TIPS FOR DESIGNING BEE RESEARCH PROJECTS

Randy Oliver ScientificBeekeeping.com

Revised 8 November 2017

Beekeepers are known for having curious and experimental minds. Since factors affecting beekeeping are continually changing, new unanswered questions are bound to arise; the beekeeper "citizen scientist" can often answer them himself by performing a well-designed experiment (and then share those results to the benefit of everyone). But for the results of any experiment to be meaningful, it is important that the experiment follow certain scientific principles. Following are some tips for running a successful experiment.

RESEARCH OBJECTIVE

You can perform a "trial" to *test the hypothesis* that this "treatment" (e.g., some sort of management technique or application of a medication) will result in a certain effect (e.g., lower mite level, greater honey production) [1]. Or, you could perform a "study" to learn something about bees or beekeeping (e.g., whether feral colonies have a higher incidence of X than do managed colonies).

#1 point—Ask yourself, what is the single specific question that you are trying to answer with this experiment? ***Write that question on the wall, and eliminate anything from your experimental design on that doesn't have to do with specifically answering that question*** (and make sure that your experimental design will unambiguously answer the question). Fill in the following blanks before you go further (take some time to think them through):

Your question _____

Your hypothesis _____

The treatment (variable) to be tested _____

The predicted effect _____

(A trial tests for supporting evidence for the hypothesis that there is a measurable effect of the treatment vs the "null hypothesis" that there is no effect from the treatment, as evidenced by no difference between the treatment and control groups at the end of the trial).

BE REALISTIC

It is a common mistake for inexperienced researchers to bite off more than they can chew. Limit your experiment to one variable (such as the type of mite treatment or management technique applied), and minimize the number of metrics (such as adult bee population, amount of brood, weight gain, varroa counts, nosema counts, etc.). **The three metrics with the most practical application to beekeepers are colony strength (number of frame interstices filled with bees), weight gain, and varroa alcohol wash counts.**

BACKGROUND RESEARCH

Perform an internet (e.g. Google Scholar) literature search of previous research, and read any previous studies that are applicable to this treatment. Such studies should also give you a model for your write up. Speak to researchers familiar with this topic.

SCIENTIFIC PAPERS

Your background research will introduce you to the format of scientific papers, which follow the typical format of:

Abstract

Title, Author, and Date

Introduction

Materials and Methods

Results

Discussion

Acknowledgements and Funding Source

Literature Cited

It would be wise to plan and execute your experiment to fit the above format (see the example at the end of this doc for more details). The scientific model is well proven, and allows you to share what you learn with others. You may also wish to present the results to one or more groups of beekeepers. Knowing the above, may I strongly suggest that you:

PLAN AHEAD AND WORK BACKWARDS

Here's a tip: **first run the experiment backwards in your head, and then on paper**. Begin by imagining exactly how you plan to present the results to an audience in a manner so that they can fully understand the significance of your experimental findings. Are you going to use scatter, line, or column graphs? Plan **and draw on paper**, the exact sort of graphs you will be using to present your findings, and how the results would look if they either support or disprove your hypothesis. **Then work backwards to set up every detail of the experimental design to produce those graphs** (e.g., what will be your x and y axes?)**.** This approach will help to save you from kicking yourself later on for wishing that you'd done something differently! However, it is unlikely that any experiment will go exactly as planned. It is the rare experiment that I wouldn't do somewhat differently the second time—read the Discussion section in published similar studies to learn from others' errors and experience. See end note [2] for further discussion.

MATERIALS AND METHODS

This is how you actually run the experiment. You are going to try to determine whether your *treatment(s)* cause any measureable differences in the test group(s) compared to an identical *control* group (which should receive a sham treatment, e.g., plain syrup, or a simple opening and smoking).

Variables--*you want the specific treatment to be the only "variable" between the groups. Every colony in the entire experiment must receive identical location, handling, feeding, etc. other than the specified "treatment."* Or if you are doing a study, how exactly did you sample and why; how did you avoid bias; how were the samples analyzed?

The M&M should be written with enough detail that another researcher could exactly duplicate your experiment (for replication).

Relevance—to be relevant to "real-life" beekeeping, the size, condition, feeding, and environment of the test colonies should approximate that of normal beekeeping practices.

The best design is generally randomized block—if there are 3 treatments and a control, then group colonies into groups of 4, and randomly assign all 4 treatments in the group (use random.org; use different randomization for each group). This will minimize location, pathogen transmission, and orientation effects.

Number of colonies (generally a minimum of 12 in each group). One of the toughest problems in bee research is the intrinsic variability between colonies. This "noise" often makes it difficult to tease out any effects due to treatment. **The more colonies in the trial (the larger the "n") the better.** Practically, I wouldn't waste my time with fewer than 12 colonies per group; 18-36 would be even better.

The Control Group(s)

Generally you will be testing one or more treatments (the fewer the better, unless you're looking for a dosage effect) vs. a "Negative Control" group that is managed in exactly the same way, but without the specific treatment variable.

This means that each time you apply a treatment to the Test group, that you must also smoke, open, disturb, or feed the Control in exactly the same way.

If the treatment is a food additive added to sugar syrup, you need to feed an identical amount of identical untreated syrup to the Control hives. If the treatment is a miticide applied in waxed paper, you need to apply a "sham treatment" of waxed paper alone to the Control group.

For example, when I tested OA/glycerin shop towels, EPA wanted results from the Treatment group (OA/gly towel), plus a "solvent control"—a towel with glycerin alone, plus a "sham control"—a towel alone.

If you're testing a miticide, you may wish to run a "positive control" of a miticide treatment that you know from experience will work—so you'll run a Positive Control group, a Treatment Group, and a Negative Control (or perhaps sham treatment) group.

Incubator trials—consult with advisors for best practices. I will soon be publishing plans for a homemade incubator as well as inexpensive cup cages.

Replication—your findings will be considered more "robust" if replicated. A single replicate is better than nothing, but it is not at all surprising to get a different result if you repeat the experiment. Typically a trial is replicated three times in order to confirm you get the same result each time (you may not). Each repeat of the experiment is called a replicate. If you ran the trial with test groups at two different locations, or repeated the trial at a different time, or if two different beekeepers duplicated the same experiment, then that would be two replicates. Statistical analysis of a single test group is considered by some as pseudoreplication, and thus lacking strong validity.

Source of colonies (queens, etc) should be randomized per treatment group. In some trials it may help to found all colonies with sister queens so as to minimize variation. If queens from different sources are used, best to use each source in one randomized block. Also make sure that the combs, amount of brood, and type and color of the hive bodies are all similar.

Grouping of hives—the further apart the hives the better, in order to minimize the "drift" of (especially, infected) bees from hive to hive. The distance factor must be weighed against changing the variable of "location" if such spacing puts some of the hives, say, into shade or some other change in the environment.

Hives should not be placed in a line (in order to minimize the effect of drifting bees) and should all have the same solar, wind, etc exposure, or exposures rotated. Use landmarks (shrubs or yard trash) in the apiary to help the bees to orient back to their proper hive. See [3] for cautions.

Generally, each pallet of hives receives one treatment, due to drift from hive to hive on that pallet (best to turn the 4 hives with their entrances facing in 4 different directions). Pallets should be placed so that each treatment group has pallets facing each compass direction.

Randomization--Treatment or control should be determined by flip of a coin or random number generator (random.org). This decision must be entirely independent of the Investigator, so that there is no bias, and should be done only after all other colony preparations have been done. Note: "arbitrary" or "haphazard" assignment of treatment is not random! One problem that I've had with the random assignment of treatments is that sometimes the coin flip or random number generator (http://random.org/) puts all hives receiving a single treatment all in one location. This is where a randomized block design helps. If you are testing only two groups (Treatment vs Control) then call each pair of hives a block, and randomly assign treatment to one of each pair.

Systematic assignment of treatment—you may instead treat every other colony, or some other systematic assignment of treatment. If so, rotate or randomly assign the order for each group (e.g., rotate the solar exposure of hives receiving each treatment—don't have all the test colonies facing the same direction). Another method is to first number and then grade all colonies for the main metric (say mite count). Then use a spreadsheet to rank all colonies by mite count from highest to lowest. Then assign treatment alternately down the list (I like this method, since it is applies the treatment equally across the range of starting points).

If you are trying to determine, for instance, the effect of a miticide, I prefer to ***stratify*** my assignment of treatments by the starting mite counts of the hives. Take mite counts from all colonies on Day 0, rank (sort in Excel) the hives by mite count, and then assign treatments alternately down the ranking (from those with the highest starting counts to the those with the lowest—this will ensure that each treatment is assigned to the same range of mite infestations). I.e., if you have 3

treatments and control, then randomly assign those 4 treatments to the 4 most-infested colonies, and continue likewise to next most-infested 4 colonies, and so on down the list.

Number the hives—you don't want wind, rain, or skunks to mess up your numbering (practical experience here). Firmly affix weatherproof numbers [4] to each hive.

Label *everything!* Don't trust your memory. Label every hive, sample bag or jar, data sheet, bottle of treatment, etc. Warning: alcohol will dissolve many inks; if you are dealing with samples preserved in alcohol, mark your labels with pencil rather than ink (many learn this the hard way).

Bias and Blinding—It is virtually impossible to avoid bias. Go out of your way to "blind" yourself as to which colonies were receiving which treatment *until all the data is collected and analyzed*. Ideally, the Investigator, field inspectors, or graders will be completely "blinded" as to which colonies received which treatments. **Do not label the hives with indication of treatment received *when they are being graded*—the inspector (grader) must not be aware of which hives received which treatment.**

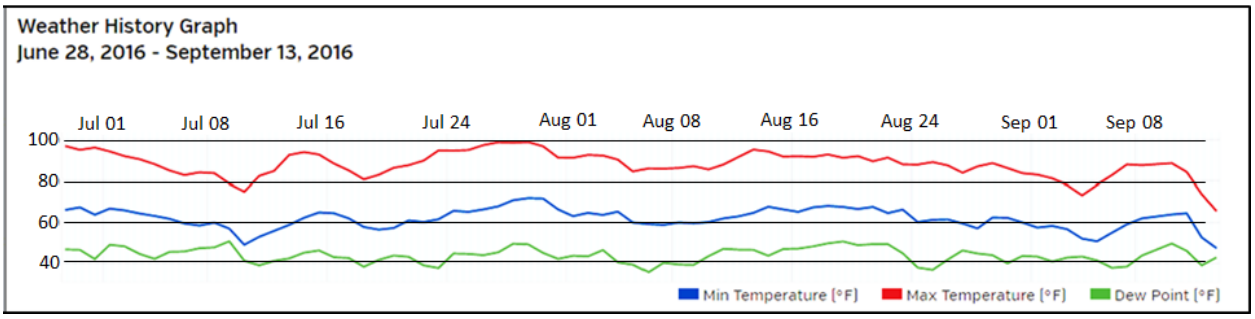
Log book.--A chronological log book, in ink, must be kept of every detail and action taken in the yard. *Many important discoveries are made after the fact by reviewing the log book.* Never trust your memory! The printing up of a log entry sheet in advance will help you to clarify which data to collect, and the form in which it should be recorded. I print the attached data entry sheet **on heavy cardstock**, and fill it with a pencil or waterproof pen in the field. All raw data should be recorded on paper, in case of computer mishaps. As soon as possible, enter the data into an Excel spreadsheet.

Treatments--Specify in advance how treatments will be delivered, and how control colonies will be given sham treatments. If one smokes and opens the treatment hives, then one must do the same for the control hives; if one feeds a treatment in syrup, then the control hives must receive an equivalent amount of untreated identical syrup.

Effect--How will you measure any effect? Examples would be regular weighing of hives, cluster size estimate, mite counts, nosema spore counts or prevalence determination, etc. Time points—Assign approximate time points for taking measurements in advance. Day 0 is the starting day (after equalization and set up)—the baseline to which future measurements will be compared. Collect data at *at least* three time points (four would be better) so as to be able to detect trends.

Grading--Specify what metrics will be measured (number of frames of bees, hive weight, etc) and how they will be measured, and at what time (Day 0, midpoints, End). The more time points at which measurements are taken, the better that a trend can be detected (generally plan at least three data time points).

The weather, dates, and location—I am often frustrated by scientific papers that don't give critical details on these variables. Be sure to include the full date at each step, and colony condition (spring buildup, honeyflow, late-summer dearth) in order that others can see under what conditions your findings are relevant. Download the weather history from www.wunderground.com, and place it in the M&M section.



Don't cherry pick your data—there may be annoying outliers or results that don't make sense to you. Exclude or censor data with the greatest caution! Decide, in writing, in advance, which criteria you will use to remove colonies (censoring) from the experiment—queenlessness, disease, starvation, etc.

Removal of colonies from the experiment—allowing colonies to collapse may flood adjacent hives with parasites. You should physically remove any colony from the test yard(s) when it has "no further chance to survive by itself with regard to the local conditions." Remove the censored colonies after dusk, or when it is too cold for any bees to be in the field.

Statistics—you should run your design by a statistician prior to beginning the experiment! The simplest results would be date entered into an Excel spreadsheet, means and medians calculated, and standard errors of each mean indicated on a column (bar) or line graph [5]. Scattergrams, histograms, and line charts can also be used. Perform the Student's T test to determine statistical significance (http://studentsttest.com/) for data that follows a "normal" (bell curve) distribution, or the Mann-Whitney test if it does not (http://www.socscistatistics.com/tests/mannwhitney/Default2.aspx). Use the Henderson-Tilton's formula to determine treatment efficacy (http://www.ehabsoft.com/ldpline/onlinecontrol.htm#HendersonTilton).

You can also run an ANOVA for comparison of means, but that's a bit more complicated than I wish to go here. In any case, don't let the statistics scare you—collect the data, and then have someone help you with the stats.

However, of even more import than statistical significance is the *size* of the effect—**was there enough difference due to the treatment to be of biological or practical relevance?** Don't rely upon statistics alone—look carefully at the bees during the course of the experiment, and look for patterns and trends in the data. The human eye/brain is wired to spot trends and patterns, which is why graphing the data is so important; statistical tests for probability check our inclination to "see" patterns where none actually exist.

In order that your experimental results can be compared to those of others, I strongly suggest that you download and read the appropriate sections in **The COLOSS BEEBOOK Volume 1, Standard methods for Apis mellifera research: http://www.coloss.org/beebook/I/introduction**

You do not necessarily need to follow their suggested guidelines (e.g. I prefer the quick and non-intrusive cluster strength grading method to the suggested Coloss methods), but you had better have good reasons for not doing so.

THE PROTOCOL

Before you start, write out the exact protocol that you plan to follow, ***in minute detail***. You can later "deviate" from the protocol if necessary, with an explanation of why. Include exactly how you will set up the experiment, apply treatments, care for the bees during the course of the experiment, and the time points at which you will collect data, and exactly which data you are going to collect, and exactly how you are going to measure it (e.g., how many bees in a sample, taken from where in the hive, at what time of day, and on what dates). Writing a detailed protocol in advance can save you a lot of decision-making in the field, and grief in the end when you say, "I wish that I would have thought that out before I started!" A few hours of careful planning can save you months of wasted effort in the field—researchers (including myself) have tons of unpublished research because they didn't plan correctly (and unfortunately, tons of research that should never have been published (especially by grad students) gets published anyway).

DATA COLLECTION

Record all your data on your paper data sheets while in the field. I suggest printing your raw data sheets on heavy cardstock paper, as field work is rough on letter paper. A sample sheet that you can copy is below.

Experiment Name

_____

Date_____

Yard_____

Investigator

_____

Time point _____

Type of data (measurement)

_____

Field Notes:

Hive	Date	Date	Date	Date
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				

ANALYSIS OF RESULTS

Type your raw data from the field data sheets into an Excel file. After doing so, read the Excel columns to a helper holding the raw data sheet to confirm that you didn't make any errors in entry. Save your raw data in a Raw Data tab, and lock it to prevent any changes.

You can then copy the raw data to a second tab for sorting and analysis (***never mess with your raw data in the Raw Data tab***). Below is an example of worked data from one of my experiments (data from only the top two treatment groups shown).

2015 late summer treatment, sorted and worked data										
			Date	7-Sep	17-Sep	16-Oct				23-Nov
			Day	-10	0	29				67
		Trtmt	**hive no**	**early Sep count**	**Sept 17 count**	**16 Oct count**	**16 Oct notes**			**23 Nov count**
		A	**12**	13	23	4				10
		A	**16**	28	19	0				
		A	**19**	35	31	3				0
		A	**24**	15	5	0				1
		A	**37**	19	19	7				4
		A	**42**	18	0	0				1
		A	**44**	11	21	0				0
		A	**51**	9	14	0				1
		A	**54**	37	22	3				2
		A	**57**	58	41	1	not all Apiguard removed			2
		A	**61**	62	70	2				1
Start count	12	A	**70**	12	18	0		H-T efficacy	% reduction	1
End count	11		mean	26.41667	23.58333	1.666667		99%	93%	2.090909
			median	18.5	20	0.5				1
			SEM	5.249038	5.214284	0.643538				0.857635
		AA	**1**	70	51	7				6
		AA	**5**	9	13	0				2
		AA	**22**	15	21	1				2
		AA	**23**	10	19	3				5
		AA	**32**	22	44	6				10
		AA	**39**	20	18	11	strong			25
		AA	**41**	11	9	11				23
		AA	**48**	22	23	3				8
		AA	**60**	21	34	0				5
		AA	**63**	13	19	4				17
		AA	**66**	8	10	2				6
	12	AA	**68**	50	48	3				
	11		mean	22.58333	25.75	4.25		97%	83%	9.909091
			median	17.5	20	3				6
			SEM	5.399857	4.269563	1.094926				2.443747

Excel will calculate means and median values. For the Standard Error of the Mean (SEM), you must enter the formula:

=STDEV(E6:E17)/SQRT(COUNT(E6:E17)) E6:E17 in this case indicates the range data for which you wish to calculate the SEM (column E, rows 6 through 17).

For efficacy calcs, I used the Henderson-Tilton's formula:

Henderson-Tilton's formula

$$\text{Corrected \%} = \left(1 - \frac{\text{n in Co before treatment} * \text{n in T after treatment}}{\text{n in Co after treatment} * \text{n in T before treatment}}\right) * 100$$

Where : n = Insect population , T = treated , Co = control

Here's an example from Excel: Efficacy =(1-(E6*F33/(F6*E33))). The $ signs lock the Control counts.

MATERIALS AND METHODS

If you are going to use abbreviations for the treatments, clearly show them in a table, not just in the text of the write up.

2015 Test of late-summer treatments (n = 12 ea group)	
Group	**Treatment**
Atop	50g Apiguard on top, in a rim, 2x (12.5 g thymol ea dose)
Amid	25g Apiguard in the broodnest, 2x (6.25 g thymol ea dose)
Tpad	6.25g total thymol in glycerin on 2 facial pads in the broodnest, 2x
MM	Single MAQS in the broodnest, 2x
MT	1 MAQS, followed by 25g Apiguard in the broodnest
C	Controls—no treatment

RESULTS

The **Results** section is where you present your results. Use tables and graphs (Table 1, Fig. 1). [Tables and figures must be cited somewhere in the text of the article.] ***Do not discuss your <u>interpretation</u> of the results yet.***

The results can be presented in a table, as below:

Table 1. Results of Late-Summer Varroa Treatments					
Date	Prelim	17-Sep	16-Oct	23-Nov	
Day	-10	0	29	67	
Treatment Group	Mean Mite Counts				Day 29 Efficacy
50g Apiguard under lid (2x)	26.4	23.6	1.7	2.1	99%
25g Apiguard between 2x	22.6	25.8	4.3	9.9	97%
6.25g thymol pad 2x	24.0	22.5	7.7	7.4	95%
Sgl MAQS 2x	23.7	20.7	0.8	2.4	99%

1 MAQS then 25g Apiguard	24.0	19.3	1.2	2.9	99%
Controls	8.6	2.9	17.9	15.3	0%

However, results are generally best presented in graphical form, since the human brain recognizes patterns in graphs far better than from data presented in tables. There are three main types of graphs that you can use—column, line, or scatter. The above data is best presented in a column graph (taken from http://scientificbeekeeping.com/a-test-of-late-summer-varroa-treatments/).

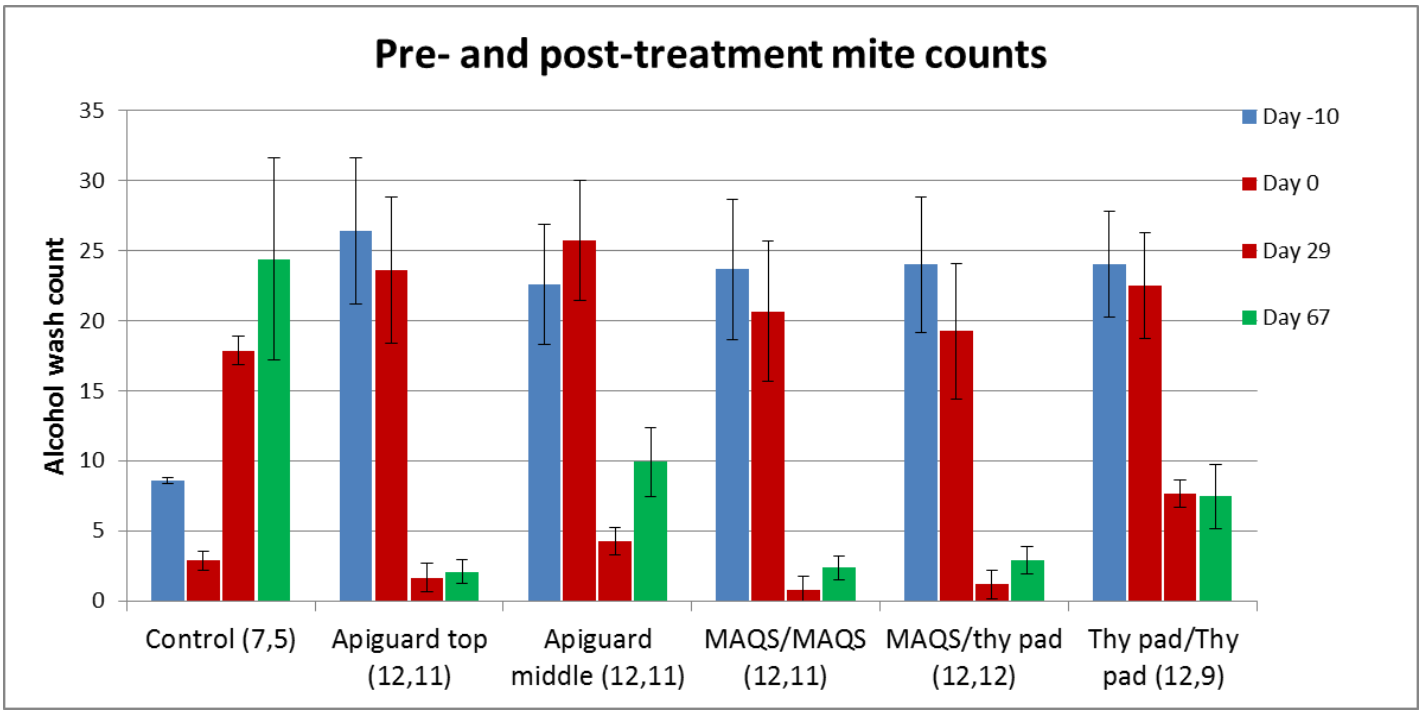
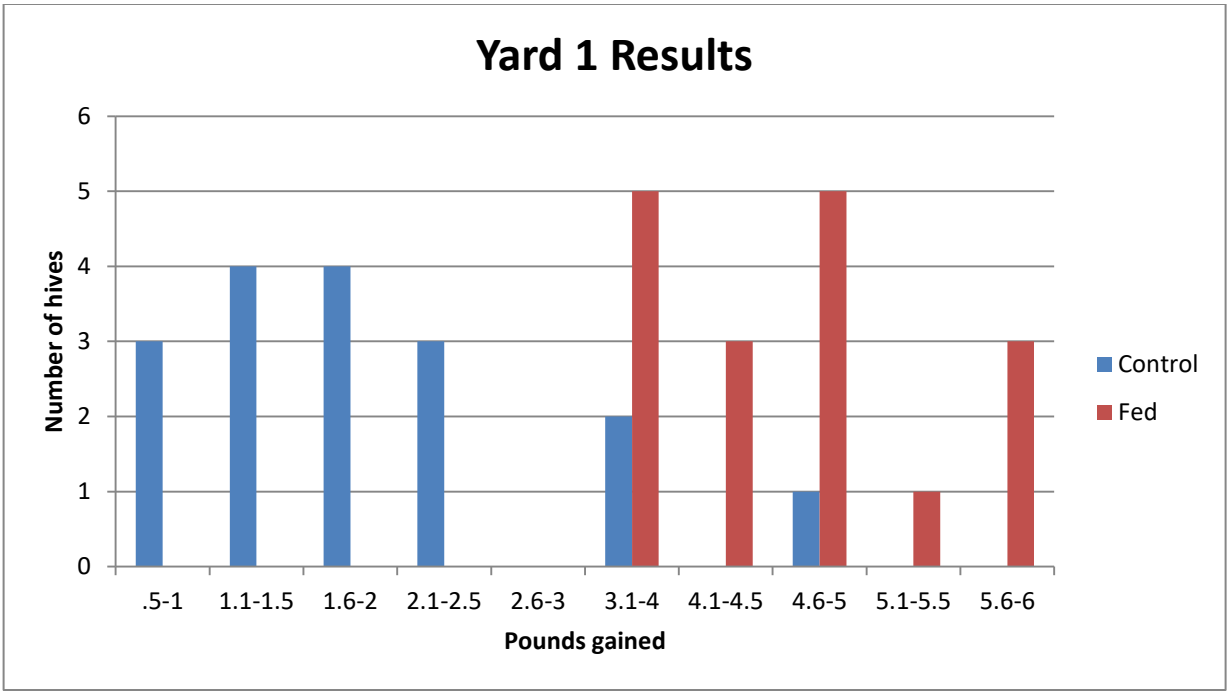


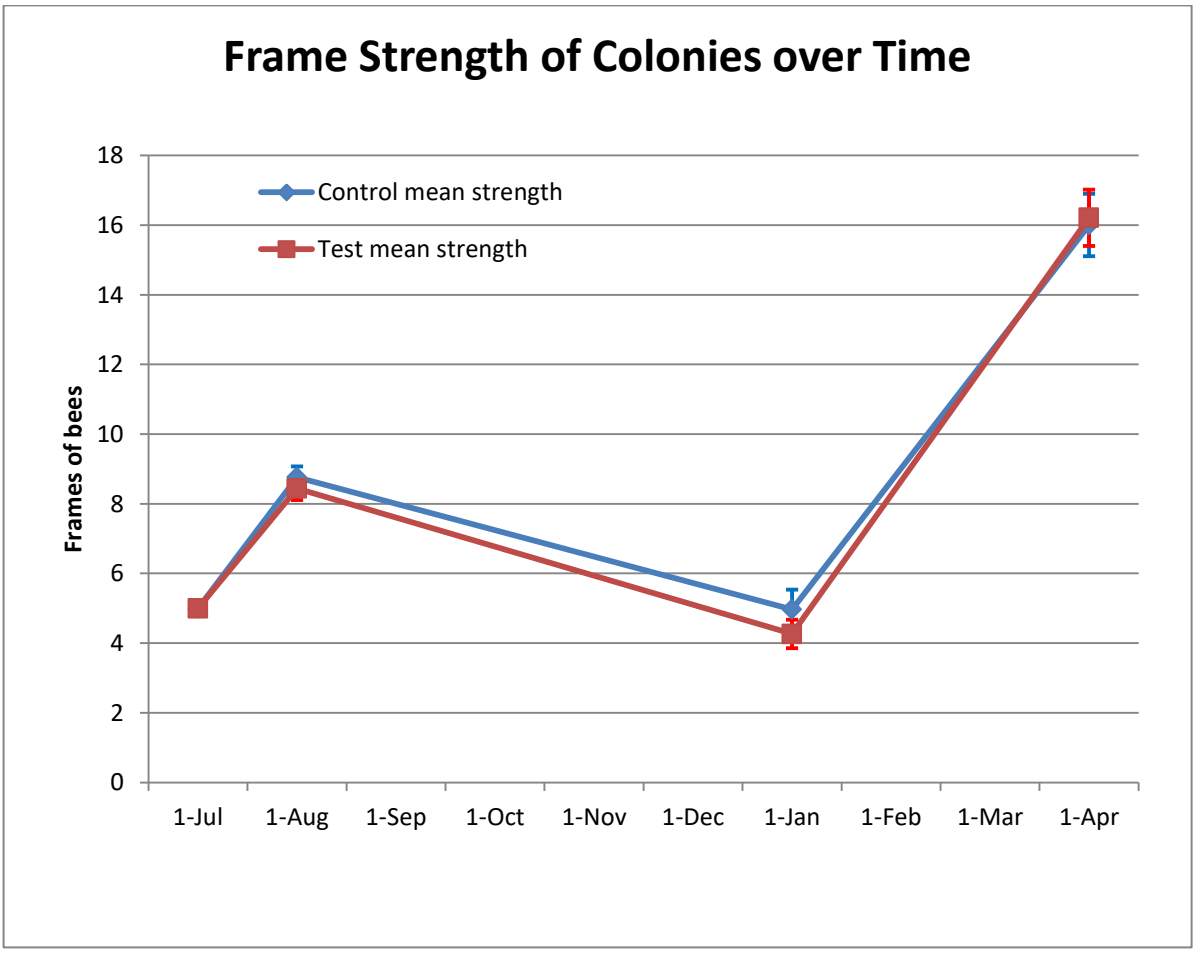
Figure 13. Mite levels climbed in the Control hives. The three most efficacious treatments were Apiguard applied twice to the top of the hive, a single MAQS strip applied twice in the broodnest, and a single MAQS followed by a thymol pad. For each group, the starting and final numbers of queenright colonies are shown in parentheses (2 Controls were removed after Day 29 due to excessive mite counts). Note the apparent colony stress from the thymol pads, as indicated by the loss of 3 colonies.

A column graph can also be used as a histogram—the distribution of number of hives falling into different categories, such as below:

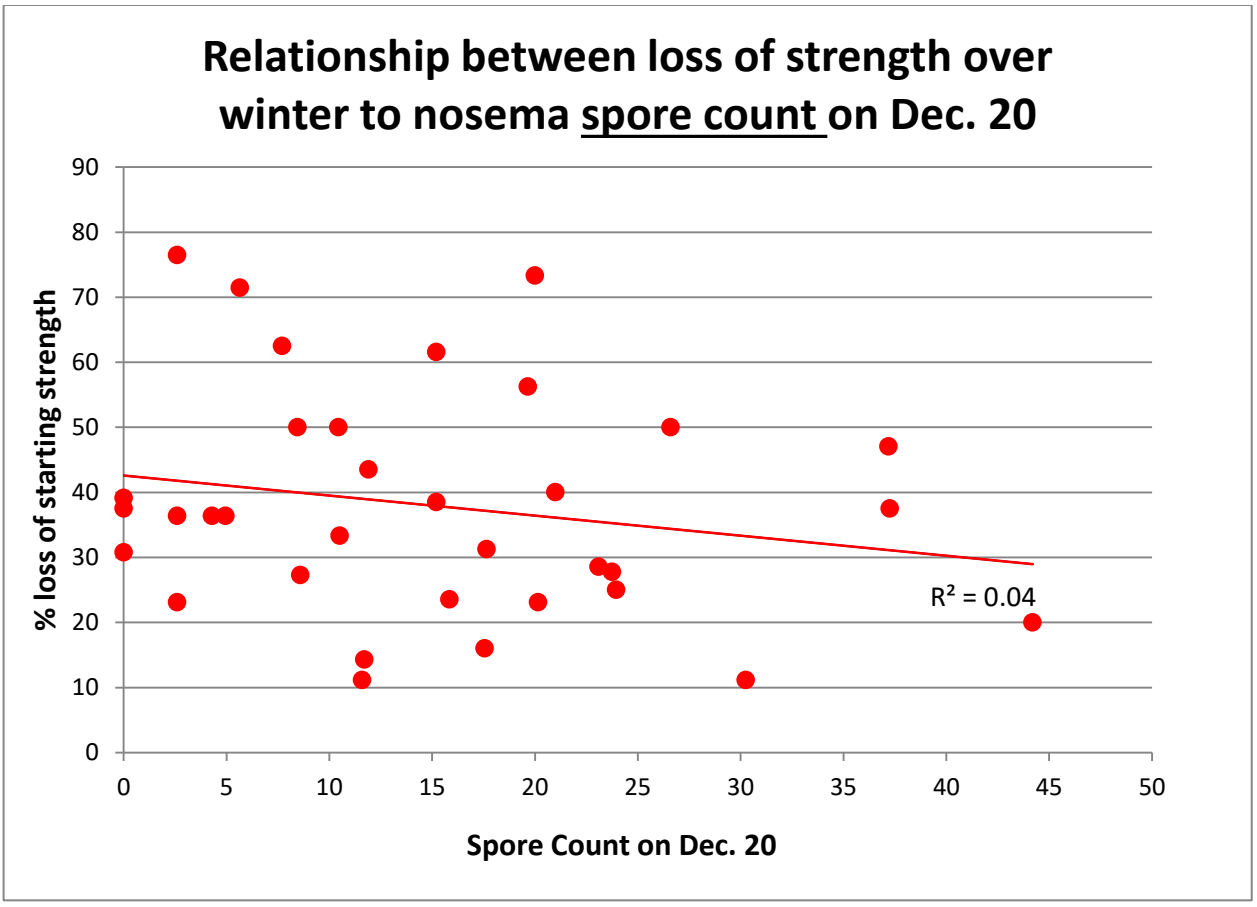


In the histogram above, it is easy to see that the fed hives tended to gain more weight than the controls. It also indicates that some of the Control hives were exceptional honey producers.

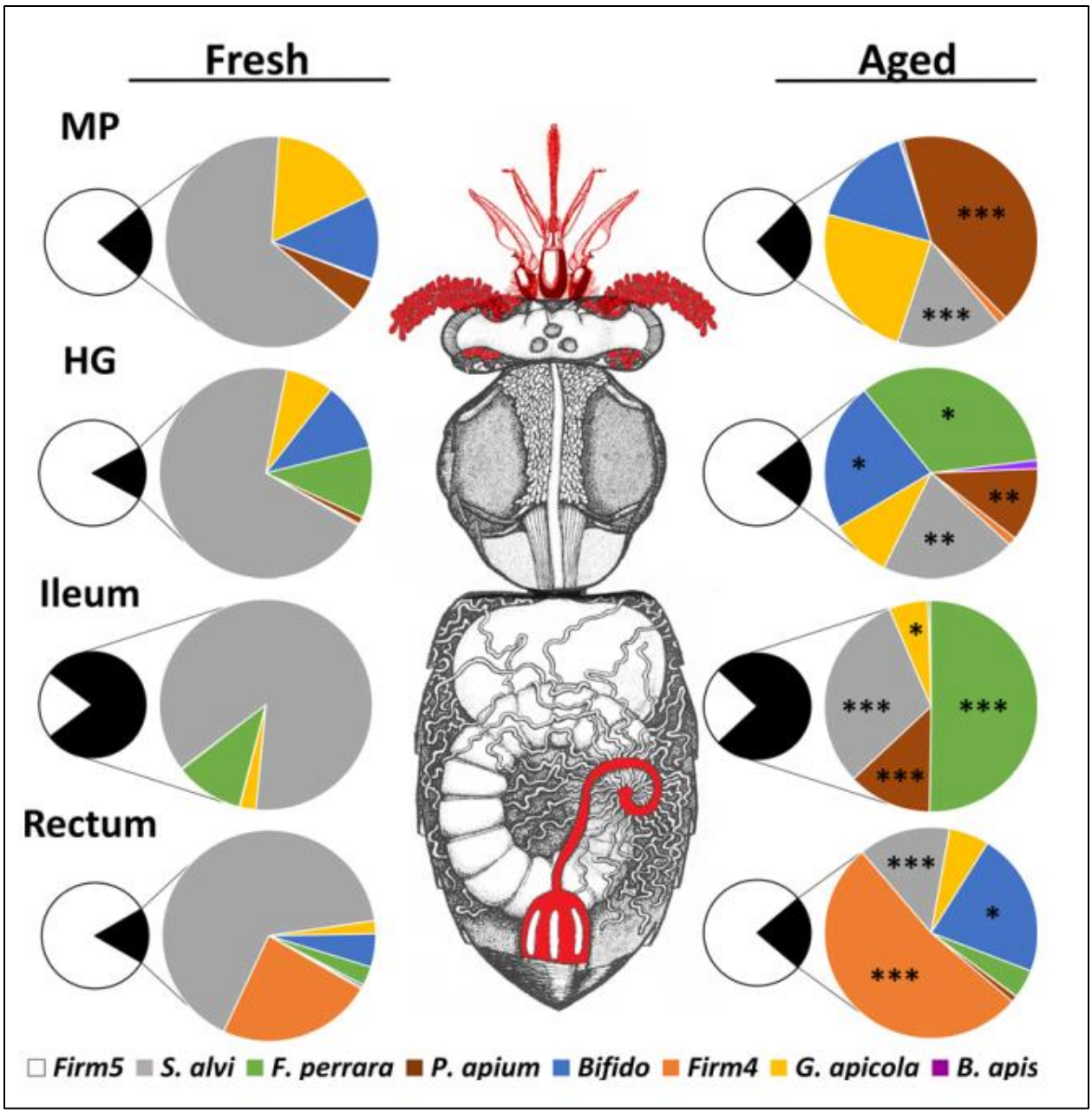
Below are examples of a line graph and a scatter graph (each for different data sets).



An example of a line graph, showing, in this case, no significant difference between the Test and Control groups over time. Note the SEM bars; since they are small, these results indicate consistency.

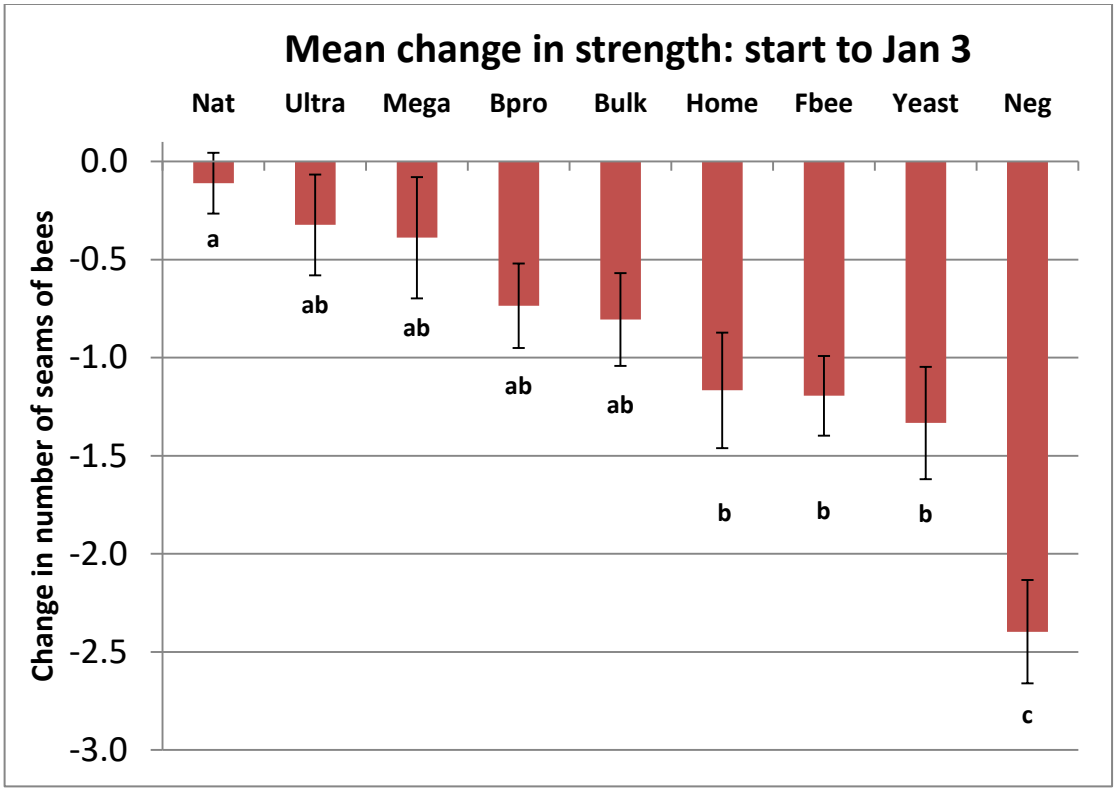


An example of a scatter plot, with a trendline (calculated by Excel). The very low r^2 value indicates that there was no consistent relationship between the variables (nosema spore counts and loss of colony strength). Had there been a strong correlation, the data points would have grouped along a line with a pronounced slope, and the r squared value of the trendline would have approached a value of 1 (typically indicative of a good correlation with a value above 0.60).



Pie charts are useful for demonstrating changes in distribution. From Maes, Rodrigues, Oliver, Ott, Anderson (2016) Diet-related gut bacterial dysbiosis correlates with impaired development, increased mortality and nosema disease in the honeybee (*Apis mellifera*). Molecular Ecology doi: 10.1111/mec.13862

You may also need to do statistical tests in order to determine whether your results for your test group were "significantly" different than those from the controls (less than 5% likely to result from chance, or better yet, less than 1% likely; p values of 0.05 or 0.01, respectively). Use the Student's t test (for data with a "normal" distribution), or the Kruskal-Wallis or Mann–Whitney U test (for nonparametric data). You may also wish to run an analysis of variance (ANOVA). For determine of efficacy of treatments, use the Henderson-Tilton formula. Calculators for these tests are available online, but it's best to get help from a trained statistician.



This is an example of results with statistically-significant differences indicated by letters (from http://scientificbeekeeping.com/a-comparative-test-of-the-pollen-sub/).

THE PARTS OF A SCIENTIFIC PAPER.

Title in Bold (*Latin names in italic)*

Key words: for computer search

by Your Name

The Date

ABSTRACT

A very brief summary of what you did, your results, and your key conclusions.

INTRODUCTION

A brief summary of the state of our knowledge on the specific subject of this experiment, including appropriate citations (Author Date) of any background research, your hypothesis, why you feel that the experiment is worth doing, and its potential practical applications. ***Be sure to clearly state the question that this experiment is trying to answer*** (every action and word of this experiment and write up should apply to the answering of this question).

MATERIALS AND METHODS

Exact and detailed instructions of what you did, such that other researchers could exactly duplicate your experiment. Use appropriate brand names and measurements. Refer to figures or tables as needed for clarity; refer to them in the text by citing them (e.g., "The feeder design included a screen to guide the bees (Fig. 1)."). **Each figure or table should contain enough information to be self explanatory without reading the text of the paper**; include a brief

descriptive title at the top, and an explanatory legend below (the legend of Tables by convention goes at the top). Describe the number of replications, the variables, and controls.

RESULTS

This is where you present your results. Explain them in writing, but also include data tables or graphs to illustrate your findings visually (cite them in the text). Calculate standard errors of the mean and levels of significance and show them in the illustrations. ***Do not discuss your interpretation of the results yet!***

DISCUSSION

This is where you discuss what you learned. Interpret the data. List practical applications, problems with the experiment, and suggest future experiments.

ACKNOWLEDGEMENTS

Who helped you, who reviewed the paper, and who funded your research (was there any conflict of interest?).

LITERATURE CITED

Use the following style:

Genersch, E and M Aubert (2010) Emerging and re-emerging viruses of the honey bee (*Apis mellifera* L.). Vet Res. 41(6): 54.

PRACTICAL ADVICE

You should be aware of the amount of tedious time involved in grading, in processing samples, and in data analysis and writing up. I have a lot of practice, and can tell you how long it takes us.

Be sure to practice taking each measurement several times at minimum before you even think about touching a hive in the experiment!!!!! You do not want to be learning or practicing on test hives or actual samples!!!!

Two of us—one tipping, one grading—can grade hives in the morning at a rate of around one hive per minute. Remember, you need a blinded grader. I prefer the relatively noninvasive "cluster grading," in early morning, before the colony has broken cluster. Gently smoke and tip the hive up off the bottom board, and count the number of frame interspaces filled with bees (I typically estimate to the nearest half or quarter frame). For colonies in a single box, the grade from below is enough (although you may count some very tiny colonies as zeroes). If there is more than one box, take a count for each box from both the top and bottom view. Then average the two counts for each box. Be sure to randomize the order in which you grade the hives in each group at each time point, so as to minimize the effect of time (clusters may start to break if it warms, thus affecting the grades).

Taking a level ½ cup sample of bees (for varroa infestation rate and/or nosema count) takes another couple of minutes per hive. Allow a couple more minutes to check for brood. You should be able to grade and sample 24 hives before noon.

Now comes the hard part—processing the samples. Allow 2-3 minutes ea for varroa counts (faster if you build a shaker table). Then 5-10 minutes for each nosema count. For an experiment with 24 hives, if you allow 15 minutes total for each sample, that would add up to 6 hours of bench work for each time point.

There are three main methods of measuring nosema: field of view, mean spore count, or prevalence (all covered at this website). The latter two are the most accurate, and take about the same amount of time. For mean spore count I suggest using 25 bees by the "ziplock method and a hemacytometer. For prevalence, perform gut squashes of 10 individual bees. Either method gives roughly the same results, but prevalence is more biologically relevant.

FIELD PHOTOS

For equalizing colonies at the start of a trial, I really like using my 5-frame method. Place three frames of bees, brood, and the queen in the middle, with equal total amounts of brood in each hive. Place a frame of honey on one side, a drawn comb on the other (each comb covered with bees). Leave the other 5 combs out. The next day, remove all covers and either add or remove bees, typically by swapping the drawn comb. Bring in shook bees from another yard to add as necessary, then close the hives. Repeat daily—removing all covers, until every hive has the same coverage of bees on the two outer comb faces—at which point they will be equalized for strength, and you can add the remaining combs (without bees).

Above hives now equalized and ready to go. In this trial, we were not concerned about drift, so placing the hives in a line was OK. Normally, arrange them to avoid drift.

Equalizing strengths and weight of test hives. Bees shaken on lids in front of the hives to allow older bees to fly off—leaving only young bees to walk in. It eliminates another variable if all hives start with roughly the same amount of bees, brood, drawn comb, and weight.

Hives marked with plastic table setting numbers.

Be sure to record environmental influences throughout the trial. In the example above, I noted that there was a strong nectar flow on, plentiful mixed high-quality pollen, plenty of jelly around the larvae (not shown). These observations can help a great deal in the Discussion.

Be sure to run sham controls. Here the white sham treatment contains exactly as much sugar as in the pollen sub patty, but without the protein and other added ingredients. This controls for the effect of the added sugar.

A photo is worth a thousand words! Take plenty of photos at each time point. The photo above clearly shows how the bees eagerly devoured the sham treatment (sugar paste), while barely touching the pollen sub patty.

When feeding more than one treatment in syrup, I add a tiny bit of food coloring to each treatment in order to prevent us from getting the feeder jars mixed up.

Grading hives at dawn (I lightened the photo). Note the flashlights to see between the frames. And most important, and undistracted data recorder.

Grading for strength at the beginning of almond bloom. I'm counting the number of frame interspaces filled with bees, to the nearest ¼ frame.

Weighing hives with a digital crane scale.

Inspecting brood frames for adverse effects from the treatment.

Two inspectors estimating bee coverage on each side of every frame (they cannot speak aloud). One will consistently grade larger than another. The two beekeepers in the back are weighing hives on a digital postal scale set on top of a hive.

Somewhat less accurate is to use a digital luggage scale to hook in the hand cleat and lift the hive very slightly from side to side or front to back (not as accurate), and then add the two weights. Confirm the repeatability of your weights first!

Counting mites from an alcohol wash. See Smokin' Hot Mite Washin'

Taking honey samples for analysis.

A homemade incubator with “bee cup” cages. I will soon publish plans for each.

END NOTES

[1] Formally, the researcher tries to disprove the *null hypothesis*—that the experimental treatment will have no consistent effect (that the treated hives do not differ at the end of the experiment from the controls).

[2] *The following is from an email exchange that I had with someone who wanted to do a study to “prove” a point:*

Who do you expect to have an interest in the results of your study? The first thing to do is to imagine that you are standing in front of that group, making a presentation of your findings and discussion.

What I'm asking you to do is to make up, *prior to starting the experiment*, dummy powerpoint slides, as though the study had gone exactly as you had hoped, in order to show yourself how you would make your presentation points. This is similar to storybooking a movie prior to starting to film it.

I do not recommend presenting Excel spreadsheets to groups—charts (graphs) are much better. I understand that you may not yet be proficient in Excel. Go ahead and draw up your presentation in pencil, how you hope the graphs will look. Pretend that you are actually making the presentation to the groups that you mention above. Allow yourself to make up data for now, as though the study gave you exactly the results that you hoped for. Although I do not recommend putting more words on a Powerpoint slide than you would put on a T-shirt, for now type out the presentation points that you would wish to make—why the experiment was relevant, what you learned from it, and how others can apply your findings.

Then see if the presentation would be robustly convincing for those groups. I mean really imagine that you are standing up in front of them!

Only then can we really see whether the study is worth doing, and exactly how to design it. I hope that you understand what I'm trying to get you to do. Too many studies are started without really thinking them through to the end, and you just waste your time and effort in collecting a bunch of data that doesn't amount to anything worthwhile!

[3] A really tough problem to surmount in field experiments with hives of bees is that if one hive gets sick, the bees tend to drift to the immediately adjacent hives and make that colony sick too! So think out your hive placement, and how treatments are assigned. If all hives on one pallet get the same treatment, and one gets sick, all will likely get sick, and then your results would suggest that the treatment caused them to get sick! On the other hand, if you have paired hives, one treated and one control, then if the treated hive gets sick, its adjacent control hive will likely get sick too. Take some time and think this through before you start!

[4] New Star 23176 1 to 50-Double Side Plastic Table Numbers, 4 by 4-Inch, Black on White. Attach with at least two staples or thumbtacks. http://www.amazon.com/dp/B00B1JQO7W?psc=1

[5] Error bars in Excel are a bit tricky. Put your calculated SEMs [=STDEV(data range)/SQRT(COUNT(same data range)] below the means in your data table. Highlight the data series in the chart, then in the Layout page, go to Error Bars, more error bar options, custom, specify value. Then highlight the calculated SEMs to use.