

# Appendix to ‘Detecting social information in a dense database of infants’ natural visual experience’

2021-06-09

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.2.1 --
## v ggplot2 3.3.2      v purrr  0.3.2
## v tibble  3.0.3      v dplyr  1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0

## Warning: package 'ggplot2' was built under R version 3.6.2
## Warning: package 'tibble' was built under R version 3.6.2
## Warning: package 'tidyr' was built under R version 3.6.2
## Warning: package 'dplyr' was built under R version 3.6.2

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(ggplot2)
library(ggthemes)

## Read in preprocessed data
load(here::here('data/preprocessed_data_2021/all_vid_data_from_bbs_all_detections.RData')) # all detect

## all bounding boxes
load(file=here::here('data/preprocessed_data_2021/bounding_box_summaries.RData'))

## Load in gold sample detections
load(here::here('data/preprocessed_data_2021/gold_sample_annotations2020-01-31.RData')) # human annotat
load(here::here('data/preprocessed_data_2021/gold_sample_from_bbs_2021_all_detections.RData'))

# all_gold_sample_frames_op_hc = load(here::here('data/preprocessed_data_2021/gold_sample_from_bbs_2021

## summary for all detections
face_hand_by_age <- all_vid_data %>%
  ungroup() %>%
  tidyr::replace_na(list(faces_and_hands=0)) %>%
  group_by(age_day_bin, child_id) %>%
  summarize(num_frames_total = sum(num_frames),
            # prop faces overall
            num_faces = sum(num_faces),
            num_hands = sum(num_hands),
            prop_faces = num_faces / num_frames_total,
            prop_hands = num_hands / num_frames_total,
            # in center FOV
```

```

    num_faces_center = sum(num_faces_center, na.rm=TRUE),
    num_hands_center = sum(num_hands_center, na.rm=TRUE),
    prop_faces_center = num_faces_center / num_frames_total,
    prop_hands_center = num_hands_center / num_frames_total,
    # detailed face info
    prop_full_faces = sum(num_full_faces, na.rm=TRUE)/num_frames_total,
    prop_faces_and_hands = sum(faces_and_hands)/num_frames_total,
    # face/hand contingency
    prop_faces_with_hands = sum(faces_and_hands, na.rm=TRUE)/num_faces,
    prop_hands_with_faces = sum(faces_and_hands, na.rm=TRUE)/num_hands
  )
}

## `summarise()` regrouping output by 'age_day_bin' (override with `.groups` argument)
all_detections <- face_hand_by_age %>%
  filter(num_frames_total > 2000) %>% # eliminate small data point that skews scaling
  select(prop_faces, prop_hands, num_frames_total, age_day_bin, child_id) %>%
  pivot_longer(cols = c(prop_faces, prop_hands), names_to = "region", values_to = "prop") %>%
  mutate(approach = "uncropped",
         region = ifelse(region == "prop_faces", "Faces", "Hands"))

# Function to evaluate detectors
evaluate_detector <- function(truth, detection) {
  if (truth == TRUE) {
    if (truth == detection) return("TP") # was face/wrist, detected face/wrist
    else return("FN") # was face/wrist, missed face/wrist
  }
  else {
    if (truth == detection) return("TN") # was not face/wrist, did not detect face/wrist
    else return("FP") # was not face/wrist, detected face/wrist
  }
}

# function to return prfs
return_prf_short = function(eval){
  tp=sum(eval == "TP")
  fp=sum(eval == "FP")
  fn=sum(eval == "FN")
  p = tp / (tp + fp)
  r = tp / (tp + fn)
  f=( 2 * p * r ) / (p + r)
  return(c(p,r,f))
}

#

# join human and OP detections
gold_sample <- gold_sample %>%
  select(vid_name, frame, face_present_ketan, hand_present_ketan) %>%
  mutate(face_present_ketan = as.logical(face_present_ketan), hand_present_ketan = as.logical(hand_present_ketan))
  right_join(all_gold_sample_frames_op) %>%
  mutate(face_eval_ketan = evaluate_detector(face_present_ketan, face_detected), hand_eval_ketan = evaluate_detector(hand_present_ketan, hand_detected))

## Joining, by = c("vid_name", "frame")

```

```

## output prfs
face_performance = return_prf_short(gold_sample$face_eval_ketan)
hand_performance = return_prf_short(gold_sample$hand_eval_ketan)

## Move to appendix
## get out summary by age for goldset hand-labeled frames
vid_info <- all_vid_data %>%
  select(child_id, vid_name, age_days, age_day_bin)

summary_by_age_gold <- gold_sample %>%
  left_join(vid_info) %>%
  group_by(age_day_bin, child_id) %>%
  summarize(num_frames_total = n(), num_faces = sum(face_present_ketan), num_hands = sum(hand_present_ketan))

## Joining, by = c("vid_name", "child_id", "age_days")
## `summarise()` regrouping output by 'age_day_bin' (override with `.groups` argument)
### Examine gold sample performance by child hands
load(here::here('data/preprocessed_data_2021/child_adult_hand_annotations_by_frame.RData'))

gold_sample_no_child_hands <- gold_sample %>%
  left_join(child_adult_hand_annotations, by=c("vid_name", "frame")) %>%
  replace_na(list(child_hand_seg = FALSE)) %>% # replace NAs with false (frames not in annotations (NAs))
  filter(child_hand_seg==FALSE) %>% # now these are counted as frames where OP didn't need to detect so
  mutate(hand_eval_adults = evaluate_detector(hand_present_ketan, hand_detected))

summary_by_age_adult_hands_gold <- gold_sample_no_child_hands %>%
  left_join(vid_info) %>%
  group_by(age_day_bin, child_id) %>%
  summarize(num_frames_total = n(), num_faces = sum(face_present_ketan), num_hands = sum(hand_present_ketan))

## Joining, by = c("vid_name", "child_id", "age_days")
## `summarise()` regrouping output by 'age_day_bin' (override with `.groups` argument)

goldset <- summary_by_age_gold %>%
  gather(region, prop, prop_faces, prop_hands) %>%
  mutate(approach = "goldset",
         region = ifelse(region == "prop_faces", "Faces", "Hands"))

goldset_adult_hands <- summary_by_age_adult_hands_gold %>%
  gather(region, prop, prop_faces, prop_hands) %>%
  mutate(approach = "goldset",
         region = ifelse(region == "prop_faces", "Faces", "Hands"))

```

## Face/hand detections relative to human annotations

```

ggplot(all_detections %>% filter(child_id %in% c('S', 'A')),
       aes(x=age_day_bin, y=prop,
           size=log10(num_frames_total),
           col=region)) +
  geom_point(alpha=.2) +

```

```

geom_smooth(span=10, aes(weight = num_frames_total), show.legend = FALSE) +
geom_smooth(data = goldset_adult_hands, span=10, aes(weight = num_frames_total), show.legend = FALSE,
            lty = 2, span=10, se = FALSE) +
geom_smooth(data = goldset, span=10, aes(weight = num_frames_total), show.legend = FALSE,
            lty = 3, span=10, se = FALSE) +
ylab('Proportion Detections') +
xlab('Age (Months)') +
ylim(0,.6) +
facet_grid(.~child_id) +
theme_few(base_size=10) +
ggthemes::scale_color_solarized(name = "") +
scale_size_continuous(name = "Detections (Log 10)") +
theme(legend.text=element_text(size=10)) +
theme(legend.position="bottom")

```

## Warning: Duplicated aesthetics after name standardisation: span

## Warning: Duplicated aesthetics after name standardisation: span

## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 2 rows containing non-finite values (stat\_smooth).

## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 17 rows containing non-finite values (stat\_smooth).

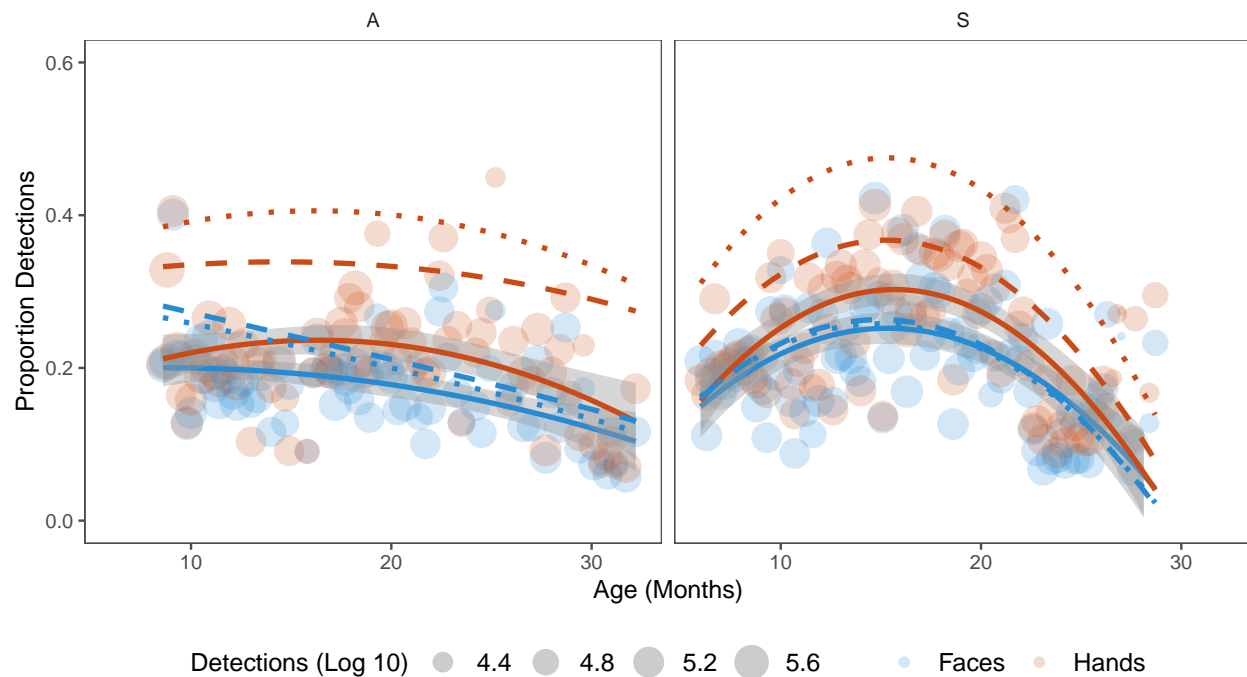


Figure 1: Proportion of faces and hands seen as a function of age for each child in the dataset. Data are binned by each week that the videos were filmed and scaled by the number of frames in that age range. Dashed lines show estimated trend lines from proportion of faces/hands in view when analyzing the gold set of frames made by human annotators. Dotted lines show trend lines from the goldset when frames when children's own hand were detected.

## Density of child vs. adults hands in the visual field

```
### Load turk hand annotations with all bounding boxes (multiple dets per frame)
load(here::here('data/preprocessed_data_2021/hand_annotations_2020-01-29.RData'))

child_hands <- hand_annotations %>%
  filter(label=="Child hand")

adult_hands <- hand_annotations %>%
  filter(label=="Adult hand")

### plot centers of the bounding boxes made for child and adult hands
child_hand_plot <- ggplot(child_hands, aes(x=center_x, y=center_y)) +
  geom_point(alpha=.1) +
  stat_density_2d(aes(fill = ..level..), geom="polygon", alpha=.8) +
  coord_fixed(ratio=1) +
  ggtitle('A. Child hand density') +
  ylim(0,480) +
  xlim(0,640) +
  ylab('') +
  xlab('') +
  theme_few(base_size=10) +
  theme(legend.position="none", axis.text.x=element_blank(), axis.ticks.x=element_blank(), axis.text.y=
  scale_y_reverse()

## Scale for 'y' is already present. Adding another scale for 'y', which
## will replace the existing scale.

##
adult_hand_plot <- ggplot(adult_hands, aes(x=center_x, y=center_y)) +
  geom_point(alpha=.1) +
  stat_density_2d(aes(fill = ..level..), geom="polygon", alpha=.8) +
  coord_fixed(ratio=1) +
  ylim(0,480) +
  xlim(0,640) +
  scale_y_reverse() +
  ylab('') +
  xlab('') +
  theme_few(base_size=10) +
  theme(legend.position="none", axis.text.x=element_blank(), axis.ticks.x=element_blank(), axis.text.y=
  ggtitle('B. Adult hand density')

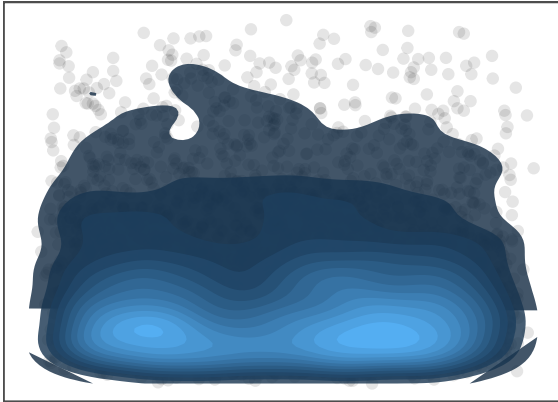
## Scale for 'y' is already present. Adding another scale for 'y', which
## will replace the existing scale.

cowplot::plot_grid(child_hand_plot, adult_hand_plot, nrow=1)
```

## Distribution of faces and hands in the visual field

We explored where in the visual field children tended to see faces and hands, suspecting that these distributions might become wider as children grow older and learn to locomote on their own, following preliminary analyses from Frank (2012). As expected, faces tended to appear in the upper visual field in contrast to hands, which tended to be more centrally located (see Figure C1). However, we found little evidence for any changes in the

A. Child hand density



B. Adult hand density

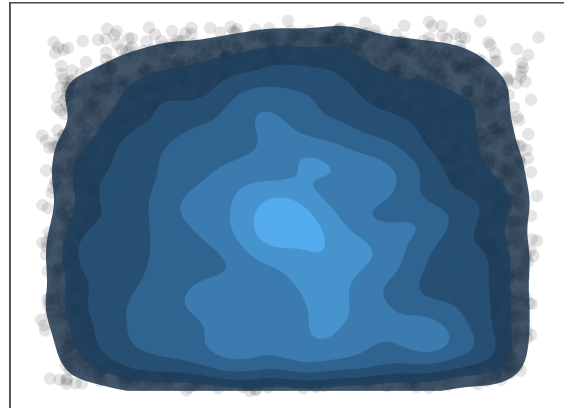


Figure 2: Density estimates for the child (left) and adult (right) hands that were detected in the 24K frame random gold set; each dot represents the center of a bounding box made by an adult participant. Brighter values indicate more detections.

positions of faces and hands across age, suggesting that this is a relatively stable property of infants' visual environment from 6 months of age.

```
bin_size=7
min_age = min(all_vid_bbs$age_days, na.rm=TRUE)
max_age = max(all_vid_bbs$age_days, na.rm=TRUE)
bin_starts = seq(min_age-1, max_age+1, bin_size)
bins = c(bin_starts, max_age)

all_vid_bbs <- all_vid_bbs %>%
  mutate(age_months = age_days/30.4) %>%
  mutate(age_day_bin = cut(age_days, bins, labels=floor(bin_starts/30.4))) %>%
  mutate(age_day_bin = as.numeric(as.character(age_day_bin))) %>%
  mutate(avg_center_y = avg_top + avg_height/2, avg_center_x = avg_left + avg_width/2) %>%
  filter(num_detect > 100)

ggplot(all_vid_bbs %>% filter(label %in% c('face', 'hand')), aes(x=avg_center_x, y = avg_center_y, size=
  geom_point(alpha=.3) +
  stat_density2d(aes(fill = label, color=label), geom="polygon", alpha=.6) +
  coord_fixed(ratio=.76) +
  # ggtitle('Face/hand density') +
  ylab('') +
  xlab('') +
  theme_few(base_size=14) +
  theme(legend.position="top", axis.text.x=element_blank(), axis.ticks.x=element_blank(), axis.text.y=e
  scale_y_reverse() +
  facet_wrap(~age_day_bin) +
  guides(size = FALSE) +
  ggthemes::scale_color_solarized(name = "") +
  ggthemes::scale_fill_solarized(name = "")
```

```
## Warning: stat_contour(): Zero contours were generated
## Warning in min(x): no non-missing arguments to min; returning Inf
## Warning in max(x): no non-missing arguments to max; returning -Inf
```







```
## Warning: stat_contour(): Zero contours were generated
## Warning in min(x): no non-missing arguments to min; returning Inf
## Warning in max(x): no non-missing arguments to max; returning -Inf
```

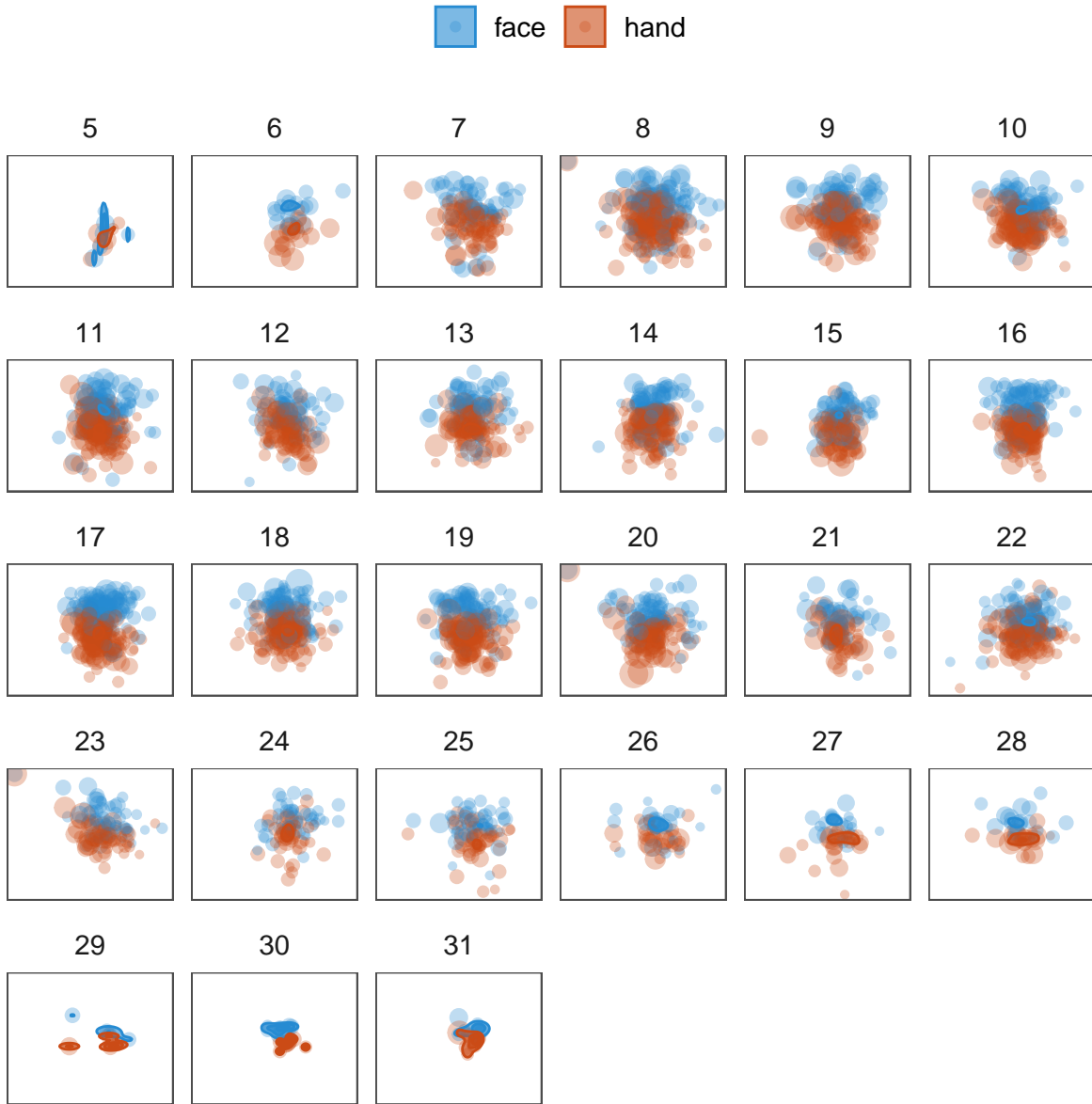


Figure 3: Each panel shows the average position of faces and hands in the visual field; each dot represents the average position from one video within a given age range.