Title
Summary of recordings of spike trains available for analysis and guide to their context and means for their analysis

Permalink
https://escholarship.org/uc/item/7sc6k27t

Author
Woody, Charles D.

Publication Date
2005-08-12

Supplemental Material
https://escholarship.org/uc/item/7sc6k27t#supplemental

SUMMARY OF RECORDINGS OF SPIKE TRAINS AVAILABLE FOR ANALYSIS

AND

GUIDE TO THEIR CONTEXT AND MEANS FOR THEIR ANALYSIS

C. D. WOODY

PROFESSOR OF PSYCHIATRY AND NEUROBIOLOGY

UCLA MENTAL RETARDATION RESEARCH CENTER

760 Westwood Plaza, Los Angeles, CA 90024

2005

WHAT IS THIS?

- This is a database of recordings of spike activity from about 5000 neurons along the auditory pathway from the cochlear nuclei to the neocortex, cerebellum, and elsewhere. Sufficient indexes and examples are available here to understand the data. The bulk of the data is on five CDs (0-4) that can be borrowed through interlibrary loan (UCLA Biomedical Library, WL102.5, W912r 2003) and copied. (See *04DataCDs.xls* for that index.) That data can also be downloaded from this website (see 12. Appendix 2), but is voluminous.
- There is also an analysis program that averages histograms of spike activity and concurrent electromyographic (EMG) activity obtained from the orbicularis oculi (eyeblink) and levator oris (nose twitch) muscles. This, too, is available here.
- Each recording contains the response to a 70 dB click* (produced by applying a 1 ms rectangular pulse to a loudspeaker) before or after conditioning blink responses to the click as a CS. A discriminative acoustic stimulus, a 70 dB hiss DS, was also presented at a time apart from the click.
- There are guides to the regions from which each recording was made and to the associated behavioral states as well as a tutorial on how to use the analysis program.
- There is an appendix listing the data format.
- To proceed please read the cited summary and guide, or continue if you're already there.

1. DISCLAIMER

   The contents (except for Appendix 2) have been checked for viruses, and much of the spike data listed in the Excel (xls) files has been run successfully on my PC, but I cannot guarantee that they or the remaining data contain nothing that would cause problems. Accordingly, I must disclaim any responsibilities for any problems that might arise from their use. If this is not understood and agreed, please do not continue.

2. GUIDE TO RECORDINGS

1. The data are divided by:
   a. Region of the brain (for some detailed reports and superaverages see "01PublSupav.xls")
   b. Cat from which data were obtained
   c. Behavioral State (Adaptation, Conditioning, Extinction, and backward conditioning)
   d. Spike activity by cell (and occasionally by cell type)
   e. Associated EMG: **EMG3 (orbicularis oculis)**. Also: EMG4 (levator oris), EMG1 (contralateral orbicularis oculis), EMG2 (contralateral levator oris). *USE CAUTION WHEN ANALYZING DATA FROM EMG 1, 2 and 4 because these recordings were not always maintained when recording spike activity*.
   f. Trials (each cell normally received 5-20 trials of click and hiss presentations)

2. Regions and folders containing the data are described in ***"04INDEX.XLS"***.
   The regions include the *MOTOR CORTEX, AUDITORY CORTEX, THALAMUS, MEDIAL GENICULATE, INFERIOR AND SUPERIOR COLLICULUS, CEREBELLUM, DENTATE, INTERPOSITUS, DORSAL AND VENTRAL COCHLEAR NUCLEUS, LATERAL LEMNISCUS NUCL., SUPERIOR OLIVARY NUCL., ANTERIOR CAUDATE NUCL., AND NUCL. RETICULARIS PONTIS CAUDALIS.* Further details are given in "04Index.xls".

3. Behavioral states are as follows:

   Adap – adaptation (presentations of the CS and DS before conditioning)

   Cond – test trials of CS and DS after conditioning with continued intermittent pairing with US to maintain CR (thus there are presentations like Adap followed by presentations with CS, US and DS – note these are in the first and second columns of the *xls* listings)

   Ext – test trials of CS and DS after conditioning without further pairing with US to allow extinction of CR

   BD – test trials of CS and DS after backward conditioning (reversed order of CS and US) with intermittent backward pairing to maintain the paradigm

*Amplified earphone recordings of the click and hiss are shown in Fig. 4 of Woody et al. *Brain Res*. 868:56-65, 2000, and their power spectrum densities are shown in Fig. 2 of Xi et al. *NeuroReport* 5:1567-1570, 1994.

4. Cells in cerebellum and DCN are classified as simple (no notation) or complex (c) based on the absence or presence of complex spikes.

   All cells are listed by an identifying number for the recorded unit, followed by R (to designate raw data).

5. Trials are listed by their range e.g., 1-10, etc.

3. PUBLICATIONS AND GOALS

01PUBLSUPAV.XLS contains a listing of publications drawn from this database and related material. Check the corresponding *ppt* file. There have been three main goals: 1) to establish the transmission pathway (by latencies of spike response onset) and understand how acoustic information is transformed relay by relay as it is transmitted to (and from) the cortex, 2) to establish how the transforms at each relay change after conditioning, and 3) to determine the correlation between the patterns of spike activity and the patterns of the conditioned stimulus and the conditioned movement, the latter measured by the EMG.

4. WHAT THE ANALYSIS PROGRAM DOES

1. Averages four groups of specified spike or EMG files.

2. Graphs the current averages as peri-stimulus time histograms (PSTHs). (Check the *ppt* files.)

3. Adjusts scaling to specification.

4. Prints or copies the graphs to slides.

5. Creates super averages from specified averages or data from many units

6. Graphs, scales, prints and copies super averages

7. Batch processing – as yet not ready for use

8. Provides detailed **_reports_** of all useful information contained within the current average's raw data files or general information about the superaverages

9. Provides error messages concerning the analyses.

Directories: ***Must be set first when using the analysis program.*** The input directory specifies the location from which data files are read. The output directory specifies where newly created averages and superaverages are to be stored.

5. ANALYSIS PROGRAM

ANALYZE.ZIP is a zip file of *ANALYZE.EXE* (must be unzipped in the WINDOWS folder before using). *ANALYZE.EXE* is the program for visualizing and analyzing the data. It is run from the WINDOWS folder. It allows one to make averaged, peri-stimulus time histograms of spike or EMG activity. Just how this works may be seen by using the tutorial. Before using it you may wish to print files: *04index.xls, 01publsupav.xls*, and *01indabbrev.xls* from the website (or CD#0).

6. TUTORIAL

(The following assumes your storage disk is labeled 'C'. If not, change the labels that follow accordingly.)

1. Create a folder "Anliz" on disk C for storing analyzed data. (C:/Anliz)
2. If on line copy the four files (3307.spk and emg, 3310.spk and emg) into a folder named "InfoShort" on disk C (C:\InfoShort), and change the word INFCOLL to InfoShort in all subsequent instructions. You can do much but not everything with the shortened set of data that you can do with INFCOLL. If you have the set of CDs, copy the data file folder INFCOLL from CD#4 to disk C (C:/INFCOLL)
3. Run analyze.exe, which may be found on line or on CD#0. *(To do this it may be necessary first to copy it into your windows folder.) (Note: Mspace is used for a different analysis. See Victor & Purpura in Section 11. That program may be sought through Mike Page, Great Island Software, 9 Ethelma Dr., Charham, MA 02633, E-mai: pagem@GreatIslandSoftware.com   Web:www.GreatIsland Software.com.   It also requires an associated mspace_original.dll file.)*

4. Wait. Then left click (L/c) *File*
5. L/c *Set Directories*
6. Set <u>input directory</u> to the data file copied from the CD, e.g.: C:/ INFCOLL (or C:/InfoShort if on line)
7. Set <u>output directory</u> to C:/Anliz   (A different directory, Anlizav, contains my indexed superaverages)
8. L/c *OK*. If not accepted, the input or output files were not validly specified
9. L/c *current averages* (<u>pinhole</u>) – not necessary if it already has a black dot
10. L/c far right arrow for *File Type*; L/c raw
11. L/c far right arrow for *Version*

     YOU SHOULD NOW BE SEEING THE NUMBERED CELLS FROM WHICH SPIKES AND CONCURRENT
     EMGs ARE AVAILABLE IN THIS DATA FILE FOLDER

If so, you may wish to proceed to A 5.

A. <u>*To make an average of spike data*</u>
   1. Run <u>analyze.exe</u> with (L/c) *current averages* (<u>pinhole</u>) filled with a dark dot
   2. Your <u>input directory</u> should still be set to C:/INFCOLL
   3. Your <u>output directory</u> should still be set to C:/Anliz or whatever file you have created and defined for
      outputs.
   4. Set 'File type' raw by L/c far right arrow and L/c raw
   <u>5</u>. Set 'Version #' 3307 by L/c far right arrow and finding and L/c 3307
   6. Set 'Trials' 1-13 by L/c far right blank box and typing 1-13
   7. Set 'Signal' spike by L/c far right arrow and L/c spike
      (Note that in all of the above instructions you have set all four (1-4) boxes to the left by adjusting the
      value at the far right. It is possible to adjust each of those four boxes individually, separately.)
   8. L/c 'Create'
   9. L/c OK; L/c 'FULL'
      You should now be seeing four histograms of the spike response to the click CS of cell 3307, trials 1-13.
      (These are peri-stimulus time histograms (PSTHs) because there is baseline activity preceding delivery of
      the stimulus at time 0 in addition to the response that follows. This cell responded rapidly in the first bin
      of 0-32 ms after the click.)
   10. <u>*Print histograms*</u> – L/c File; L/c 'Print' (parameters can be changed in GRAPH, Page Setup)
   11. <u>*Copy to power point*</u> – L/c Edit; L/c copy; L/c bitmap (make check disappear); L/c OK … Then, paste
       as a slide in your opened ppt file.
   12. <u>*Copy*</u> any one histogram graph to <u>text file</u> listing the <u>*numeric values*</u> that were graphed (and thence to
       an xls file) by L/c diagonal arrow in box at left, right clicking the one graph to be transferred, L/c graph,
       L/c save graph data … , define location to be stored, and name the text file. (To make an xls listing,
       open the text file, select all, copy, close, then open the appropriate xls file and paste in the desired
       location. This is *very handy for doing comparative statistics* in which different graphs are tested for
       correlation.)
   13. L/c control panel
   14. Change histograms 1&2 'Time Scale' to 80 ms by L/c appropriate arrows and L/c 80
   15. L/c Create
   16. L/c OK
   17. L/c FULL
       You can now see that the initial rise in activity >3Z occurs 12-16 ms after the click, and the peak
       response occurs 16-20 ms after the click. *(If one adds up the data in histogram 1, it will match the data at the
       corresponding times in histogram 3.)*
   18. L/c control panel
   19. Check spontaneous firing rate and standard deviation values
   20. L/c current X – dotted horizontal line is the mean spontaneous firing rate; solid horizontal line is 3Z
       (3 std. dev.) above mean spontaneous firing rate
   21. L/c control panel, scroll down towards bottom, and note where Z can be changed to none. 1, 2 or 3
   22. Scroll to top and note where histograms are labeled 1,2,3,4 and All
   23. L/c right arrow for Stimulus, histogram 2, and L/c DS
   24. Do the same for histogram 4
   25. L/c Apply
   26. L/c current X
       You should now have changed both histograms to the right to the spike activity in response to the hiss
       DS; i.e. a different response than to the click CS.
       NOTE further that this was done by modifying the already performed calculations on the specified cell

and trials (cell 3307, trials 1-13). This saves you the time it would take to recompute using the *CREATE* approach taken earlier. You can also rescale parameters in the graphed data using *APPLY*.

27. Change the time scales of histograms 1 and 2 to 1280; *apply*; and compare the result with histograms 3 and 4. (Time scales may be set to 80, 160, 320, 640 or 1280 ms.) Note the amplitude scales on the graphs are different. Set those of Histograms 1&2 ('Graph Max' and 'Graph Min') to the values specified in the control panel for histograms 3 & 4, and L/c 'apply'. In this way you can adjust the scaling to whatever you want.

28. Finally, *print a report* (which lists information about the analysis that was performed) by L/c report, L/c current data files, L/c File, L/c print

B. To make an average of EMG data (EMG3 is the preferred set of emg data, cf. 6.INDEX)

This is done in the same way as in part A, except that one sets the 'Signal' to EMG3 instead of spike. (You may want to try this for version 3307, trials 1-13, by making the appropriate changes in the control panel settings and L/c Create. *Caution must be exercised in analyzing EMGs because the data have not yet been checked for absence of flaws in transcription that will cause failures.* These flaws may cause generation of an error message indicating corrupt data or inability to find a particular data file. One must use this information with trial and error approaches to find and eliminate the flawed data.

Error messages stating that a conditioning trial (or otherwise) has been erroneously included will generally be of small numeric consequence unless a conditioning trial with delivery of the US (accompanied by a large artifact occurring 570 ms after the CS) is included. This will usually be evidenced by an abrupt discontinuity in the average occurring at that time. Again, trial and error inspection of smaller portions of the data can detect and eliminate such errors.

C. To make an average of averages - SUPERAVERAGES

A *superaverage* is the term used to connote any average of averaged trials from more than one single unit. One can average (*super*) data from averages themselves, or one can sequentially average (*raw*) trials from many cells. Since the formats are not the same, *the analysis program must be told whether the trials are raw data or super data*. In addition *the analysis program must be switched from current average to superaverage mode.*

   Note that the data that are to be superaveraged may be specified by entering the information manually in an average specification window in the form indicated (after the specification window has been opened) by L/c help, or the data may be specified by entering prespecified information from the xls data information files. We will try each way as the tutorial continues:

1. Run analyze.exe with (L/c) *Super averages* (pinhole) filled with a dark dot
2. Your input directory should still be set to C:/INFCOLL (L/c File; L/c Set Directories)
3. Your output directory should still be set to C:/Anliz or whatever file you have created for outputs.
4. Set *Signal* to spike (by L/c far right arrow; L/c spike)
5. L/c *create*
   You should now see the '*averaging specification window*'.

We will now try two different ways of specifying the data to be superaveraged:
a. Manual entry
   1. Type the following in the average specification window:
3307R1-13	If using InfoShort type:	3307R1-13
3310R1-22		3310R1-22
3311R1-21	and disregard anything to do with the files you don't have (i.e., 3311,
3312R1-20	3312, and 3313)
3313R1-19	

   There must be no blank lines in between.
   2. L/c OK; L/c OK; L/c FULL
      You should now see four histograms showing the average PSTH of trials 1-13 from cell 3307 + trials 1-22 from cell 3310 + trials 1-21 from cell 3311 + trials 1-20 from cell 3312 + trials 1-19 from cell 3313. Note that the histograms show an identifying number for the new superaverage you have created. (You should be able to find this number in your output directory, and recover the superaverages in the manner described in section D.)
      Note that the total number of cells (C:5) from which the superaverage is made is given at the top of each histogram as is the total number of trials (T:93).

3. Note that you can change parameters using '*APPLY*' as when averaging spike data from a single cell. (See instructions A10-28.)

b. *XLS file entry*
   1. Open the xls file '01InfCoTEST.xls' (only available on set of CDs)
   2. Select all sixteen listings in the first column; COPY; CLOSE; EXIT. (In the previously opened analyze program, L/c superaverage pinhole and L/c Create.)
   3. Position the cursor in the upper left of the average specification window
   4. Press Ctl V

   You should have pasted information about cells 3307-3313, 3370-3377, and 3400-3402 into the average specification window

There must be no blank lines in between cells.

   5. L/c OK; OK; FULL
      You should now see four histograms showing the average PSTH of trials 1-13 from cell 3307 + trials 1-22 from cell 3310 + … up to and including   + trials 1-17 from cell 3402. (The top of the graph should show 16 cells (C:16) and 287 trials (T:287) as having been analyzed.)

If you have all spikes (solid bars in the histograms), try going back to the control panel and specifying EMG3 as the signal for histograms 3&4; then L/c Create; then L/c Restore Spikes; then L/c OK,OK, FULL. You should now see the spike activity above and the concurrently recorded EMG (hollow bars) activity of the orbicularis muscles below. Changes 'Stimulus' to DS for histograms 2&4; then L/c 'Apply'.

D. *Displaying previously made averages*

   You may display previously made superaverages (e.g., 04Index.xls, 01publsupav.xls; c.f., 3556.avg, 01publsupavgraphs.ppt, 04Indexsupavs.ppt) by resetting the input directory as appropriate (e.g., to your output directory and then making the version # on the control panel agree with the V specified at the top of the last set of histograms). Follow the procedures for making an average of spike data but:
   (4.) Set file type Super by L/c far right arrow and L/c super
   (5.) Set version number to the identification number of the desired average
   (6.) Omit - (you do not need to specify trials since this is an average weighted by cells)
   (7.) Set histogram# 1,2,3 4 and stimulus CS,DS,CS,DS or whatever is appropriate
   (8.) L/c Create
   Don't forget to reset the Input Directory if you return to INFCOLL.

7. SUGGESTIONS FOR ANALYZING DATA

   Many types of analyses can be done, and most can be performed more easily if the *xls* guide files are used. To understand how the guide files are constructed, print a copy of 01Infcoll.xls.

Infcoll.xls: This is a list of all the cells from the inferior colliculus from which spike activity was recorded. It is divided by cat and behavioral state. (The old, intermediate storage disks are listed but may be ignored.) As seen in *04Index.xls* and *04DataCDs.xls*, the data are available in a file named InfcolL, found on CD4. By copying that file to your hard disk, and specifying it as the 'Input directory' (see Tutorial), you should be able to see the cell numbers containing the raw data (i.e.,3307, 3310, 3311, …). By using the highlighting transfer approach described in the tutorial (*XLS file entry*, section C, subsection b), you can make averages of these data.

If you average spiking in cells 66-111; 135-720; 1122-1127; 3307-4220; and 1131-1150, you should obtain a superaverage of 294 cells and 3970 trials similar to that in superaverage file 5932 (located in anlizav). If you average their associated EMG activity, you must omit cells 1131-1150 because they did not have associated EMG recordings. The superaverage of the EMG for cells 66-111, 135-720, 1122-1127, and 3307-4220 has 280 "cells" and 3846 trials, and is shown in superaverage file 5934 (located in anlizav).

There are also more advanced means of manipulating the grouping of the data within the Excel files by using Excel pivot table maneuvers.

By these means one can compare spiking in naïve versus conditioned behavioral states, first five versus last five post-conditioning trials (to assess extinction), spiking from animal to animal, region by region (*04index.xls*), and in some regions cell type by cell type.

Some of the xls guide files contain additional columns listing conditioning trials on which the US was delivered. And a few files have separate columns listing the EMG trials. (The latter are rather superfluous since the trial numbers for the spikes are usually the same as the trial numbers for the concurrent EMG data. The difference is that the latter were frequently grouped as one average of 3 trials if no response was detected [to save PDP-11 computer storage space].Thus you can only analyze those trials as a group (i.e. always in combination), but their mean is formatted as though it were raw data so they are treated as 'raw trials'.)

8. LIMITATIONS

There are several limitations that should be noted. First, while the present Adap/Cond data have been checked for compatibility for incorporation into the histograms, there may be unrecognized incompatibilities. (There is an appendix below listing the intended data format.) Second, the studies, though extensive, are obviously incomplete. Third, the studies are flawed (hopefully just a little) by human inconsistency. (For example, the aim in training was to generate short latency, discriminative responses to the click, but this was not always accomplished. Only in part can this be attributed to natural inconsistencies among the animals.) Fourth, analogous human inconsistency is responsible for occasionally setting the spike recognition threshold too low or too high and for occasionally delivering a brief (tens of ms) current pulse through the recording electrode while spike detection was in progress. Both appear to have been done randomly and infrequently.

9. ACKNOWLEDGEMENT

10. APPENDIX : FORMAT OF DATA AND DETAILS OF CURRENT ANALYSES

The analysis programs are used:

1. To construct and store average PST histograms of 640 2ms bins (1280ms) plus 40 extra 2ms buffer bins (80 ms)
    a. of fixed-format raw data of 680 2 ms data points per trial for up to 99 repeated trials per cell
        i) of spike activity (requires one extra reformatting step)
        ii) of concurrent EMG activity
    b. of the fixed-format averages resulting from 1a ("superaverages")

2. To display on a monitor screen and then (optionally) print the histogram averages of part 1a or b in any of the following five choices of fixed, 40 bin formats:
        data lengths 80 160 320 640 1280 ms
        bin widths   2   4   8   16   32 ms
    a. with x & y axis scales displayed and numbered
    b. with a broken horizontal line set to the pre-stimulus baseline mean (average of 200 2ms data points- the 40 extra above plus 160 of the 640 that are displayed in the 1280ms histogram). (The 40 precede the160, and, together, reflect the 400 ms preceding stimulus delivery.)
    c. with a solid horizontal line set 3 std. dev.(Z) above the mean of the 200 data points in b
    d. optionally, with the horizontal line set 1 or 2 std. dev. above the mean
    e. with computed and displayed maximum and minimum values of each histogram
    f. with a "menu" to allow choices for histogram analysis and printing

i) The display shows 4 histograms at one time:

```
-----  -----
-----  -----
```

ii) The default scaling of magnitude fills the histogram space. Optional scaling allows the investigator to set the maximum and minimum ranges of magnitude displayed.

iii) The display prints the number of trials averaged and additional, available numeric data (for averages of raw data from one cell) or the number of cells averaged (for superaverages), and indicates whether CSs or DSs are the stimuli for which the data are shown (each data set is so labeled).

Definitions:

1. Raw data file – spike or EMG data files
2. Super average – an average of a group of averaged files
3. Current average – an average of selected raw trials or of a super average
4. Control panel – allows the user to set parameters for averaging, graphed display, and copying of i) the current average, ii) super averages, and iii) batch printing.
5. Graphs – 4 histograms on a page. The graphs data can be from i) the current average, ii) super averages, and iii) batch printing.
6. Signal – source for the data: EMG1, EMG2, EMG3, EMG4, or spikes.
7. Stimulus - Conditioned stimulus (CS) or Discriminative stimulus (DS). Trial data, average data, and super average data are aligned on the stimuli.
8. Version number - uniquely (in theory) identifies a data file within a stipulated directory

FILE FORMAT

The data are stored in binary coded format (16 bit integers = 2 bytes per data point, no floating point) in _three types of files_:

```
00000.emg, which contains raw EMG data,
00000.spk, which contains raw spike data, and
00000.avg, which contains superaverage data.
```

Sets of records are distinguished by version numbers within the files.

In the EMG and SUPERAVERAGE files each set of records is 2560 words (10 blocks) long, while in the SPIKE files each set of records is 1024 (4 blocks) long.

SUPERAVERAGE files:

1. SUPERAVERAGE records contain averages of histograms of spike trains (or EMG data) accompanied by a statement of the number of cells (or trials) averaged.

2. Data for spike train averages: (average spikes per 2 ms) * 100 *(number of cells)
   Data for EMG averages: (average across trials)

3. The first set of records (blocks 1-10) always contains descriptions (identifying parameters) of what is to come in each subsequent group of four, ten-block sets of records contained in the file. There are "groups of four sets" because there are four (#1,2,3,4) different histogram averages in each file.

4. Thus,
   the second set (blocks 11-20) contains data from HISTOGRAM AVERAGE #1
   the third set (blocks 21-30) contains data from HISTOGRAM AVERAGE #2;
   the fourth set (blocks 31-40) contains data from HISTOGRAM AVERAGE #3;
   the fifth set (blocks 41-50) contains data from HISTOGRAM AVERAGE#4.

5. Thus,

File	Histogram	Start Block	End Block	Start Word	End Word
1	ID	1	11	1	2560
1	1	11	20	2561	5120
1	2	21	30	5121	7680
1	3	31	40	7681	10240
1	4	41	50	10241	12800
2	ID	51	60	12801 etc	
2	1	61	70		

2	2	71	80
2	3	81	90
2	4	91	100

6. *Identifying Parameters* (contained in blocks1-10)

   [2 byte integers; 10 blocks has 5120 bytes or space for up to 2560 integers]

    word 101 - 99

    word 102 - 99

    word 103 - 99

    word 104 - 99 - four 99s designate SUPERAVERAGE file

    word 105 - unused

    word 106 - disk # data were stored on

    word 107 - file version # (versions in octal)

    word 108 - 5 (for spike data)

    word 109 - First trial averaged in this record

    word 110 - Last trial averaged in this record

    words 111-115, etc. - Every successive five words is for next record showing information 106-110 as above

    Four successive "98"s designate end of identifying information

    Word 1005 - 1=ISIs stored 0=old data before ISIs were stored

Each 2560 words contains 2500 data samples plus 60 words at end containing <u>additional descriptive parameters</u>.

7. *Additional Descriptive Parameters* (contained in 60 words beginning with the 2501st word)

    2501 - # of cells averaged

      2 - paradigm # (1-adap/ext, 2-cond, 3-delayed HS)

      3 - unused

      4 - unused

    5-11 - interstimulus intervals in ms (for coding in 15-20)

  12-13 14 - unused

    15-20 - coding of stimuli for interstimulus intervals (1=CS, 2=US, 4=HS, 8=DS, 16=bridge,

        64=prepulse, 255=shift to second stimulus)

Thus,

 5 - 450 and 15 - 64 means a prepulse was followed 450 ms later by

 6 - 2530 and 16 -   1 a CS, followed 2530 ms by

 7 - 1470 and 17 - 16 a bridge followed 1470 ms by

 8 - 400 and 18 - 255 a shift to data preceding by 400 ms delivery of the second stimulus

 9 - 2530 and 19 - 8 a DS, followed 2530ms by

10 - 400 and 20 - 16 a bridge, followed 400 ms by

11 - 2217 and 2 - 0 nothing since the record of data was concluded.

[-------3S = 1500 data samples-------]   [---2S = 1000

data samples--]

  450   2530   1470   400   2530

----------I----------------------------- b-I ---------------I-------I------------------------------------bI

     CS   bridge   DS   bridge
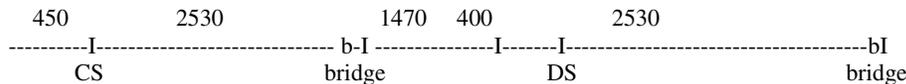
One data sample per 2 ms; 1500 data samples (3S) around first stimulus,1000 around second stimulus.

SPIKE files:

  1. SPIKE records contain data from a number of single trials that are easily distinguished by accompanying, identifying "first" and "last" trial numbers.

  2. Spike data are stored as the times of occurrence (in millisecs) from the beginning of the period of data collection surrounding each of two stimuli.

    For the first stimulus (usually the CS), collection begins at the start of the trial, 450 ms before CS delivery.

      (Thus, 1=450 ms before CS delivery, 451=1 ms after CS delivery, etc.)

    For the second stimulus (usually the DS) collection begins 400 ms before DS delivery. See words 1001, 1002, below. (Thus, 1=400 ms before DS delivery, 451=51 ms after DS delivery, etc.)

3. The first set of records (blocks 1-4) always contains descriptions (identifying parameters) of what is to come in the subsequent sets of 4 blocks, each 4 block set containing one trial of spike data.

4. *Identifying Parameters* (contained in blocks1-4)
      [2 byte integers; 4 blocks has 2048 bytes or space for up to 1024 integers]
      words 1-4 - Cell Identification Number, repeated four times
      word 101 - disk # data was stored on
      word 102 - file version # (versions in octal)
      word 103 - # "5" (indicates that this is spike data)
      word 104 - First trial stored in this record
      word 105 - Last trial stored in this record (or "0" if only a single trial is stored)
      words 106-110, etc. - Every successive five words is for next trial group, showing information as above

5. Typically,
    the second set of records (blocks 5-8) contains trial #1 (unless otherwise specified in the identifying parameters);
    the third set (blocks 9-12) contains trial #2;
    the fourth set (blocks 13-16) contains trial #3;
        etc.

6. Thus, the beginning of data for trial n is:
    (n * 2048) + 1 bytes
    (n * 1024) + 1 words

7. The fourth block of each set of records for each trial contains additonal descriptive parameters:
    word 1001 - position within spike data record of beginning of second data segment
    word 1002 - position within spike data record of end of second data segment
    word 1003 - blank or ? redundant 1
    word 1004 - paradigm # (1=adap/ext, 2=cond, 3=delayed HS)
    word 1005-1014 - interstimulus intervals in ms
    word 1015-1024 - designates which stimuli occur (1=CS, 2=US, 4=HS, 8=DS, 16=bridge, 64=prepulse)

EMG files:

1. EMG records contain data from a number of single trials and/or averages of 3 trials that are easily distinguished by accompanying, identifying "first" and "last" trial numbers.

2. The first set of records (blocks 1-10) always contains descriptions (identifying parameters) of what is to come in each subsequent group of four sets of records contained in the file. There are "groups of four sets" because there are four (#1,2,3,4) different EMG leads.

3. Thus, typically,
    the second set (blocks 11-20) contains trials from EMG #1 (unless otherwise specified in identifying parameters);
    the third set (blocks 21-30) contains trials from EMG #2;
    the fourth set (blocks 31-40) contains trials from EMG #3;
    the fifth set (blocks 41-50) contains trials from EMG#4.

4. And, thus, for an EMG file containing 10 trials of which trials 1-3 are an average, 4 a single trial, 5-7 an average, and 8-10 an average:
    Average of trials 1-3 - blocks 11-20 (EMG1), 21-30 (EMG2), ... , 41-50 (EMG4)
    Trial 4 - blocks 51-60 (EMG1), 61-70 (EMG2), ... , 81-90 (EMG4)
    Average of trials 5-7 - blocks 91-100 (EMG1), 101-110 (EMG2), ... , 121-130 (EMG4)
    Average of trials 8-10 - blocks 131-140 (EMG1), 141-150 (EMG2), ... ,161-170 (EMG4)

5. Thus, data for a trial group + EMG starts at:
    (Trial Group # - 1) * 10240
      + (EMG# - 1) * 2560
         + 2561

Trial Group	EMG	Start Block	End Block	Start Word	End Word
1	1	11	20	2561	5120
1	2	21	30	5121	7680
1	3	31	40	7681	10240
1	3	41	50	10241	12800
2	1	51	60	12801 etc	
2	2	61	70		

6. *Identifying Parameters* (contained in blocks1-10)
>   [2 byte integers; 10 blocks has 5120 bytes or space for up to 2560 integers]
>   word 101 - disk # data waas stored on
>   word 102 - file version # (versions in octal)
>   word 103 - EMG # (1,2,3, or 4)
>   word 104 - First trial stored in this record
>   word 105 - Last trial stored in this record (or "0" if same as first trial)
>   words 106-110, etc. - Every successive five words is for next record showing information as above

Each 2560 words contains 2500 data samples plus 60 words at end containing *additional descriptive parameters*.

7. *Additional Descriptive Parameters* (contained in 60 words beginning the 2501st word)

>   1 - paradigm # (1-adap/ext, 2-cond, 3-delayed HS)
>   2 - unused in EMG
>   3 - unused in EMG
>   4-11 - interstimulus intervals in ms (for coding in 14-19)
>   12-13 - unused
>   14-19 - coding of stimuli for interstimulus intervals (1=CS, 2=US, 4=HS, 8=DS, 16=bridge,
>   64=prepulse, 255=shift to second stimulus)

Thus,
> 4 - 450   and 14 - 64   means a prepulse was followed 450 ms later by
> 5 - 2530 and 15 - 1   a CS, followed 2530 ms by
> 6 - 1470 and 16 - 16   a bridge followed 1470 ms by
> 7 - 400   and 17 - 255 a shift to data preceding by 400 ms delivery of the second stimulus
> 8 - 2530 and 18 - 8   a DS, followed 2530ms by
> 9 - 400   and 19 - 16   a bridge, followed 400 ms by
> 10 -   0 and 20 - 0   nothing since the record of data was concluded.


```
[-------3S = 1500 data samples-------]        [---2S = 1000 data samples--]

  450           2530              1470    400          2530
----------I----------------------------- b-I---------------I-------I--------------------------------bI
     CS                         bridge            DS                        bridge
```
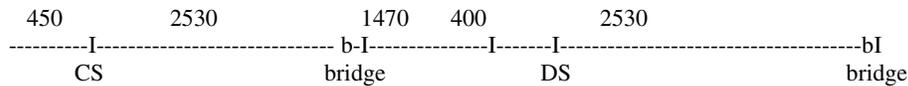
One EMG data sample per 2 ms; 1500 data samples (3S) around first stimulus, 1000 around second stimulus.

11. REFERENCES TO MORE ADVANCED ANALYSES

Victor, JD, Purpura KP (1997) Metric-space analysis of spike trains: theory, algorithms and application. Network: Comput Neural Syst 8:127-164.

Van Steveninck de Ruyter, RR, Lewen GD, Strong SP, Koberle R, Bialek W (1997) Reproducibility and variability in neural spike trains. Science, 275:1805-1808.

Bialek, W, de Ruyter van Steveninck R, Warland D (1991) Reading a neural code. Science 252:1854-1857.

Anderson, TW (1984) An Introduction to Multivariate Statistical Analysis. New York: Wiley.

Fukunaga K (1990) Statistical Pattern Recognition. 2nd ed. San Diego: Academic Press.

WEBSITE or CD 0

Preferred data indexes	**Raw data indexes**	**MAJOR INDEXES or GUIDES**
01AudctxCond.xls	01audctx.xls	01IndAbbrev.xls
01AudctxNaive.xls	01cbelFLOC.xls	04Index.xls
01caudate.xls	01cbelL.xls	01PublSupav.xls
01cbelFLOCADAP.xls	01cbelM.xls	05INDEX.doc (Website only)
01cbelFLOCCOND.xls	01cbelSim.xls	04DataCDs.xls
01cbelLADAP.xls	01DCN.xls	01Index.xls (old index)
01cbelLCOND.xls	01DCN2.xls	
01cbelMADAP.xls	01dentate.xls	**PROGRAMS**
01cbelMCOND.xls	01geniculate.xls	Analyze.exe
01cbelSimADAP.xls	01InfColl.xls	Analyze.zip
01cbelSimCOND.xls	01Intpos.xls	
01DCNCond.xls	01LatLem.xls	**POWER POINT FILES**
01DCNNaive.xls	01motctx.xls	01IndexSupavs.ppt
01DentateCond.xls	01retponc.xls	04IndexSupavs.ppt
01DentateNaive.xls	01SupColl.xls	01publSupavGraphs.ppt
01geniculatecond.xls	01SupOliv.xls	**SPECIAL DATA**
01geniculatenaive.xls	01ThalamusA.xls	Anlizav (contains
01InfColCond.xls	01ThalamusM.xls	*superaverages*)
01InfColNaive.xls	01ThalamusP.xls	PublSupav (contains
01InfCoTEST.xls	01VCN2.xls	*superaverages*)
01LatLemallCOND.xls		
01motctxcond.xls		
01motctxnaive.xls		**RAW DATA**
01retponcCOND.xls		CbelSimL
01retponcADAP.xls		Infcotest
01SupColl.xls		Infoshort
01SupOlivallBD.xls		
01ThalamusACond.xls		
01ThalamusANaive.xls		
01ThalamusM.xls		
01ThalamusP.xls		
01VCNcond.xls		
01VCNNaive.xls		
02AudctxCond.xls		
02DentCond.xls		

Generally, the preferred data files have had data that disrupt the analytic program excluded -- especially those marked Adap (adaptation) or Cond (conditioned).

12. Appendix 2

The database of recordings of spike activity from CDs 0-4 is also available in the initial, 'related files' listings as downloadable, self-executable data files: e.g., CbelFlocL.exe, CbelLatL.exe, etc. **BEWARE. Because they are so voluminous, some can take many hours to download.** Thus the CDs are available by interlibrary loan (as noted earlier), and this is the preferred means for accessing them. (The file *04DataCDs.xls* lists each database file and its size in megabytes.)