peekbank

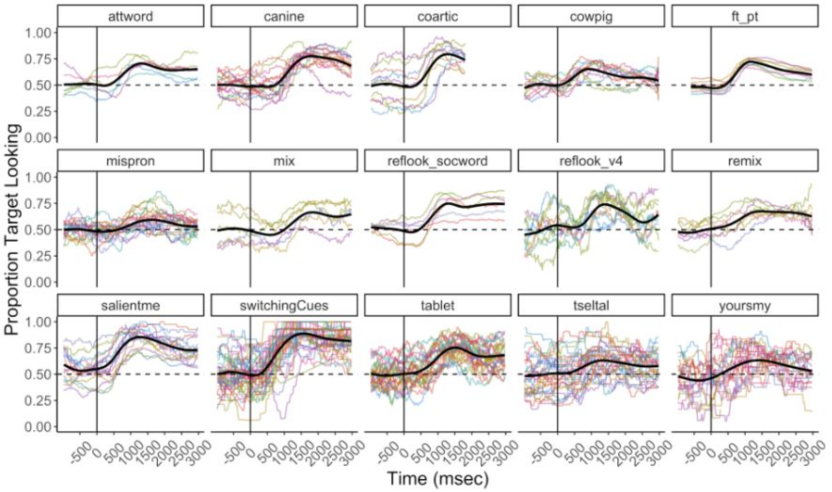
Demo

A flexible and reproducible interface to developmental eyetracking datasets

What is peekbank?

peekbank is a flexible and reproducible interface to developmental eyetracking datasets.

The Peekbank project is an open database storing eye-tracking datasets on children's word recognition in a well-documented, easily accessible, tabular format. It also provides processing tools for standardizing eye-tracking data across data sources (`peekds` R package), interfaces for accessing the database (`peekbankr` R package), and applications for visualizing the data ([Peekbank Shiny App]).



Roadmap

1. Peekbank
2. Peekbank website
3. PeekbankR

Peekbank

- An open database of developmental eye-tracking studies on children's word recognition
- Contains datasets from looking-while-listening tasks across different labs
- Aims to address theoretical and methodological challenges in measuring vocabulary development
- Enables analysis at a large scale

Peekbank Framework

- Consists of three main components:
  - Processing raw experimental datasets
  - Populating a relational database
  - Providing an interface to the database
- Uses a common, tidy format to standardize eye-tracking data across studies
- Open source and under active development on GitHub

Interactive Data Visualization with Peekbank Shiny App

- Web-based tool for visualizing Peekbank data: https://peekbank-shiny.com/
- Visualizations can be filtered by age, dataset, condition, etc.
- Useful for quick exploration before custom analyses

Interactive Data Visualization with Peekbank Shiny App

data from the Peekbank database. Specifically, users can visualize:

1. the *time course of looking data* in a profile plot depicting infant target looking across trial time
2. *overall accuracy*, defined as the proportion target looking within a specified analysis window
3. *reaction times* in response to a target label, defined as how quickly participants shift fixation to the target image on trials in which they were fixating on the distractor image at onset of the target label
4. an *onset-contingent plot*, which shows the time course of participant looking as a function of their look location at the onset of the target label

Zettersten et al. (2023): https://link.springer.com/article/10.3758/s13428-022-01906-4

Interactive Data Visualization with Peekbank Shiny App

Visualizations include:

Data type:

- Target looking time
- Overall accuracy
- Reaction time

Graph type:

- Time-course plot
- Histogram

try it out!

Current Datasets:

https://langcog.github.io/peekbank-website/docs/contributors/

Accessing PeekbankR

1. Install peekbankr: https://github.com/langcog/peekbankr
2.

```
> # install.packages("remotes") # uncomment when necessary
> remotes::install_github("langcog/peekbankr")
```

Peekbank Framework

- Data schema includes linked tables tracking different types of information in a relational database:
- There are several different `get_` functions that you can use to extract different types of data from the peekbank-db:
  - get_datasets()
  - get_subjects()
  - get_administrations()
  - get_trials()
  - get_stimuli()
  - get_aoi_region_sets()
  - get_aoi_timepoints()
  - get_xy_timepoints()

https://langcog.github.io/peekbank-website/docs/data-access/

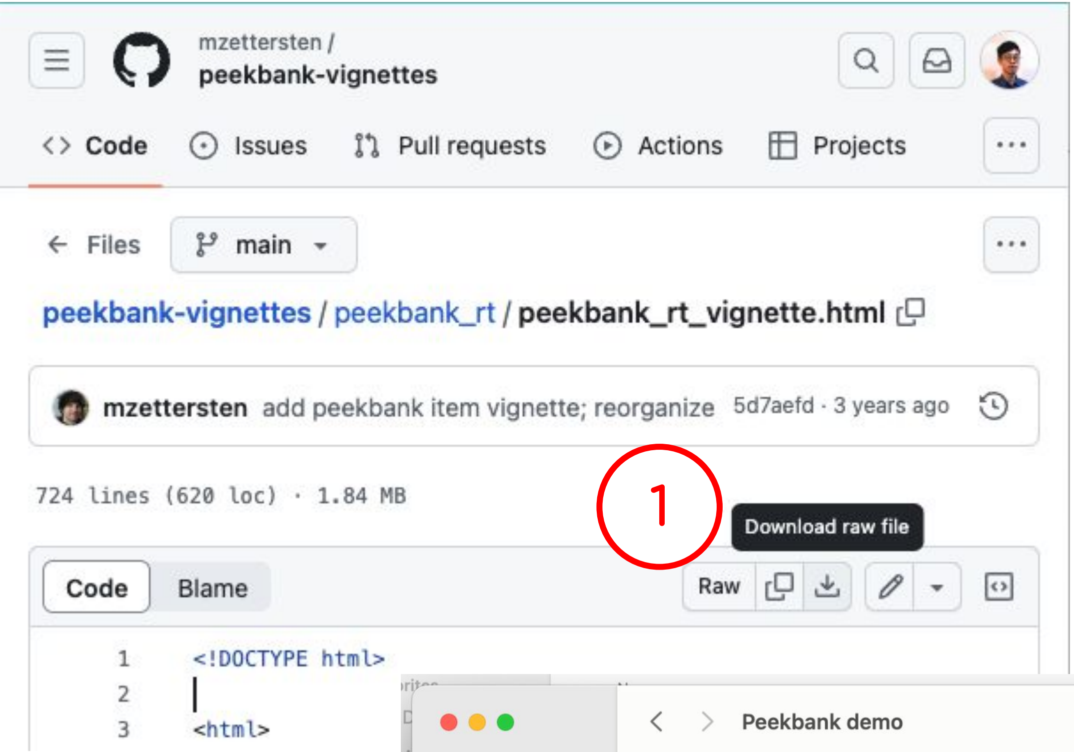
Use Case 1: Effect size

- To what extent can we estimate the effect size of above-chance looking at the target?
- How does it differ across development (e.g., age)?

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

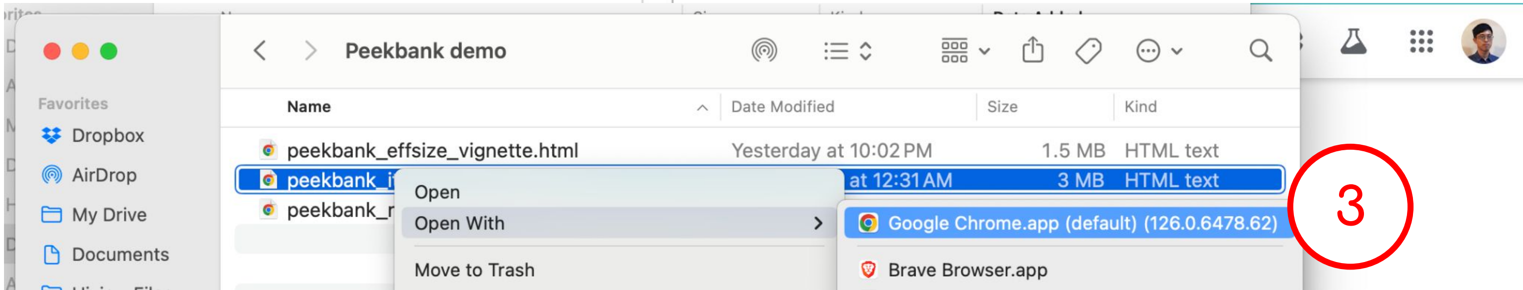
Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html



2

select where to save you file



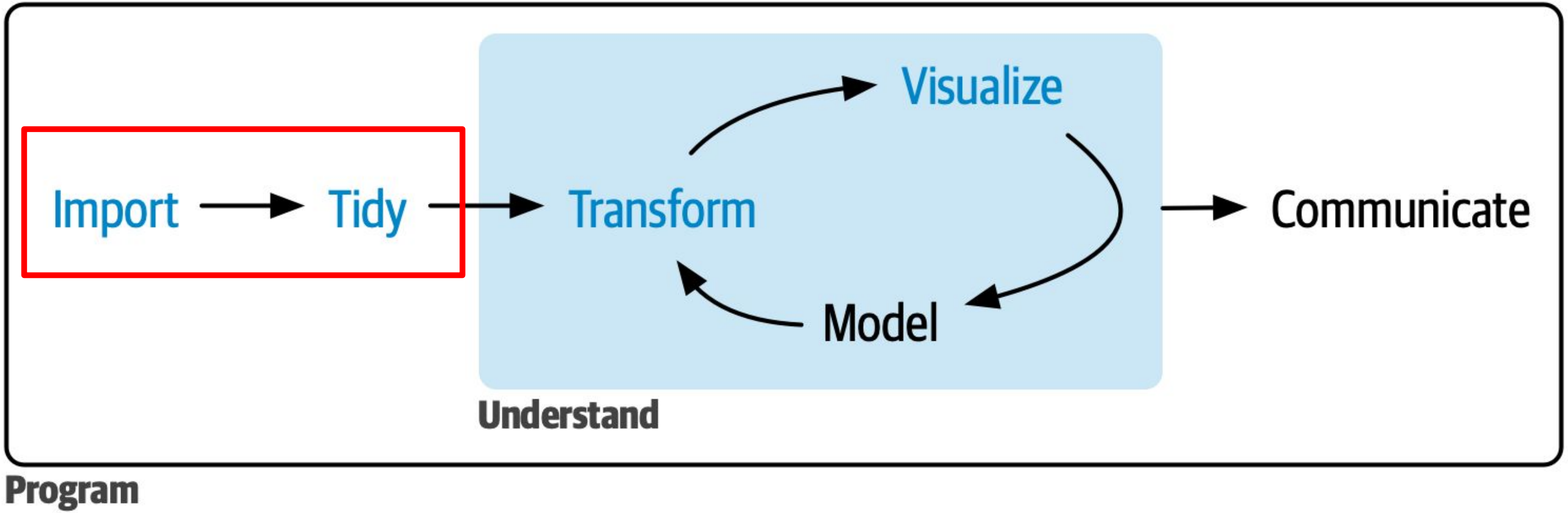
Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

Preliminaries

```
> library(peekbankr)
> library(tidyverse)
> library(here)
> library(lme4)
> library(lmerTest)
> library(effectsize)
> library(metafor)
```

Data Analysis Pipeline



Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

get all of the tables you need with **get_**

```
> datasets <- get_datasets()
> administrations <- get_administrations()
> subjects <- get_subjects()
> aoi_timepoints <- get_aoi_timepoints()
> stimuli <- get_stimuli()
> trial_types <- get_trial_types()
> trials <- get_trials()
```

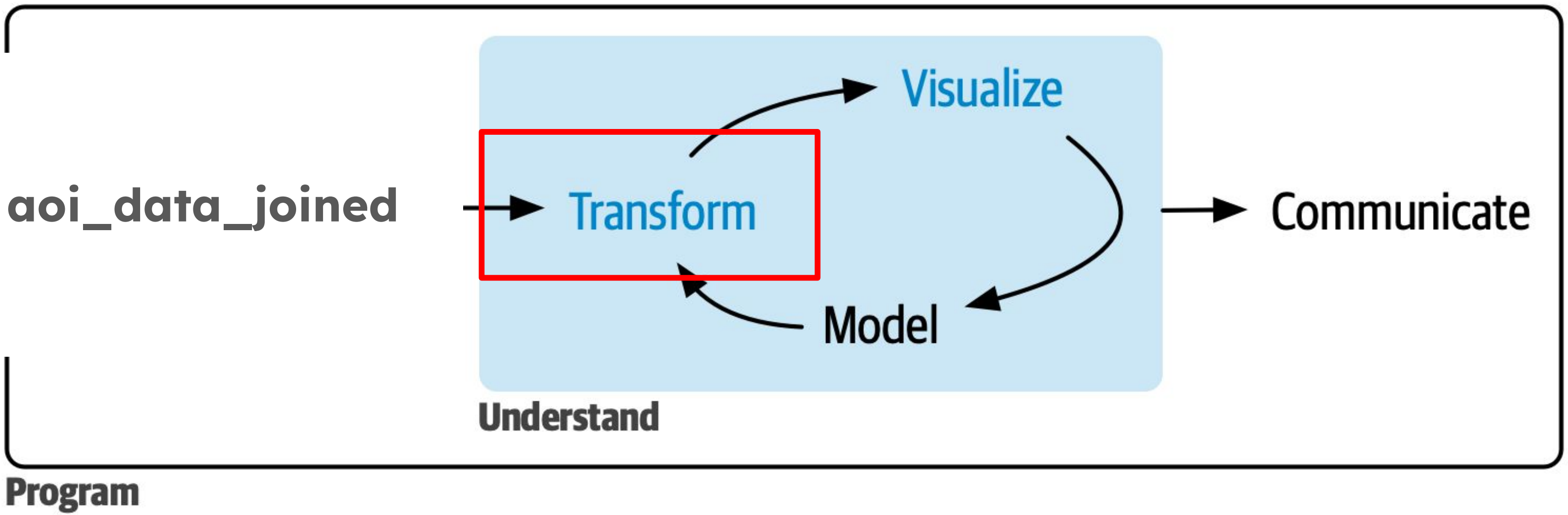
Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

join all data

```
> aoi_data_joined <- aoi_timepoints %>%
      right_join(administrations) %>%
      right_join(subjects) %>%
      right_join(trials) %>%
      right_join(trial_types) %>%
      right_join(datasets) %>%
      mutate(stimulus_id = target_id) %>% #add a second join for
  distractor info
      right_join(stimuli)
```

Data Analysis Pipeline



Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

Average trial-level data

Goal here is to average looking performance for each trial across a critical window (t_min and t_max). We also set a threshold for how much looking data must be included in order for the trial to merit inclusion.

```
#### PARAMETERS TO SET ####
#critical window dimensions roughly consistent with e.g., Swingley & Aslin, 2002
t_min <- 300
t_max <- 2000
#proportion missing trials threshold (any trial in which over half of the critical window missing is lo
oking data is excluded )
max_prop_missing <- 0.5
#min/max age (in mos)
min_age <- 9
max_age <- 27
#age bin size (number of months per bin)
age_bin_size <- 6


by_trial_means <- aoi_data_joined %>%
  #restrict to english datasets (this is just because there are so few non-English datasets atm)
  filter(native_language == "eng") %>%
  #restrict age range
  filter(age > min_age, age <= max_age) %>%
  # familiar target items only %>%
  filter(stimulus_novelty == "familiar") %>%
  #window of analysis
  filter(t_norm >= t_min, t_norm <= t_max) %>%
  #bin ages (can adjust size of age bins here)
  mutate(age_binned = cut(age, seq(min_age,max_age,age_bin_size))) %>%
  group_by(dataset_name,subject_id, trial_id, english_stimulus_label,
           age, age_binned) %>%
  summarise(prop_target_looking = sum(aoi == "target", na.rm = TRUE) /
              (sum(aoi == "target", na.rm=TRUE) +
                 sum(aoi=="distractor", na.rm=TRUE)),
            prop_missing = mean(aoi %in% c("missing","other"), na.rm = TRUE)) %>%
  #remove trials with insufficient looking to target or distractor
  filter(prop_missing<=max_prop_missing)
```

- Setting parameters
- Specified a time window
- Average (summarise) trial performance
- Only trials with sufficient looking data is included

Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

Average within subjects, by-dataset

One could consider excluding participants based on the number of trials a participant contributes overall here.

```
by_subj_means <- by_trial_means %>%
  group_by(dataset_name,subject_id,age_binned) %>%
  summarise(
    trial_num=n(),
    avg_target_looking =mean(prop_target_looking,na.rm=TRUE)
  )
```

- enable exclusion of participants based on the number of trials they contribute

Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

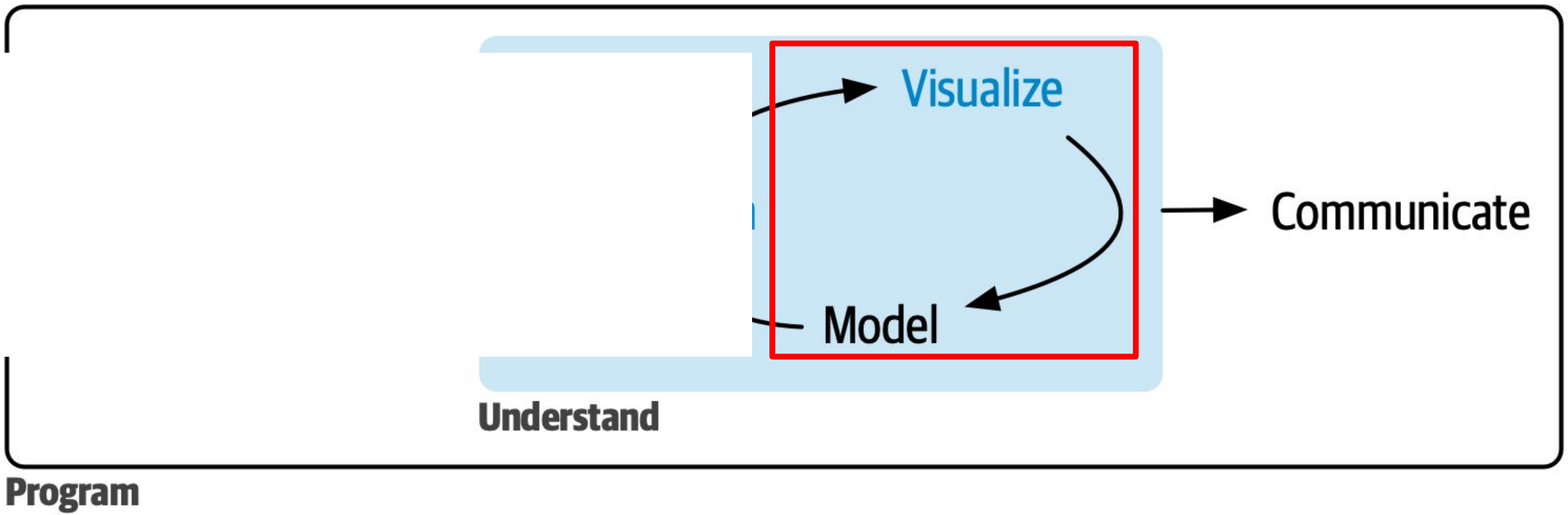
Average across subjects - by dataset and age bin, by dataset, and by age bin

```
#make this a function so we can use map
compute_cohens_d <- function(current_data) {
  temp <- cohens_d(avg_target_looking-0.5~1,data=current_data)
  temp
}

by_dataset_age_means <- by_subj_means %>%
  group_by(dataset_name, age_binned) %>%
  mutate(subj_n=n()) %>%
  group_by(dataset_name, age_binned,subj_n) %>%
  #filter to at least 5 subjects
  filter(subj_n>=5) %>%
  mutate(
    mean_target = mean(avg_target_looking),
    sd_target=sd(avg_target_looking),
    d_target = (mean_target-0.5)/sd_target
  ) %>%
  group_by(dataset_name, age_binned,subj_n,mean_target,sd_target,d_target) %>%
  nest() %>%
  mutate(cohens_d = purrr::map(data,compute_cohens_d)) %>%
  select(-data) %>%
  unnest(cols=c(cohens_d)) %>%
  ungroup() %>%
  mutate(
    chance=0.5
  )
```

- Average across subjects
- Here we focus on **by data set and age**
- Compute effect size (Cohen's d)

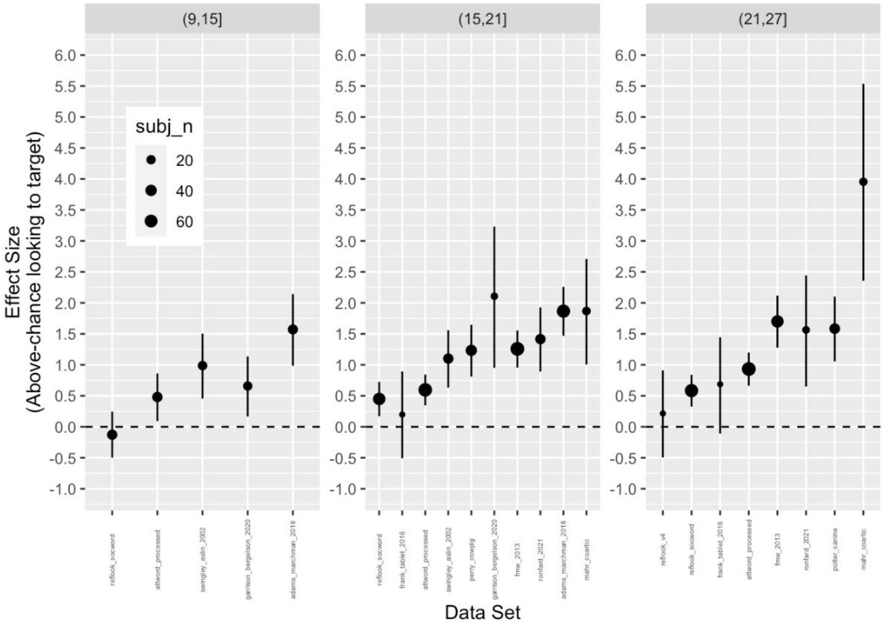
Data Analysis Pipeline



Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

```r
ggplot(by_dataset_age_means,aes(reorder(dataset_name,Cohens_d,mean),Cohens_d))+
  geom_hline(yintercept=0,linetype="dashed")+
  geom_point(aes(size=subj_n))+
  geom_errorbar(aes(ymin=CI_low,ymax=CI_high),width=0)+
  scale_size(range = c(1, 3))+
  #geom_point()+
  theme(axis.text.x=element_text(angle=90,size=4,vjust=0.5))+
  #theme(legend.position="none")+
  xlab("Data Set")+
  ylab("Effect Size\n(Above-chance looking to target)")+
  facet_wrap(~age_binned,nrow=1,scales = "free")+
  theme(legend.position=c(0.1,0.7))+
  scale_y_continuous(breaks=seq(-1,6,0.5),limits=c(-1,6))
```



- Plotting effect sizes by dataset and age

Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

Meta-analysis of specific age group

9-15-month olds | 15-21-month olds | 21-27-month olds

```
#using the metafor package
#this is a little hacky, in order to get the effect size for a one-sample test against chance
effect_sizes_9_15 <- escalc(measure="SMD",m1i=mean_target,m2i=chance,sd1i=sd_target,sd2i=sd_target,n1i=
subj_n,n2i=subj_n,data=filter(by_dataset_age_means,age_binned=="(9,15]"),slab=dataset_name)
meta_model <- rma(yi,vi,data=effect_sizes_9_15)
meta_model
```

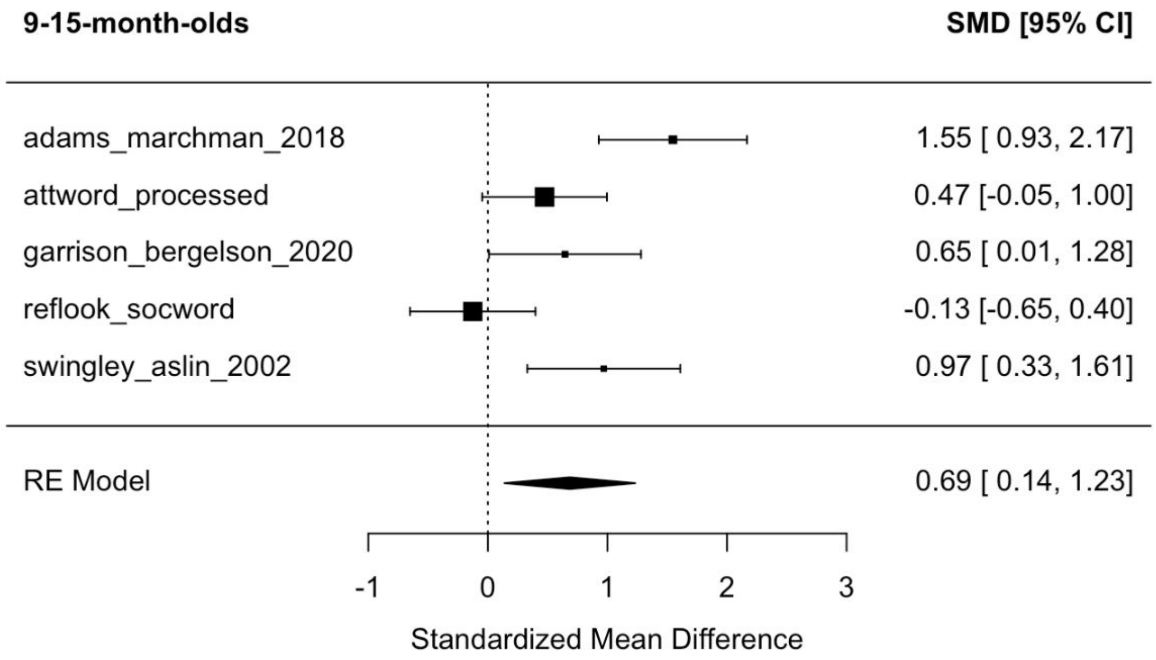
```
##
## Random-Effects Model (k = 5; tau^2 estimator: REML)
##
## tau^2 (estimated amount of total heterogeneity): 0.2998 (SE = 0.2757)
## tau (square root of estimated tau^2 value):      0.5475
## I^2 (total heterogeneity / total variability):   77.20%
## H^2 (total variability / sampling variability):  4.39
##
## Test for Heterogeneity:
## Q(df = 4) = 17.8439, p-val = 0.0013
##
## Model Results:
##
## estimate      se    zval    pval   ci.lb   ci.ub
##   0.6853  0.2793  2.4541  0.0141  0.1380  1.2327  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Use Case 1: Effect size

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_effsize/peekbank_effsize_vignette.html

Forest Plot

```
forest(meta_model,header="9-15-month-olds")
```



Use Case 2: Item level analysis

- Which specific words or semantic categories show the clearest developmental progression across different ages and language datasets?
- Can these developmentally sensitive words be used as markers to assess a child's language abilities at the individual level?

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_items/peekbank_item_vignette.html

Use Case 2: Item level analysis

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_items/peekbank_item_vignette.html

```
#compute baseline looking (for baseline-corrected means)
by_trial_baseline <- aoi_data_joined %>%
  #restrict to english datasets (this is just because there are so few
non-English datasets atm)
  filter(native_language == "eng") %>%
  #restrict age range
  filter(age > 12, age <= 60) %>%
  # familiar target items only %>%
  filter(stimulus_novelty == "familiar") %>%
  #window of analysis
  filter(t_norm >= baseline_window[1], t_norm <= baseline_window[2]) %
>%
  #bin ages (can adjust size of age bins here)
  mutate(age_binned = cut(age, seq(12,60,age_bin_size))) %>%
  rename(target_label = english_stimulus_label) %>%
  group_by(dataset_name,subject_id, trial_id, target_label,
           age, age_binned) %>%
  summarise(
   baseline_n=n(),
   baseline_ms=baseline_n*25,
    baseline_looking = sum(aoi == "target", na.rm = TRUE) /
              (sum(aoi == "target", na.rm=TRUE) +
                 sum(aoi=="distractor", na.rm=TRUE)),
           prop_baseline_missing = mean(aoi %in% c("missing","othe
r"), na.rm = TRUE)) %>%
  #remove trials with insufficient looking to target or distractor
  filter(prop_baseline_missing<=max_prop_missing& baseline_ms>=500)

#combine
by_trial_target_means <- by_trial_means %>%
  left_join(by_trial_baseline) %>%
  mutate(corrected_target_looking=prop_target_looking-baseline_lookin
g)
```

Compute looking during baseline

Adjust target looking by subtracting baseline

Use Case 2: Item level analysis

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_items/peekbank_item_vignette.html

Average within subjects, by-item and by-dataset

One could consider excluding participants based on the number of trials a participant contributes overall here.

```
by_subj_item_means <- by_trial_target_means %>%
  group_by(dataset_name,subject_id, target_label,
           age, age_binned) %>%
  summarise(
    trial_num=n(),
    avg_target_looking = mean(prop_target_looking,na.rm=TRUE),
    avg_corrected_target_looking=mean(corrected_target_looking,na.rm=T
RUE)
  )
```

- enable exclusion of participants based on the number of trials they contribute

Use Case 2: Item level analysis

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_items/peekbank_item_vignette.html

Average across subjects - by item, dataset and age bin

```
by_item_means <- by_subj_item_means %>%
  group_by(dataset_name, target_label,age_binned) %>%
  summarise(
    subj_n=n(),
    target_looking = mean(avg_target_looking,na.rm=TRUE),
    corrected_looking = mean(avg_corrected_target_looking,na.rm=TRUE)
  )
```

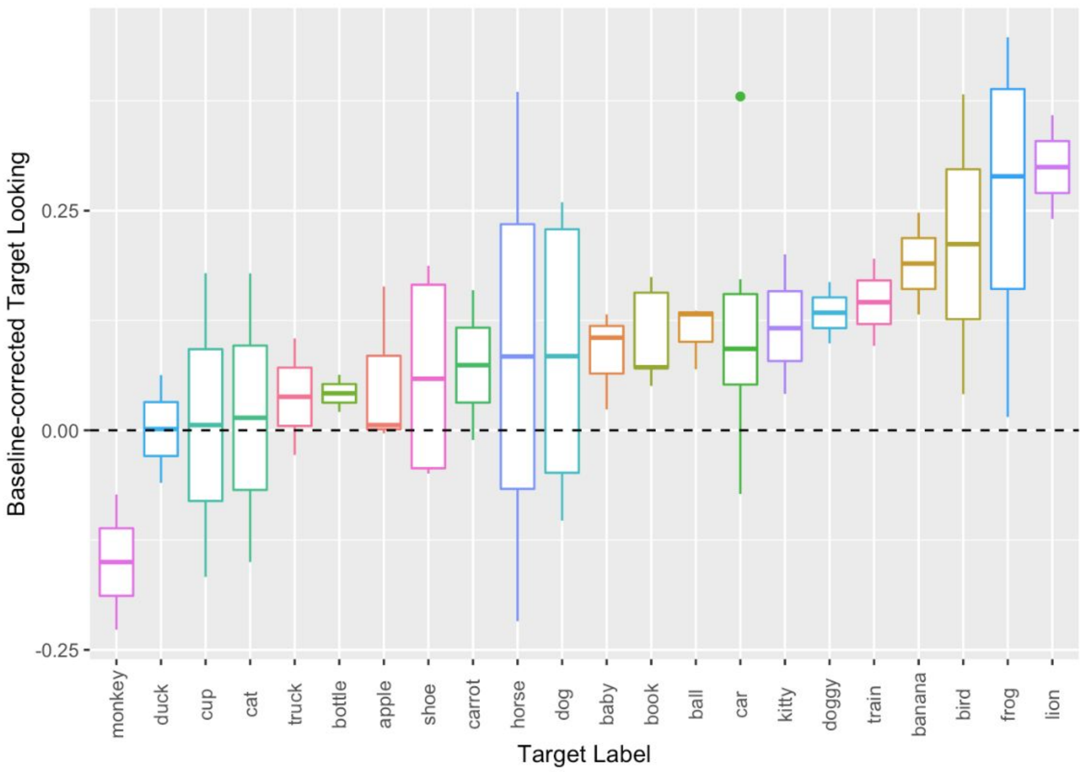
- focus on item level information

Use Case 2: Item level analysis

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_items/peekbank_item_vignette.html

Baseline-corrected target accuracy

```
ggplot(filter(by_item_means,age_binned=="(12,18]"&dataset_num>1&!is.na
(corrected_looking)),aes(reorder(target_label,corrected_looking,mean),
corrected_looking,color=target_label))+
  geom_boxplot()+
  #geom_point()+
  theme(legend.position="none")+
  theme(axis.text.x=element_text(angle=90,size=9,vjust=0.5))+
  xlab("Target Label")+
  ylab("Baseline-corrected Target Looking")+
  geom_hline(yintercept=0,linetype="dashed")
```

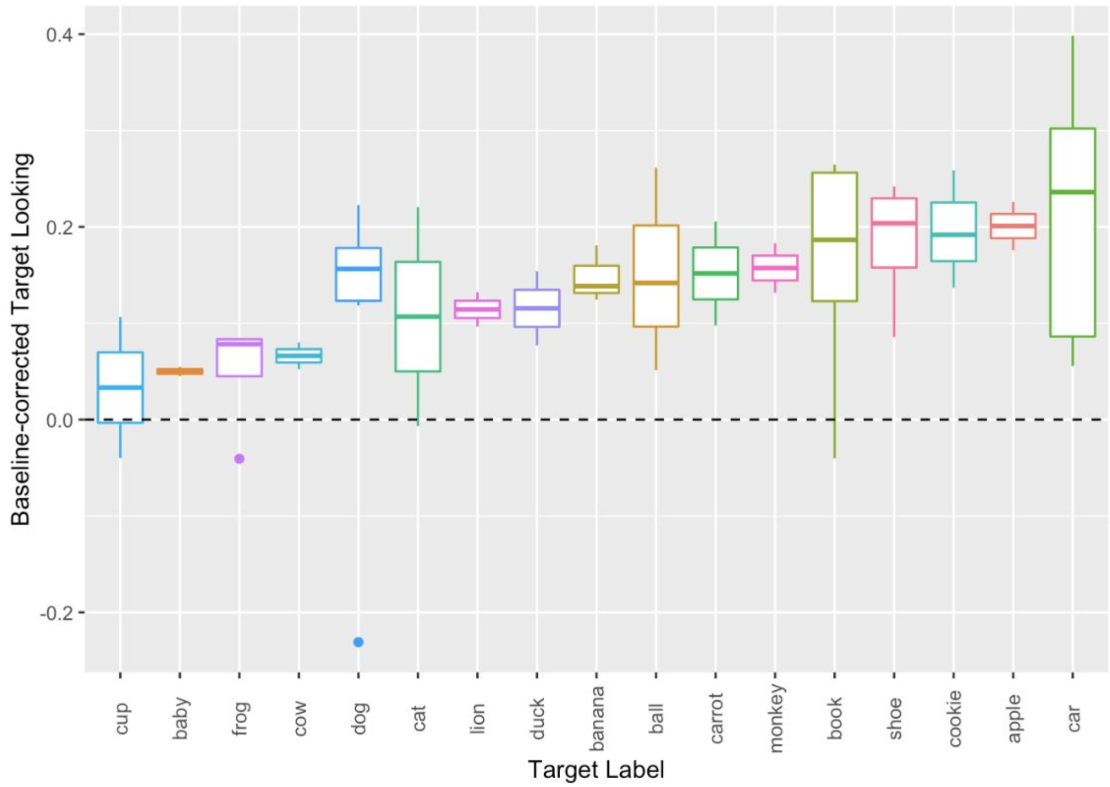


Use Case 2: Item level analysis

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_items/peekbank_item_vignette.html

Baseline-corrected target accuracy

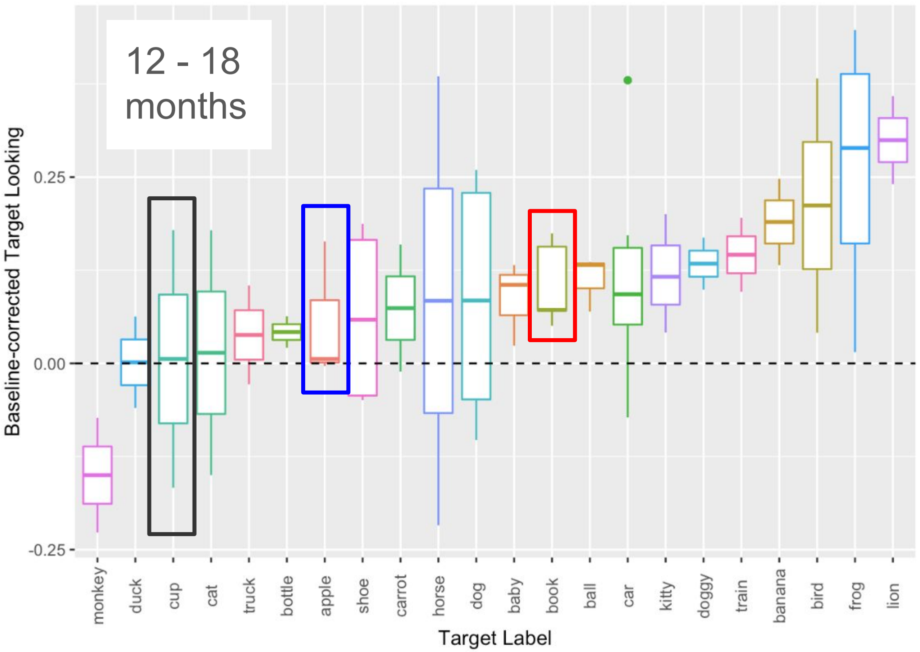
```
ggplot(filter(by_item_means,age_binned=="(18,24]"&dataset_num>1&!is.na
(corrected_looking)),aes(reorder(target_label,corrected_looking,mean),
corrected_looking,color=target_label))+
  geom_boxplot()+
  #geom_point()+
  theme(legend.position="none")+
  theme(axis.text.x=element_text(angle=90,size=9,vjust=0.5))+
  xlab("Target Label")+
  ylab("Baseline-corrected Target Looking")+
  geom_hline(yintercept=0,linetype="dashed")
```

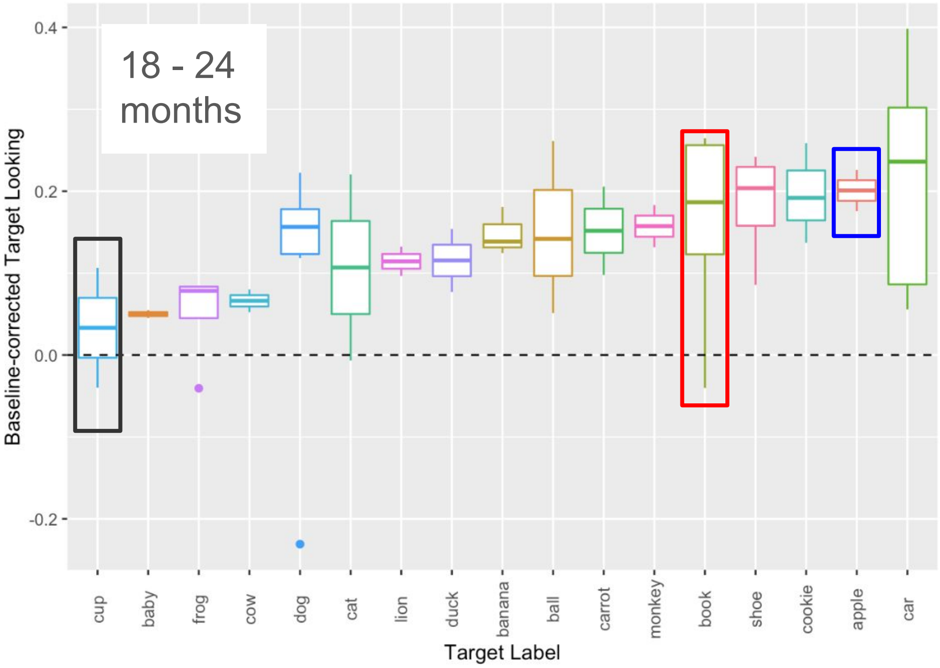


Use Case 2: Item level analysis

https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_items/peekbank_item_vignette.html

For example:





Use Case 3: Reaction Time

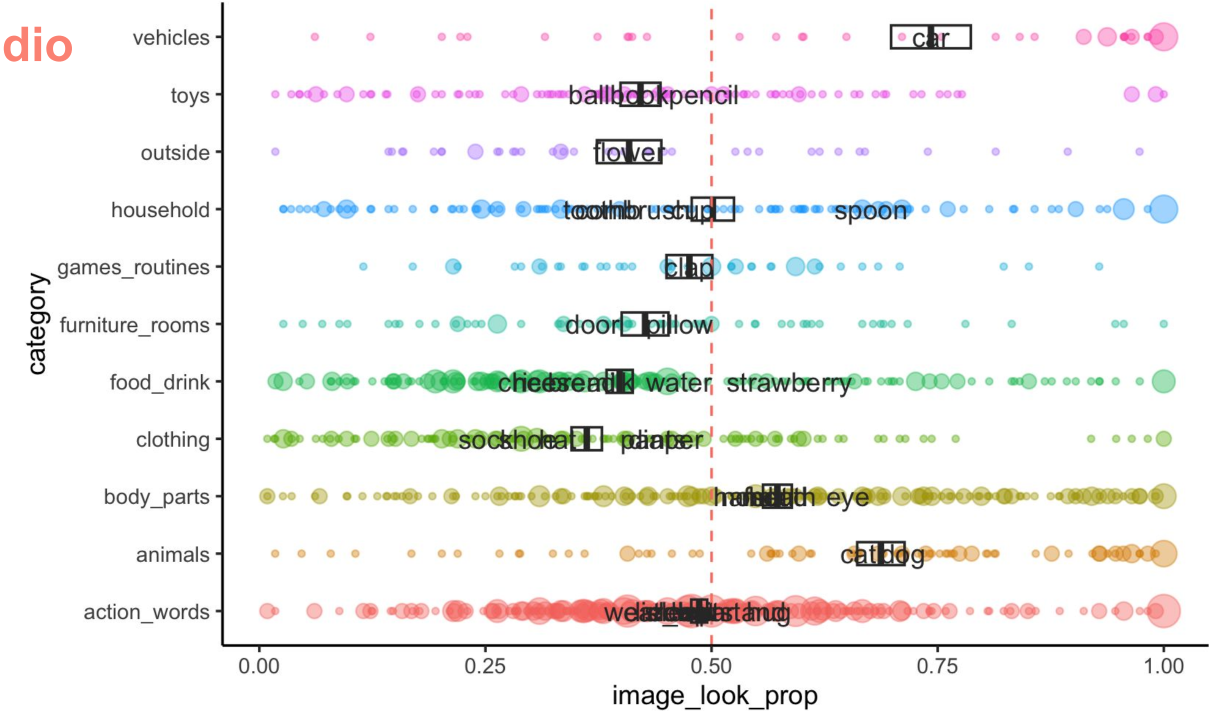
https://github.com/mzettersten/peekbank-vignettes/blob/main/peekbank_rt/peekbank_rt_vignette.html

Other potential use case

- for example, current effort in our lab

Perceptual biases

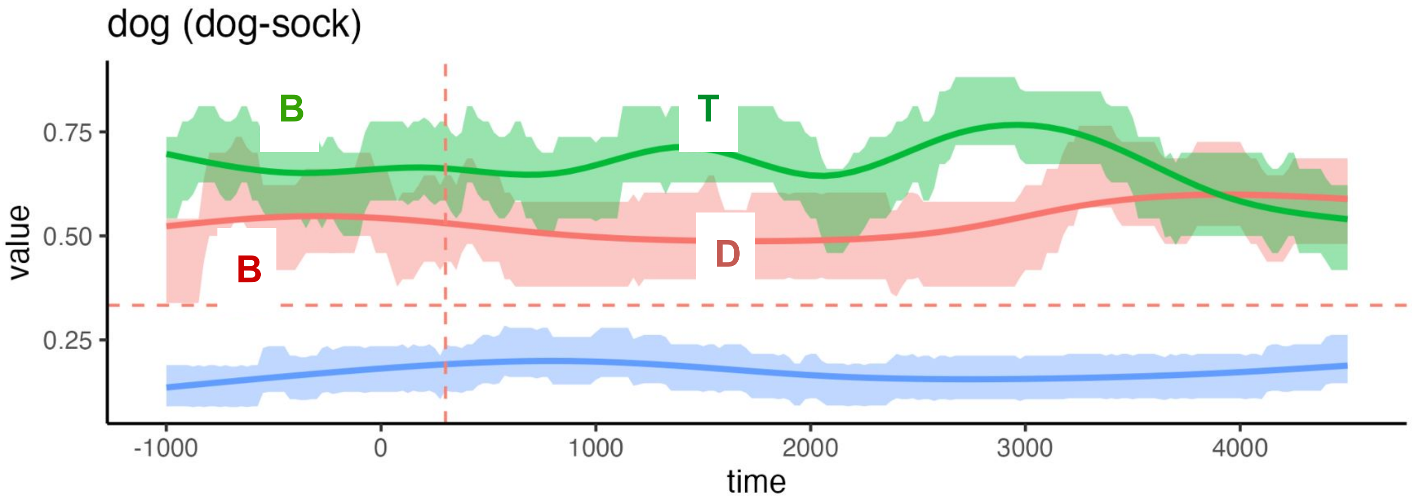
before audio prompt



Signal Processing Challenges in Language Assessments

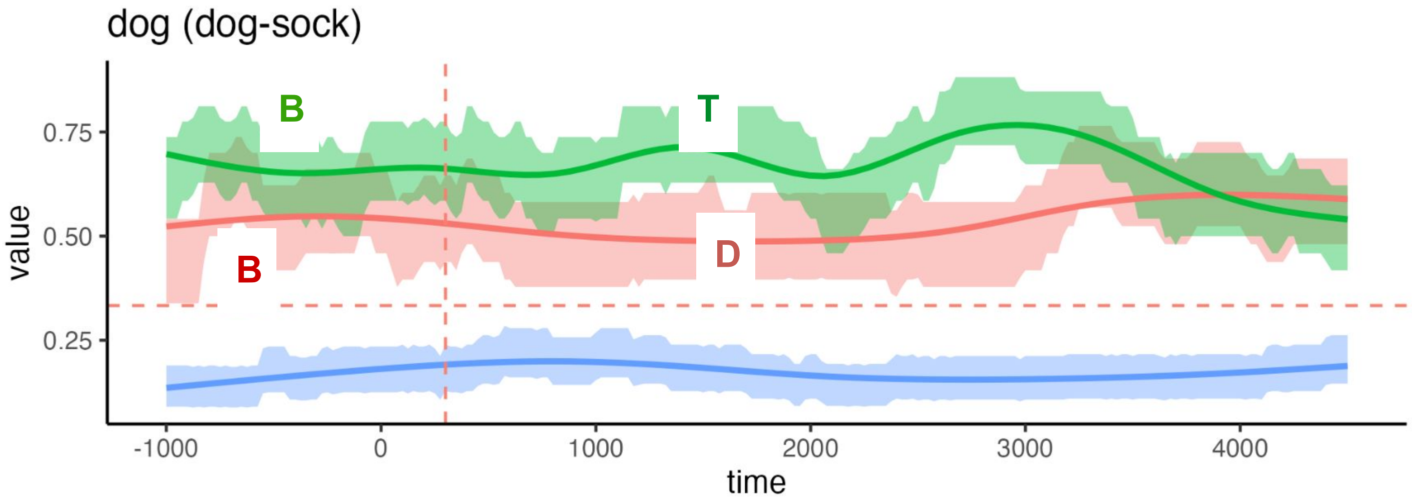
- **Pre-audio baselines:**
  - Measure infant's gaze before audio cue to provide reference point
  - Help distinguish random vs. audio-triggered gaze shifts
- **Image distractors:**
  - Prevent bias towards particular stimulus
  - Use multiple images to assess genuine audio-cued gaze direction
- Determining signal processing techniques that yield least noise and most convergence with word recognition

Methods - Eye-tracking Task



- Target look during **Test**: **T**
- Target look during **Test** - during **Baseline**: **T - B**
- Image as **Target** - Image as **Distractor**: **T - D** (requires at least a yoked pair)

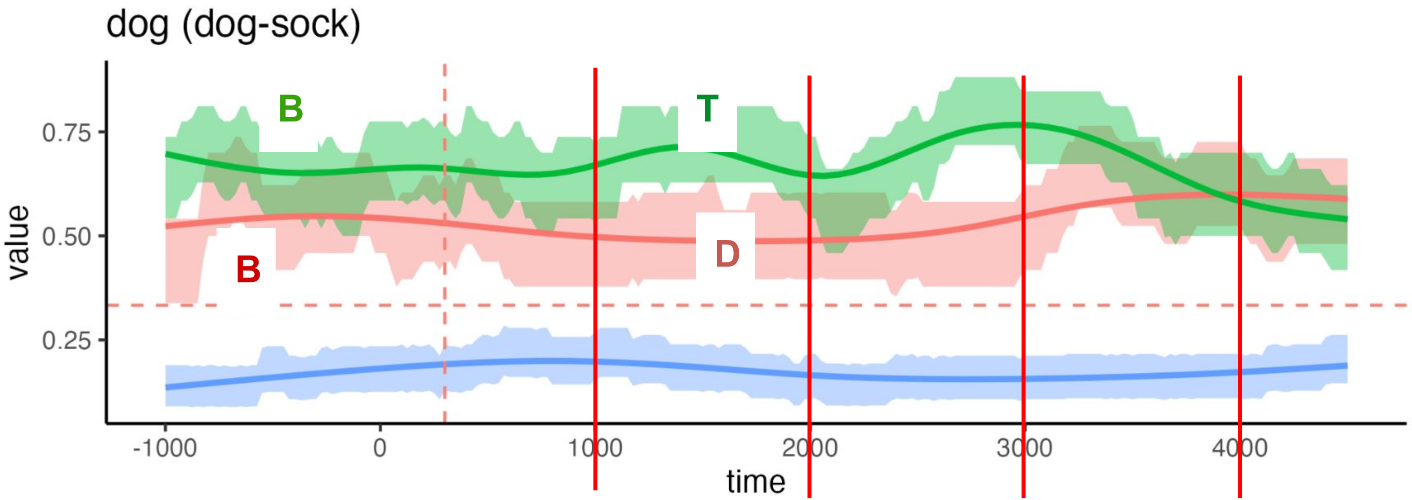
Methods - Eye-tracking Task



- high **T**
- higher **T** than **B**
- higher $\mathbf{T}_{\mathbf{image}}$ than $\mathbf{D}_{\mathbf{image}}$

≈ target recognition

Methods - Eye-tracking Task



many combinations
(time window × method)

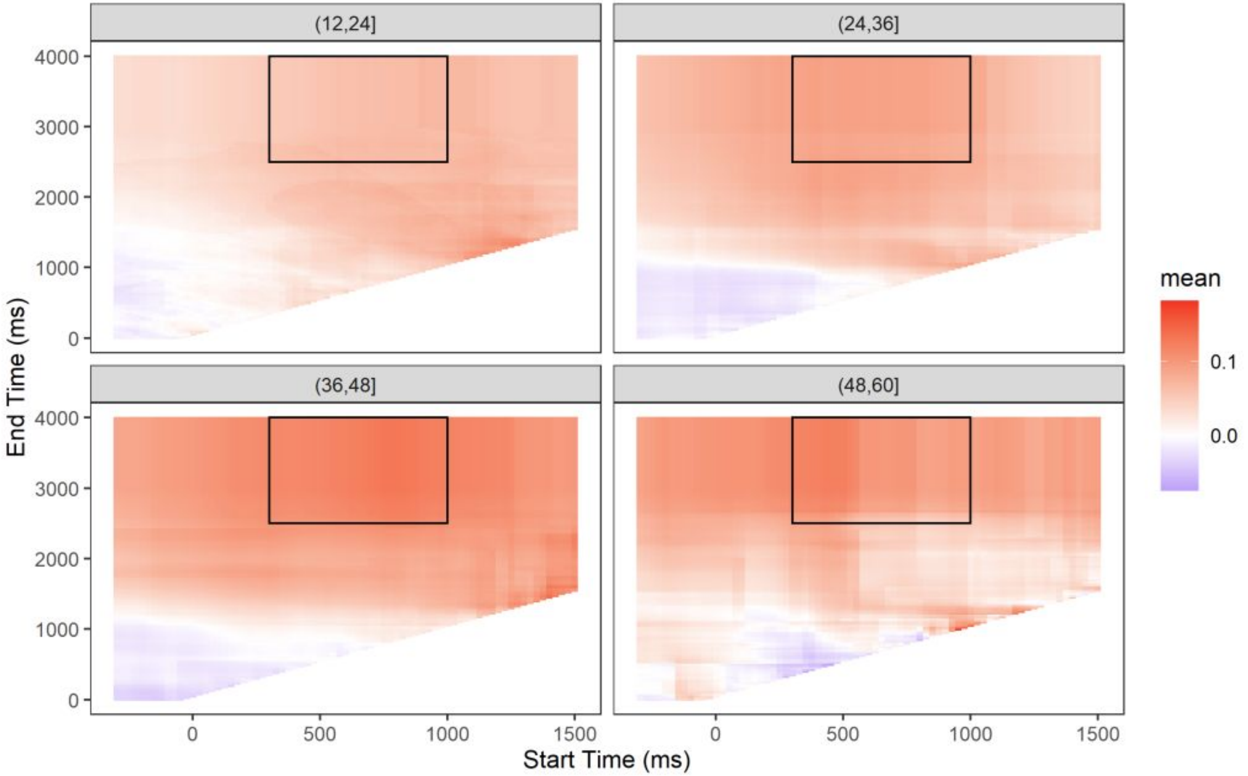
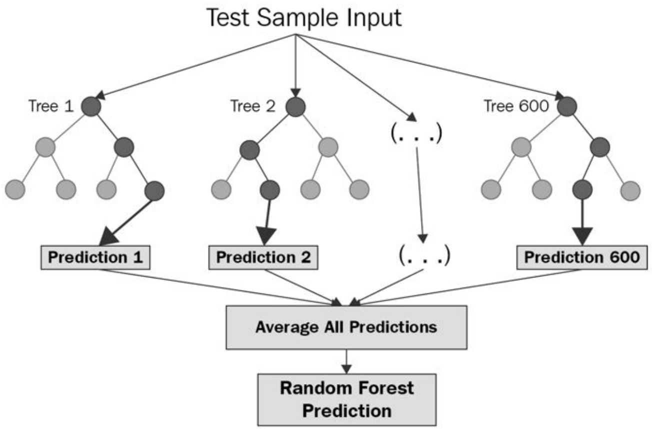


Figure 4: Participants' average inter-item correlation for proportion of looking time to familiar targets, as a function of window start time and end time, with each facet showing a different age group. More positive (red) correlations are more desirable, and blue/white represent start/end time combinations that researchers should avoid. Black lines highlight the region (start time: [300, 1000], end time: [2500, 4000]) in which IICs tend to be highest (mean = .093; range = [.042 - .131]).

Zettersten et al. (2021)

https://escholarship.org/content/qt9j05h4n1/qt9j05h4n1_noSplash_3c8f0548a4836b37c7c6f5a348cb9f2a.pdf

Random Forest Algorithm



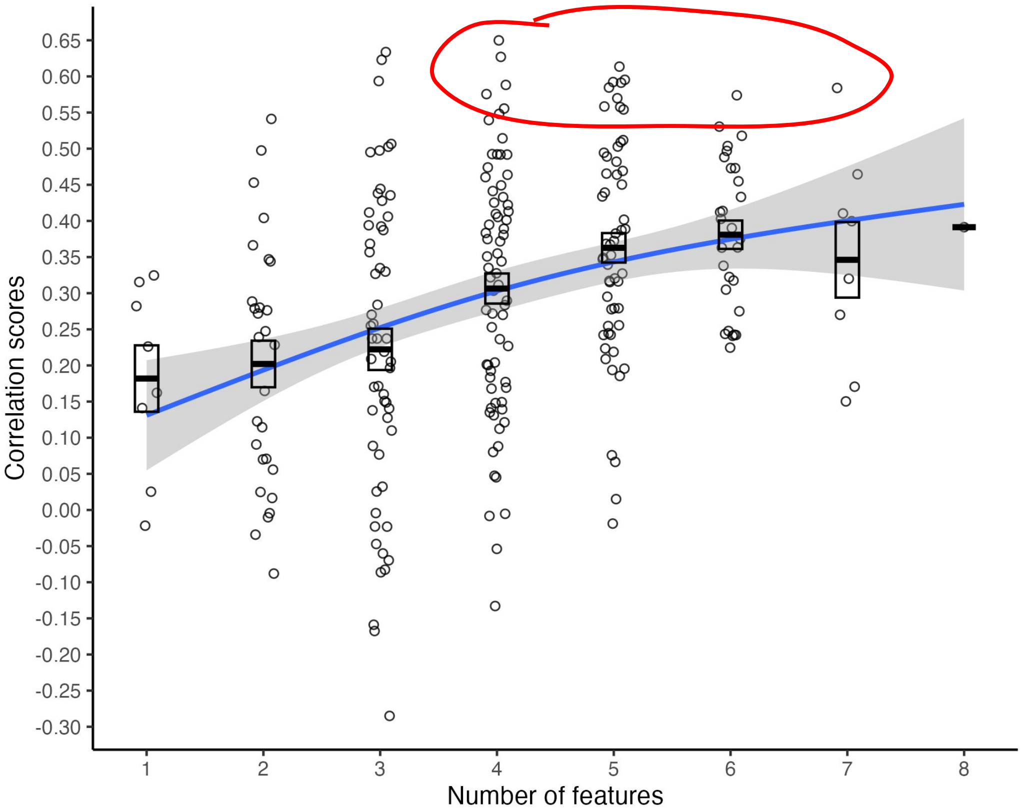
source.

- An **Ensemble** Machine Learning Method
- Builds multiple **trees**, each using different **feature subsets**.
- Handles high-dimensional, complex data.
- We incorporate a wide range of gaze metrics for a more refined analysis of target recognition.

Utilising big data

- with many potential gaze metrics to consider
- ML/AI model such as random forests are well-suited for handling high-dimensional, complex gaze data but require larger data to train
- the emergence of large, open datasets like Peekbank provides an opportunity to gain new insights into early language development by leveraging the power of big data.

Optimal number of features



Optimal feature combinations

Best ten combinations:

Feature combination	Number of features	Correlation
1_sound:2_sound:diff_1:diff_3	4	.650
2_sound:diff_1:diff_3	3	.634
1_sound:3_sound:diff_1:diff_3	4	.627
1_sound:diff_1:diff_3	3	.623
1_sound:2_sound:diff_1:diff_2:diff_3	5	.614
2_sound:diff:diff_1:diff_2:diff_3	5	.595
1_sound:2_sound:diff_3	3	.594
sound:2_sound:diff_1:diff_2:diff_3	5	.592
1_sound:3_sound:diff_1:diff_2:diff_3	5	.591
2_sound:diff_1:diff_2:diff_3	4	.588

Comparisons with rule-based approach

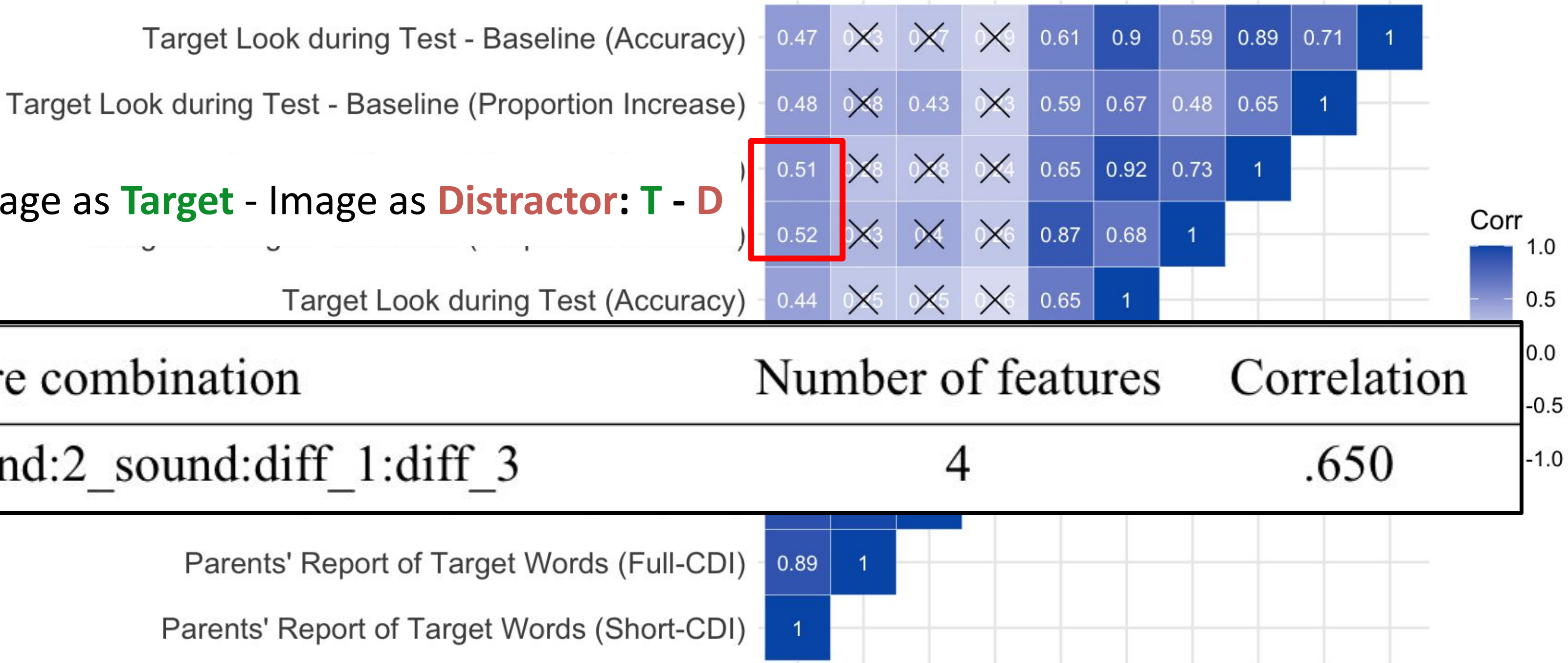
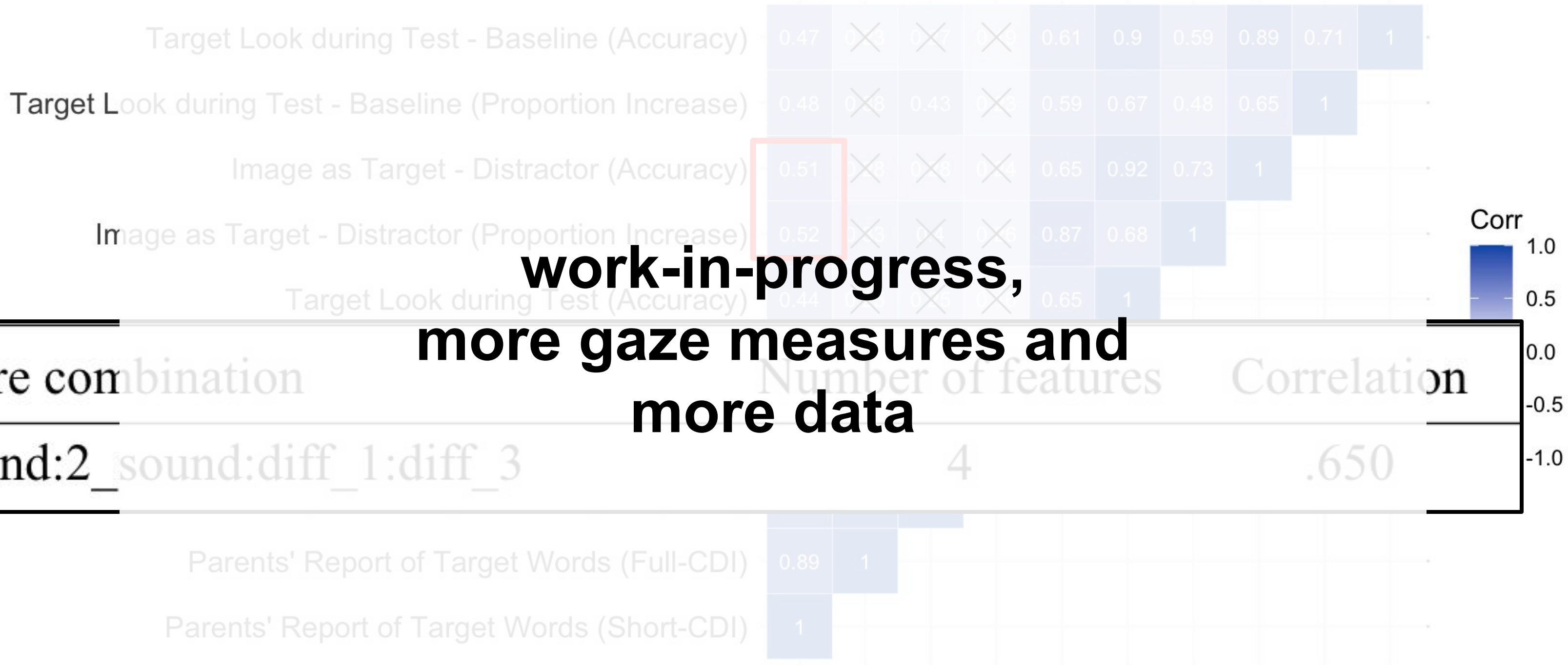


Image as **Target** - Image as **Distractor**: **T** - **D**

Feature combination	Number of features	Correlation
1_sound:2_sound:diff_1:diff_3	4	.650

Comparisons with rule-based approach



work-in-progress,
more gaze measures and
more data

Feature combination	Number of features	Correlation
1_sound:2_sound:diff_1:diff_3	4	.650