Analyzing Visual World Eye-tacking Data: A Hands-on Introduction

Yuhang Xu/徐雨航

University of Rochester

yuhang.xu@rochester.edu

Plan

- **2:30-3:20 “Getting the data”**
- **3:20-3:30 Break**
- **3:30-4:30 “Analyzing the data ”**
- **4:30-5:00 Q & A**
- https://xyuhang.wordpress.com

Qs you may have…

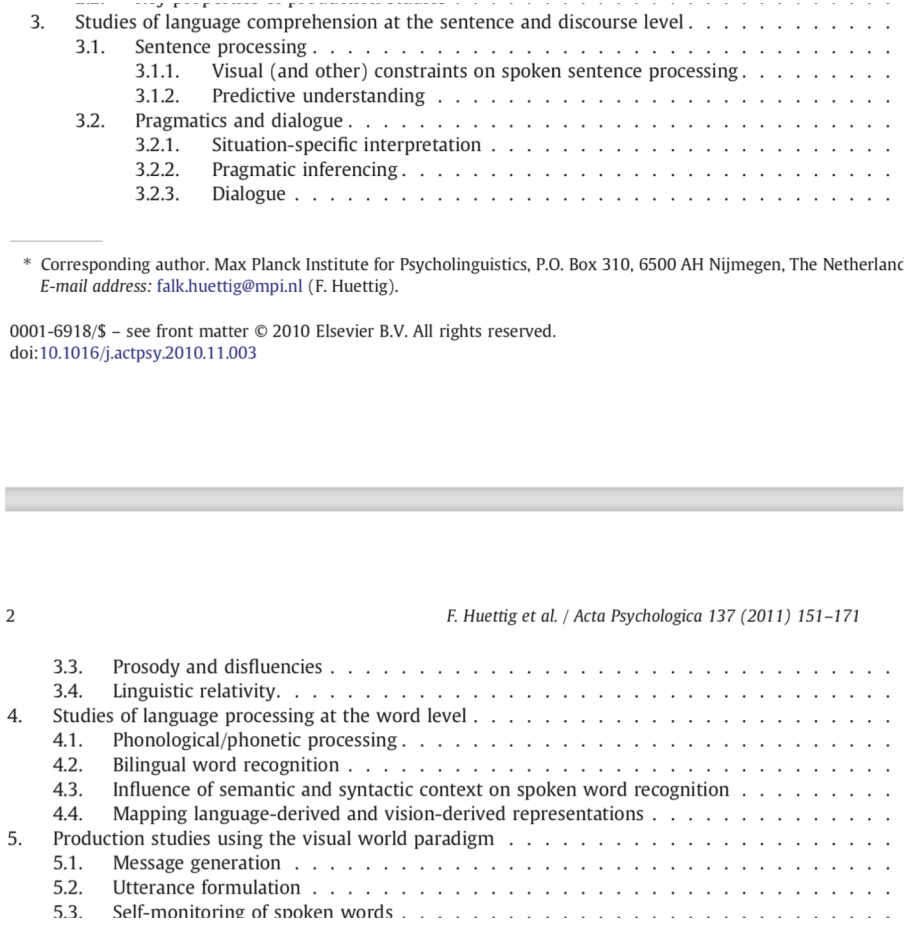
- **Do I really need the VWP?**

- **For example, the VWP cannot be used for the study of**
  - (a) language that is not at least partially about the visual world;
  - (b) language that is about events and entities that cannot easily be depicted
  - (c) written language. (reading…)

    - Salverda, A.P., & Tanenhaus, M.K. (in press). The visual world paradigm. In A.M.B. de Groot & P. Hagoort (Eds.), *Research methods in psycholinguistics: A practical guide*. Malden, MA: Wiley-Blackwell.

But…

Applications of the VWP



3. Studies of language comprehension at the sentence and discourse level
   - 3.1. Sentence processing
     - 3.1.1. Visual (and other) constraints on spoken sentence processing
     - 3.1.2. Predictive understanding
   - 3.2. Pragmatics and dialogue
     - 3.2.1. Situation-specific interpretation
     - 3.2.2. Pragmatic inferencing
     - 3.2.3. Dialogue

* Corresponding author. Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlanc
E-mail address: falk.huettig@mpi.nl (F. Huettig).

0001-6918/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2010.11.003

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. Acta Psychologica, 137, 151–171.

Qs you may have…

- I don't have a fancy eye-tracker…
- I do field work…

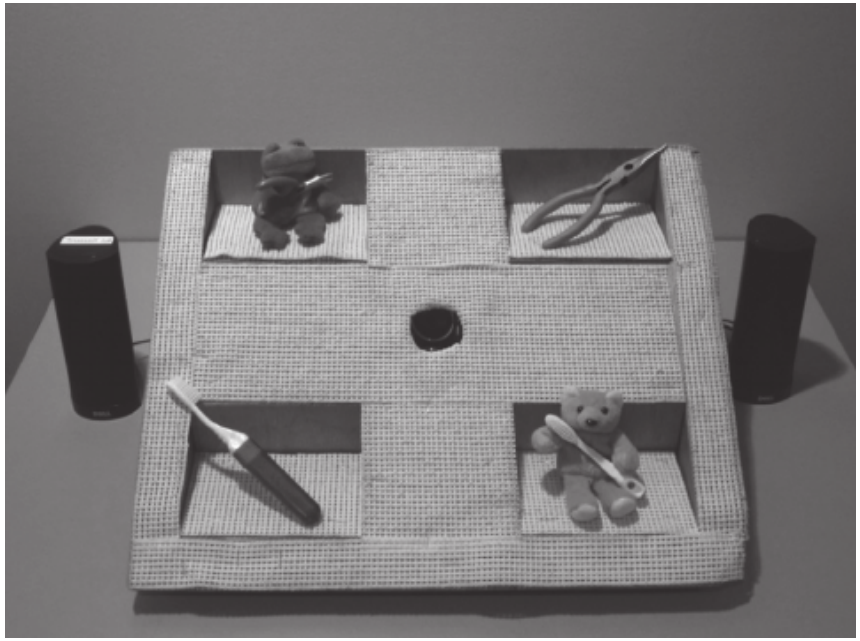


Figure . A set-up for the 'Poor Man's' eye-tracker Pozzan and Trueswell

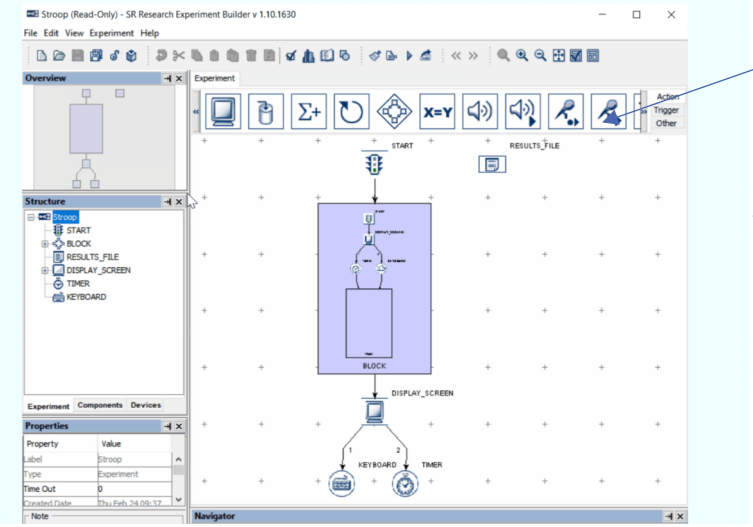
Qs you may have…

- **Why do we need to cover both “experiment programming” and “data analysis”**

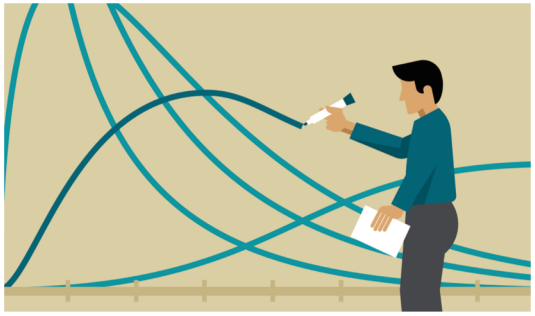
- **Examples:**

Example:

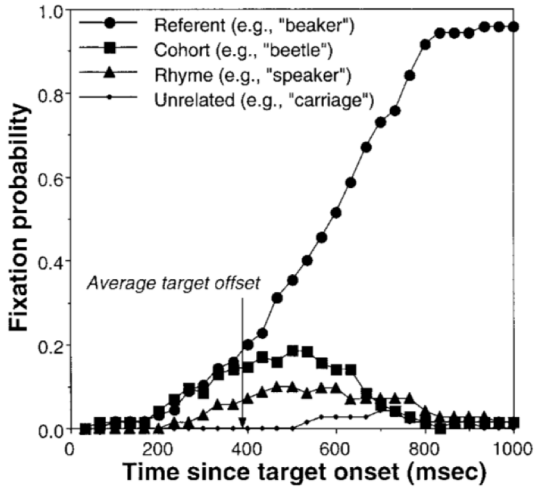
- **When Implementing an visual world experiment using Experiment Builder…**

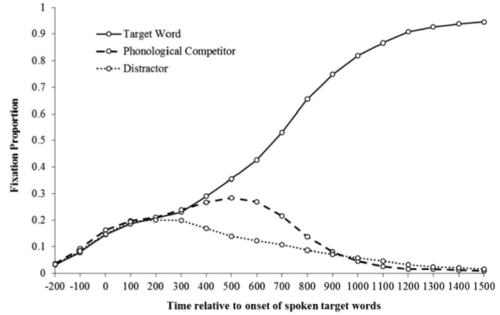


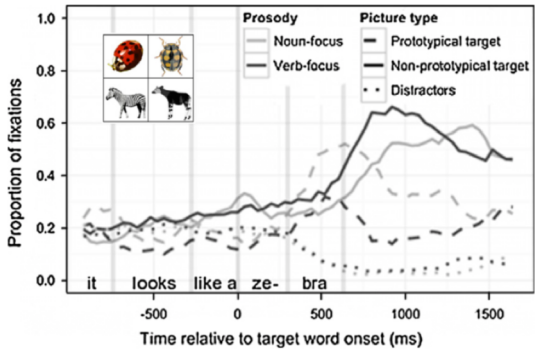
What to include?

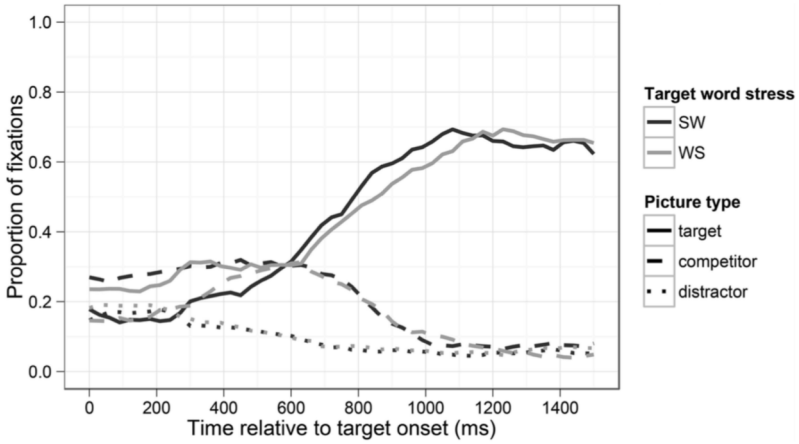


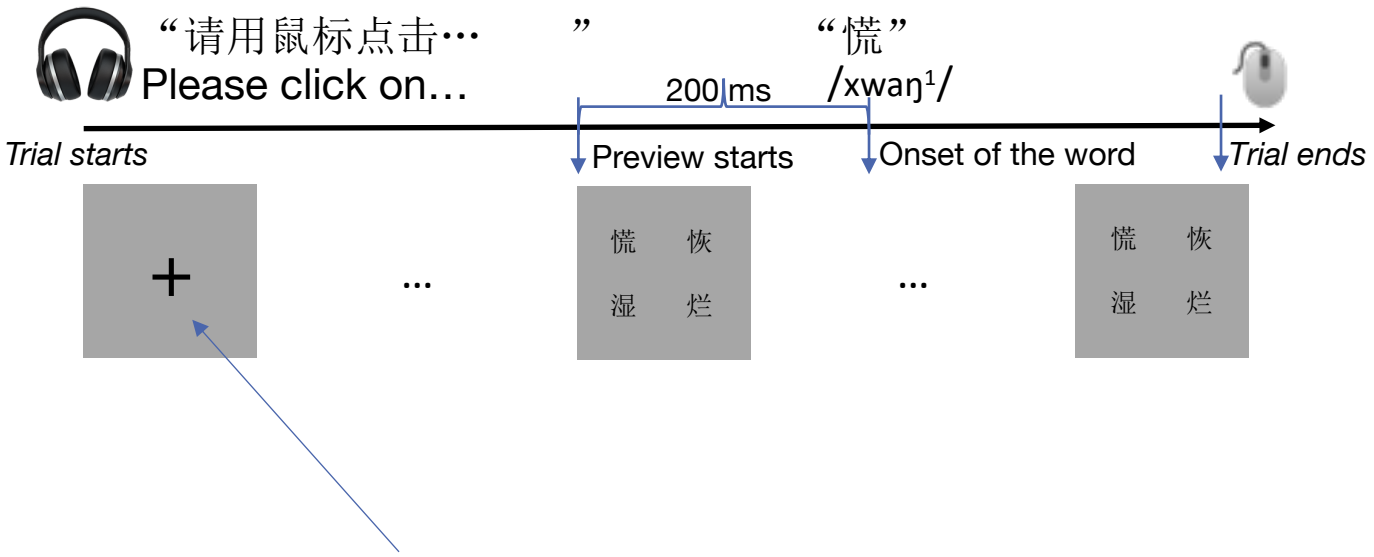
Example : “+”


Referent (e.g., "beaker")
Cohort (e.g., "beetle")
Rhyme (e.g., "speaker")
Unrelated (e.g., "carriage")
Fixation probability
Average target offset
Time since target onset (msec)


Target Word
Phonological Competitor
Distractor
Fixation Proportion
Time relative to onset of spoken target words

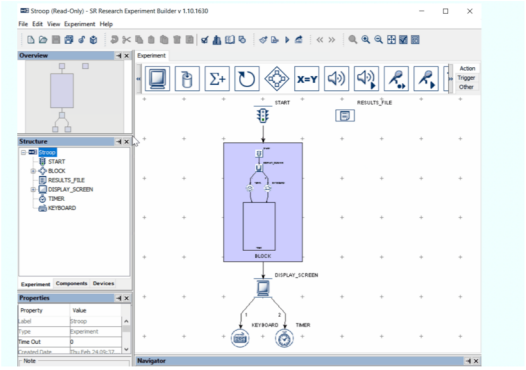

Prosody
Noun-focus
Verb-focus
Picture type
Prototypical target
Non-prototypical target
Distractors
Proportion of fixations
it looks like a ze- bra
Time relative to target word onset (ms)


Proportion of fixations
Time relative to target onset (ms)
Target word stress
SW
WS
Picture type
target
competitor
distractor


“请用鼠标点击…
Please click on…
”
“慌”
/xwaŋ¹/
200 ms
Trial starts
Preview starts
Onset of the word
Trial ends
+
…
慌 恢
湿 烂
…
慌 恢
湿 烂

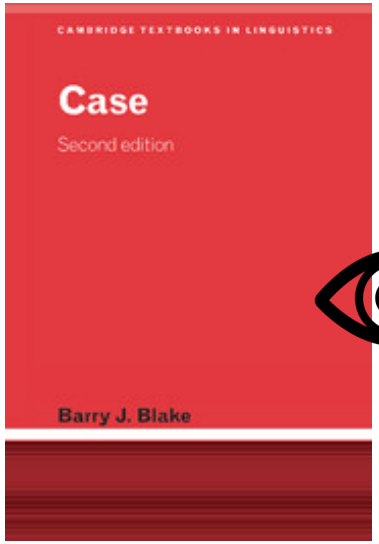
Example : “+”

- **“Unless there are other compelling reasons, we recommend against instructing participants to fixate a specific location at the start of a trial (e.g. by using a fixation cross). Maintaining fixation is resource-intensive. Moreover, asking participants to control their initial fixation can reduce the number of eye movements, with some participants maintaining fixation until before they begin to perform an action.”**



Style of this tutorial

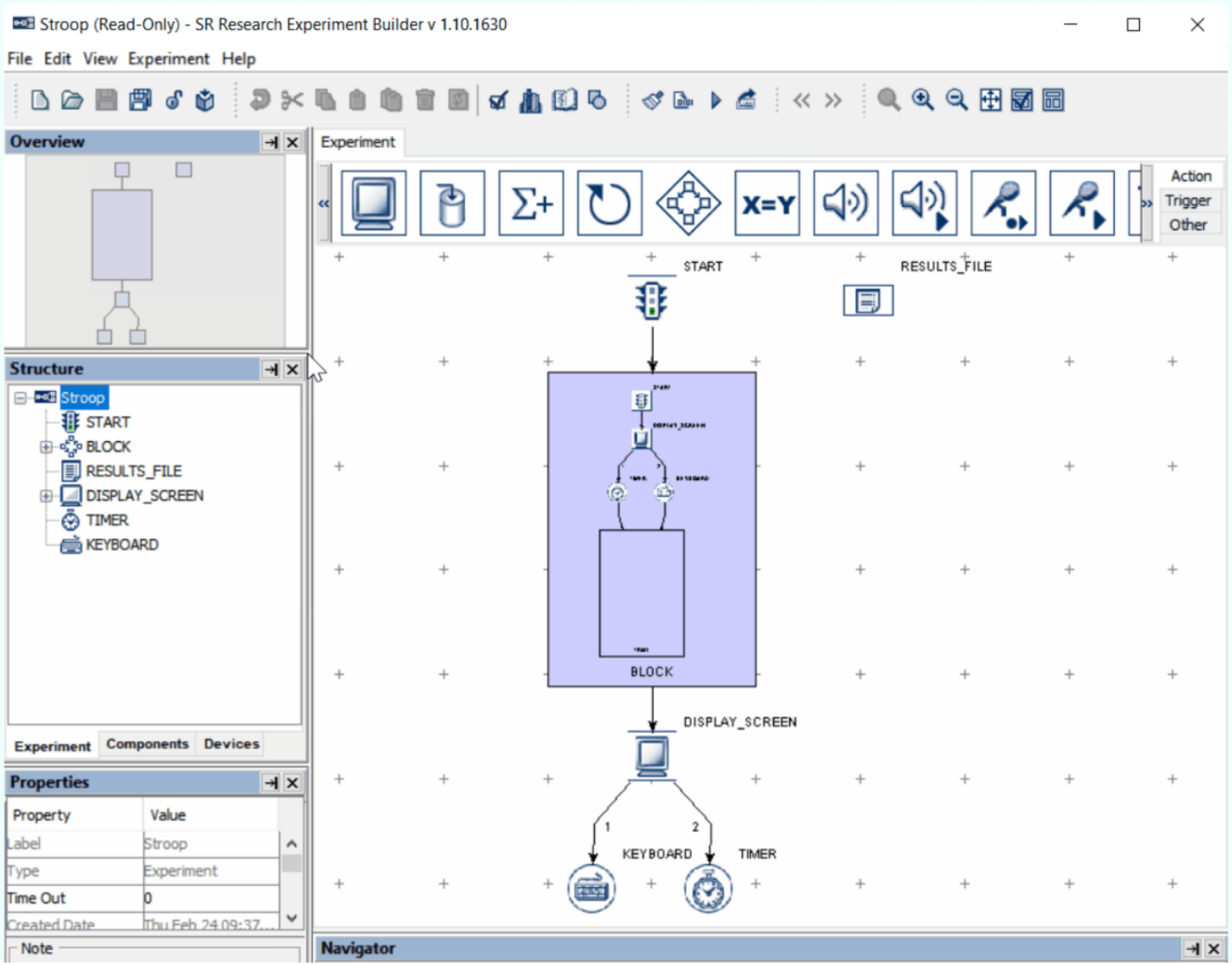
- **Two extremes…**



yuhang.xu@rochester.edu

Part 1

Getting the data


Stroop (Read-Only) - SR Research Experiment Builder v 1.10.1630
File Edit View Experiment Help
Overview
Structure
Stroop
START
BLOCK
RESULTS_FILE
DISPLAY_SCREEN
TIMER
KEYBOARD
Experiment
Components
Devices
Properties
Property
Value
Label
Stroop
Type
Experiment
Time Out
0
Created Date
Thu Feb 24 09:37...
Note
Experiment
Action
Trigger
Other
START
RESULTS_FILE
BLOCK
DISPLAY_SCREEN
1
2
KEYBOARD
TIMER
Navigator

Experiment Builder

- **A graphical programming environment for building experiments.**
- **Conceptually similar to E-Prime.**
- **Tightly integrated with Eyelink eye trackers.**

- **Files->Examples**

- **EB Project Checklist**
  - Chapter 16 of *Experiment Builder User Manual*
  - Read it. Follow it.
-

Why EB?

- **Xbuilder**
  - In house software
  - Tanenhause Lab only

- **Experiment Builder** 💓
  - SR-Research ”build-in” with the eyelink eye-tracker
  - Eyelink+EB
- **MATLAB**
  - MATLAB along with the Psychophysics Toolbox
  - http://sarahbrownschmidt.com/lab/
  - Sarah Brown-Schmid. Vanderbilt University.
  - Eyelink+ MATLAB
- **E-Prime**
  - The University of Tokyo
  - 伊藤たかね & 陳姿因
  - Tobi+E-Prime

- **Other software with other eye-trackers?**
  - SMI?

Experiment Builder

- **Workflow**
- **Components**
- **Useful tips**

SR Research Support

- **Main web site: https://www.sr-research.com**
- **Support Forum: https://www.sr-support.com/forum**
- **You'll need to create an account.**

Experiment Builder

- **Demo**
- **https://www.sr-support.com/forums/**
- **VWP Template**

Example E-Builder projects

+ SUBSCRIBE

TOPICS | LATEST ACTIVITY | MY SUBSCRIPTIONS | PHOTOS

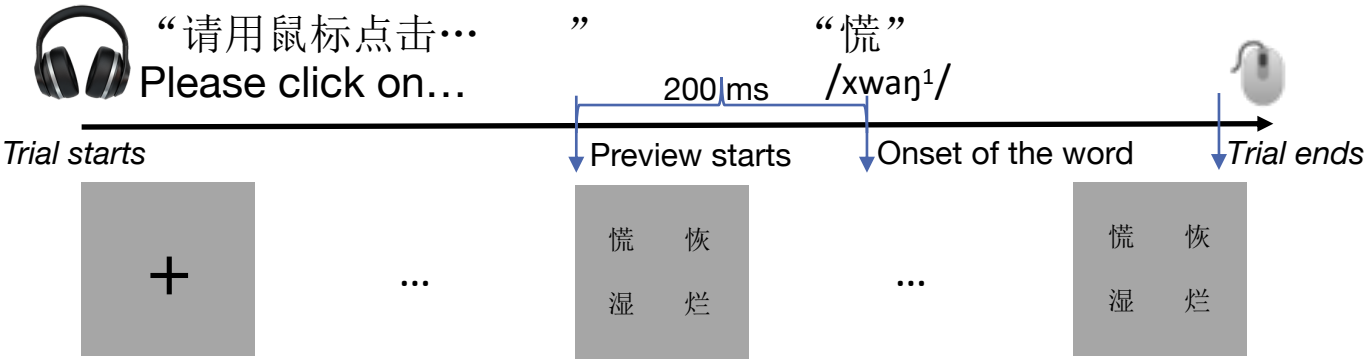
Page 1 of 7 | Filter

Topics	Statistics	Last Post
Closed, Sticky: Experiment Builder Examples Table of Contents Started by SR-William, 06-19-2017, 07:41 PM	0 responses 719 views 0 likes	by SR-William 06-19-2017, 07:41 PM
Closed, Sticky: Visual World Paradigm Started by jiye, 12-05-2006, 12:19 PM	4 responses 4,583 views 0 likes	by marcus 03-17-2017, 07:04 PM
Sticky: Collecting Participant Information Started by jiye, 06-20-2008, 06:09 PM	0 responses 1,375 views 0 likes	by jiye 06-20-2008, 06:09 PM
Closed, Sticky: Response Checking and RT calculation (with keyboard/EyeLink button box) Started by jiye, 10-08-2007, 02:41 PM	1 response 1,095 views 0 likes	by jiye 10-08-2007, 02:43 PM
Manipulating variables with Python functions without Custom Class Started by SR-William, 04-06-2018, 09:33 PM	0 responses 38 views	by SR-William 04-06-2018, 09:33 PM

Experiment Builder

- **Let's go to the Experiment Builder**

- **Listen carefully…**
- **No notes on this part…**



Next,

- **Extracting relevant sample reports from raw eye-tracking data**

Data Viewer

- **“sample report”**

Sample report features to include

RECORDING_SESSION_LABEL
DATA_FILE
IP_DURATION
IP_END_TIME
IP_LABEL
IP_START_TIME
LEFT_GAZE_X
LEFT_GAZE_Y
LEFT_INTEREST_AREA_ID
RIGHT_GAZE_X
RIGHT_GAZE_Y
RIGHT_INTEREST_AREA_ID
SAMPLE_INDEX
SAMPLE_MESSAGE
TIMESTAMP
TRIAL_INDEX
TRIAL_LABEL
TRIAL_START_TIME
all of your experimental columns

Part 2

Analyzing the data

mixed effects models

Motivation for mixed effects models

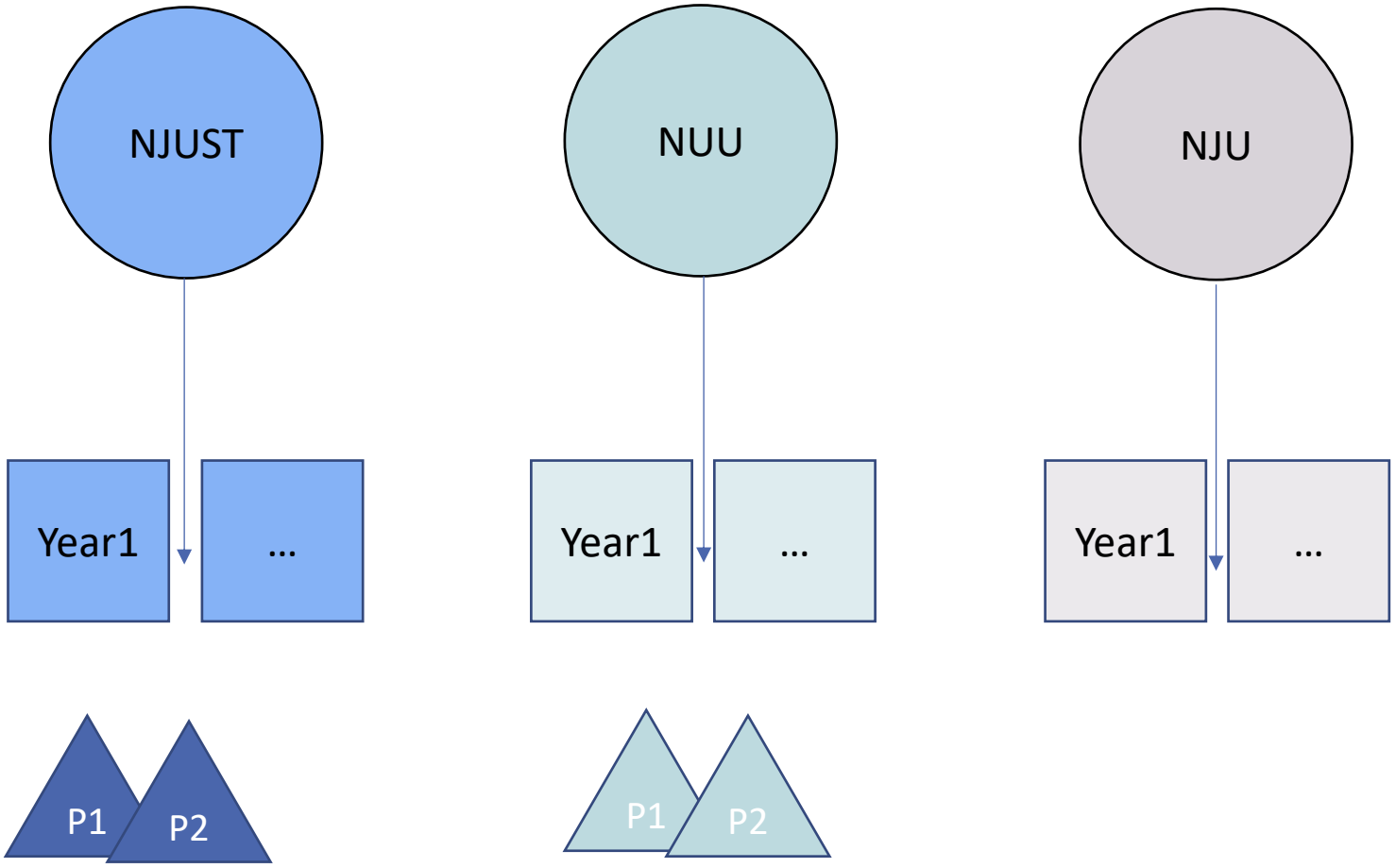
- **Problems with ANOVA**
  - Multiple random effects (subjects, items)
  - Categorical outcomes
  - Continuous predictors

Motivation for mixed effects models

- **Inferential statistics you may be familiar with:**
  - ANOVA
    - Regression
    - Correlation
- **All of these methods involve random sampling out of a larger population**

Motivation for mixed effects models

- **Standard assumption: All observations are *independent***
  - Subject 1's score doesn't tell Subject 1 us anything Subject 2 about Subject 2's
- **Independence assumption is fair if we randomly sample 1 person at a time**
  - e.g., you recruit 100 undergrads from the NJUST Subject Pool
- **But many research designs involve more complex sampling procedures**



One way to describe what's going on here is that there are several levels of sampling, each **nested** inside each other

Nested Random Effects

- **Two challenges:**
  - Statistically, we need to take account for this non independence (similarity)
    - Even a small amount of non-independence can lead to spurious findings (Quené & van den Bergh, 2008
  - We might want to *characterize* differences at each level!

Crossed Random Effects

- **Items you used**

- **Experimental / research materials are often *sampled* out of population**

Appendix

Test Items

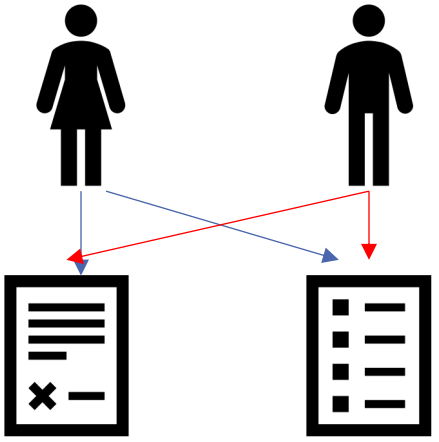
The temporal-manipulation syllable in each stimulus context is italicized. The words beginning with strong–weak (SW) and weak–strong (WS) stress patterns are in bold.

Many city people *gave* those two . . .	
SW: **alligators** a wide berth.	WS: **alpacas** carrots and hay.
Jenny's helpful neighbor *saw* her two . . .	
SW: **bunnies** escape from their cage.	WS: **bananas** in her driveway.
The teacher's brother *took* that . . .	
SW: **canteloupe** out of the fridge.	WS: **canteen** into the desert.
Mr. Johnson *gave* his . . .	
SW: **cashews** to his students.	WS: **cashier** the wrong credit card.
The seven students rarely *used* their . . .	
SW: **compasses** to find their camp.	WS: **computers** during their lectures.
Betsy never *found* that . . .	
SW: **corset** very comfortable.	WS: **corsage** she was looking for.
Many people *saw* that . . .	
SW: **eagle** around campus.	WS: **eclipse** of the sun.

(*Appendix continues*)

Crossed Random Effects

- **we are sampling two things—subjects and items**
- ***each* subject gets *each* item**
  - Crossed random effects
- **Still, problem is that we have multiple random effects (things being sampled)**



OLD ANOVA solution: Do 2 analyses

Note: not real data

Subject	Control	Primed
Scott	1305	1122
Tessa	1100	905
Natasha	950	900
Chuck	1070	988

SUBJECT ANALYSIS

$F_1(1,3) = 18.31, p < .05$

Item	Control	Primed
Knight	1100	883
Monkey	979	930
Bird	1114	912
Vacation	1094	830
Pirate	1251	939

ITEM ANALYSIS

$F_2(1,4) = 22.45, p < .01$

- **F1 F2**

- ***Subjects analysis*:**
  - Compare each subject (averaging over all of the items)
  - Does the effect generalize across subjects?
- ***Items analysis:***
  - Compare each item (averaging over all of the subjects)
  - Does the effect generalize across items?

- **Problem: We now have 2 different sets of results.**
  - Might conflict!
  - Possible to combine them with *min F'*, but not widely used

Problem 2: Categorical Data

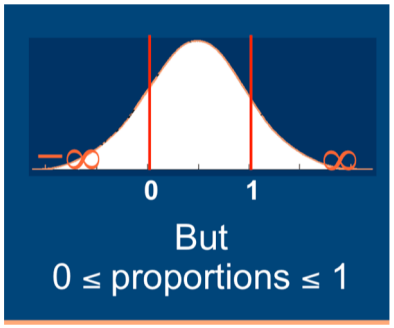
- **ANOVA assumes our response is *continuous***
  - *E.g., RT= 2890ms*

- **But, we often want to look at *categorical* data**

Region fixated in eye-tracking experiment

Problem 2: Categorical Data

- **Analyze proportions**
  - Maybe with some transformation
  - (e.g., arcsine, logit)



- **Violates assumptions of ANOVA**
  - ANOVA assumes normal distribution, which has infinite tails
  - But proportions are clearly *bounded*
  - Model could predict impossible values like 150%
- **Can lead to:**
  - - Spurious effects
  - (Type I error)
  - - Missing a true effect
  - (Type II error)

Problem 3: Continuous Predictors

- **Many interesting independent variables vary continuously**
  - e.g., Word frequency; Second language proficiency; reading skill; SAT scores
- **ANOVAs require division into categories**
  - e.g., median split
- **Problem:**
  - Can only ask “is there a difference?”, not *form* of relationship
  - Loss of statistical power (Cohen, 1983)

Mixed Effects Models to the Rescue!

- **Biggest contribution of mixed-effects models is to incorporate *multiple random effects* into the *same analysis***

- **ANOVA**: Unit of analysis is ***cell mean***

SUBJECT	PRIMED	CONTROL
Scott	960	1214
Ben	913	1003

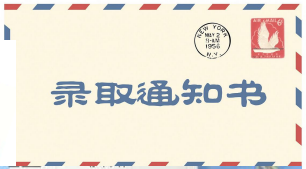
- **Mixed effects models**: Unit of analysis is ***individual trial***!

RT	PRIME?	SUBJECT	ITEM
1224	Yes	Scott	eagle
683	No	Scott	penguin
432	Yes	Scott	pitohui
892	No	Scott	robin
1028	Yes	Ben	penguin

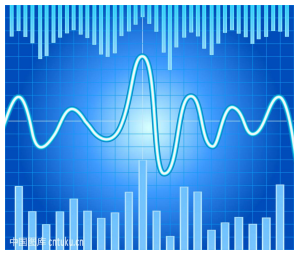
Mixed Effects Models

- **Model outcome using regression-like approach**
  - Look at *individual trials/observations* (not means)
- **In a regression, easy to include independent variables that are continuous**
- **Link functions allow us to relate model to DV that *isn't* normally distributed**
- `lmer( RT ~ Class+Frequency + (1|Subject) + (1|Item), lexdec )`
-

Y



Fixed effects



Random

How about VWP data

- **outcome**
  - 0,1 look or not
- **Predictor**
  - Time-window
  - *Continuous OR Categorical*

eyetrackingR

- *eyetrackingR* was created by Jacob Dink and Brock Ferguson. Jacob and Brock received their PhD's in Cognitive Science from Northwestern University and are now co-founders of Strong Analytics, a data science consulting and development firm.

- **Each eyetrackingR vignette uses the *eyetrackingR* package to analyze real data from a simple 2-alternative forced choice (2AFC) word recognition task administered to 19- and 24-month-olds. On each trial, infants were shown a picture of an animate object (e.g., a horse) and an inanimate object (e.g., a spoon). After inspecting the images, they disappeared and they heard a label referring to one of them (e.g., "The horse is nearby!"). Finally, the objects re-appeared on the screen and they were prompted to look at the target (e.g., "Look at the horse!").**
-

eyetrackingR-Demo

Q & A

- yuhang.xu@rochester.edu