

Instructor: Paul Wolfgang
Faculty sponsor: Yuan Shi, Ph.D.

Research work was completed in collaboration with Michael Tobia, Kevin L. Brown, Ph.D., Diana S. Woodruff-Pak, Ph.D.

EBCC Data Analysis Tool (EBCC DAT)

Introduction

This project was conceived after completing two semesters of independent study research on eye blink classical conditioning in Dr. Woodruff-Pak's lab at Temple University as part of the neuroscience specialization program. The work involved analyzing a series of datasets collected from studies using eye blink conditioning methods. Several ideas on how the process of data analysis and interpretation could be improved were generated during that time, and a conclusion was reached that a single user-friendly application incorporating data visualization and data processing functionalities was needed. EBCC DAT was developed to meet this need and develop a flexible framework for future applications in data analysis. This application is focused on visualizing data in three dimensions, computing averages, as well as computing measures and thus processing and summarizing data. While the focus of the project was on eye blink classical conditioning research, and therefore it was specifically tested only on a limited range of data formats, the design of the application is flexible enough that it can be extended to other similar kinds of data. To fully understand the purpose and methods of this project, an overview of eye blink classical conditioning research techniques and data is necessary.

Overview of Eyeblink Classical Conditioning

Eyeblink classical conditioning utilizes the natural tendency of organisms to associate neutral stimuli with reflex-eliciting stimuli. A neutral stimulus is something that does not normally produce any specific response, aside from a general reaction such as a startle response that organisms experience when a loud noise is heard. A startle reaction is not a specific response because any number of sudden events can trigger it, such as a flash of light, someone's voice, etc., and no specific reflex (such as an eyeblink) is triggered by the stimulus. A neutral stimulus is called conditioned stimulus (CS). 1kHz tone at 80 dB is a typical CS. CS does not elicit an eyeblink reflex in unconditioned subjects, but through eyeblink conditioning can become associated with another stimulus that does. A stimulus that elicits a reflex is called an unconditioned stimulus (US), and when it is administered at the same time as CS, it becomes associated with CS in such a way that CS alone will then begin eliciting the reflex.

Eyeblink conditioning is very useful as an assessment tool for measuring and comparing cognitive performance between groups of subjects, because the underlying brain structures and function have been extensively researched and well understood, and

the methodology has been standardized. It can be used in studies ranging from evaluating the effectiveness of medications for treating Alzheimer's disease, to pinpointing the location of certain kinds of cognitive functions within the brain.

Typical eyeblink conditioning research involves mouse, rabbit, or human subjects. In mice US is a mild shock to the muscles around the eye administered by surgically implanted electrodes while in rabbits and in humans US is a corneal air puff of around 5 psi. US elicits an eyeblink reflex in a subject that over time becomes associated with the CS. This association between US and CS over time causes CS alone to trigger an eyeblink. The degree to which CS triggers a response is an indicator of the degree to which the underlying neural substrate supports the formation of the association. To assess the level of the CS, trials without a US (called CS-only trials) are used. Trials containing both a CS and a US are called paired trials.

Eye blink conditioning is primarily focused on assessing the function of the hippocampus and the cerebellum, as these are the two main regions of the brain involved in eyeblink conditioning. By varying the methods used in a study different aspects of these connected brain systems can be measured. Importantly, there are two main ways to combine US and CS: delay conditioning and trace conditioning.

In delay eyeblink conditioning CS and US overlap. CS onset precedes US onset and both CS and US terminate at the same time. For the CS to become associated with the US, cerebellum alone is sufficient, so this type of conditioning is used when the cerebellum function is being assessed without hippocampal involvement.

In trace eyeblink conditioning CS terminates before US onset. The period between CS termination and US onset is the trace interval. Hippocampal function is necessary in addition to cerebellar function for trace conditioning to succeed, so this type of conditioning is used when hippocampus is being targeted for assessment.

Figure 1 illustrates the differences between the delay and trace eyeblink conditioning.

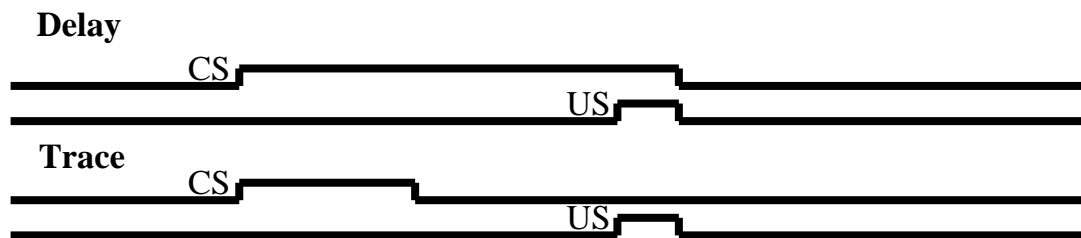


Figure 1. Delay vs. trace eyeblink conditioning

Description of Eyeblink Conditioning Data

These various ways of assessing subjects using eyeblink conditioning usually conform to a standard protocol, in which a series of brief trials are administered over a period of less than one hour. This series of trials makes up a session, and the collected data is grouped together into a dataset. Sessions are repeated over a number of days, and upwards of 20 datasets, each consisting of up to 100 trials, are collected. After the data is collected it is analyzed for dependent measures, which generally include the percentage of conditioned responses (%CR) produced by the subject, where a conditioned response

(CR) is defined in a way that is specific to each study but is generally intended to reflect a recognizable response from the subject to the CS.

Data collected during eyeblink conditioning studies is obtained from instruments that detect a response from the subject. These instruments output a time series data, i.e. a sequence of samples taken at discrete evenly spaced intervals reflecting the state of the subject's eyelid. In mice, where electrodes are implanted in the muscles around the eyelid, electrical potential is measured at the electrodes and converted into a sequence of amplitude measurements reflecting the degree to which the muscle has contracted. In humans and rabbits, an infrared sensor positioned in front of the eye detects the amount of light that is reflected from the eye and produces amplitude measurements that show this variation and thus indicate the state of the eyelid.

While specific parameters such as the length of trials, the number of trials per session, the number of sessions, and the timing of onset of CS and US vary from study to study, the methodology underlying these studies dictates that these parameters stay close to a specific range. In a typical eyeblink conditioning study a single trial ranges between 1 and 3 seconds, and the sampling rate can be anywhere between 300 and 3000Hz with 1000Hz being commonly used.

Each session consists of multiple trials, generally on the order of 100, and these are grouped into a dataset that is then stored as a single file on the computer system. A study can consist of about 20 sessions, for a total of 20 datasets and 2000 trials. Delay and trace conditioning studies differ in the onset and termination times of the CS and the US, and it is best to illustrate these differences with a comparison of parameters from two real studies used when testing EBCC DAT:

In the delay eyeblink conditioning mouse study,

- A single trial lasts 1350 ms
- 100 trials are administered in one session
- 20 sessions are administered during the study
- CS onset at 250 ms
- US onset at 750 ms
- US and CS co-terminate at 850 ms

In the trace eyeblink conditioning mouse study,

- A single trial lasts 1350 ms
- 100 trials are administered in one session
- 14 sessions are administered during the study
- CS onset at 250 ms
- CS termination at 500 ms
- US onset at 750 ms
- US termination at 850 ms

Since US involves actively administering the stimulus to the eye which is picked up by the instruments and obscures the subject's response, some of the trials are given without the US. Therefore each session consists of paired and CS-only trials. During paired trials both CS and US are administered, whereas during CS-only trials US is not given, which allows CR to be measured without interference from the US. In the above studies, 90 trials are paired and 10 are CS-only.

Goals

The main goal of this project is to facilitate the interpretation and analysis of data that is collected during eyeblink conditioning studies. Multiple dependent measures have to be extracted from the data and this process is greatly enhanced by features such as 2D and 3D data visualization, data set averaging, and other data processing capabilities combined in a single user friendly application.

EBCC DAT was designed with a number of goals in mind. Providing data visualization capabilities was the main goal of this project. Data visualization allows individual trials to be viewed as graphs, or combined on a single screen and viewed as a sequence of waveforms constituting a dataset. This enables the researcher to visually identify certain characteristics present in each dataset that are helpful when considering the methods to be used when calculating dependent measures.

Three levels of visualization were necessary:

1. Trials had to be rendered on a two-dimensional graph and ability to scroll through the trials had to be implemented so that each individual trial could be examined in detail.
2. Trials had to be rendered on a three-dimensional graph as a sequence of waveforms, each waveform representing a single trial. User then needed to be able to rotate, zoom, translate, and scale the graph.
3. Datasets had to be rendered on a three-dimensional graph as a sequence of waveforms, each waveform representing a single dataset. CS-only, paired, and total average waveforms needed to be computed for each dataset and the results displayable on the graph.

Allowing average waveforms to be derived from each datasets was necessary to enable the tracking of changes across sessions during a study.

Rendering performance had to be satisfactory, which becomes an issue when a large number of data points are represented using a 3D software library such as Java3D API.

To allow amplitude of waveform to be estimated from the graph, a topographic color mapping was required. Topographic maps smoothly map each point on the curve to a specific color value that corresponds to some function defined over that curve; in this case the function was the curve's amplitude. By reflecting the amplitude in the color of the curve a comparison within and between graphs could easily be made.

In addition to data visualization, another goal of this project was to build a framework for a robust data processing application suitable to a variety of file formats. Instruments for eyeblink conditioning produce various file formats which nevertheless contain data that is uniform in its underlying structure. Therefore it was essential to design EBCC DAT to be independent of specific data formats and systems that produce them. A uniform data model had to be designed to represent eyeblink classical conditioning data imported from different file formats regardless of its origin.

Data processing capabilities required a method for specifying dependent measures, computing these measures over trials, and producing a trial summary in a convenient format.

Approach

EBCC DAT was developed iteratively, with a number of versions produced during the development. Requirements were revised a number of times in response to feedback from members of the lab.

Data collected in mouse and human eyeblink conditioning sessions was used for testing the software. Results were compared with those produced by other tools available in the lab. Visualization components of the application were tested by comparing graphs to numeric values and the usability of the tool was evaluated during lab meetings and demonstrations of the program's functions.

Results

Design of EBCC DAT

The design of the application involved several choices. Java was a natural choice because of its platform independence and a number of software libraries ideally suited for this project. Leveraging AWT and Swing to create a GUI was made easier due to the versatility of NetBeans IDE. To create 2D plotting capabilities, AWT's built-in features were used. For 3D rendering, Java3D API was chosen as the most suitable software library. In order to read Excel spreadsheets and provide a transparent way for the user to specify measures POI HSSF API was chosen.

The design of EBCC DAT can be separated into five parts:

1. User Interface
2. Graphics Framework
3. Measures Framework
4. Data Model
5. IO Framework

The user interface was designed to provide direct access to application's functionality. Core functionality is accessible through the main application window that provides menus for opening datasets, performing operations on these datasets, and viewing a 2D graph of trials contained in the active dataset. Through the main window's menu 3D views of datasets can be opened, which creates separate window for each new view. To define data processing parameters, a measures and periods window can be opened which provides an interface to create, export, and import a list of measures and periods. Measures and periods window provides user with access to the measures framework of the application. Data set groups window is a component of the application that is currently still in development and is intended to improve the functionality of averaging datasets.

The graphics framework fulfills the data visualization requirements of this application. JPanelPlot3D is a class that paints a rendering of data supplied to it, and responds to events by rotating, translating, zooming, and scaling the rendering. Trial class, which is a part of the data model, provides a view to the data contained in it via an inner class instance of either LineTrial3D or StripTrial3D. That view is passed to JPanelPlot3D to enable it to render the data contained in the Trial.

The measures framework allows measures to be specified over a trial, either by creating a subclass of Measure class or utilizing the flexibility of MeasureCustomFormula and MeasureArrayCustomFormula classes, both of which allow measures to be specified as Excel-style formulas.

The data model is used to represent eye blink conditioning data in a uniform way within the application regardless of the source of data, while allowing for variation of certain parameters in data representation. Data is represented in a hierarchy of datasets and trials, with periods used to connect the data model and the measures framework. Periods define an subinterval of time in a trial, and are used by measures when computing a summary of trials.

The IO Framework provides input and output facilities to allow the application to be easily extended to incorporate additional file formats. DataSetReader, TrialReader, and AbstractSampleReader define operations for reading a sequence of samples from an input stream and storing these in the data model. AbstractSampleReader as well as TrialReader and DataSetReader can be extended to specify how to convert data to Trials and DataSets for a given file format. Two file formats, an Excel spreadsheet format for storing mouse datasets and a binary “.RAW” format produced by San Diego Instruments system for storing human datasets are currently implemented using this method.

Applications

Renderings of eyeblink conditioning data are frequently used in research publications to illustrate data that was collected and methods that were used to analyze it. Views of averaged datasets provide a way to illustrate differences between two groups of subjects, as well as showing aspects of eyeblink conditioning data that are otherwise difficult to perceive.

Observations that can be made from renderings of eyeblink conditioning data include distinguishing CS-only trials from paired trials, distinguishing startle responses from CRs, and the timing and amplitudes of the responses to the CS. Especially in averaged trials, observations of timings and amplitudes can clearly indicate the degree to which CS and US become associated for a given subject.

In addition to serving as a tool for creating renderings for publication purposes, EBCC DAT enables a detailed inspection of data that may reveal previously unnoticed features. Visual inspection thus can provide new insight and enable one to refine data analysis methodology to detect differences in measures that were previously not analyzed.

EBCC DAT can be used to produce summaries of datasets with measures calculated over trials. By allowing these measures to be defined using Excel-like syntax, a smooth transition from analysis involving Excel spreadsheets is provided. Summaries are generated in Excel format, so results can easily be imported into other software for further statistical analysis.

Conclusions

As a result of work involving EBCC DAT a better understanding of methods involved in data analysis and visualization was gained. Significant insight into 3D

programming techniques was obtained, and a fully functional application for 3D data visualization was developed. While EBCC DAT is fully functional and fulfills the requirements of this project, development will continue to further extend this application. A number of refinements remain to be implemented, and additional features will make the application significantly more powerful. Among desirable additions to EBCC DAT is the addition of 3D labeling of axes, placement of markers for orientation, enhancement to the user interface, and additional data processing capabilities. Because of its flexible design, EBCC DAT could also be extended for processing of data from research domains other than eye blink classical conditioning. A further improvement to the application would be the addition of a database, which would centralize data storage and reduce difficulties associated with associating metadata with datasets stored as separate files.