

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

2022-01-05

```
## 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
```

Face/hand detections relative to human annotations

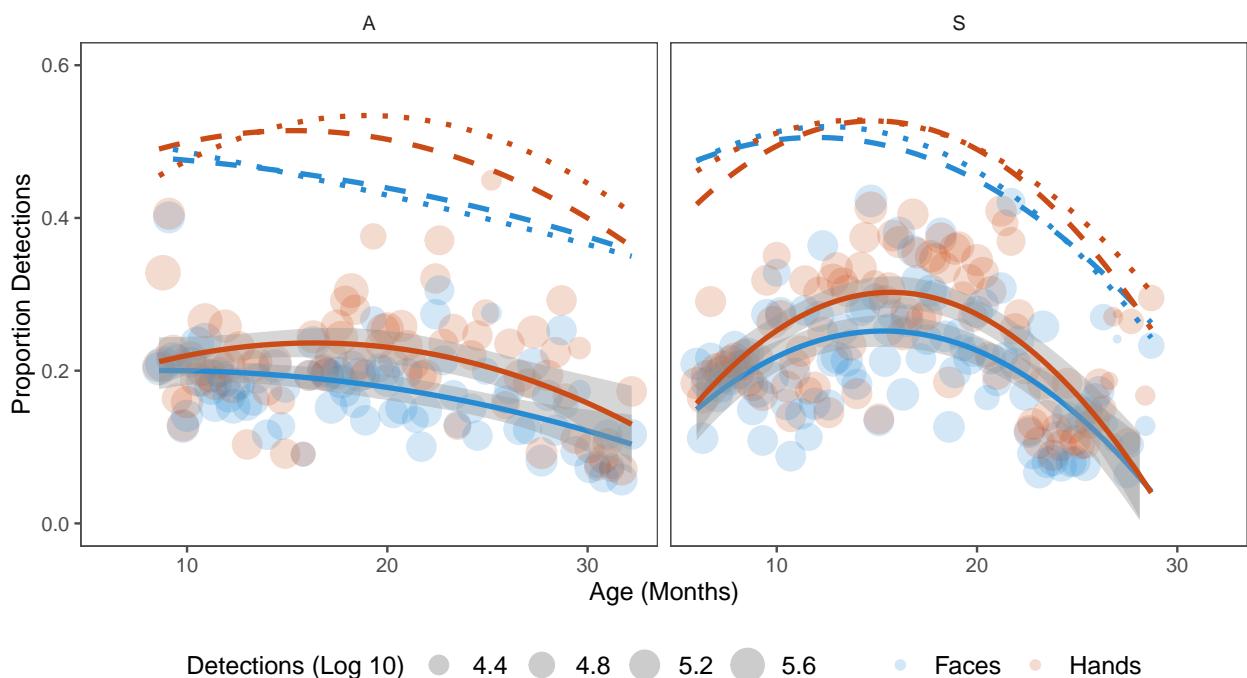
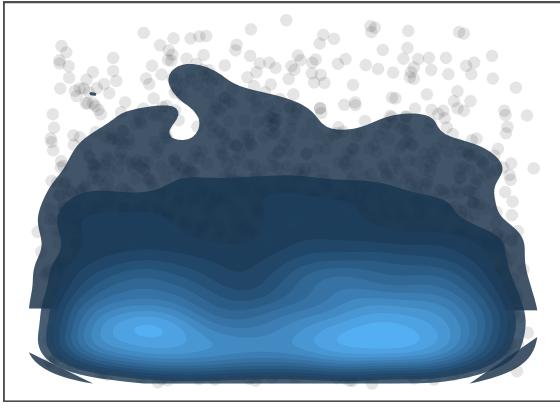


Figure 1: Proportion of faces and hands seen as a function of age for each child in the dataset. Data are binned by each week that the videos were filmed and scaled by the number of frames in that age range. Dashed lines show estimated trend lines from proportion of faces/hands in view when analyzing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

A. Child hand density



B. Adult hand density

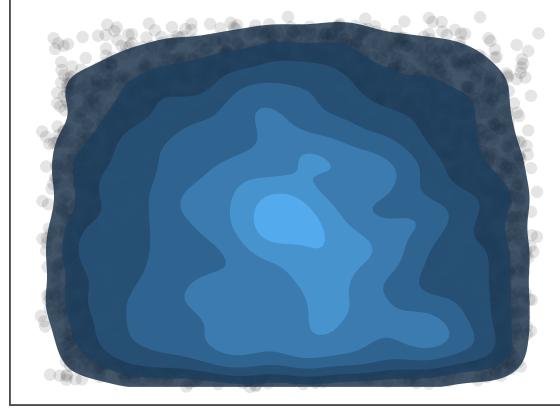


Figure 2: Density estimates for the child (left) and adult (right) hands that were detected in the 24K frame random gold set; each dot represents the center of a bounding box made by an adult participant. Brighter values indicate more detections.

Density of child vs. adults hands in the visual field

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

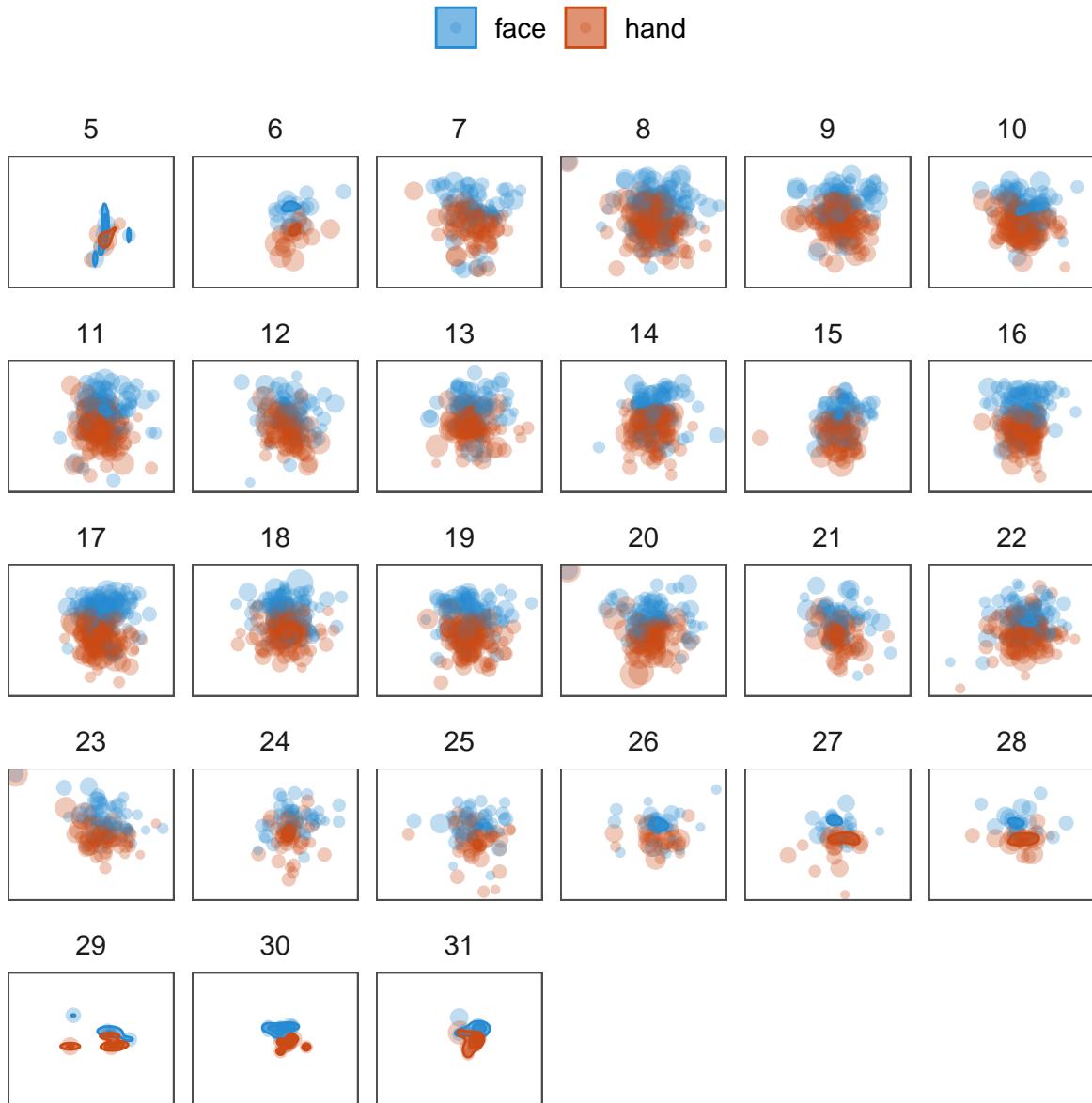
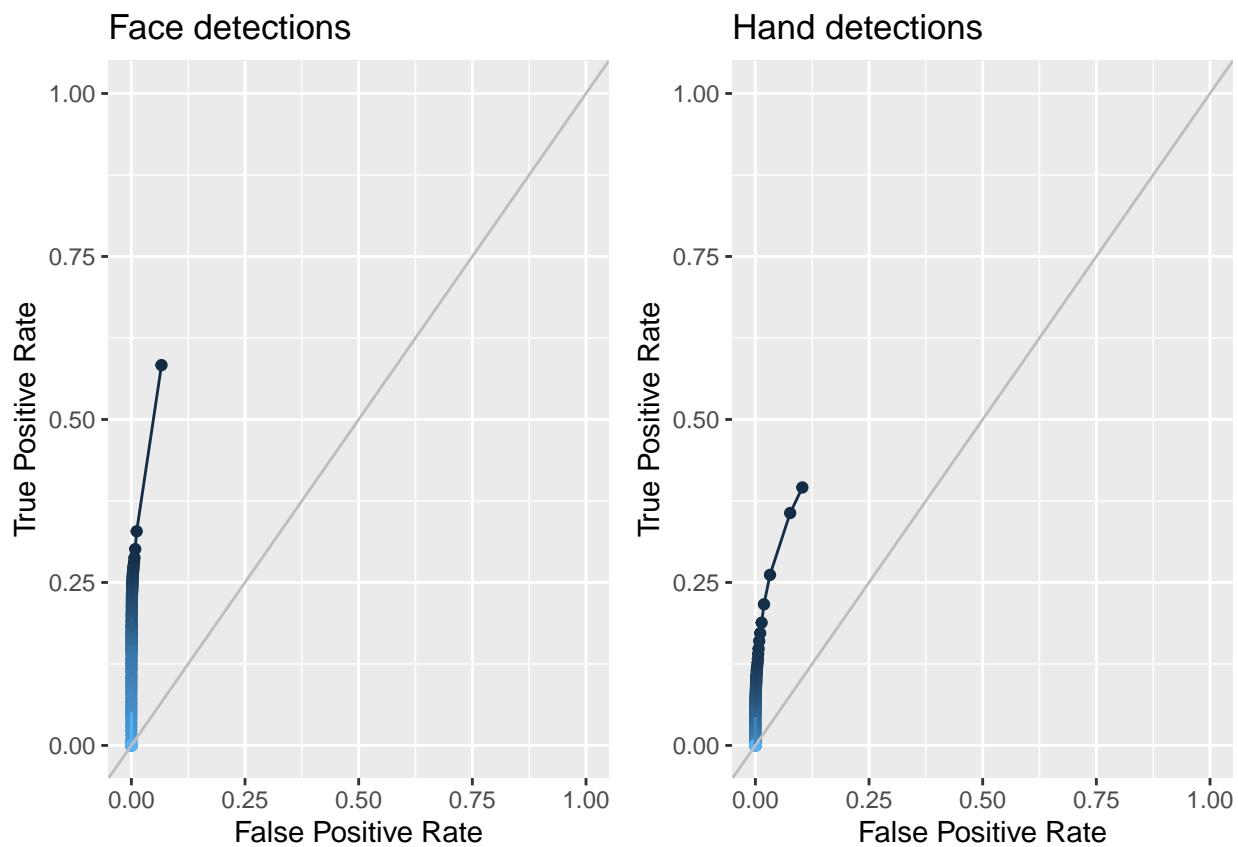


Figure 3: Each panel shows the average position of faces and hands in the visual field in a videos from a given age range, i.e., videos when children in the dataset were 6-31 month-old. Each dot represents the average position from one video within a given age range.

Detection threshold estimations



OpenPose Error Analyses

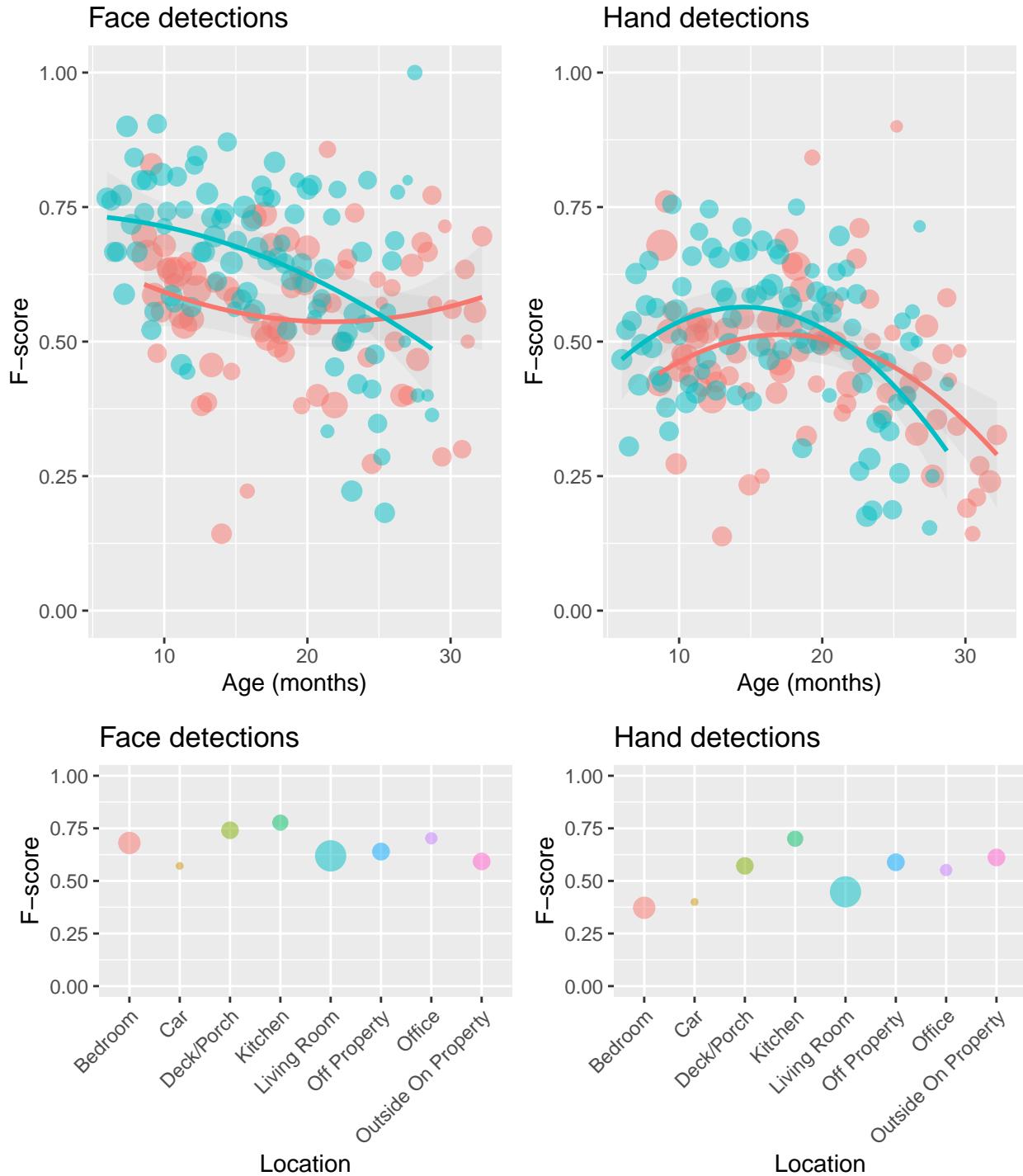


Figure 4: F-score of OpenPose hand and face detections in the gold sample by location. Size indicates number of detections per location: minimum = 55 (car); maximum = 1753 (living room).

OpenPose Annotation Examples

Missed Detections (Faces)



False Alarms (Faces)



Missed Detections (Hands)



False Alarms (Hands)



Figure 5: Example false alarms/misses for faces/hands randomly sampled from the larger set of manual annotations.

Activity by Location

Motor Milestones Analysis Details

We examine whether the SAYcam participants' age of reaching particular developmental milestones (sitting, cruising, and walking) significantly influenced the proportion of hands or faces that they saw, using a mixed-effects regression with per-child random intercepts.

We tried including random slopes by age and milestone, but the models did not converge and/or were singular. Thus, for each milestone the R syntax was: `\texttt{prop_faces ~ age * milestone + (1 | child_id)}`. Children's age of achieving each milestone was dummy-coded (i.e., 0=milestone not reached; 1=milestone reached).

Sitting

Children's age at which they began sitting did not significantly predict changes in either the amount of faces or hands that they saw.

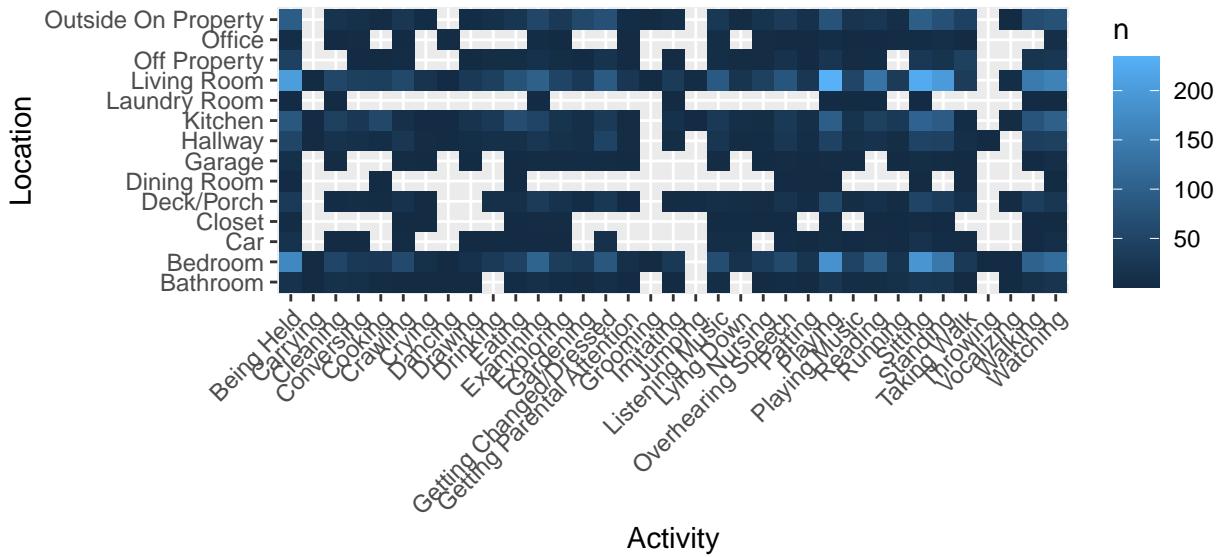


Figure 6: Counts of activities by location, including multilocation videos.

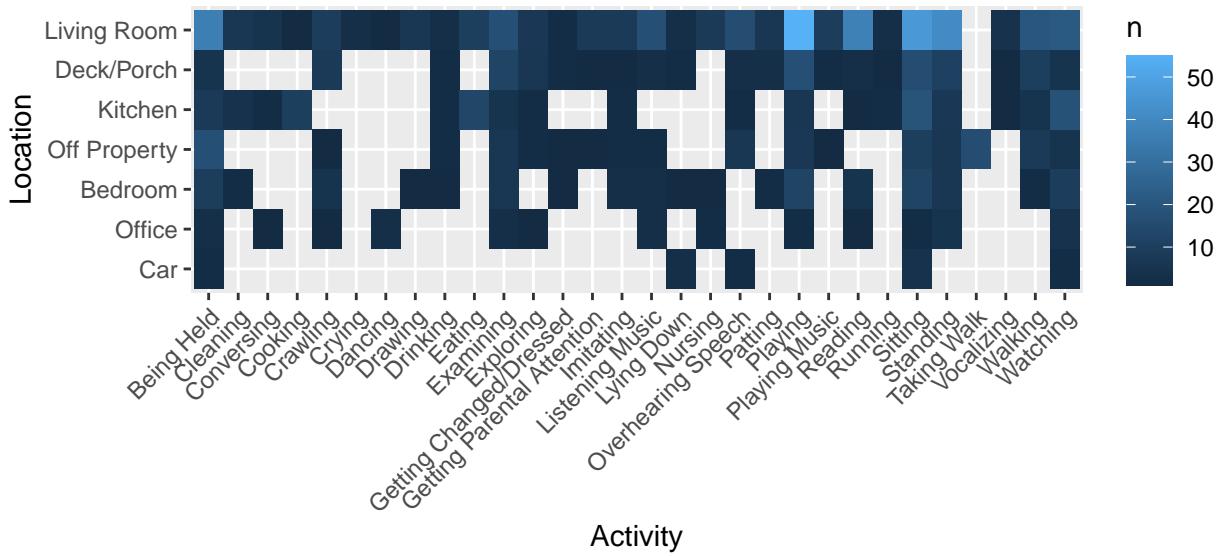


Figure 7: Counts of activities by location in videos taking place in only a single location.

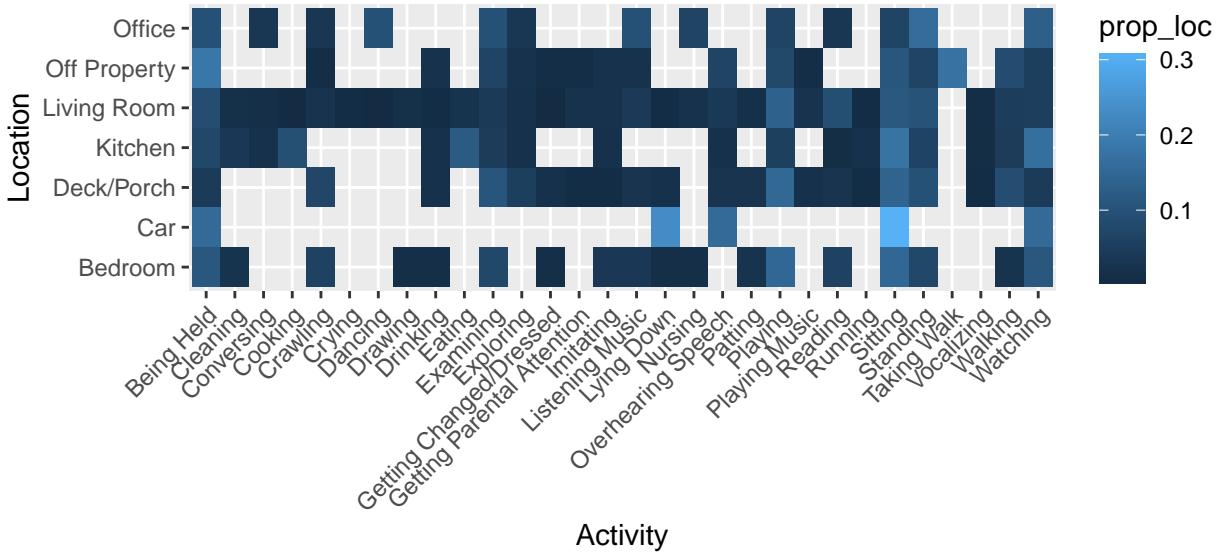


Figure 8: Probability of different activities, given location (in single-location videos).

Table 1: Change in Face and Hand Prevalence by Ability to Sit

	Dependent variable:	
	Proportion of Faces	Proportion of Hands
	(1)	(2)
age	0.176 (0.127)	0.152 (0.142)
sit	1.388 (0.963)	1.231 (1.070)
age:sit	-0.179 (0.127)	-0.154 (0.142)
Constant	-1.133 (0.963)	-0.950 (1.070)
Observations	235	235
Log Likelihood	282.594	257.147
Akaike Inf. Crit.	-553.188	-502.294
Bayesian Inf. Crit.	-532.430	-481.537

Note:

*p<0.1; **p<0.05; ***p<0.01

Cruising

As children became able to cruise, they saw significantly more faces ($\beta = 0.27$) and hands ($\beta = 0.26$). For faces, there was a significant negative interaction of age and cruising ($\beta = -0.03$, $p = 0.025$), and a marginal positive effect of age ($\beta = 0.03$, $p = 0.055$).

Table 2: Change in Face and Hand Prevalence by Ability to Cruise

	<i>Dependent variable:</i>	
	Proportion of Faces	Proportion of Hands
	(1)	(2)
age	0.026* (0.014)	0.023 (0.015)
cruise	0.273** (0.111)	0.258** (0.123)
age:cruise	-0.030** (0.014)	-0.026* (0.015)
Constant	-0.004 (0.110)	0.047 (0.123)
Observations	235	235
Log Likelihood	282.014	257.138
Akaike Inf. Crit.	-552.029	-502.276
Bayesian Inf. Crit.	-531.271	-481.518

Note:

*p<0.1; **p<0.05; ***p<0.01

Walking

As children became able to walk, they saw marginally more hands ($\beta = 0.09$, $p = 0.052$).

Table 3: Change in Face and Hand Prevalence by Ability to Walk

	<i>Dependent variable:</i>	
	Proportion of Faces	Proportion of Hands
	(1)	(2)
age	-0.004 (0.004)	-0.002 (0.004)
walk	0.051 (0.042)	0.091* (0.047)
age:walk	-0.001 (0.004)	-0.003 (0.004)
Constant	0.248*** (0.037)	0.257*** (0.043)
Observations	235	235
Log Likelihood	279.841	256.940
Akaike Inf. Crit.	-547.682	-501.881
Bayesian Inf. Crit.	-526.924	-481.123

Note:

*p<0.1; **p<0.05; ***p<0.01