Appendix A

Face/hand detections relative to human annotations

Chart, diagram

Description automatically generated

*Figure A1*. 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.

Appendix B

Density of child vs. adults hands in the visual field

A picture containing text

Description automatically generated

*Figure B1* . 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.

Appendix C

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.

*Graphical user interface

Description automatically generatedFigure C1*. Each panel shows the average position of faces and hands in the visual field in a video from a given age range, i.e., videos when children in the dataset were 6-31 months old. Each dot represents the average position from one video within a given age range.