The everyday distribution of infant visual ecology

Swapnaa Jayaraman

Indiana University, Bloomington
If you were to teach someone what a *cup* is, how would you do it?
All the cups in the world?

Broad, representative sample?

Or, a few good examples?

What kind of distribution do we learn from?
The distribution of the things in the infant’s ecology is the information on which visual object recognition is built.
How can we describe this visual ecology?
Hi baby!

It’s a duck

Want milk?
Hi baby!

What is the VISUAL availability of objects in everyday life?

It’s a duck
Visual objects that babies learn very early

1. Faces

The very first objects babies learn

Convey information about the emotional and attentional states of social partners
Visual objects that babies learn very early

2. Basic level categories*

Babies start to learn and reliably identify other objects as they grow

* The categories named by early learned nouns
What are their visual properties?

How can we measure these properties?
Sampling from natural visual environments

Study 1* & Study 2*: Faces and hands
Video recordings from infant’s perspective

Study 3: Everyday Objects
Photographs of objects in infant environments

*Developmental studies
Huge advancements in studying adult visual ecology

Why do we need to study it developmentally?

**Structure of visual world for everyone**

Why not?
The developmental perspective

- Prone
- Prone, chest up
- Rolls over
- Sits up
- Crawls
- Stands (with support)
- Pulls to stand
- Cruises
- Stands alone
- Walks alone

Adapted from Adolph, Karasik, & Tamis-LeMonda, 2010
There is not one ecology; there is a changing visual ecology with development.
Study 1 & 2: Faces and hands

- What is in front of baby’s head
- 75 infants
- 0-2 years
- 342 hours of video
- 31 million frames
- 150,000 frames coded
Coding and analysis

Human coders were asked to judge each image:

“Do you see a face or a face part in this image?”

And separately:

“Do you see a hand in this image?”
Study 1: Faces in early visual environments

Early visual experiences dense with face information

$R^2 = .42, F(1, 20) = 16.11. p < .001$
We also coded

Identities of faces (whose face)

Distance of face from infant

The view of face (frontal, profile)
Identities: the 3 most frequent people accounted for almost all the faces for the very young infants

\[ R^2 = .14, F(1, 20) = 4.516. p < .05 \]
\[ R^2 = .16, F(1, 20) = 5.24. p < .05 \]
\[ R^2 = .23, F(1, 20) = 7.413. p < .05 \]
Distance: For very young infants most faces are close to them (within 2 feet)

Consistent with their visual abilities

$R^2 = .37, F(1, 20) = 13.61, p < .05$
Views: For everybody, most faces present frontal views
Ecology of faces in early environments

Matters to understanding visual object recognition and development

1. **The ecology changes**
   (not developmentally stationary sampling of the world)

2. **Prototypical faces** densely populate the visual world of young infants:
   Close, frontal views, of very few people, and lots of face time

   *This is the information on which the earliest visual development feeds*

3. **Variation around this prototype expands**
   – incrementally – over the first year
Study 2: Hands in early visual environments
Lots of faces early, but it’s not just about people nearby.
Of Hands, Own Hand

Across all ages, of all hands:

76% were touching an object

48% were holding an object

Developmental shift from faces to manual actions on objects
Study 3: Everyday objects
Early faces present a distribution in which a few individuals dominate.

Do objects present a more uniform distribution?
Early evidence from a pilot study

Just 8 categories
No head camera
Parent took photo when someone said one of the 8 words

Where is her cup?

12 mo, n=10
Object categories

Cup  Duck
Hat  Car
Chair  Airplane
Dog  Donkey

How many distinct instances in a week?
The everyday distribution

How many tokens of each category?

<table>
<thead>
<tr>
<th>Category</th>
<th>Median Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>airplane</td>
<td>25</td>
</tr>
<tr>
<td>car</td>
<td>20</td>
</tr>
<tr>
<td>chair</td>
<td>15</td>
</tr>
<tr>
<td>cup</td>
<td>10</td>
</tr>
<tr>
<td>hat</td>
<td>5</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
</tr>
<tr>
<td>dog</td>
<td>20 dogs in a week</td>
</tr>
<tr>
<td></td>
<td>5 cups in a week</td>
</tr>
</tbody>
</table>

Median Frequency across all children
Of the sampled distribution, what proportion is the most frequent token?

- 5 cups in a week mostly about 1 cup
- 20 dogs in a week mostly about 1 dog
Objects: the everyday distribution

Not unique to faces!
Common theme for visual statistics of objects in natural environments: Zipfian distribution

The same with natural language corpuses, and other data from physical and social sciences
Infant visual ecology – what have we learned?

1. The contents and properties change with development, in systematic ways. Though not shown yet, they are likely to influence visual development.

2. Faces present prototypical dense views early in infancy: close, big in the view, of a very few people, and show frontal views with both eyes. These properties may be critical to building specialized face processing.

3. Hands are ubiquitous and in contact with objects and come to dominate early in the second year. Hands direct visual attention and convey causal information to babies. These early hands on objects are the data on which that knowledge is likely built.

4. Faces are not special in their Zipfian distribution. Early object experiences also start with distributions in which a very few individual instances dominate. Early visual development is built on this everyday distribution.
There is much more to do
Universality of infant environments
Low level visual environments

Intensity

RMS Contrast
Optic flow
Figure 1 shows example flow fields (a,b), normalized speed histograms (c-j), and normalized motion direction histograms (k-s). Speed bin widths are 2.25 deg/s; Direction bin widths are 22.5 deg.
### U.S.
- Median: 7.1 s
- Mean: 12 s

### India
- Median: 8.3 s
- Mean: 14.4 s

### Segment Duration (min)

<table>
<thead>
<tr>
<th>Moving</th>
<th>Stationary</th>
</tr>
</thead>
</table>

### a.
- Static: 025TT, $r^2 = 0.95$
- Moving: 025TT, $r^2 = 0.92$

### b.
- Static: 053KD, $r^2 = 0.99$
- Moving: 053KD, $r^2 = 0.86$

### c.
- Static: 025TA, $r^2 = 0.99$
- Moving: 025TA, $r^2 = 0.94$

### d.
- Static: 054AJ, $r^2 = 0.82$
- Moving: 054AJ, $r^2 = 0.92$
# Time sampling study

<table>
<thead>
<tr>
<th>Age</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Median</th>
<th>Mode</th>
<th>Range</th>
<th>Total*</th>
<th>Prop. of times when at least 1 person was present</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 months</td>
<td>2.17</td>
<td>1.38</td>
<td>2.2</td>
<td>1.9</td>
<td>0.68 - 4.10</td>
<td>44.0</td>
<td>.961</td>
</tr>
<tr>
<td>6 months</td>
<td>1.85</td>
<td>0.95</td>
<td>1.9</td>
<td>1.8</td>
<td>1.00 - 3.14</td>
<td>39.2</td>
<td>.997</td>
</tr>
<tr>
<td>12 months</td>
<td>2.41</td>
<td>1.57</td>
<td>2.2</td>
<td>2.2</td>
<td>1.06 - 3.88</td>
<td>56.5</td>
<td>.970</td>
</tr>
<tr>
<td>18 months</td>
<td>2.37</td>
<td>1.13</td>
<td>2.3</td>
<td>2.1</td>
<td>1.11 - 4.00</td>
<td>56.0</td>
<td>.992</td>
</tr>
<tr>
<td>24 months</td>
<td>2.30</td>
<td>1.42</td>
<td>2.3</td>
<td>2.4</td>
<td>1.00 - 4.22</td>
<td>55.4</td>
<td>.954</td>
</tr>
</tbody>
</table>
Head camera study
Thank you!

Collaborators

Linda Smith  
*Indiana U*

Rick Gilmore  
*Penn State*

Florian Raudies  
*HP Labs*

Caitlin Fausey  
*U Oregon*

Char Wozniak  
*Indiana U*

Ariel La  
*U Wisconsin*

Technicians

Participants

Research assistants  
Braden King, Joseph Imbrock, Shruthee Rajendran, Stephanie Krupa, Megan Gibbons, Remy Snead, Edward Rach, Minyi Zheng, Connor Courtney, Olivia Haas, Alexis Jones, and Afiah Hasnie

Funding sources  
NICHHD 28675, NIH NRSA HD007475 - 18, NIH R21HD068475
Extra slides
Face time

Frequent faces
Displaying both eyes

Close-up faces
Early multimodal environments
Arrival rates of faces

Considering infants in each age group as one *super-baby*
Measuring distance of face from infant

2 ft

4 ft
Three signature properties

Belong to a few individuals
Appear close to the infant
Display both eyes

Dad (one of top 3 identities)
6 feet away
Doesn’t display both eyes

Cluster score = 1

Dad (one of top 3 identities)
1.5 feet away
Displays both eyes

Cluster score = 3
Proportion of faces with high (2 or 3) scores

$R^2 = .54$, $F(1, 20) = 23.28$. $p < .001$
Cumulative waking hours and face hours

Age in months