Figure 1: (A) shows an example of text with features and categories annotated by BERT, (B) shows how the co-occurrence is calculated for an example annotated sentence, "The green frog croaks", and (C) shows the entries in the co-occurrence matrix after having processed the first sentence.
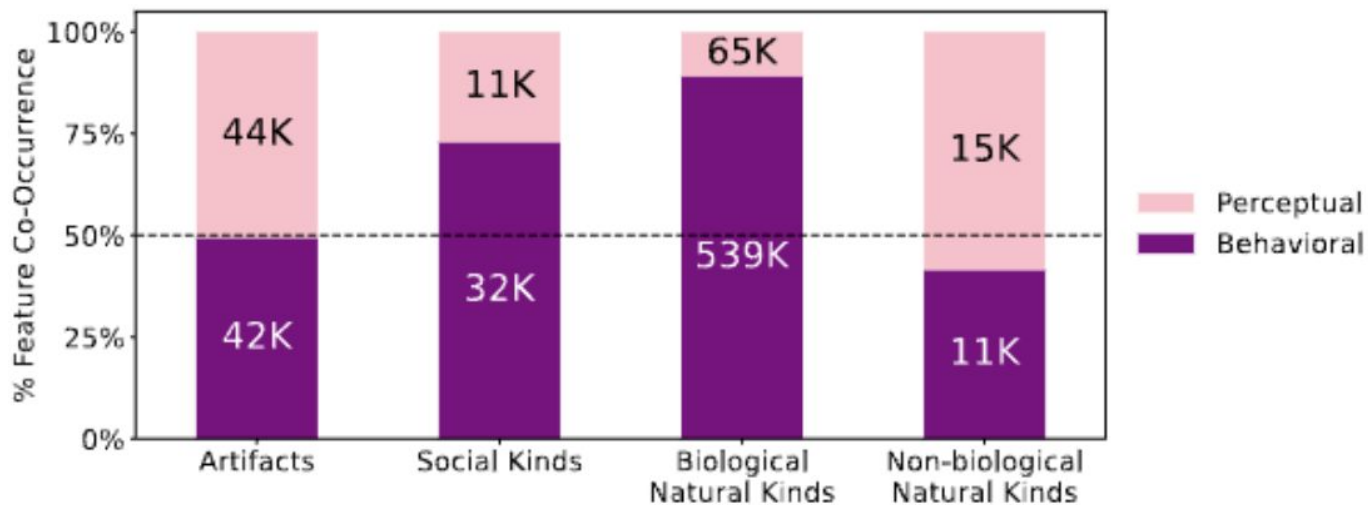
Figure 2: Percentage of co-occurring perceptual (pink) versus behavioral (purple) words for artifacts, social, biological, and non-biological kinds.

Table 1: Overall top 5 words for each feature and kind. The total word count across all corpora is included for each word.

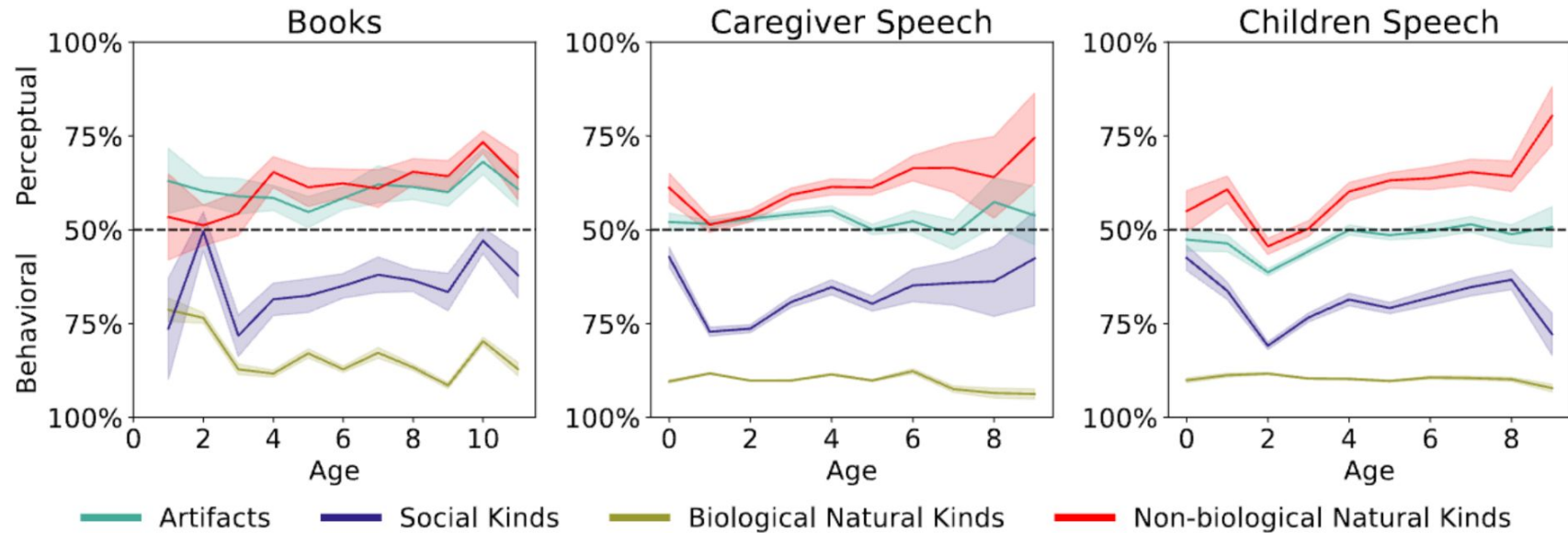| Perceptual	Behavioral	Artifacts	Social kinds	Biological Natural Kind
one: 57222	go: 101032	house: 19391	mother: 57643	baby: 27319
little: 32429	get: 80659	book: 18307	father: 27418	hand: 8444
big: 21616	see: 65008	car: 12800	boy: 9073	dog: 8351
two: 20287	look: 56378	toy: 7538	girl: 8336	cat: 7799
three: 9905	put: 48150	cookie: 7515	friend: 4394	bear: 7373

Figure 3: Change in relative percentage of co-occurrence between perceptual (top) and behavioral (bottom) features across age for artifacts (green), social kinds (purple), biological natural kinds (yellow), and non-biological natural kinds (red). For example, for non-biological natural kinds, the percentage of co-occurence with perceptual features increases with age. *Note:* In all of the figures, the ribbons show 95% bootstrapped confidence intervals.

Table 2: Bayesian regression models of the effect of age on each category for each source. *Note*: We always report the mean of the posterior estimate together with the 95% credible intervals in brackets.

	Books	Children's speech	Caregivers' speech
Artifacts	0.003 [−0.004, 0.011]	0.007 [−0.001, 0.016]	0.001 [−0.006, 0.008]
Social Kinds	0.009 [−0.009, 0.028]	−0.004 [−0.023, 0.015]	0.009 [−0.007, 0.027]
Biological Natural Kinds	−0.010 [−0.023, 0.002]	−0.001 [−0.004, 0.001]	−0.004 [−0.009, 0.001]
Non-biological Natural Kinds	0.016 [0.007, 0.025]	0.024 [0.007, 0.041]	0.018 [0.007, 0.028]

Table 3: Top 5 words for each non-biological natural kind subcategory.

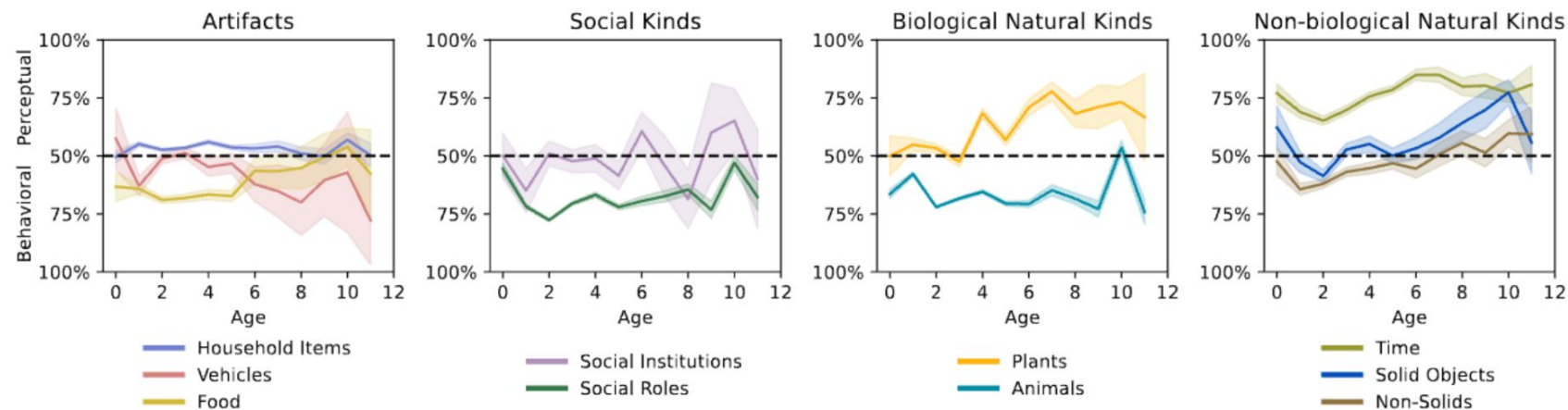| Time		Non-Solids		Solid Objects	
time:	11491	water:	9202	moon:	3072
day:	8116	light:	3082	ground:	1061
night:	7411	sun:	2532	rock:	1038
today:	5717	snow:	1684	hill:	798
minute:	4536	sky:	1427	mountain:	655

Figure 4: Change in relative percentage of co-occurrence between perceptual and behavioral features over age for subcategories of artifacts, social kinds, biological natural kinds, and non-biological natural kinds.
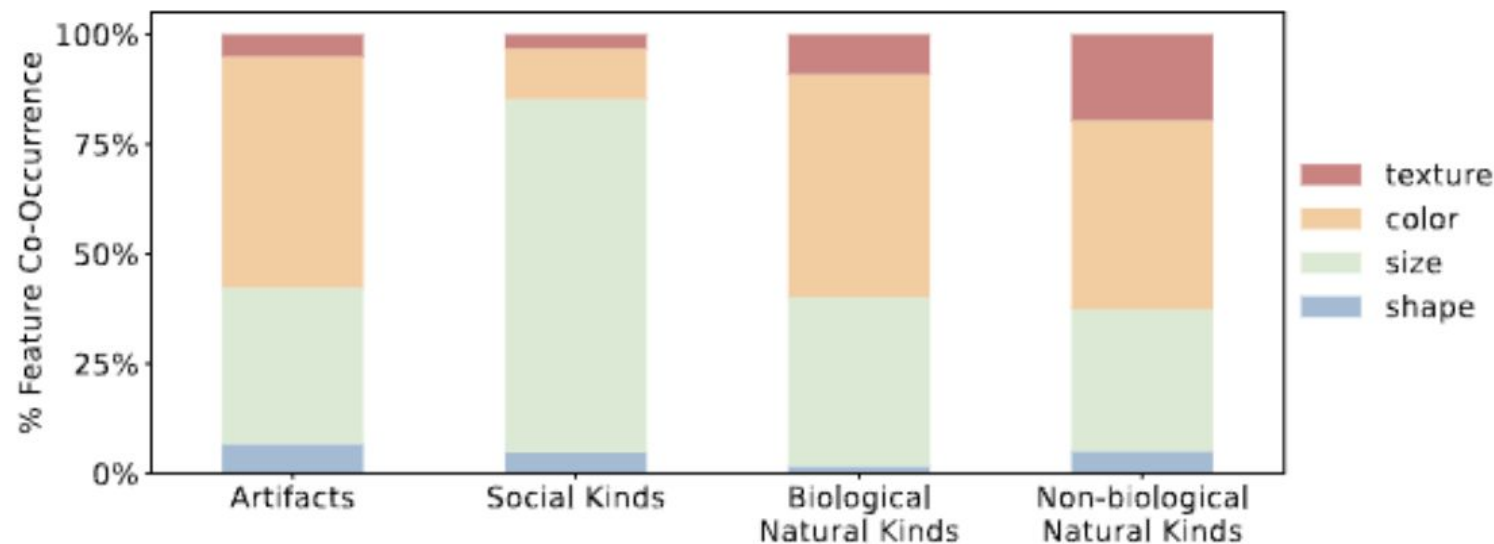
Figure 5: Percentage of co-occurring perceptual subcategories for Artifacts, Non-Biological Natural Kinds, Biological Natural Kinds, and Social Kinds.

Table 4: Top 5 words for each perceptual subcategory.

Size		Color		Texture		Shape	
big:	21616	red:	7884	hard:	3061	round:	1052
small:	2523	blue:	6851	wet:	2026	square:	909
tiny:	1547	green:	6588	soft:	1361	triangle:	379
huge:	442	yellow:	4701	dry:	753	rectangle:	184
large:	429	orange:	3725	sticky:	746	circle:	158