Herein is the analysis of two small data sets that will provide a check to make sure that everything is working properly before you start analyzing data where the outcome is unknown. The first set uses a data set from Windaq. The organism is the Asian Citrus Psyllid. Thus expect waveforms NP (non-probing), C (probing), D (contact with phloem), E1 (salivation in phloem), E2 (Phloem ingestion), G (xylem ingestion). After finishing this set we will look at some aphid data gathered using Probe. I don't remember the species of aphids that was used for these recordings. Five of the aphids are some data that Alberto Fereres gave to me, and the remainder comes from Elaine Backus. I have arbitrarily assigned the first five aphids as treatment 1 and the rest as treatment 2. The expected waveforms are NP, C, pd (potential drop), E1, E2, G, F (stylet derailment), and E1e (extracellular salivation).

Note: To English buffs everywhere. I use quotation marks to denote exact text. While punctuation typically goes inside the quotation marks, I cannot follow this convention because "psyllid data" is different from "psyllid data." and the period has meaning for many computers.

Part 1) Analysis of psyllid behavior. The data are in the folder "Psyllid Data". Within this folder are ten insects, the first five are treatment 1, and the rest are treatment 2. Copy these files into a separate folder on your computer. This is what you will have when you start the analysis of your data. There is also another folder that contains programs and output. Since each program requires some modification (at least the file names), I have saved copies of the program in each folder.

> NOTE: Insect numbers are of the format a01. I use 01 for the number 1 so that if these are sorted insect 2 will come after insect 1. Sorting in many cases is by character not number, and in this case insect 10 will come after insect 1.

Open the file manipulation program, follow the tutorial to change file names and insect numbers. Run the program. The raw data is now in "ExampleW.csv". We now need to check the file for errors using the error checking program. We get the following output:

Frequency Table of Waveform Event Transitions

The FREQ Procedure

waveform	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C	234	44.32	234	44.32
D	18	3.41	252	47.73
E1	35	6.63	287	54.36
E2	20	3.79	307	58.14
G	39	7.39	346	65.53
NP	182	34.47	528	100.00

Frequency Table of Waveform Event Transitions

The FREQ Procedure

trans1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C to D	18	3.47	18	3.47

C to G	39	7.53	57	11.00
C to NP	172	33.20	229	44.21
D to C	2	0.39	231	44.59
D to E1	16	3.09	247	47.68
E1 to C	15	2.90	262	50.58
E1 to E2	20	3.86	282	54.44
E2 to E1	19	3.67	301	58.11
G to C	39	7.53	340	65.64
NP to C	178	34.36	518	100.00

There is some additional output that comes after these two tables. You are welcome to look at it, but it is not important for this task.

We examine the two frequency tables. All the expected waveforms are present, and all of the listed transitions are possible. Thus there are no obvious errors.

It is suggested that at this point you use Notepad, or Wordpad to open the data file. Make a change, like delete one of the observations and save using a different name. Then run Error Checker on that file. Go back to the original data and make one of the values negative and save using a different name. Then run Error Checker on that file. This will help you see what happens when there are problems.

The last waveform in all the files has a shorter duration than if it had ended naturally. Given that these are long recordings (over 15 hours), it would be nice to remove these values. Also, these insects were recorded at different times. Our target duration is 24H, but due to various constraints we don't typically get 24H in all recordings. So we will apply the following steps:

1) Find the duration of the shortest recording. Record a time that is a couple seconds shorter than this value as the cut-off.
2) Find the behavior that starts before this value and ends after this value.
   a. If this behavior is NP retain this value and delete everything that follows. This will distort the duration of NP, but mostly we don't care as much about this behavior as we do about other behaviors. If you disagree with this then change this statement:

```
if sumstart < cutoff then do; if sumend > cutoff then do; if
marker1=0 then do; if waveform ne "NP" then marker2=1; end; end;
end;
```

   by deleting the "If waveform ne "NP" then" and also deleting one "end;" statement.
   b. If this behavior is not NP, and if there are more behaviors that follow, then retain this behavior.
   c. If this behavior is not NP, and it is the last behavior in the file then delete it.
3) This method does result in a loss of data because the presence of events that end unnaturally only affects estimates of duration. This method changes both durations and counts. This problem could be avoided by running the analysis on the trimmed and untrimmed data. However, a note should be included in any manuscript if this approach is used because the

means will not appear to be correct. The total duration (trimmed) was 100 hours. The count (untrimmed) was 10. The average duration was therefore 18 hours.

We need to know the duration of the shortest recording. Open the ExampleW file using Excel. You will see three columns of data that fill column A, column B, and column C. Row 1 will have titles. The data start in row 2. In column D, row 2 type in "=if(a2=a1,c2+d1,c2)" and in column E row 2 type in "if(a2=a3,"",d2)" and then select all cells in column D and E from row 2 to the end of your data. Fill down. In column E, row 1 type in "=min(e1:e455)". Excel should show a value of 82772.94 (just under 23 hours). If you want to save this, save a copy under a different name. If you accidentally save the file, open it again, delete columns D and E, then save again. In any case, the file should now be closed in its unaltered state.

Open program Trimmer. In addition to changing the file names as usual, also change the cutoff value to something slightly less than the value that you found in Excel. Something like this will work:

```
data one; set one; cutoff=82770;
```

I typically change the output file name to something slightly different. ExampleWT for ExampleW trimmed. Run the program.

If you forgot to close the Excel file then you might only see the SAS log, and examining the output might give you something like this:

C	C	NP to C a02	C	399.2
NP	NP	C to NP a02	NP	320.8
C	C	NP to C a02	C	248
NP	NP	C to NP a02	NP	96

Go back, close the Excel file, and rerun. Now you should see a printout of the output data, and opening the ExampleWT.csv file will give you something that looks the same as the original file, but only 515 rows of data rather than 529 rows.

If you want, you can go back to the error checking program, to get frequency tables for each treatment. Make sure to change the file name to ExampleWt.csv. Also activate the line:

```
Data one; set one; *if substr(insectno,1,1)="e" then output;
```

So that it looks like this:

```
Data one; set one; if substr(insectno,1,1)="a" then output;
```

Run the program and save the output. Then change it to:

```
Data one; set one; if substr(insectno,1,1)="b" then output;
```

Here is what you should get:

For treatment A:

Frequency Table of Waveform Event Transitions

The FREQ Procedure

waveform	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C	146	46.35	146	46.35
D	5	1.59	151	47.94
E1	12	3.81	163	51.75
E2	7	2.22	170	53.97
G	22	6.98	192	60.95
NP	123	39.05	315	100.00

Frequency Table of Waveform Event Transitions

The FREQ Procedure

trans1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C to D	5	1.61	5	1.61
C to G	22	7.10	27	8.71
C to NP	118	38.06	145	46.77
D to E1	5	1.61	150	48.39
E1 to C	5	1.61	155	50.00
E1 to E2	7	2.26	162	52.26
E2 to E1	7	2.26	169	54.52
G to C	22	7.10	191	61.61
NP to C	119	38.39	310	100.00

For treatment B:

Frequency Table of Waveform Event Transitions

The FREQ Procedure

waveform	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C	81	40.70	81	40.70
D	13	6.53	94	47.24
E1	22	11.06	116	58.29
E2	12	6.03	128	64.32
G	17	8.54	145	72.86
NP	54	27.14	199	100.00

Frequency Table of Waveform Event Transitions

The FREQ Procedure

trans1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C to D	13	6.70	13	6.70
C to G	17	8.76	30	15.46
C to NP	49	25.26	79	40.72
D to C	2	1.03	81	41.75
D to E1	11	5.67	92	47.42
E1 to C	9	4.64	101	52.06
E1 to E2	12	6.19	113	58.25
E2 to E1	11	5.67	124	63.92
G to C	17	8.76	141	72.68
NP to C	53	27.32	194	100.00

Now open the Analysis program, change the file name, and change the file name in the ODS statement. Turn off the transformations (if needed). I chose to name the output file "ExampleWT AnalysisNT". The transformations should all be in green letters:

```
************************************************************************
************************************************************************
************************************************************************
************************************************************************
*******                                                       **********
*******                                                       **********
*******                                                       **********
*******                  TRANSFORMATIONS                      **********
*******                                                       **********
*******                                                       **********
************************************************************************
************************************************************************
************************************************************************
************************************************************************;

/*
************************************************************************
************************************************************************
**** Here is a good generic set of transformations                 ****
**** Counts are sqrt transformed, durations are log transformed    ****
************************************************************************
************************************************************************;
Data Ebert; Set Ebert;
PrcntPrbC = arsin(sqrt(PrcntPrbC/100));
PrcntPrbE1 = arsin(sqrt(PrcntPrbE1/100));
```

The key here is the "/*".

Run the program. Then change the ODS file name. I chose "ExampleWT AnalysisT".

To activate transformations, delete the only example of "/*" in this portion of green text.

It should now look like:

```
*************************************************************************
*************************************************************************
*************************************************************************
*************************************************************************
*******                                                       **********
*******                                                       **********
*******                                                       **********
*******                   TRANSFORMATIONS                     **********
*******                                                       **********
*******                                                       **********
*************************************************************************
*************************************************************************
*************************************************************************
*************************************************************************;


*************************************************************************
*************************************************************************
**** Here is a good generic set of transformations                  ****
**** Counts are sqrt transformed, durations are log transformed     ****
*************************************************************************
*************************************************************************;
Data Ebert; Set Ebert;
PrcntPrbC = arsin(sqrt(PrcntPrbC/100));
PrcntPrbE1 = arsin(sqrt(PrcntPrbE1/100));
```

Run the program.

Open the two output files in a word processing program like Microsoft Word®. Wait a few minutes to allow the program to process the file (let it finish the page count, if this is done automatically). Select all, then copy and paste into a spreadsheet program like Microsoft Excel®. I paste the non-transformed results starting in column A row 1. I paste the transformed results into the same worksheet starting in column L row 1. Give the computer time to finish each task. This is very important if you have several treatments, and it can take a couple of minutes. It should be fairly speedy with this example.

Since this output is a mean comparison, you can get the spreadsheet to extract only the results that you want to see. Mostly, this involves the means, standard errors, degrees of freedom, F statistic, and significance.

Find the first instance in column G, where you see "Pr > F". In this case this should be row 46. Copy this text. (This is important because (for whatever reason) you cannot type this in and have the next step work.)

In column U row 46 type in =if($G46="Pr > F",$B6,"") BUT when you get to the point where you would type in Pr > F, use the paste (control-P is the keyboard shortcut) command. Then type in the ending parentheses and the remainder of the statement. You should now see "TmFrstPrbFrmStrt" in this cell.

Select this cell and cells in columns V through AC. Then fill right (Control-R is the keyboard shortcut). The mean for treatment 1 is in column B, row 51 (cell B51), while the standard error is in cell C52. The mean and standard error for treatment 2 is in the row below that B52, and C52 respectively. In V46 change the statement “=if($G46=”Pr > F”,$B6,””)” to look like this “=if($G46=”Pr > F”,$B51,””)” where the only difference is the highlighted portion. In the remaining cells, change each $B6 to whatever cell has the information that you want. This example shows mean, standard error for treatment 1, mean and standard error for treatment 2, numerator degrees of freedom, denominator degrees of freedom, F value, and the probability of getting an F value as large or larger given that the null hypothesis is true.

You should now have row 46 in columns U through AC filled with the following:

TmFrstPrbFrmStrt	1156.32	2779.1	6426.27	2485.71	1	7	1.13	0.3229

Fill down to the end of the SAS output.

Copy columns U through AA. Use the paste special command to paste values into a new worksheet. This will fill columns A through column I. There will still be a large number of blank rows that we need to permanently remove. There are at least two ways to do this.

1) Select columns A through I and go to filter (Editing menus, “Sort and Filter” tab). Select filter. In row 1, each cell has a down arrow icon. Select this, go to the bottom of the list, and uncheck the box labeled “(blanks)”. The blank lines will disappear, but they are still there. Select all your data, copy and paste values. Now the blank lines are gone. NOTE: If you paste into the same worksheet, then you will need to go back to the box labeled “(blanks)” and check the box again. Otherwise most of the data will be hidden in the blank rows that are not being displayed.
2) In column J type in the number 1. Find the end of the data by dragging the slide bar to the bottom. Holding down the shift key, select the cell in column J that is next to the last set of data. This should select all the cells from row 1 to row 7331. In the Edit menu select Fill, then series. In the pop-up menu make sure that the “step value” is 1, and click ok. Select columns A through J (in that order), and then choose “sort and filter” from the edit menu. In the menu that appears choose “sort Z to A” All the results are now at the top, so copy them (Cells A1 through J98), and paste them into a new worksheet. Click anywhere to deselect them, then select columns J through A (in that order). Go back to the Edit menu, then “sort and filter” and finally “sort A to Z”. The last step is to delete the contents of column J.

The second approach is a bit more complex, but the techniques involved can be useful elsewhere.

You should end up with a table that looks something like this:

TmFrstPrbFrmStrt	1010.12	2327.46	6426.27	2327.46	1	8	1.55	0.249
CtoFrstE1	8	2.6161	5.75	1.8498	1	4	1.06	0.3606
NumF	0	.	0	.	1	8	.	.
DurFrstPrb	9377.65	4136.26	729.7	4136.26	1	8	1.81	0.2155
DurScndPrb	5720.38	3358.48	3003.48	3358.48	1	8	0.01	0.929
ShrtCbfrE1	772.44	1304.83	1744.03	922.65	1	4	0.2	0.6806
DurScndZ	1647.87	1857.62	6740.22	1857.62	1	8	3.97	0.0813
DurNnprbBfrFrstE1	18158	5970.86	26133	5970.86	1	8	3	0.1213
TmStrtEPGFrstE	53300	9365.17	51522	9365.17	1	8	0.13	0.7255
TmFrmFrstPrbFrstE	53094	0	45115	0	1	8	0.04	0.8479
TmBegPrbFrstE	1114.26	1614.89	3041.55	1141.9	1	4	0.39	0.564
NumG	4.4	0.9798	3.4	0.9798	1	8	0.03	0.8597
DurG	7218	1979.91	5386.37	1770.89	1	7	0.07	0.8042
MeanG	1384.01	591.93	2064.91	529.44	1	7	1.34	0.2856
NumPrbsAftrFrstE	3.6	2.1471	3.6	2.1471	1	8	0.06	0.8201
NmbrShrtPrbAftrFrstE	0.6	0.4583	0.4	0.4583	1	8	0.02	0.9026
NumE1	2.4	1.7635	4.4	1.7635	1	8	0.99	0.3484
NumLngE1BfrE2	0	.	0	.	1	8	.	.
NumSnglE1	-2.8E-17	0.1414	0.2	0.1414	1	8	1	0.3466
NumE2	1.4	1.077	2.4	1.077	1	8	0.41	0.5378
NumLngE2	1.2	0.9592	2	0.9592	1	8	0.34	0.5742
DurFirstE	18512	3499.92	5865.4	2474.82	1	4	1.17	0.34
CntrbE1toE	1.6565	29.659	27.4495	20.9721	1	4	0.97	0.3797
DurE1FlwdFrstSusE2	266.32	158.77	100.75	129.64	1	3	0	0.9614
DurE1FlldFrstE2	266.32	158.77	100.75	129.64	1	3	0	0.9614
PotE2Indx	54.4479	15.3886	43.5008	12.5647	1	3	0.54	0.5159
TtlDurE	37394	6799.53	16530	6799.53	1	4	1.17	0.3394
TtlDurE1	617.8	312.39	353.32	220.89	1	4	0	0.95
TtlDurE1FlldSusE2	511.68	279.69	194.03	228.37	1	3	0.01	0.9401
TtlDurE1FlldE2	520.8	319.91	321.81	261.21	1	3	0.02	0.8938
TtlDurE1FllwdE2PlsE2	37297	7089.8	21891	7089.8	1	3	1.58	0.2978
TtlDurE2	36776	7165.17	21569	7165.17	1	3	1.56	0.3007
MnDurE1	81.46	43.7931	87.017	30.9664	1	4	0.09	0.7752
MnDurE2	12971	3569.87	4927.75	2914.78	1	3	3.46	0.16
NumPrbs	23.8	7.025	10.6	7.025	1	8	1.33	0.282
NmbrC	29.2	7.3478	16.2	7.3478	1	8	1.3	0.2871
NmbrShrtC	3.2	1.4387	1	1.4387	1	8	0.77	0.4044
NumNP	24.6	7.0463	10.8	7.0463	1	8	1.65	0.2345
NmbrPDL	0	.	0	.	1	8	.	.
NmbrPDS	0	.	0	.	1	8	.	.
NmbrE1e	0	.	0	.	1	8	.	.

TtlDurC	34416	6836.5	29871	6836.5	1	8	0.19	0.673
TotDurNnPhlPhs	47502	5704.12	62221	5704.12	1	4	1.32	0.3153
TtlDurNP	25632	5774.81	30961	5774.81	1	8	1.36	0.2766
TtlPrbTm	55194	6581.28	48739	6581.28	1	8	0.16	0.704
MnDurNP	1760.39	1254.98	3932.51	1254.98	1	8	2.36	0.1631
MnDurC	2111.65	967.67	1746.03	967.67	1	8	0.08	0.7889
TmFrstSusE2	53406	9351.5	51778	9351.5	1	8	0.13	0.7247
TmFrstSusE2FrstPrb	53200	0	45540	0	1	8	0.04	0.8523
TmFrstSusE2StrtPrb	1380.59	1605.42	2528.41	1310.82	1	3	0.03	0.8838
TmFrstE2StrtEPG	53406	9351.5	51778	9351.5	1	8	0.13	0.7247
TmFrstE2FrmFrstPrb	53200	0	45540	0	1	8	0.04	0.8523
TmFrstE2FrmPrbStrt	1380.59	1605.42	2528.41	1310.82	1	3	0.03	0.8838
TtlDurNp1	1562.15	399.56	3476.64	399.56	1	8	3.51	0.0979
TtlDurNp2	1347.74	606.65	2959.17	606.65	1	8	1.57	0.245
TtlDurNp3	1306.27	755.76	2121.96	755.76	1	8	1.49	0.2567
TtlDurNp4	814.4	578.24	1412.48	578.24	1	8	2.83	0.1313
TtlDurNp5	508.49	471.44	2134.02	471.44	1	8	4.95	0.0567
TtlDurNp6	996.42	732.45	1759.8	732.45	1	8	2.45	0.1565
NumPDS1	0	.	0	.	1	8	.	.
NumPDS2	0	.	0	.	1	8	.	.
NumPDS3	0	.	0	.	1	8	.	.
NumPDS4	0	.	0	.	1	8	.	.
NumPDS5	0	.	0	.	1	8	.	.
NumPDS6	0	.	0	.	1	8	.	.
MnDurPdS1	0	.	0	.	1	8	.	.
MnDurPdS2	0	.	0	.	1	8	.	.
MnDurPdS3	0	.	0	.	1	8	.	.
MnDurPdS4	0	.	0	.	1	8	.	.
MnDurPdS5	0	.	0	.	1	8	.	.
MnDurPdS6	0	.	0	.	1	8	.	.
NumF1	0	.	0	.	1	8	.	.
NumF2	0	.	0	.	1	8	.	.
NumF3	0	.	0	.	1	8	.	.
NumF4	0	.	0	.	1	8	.	.
NumF5	0	.	0	.	1	8	.	.
NumF6	0	.	0	.	1	8	.	.
TtlDurF1	0	.	0	.	1	8	.	.
TtlDurF2	0	.	0	.	1	8	.	.
TtlDurF3	0	.	0	.	1	8	.	.
TtlDurF4	0	.	0	.	1	8	.	.
TtlDurF5	0	.	0	.	1	8	.	.
TtlDurF6	0	.	0	.	1	8	.	.
NumPrb1	2	0.4243	0.6	0.4243	1	8	6.26	0.0369

NumPrb2	2	0.5292	0.6	0.5292	1	8	2.17	0.1787
NumPrb3	2	0.7616	0.6	0.7616	1	8	1.52	0.253
NumPrb4	1.4	0.2236	0.8	0.2236	1	8	2.66	0.1416
NumPrb5	2.4	0.9	1.4	0.9	1	8	0.78	0.4016
NumPrb6	1.4	0.4796	1	0.4796	1	8	0.37	0.5607
TmEndPDBegE1FllwdSusE2	60967	4066.69	71071	3320.44	1	3	3.7	0.15
maxE2	18525	3706.69	9889.1	3026.5	1	3	3.26	0.1689
DurNpFllwFrstSusE2	241.63	555.25	837.6	453.36	1	3	0.69	0.4667
PrcntPrbC	64.0906	13.3759	59.2641	13.3759	1	8	0.1	0.756
PrcntPrbE1	0.3334	0.29	0.725	0.29	1	8	1.57	0.2452
PrcntPrbE2	25.3767	15.487	23.4923	15.487	1	8	0.01	0.9312
PrcntPrbF	0	.	0	.	1	8	.	.
PrcntPrbG	10.1227	5.4718	15.5733	5.4718	1	8	0.73	0.4165
PrcntE2SusE2	90	17.6383	83.3333	14.4016	1	3	0.01	0.942

The values are the mean value of the variables as calculated for each insect (these are “per insect” from values calculated “by insect”).

Since the data are in seconds, these means of durations are in seconds. The data is not accurate to hundredths of a second. You can use the format cells command in Excel to remove the extra precision. This is as far as this tutorial can go. Deciding on what is significant and what to present is left to the scientist performing the data analysis. We will provide some suggestions in a different tutorial.

Bonus Material

The task is to print out a list of all the results for each insect to facilitate a comparison between the output of this program and that of other programs that do not provide a complete analysis. In the base SAS you will need to put a print statement at the end of the program.

Data Ebert; Set Ebert; Proc Print; Run;

You will then need to either run the entire program, or if you have already run the program you can highlight this section and run only this section.

If you use Enterprise Guide, there will be an “Output Data” tab next to the Log tab. Below the “Program” tab there will be a name with a small down arrow. Click on the down arrow to see a list of all the data sets that are in memory (there should be 12 of them). Mouse click on the one named Ebert. Select all, copy and paste into Excel. The only problem here is that the names of the variables do not get copied.

One way to get the variable names included is to use Proc Export:

```
proc export data=Ebert outfile='C:\Users\Location 3\Ebert.csv'
dbms=csv replace;
```

Before comparing output, make sure that all transformations have been turned off. Then export the dataset Ebert. Open the file in Excel. Copy the data, and then click on the down arrow under Paste menu. Select “Paste Special” and from that pop-up menu select “Transpose.” This is now in the format that will match what you will get from the Sarria workbook.

To get the data into Sarria you will have to recode all the behaviors. You can do this in SAS or in Excel. Here is one approach:

1) Open the data file “ExampleWT.csv” using Excel or other spreadsheet program.
2) There are two major tasks: convert waveforms to a numeric code, and change the duration to a time from beginning of file.
    a. To convert to a numeric code go to cell D2 and type in “=if(A2=”NP”,1,””)” In the next few columns type in the same formula, but change the behavior and change the number to match the codes in the Sarria workbook. NP=1, C=2, E1=4, E2=5, and G=7. There is no code for D, so I chose D=11. In cell J2 type in “=Sum(D2,I2)” and then fill down to the end of the file {I would save this in a special place to avoid having to do it over again later}. Copy and paste values into column A, and delete columns D through J. NOTE code 11 is treated as PDL in Sarria. This will alter some calculations!
    b. Find the end of each insect and insert two empty rows.
    c. In cell D2 type in 0.
    d. In cell D3 type in “=D2+c2” and fill down.
    e. Go to the end of each insect. In the first blank line at the end of the insect type in 12 for the behavior. This is the end of file code used by the Sarria workbook.
    f. At the start of the next insect type in a 0.
3) Copy and paste values each insect into the Sarria workbook. Then follow the instructions for running the workbook.

The first four insects give the following output (Font is at 6 point to make everything fit. It is recommended to copy this table and paste into Excel, then adjust the font size):

OUTPUT FROM EBERT 1.0

	a01	a02	a06	a07
TmFrstPrbFrmStrt	425.3	602.72	2644	37.76
CtoFrstE1	9	7	.	.
NumF	0	0	0	0
DurFrstPrb	3676.08	399.2	335.2	42293.63
DurScndPrb	1863.36	248	1438.4	24121.6
ShrtCbfrE1	487.52	1057.36	.	.
DurScndZ	1299.84	320.8	117.6	5427.2

OUTPUT FROM SARRIA

	a01	a02	a06	a07
Time to 1st probe from start of EPG	425.3	602.72	2644	37.76
Number of probes to the 1st E1	9	7		
Number of F	0	0	0	0
Duration of 1st probe	3676.08	399.2	335.2	42293.63
Duration of 2nd probe	1863.36	248	1438.4	24121.6
Duration of the shortest C wave before E1	487.52	1057.36		
Duration of the second nonprobe period	1299.84	320.8	117.6	5427.2

TtlDurF	.	.	.	.	Total duration of F				
DurNnprbBfrFrstE1	5610.49	3203.52	10740.8	8733.76	Duration of nonprobe period before the 1st E	5610.49	3203.52	10740.8	8733.76
meanpd	.	.	.	.	Mean duration of pd	44.53333	48.945		
meanPDL	.	.	.	.	Mean duration of pd-L	44.53333	48.945		
meanPDS	.	.	.	.	Mean duration of pd-S				
meanNPdPrb	.	.	.	.	Average number of pd per probe	0.142857	0.153846		
meanF	.	.	.	.	Mean duration of F				
TmStrtEPGFrstE	24469.83	7687.73	76030.24	75148.99	Time from start of EPG to 1st Ey	24469.83	7687.73	76030.24	75148.99
TmFrmFrstPrbFrstE	24044.53	7085.01	76030.24	75148.99	Time from 1st probe to 1st Ey	24044.53	7085.01	73386.24	75111.23
TmBegPrbFrstE	841.92	1386.61	.	.	Time from the beginning of that probe to 1st Ey	841.92	1386.61		
NumG	5	0	6	5	Number of G	5	0	6	5
DurG	6020.96	.	4681.6	15700.16	Duration of G	6020.96		4681.6	15700.16
MeanG	1204.192	.	780.2667	3140.032	Mean duration of G	1204.192		780.2667	3140.032
NumPrbsAftrFrstE	12	6	0	0	Number of probes after 1st E	12	6	0	0
NmbrShrtPrbAftrFrstE	3	0	0	0	Number of probes (shorter than 3 minutes) after 1st E	3	0	0	0
NumE1	8	4	0	0	Number of E1	8	4	0	0
NumLngE1BfrE2	0	0	0	0	Number of E1 (longer than 10 minutes) followed by E2	0	0	0	0
NumSnglE1	0	0	0	0	Number of single E1	0	0	0	0
NumE2	5	2	0	0	Number of E2	5	2	0	0
NumLngE2	4	2	0	0	Number of sustained E2 (longer than 10 minutes)	4	2	0	0
DurFirstE	13458.4	23564.8	.	.	Duration of 1st E	13458.4	23564.8		
CntrbE1toE	3.132336	0.180673	.	.	Contribution of E1 to phloem phase (%)	3.132336	0.180673		
DurE1FlwdFrstSusE2	523.52	9.12	.	.	Duration the E1 followed by first sustained E2 (>10 min)	523.52	9.12		
DurE1FlldFrstE2	523.52	9.12	.	.	Duration the E1 followed by the first E2	523.52	9.12		
PotE2Indx	60.58456	48.31132	.	.	Potential E2 index	60.58456	48.31132		
TtlDurE	37283.36	37504.24	.	.	Total duration of E	37283.36	37504.24		
TtlDurE1	1167.84	67.76	.	.	Total duration of E1	1167.84	67.76		
TtlDurE1FlldSusE2	986.08	37.28	.	.	Total duration of E1 followed by sustained E2 (>10 min)	986.08	37.28		
TtlDurE1FlldE2	1004.32	37.28	.	.	Total duration of E1 followed by E2	1004.32	37.28		
TtlDurSnglE1	.	.	.	.	Total duration of single E1				
TtlDurE1FllwdE2PlsE2	37119.84	37473.76	.	.	Total duration of E1 followed by E2 and E2	37119.84	37473.76		
TtlDurE2	36115.52	37436.48	.	.	Total duration of E2	36115.52	37436.48		
MnDurE1	145.98	16.94	.	.	Mean duration of E1	145.98	16.94		
MnDurE2	7223.104	18718.24	.	.	Mean duration of E2	7223.104	18718.24		
NumPrbs	21	13	60	2	Number of probes	21	13	60	2
NmbrC	29	15	66	7	Number of C	26	13	66	7
NmbrShrtC	3	1	11	0	Number of short probes (C<3 minutes)	3	1	11	0
NumNP	21	14	61	3	Number of np	21	14	61	3
NmbrPD	.	.	.	.	Number of pd	3	2	0	0
NmbrPDL	0	0	0	0	Number of pd-L	3	2	0	0

NmbrPDS	0	0	0	0	Number of pd-S	0	0	0	0
NmbrE1e	0	0	0	0	Number of E1e	0	0	0	0
TtlDurC	33252.32	9316.32	60607.84	50715.07	Total duration of C	33385.92	9414.21	60607.84	50715.07
TtlDurE1e	.	.	.	.	Total duration of E1e				
TotDurNnPhlPhs	47321.75	47682.69	.	.	Total duration of no phloematic phase	47321.75	47682.69		
TtlDurNP	7914.87	38268.48	10740.8	8733.76	Total duration of np	7914.87	38268.48	10740.8	8733.76
TtlDurPD	.	.	.	.	Total duration of pd	133.6	97.89		
TtlDurPDL	.	.	.	.	Total duration of pd-L	133.6	97.89		
TtlDurPDS	.	.	.	.	Total duration of pd-S				
TtlPrbTm	76690.24	46918.45	65289.44	66415.23	Total probing time	76690.24	46918.45	65289.44	66415.23
MnDurNP	376.8986	2733.463	176.0787	2911.253	Mean duration of np	376.8986	2733.463	176.0787	2911.253
MnDurC	1146.632	621.088	918.3006	7245.01	Mean duration of C	1284.074	724.17	918.3006	7245.01
TmFrstSusE2	24993.35	7696.85	76030.24	75148.99	Time to from start of EPG 1st sustained E2 (10 minutes)y	24993.35	7696.85	76030.24	75148.99
TmFrstSusE2FrstPrb	24568.05	7094.13	76030.24	75148.99	Time from 1st probe to 1st sustained E2 (10 minutes)y	24568.05	7094.13	73386.24	75111.23
TmFrstSusE2StrtPrb	1365.44	1395.73	.	.	Time from the beginning of that probe to 1st sustained E2 (10 minutes)y	1365.44	1395.73		
TmFrstE2StrtEPG	24993.35	7696.85	76030.24	75148.99	Time from start of EPG to 1st E2y	24993.35	7696.85	76030.24	75148.99
TmFrstE2FrmFrstPrb	24568.05	7094.13	73386.24	75111.23	Time from 1st probe to 1st E2y	24568.05	7094.13	73386.24	75111.23
TmFrstE2FrmPrbStrt	1365.44	1395.73	.	.	Time from the beginning of that probe to 1st E2y	1365.44	1395.73		
TtlDurNp1	425.3	2100.8	2761.6	37.76	Total duration of np during the 1st hour	425.3	2100.8	2761.6	37.76
TtlDurNp2	1299.84	1102.72	736.16	0	Total duration of np during the 2nd hour	1299.84	1102.72	736.16	0
TtlDurNp3	1988.64	0	942.72	0	Total duration of np during the 3rd hour	1988.64	0	942.72	0
TtlDurNp4	336	0	611.2	0	Total duration of np during the 4th hour	336	0	611.2	0
TtlDurNp5	59.75	0	786.88	0	Total duration of np during the 5th hour	59.75	0	786.88	0
TtlDurNp6	1252.32	0	129.76	0	Total duration of np during the 6th hour	1252.32	0	129.76	0
NumPDS1	0	0	0	0	Number of pd-S during the 1st hour	0	0	0	0
NumPDS2	0	0	0	0	Number of pd-S during the 2nd hour	0	0	0	0
NumPDS3	0	0	0	0	Number of pd-S during the 3rd hour	0	0	0	0
NumPDS4	0	0	0	0	Number of pd-S during the 4th hour	0	0	0	0
NumPDS5	0	0	0	0	Number of pd-S during the 5th hour	0	0	0	0
NumPDS6	0	0	0	0	Number of pd-S during the 6th hour	0	0	0	0
MnDurPdS1	0	0	0	0	Average duration of pd-S during the 1st hour				
MnDurPdS2	.	.	.	.	Average duration of pd-S during the 2nd hour				
MnDurPdS3	.	.	.	.	Average duration of pd-S during the 3rd hour				
MnDurPdS4	.	.	.	.	Average duration of pd-S during the 4th hour				
MnDurPdS5	.	.	.	.	Average duration of pd-S during the 5th hour				
MnDurPdS6	.	.	.	.	Average duration of pd-S during the 6th hour				
NumF1	0	0	0	0	Number of F during the 1st hour	0	0	0	0
NumF2	0	0	0	0	Number of F during the 2nd hour	0	0	0	0
NumF3	0	0	0	0	Number of F during the 3rd hour	0	0	0	0
NumF4	0	0	0	0	Number of F during the 4th hour	0	0	0	0
NumF5	0	0	0	0	Number of F during the 5th hour	0	0	0	0

NumF6	0	0	0	0	Number of F during the 6th hour	0	0	0	0
TtlDurF1	0	0	0	0	Total duration of F during the 1st hour	0	0	0	0
TtlDurF2	0	0	0	0	Total duration of F during the 2nd hour	0	0	0	0
TtlDurF3	0	0	0	0	Total duration of F during the 3rd hour	0	0	0	0
TtlDurF4	0	0	0	0	Total duration of F during the 4th hour	0	0	0	0
TtlDurF5	0	0	0	0	Total duration of F during the 5th hour	0	0	0	0
TtlDurF6	0	0	0	0	Total duration of F during the 6th hour	0	0	0	0
NumPrb1	1	4	2	1	Number of probes during the 1st hour	1	4	2	1
NumPrb2	2	3	4	1	Number of probes during the 2nd hour	2	3	4	1
NumPrb3	2	1	6	1	Number of probes during the 3rd hour	2	1	6	1
NumPrb4	2	1	2	1	Number of probes during the 4th hour	2	1	2	1
NumPrb5	2	1	7	1	Number of probes during the 5th hour	2	1	7	1
NumPrb6	3	1	2	1	Number of probes during the 6th hour	3	1	2	1
TmFrstCFrstPD	.	.	.	.	Time from the beginning of the 1st probe to first pd	24014.45	7034.24		
TmEndLstPDEndPrb	.	.	.	.	Time from the end of the last pd to the end of the probe	13495.84	23680.72		
SumPDII1	.	.	.	.	Total duration of subphase II1 fo the pd				
SumPDII2	.	.	.	.	Total duration of subphase II2 fo the pd				
SumPDII3	.	.	.	.	Total duration of subphase II3 fo the pd				
TmEndPDBegE1FllwdSusE2	67119.81	54815.09	.	.	Time from the end of the last pd to the beginning of the E1 followed by the sustained E2 (>10 min)				
TmLstPdEndRcrd	.	.	.	.	Time from the end of the last pde to the end of the EPG record (Z)	31120.34	30371.84		
TmLstE1EndRcrd	.	.	.	.	Time from the beginning of E1 to the end of the EPG record (Z)				
TmLstE2EndRcrd	.	.	.	.	Time from the beginning of E2 to the end of the EPG record (Z)				
maxE2	13493.44	23555.68	.	.	Duration of the longest E2	13493.44	23555.68		
DurNpFllwFrstSusE2	226.3	256.96	.	.	Duration of np just after the probe of the first sustained E2	226.3	256.96		
DurTrmNpFllwFrstSusE2	.	.	.	.	Duration of np just after the probe of the first sustained E2 if it lasts untill the end of the recording				
PrcntPrbC	43.35926	19.85641	92.82947	76.3606	% of probing spent in C	43.53347	20.06505	92.82947	76.3606
PrcntPrbE1	1.522801	0.144421	0	0	% of probing spent in E1	1.522801	0.144421	0	0
PrcntPrbE2	47.09272	79.79053	0	0	% of probing spent in E2	47.09272	79.79053	0	0
PrcntPrbF	0	0	0	0	% of probing spent in F	0	0	0	0
PrcntPrbG	7.851012	0	7.170532	23.6394	% of probing spent in G	7.851012	0	7.170532	23.6394
PrcntE2SusE2	80	100	.	.	% E2 >10 min	80	100		

There are two types of highlighted text. Text highlighted in peach includes numbers that are in the wrong place because D is coded as a type of pd in the Sarria workbook. Text highlighted in green has values that are a disagreement between Ebert 1.0 and the Sarria workbook. The Sarria workbook calculates the number of C (pathway) by counting the number of C (event) and subtracting the number of pd. The duration of C (pathway) is calculated as the duration of all C events plus the duration of all pd events. Thus the duration of C (pathway) for C-pd-C is the sum of one pd and two C events. The duration of C (pathway) for C-pd-C-pd is the sum of two pd and two C events. The number of C events in the first instance is 1, but it is zero in the second instance. Thus the Sarria workbook finds a different number of C events than does Ebert 1.0, and the two programs will not agree on the result of those variables that involve a count of the number of C (pathway) events.

Using Data from Probe

The overall methodology is very similar to what was described above. One key difference is that there is a different program for file manipulation. However, in terms of what needs to be done to make the program work, the tasks are exactly the same: change file names, change insect numbers, and change the output file name. Note that the file from Probe will contain numeric codes for the waveforms, but the output from the SAS programs will contain character codes. There is a folder with eight aphids that we will use as an example. For lack of a more creative name, the folder is called “Aphid”.

FileManiP 080714: Change the file names for all infile statements. The infile statements look like this:

```
infile 'C:\Users\Location 3\Control 2.ana' dsd dlm='09'x truncover;
```

The next place that may need adjustment is the insect number. The statement looks like this:

```
insect1="b2";
```

The second file (=second insect), and all other files use a set of statements that can be copied + pasted over and over for as many files as you have. The code is like this:

```
data one;
infile 'C:\Users\Location\Control 2.ana' dsd dlm='09'x truncover;
input @; _infile_=compress(translate(_infile_,'.',','),'"'); input a b
c ;

data one; set one; drop c; dur=0;
insect1="b2";
data one; set one; retain holder1 in0;
if in0 ne insect1 then do; holder1=0; in0=insect1; dur=b; end; else
dur=b-holder1; holder1=b;
data two; set one; insectno=insect1; waveform=a; duration=dur;
data two; set two; drop a b holder1 in0 dur insect1;
data two; set two; retain holder1 in0;
if in0 ne insectno then do; in0=insectno; holder1=0; end; else
wave1=holder1; holder1=waveform;
data two; set two; if wave1 ne "." then output; data two; set two;
waveform=wave1;
```

```
data two; set two; drop in0 wave1 holder1; data two; set two; proc
append base=allsets data=two;
proc datasets nolist nodetails; delete one two;
```

For each copy of this code, make sure that the file name and insect number match the correct file.

Finally, go to the end of the program and change the export file in this statement:

```
proc export data=allsets outfile='C:\Users\Location 3\AphidRaw.csv'
dbms=csv replace;
```

Run the program.

You should now have a new file called "AphidRaw.csv" that contains three columns of data: insect number, waveform, duration.

The remaining tasks are now the same as previously described.

Open the file manipulation program, follow the tutorial to change file names and insect numbers. Run the program. The raw data is now in “AphidRaw.csv”. We now need to check the file for errors using the error checking program. We get the following output:

Frequency Table of Waveform Event Transitions

The FREQ Procedure

waveform	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C	1120	49.62	1120	49.62
E1	20	0.89	1140	50.51
E2	9	0.40	1149	50.91
F	6	0.27	1155	51.17
G	2	0.09	1157	51.26
NP	95	4.21	1252	55.47
PD	1005	44.53	2257	100.00

Frequency Table of Waveform Event Transitions

The FREQ Procedure

trans1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C to E1	18	0.80	18	0.80
C to F	6	0.27	24	1.07
C to G	2	0.09	26	1.16
C to NP	87	3.87	113	5.02
C to PD	1005	44.69	1118	49.71
E1 to C	11	0.49	1129	50.20
E1 to E2	9	0.40	1138	50.60
E2 to C	5	0.22	1143	50.82
E2 to E1	2	0.09	1145	50.91
F to C	6	0.27	1151	51.18
G to C	2	0.09	1153	51.27
NP to C	93	4.14	1246	55.40
PD to C	1002	44.55	2248	99.96
PD to NP	1	0.04	2249	100.00

There is some additional output that comes after these two tables. You are welcome to look at it, but it is not important for this task.

We examine the two frequency tables. All the expected waveforms are present, and all of the listed transitions are possible. Thus there are no obvious errors.

It is suggested that at this point you use Notepad, or Wordpad to open the data file. Make a change, like delete one of the observations and save using a different name. Then run Error Checker on that file. Go back to the original data and make one of the values negative and save using a different name. Then run Error Checker on that file. This will help you see what happens when there are problems.

The last waveform in all the files has a shorter duration than if it had ended naturally. Given that these are shorter recordings (2.8 to 5.99 hours), the loss of even a single data point becomes more critical. Therefore we will not use the trimming program. The estimated means will all be biased, but it could be argued that they will be biased either way if you also assume that previous durations will be of shorter duration than longer durations. You could also argue that this approach preserves the count data. There is no perfect universal solution to this problem.

If you want, you can go back to the error checking program, to get frequency tables for each treatment. Activate the line:

```
Data one; set one; *if substr(insectno,1,1)="e" then output;
```

So that it looks like this:

```
Data one; set one; if substr(insectno,1,1)="a" then output;
```

Run the program and save the output. Then change it to:

```
Data one; set one; if substr(insectno,1,1)="b" then output;
```

Here is what you should get:

For treatment A:

Frequency Table of Waveform Event Transitions

The FREQ Procedure

waveform	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C	665	49.52	665	49.52
E1	14	1.04	679	50.56
E2	8	0.60	687	51.15
F	4	0.30	691	51.45
G	2	0.15	693	51.60
NP	32	2.38	725	53.98
PD	618	46.02	1343	100.00

Frequency Table of Waveform Event Transitions

The FREQ Procedure

trans1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C to E1	13	0.97	13	0.97
C to F	4	0.30	17	1.27
C to G	2	0.15	19	1.42
C to NP	27	2.02	46	3.44
C to PD	618	46.19	664	49.63
E1 to C	6	0.45	670	50.07
E1 to E2	8	0.60	678	50.67
E2 to C	5	0.37	683	51.05
E2 to E1	1	0.07	684	51.12
F to C	4	0.30	688	51.42
G to C	2	0.15	690	51.57
NP to C	31	2.32	721	53.89
PD to C	617	46.11	1338	100.00

For treatment B:

Frequency Table of Waveform Event Transitions

The FREQ Procedure

waveform	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C	455	49.78	455	49.78
E1	6	0.66	461	50.44
E2	1	0.11	462	50.55
F	2	0.22	464	50.77
NP	63	6.89	527	57.66
PD	387	42.34	914	100.00

Frequency Table of Waveform Event Transitions

The FREQ Procedure

trans1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C to E1	5	0.55	5	0.55
C to F	2	0.22	7	0.77
C to NP	60	6.59	67	7.35
C to PD	387	42.48	454	49.84
E1 to C	5	0.55	459	50.38

E1 to E2	1	0.11	460	50.49
E2 to E1	1	0.11	461	50.60
F to C	2	0.22	463	50.82
NP to C	62	6.81	525	57.63
PD to C	385	42.26	910	99.89
PD to NP	1	0.11	911	100.00

Now open the Analysis program, change the file name, and change the file name in the ODS statement. Turn off the transformations (if needed). I chose to name the output file “ExampleWT AnalysisNT”. The transformations should all be in green letters:

```
***********************************************************************
***********************************************************************
***********************************************************************
***********************************************************************
*******                                                      **********
*******                                                      **********
*******                                                      **********
*******                   TRANSFORMATIONS                    **********
*******                                                      **********
*******                                                      **********
***********************************************************************
***********************************************************************
***********************************************************************
***********************************************************************;

/*
***********************************************************************
***********************************************************************
**** Here is a good generic set of transformations                 ****
**** Counts are sqrt transformed, durations are log transformed    ****
***********************************************************************
***********************************************************************;
Data Ebert; Set Ebert;
PrcntPrbC = arsin(sqrt(PrcntPrbC/100));
PrcntPrbE1 = arsin(sqrt(PrcntPrbE1/100));
```

The key here is the “/*”.

Run the program. Then change the ODS file name. I chose “ExampleWT AnalysisT”.

To activate transformations, delete the only example of “/*” in this portion of green text.

It should now look like:

```
***********************************************************************
***********************************************************************
***********************************************************************
***********************************************************************
*******                                                      **********
*******                                                      **********
*******                                                      **********
*******                   TRANSFORMATIONS                    **********
```

```
*******                                                                **********
*******                                                                **********
*************************************************************************
*************************************************************************
*************************************************************************
*************************************************************************;


*************************************************************************
*************************************************************************
**** Here is a good generic set of transformations                      ****
**** Counts are sqrt transformed, durations are log transformed         ****
*************************************************************************
*************************************************************************;
Data Ebert; Set Ebert;
PrcntPrbC = arsin(sqrt(PrcntPrbC/100));
PrcntPrbE1 = arsin(sqrt(PrcntPrbE1/100));
```

Run the program.

Open the two output files in a word processing program like Microsoft Word®. Wait a few minutes to allow the program to process the file (let it finish the page count, if this is done automatically). Select all, then copy and paste into a spreadsheet program like Microsoft Excel®. I paste the non-transformed results starting in column A row 1. I paste the transformed results into the same worksheet starting in column L row 1. Give the computer time to finish each task. This is very important if you have several treatments, and it can take a couple of minutes. It should be fairly speedy with this example.

Since this output is a mean comparison, you can get the spreadsheet to extract only the results that you want to see. Mostly, this involves the means, standard errors, degrees of freedom, F statistic, and significance.

Find the first instance in column G, where you see "Pr > F". In this case this should be row 46. Copy this text. (This is important because (for whatever reason) you cannot type this in and have the next step work. The result seems to have the same form in either case, but the result is very different.)

In column U row 46 type in =if($G46="Pr > F",$B6,"") BUT when you get to the point where you would type in Pr > F, use the paste (control-P is the keyboard shortcut) command. Then type in the ending parentheses and the remainder of the statement. You should now see "TmFrstPrbFrmStrt" in this cell.

Select this cell and cells in columns V through AC. Then fill right (Control-R is the keyboard shortcut). The mean for treatment 1 is in column B, row 51 (cell B51), while the standard error is in cell C52. The mean and standard error for treatment 2 is in the row below that B52, and C52 respectively. In V46 change the statement "=if($G46="Pr > F",$B6,"")" to look like this "=if($G46="Pr > F",$B51,"")" where the only difference is the highlighted portion. In the remaining cells, change each $B6 to whatever cell has the information that you want. This example shows mean, standard error for treatment 1, mean and standard error for treatment 2, numerator degrees of freedom, denominator degrees of freedom, F value, and the probability of getting an F value as large or larger given that the null hypothesis is true.

You should now have row 46 in columns U through AC filled with the following:

TmFrstPrbFrmStrt	261.61	109.79	220.19	173.59	1	5	.11	.759

Fill down to the end of the SAS output.

Copy columns U through AA. Use the paste special command to paste values into a new worksheet. This will fill columns A through column I. There will still be a large number of blank rows that we need to permanently remove. There are at least two ways to do this.

3) Select columns A through I and go to filter (Editing menus, "Sort and Filter" tab). Select filter. In row 1, each cell has a down arrow icon. Select this, go to the bottom of the list, and uncheck the box labeled "(blanks)". The blank lines will disappear, but they are still there. Select all your data, copy and paste values. Now the blank lines are gone. NOTE: If you paste into the same worksheet, then you will need to go back to the box labeled "(blanks)" and check the box again. Otherwise most of the data will be hidden in the blank rows that are not being displayed.
4) In column J type in the number 1. Find the end of the data by dragging the slide bar to the bottom. Holding down the shift key, select the cell in column J that is next to the last set of data. This should select all the cells from row 1 to row 7331. In the Edit menu select Fill, then series. In the pop-up menu make sure that the "step value" is 1, and click ok. Select columns A through J (in that order), and then choose "sort and filter" from the edit menu. In the menu that appears choose "sort Z to A" All the results are now at the top, so copy them (Cells A1 through J98), and paste them into a new worksheet. Click anywhere to deselect them, then select columns J through A (in that order). Go back to the Edit menu, then "sort and filter" and finally "sort A to Z". The last step is to delete the contents of column J.

The second approach is a bit more complex, but the techniques involved can be useful elsewhere.

You should end up with a table that looks something like this:

TmFrstPrbFrmStrt	261.61	109.79	220.19	173.59	1	5	0.11	0.759
CtoFrstE1	4.4	1.873	6.5	2.9614	1	5	0.53	0.5003
NumF	0.8	0.3399	0.6667	0.4389	1	6	0	0.9728
DurFrstPrb	4680.79	1821.67	1612.23	2351.77	1	6	0.32	0.5943
DurScndPrb	527.65	193.15	198.12	249.36	1	6	0.74	0.4238
ShrtCbfrE1	3660.85	739.55	2498.58	1169.33	1	5	0.7	0.4412
DurScndZ	256.12	153.82	391.54	198.59	1	6	0.08	0.7849
TtlDurF	2668.24	892.86	2857.98	1093.52	1	3	0	0.9526
DurNnprbBfrFrstE1	710.76	659.29	2411.34	851.14	1	6	1.68	0.2422
meanpd	5.2649	6.2737	19.0899	8.0993	1	6	1.67	0.2434

meanPDS	5.2649	6.2737	19.0899	8.0993	1	6	1.67	0.2434
meanNPdPrb	22.2968	6.0257	16.2936	7.7791	1	6	1.19	0.3172
meanF	2286.78	995.34	2857.98	1219.04	1	3	0.18	0.7012
TmStrtEPGFrstE	7303.62	3036.46	10418	3920.05	1	6	0.33	0.5862
TmFrmFrstPrbFrstE	7042.01	3032.22	10271	3914.58	1	6	0.34	0.5809
TmBegPrbFrstE	3660.85	739.55	2498.58	1169.33	1	5	0.7	0.4412
NumG	0.4	0.3266	1.11E-16	0.4216	1	6	0.56	0.4816
NumPrbsAftrFrstE	1.8	0.6182	1.3333	0.7981	1	6	0.22	0.6561
NmbrShrtPrbAftrFrstE	0.6	0.2494	0.6667	0.322	1	6	0.03	0.8754
NumE1	2.8	1.0132	2	1.3081	1	6	0.69	0.4389
NumLngE1BfrE2	0.2	0.2211	0.3333	0.2854	1	6	0.14	0.7246
NumSnglE1	1	0.5963	1.3333	0.7698	1	6	0.08	0.7849
NumE2	1.6	0.5735	0.3333	0.7404	1	6	2.25	0.184
NumLngE2	0.8	0.3399	0.3333	0.4389	1	6	0.59	0.4721
DurFirstE	4141.32	3061.18	667.22	4840.14	1	5	0.47	0.5235
CntrbE1toE	39.8506	17.724	85.1326	28.0241	1	5	1.14	0.335
DurE1FlwdFrstSusE2	1350.2	1067.89	3466.14	1849.65	1	2	0.66	0.5011
DurE1FlldFrstE2	1027.09	821.34	3466.14	1642.68	1	3	1.25	0.345
PotE2Indx	43.8914	19.9932	34.951	39.9863	1	3	0.02	0.8897
TtlDurE	6904.12	2878.95	4547.59	4552.01	1	5	0.05	0.8344
TtlDurE1	1576.67	1070.85	3391.09	1693.16	1	5	1.21	0.3215
TtlDurE1FlldSusE2	1360.37	1064.1	3466.14	1843.07	1	2	0.65	0.5049
TtlDurE1FlldE2	1125.37	801.64	3466.14	1603.29	1	3	1.07	0.3776
TtlDurSnglE1	639.94	349.76	927	428.36	1	3	0.51	0.5282
TtlDurE1FllwdE2PlsE2	7784.68	3565.16	5779.14	7130.32	1	3	0.01	0.9201
TtlDurE2	6659.32	3840.51	2313	7681.02	1	3	0.09	0.7813
MnDurE1	540.38	247.65	1204.78	391.57	1	5	1.88	0.2287
MnDurE2	5468.6	4104.98	2313	8209.96	1	3	0	0.9745
NumPrbs	6.2	5.7411	21	7.4117	1	6	2.88	0.1406
NmbrC	9.6	5.1635	23.3333	6.6661	1	6	3.09	0.1292
NmbrShrtC	2	5.325	14.3333	6.8745	1	6	2.39	0.1732
NumNP	6.4	5.5714	21	7.1926	1	6	2.95	0.1369
NmbrPD	123.6	34.029	129	43.9313	1	6	0.01	0.9367
NmbrPDL	0	.	0	.	1	6	.	.
NmbrPDS	123.6	34.029	129	43.9313	1	6	0.01	0.9367
NmbrE1e	0	.	0	.	1	6	.	.
TtlDurC	10793	1898.84	9980.6	2451.39	1	6	0.01	0.9265
TotDurNnPhlPhs	14651	2872.31	17003	4541.51	1	5	0.27	0.6232
TtlDurNP	1090.88	692.94	2852.72	894.58	1	6	2.03	0.2044
TtlDurPD	653.8	133.24	952.13	172.01	1	6	1.79	0.2295
TtlDurPDS	653.8	133.24	952.13	172.01	1	6	1.79	0.2295
TtlPrbTm	20464	2207.86	14918	2850.34	1	6	2.18	0.1905

MnDurNP	155.51	43.0999	155.5	55.6417	1	6	0.04	0.8554
MnDurC	1151.8	251.06	723.85	324.11	1	6	1.44	0.275
TmFrstSusE2	13380	3467.37	15564	4476.36	1	6	0.43	0.5363
TmFrstSusE2FrstPrb	13306	3477.56	15487	4489.52	1	6	0.43	0.5344
TmFrstSusE2StrtPrb	4703.13	1225.07	6662.9	2121.88	1	2	0.52	0.5453
TmFrstE2StrtEPG	10520	3114.97	15564	4021.41	1	6	1.45	0.2733
TmFrstE2FrmFrstPrb	10258	3152.39	18021	4984.37	1	5	1.87	0.2302
TmFrstE2FrmPrbStrt	5153.05	976.13	6662.9	1952.26	1	3	0.4	0.5741
TtlDurNp1	344.99	228.68	728.6	295.23	1	6	0.72	0.4281
TtlDurNp2	187.59	250.67	820.5	323.62	1	6	1.75	0.2335
TtlDurNp3	262.35	220.85	889.71	285.12	1	6	1.32	0.2943
TtlDurNp4	12.288	10.9907	0	17.3779	1	5	0.36	0.5761
TtlDurNp5	26.26	29.941	65.65	47.3409	1	5	0.49	0.5133
TtlDurNp6	257.4	219.61	555.22	347.23	1	5	0.97	0.3693
NumPDS1	27.2	8.6772	12	11.2022	1	6	0.72	0.4301
NumPDS2	21.8	11.8201	28.6667	15.2597	1	6	0.13	0.7301
NumPDS3	23.4	11.6636	29	15.0577	1	6	0.18	0.6877
NumPDS4	17.6	11.936	38.5	18.8725	1	5	1.2	0.3238
NumPDS5	13.6	10.2795	22.5	16.2533	1	5	0.12	0.7393
NumPDS6	20	8.3964	28.5	13.2759	1	5	0.58	0.4824
MnDurPdS1	4.4634	2.7191	7.704	3.5103	1	6	0.02	0.9028
MnDurPdS2	5.2739	13.4385	34.7668	17.3491	1	6	1.5	0.2663
MnDurPdS3	5.4538	0.5362	5.5115	0.5362	1	4	0	0.9501
MnDurPdS4	5.2043	0.3766	4.8146	0.4613	1	3	0.4	0.5709
MnDurPdS5	5.1238	0.3201	4.8038	0.4526	1	1	0.33	0.6667
MnDurPdS6	6.0054	1.2864	6.5689	1.5755	1	3	0.07	0.8033
NumF1	0.4	0.2494	0.6667	0.322	1	6	0.43	0.537
NumF2	0.2	0.2211	0.3333	0.2854	1	6	0.14	0.7246
NumF3	0	.	0	.	1	6	.	.
NumF4	0.4	0.3578	0	0.5657	1	5	0.36	0.5761
NumF5	0.2	0.1789	0	0.2828	1	5	0.36	0.5761
NumF6	0	.	0	.	1	5	.	.
TtlDurF1	936.1	652.98	1560.17	842.99	1	6	0.43	0.5375
TtlDurF2	207.09	228.94	345.14	295.56	1	6	0.14	0.7246
TtlDurF3	0	.	0	.	1	6	.	.
TtlDurF4	421.13	376.67	5.68E-14	595.56	1	5	0.36	0.5761
TtlDurF5	36.628	32.7611	2.13E-14	51.7998	1	5	0.36	0.5761
TtlDurF6	0	.	0	.	1	5	.	.
NumPrb1	2.6	1.6653	5	2.1499	1	6	0.67	0.4448
NumPrb2	1.8	1.0242	4.3333	1.3222	1	6	2.13	0.1949
NumPrb3	2.6	4.0078	12.3333	5.174	1	6	2.46	0.168

NumPrb4	1.2	0.1789	1	0.2828	1	5	0.36	0.5761
NumPrb5	1.2	0.228	1.5	0.3606	1	5	0.49	0.5133
NumPrb6	1.8	0.3633	2.5	0.5745	1	5	1.1	0.3423
TmFrstCFrstPD	1158.33	1112.01	2025.39	1435.6	1	6	0.85	0.3911
TmEndLstPDEndPrb	1092.91	730.61	102.59	943.22	1	6	1.16	0.3235
TmEndPDBegE1FllwdSusE2	4338.75	4292.92	46.7	7435.55	1	2	0.25	0.6667
TmLstPdEndRcrd	9093.71	4241.9	783.19	5998.95	1	4	1.28	0.3212
TmLstE2EndRcrd	11042	6683.81	0	0	0	.	.	.
maxE2	6225.51	3913.37	2313	7826.73	1	3	0.2	0.6851
DurNpFllwFrstSusE2	87.57	36.3	51.27	51.336	1	1	0.33	0.6667
PrcntPrbC	53.1862	9.9006	75.2598	12.7816	1	6	2.26	0.1833
PrcntPrbE1	7.6828	5.2276	11.0734	6.7488	1	6	0.06	0.8108
PrcntPrbE2	25.1941	12.6799	3.6809	16.3697	1	6	1.45	0.2733
PrcntPrbF	8.2314	4.5384	9.9859	5.859	1	6	0.04	0.8435
PrcntPrbG	5.7054	4.6585	4.44E-16	6.014	1	6	0.56	0.4816
PrcntE2SusE2	62.5	23.9357	100	47.8714	1	3	0.49	0.534

The values are the mean value of the variables as calculated for each insect (these are "per insect" from values calculated "by insect").

Since the data are in seconds, these means of durations are in seconds. The data is not accurate to hundredths of a second. You can use the format cells command in Excel to remove the extra precision. This is as far as this tutorial can go. Deciding on what is significant and what to present is left to the scientist performing the data analysis. We will provide some suggestions in a different tutorial.

Bonus Material

The task is to print out a list of all the results for each insect to facilitate a comparison between the output of this program and that of other programs that do not provide a complete analysis. In the base SAS you will need to put a print statement at the end of the program.

Data Ebert; Set Ebert; Proc Print; Run;

You will then need to either run the entire program, or if you have already run the program you can highlight this section and run only this section.

If you use Enterprise Guide, there will be an "Output Data" tab next to the Log tab. Below the "Program" tab there will be a name with a small down arrow. Click on the down arrow to see a list of all the data sets that are in memory (there should be 12 of them). Mouse click on the one named Ebert. Select all, copy and paste into Excel. The only problem here is that the names of the variables do not get copied.

One way to get the variable names included is to use Proc Export:

```
proc export data=Ebert outfile='C:\Users\Location 3\Ebert.csv'
dbms=csv replace;
```

Before comparing output, make sure that all transformations have been turned off. Then export the dataset Ebert. Open the file in Excel. Copy the data, and then click on the down arrow under Paste menu. Select "Paste Special" and from that pop-up menu select "Transpose." This is now in the format that will match what you will get from the Sarria workbook.

To get the data into Sarria you will have to recode all the behaviors. You can do this in SAS or in Excel. Here is one approach:

4) Open the data file "ExampleWT.csv" using Excel or other spreadsheet program.
5) There are two major tasks: convert waveforms to a numeric code, and change the duration to a time from beginning of file.
    a. To convert to a numeric code go to cell D2 and type in "=if(A2="NP",1,"")" In the next few columns type in the same formula, but change the behavior and change the number to match the codes in the Sarria workbook. NP=1, C=2, E1=4, E2=5, and G=7. There is no code for D, so I chose D=11. In cell J2 type in "=Sum(D2,I2)" and then fill down to the end of the file {I would save this in a special place to avoid having to do it over again later}. Copy and paste values into column A, and delete columns D through J. NOTE code 11 is treated as PDL in Sarria. This will alter some calculations!
    b. Find the end of each insect and insert two empty rows.
    c. In cell D2 type in 0.
    d. In cell D3 type in "=D2+c2" and fill down.
    e. Go to the end of each insect. In the first blank line at the end of the insect type in 12 for the behavior. This is the end of file code used by the Sarria workbook.
    f. At the start of the next insect type in a 0.
6) Copy and paste values each insect into the Sarria workbook. Then follow the instructions for running the workbook.

The first four insects give the following output (Font is at 4 point to make everything fit. It is recommended to copy this table and paste into Excel, then adjust the font size):

Warning: In using the Sarria workbook there may be some trouble with European versus American number formats. One solution is to do a global replace finding commas and replacing them with periods.

EBERT					Sarria				
insectno	a1	a2	a3	a4	insectno	a1	a2	a3	a4

TmFrstPrbFrmStrt	93.11	232.28	208.1	42.8	Time to 1st probe from start of EPG	93.11	232.28	208.1	42.8
CtoFrstE1	7	3	10	1	Number of probes to the 1st E1	7	3	10	1
NumF	0	1	1	0	Number of F	0	1	1	0
DurFrstPrb	101.54	57.6	4452.69	8870.27	Duration of 1st probe	101.54	57.6	4452.69	8870.27
DurScndPrb	51.55	476.22	68.54	825.74	Duration of 2nd probe	51.55	476.22	68.54	825.74
ShrtCbfrE1	2409.09	1992.22	3004.94	4452.94	Duration of the shortest C wave before E1	2409.09	1992.22	3004.94	4452.94
DurScndZ	191.79	27.9	811.01	123.87	Duration of the second nonprobe period	191.79	27.9	811.01	123.87
TtlDurF		1539	4176.95		Total duration of F		1539	4176.95	
DurNnprbBfrFrstE1	422.8	319.47	2036.95	42.8	Duration of nonprobe period before the 1st E	422.8	319.47	2036.95	42.8
meanpd	5.067	5.41475	4.720542	5.203692	Mean duration of pd	5.067	5.432362	4.720542	5.203692
meanPDL					Mean duration of pd-L				
meanPDS	5.067	5.41475	4.720542	5.203692	Mean duration of pd-S	5.067	5.432362	4.720542	5.203692
meanNPdPrb	7.142857	33.33333	15.09091	21.66667	Average number of pd per probe	7.142857	33.16667	15.09091	21.66667
meanF		1539	4176.95		Mean duration of F		1539	4176.95	
TmStrtEPGFrstE	3802.83	2845.51	18197.2	4495.74	Time from start of EPG to 1st Ey	3802.83	2845.51	18197.2	4495.74
TmFrmFrstPrbFrstE	3709.72	2613.23	17989.1	4452.94	Time from 1st probe to 1st Ey	3709.72	2613.23	17989.1	4452.94
TmBegPrbFrstE	2409.09	1992.22	3004.94	4452.94	Time from the beginning of that probe to 1st Ey	2409.09	1992.22	3004.94	4452.94
NumG	0	0	0	0	Number of G	0	0	0	0
DurG					Duration of G				
MeanG					Mean duration of G				
NumPrbsAftrFrstE	0	3	1	2	Number of probes after 1st E	0	3	1	2
NmbrShrtPrbAftrFrstE	0	1	1	0	Number of probes (shorter than 3 minutes) after 1st E	0	1	1	0
NumE1	1	5	1	5	Number of E1	1	5	1	5
NumLngE1BfrE2	0	1	0	0	Number of E1 (longer than 10 minutes) followed by E2	0	1	0	0
NumSnglE1	0	3	1	1	Number of single E1	0	3	1	1
NumE2	1	1	0	4	Number of E2	1	1	0	4
NumLngE2	1	1	0	2	Number of sustained E2 (longer than 10 minutes)	1	1	0	2
DurFirstE	17766.39	18.02	1316.41	1309.02	Duration of 1st E	17766.39	18.02	1316.41	1309.02
CntrbE1toE	0.22869	70.26518	100	13.68873	Contribution of E1 to phloem phase (%)	0.22869	70.26518	100	13.68873
DurE1FlwdFrstSusE2	40.63	3466.14		543.84	Duration the E1 followed by first sustained E2 (>10 min)	40.63	3466.14		543.84
DurE1FlldFrstE2	40.63	3466.14		543.84	Duration the E1 followed by the first E2	40.63	3466.14		543.84
PotE2Indx	100	34.95098		35.40995	Potential E2 index	100	34.95098		35.40995
TtlDurE	17766.39	7778.76	1316.41	6782.88	Total duration of E	17766.39	7778.76	1316.41	6782.88
TtlDurE1	40.63	5465.76	1316.41	928.49	Total duration of E1	40.63	5465.76	1316.41	928.49
TtlDurE1FlldSusE2	40.63	3466.14		574.35	Total duration of E1 followed by sustained E2 (>10 min)	40.63	3466.14		574.35
TtlDurE1FlldE2	40.63	3466.14		862.65	Total duration of E1 followed by E2	40.63	3466.14		862.65
TtlDurSnglE1		537.58	1316.41	65.84	Total duration of single E1		537.58	1316.41	65.84
TtlDurE1FllwdE2PlsE2	17766.39	5779.14		6717.04	Total duration of E1 followed by E2 and E2	17766.39	5779.14		6717.04
TtlDurE2	17725.76	2313		5854.39	Total duration of E2	17725.76	2313		5854.39
MnDurE1	40.63	1093.152	1316.41	185.698	Mean duration of E1	40.63	1093.152	1316.41	185.698
MnDurE2	17725.76	2313		1463.598	Mean duration of E2	17725.76	2313		1463.598
NumPrbs	7	6	11	3	Number of probes	7	6	11	3
NmbrC	7	11	13	7	Number of C	7	11	13	7

NmbrShrtC	4	2	3	0	Number of short probes (C<3 minutes)	4	2	3	0
NumNP	7	6	12	3	Number of np	7	6	12	3
NmbrPD	50	200	166	65	Number of pd	50	199	166	65
NmbrPDL	0	0	0	0	Number of pd-L	0	0	0	0
NmbrPDS	50	200	166	65	Number of pd-S	50	199	166	65
NmbrE1e	0	0	0	0	Number of E1e	0	0	0	0
TtlDurC	3380.03	11628.33	12980.36	14534.66	Total duration of C	3380.03	11628.33	12980.36	14534.66
TtlDurE1e					Total duration of E1e				
TotDurNnPhlPhs	3802.83	13784.22	20220.98	14789.88	Total duration of no phloematic phase	3802.83	13784.22	20220.98	14789.88
TtlDurNP	422.8	616.89	3063.67	255.22	Total duration of np	422.8	616.89	3063.67	255.22
TtlDurPD	253.35	1082.95	783.61	338.24	Total duration of pd	253.35	1081.04	783.61	338.24
TtlDurPDL					Total duration of pd-L				
TtlDurPDS	253.35	1082.95	783.61	338.24	Total duration of pd-S	253.35	1081.04	783.61	338.24
TtlPrbTm	21146.42	20946.09	18473.72	21317.54	Total probing time	21146.42	20946.09	18473.72	21317.54
MnDurNP	60.4	102.815	255.3058	85.07333	Mean duration of np	60.4	102.815	255.3058	85.07333
MnDurC	482.8614	1057.121	998.4892	2076.38	Mean duration of C	482.8614	1057.121	998.4892	2076.38
TmFrstSusE2	3843.46	14945.14	21537.39	5039.58	Time to from start of EPG 1st sustained E2 (10 minutes)y	3843.46	14945.14	21537.39	5039.58
TmFrstSusE2FrstPrb	3750.35	14712.86	21537.39	4996.78	Time from 1st probe to 1st sustained E2 (10 minutes)y	3750.35	14712.86	21329.29	4996.78
TmFrstSusE2StrtPrb	2449.72	6662.9		4996.78	Time from the beginning of that probe to 1st sustained E2 (10 minutes)y	2449.72	6662.9		4996.78
TmFrstE2StrtEPG	3843.46	14945.14	21537.39	5039.58	Time from start of EPG to 1st E2y	3843.46	14945.14	21537.39	5039.58
TmFrstE2FrmFrstPrb	3750.35	14712.86	21329.29	4996.78	Time from 1st probe to 1st E2y	3750.35	14712.86	21329.29	4996.78
TmFrstE2FrmPrbStrt	2449.72	6662.9		4996.78	Time from the beginning of that probe to 1st E2y	2449.72	6662.9		4996.78
TtlDurNp1	422.8	319.47	208.1	42.8	Total duration of np during the 1st hour	422.8	319.47	208.1	42.8
TtlDurNp2	0	0	937.95	0	Total duration of np during the 2nd hour	0	0	937.95	0
TtlDurNp3	0	213.71	759.6	212.42	Total duration of np during the 3rd hour	0	213.71	759.6	212.42
TtlDurNp4	0	0	0	0	Total duration of np during the 4th hour	0	0	0	0
TtlDurNp5	0	0	131.3	0	Total duration of np during the 5th hour	0	0	131.3	0
TtlDurNp6	0	83.71	1026.72	0	Total duration of np during the 6th hour	0	83.71	1026.72	0
NumPDS1	46	24	0	53	Number of pd-S during the 1st hour	46	24	0	53
NumPDS2	4	64	14	12	Number of pd-S during the 2nd hour	4	64	14	12
NumPDS3	0	59	26	0	Number of pd-S during the 3rd hour	0	59	26	0
NumPDS4	0	16	61	0	Number of pd-S during the 4th hour	0	16	61	0
NumPDS5	0	0	45	0	Number of pd-S during the 5th hour	0	0	45	0
NumPDS6	0	37	20	0	Number of pd-S during the 6th hour	0	36	20	0
MnDurPdS1	5.080652	4.410417	0	5.241321	Average duration of pd-S during the 1st hour	5.080652	4.410417		5.241321
MnDurPdS2	4.91	4.593125	4.862143	5.0375	Average duration of pd-S during the 2nd hour	4.91	4.593125	4.862143	5.0375
MnDurPdS3		4.941525	4.953077		Average duration of pd-S during the 3rd hour		4.941525	4.953077	
MnDurPdS4		5.116875	4.512295		Average duration of pd-S during the 4th hour		5.116875	4.512295	
MnDurPdS5			4.803778		Average duration of pd-S during the 5th hour			4.803778	
MnDurPdS6		8.370811	4.767		Average duration of pd-S during the 6th hour		8.550278	4.767	
NumF1	0	1	1	0	Number of F during the 1st hour	0	1	1	0
NumF2	0	0	1	0	Number of F during the 2nd hour	0	0	1	0
NumF3	0	0	0	0	Number of F during the 3rd hour	0	0	0	0

NumF4	0	0	0	0	Number of F during the 4th hour	0	0	0	0
NumF5	0	0	0	0	Number of F during the 5th hour	0	0	0	0
NumF6	0	0	0	0	Number of F during the 6th hour	0	0	0	0
TtlDurF1	0	1539	3141.52	0	Total duration of F during the 1st hour	0	1539	3141.52	0
TtlDurF2	0	0	1035.43	0	Total duration of F during the 2nd hour	0	0	1035.43	0
TtlDurF3	0	0	0	0	Total duration of F during the 3rd hour	0	0	0	0
TtlDurF4	0	0	0	0	Total duration of F during the 4th hour	0	0	0	0
TtlDurF5	0	0	0	0	Total duration of F during the 5th hour	0	0	0	0
TtlDurF6	0	0	0	0	Total duration of F during the 6th hour	0	0	0	0
NumPrb1	7	3	1	1	Number of probes during the 1st hour	7	3	1	1
NumPrb2	1	1	5	1	Number of probes during the 2nd hour	1	1	5	1
NumPrb3	1	2	5	3	Number of probes during the 3rd hour	1	2	5	3
NumPrb4	1	1	1	1	Number of probes during the 4th hour	1	1	1	1
NumPrb5	1	1	2	1	Number of probes during the 5th hour	1	1	2	1
NumPrb6	1	3	2	1	Number of probes during the 6th hour	1	3	2	1
TmFrstCFrstPD	304.3	121.73	5268.41	12.92	Time from the beginning of the 1st probe to first pd	304.3	121.73	5268.41	12.92
TmEndLstPDEndPrb	35.04	243.3	44.47	4653.53	Time from the end of the last pd to the end of the probe	35.04	243.3	44.47	4653.53
SumPDII1					Total duration of subphase II1 fo the pd				
SumPDII2					Total duration of subphase II2 fo the pd				
SumPDII3					Total duration of subphase II3 fo the pd				
TmEndPDBegE1FllwdSusE2	44.96	46.7		12924.58	Time from the end of the last pd to the beginning of the E1 followed by the sustained E2 (>10 min)	44.96			12924.58
TmLstPdEndRcrd	17811.35		1173.93	17313.22	Time from the end of the last pde to the end of the EPG record (Z)	17811.35	47.01	1173.93	17313.22
TmLstE1EndRcrd	.	.	.	.	Time from the beginning of E1 to the end of the EPG record (Z)				
TmLstE2EndRcrd	17725.76			4358.13	Time from the beginning of E2 to the end of the EPG record (Z)	17725.76			4358.13
maxE2	17725.76	2313		4358.13	Duration of the longest E2	17725.76	2313		4358.13
DurNpFllwFrstSusE2		51.27		123.87	Duration of np just after the probe of the first sustained E2		51.27		123.87
DurTrmNpFllwFrstSusE2					Duration of np just after the probe of the first sustained E2 if it lasts untill the end of the recording				
PrcntPrbC	15.98393	55.51552	70.26392	68.18169	% of probing spent in C	15.98393	55.51552	70.26392	68.18169
PrcntPrbE1	0.192137	26.09442	7.125852	4.355521	% of probing spent in E1	0.192137	26.09442	7.125852	4.355521
PrcntPrbE2	83.82393	11.04263	0	27.46278	% of probing spent in E2	83.82393	11.04263	0	27.46278
PrcntPrbF	0	7.347433	22.61023	0	% of probing spent in F	0	7.347433	22.61023	0
PrcntPrbG	0	0	0	0	% of probing spent in G	0	0	0	0
PrcntE2SusE2	100	100		50	% E2 >10 min	100	100		50

Text highlighted in green has values that are a disagreement between Ebert 1.0 and the Sarria workbook. The issue here is that the Sarria workbook appears to have ignored the last pd for insect 2. This is an unusual insect because the last pd is also the last behavior in that recording.

These are recordings from aphids, and we did not have to trick the program into working with a behavior that was not found in aphids (unlike for Psyllids that have the D waveform not found in aphids.). Also possibly a little luck. At any rate both programs arrived at the same count for the number of C (pathway).

This concludes this part of the tutorial. You should now be comfortable with conducting an analysis of your data. Now there are some choices that you will need to make while doing these analyses. These are the choices that make your experiment unique. All we can do is point out some issues, and possibly some consequences of different choices. Please see “EPG Analysis Choices” for some discussion.