calculations (e.g. area of plot)
  - ✓ Wrong plot, transcription, misread scale
- **Variable soil characteristics within a field.**
  - ✓ Soil texture, drainage, compaction, elevation
- **Within-field variability for insect & disease damage, herbicide injury, weather, etc.**
- **From year to year, weather variability creates error, especially as it interacts with other factors.**

- **Your challenge is …**
  - ✓ To sort out the true yield effects of the treatments from those effects caused by error.
    - ❑ You can never be 100% certain that yield differences in a trial are solely due to the treatments being evaluated.
  - ✓ Fortunately, that's why statistical analysis was invented!

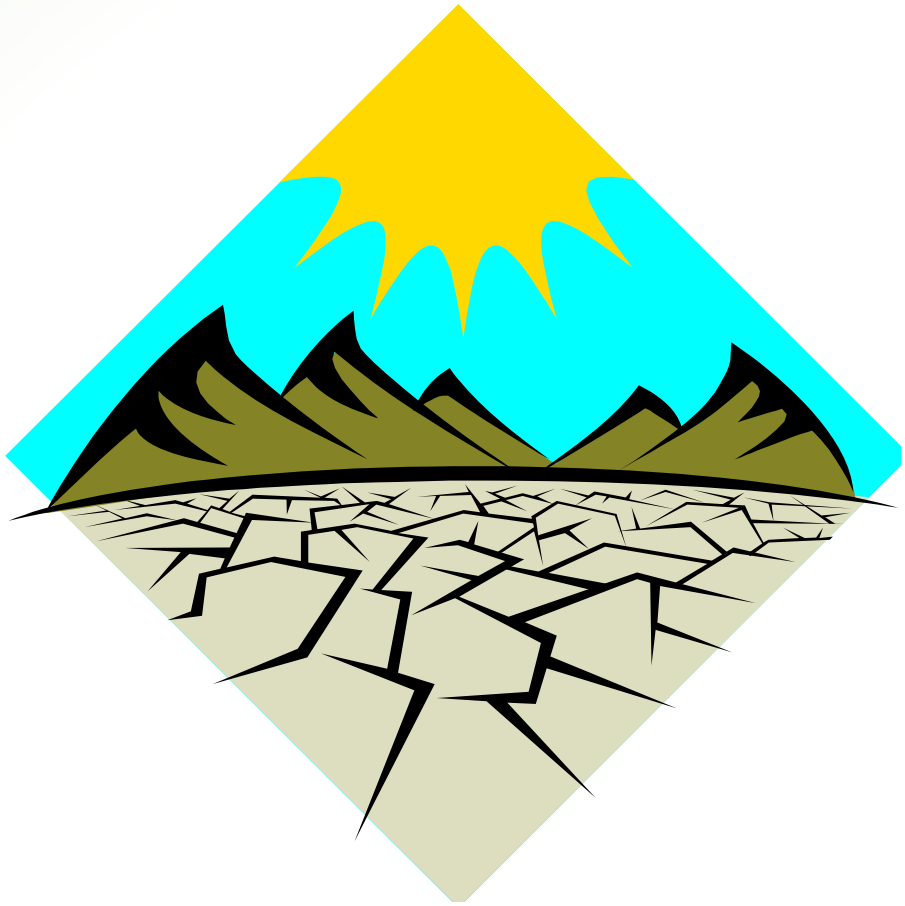
Sources of Error

- **Adjustments in Site-Years**
  - ✓ 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
  - ✓ If variability is expected to be high over both soils and years, then you will need a lot of sites and years

- **Genetics and Response Variability**
  - ✓ If variability is expected to be high over varieties/hybrids, you have a problem
  - ✓ We know that some inputs might well vary with genetics (e.g., Bt and insecticides):
    - ❑ Some may be predictable, like plant population and standability
    - ❑ Many possible interactions are not predictable, any more than responses to inputs in general
    - ❑ Most seed companies do not want to deal with inputs that depend on genetics: this can restrict sales, and it seldom has an identifiable genetic mechanism to breed for (or against)

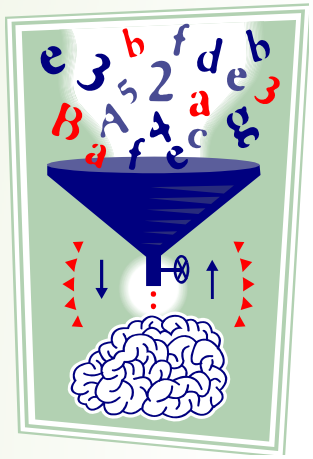
Should/Can We Ever Discard Data?

- Not unless we know for sure that what happened to make us want to throw out that location will never, ever happen again; weather is generally off-limits as a reason to discard, but manageable things, like stand, probably are not
- Maybe, if "uncommon" factors limited yields severely - say to 50 percent of normal
- Not unless it's done before harvest, or:
- We don't trust the data after it's been taken



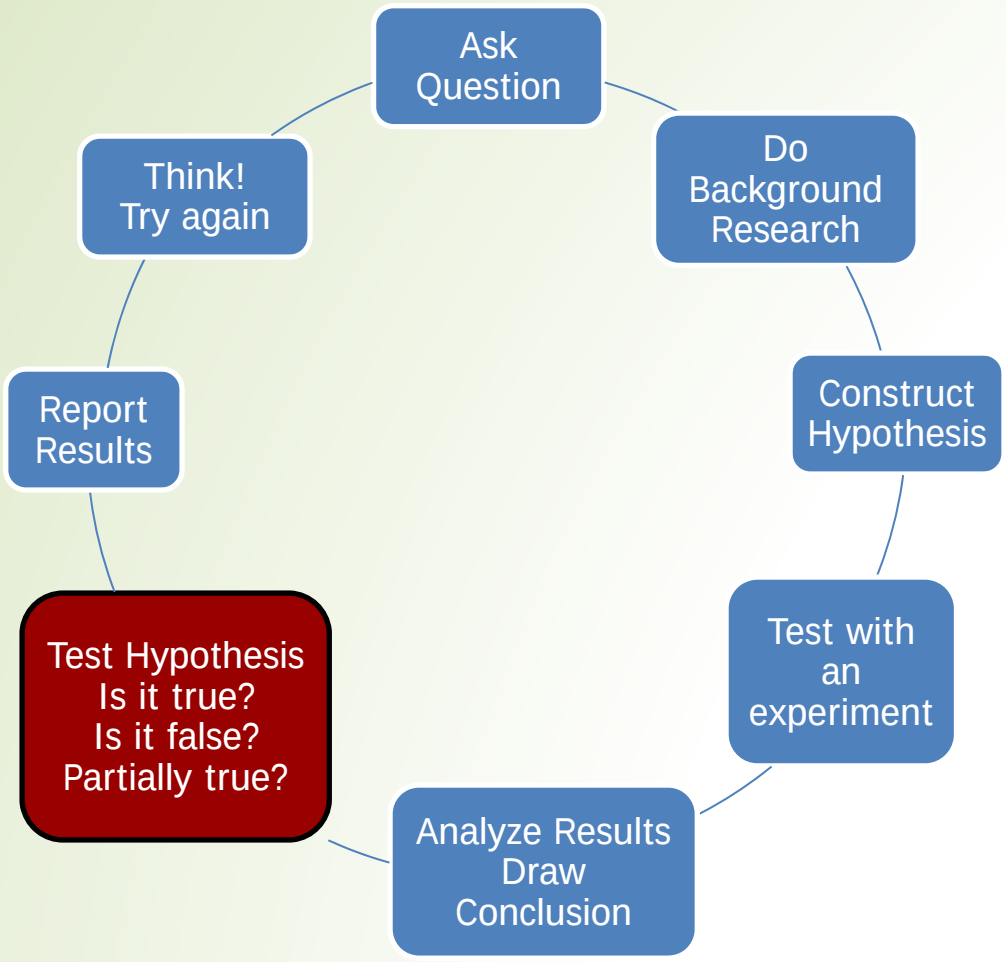
Statistics

- **"If the experiment is done well enough, there's not real need for statistics" – a famous statistician**
- **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



- **If we don't "need" stats, why bother?**
  - ✓ We have to use somewhere what we learned in school (it's what scientists do)
  - ✓ For individual (replicated) trials, stats provide an extra perspective on the results
  - ✓ Across site-years, stats add value to a prediction
- **Just because stats are done does not make it good research**
- **Stats do NOT substitute for the large amount of data (site-years) that good research always requires**

The Scientific Method



- **Testing the Hypothesis (H_o) –**
  - ✓ In Words: The population mean of one treatment = the population mean of another treatment
  - ✓ In Formula: H_o: $u_1 = u_2$
  - ✓ Fact: The hypothesis may be True or False

Hypothesis Testing

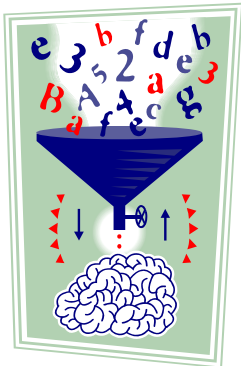
$H_o: u_1 = u_2$

- "Reject the Null Hypothesis": Conclusion is there is a difference between the treatments for the trait measured.
- "Do not reject the Null Hypothesis": Conclusion is there is no difference between the treatments for the trait measured.

Real condition of the Null Hypothesis	REJECT	DO NOT REJECT
TRUE ($u_1 = u_2$)	Type I Error	No Error
FALSE ($u_1 \neq u_2$)	No Error	Type II Error

Type I Error

- **The act of rejecting a true null hypothesis:**
  - ✓ This leads to the conclusion that a treatment effect exists when in fact none does
  - ✓ Probability of committing a Type I error = *a* ("alpha")
    - ❑ Experimenter (You) sets alpha
    - ❑ One normally sets alpha low so confidence is high when the null hypothesis is rejected.
    - ❑ When a treatment doesn't cost more than another (e.g. choice between two hybrids) then alpha can be increased
    - ❑ When a treatment is very expensive, then alpha should be low to increase our confidence
  - ✓ Confidence of the statistical decision is 1 – alpha
  - ✓ "False positive"
- **Consequence is we might adopt a practice that does not pay off.**



Type II Error

- **The act of not rejecting a false null hypothesis:**
  - ✓ This leads to the conclusion that a treatment effect does not exist when in fact one does
  - ✓ Chances of committing a Type II error increase as alpha is decreased
  - ✓ "False negative"
- **Consequence is we don't adopt a practice that might have paid off.**

The t Test

$$t = \frac{\text{Difference between two treatment means}}{\text{Standard deviation of the difference between two means}}$$

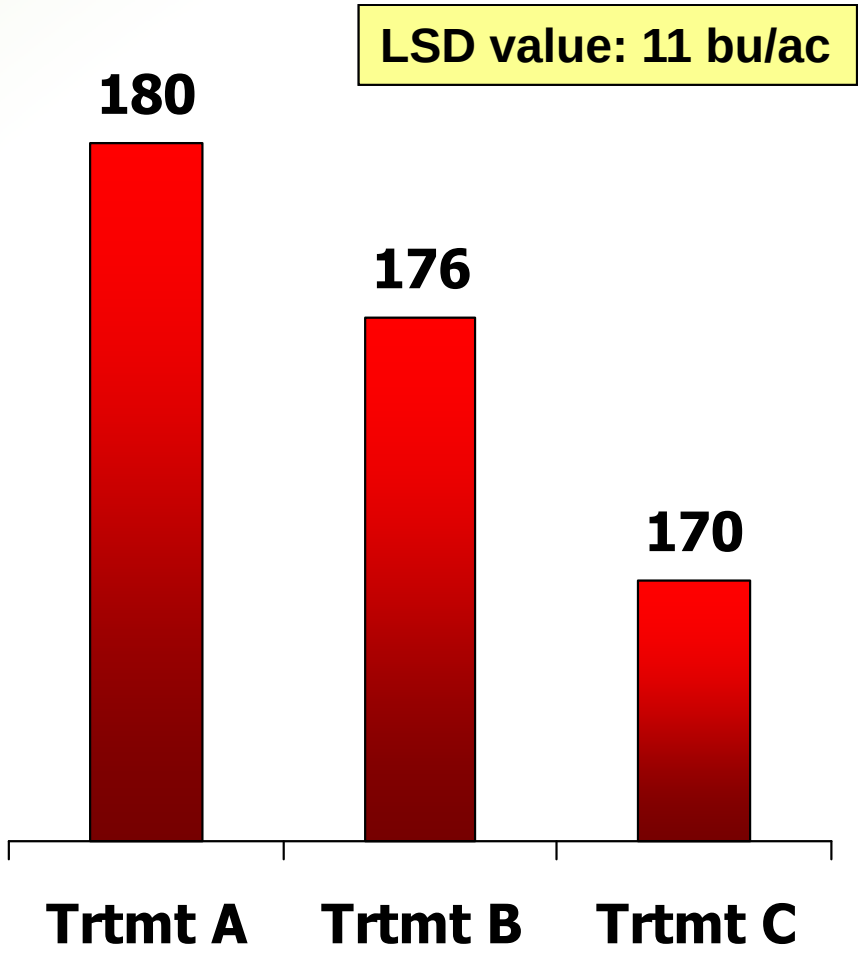
- **When the number of replications is increased:**
  - ✓ Get a better estimate of the treatment means
  - ✓ Reduces the size of the standard deviation of the difference between two means
  - ✓ Increases the size of the t value
  - ✓ Reduces the tabled value for significance at any level of confidence
  - ✓ Increases the chance of detecting treatment differences.

Comparing treatments…

- **Statistical analysis allows for the calculation of a value that is used to estimate whether the difference between two treatments is due to treatment effects or is simply error.**
  - ✓ Least Significant Difference (LSD)
  - ✓ If two treatment means differ by more than the LSD value, then conclude that the difference is due to treatment effects AND that similar results will be observed in the future.
  - ✓ If two treatment means differ by less than the LSD value, then conclude that the difference is due to random chance or error AND may not be observed again in the future.

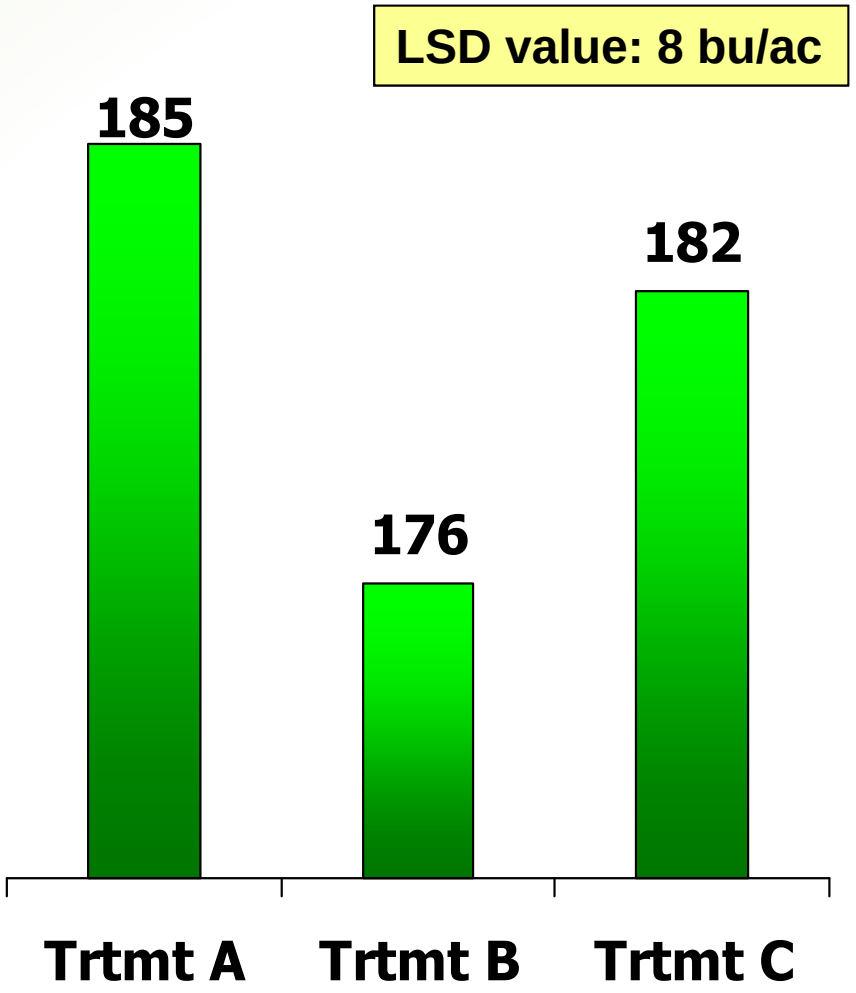
Example of using LSD values

- **Conclusion: None of the pairs of treatment means differ by more than the LSD value, so you must conclude…**
  - ✓ Treatment effects are similar,
  - ✓ Observed differences likely due to error, and
  - ✓ Observed treatment trends would NOT repeat in subsequent trials.



Example of using LSD values

- **Conclusion: Treatment A significantly out-yielded Treatment B and will likely do so again in future field trials.**
  - ✓ But Treatment C did not yield significantly different than the other two.

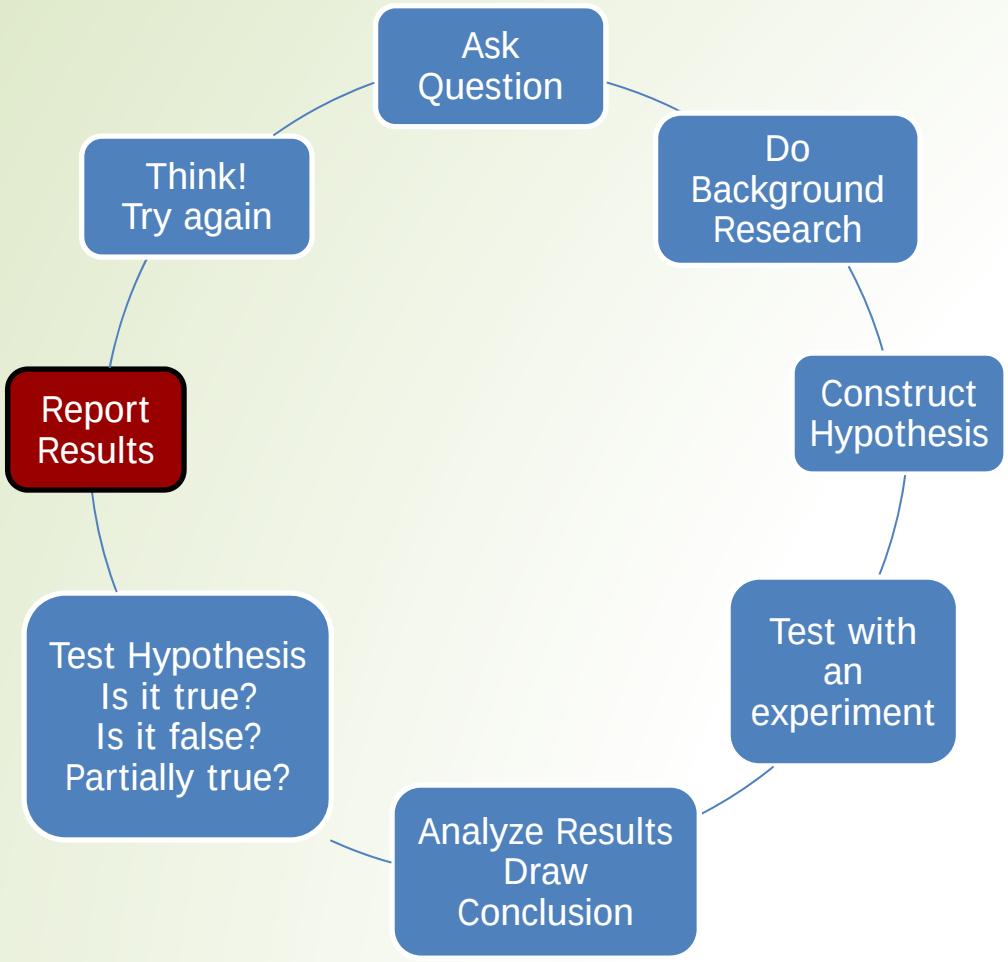


Consequences of Decisions

- **Determine there is a difference between treatments:**
  - ✓ If there really is, it's a CORRECT decision
  - ✓ If there isn't, the consequence is the cost of the treatment lowers profitability
- **Determine there is no difference between treatments:**
  - ✓ If really no difference, it's a CORRECT decision
  - ✓ If really a difference, the grower loses potential profitability

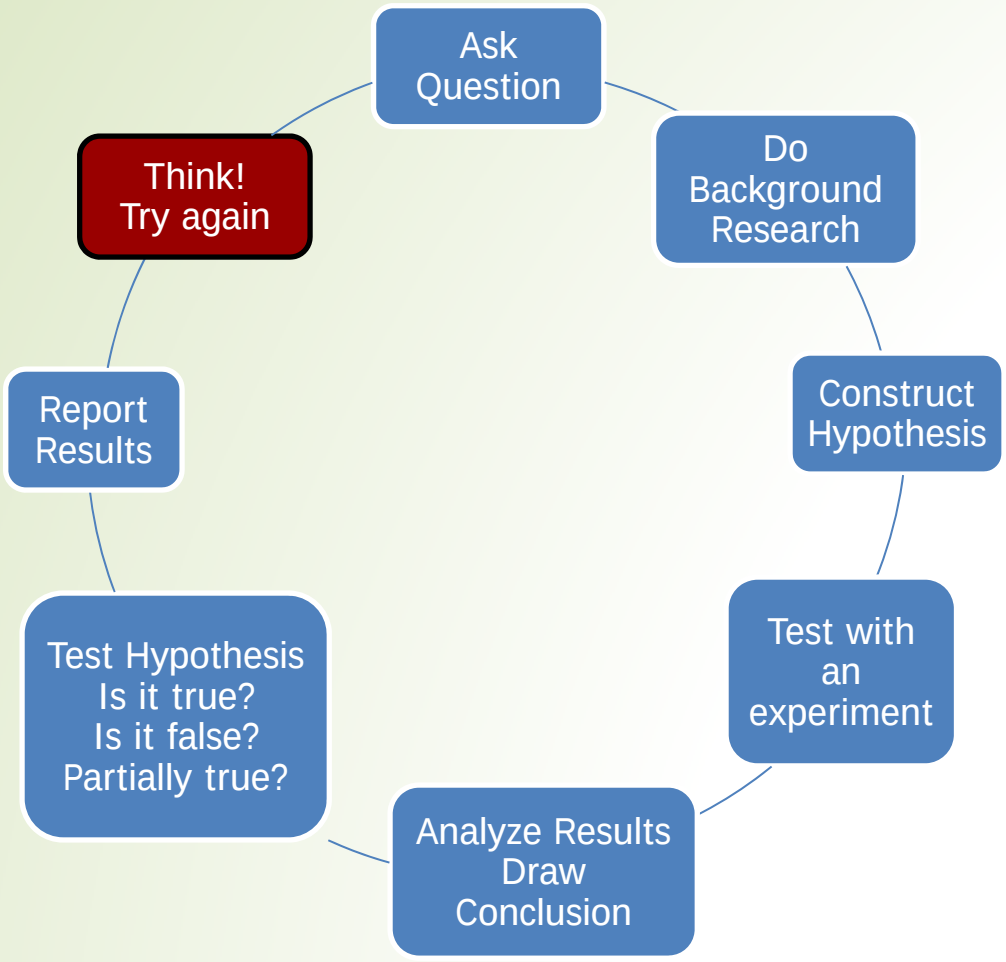
- **Grower Production Decisions**
  - ✓ Should be GOOD decisions
    - ❑ Correct decisions
    - ❑ Profitable decisions

The Scientific Method



- **Communicate Your Results:**
  - ✓ To complete your project communicate your results to others in a scientific journal or by presenting results on a poster at a scientific meeting.
- **Clearly written text**
- **Graphics**
  - ✓ Pictures
  - ✓ Tables
  - ✓ Graphs

The Scientific Method

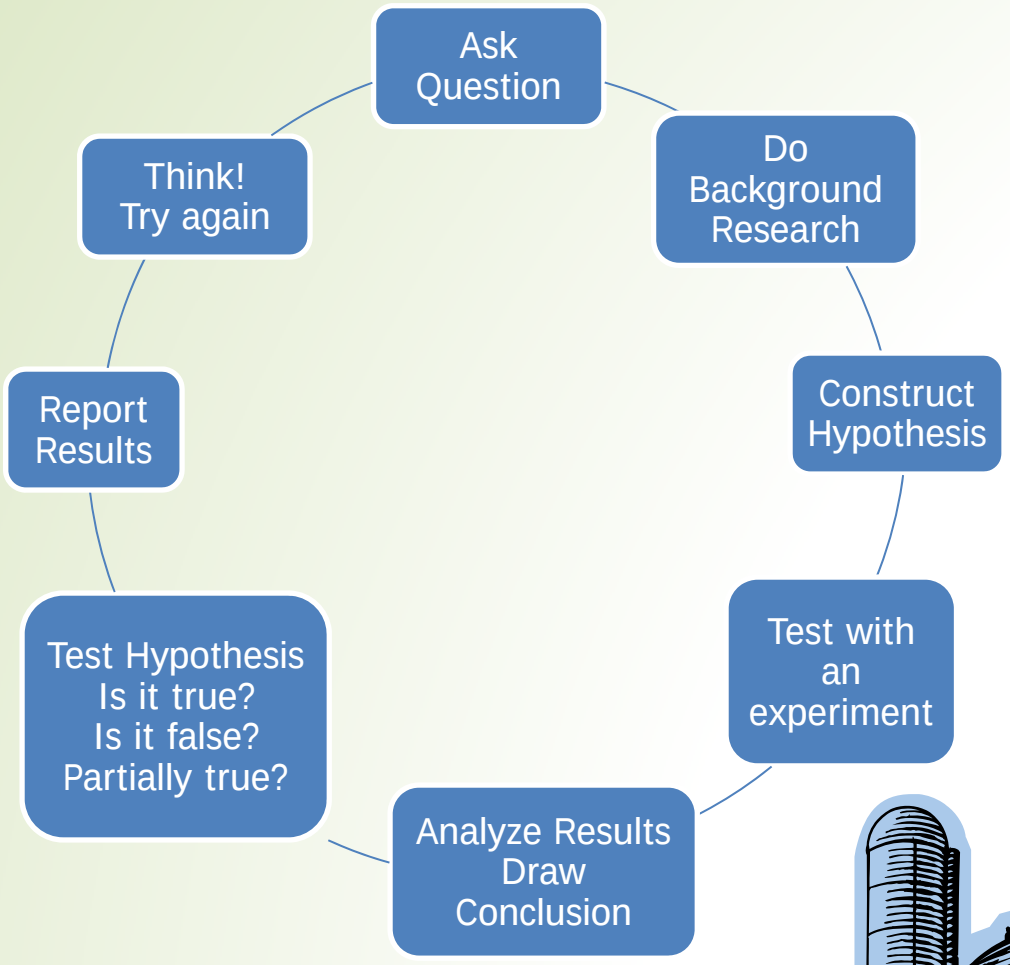


- **Keep trying. For most questions, there is rarely a final answer.**
- **If enough experiments are performed, then research can lead to science (knowledge), but only if it's done well.**
  - ✓ Done "well" means using accepted scientific methods, which often include statistics
  - ✓ If not done well, can lead to more harm than good.
- **What part of the question hasn't been examined before?**

Scientific ideas based on the amount of confirmed experimental evidence

- **Hypothesis**
  - ✓A statement that uses a few observations
  - ✓An idea based on observations without experimental evidence
- **Theory**
  - ✓Uses many observations and has loads of experimental evidence
  - ✓Can be applied to unrelated facts and new relationships
  - ✓Flexible enough to be modified if new data/evidence introduced
- **Law**
  - ✓Stands the test of time, often without change
  - ✓Experimentally confirmed over and over
  - ✓Can create true predictions for different situations
  - ✓Has uniformity and is universal
- You may also hear about the term "**model**." A model is a scientific statement that has some experimental validity or is a scientific concept that is only accurate under **limited situations**. Models do not work or apply under all situations in all environments. They are not universal ideas like a law or theory.

Conclusions



- **"Planned comparison:" careful choice 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**
- **"Adequate" number of fields and years (Replication and Randomization)**
- **"Prediction:" a statement of likelihood of expected results from use of a particular input**

Final Point:

- **"We know our product works, but you have to know what you're doing/choose the right conditions/use it properly in order to make it work for you."**

- **<u>Likely translation</u>: "Our research, if we did any, shows great variation in response, sometimes including yield loss. We don't know how to tell you how to make positive returns more likely, but we can certainly tell you how you could have done a better job when it doesn't work (though we'd rather you not know when that happens). We are working on an improved formulation. Thank you for your business, and good luck."**

- **"There is a sucker born every minute." - P.T. Barnum**

Thanks for your attention!
Questions?

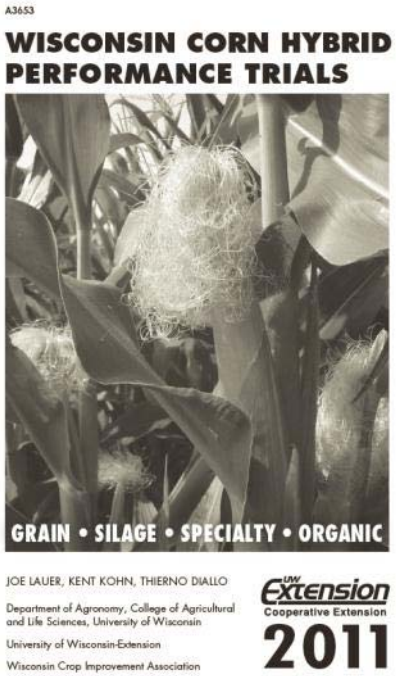
Website: http://corn.agronomy.wisc.edu

To subscribe to season updates:
http://corn.agronomy.wisc.edu/Web/Subscribe.htm

The Corn Agronomy Blog
http://WiscCorn.blogspot.com/

Twitter
http://twitter.com/#!/WiscCorn

Google+
https://plus.google.com/118235681838809938708/about



http://corn.agronomy.wisc.edu/HT/

http://corn.agronomy.wisc.edu/PEPS/