

## Instructions for Using the Excel Implementation of Rate Estimation Theory

This spreadsheet carries out the calculation of rate estimates and odds for acquisition and extinction as detailed in Rate Estimation Theory (RET). It illustrates and implements the mathematical development of this theory of conditioning given in the following references:

Gallistel, C. R. (1990). The organization of learning (Cambridge, MA, Bradford Books/MIT Press).

Gallistel, C. R., and Gibbon, J. (2000). Time, rate and conditioning. *Psychological Review* 107, 289-344.

Gallistel, C. R., and Gibbon, J. (2002). The symbolic foundations of conditioned behavior (Mahwah, NJ, Lawrence Erlbaum Associates).

This spreadsheet should run on Excel 5 or higher on either a Mac or a PC. If you encounter problems, please notify me: [galliste@ruccs.rutgers.edu](mailto:galliste@ruccs.rutgers.edu)

### Basic Steps

- 1) Open the workbook named RET.xls
- 2) Check that it opened to the "Lng Model" sheet (look at sheet tabs at bottom left of workbook window). Check also that the cell in the upper-left corner of the window is A1. If it isn't, use the scroll bars to make it so, that is, scroll to top of sheet and to left edge of sheet.
- 3) The seventh row and down of the first six columns (Columns A-F) are the input cells. To calculate the rate estimates and other quantities for a given experimental protocol, you enter an experimental protocol into these cells. The other cells on the sheet have been protected so that you cannot accidentally overwrite them. Also, the workbook itself is locked, so that you cannot save changes to it. If something goes wrong, just close the workbook and reopen it.
- 4) To see what protocols look like and to get ready-made the protocols for many well-known conditioning experiments, click on the Protocols worksheet (the sheet tabs are at the bottom left of the workbook window). To understand these protocols and to make ones of your own, see the section below on "Interpreting Protocols."
- 5) To enter a ready-made protocol into the Lng Model, select it, copy it, and paste it in. To select a protocol, begin at the grayed blank cell at the upper-most left in Row 3, the row immediately below the rows with column-heading text, go over 6 columns to the US column and down 104 rows to Row 107. Copy this block of cells (a block that is 6 columns wide and 104 rows long). Go back to Lng Model sheet, select the uppermost, leftmost input cell (Cell A3, which is grayed) and paste. Don't worry about pasting into the wrong location. The sheet won't let you, because all but the input cells are protected.

6) To see the results, scroll straight down until the graphs appear (Row 140). (See below for a guide to these graphs.)

7) To avoid scrolling back and forth between input and results, open additional windows, size them to show only what you want to see, and tile them. The most convenient configuration is to have a window on the left open to the top of the input columns and showing only those columns, a window in the middle, showing the graphs of the results, and a window on the right showing the ready-made protocols. You may also want a window open on the matrices and the numerical results, which are located in Rows 115-136.

8) Saving results. The workbook is locked. To save results, use the Save As command and give the workbook another name. The copy saved under this other name is not locked.

### **Interpreting and Making Protocols**

A protocol specifies the times at which events happen during training. These times are in the leftmost column of the protocol. An event is the onset or offset of a CS or the occurrence of a US. If a given CS is on or comes on at the time of an event, then there is a 1 in the column for that CS. Thus, the onset of a CS is indicated by the appearance of a 1 in the column for that CS and the offset of a CS is indicated by the disappearance of 1's in that column. If a CS remains on through several other events, for example, the coming on and going off of other CS's or the occurrence of a US, then there will an uninterrupted sequence of 1's in that column from the time of onset of that CS until the time of its offset. CS1 is always the background (the experimental chamber). Note that a 1 appears in the column for CS1 at event time 0 and for every event time thereafter, because event time 0 is when the animal is placed in the training situation and the chamber is present at all times thereafter. When an isolated 1 appears in the column for a CS, it means that the CS came on at the time indicated for that row and went off at the time indicated for the next row. This interpretation does not, however, apply to USs, because USs are assumed to be point events. Their duration is assumed to be negligible; hence, they do not have separate on and off times. When a 1 appears in the US column, it means that a US "happened" at that time. All CSs marked with a 1 in the row *preceding* the row (event time) in which a US occurred are treated as present when the US occurred, even if they are not marked by a 1 in the row (at the time) where the US occurred. If a CS is marked with a 1 in the preceding row but not in the row for the US that means that the offset of the CS coincided with the US (a common occurrence in classical conditioning protocols). The model treats such CSs as having been present when the US occurred.

### **Making your own protocols**

You can enter protocols directly into the data entry columns of the learning model. But, pull down the Options Menu and select Calculate and then Manual before you do so. Otherwise, the model will recalculate the rates every time you key in a single entry. Once you have set the sheet for manual calculation, you can

key in one or more trials and then tell it to calculate, then key in more trials, and so on. A preferable way to enter your own protocols is to open a new worksheet, create your protocols on this new worksheet, then copy and paste them into the input columns of the learning model. You can also paste in chunks of protocol corresponding to one or a few trials in order to watch the progress of learning. Just paste successive chunks below one another in the input columns of the Lng Model.

## Guide to the Graphs

[See picture on next page]

The three column graphs at top left of the array of graphs show, from left to right, the temporal totals from which the coefficients of the temporal matrix are computed (T1,T2,T3,T4, T1.2, etc.), the number of USs that have occurred in the presence of each CS, and finally, the computed rates of US occurrence predicted by each CS (the corrected rates). Below these column graphs is the graph of C(p), the confidence function, which maps the computed likelihood that a CS affects the US into a confidence that it does. The computed odds that CS2, CS3, and or CS4 affect the rate of US occurrence are shown by a bar graph below the C(p) function. The extents of these bars give the x-values for the C(p) function. The y-values, that is, the confidences are given by the bar graph to the left of the C(p) function.

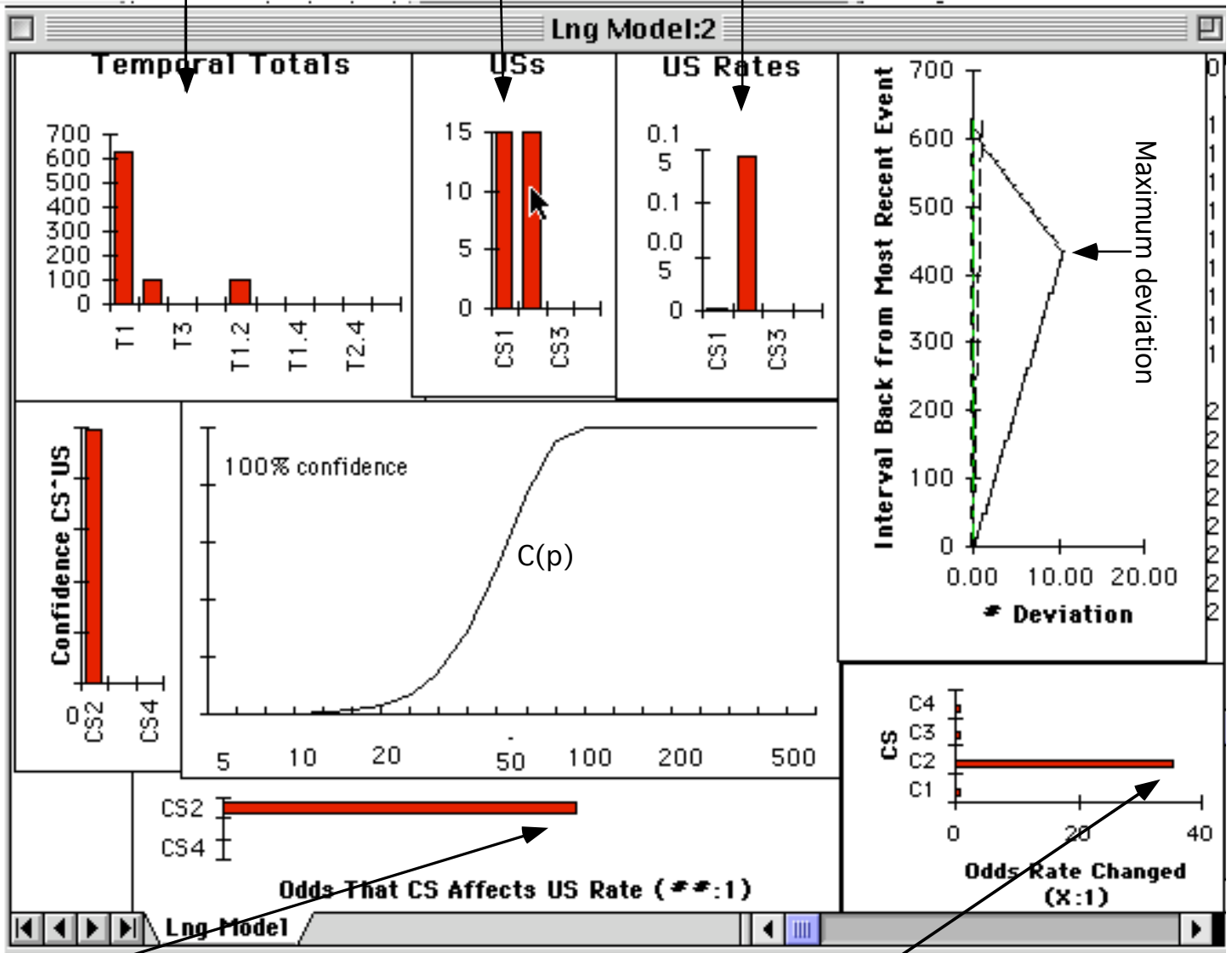
To the right of the C(p) graph is the graph of the analysis for stationarity (changes in rates of reinforcement). The vertical axis of this graph, which is conceptually its abscissa, is the length of the look backwards in time from the present moment (which is at 0 at the bottom of the vertical axis). The horizontal axis (the conceptual ordinate) gives the number of USs attributed to each CS over the backward interval in question minus the number of USs that should have been credited to its account over the same interval. The maximum deviation (high point) in the plot of the deviation versus time indicates the point at which the change in rate of reinforcement, if there has been one, is most likely to have occurred. When there has been no change in rate, the "deviations" bounce around between 0 and 1. (Deviations less than 1 are meaningless.)

The odds that there has been a change (bar graph at bottom right of window) are computed by comparing the rate estimate before the point of maximum deviation to the rate estimate after the point of maximum deviation.

Cumulative Exposures to individual CSs [T1, T2 (not labeled), T3, & T4 (not labeled)] and to pairwise combinations (T1.2, etc.)

Number of USs delivered in presence of each CS (2 & 4 not labeled)

Estimated rates of reinforcement (#s on ordinate distorted by compression)



The odds that CS2 affected the rate of reinforcement are almost 100:1. Of course, it only affected the rate before extinction began.

At the point in the simple extinction protocol where this graph was made, the odds that the rate of reinforcement of CS2 had changed were almost 40:1

## Understanding the Computations

There are two reasons for supplying this spreadsheet implementation of RET to the interested scientific public. The first is to enable those who don't want to be bothered with the mathematics to determine what RET predicts. The second is to make the underlying computations accessible to those who want to understand them.

You can scroll all over the Lng Model worksheet, studying the computations underlying the results in as much detail as you want.

In studying the computations, it is perhaps best to begin with the temporal coefficient matrices, which you will find in H117:K127. The principal matrix is H115:K118. By inverting this matrix (or its submatrices) and multiplying the inverted matrix times the column vector consisting of the uncorrected rate estimates, the sheet calculates the rate estimates (corrected rate vectors). If you click on cells of this matrix, you will see (in symbolic form, e.g., " $=T1.2/T1$ ") the ratio of the two temporal totals that constitute that particular coefficient in the matrix. These totals come from the top of the summing or accumulator columns (Columns G to P). These columns accumulate the subject's exposure to each of the four possible CSs and to each pairwise combination of those CSs. The uncorrected rate vector is in S115:S118. If you click on one of these cells, you will see in symbolic form, the two quantities that determine a particular uncorrected rate estimate. For example, clicking on S115 shows the formula " $=N1/T1$ ", which tells you that the uncorrected rate estimate for the background is the total number of reinforcers experienced in the presence of the background divided by the total exposure to the background. (Because the background is always present, the first of these is simply the total number of reinforcers and the second the total amount of conditioning time.)

The determinants accompanying the main matrix and the various submatrices indicate which matrices were actually used to calculate rate estimates. (Only matrices with non-zero determinants can be inverted). In protocols where there is more than one possible additive solution (e.g., the overshadowing and relative validity protocols), more than one of the submatrices is invertible, so more than one solution is generated. The optimal solution, the one that minimizes the number of predictors, is determined after calculating the sum of the absolute values of the rate estimates in each solution. These sums are in the cells next to labels like "ABS123" or "ABS134." Underneath these sums are "FLAG" cells. These contain conditionals that compare the sums of the different absolute values and raise a flag (set a value equal to 1) or don't raise a flag (set a value equal to 0), depending on the outcome of the comparison. The flags are surveyed in G132:L135 to determine the ultimate solution, which appears in N132:N135.

To the right of the solution vector are the cells that calculate the odds that CSs 2 through 4 affect the rate of reinforcement. If any one of these CSs does affect the rate of reinforcement, then the rate observed when that CS is present together with the background CS will be different than the rate observed when the background alone is present--unless the rate estimate in question is negative and the background rate is essentially zero, in which case the sum of these two

estimates may well be negative. To avoid negative rate ratios (hence, negative F values) from inhibitory CSs, this calculation takes the absolute value of a rate estimate before adding it to the background rate estimate. Thus, for example, for CS2, we have " $(ABS(Rate2)+Rate1)$ " as the estimated rate when that CS is present along with the background (CS1). This estimate is compared to the background rate alone by means of a ratio. This explains the formula in Cell P135, which is " $=IF(ABS(Rate2)<0.001,0.1,(ABS(Rate2)+Rate1)/Rate1)$ ". (The conditional is necessary to take care of bogus very small rate estimates that arise from rounding errors in the matrix inversion operation.) This formula gives the ratio of two rate estimates. Such ratios obey the F distribution. In Cells R132:R135, the degrees of freedom for the two rates in the rate ratio are calculated. The degrees of freedom for a given rate estimate are twice the number of events on which that rate estimate is based. The number of events on which a rate estimate is based is obtained by multiplying the rate estimate times the amount of time that that rate was in effect. The probability that there has not been a change is calculated in T133:T135, under the (somewhat confusing) heading of "1-p". Note that this calculation uses the FIDISTRIBUTION function, with the rate ratio as the first argument, and the two degrees of freedom as the second and third arguments. The complementary probability, namely, that there has been a change is calculated in S133:S135, under the heading "p".

### **Modifying the model to take longer protocols, or shortening it so that it doesn't take up so much room**

The temporal and numerical summing columns (Columns G through U down to Row 91) have been set up to work retroactively, so that the total time or number appears at the top rather than at the bottom of a column. In order to expand the sheet to deal with longer, more complex protocols, or to contract it so that it takes up less room, one can delete or add rows from or into the region between Rows 6 and 91. However, in deleting and adding rows into this region be sure always to keep the bottom row of this region (which is now Row 91) as is!!! This row must contain  $=#NA()$  in the event time column and it must have 0's instead of formulae in all the summation columns (Columns G through U). All other cells in a summation column (all the cells above Row 91) have calculational formulas. Deleting rows below a given row won't perturb the operation of the formulas in the rows above the deletion, provided always that at the very bottom of the column the last row is simply 0. To add rows, just insert them between the last formula row and the bottom (all 0's) row and then fill down from the formula row immediately above the newly inserted blank rows, taking care not to fill into the bottom row with the 0's. To delete rows, just select the rows you want to delete, taking care not to include Row 91 in the selection, and issue the delete rows command. Needless to say, before contemplating surgery on the sheet, it would be good to make a backup copy. In fact, it would be a good idea no matter what you contemplate.

**Starting with initial totals**

In some experiments—e.g., latent inhibition experiments or experiments in which there is a period during which the CS and US are randomly related (learned irrelevance phase) prior to the onset of their being paired—it is desirable to begin with some initial values of the cumulative temporal and numerical totals, representing the accumulations during the initial phase, and then calculate trial by trial what happens thereafter. This can be done by entering the initial totals into the shaded cells in row 4 of the model, then entering into the shaded cell for the '0' event time, the (nonzero) time beyond which you are going to enter data trial by trial. (Be sure to enter a 1 in the C1 column for this first entry, too.). Thus, for example, suppose you wanted to assume an initial latent inhibition phase of 10 trials on which CS2 comes on (for 2 minutes each trial) but there is no US. Then, on the 11th trial and afterwards, each occurrence of CS2 terminates with an occurrence of the US. Assume a 10 minute ITI, so that the period (C) from one onset of CS2 to the next is 12 minutes. In the 130 minutes prior to the onset of the CS that marks the start of the 11th trial, an initial total of 130 minutes will have accumulated for TC1, 20 minutes for TC2, and 20 minutes for T1.2. Enter these totals into the shaded cells in row 4 above the columns wherein these respective totals would have accumulated. Enter 130 into the shaded cell that normally gets event time 0; enter 1 into the C1 column in that row; and proceed from there. (In the control condition, the two 20 minute totals become 0.)

The capability to enter totals directly can also be used to bypass protocol composition altogether. If you are sure you know what each accumulation would be in an experiment for which you want a prediction, just enter these totals into the appropriate shaded cells in row 4 and tell the sheet to calculate.