table_materials

anjie

2023-01-14

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(here)
```

```
## here() starts at /Users/caoanjie/Desktop/projects/metalabr_exp
```

```
library(metafor)
```

```
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
##
## Loading required package: metadat
##
## Loading the 'metafor' package (version 3.8-1). For an
## introduction to the package please type: help(metafor)
```

```
library(kableExtra)
```

```
## Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output
## %in% : 'length(x) = 2 > 1' in coercion to 'logical(1)'

##
## Attaching package: 'kableExtra'
##
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
library(xtable)
library(ggthemes)

source(here("helper/summary_table_help.r"))

d <- read.csv(here("data/clean_data.csv"))
summary_d <- readRDS(here("cached_data/summary_d.Rds"))
```

summary table

- name of phenomenon
- n studies
- n effects
- n kids
- age range
- mean age
- average effect size

```
summary_d_print <- summary_d %>%
  mutate(es_print =
           paste0(round(es, 2), " [", round(es_lb, 2), ",", round(es_ub, 2), "]"),
         mean_age_print = round(mean_age, 2),
         n_kids_print = round(n_kids, 0),
         age_range_print = paste0("[", round(min_age, 2), ",", round(max_age, 2), "]")
           ) %>%
  select(ds_clean, n_studies, n_effects, n_kids_print, mean_age_print, age_range_print, es_print ) %>%
  arrange(ds_clean)
```

```
summary_d_print %>%
  kable( "latex", align = "l", digits = 2,
      row.names = FALSE,
      col.names = c("Reseach Area",
                    "N Studies",
                    "N effect sizes",
                    "N participants",
                    "Mean age (months)",
                    "Age range",
                    "Average Effect Size"
                    )) %>%
kable_styling(full_width = TRUE,
              bootstrap_options = c("condensed"))
```

Reseach Area	N Studies	N effect sizes	N participants	Mean age (months)	Age range	Average Effect Size
Abstract rule learning	20	95	1123	7.76	[3.91,13.47]	0.22 [0.07,0.37]
Categorization bias	9	80	382	51.18	[15.56,336.01]	0.25 [-0.54,1.05]
Cross-situational word learning	16	48	2241	56.57	[12.23,131.01]	0.67 [0.5,0.84]
Familiar word recognition	16	34	586	10.93	[4.86,15.01]	0.54 [0.38,0.69]
Gaze following (combined)	27	81	1407	11.05	[3.2,23.99]	0.81 [0.61,1.01]
Infant directed speech preference	21	83	985	6.30	[0.07,18.59]	0.47 [0.28,0.65]
Label advantage in concept learning	16	100	1644	10.96	[3,17.94]	0.36 [0.23,0.48]
Language discrimination and preference	36	153	2060	4.43	[0.05,11.56]	-0.13 [-0.26,0]
Mispronunciation sensitivity	32	249	2122	20.50	[6.83,44.22]	0.45 [0.24,0.66]
Mutual exclusivity	45	131	2222	35.17	[14.63,126]	1.27 [0.99,1.56]
Natural speech preference	19	55	786	4.60	[0.05,12.5]	0.44 [0.23,0.65]
Online word recognition	6	14	330	20.46	[14.98,30]	1.37 [0.78,1.96]
Prosocial agents	26	61	1244	12.82	[4.57,31.54]	0.4 [0.29,0.52]
Simple arithmetic competences	6	14	369	5.85	[4.8,9.17]	0.25 [0.04,0.46]
Sound symbolism	11	44	425	12.70	[4.14,37.98]	0.16 [-0.01,0.33]
Statistical sound category learning (habituation)	6	11	350	6.73	[2.5,11.49]	0.56 [0.19,0.93]
Statistical word segmentation	31	103	804	9.08	[5.88,18.66]	-0.08 [-0.18,0.02]
Switch task	47	143	2764	15.99	[12.34,24.9]	-0.16 [-0.25,-0.06]
Symbolic play	31	196	7148	28.46	[11,67.51]	0.63 [0.53,0.72]
Syntactic bootstrapping	18	60	832	24.93	[14.9,42]	0.24 [0.03,0.44]
Vowel discrimination	33	143	2418	7.65	[0.1,29.97]	0.59 [0.43,0.75]

```r
colnames(summary_d_print) <-c("Reseach Area",
                    "N Studies",
                    "N effect sizes",
                    "N participants",
                    "Mean age (months)",
                    "Age range",
                    "Average Effect Size")


table <- xtable::xtable(as.matrix(summary_d_print),
                        caption = "LMAO")

print.xtable(table, include.rownames = FALSE, comment = FALSE,
             size="\\fontsize{8pt}{9pt}\\selectfont")
```

```
## \begin{table}[ht]
## \centering
## \begingroup\fontsize{8pt}{9pt}\selectfont
## \begin{tabular}{lllllll}
##   \hline
## Reseach Area & N Studies & N effect sizes & N participants & Mean age (months) & Age range & Average
##   \hline
## Abstract rule learning & 20 &  95 & 1123 &  7.76 & [3.91,13.47] & 0.22 [0.07,0.37] \\
##   Categorization bias &  9 &  80 &  382 & 51.18 & [15.56,336.01] & 0.25 [-0.54,1.05] \\
##   Cross-situational word learning & 16 &  48 & 2241 & 56.57 & [12.23,131.01] & 0.67 [0.5,0.84] \\
##   Familiar word recognition & 16 &  34 &  586 & 10.93 & [4.86,15.01] & 0.54 [0.38,0.69] \\
##   Gaze following (combined) & 27 &  81 & 1407 & 11.05 & [3.2,23.99] & 0.81 [0.61,1.01] \\
##   Infant directed speech preference & 21 &  83 &  985 &  6.30 & [0.07,18.59] & 0.47 [0.28,0.65] \\
##   Label advantage in concept learning & 16 & 100 & 1644 & 10.96 & [3,17.94] & 0.36 [0.23,0.48] \\
##   Language discrimination and preference & 36 & 153 & 2060 &  4.43 & [0.05,11.56] & -0.13 [-0.26,0] \
##   Mispronunciation sensitivity & 32 & 249 & 2122 & 20.50 & [6.83,44.22] & 0.45 [0.24,0.66] \\
##   Mutual exclusivity & 45 & 131 & 2222 & 35.17 & [14.63,126] & 1.27 [0.99,1.56] \\
##   Natural speech preference & 19 &  55 &  786 &  4.60 & [0.05,12.5] & 0.44 [0.23,0.65] \\
##   Online word recognition &  6 &  14 &  330 & 20.46 & [14.98,30] & 1.37 [0.78,1.96] \\
##   Prosocial agents & 26 &  61 & 1244 & 12.82 & [4.57,31.54] & 0.4 [0.29,0.52] \\
##   Simple arithmetic competences &  6 &  14 &  369 &  5.85 & [4.8,9.17] & 0.25 [0.04,0.46] \\
##   Sound symbolism & 11 &  44 &  425 & 12.70 & [4.14,37.98] & 0.16 [-0.01,0.33] \\
##   Statistical sound category learning (habituation) &  6 &  11 &  350 &  6.73 & [2.5,11.49] & 0.56 [
##   Statistical word segmentation & 31 & 103 &  804 &  9.08 & [5.88,18.66] & -0.08 [-0.18,0.02] \\
##   Switch task & 47 & 143 & 2764 & 15.99 & [12.34,24.9] & -0.16 [-0.25,-0.06] \\
##   Symbolic play & 31 & 196 & 7148 & 28.46 & [11,67.51] & 0.63 [0.53,0.72] \\
##   Syntactic bootstrapping & 18 &  60 &  832 & 24.93 & [14.9,42] & 0.24 [0.03,0.44] \\
##   Vowel discrimination (native) & 33 & 143 & 2418 &  7.65 & [0.1,29.97] & 0.59 [0.43,0.75] \\
##   Vowel discrimination (non-native) & 15 &  49 &  600 &  8.24 & [2.41,17.78] & 0.65 [0.2,1.1] \\
##   Word Segmentation (combined) & 74 & 315 & 2910 &  9.24 & [4.46,24.74] & 0.2 [0.14,0.26] \\
##    \hline
## \end{tabular}
## \endgroup
## \caption{LMAO}
## \end{table}
```

curve table

```
age_df <- readRDS(here("cached_data/age_models_df.Rds"))

age_df_wide <- age_df %>%
  filter(ic == "AICc") %>%
  rename(Dataset = dataset) %>%
  select(REML, model_spec_clean, Dataset) %>%
  pivot_wider(names_from = model_spec_clean,
              values_from = REML)

age_df_wide$min <- apply(age_df_wide[c('Linear','Log', 'Quadratic', 'Const')], 1, min)


age_df_wide %>%
  mutate(across(where(is.numeric), round, 2)) %>%
  mutate_all(~cell_spec(.x, bold  = (.x == min))) %>%
  select(-min) %>%
  rename(Constant = Const) %>%
  kable(digits = 2, escape = F)
```

Dataset	Linear	Log	Quadratic	Constant
Label advantage in concept learning	169.48	**168.53**	170.16	170.89
Vowel discrimination (native)	256.49	256.13	256.78	**255.15**
Vowel discrimination (non-native)	73.25	73.36	73.15	**71.69**
Statistical word segmentation	128.84	129.01	128.62	**127.5**
Online word recognition	**46.5**	46.73	46.65	48.72
Mutual exclusivity	421.6	**415.85**	432.76	453.07
Sound symbolism	58.2	**58.16**	58.83	61.04
Categorization bias	300.29	**299.99**	300.37	300.9
Familiar word recognition	27.46	28.32	**27.18**	28.86
Abstract rule learning	140.8	141.34	**140.47**	140.91
Switch task	204.79	204.81	204.73	**203.67**
Mispronunciation sensitivity	620.05	628.16	**613.67**	644.4
Prosocial agents	82.16	81.95	82.23	**80.08**
Simple arithmetic competences	22.91	23.01	22.81	**16.26**
Symbolic play	234.15	234.11	234.13	**233.57**
Natural speech preference	111.4	112.01	**110.97**	111.83
Cross-situational word learning	79.81	81.62	**79.7**	83.71
Language discrimination and preference	264.7	265.59	**262.65**	264.95
Syntactic bootstrapping	107.28	**106.99**	107.57	107.47
Statistical sound category learning (habituation)	33.47	34.54	32.94	**30.46**
Gaze following (combined)	151.53	159.88	**149.47**	193.2
Word Segmentation (combined)	328.83	328.6	329.16	**327.55**
Infant directed speech preference	70.17	70.87	70.06	**69.13**

moderators plot

```r
best_fit_bmeasure_model_df <- readRDS(here("cached_data/best_fit_behavioral_measure_df.Rds"))
best_fit_exposure_phase_df <- readRDS(here("cached_data/best_fit_exposure_phase.Rds"))
best_fit_real_visual_df <- readRDS(here("cached_data/best_fit_real_visual_df.Rds"))
best_fit_real_aud_df <- readRDS(here("cached_data/best_fit_real_aud_df.Rds"))
best_fit_major_author_df <- readRDS(here("cached_data/best_fit_major_author_df.Rds"))
```

```r
# baseline is looking
bmeasure_df <- best_fit_bmeasure_model_df %>%
  filter(grepl("behavioral_measure", term)) %>%
  mutate(term_print = case_when(
    grepl("other", term) ~ "Behavioral Measure - Other",
    TRUE ~ "Behavioral Measure - Sucking"
  )) %>%
  select(dataset, term_print, estimate, std.error)

# baseline is conditionining

ep_df <- best_fit_exposure_phase_df %>%
  filter(grepl("exposure_phase", term)) %>%
  mutate(term_print = case_when(
    grepl("habituation", term) ~ "Exposure phase - Habituation",
    grepl("test_only", term) ~ "Exposure phase - Familiarization",
    grepl("familiarization", term) ~ "Exposure phase - Test only"
  )) %>%
  select(dataset, term_print, estimate, std.error)

rv_df <- best_fit_real_visual_df %>%
  filter(grepl("visual", term)) %>%
  mutate(term_print = "Visual stimulus type - Representation") %>%
  select(dataset, term_print, estimate, std.error)


av_df <- best_fit_real_aud_df %>%
  filter(grepl("aud", term)) %>%
  mutate(term_print = "Auditory stimulus type - Natural") %>%
  select(dataset, term_print, estimate, std.error)

ma_df <- best_fit_major_author_df %>%
  filter(grepl("author", term)) %>%
  mutate(term_print = "By Major Author") %>%
  select(dataset, major_author, term_print, estimate, std.error)
```

```r
all_mod_df <- bind_rows(bmeasure_df, ep_df, rv_df, av_df, ma_df) %>%
  mutate(lb = estimate - std.error  * 1.96,
         ub = estimate + std.error * 1.96) %>%
  separate(term_print, into = c("type", "group"), sep = "-")
```

```
## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 28 rows [41, 42,
## 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, ...].
```

bmeasure

behavioral measure

```
bm_p <-
all_mod_df %>%
  filter(grepl("Behavioral", type)) %>%
  ggplot(aes(x = reorder(dataset,estimate), y = estimate, color = group)) +
  geom_pointrange(aes(y = estimate, ymin = lb, ymax = ub),
                  position = position_dodge(width = .4)) +
  geom_hline(yintercept = 0, linetype = "dashed")+
  coord_flip() +
  xlab("") +
  ylab("") +
  labs(title = "Behavioral Measure (Baseline: looking)") +
  theme_few() +
  theme(
    legend.position = c(0.8, 0.15),
    legend.title = element_blank(),
    legend.background = element_rect(fill = NA, color = NA),
    legend.text=element_text(size=6),
    plot.title = element_text(hjust = 0.5, size = 10)
  )
```

exposure phase

```
ep_mod_p <-
all_mod_df %>%
  filter(grepl("Exposure", type)) %>%
  ggplot(aes(x = reorder(dataset,estimate), y = estimate, color = group)) +
  geom_pointrange(aes(y = estimate, ymin = lb, ymax = ub),
                  position = position_dodge(width = .4)) +
  geom_hline(yintercept = 0, linetype = "dashed")+
  coord_flip() +
    scale_x_discrete(position = "top")+
  xlab("") +
  ylab("") +
  labs(title = "Exposure phase (Baseline: Conditioning)") +
  theme_few() +
  theme(
    legend.position = c(0.8, 0.15),
    legend.title = element_blank(),
    legend.background = element_rect(fill = NA, color = NA),
    legend.text=element_text(size=6),
    plot.title = element_text(hjust = 0.5, size = 10)
  )
```

visual

```
vr_mod_p <-
all_mod_df %>%
  filter(grepl("Visual stimulus", type)) %>%
  ggplot(aes(x = reorder(dataset,estimate), y = estimate, color = group)) +
  geom_pointrange(aes(y = estimate, ymin = lb, ymax = ub),
                  position = position_dodge(width = .4)) +
  geom_hline(yintercept = 0, linetype = "dashed")+
  coord_flip() +
  scale_x_discrete(position = "top")+
  xlab("") +
  ylab("") +
  labs(title = "Visual stimulus type (Baseline: real stimulus)") +
  theme_few() +
  theme(
    legend.position = c(0.8, 0.1),
    legend.title = element_blank(),
    legend.background = element_rect(fill = NA, color = NA),
    legend.text=element_text(size=6),
    plot.title = element_text(hjust = 0.5, size = 10)
  )
```

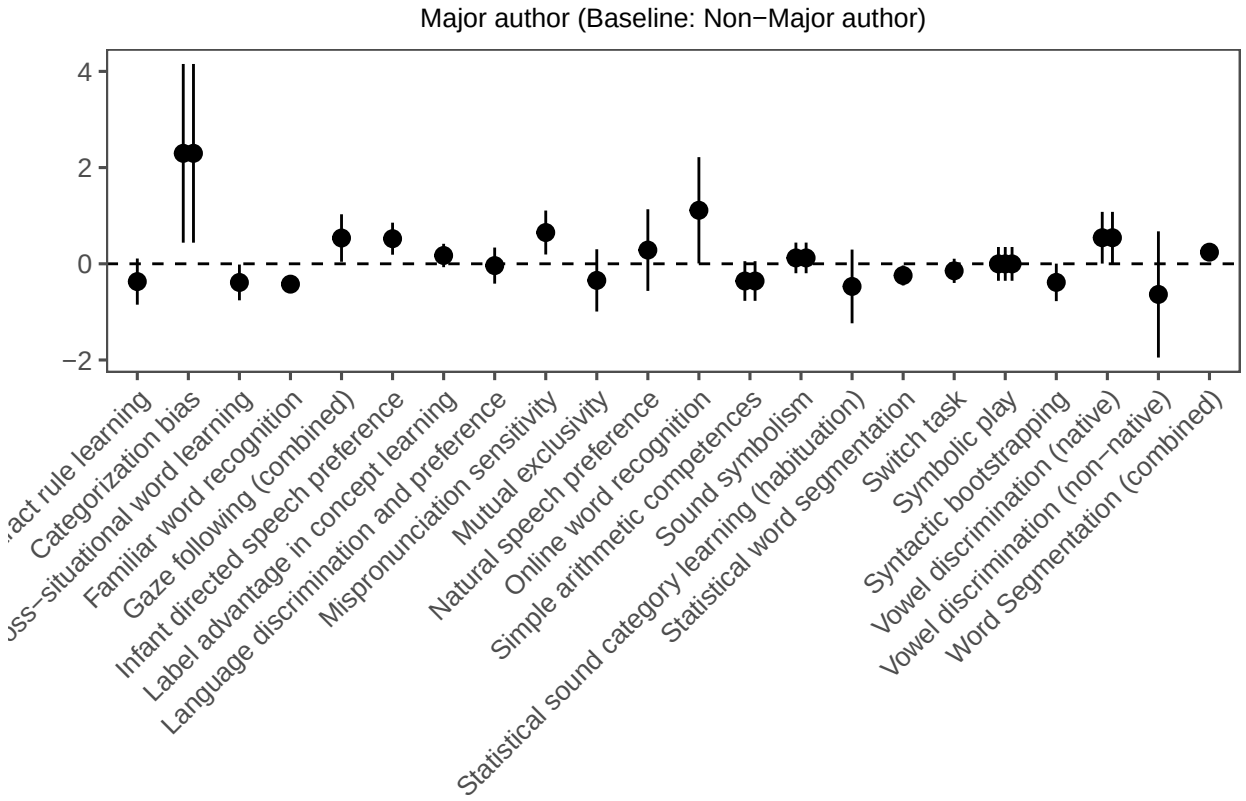
auditory

```
ar_mod_p <- all_mod_df %>%
  filter(grepl("Auditory stimulus", type)) %>%
  ggplot(aes(x = reorder(dataset,estimate), y = estimate, color = group)) +
  geom_pointrange(aes(y = estimate, ymin = lb, ymax = ub),
                  position = position_dodge(width = .4)) +
  geom_hline(yintercept = 0, linetype = "dashed")+
  coord_flip() +
  xlab("") +
  ylab("") +
  labs(title = "Auditory stimulus type (Baseline: artificial stimulus)") +
  theme_few() +
  theme(
    legend.position = c(0.8, 0.1),
    legend.title = element_blank(),
    legend.background = element_rect(fill = NA, color = NA),
    legend.text=element_text(size=6),
    plot.title = element_text(hjust = 0.5, size = 10)
  )
```

Major author?

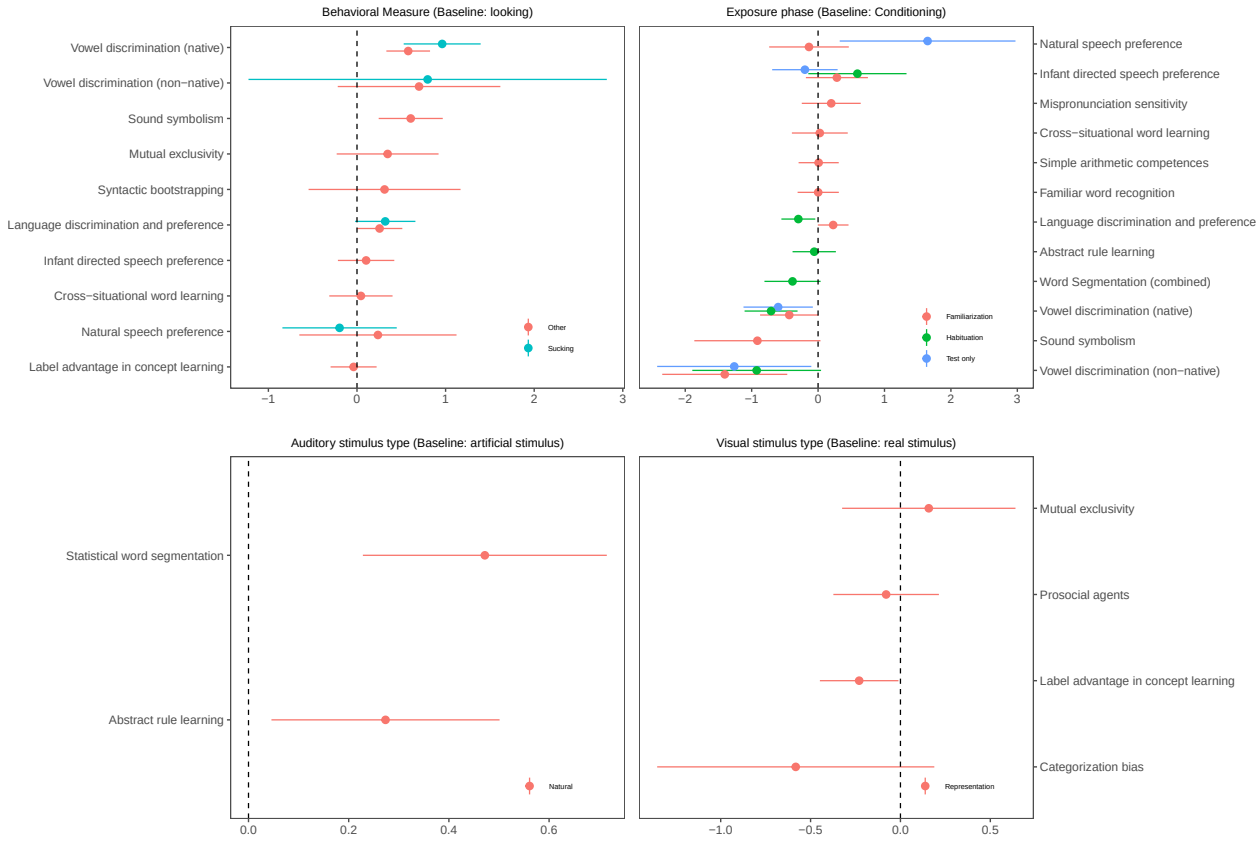
```
ma_p <-
all_mod_df %>%
  filter(grepl("Author", type)) %>%
```

```
  ggplot(aes(x = dataset, y = estimate, group = major_author)) +
  geom_pointrange(aes(y = estimate, ymin = lb, ymax = ub),
                  position = position_dodge(width = .4)) +
  geom_hline(yintercept = 0, linetype = "dashed")+
  xlab("") +
  ylab("") +
  labs(title = "Major author (Baseline: Non-Major author)") +
  theme_few() +
  theme(
    legend.position = c(0.8, 0.1),
    legend.title = element_blank(),
    legend.background = element_rect(fill = NA, color = NA),
    legend.text=element_text(size=6),
    plot.title = element_text(hjust = 0.5, size = 10),
    axis.text.x = element_text(angle = 45, hjust=1)
  )

ma_p
```



putting things together

```
library(patchwork)

(bm_p | ep_mod_p) / (ar_mod_p  | vr_mod_p)
```



devo curves

```
all_age_pred <- readRDS(here("cached_data/min_age_pred_df.RDS"))

age_range_df <- d %>%
  group_by(ds_clean) %>%
  summarise(min_age = min(mean_age_months, na.rm = TRUE),
            max_age = max(mean_age_months, na.rm = TRUE)) %>%
  rename(ds_name = ds_clean)


categorized_age_pred_df <- all_age_pred %>%
  select(pred, ds_name) %>%
  group_by(ds_name) %>%
  mutate(mean_age_months = (row_number() + 1) / 2) %>%
  left_join(age_range_df, by = c("ds_name")) %>%
  filter(mean_age_months > min_age,
         mean_age_months < min(36, max_age)) %>%
  mutate(
    ds_type = case_when(
      ds_name %in% c( "Language discrimination and preference",
                      "Mispronunciation sensitivity",
                      "Natural speech preference", "Sound symbolism", "Statistical sound category learni
                      "Vowel discrimination (native)",
```

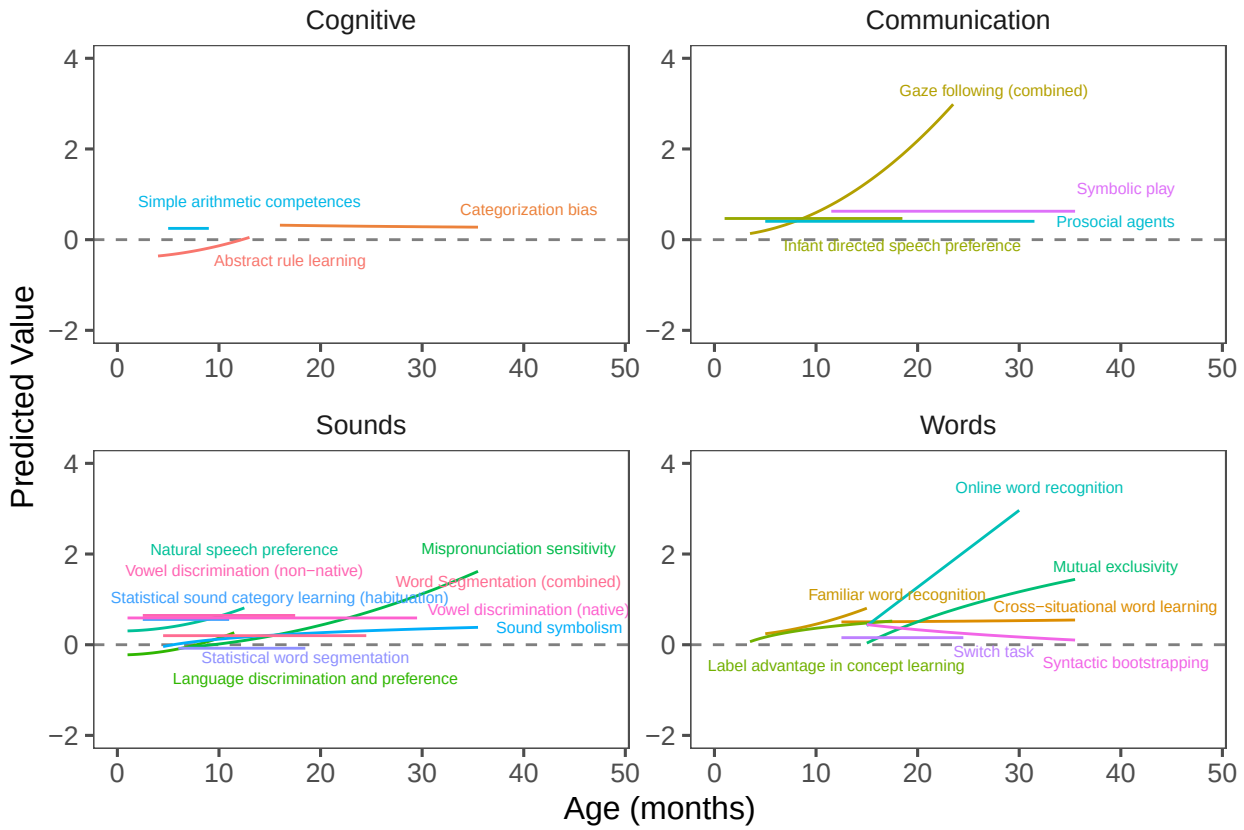
```
                      "Vowel discrimination (non-native)" ,"Word Segmentation (combined)",
                      "Statistical word segmentation") ~ "Sounds",

      ds_name %in% c("Familiar word recognition",
                      "Cross-situational word learning",
                       "Label advantage in concept learning",
                      "Mutual exclusivity",
                      "Online word recognition",
                      "Syntactic bootstrapping",
                      "Switch task") ~ "Words",

      ds_name %in% c("Gaze following (combined)",
                      "Prosocial agents",
                      "Symbolic play",
                      "Infant directed speech preference") ~ "Communication",

      ds_name %in% c("Abstract rule learning",
                      "Categorization bias",
                      "Simple arithmetic competences") ~ "Cognitive"
     )
  ) %>%
  mutate(pred = case_when(
    ds_name == "Abstract rule learning" ~ -pred,
    ds_name == "Switch task" ~ -pred,
    TRUE ~ pred
  )) %>%
  select(ds_name, ds_type, mean_age_months, pred)



label_df <- categorized_age_pred_df %>%
  filter(mean_age_months == max(mean_age_months))

label_df$x = c(
  15 - 3,
  29.5 + 11, #"Vowel discrimination (native)"
  17.5 - 5, #"Vowel discrimination (non-native)"
  18.5, #"Statistical word segmentation"
  30 + 2, #"Online word recognition"
  35.5 + 4, #Mutual exclusivity ,
  35.5 + 8, #Sound symbolism,
  35.5 + 5 , # Categorization bias
  15 + 3,#Familiar word recognition
  13 + 4,#Abstract rule learning
  24.5+ 3, #Switch task,
  35.5 + 4,#"Mispronunciation sensitivity"
  31.5 + 8, #Prosocial Agents
  9 + 4, #"Simple arithmetic competences"
  35.5 + 5,#Symbolic play
  12.5 , #"Natural speech preference"
  35.5 + 3, # "Cross-situational word learning"
  11.5 + 8, # "Language discrimination and preference"
  35.5 + 5, # "Syntactic bootstrapping"
```

```
  11 + 5, #"Statistical sound category learning (habituation)"
  23.5 + 4, #"Gaze following (combined)"
  24.5 + 14, #"Word Segmentation (combined)"
   18.5# "Infant directed speech preference"

)
label_df$y = c(
  0.05004268 - 0.5,
  0.5901183 + 0.2, #"Vowel discrimination (native)"
  0.6457766 + 1, #"Vowel discrimination (non-native)"
  -0.07808538 - 0.2, #"Statistical word segmentation"
  2.964331 + 0.5, #"Online word recognition"
  1.44 + 0.3,#Mutual exclusivity
  0.383, #Sound symbolism,
  0.275 + .4, # Categorization bias
  0.805 + 0.3,#Familiar word recognition
  0.05 - 0.5 ,#Abstract rule learning
  0.156 - 0.3,#Switch task
  1.62 + 0.5,#"Mispronunciation sensitivity"
  0.405, #Prosocial agents
  0.249 + 0.6,#"Simple arithmetic competences"
 0.627 + 0.5, #Symbolic play
  0.811 + 1.3, #"Natural speech preference"
  0.545 + .3, # "Cross-situational word learning"
  0.265 - 1, # "Language discrimination and preference"
  0.103 - 0.5, # "Syntactic bootstrapping"
  0.557 + 0.5, #"Statistical sound category learning (habituation)"
  2.99 + 0.3, #"Gaze following (combined)"
 0.2 + 1.2, #"Word Segmentation (combined)"
 0.467 - 0.6# "Infant directed speech preference"

)

p <- categorized_age_pred_df %>%
  ggplot(aes(x = mean_age_months, y = pred, group = ds_name, color = ds_name)) +
  geom_hline(yintercept = 0, linetype = "dashed", alpha = .5) +
  geom_line(alpha = 1, width = 2)+
  geom_text(data = label_df, aes(x = x, y = y, label = ds_name),
                  size = 2.2,
                 segment.color = NA) +
  xlim(0, 48) +
  ylim(-2, 4) +
  facet_wrap(~ds_type, scales = "free") +
  theme_few()+
  theme(legend.position = "none") +
  xlab("Age (months)") +
  ylab("Predicted Value")
```

```
## Warning: Ignoring unknown parameters: width
```

```
## Warning: Ignoring unknown parameters: segment.colour
```

p



label_df$ds_name label_df$mean_age_months label_df$pred