# Finding structure in time

## ICIS Preconference Workshop 2018

## In memory of Jeff Elman


Tian Linger Xu

txu@indiana.edu

Indiana University Bloomington

Ralf F. A. Cox

r.f.a.cox@rug.nl

University of Groningen

Kaya de Barbaro

kaya@austin.utexas.edu

The University of Texas at Austin

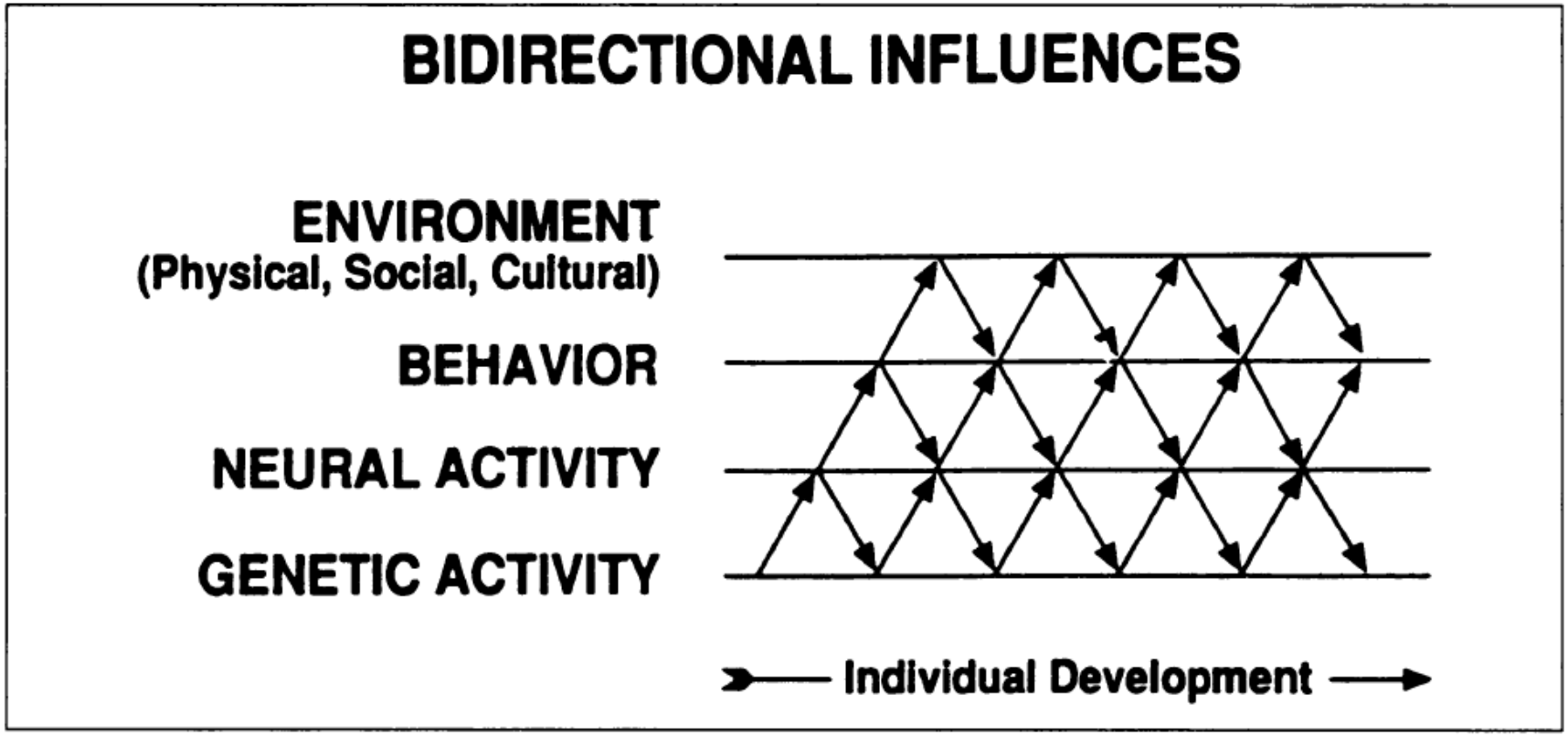
Drew Abney

dhabney@indiana.edu

Indiana University Bloomington

# Why time?

Development is an inherently temporal phenomenon




Gottlieb

# Why time?

- Development is inherently a temporal phenomenon
  - Outcomes emerge from transactive relationships between determinants internal and external to the individual

- Time is central to many key developmental processes
  - Responsiveness
  - Synchrony
  - Regulation
  - Resilience
  - Feedback
  - Learning

# Why time? What structure?

**internal states**

**gaze**

**language**

**affect**

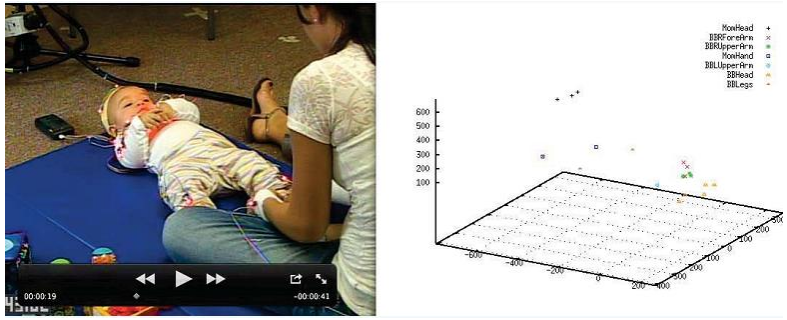
**physiology**

**hands**

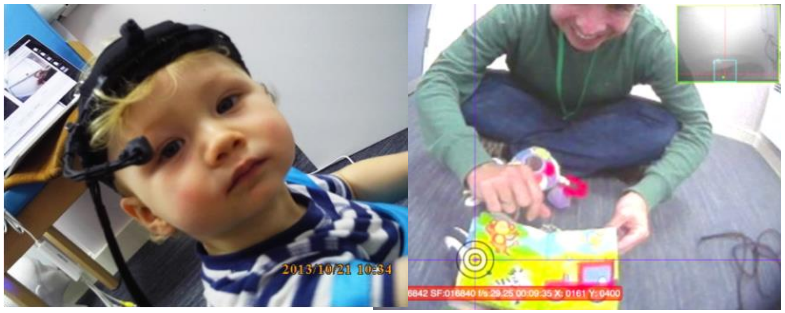
**locomotion**

Timing and structure of activity on micro timescale has consequences on longitudinal timescale

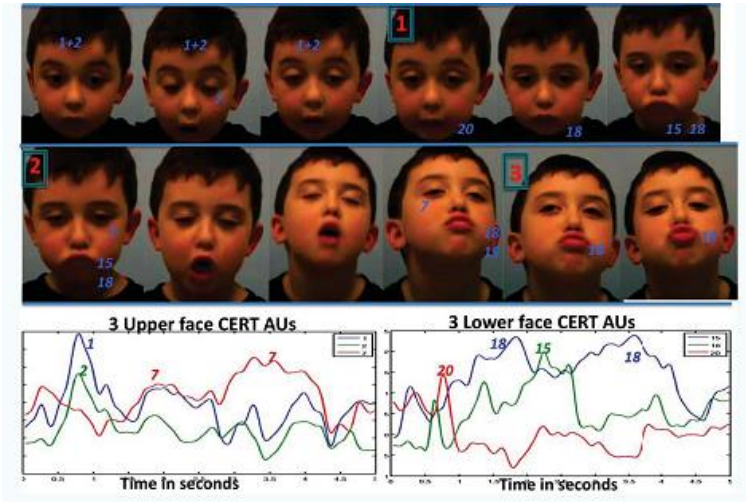
# New opportunities for data collection



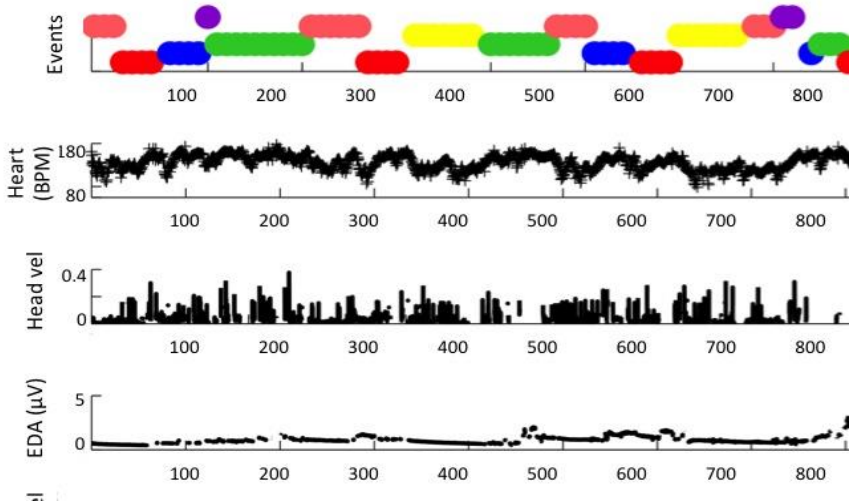
Sensorimotor activity



Gaze

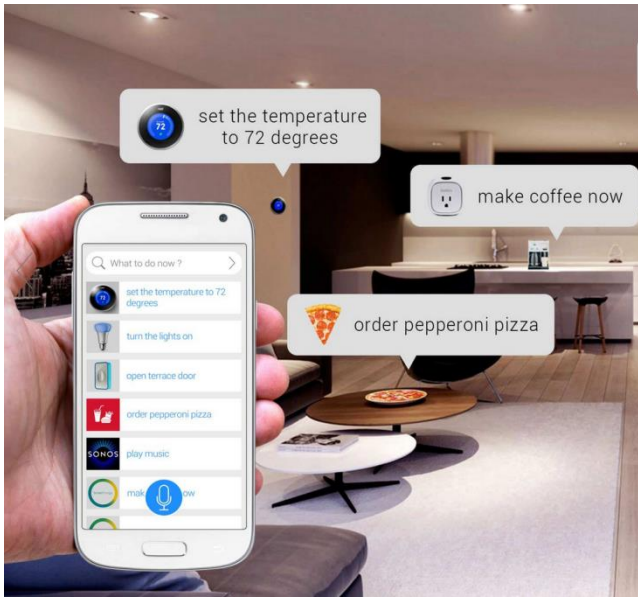


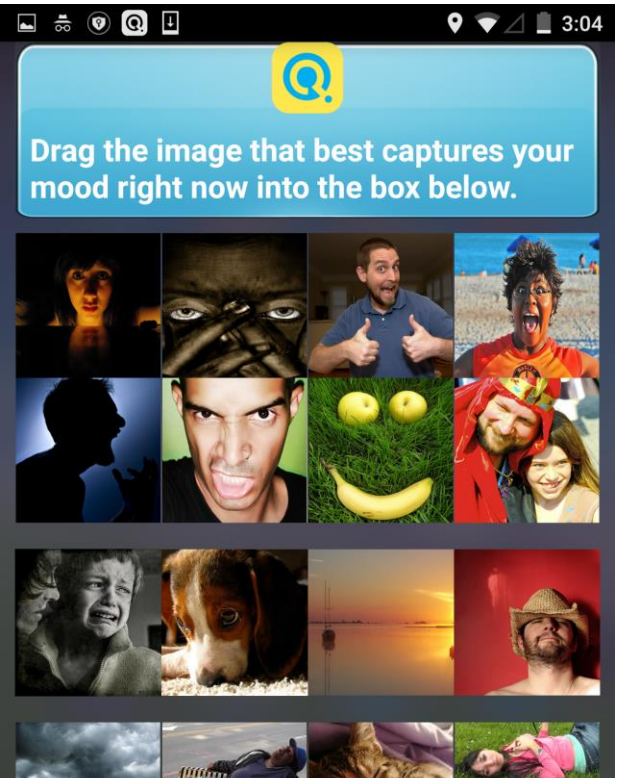

Facial expression

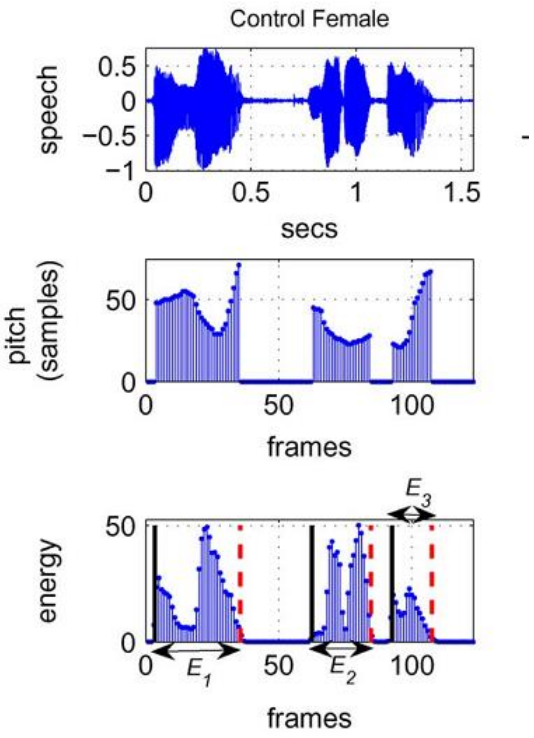


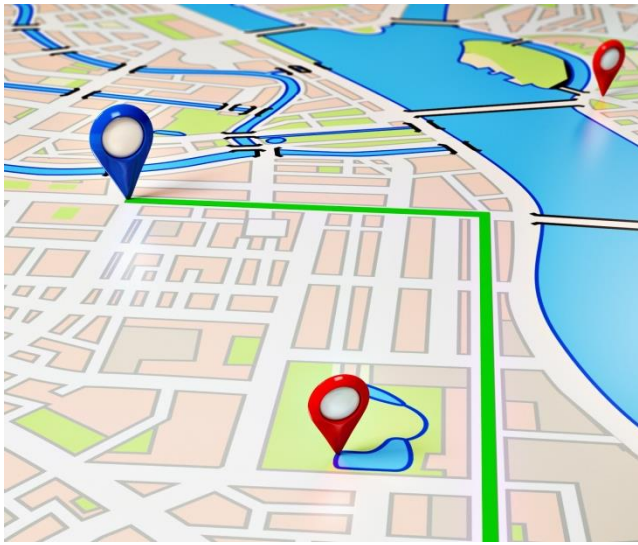

Physiology


fitbit
Discover a whole
new world of fitness.


set the temperature
to 72 degrees
make coffee now
order pepperoni pizza


3:04
Drag the image that best captures your
mood right now into the box below.


Control Female
speech
0.5
0
−0.5
−1
0
0.5
1
1.5
secs
pitch
(samples)
50
0
0
50
100
frames
energy
50
0
$E_1$
$E_2$
$E_3$
0
50
100
frames



# New challenges

How can we find structure in time?

# Goals for code paper

- Provide a sampling of techniques for characterizing temporal structure of high-density multimodal data
- Provide scripts and exercises so you can confidently use these tools with your own datasets
- Allow participants across a range of experience with Matlab (or programming) to learn something new

# Overall structure of modules

- Modules 1 & 2

Basic qualitative and quantitative techniques for discovery of structure in high-density data

Data visualization and clustering analyses

- Modules 3 & 4

Advanced algorithms for quantifying patterns of interaction between individuals or modalities

Recurrence Quantification and Granger Causality

# Visualization of high-density activity data

# Why visualize?

- Visualizations of raw data can provide insight
  - New, richer data: new, richer understanding of phenomenon
  - Theory is often underspecified as to dynamics of behavior

Visualizations can suggest

- Relevant analyses
- How to operationalize underspecified phenomena

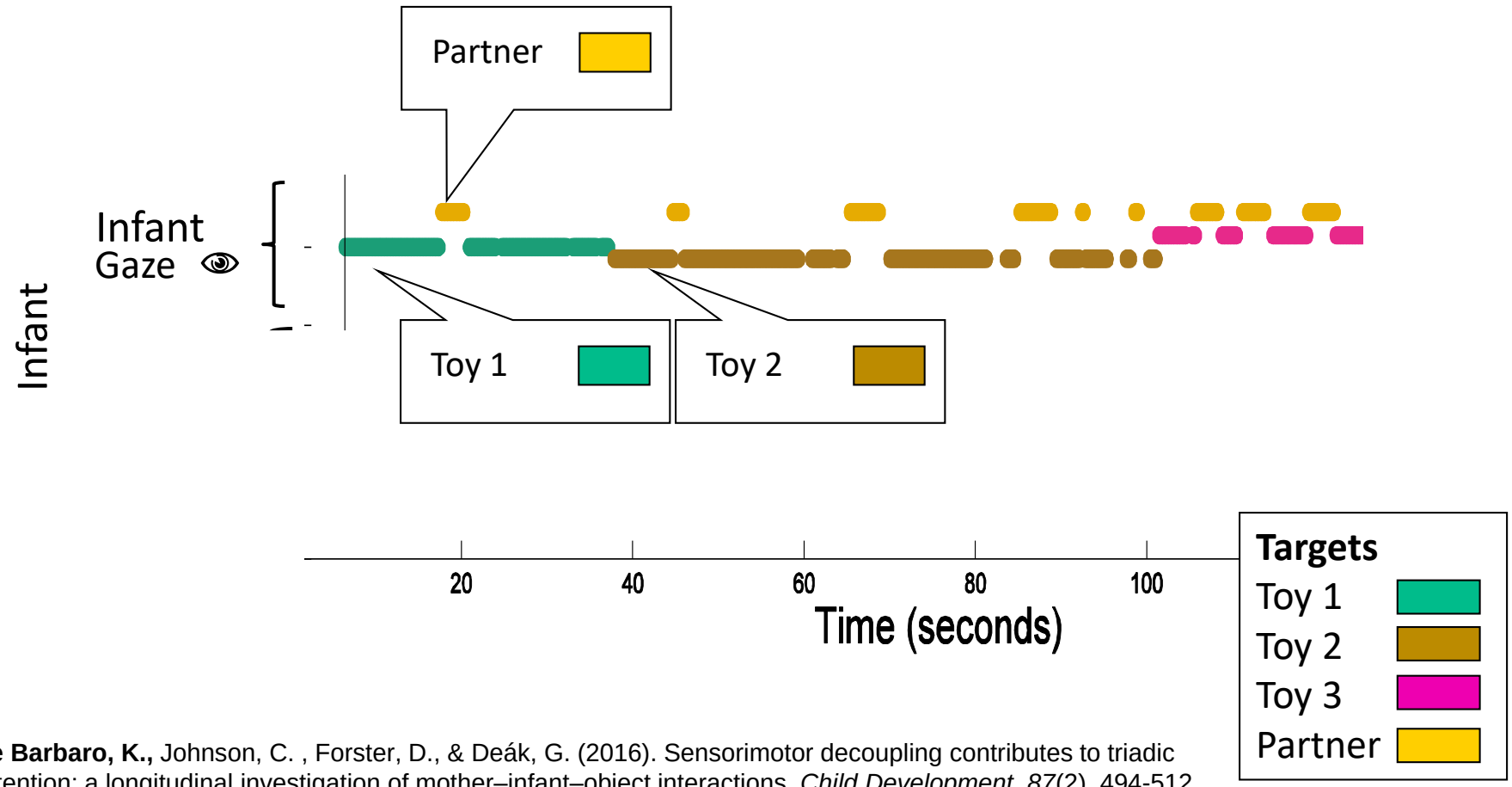
- Visualizations can also be used during analyses
  - Double check that you are capturing what you think you are

# Goals of visualization module

- Provide opportunity to make and manipulate simple plots in MATLAB for common data
  - Event streams from video annotations
  - Continuous timeseries data
- Provide you with confidence and tools to work with your data in Matlab
  - Script your own plots
  - Combine simple techniques to build more complex plots

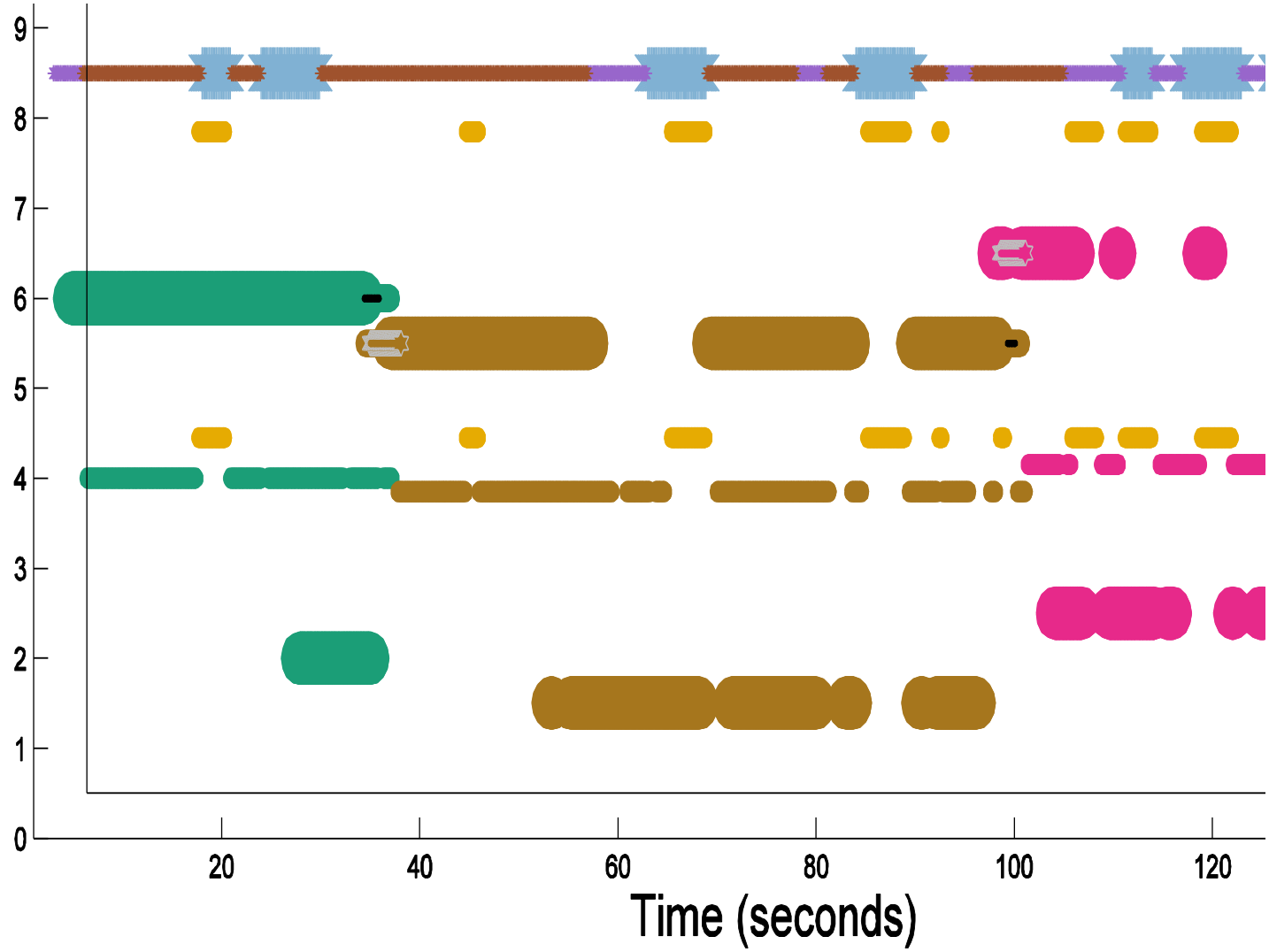
# Categorical datastreams

## 4-mo dyad: 100 sec of free play with objects

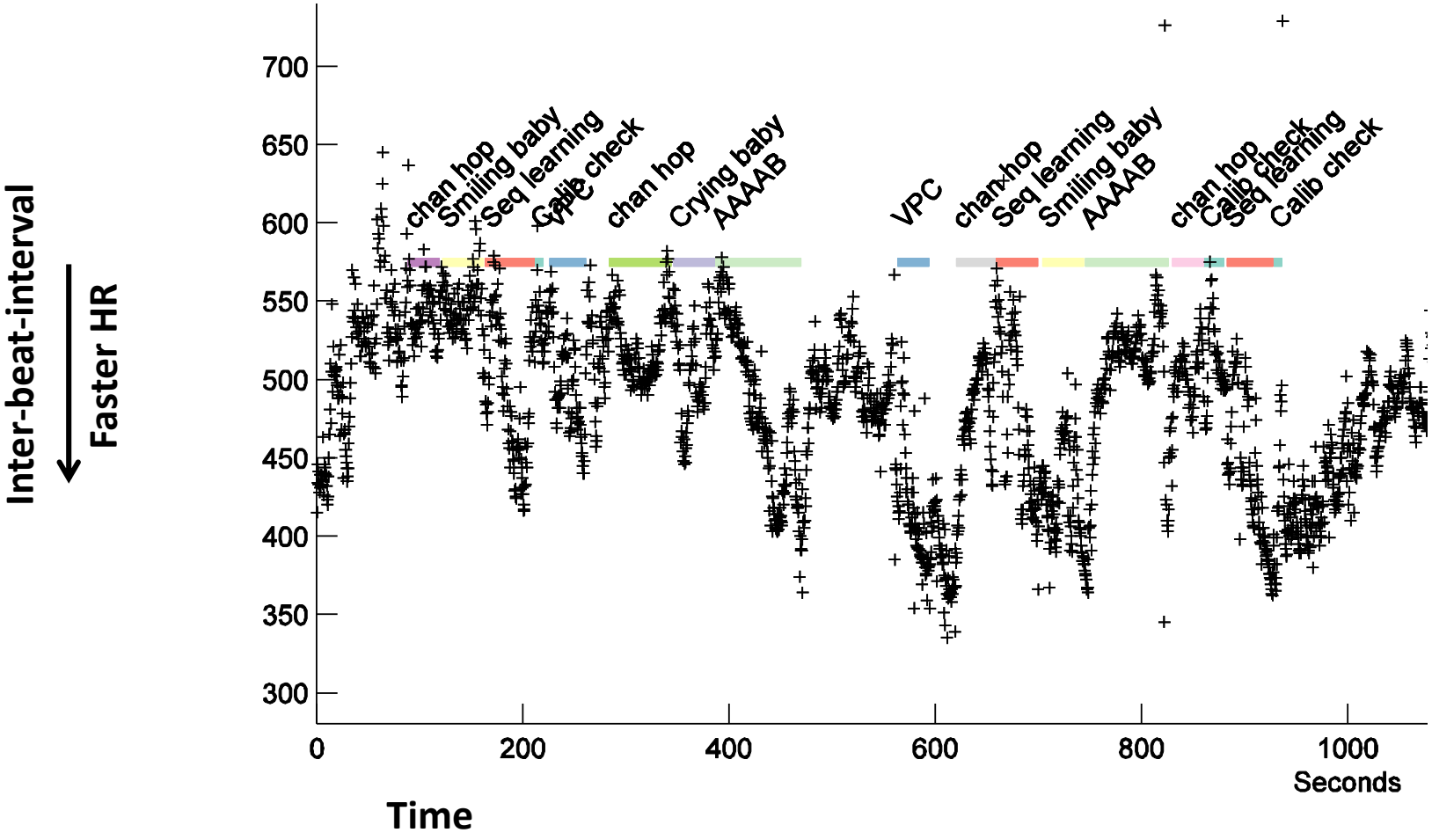



**de Barbaro, K.,** Johnson, C. , Forster, D., & Deák, G. (2016). Sensorimotor decoupling contributes to triadic attention: a longitudinal investigation of mother–infant–object interactions. *Child Development, 87*(2), 494-512.

# Same plot directly out of Matlab

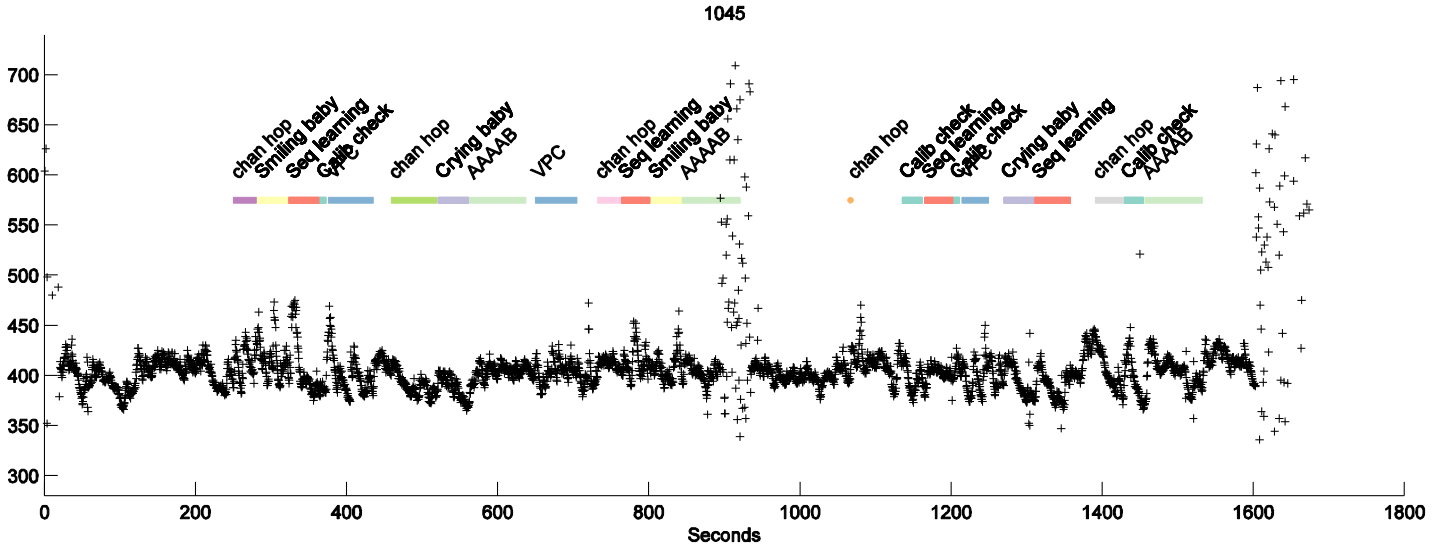


# Arousal dynamics across the session

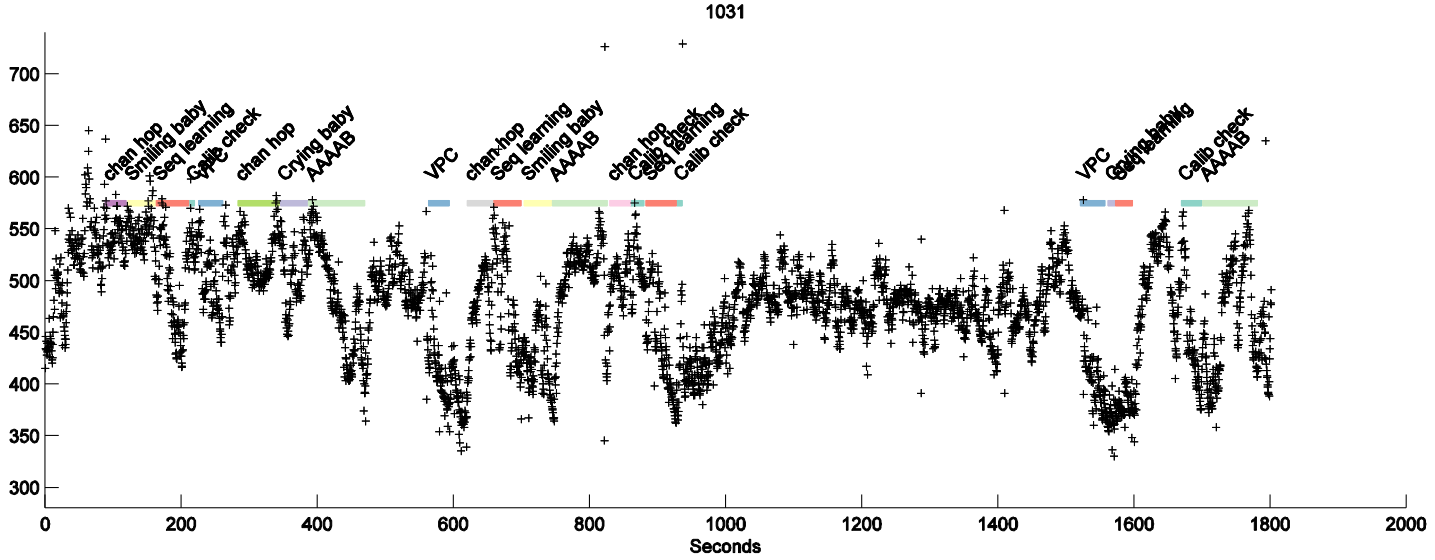


# Individual differences in arousal dynamics across the session

**Baby 1**

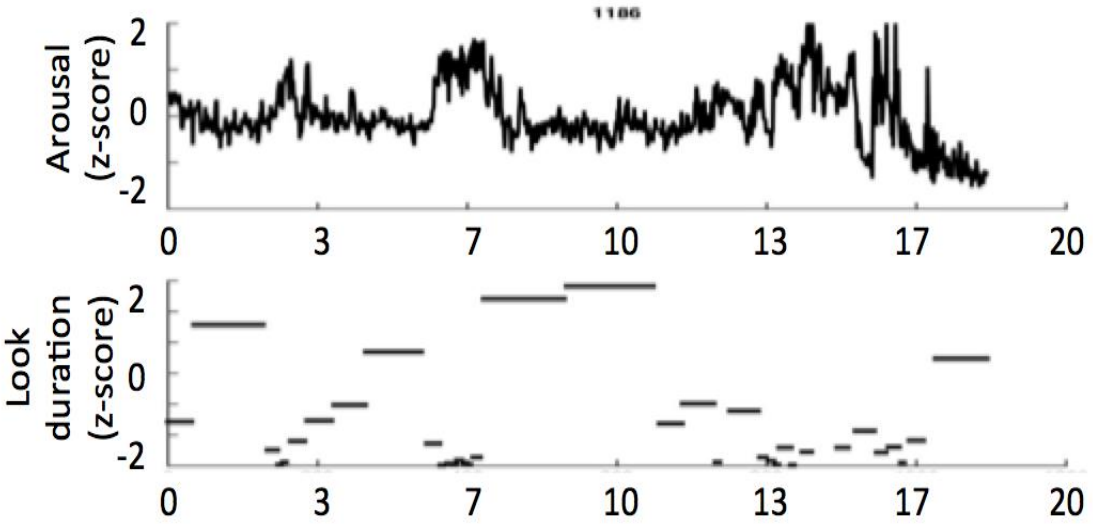



**Baby 2**




Faster HR

# Combining continuous and categorical datastreams




**de Barbaro, K.**, Wass, S. & Clackson, K. (2017). Infant attention is dynamically modulated with changing arousal levels. *Child development*, *88*(2), 629-639.

# Combining continuous and categorical data

Ahsan Hussain

Mom mood

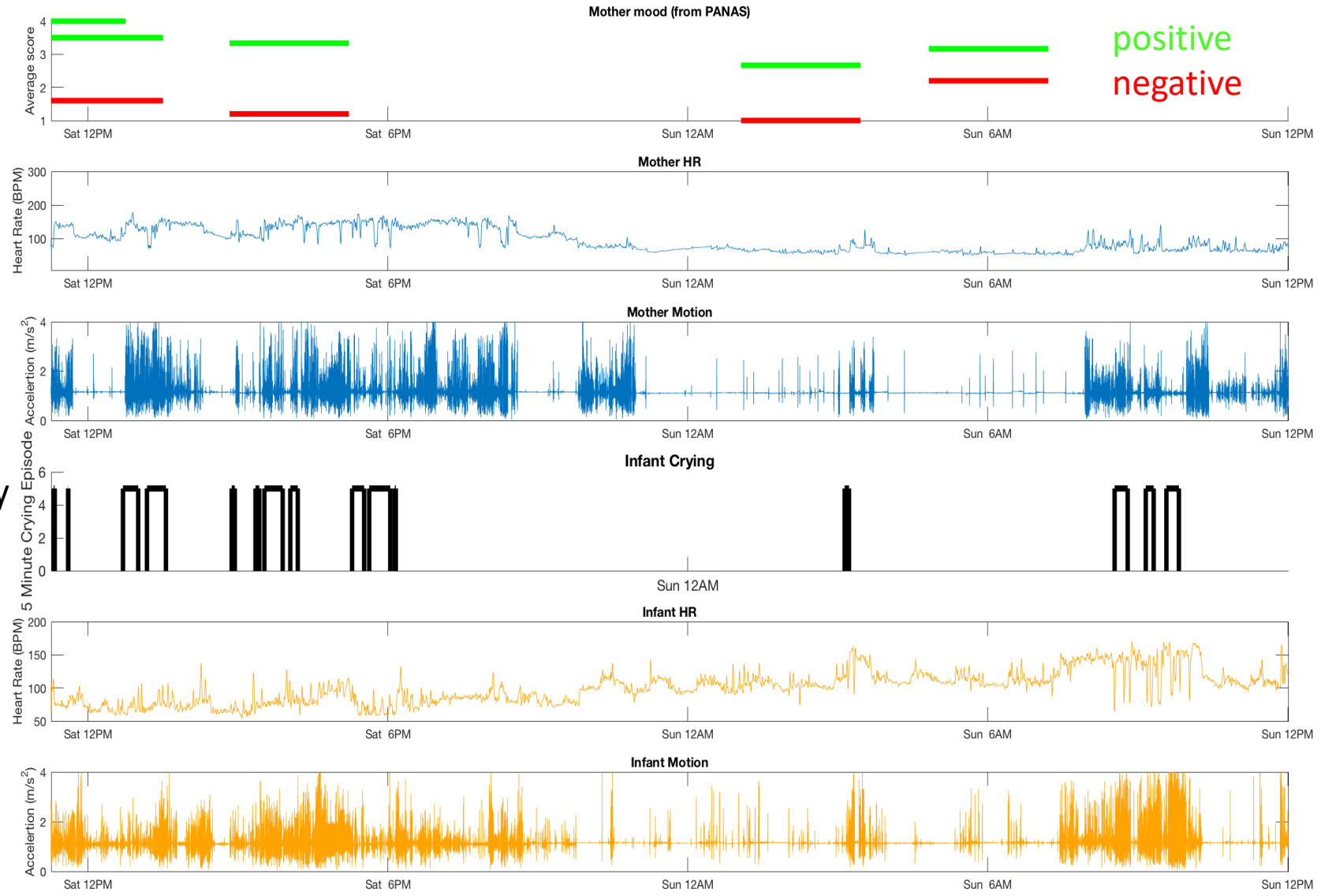
Mom HR

Mom motion

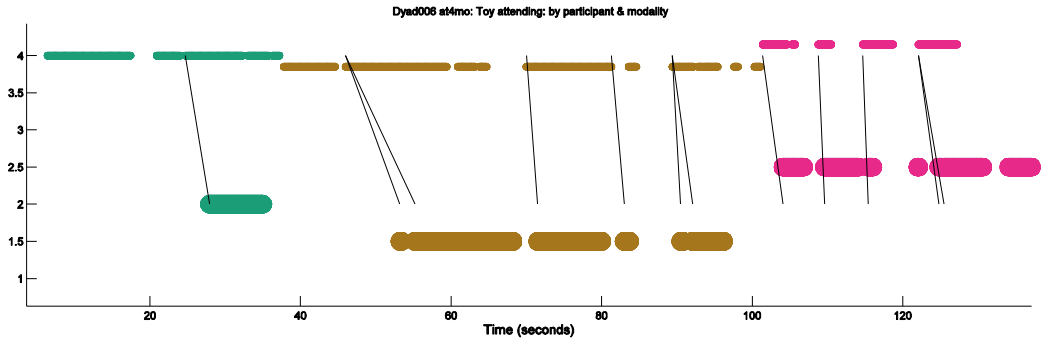
Baby cry (LENA)

Baby HR

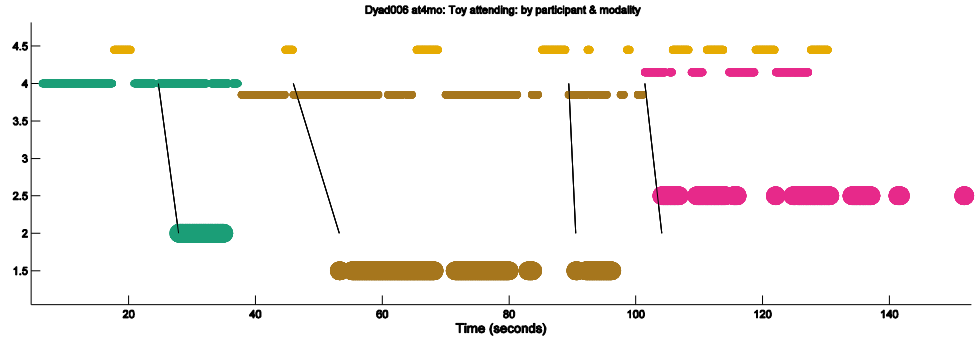
Baby motion



# Timing of gaze relative to hands

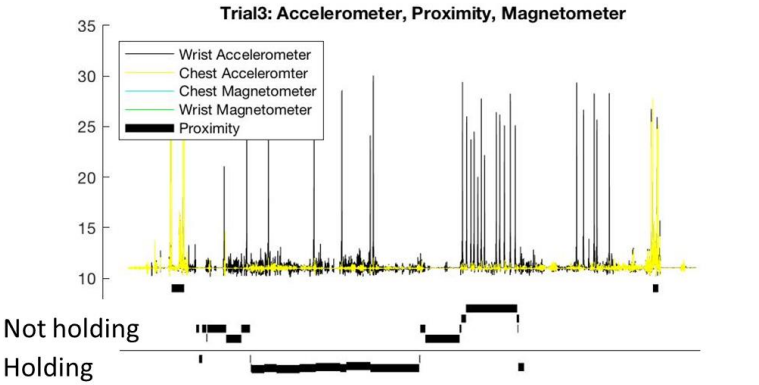


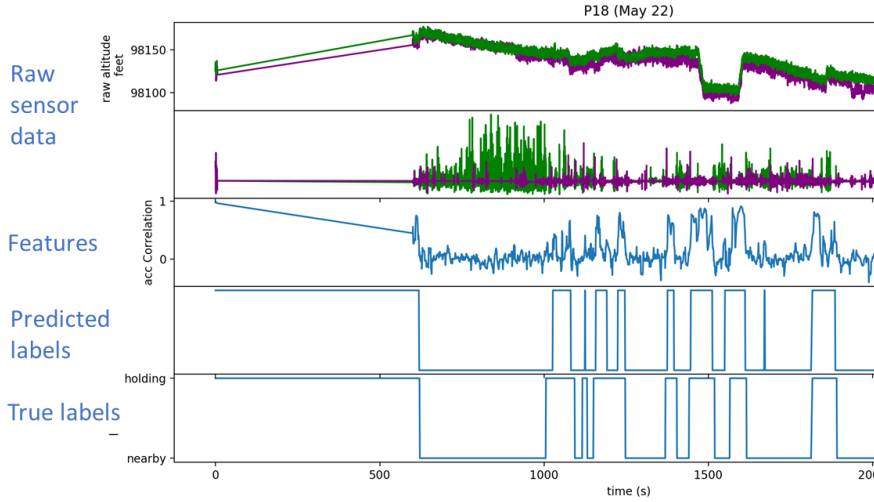

Consider every instance of manual contact?

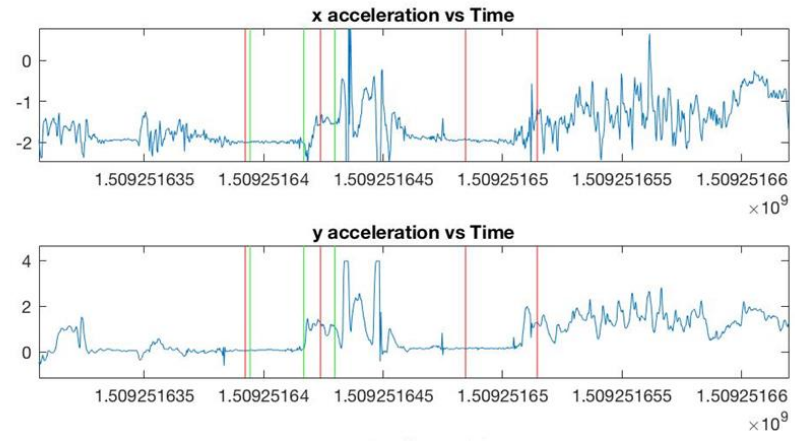


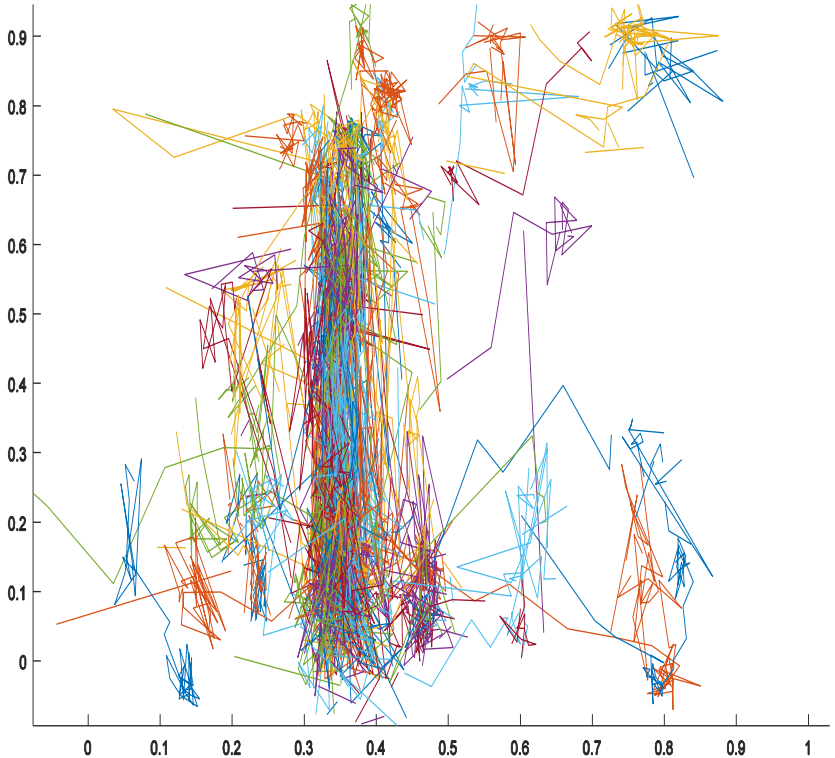

Compare manual contact to “first” contact with object

# Endless possibilities



# Before we plot …

- Cover some basic programming in matlab

- Consider the structure of your data
- Consider how time is represented in your data

- Introduce you to plot logic and syntax

# Helpful resources for Matlab

Link to cheat sheet

- http://web.mit.edu/18.06/www/MATLAB/matlab-cheatsheet.pdf

Link to basic Matlab tutorial

- https://www.math.utah.edu/~eyre/computing/matlab-intro/

# Basic programming

**Working with numeric data structures**

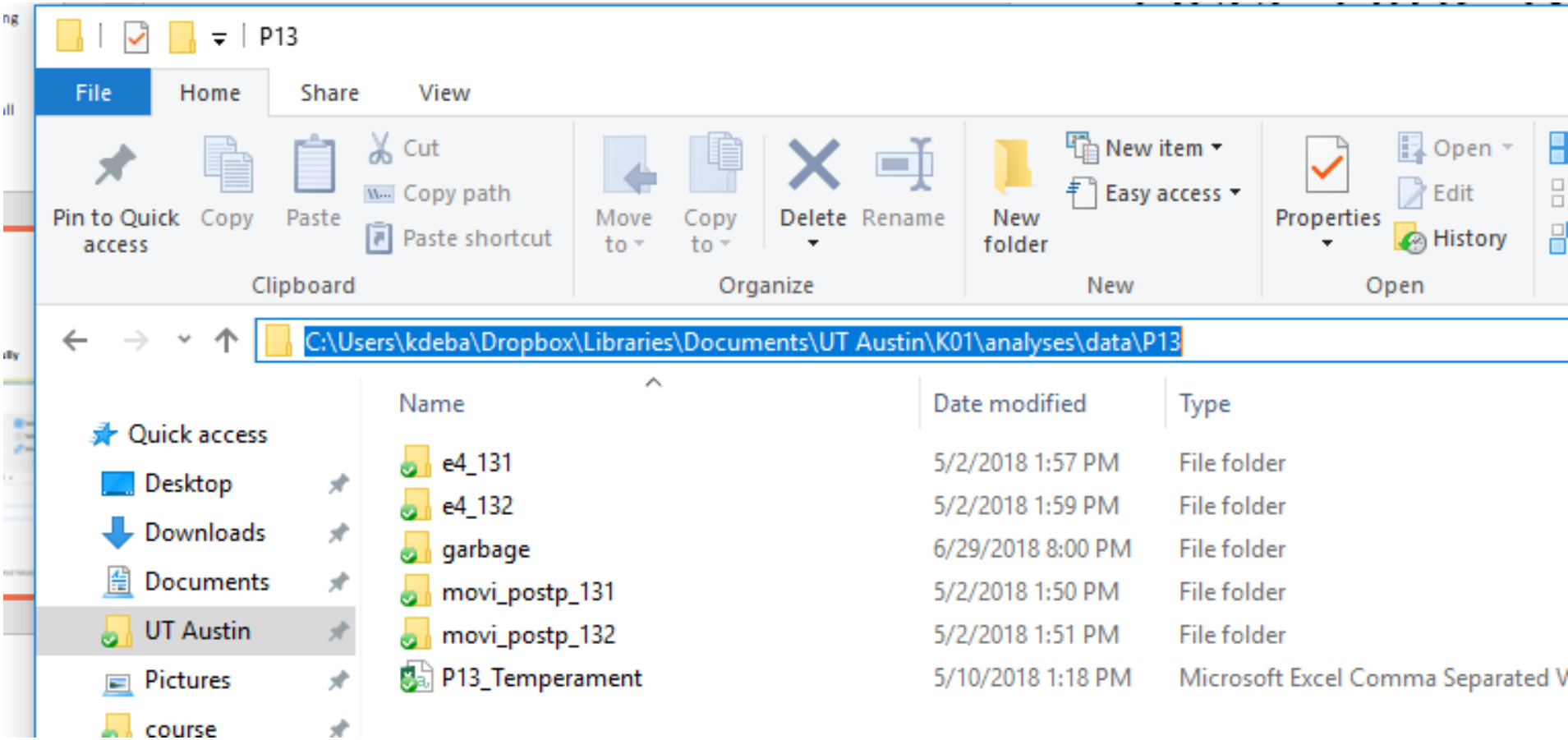
- Storing data in an array
- Basic referencing
- Logical indexing

**Basic functions and commands**

- csvread
- size
- for loops
- readtable
- cellfind

```
% time for some matlab!
ProgrammingBasics.m
```

# Accessing data via Matlab



# Basic data structures

- Timeseries
- Event sequences
- … spike trains
- etc

Data structure determines

- Methods for visualization
- What kinds of analyses are possible

You can transform your data from event sequences to timeseries and vice versa depending on your needs

# Event sequence

A sequence of events each indicated by an onset and offset

**Examples**

- Sequence of infant gaze targets, with onset and offset of each look
- Sequence of dyadic interaction states, with onset and offset of each state

# Timeseries

A sequence of data taken at successive equally spaced points in time

**Examples**

- Hour-by-hour count of words heard in a day
- Second-by-second average heart rate
- Frame-by-frame the presence or absence of mutual gaze

# Event sequence

A sequence of events each indicated by an onset and offset

Variables - data_events

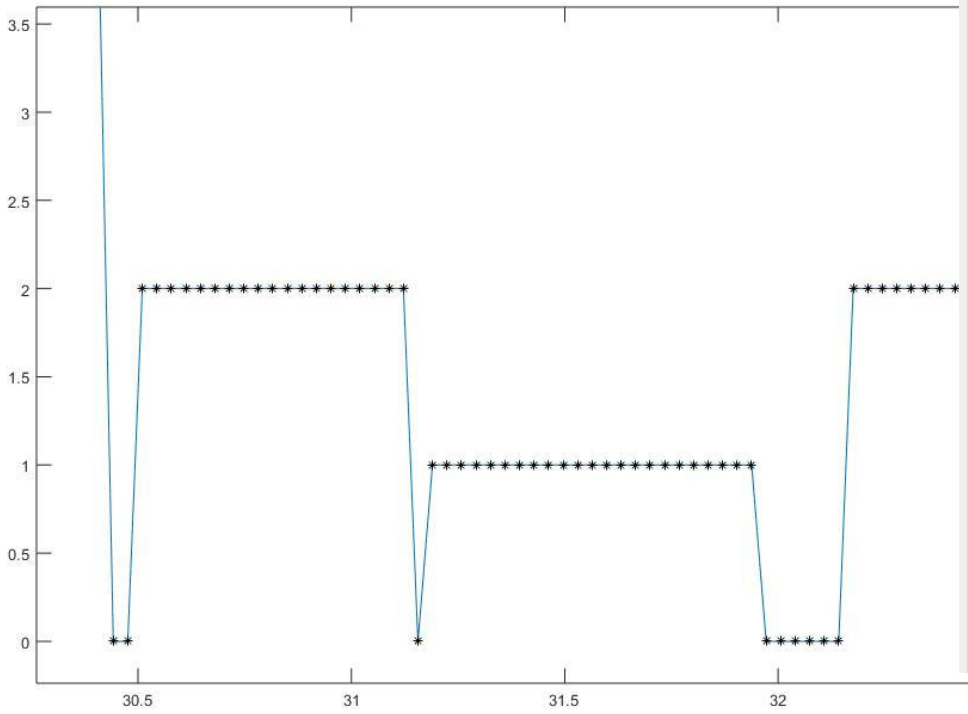
data_stream | data_events

87x3 double

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 30 | 30.4330 | 4 | |
| 2 | 30.5000 | 31.1330 | 2 | |
| 3 | 31.1670 | 31.9670 | 1 | |
| 4 | 32.1670 | 33.7330 | 2 | |
| 5 | 33.8000 | 34.4330 | 1 | |
| 6 | 34.5000 | 35 | 3 | |
| 7 | 35.1330 | 35.5330 | 2 | |
| 8 | 35.6330 | 37.0670 | 4 | |
| 9 | 37.1670 | 39.4330 | 2 | |
| 10 | 39.6000 | 40.2670 | 2 | |
| 11 | 41.4330 | 42.7330 | 2 | |
| 12 | 43.2330 | 46.7000 | 1 | |
| 13 | 46.8330 | 51.5000 | 2 | |
| 14 | 53.2330 | 56.0330 | 1 | |
| 15 | 56.1670 | 56.6670 | 3 | |
| 16 | 56.7000 | 56.8000 | 2 | |
| 17 | 57.1670 | 58.4670 | 1 | |
| 18 | 61.8330 | 62.8670 | 3 | |
| 19 | 63.1670 | 63.6670 | 1 | |
| 20 | 63.8000 | 66.7330 | 3 | |
| 21 | 68.6330 | 69.6670 | 1 | |
| 22 | 70 | 70.3670 | 1 | |
| 23 | 70.6330 | 71.2670 | 3 | |
| 24 | 71.3000 | 76.2000 | 2 | |
| 25 | 77.0670 | 77.8000 | 2 | |
| 26 | 78.2330 | 78.5000 | 3 | |

# Timeseries

A sequence of data taken at successive equally spaced points in time




Variables - data_stream

data_stream | data_events

7183x2 double

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 30 | 4 | | |
| 2 | 30.0340 | 4 | | |
| 3 | 30.0680 | 4 | | |
| 4 | 30.1020 | 4 | | |
| 5 | 30.1360 | 4 | | |
| 6 | 30.1700 | 4 | | |
| 7 | 30.2040 | 4 | | |
| 8 | 30.2380 | 4 | | |
| 9 | 30.2720 | 4 | | |
| 10 | 30.3060 | 4 | | |
| 11 | 30.3400 | 4 | | |
| 12 | 30.3740 | 4 | | |
| 13 | 30.4080 | 4 | | |
| 14 | 30.4420 | 0 | | |
| 15 | 30.4760 | 0 | | |
| 16 | 30.5100 | 2 | | |
| 17 | 30.5440 | 2 | | |
| 18 | 30.5780 | 2 | | |
| 19 | 30.6120 | 2 | | |
| 20 | 30.6460 | 2 | | |
| 21 | 30.6800 | 2 | | |
| 22 | 30.7140 | 2 | | |
| 23 | 30.7480 | 2 | | |
| 24 | 30.7820 | 2 | | |
| 25 | 30.8160 | 2 | | |
| 26 | 30.8500 | 2 | | |
| 27 | 30.8840 | 2 | | |
| 28 | 30.9180 | 2 | | |
| 29 | 30.9520 | 2 | | |
| 30 | 30.9860 | 2 | | |

# Basic plot logic and syntax

*Scatterplot for timeseries data*

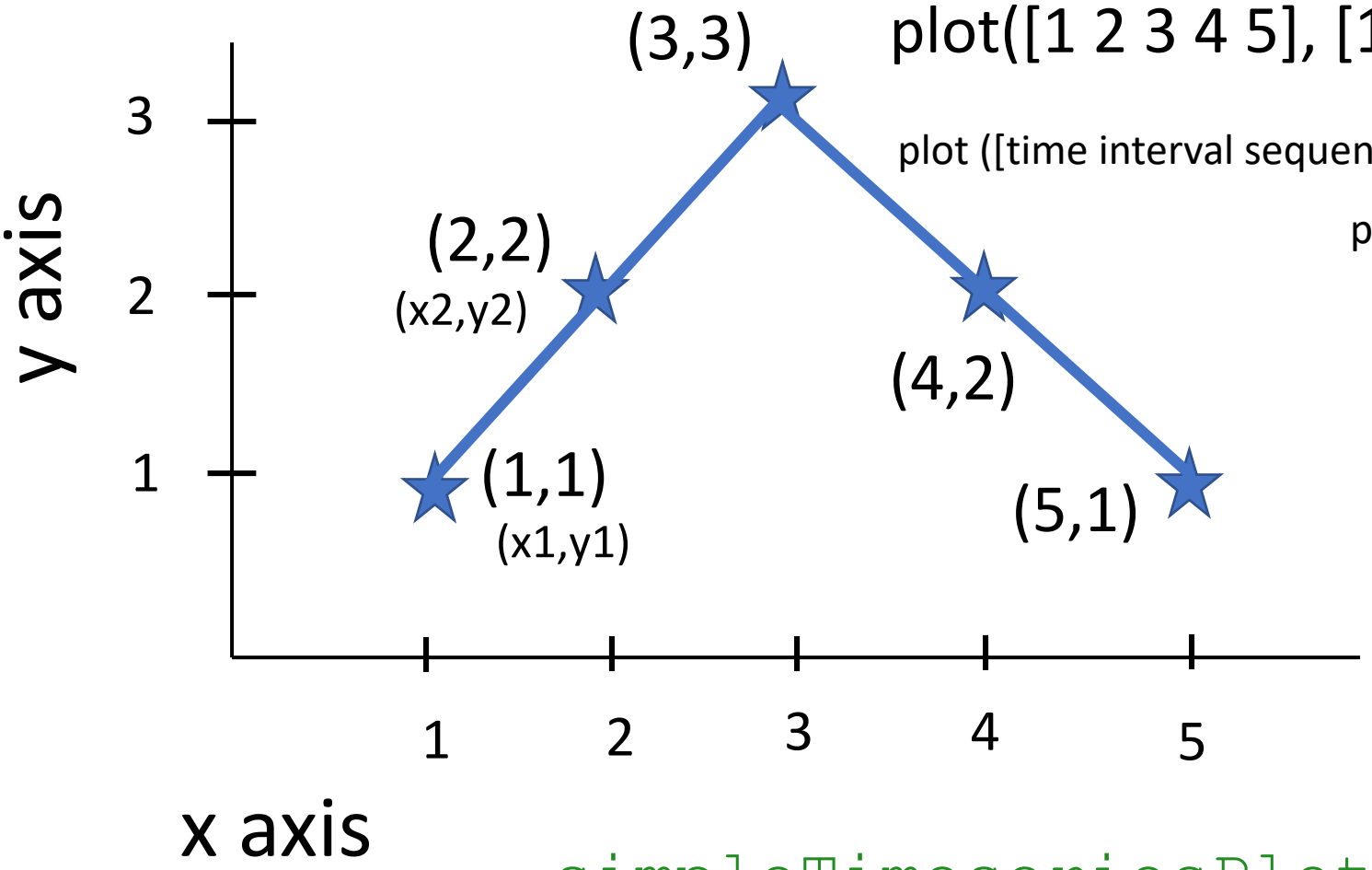
plot(xvals,yvals)

plot(1,1)

plot([1 2 3 4 5], [1 2 3 2 1])

plot ([time interval sequence], [values])

plot ([values])

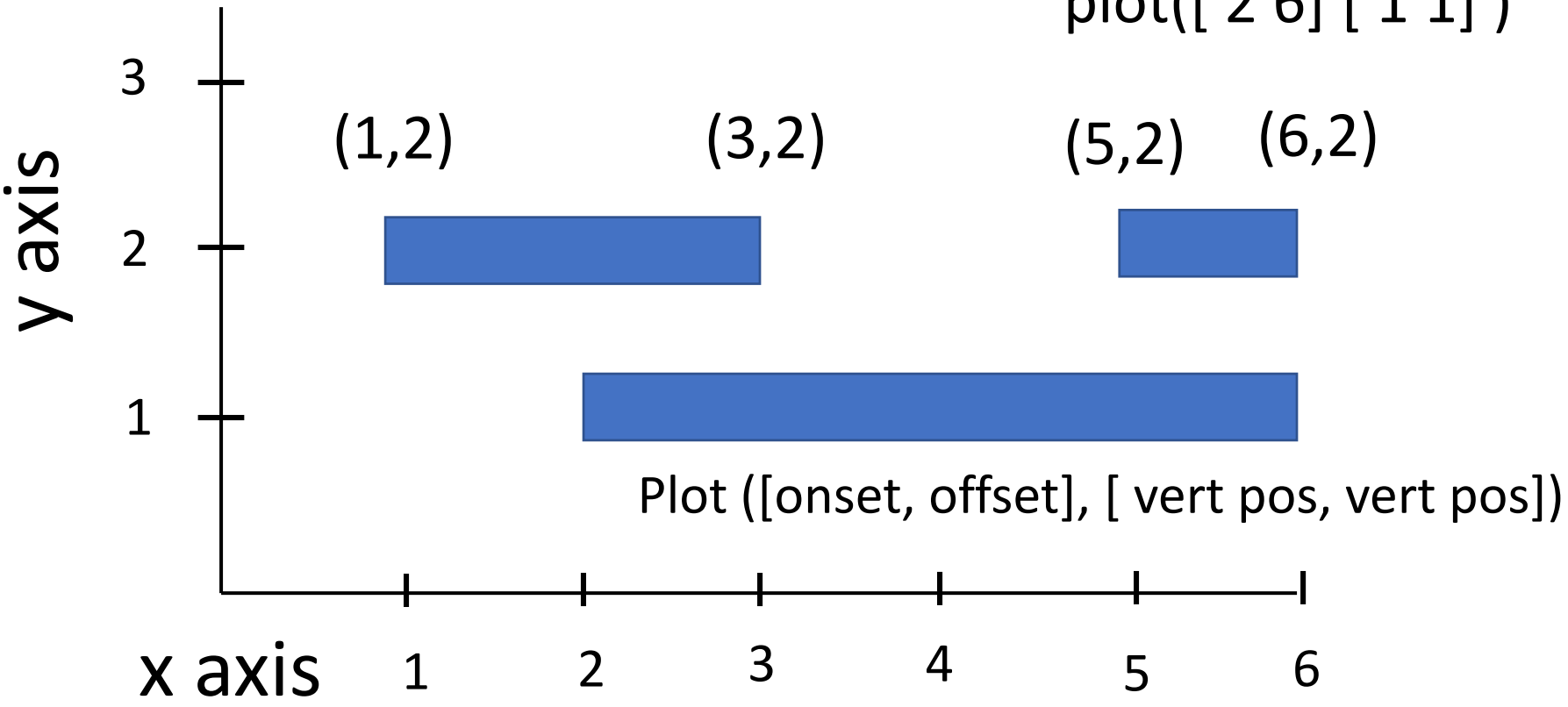



simpleTimeseriesPlots.m

# Basic plot logic and syntax

*Series of line plots for event sequence data*

```
plot(xvals,yvals)
plot([x1 x2] ,[ y1 y2])
plot([ 1 3] [ 2 2] )
plot([ 5 6] [ 2 2] )
plot([ 2 6] [ 1 1] )
```




simpleAnnotationPlots.m

# Basic plot syntax: review

**plot (xvals, yvals)**

**Timeseries: scatterplot!**

**Plot ([time interval sequence], [values])**

**Plot ([values])  % plot (yvals)**

**Event sequence: drawing one “line” per event**

**For each event:**

**Plot ([onset, offset], [ vertical position, vertical position] )**


Event data warning !

# If you want to import event data from common annotation software… Be careful!

Matlab was not designed to work with any data containing “strings”– i.e. files that include any letters rather than numbers (even in the headers!)

1. Thus, annotation software outputs (e.g. DataVyu, Elan, Mangold Noldus files) are difficult to work with in Matlab
2. And it is best to translate all outputs into numeric data

You can modify the script annotationImport.m to translate typical annotation outputs files into clean event sequences (i.e. just numbers)

# Also: categorical event sequences are common data for behavioral science – but—

1. Many analyses are restricted to data in timeseries format
2. Timeseries are much easier to work with in matlab

If you have CLEAN event sequence data, you can use:

Events2Timeseries_fromAnnotationImportScript.m

to convert event sequence data into timeseries data

** CLEAN = one column of onsets, one column of offsets, one column of events in a CSV file

Plus, not a single letter (!) in the whole CSV file

E.g. 'cevent_eye_roi_child.csv' in the data folder

Or – any output from annotationImport.m

# Available in this module:

1. Scripts to learn basic data manipulation

   `ProgrammingBasics.m`

2. Scripts to import annotation software outputs into Matlab

   `annotationImport.m`

3. Scripts to transform CLEAN event sequence data into timeseries data

   `events2Timeseries_fromAnnotationImportScript.m`

4. Scripts for making your own plots!

   Lots! Detailed in `READ ME.m`

# Thank you!

Special thanks to Gedeon Deak (UCSD) and Sherryl Goodman (Emory University) for sharing data samples

**Kaya de Barbaro, PhD**

kaya@austin.utexas.edu

Assistant Professor

Department of Psychology

The University of Texas at Austin