Developmental changes in the semantic part structure of drawn objects

Holly Huey
Department of Psychology
UC San Diego
hhuey@ucsd.edu

Bria Long
Department of Psychology
Stanford University
bria@stanford.edu

Justin Yang
Department of Psychology
UC San Diego
juy003@ucsd.edu

Kaylee George
Department of Psychology
Stanford University
krgeorge@stanford.edu

Judith E. Fan
Department of Psychology
UC San Diego
jefan@ucsd.edu

Abstract

Children produce increasingly more recognizable drawings of object concepts throughout childhood. What drives this improvement? Here we explore the role of children's ability to include relevant parts of those objects in their drawings. We crowdsourced part tags for every pen stroke in 2,160 drawings of 16 common object categories that had been produced by children between 4 and 8 years old. These part decompositions revealed both substantial variation in the number and kind of parts that children emphasized, as well as a non-monotonic relationship between the number of parts that children drew and how recognizable their drawing was. Taken together, our findings advance current understanding of how children's part compositionality in their drawings both drives and constrains recognition of depicted objects. We plan to publicly release these data to catalyze further investigation of how children's drawings change across development.

Keywords: object representations; child development; compositionality; drawing interpretation; visual communication

Introduction

Children's drawings of object concepts change dramatically and progressively across development (Piaget, 1929; Kellogg, 1969; Karmiloff-Smith, 1990; Fury, Carlson, & Sroufe, 1997), beginning with seemingly meaningless scribbles but eventually emerging into rich and recognizable depictions of the objects in the world around them. How do children learn to translate their knowledge of object concepts into structured, semantically meaningful drawings? Researchers have long hypothesized that this protracted representational ability of children to depict objects—including their constituent parts and their relational structure—reflects a core correspondence between children's visual perception and the organization of their semantic knowledge (Minsky & Papert, 1972). Indeed, children's visual production abilities evolve in tandem with their maturing perceptual systems (Bova et al., 2007; Natu et al., 2016) and increasingly rich semantic knowledge about the features of object categories (Tversky, 1985).

Recently, Long, Fan, Chai, and Frank (2021) investigated how children produce and recognize drawings of common object categories across childhood. Children progressively included diagnostic features of those objects in their drawings and capitalized on those same features when recognizing each other's drawings. Further, children's ability to produce increasingly recognizable drawings was not fully explained by their increasing visuomotor control, hinting at a much more extended development of children's visual concept learning than previously thought. However, this work leaves open key questions about how children's visual concepts are changing—and about the actual features of children's drawings that drive this improvement in their recognizability. In the present paper, we examine the part structure in the drawings collected in Long et al. (2021) by crowdsourcing semantic labels for the parts of objects included in children's drawings.

Mounting literature using drawings to explore child development has established that children's drawings improve along several dimensions, becoming more visually distinctive across different object categories (Long, Fan, & Frank, 2018), more inclusive of diagnostic properties (Sitton & Light, 1992; Barrett & Light, 1976; Bremner & Moore, 1984) and relations (Light & Simmons, 1983), and more spatially structured (Karmiloff-Smith, 1990). In particular, these findings often highlight that as children's knowledge of object concepts improve, so too do their drawings, converging on the notion that children are systematically biased to draw what they know rather than what they see (see "intellectual realism" in Luquet (1927)). However, while these prior studies have documented informative gains in children's representational abilities, they have yet to characterize a critical axis of variability that may explain the changing recognizability of children's drawings: an explicit examination of the parts of objects that children include in their drawings, and how variability in such visual prioritization may impact recognition of the drawn object.

Investigating children's knowledge about parts as a window into the structure of their developing semantic knowledge is itself not a new approach (Tversky, 1989). Because objects can be parsed into their constituent parts based on both the perceptual and functional salience of those parts, "*partonomic knowledge*" is theorized to underscore how people conceptually represent those objects (Tversky & Hemenway, 1984). Case-study evidence demonstrate that children increasingly draw more diagnostic parts of objects (e.g., handles on mugs, see Bremner and Moore (1984)) and people (Chappell & Steitz, 1993) as they learn how to discriminate what characterizes those categories. These findings resonate with work showing that children's recognition abilities cooccur with their ability to visually process the part configurations of objects (Juttner, Muller, & Rentschler, 2006; Juttner, Wakui, Petters, & Davidoff, 2016). Although these studies offer initial insights on how children express their semantic

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

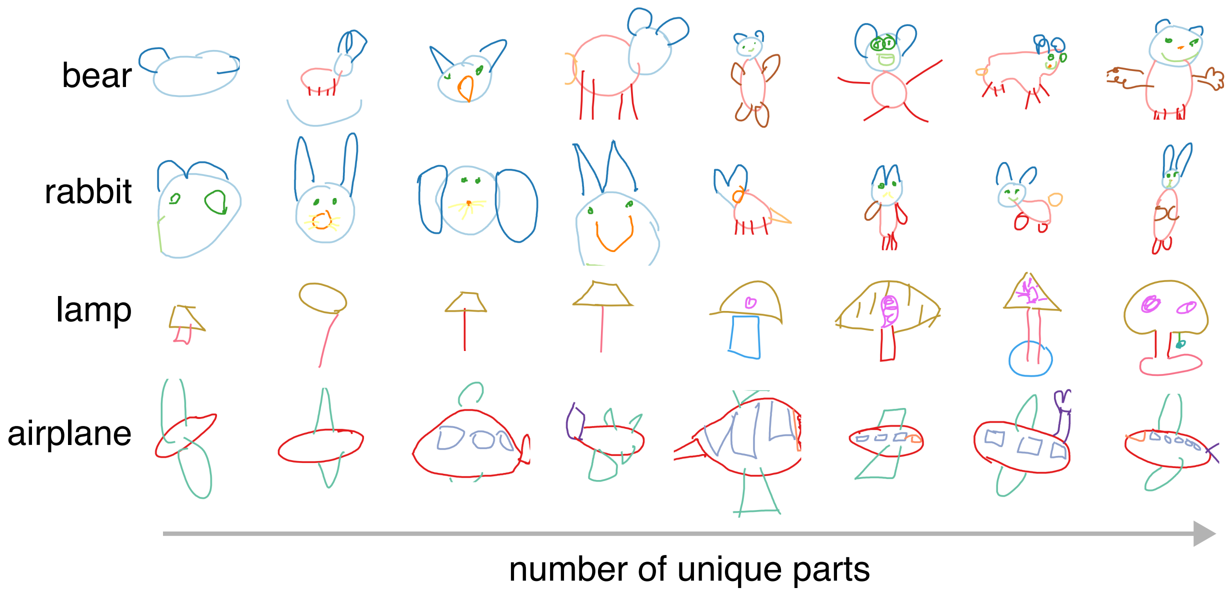

Figure 1: Example drawings across children from four object categories, increasing in the number of drawn unique parts. Colors added for illustrative purpose to show each part annotation attributed to drawings.

knowledge in visual form, they have only been able to do so for targeted object categories (e.g., mugs, people) and often with a narrow range in the ages of children. In part, this gap is driven by a limited ability to capture the salient parts of children's drawings: prior studies have often relied heavily on manual annotations for singular categories (Bremner & Moore, 1984; Barrett & Light, 1976), which has inherently limited the sample sizes necessary to explore large-scale developmental trends in drawing behavior; other studies have employed modern deep learning models to characterize feature distinctiveness of drawings (Yamins et al., 2014; Fan, Yamins, & Turk-Browne, 2018; Long et al., 2021) but which cannot explicitly encode part-level recognition of objects.

We overcome these challenges by developing a drawing annotation paradigm to develop quantitative estimates of the parts of objects that children drew in the large-scale drawing dataset collected by Long et al. (2021). Here we investigate: (1) what part information children choose to include in their drawings of common object categories; (2) whether older children include more object parts in their drawings; and (3) how variation in the number of depicted object parts impacts the ability of other children to recognize drawn objects. By measuring the semantic changes in children's prioritization of different object parts, as well as the downstream recognition they support, our work advances current understanding of how children progressively improve in their ability to convey semantically meaningful information in visual form.

Methods

Developmental drawing dataset

Drawings We first obtained line drawings of common object categories previously collected by Long et al. (2021), in which young children played a drawing game on a free-standing kiosk of a local children's museum. These categories included a variety of animals and inanimate objects, and were chosen from categories that are often drawn by children (e.g., dog, cup) and also less frequently drawn (e.g., camel, lamp). On each trial, children were verbally prompted to depict an object and had 30 seconds to draw on a touch-screen canvas of the kiosk with their finger. The drawings were encoded as a sequence of vector graphic "strokes" and parameterized by a sequence of cubic Bezier curves (i.e., splines). Here we define a stroke as the mark left by a child's finger between being "placed onto" onto the digital canvas of the tablet and "lifted up" from the canvas.

From the larger dataset, we annotated 2,160 drawings of 16 categories that were produced by 4 to 8-year-old children (N=1481 children), including 560 drawings of small animals (rabbit, dog, fish, bird); 600 drawings of large animals (camel, tiger, sheep, bear); 500 drawings of vehicles (airplane, boat, car, train); and 500 drawings of small household objects (bottle, lamp, hat, cup). We chose these drawings to annotate, because this subset was additionally used to test a different set of children's recognition of the drawn objects (see below).

Drawing recognizability The original dataset also included recognizability scores for each drawing, generated by a separate set of children (N=1,789 children, aged 3-10 years, M_{age} = 5.49 years). These recognizability scores were generated from four different "guessing games" for each broad object category (small animals, large animals, vehicles, household objects). Children were presented with a drawing on the same museum kiosk, and asked to guess which of four object categories each drawing represented (Long et al., 2021). Children indicated their guess using touchscreen buttons with

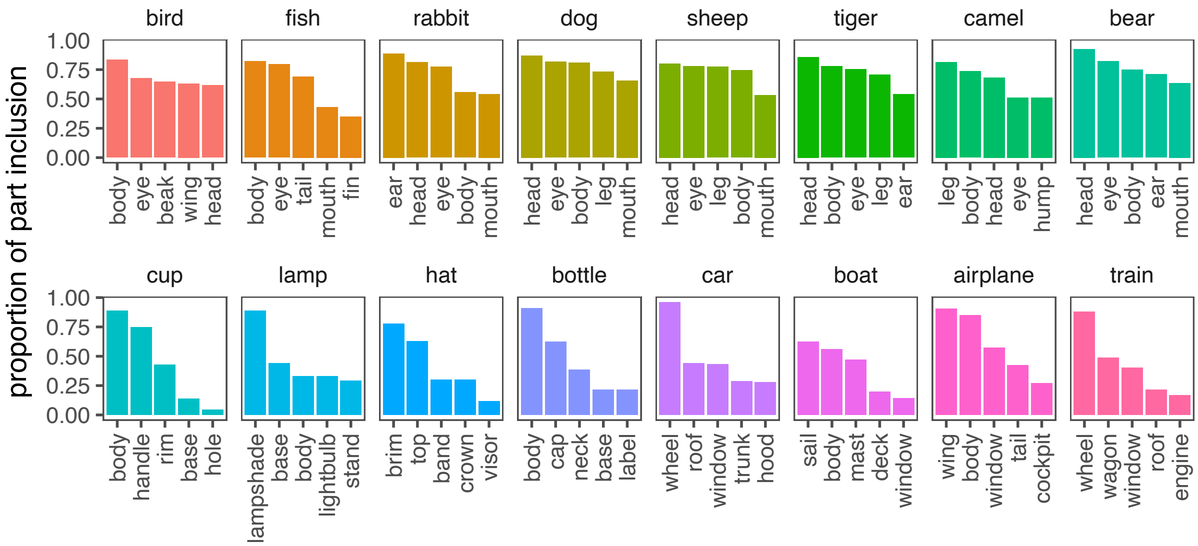

Figure 2: Distribution of the five most frequently drawn object parts in each object category.

canonical photographs of each possible category. All distractors were from the same broad category (e.g., vehicles).

Crowdsourced part decompositions

Our next goal was to generate a comprehensive set of part decompositions for each of the 16 object categories. To accomplish this, we designed a web-based crowdsourcing platform and recruited 50 English-speaking adult participants (25 female; M_{age} = 27.1 years) from Prolific to identify the basic parts of objects for each category. All participants of this and the following task provided informed consent in accordance with University of California San Diego's IRB.

On each trial, participants were cued with a text label of an object category and asked to list 3 to 10 object parts that came to mind (e.g., head, leg, tail, etc. for "tiger"). Participants were instructed to write only concrete parts of an object (e.g., "tail") rather than abstract attributes (e.g., "tufted"), to use common names of parts rather than technical jargon (e.g., "prehensile"), and to generate as complete a part list as they could for each object category. We applied lemmatization to the resulting part decompositions to remove redundant part labels, such as "hoof" and "hoofs", and manually edited part labels that were spelled incorrectly or with alternative spellings. We then selected the top 10% of part names that were most frequently listed. This generated a total of 82 object parts with a range of 5-13 possible parts per object category.

Semantic part annotation task

To systematically measure the semantic part information conveyed in children's drawings, we developed a web-based annotation paradigm adapted from previous research (Mukherjee, Hawkins, & Fan, 2019; Huey, Walker, & Fan, 2021) to assign a part label to each pen stroke constituting every drawing (Fig. 1).

Annotators 1,034 English-speaking adult participants (457 female; M_{age} = 33.9 years) were recruited from Prolific and completed the semantic annotation task. We excluded data from 78 additional participants for experiencing technical difficulties with the web interface (N=11) and for having low accuracy on our attention-check trial (N=67). Data collection was stopped when every drawing had received annotations from at least three annotators.

Procedure Each annotator was presented with a set of 8 drawings randomly sampled from the drawing dataset but consistent within the same broad object category (i.e., small animals, large animals, vehicles, small household objects). Each drawing was accompanied by the name of its object category (e.g., "airplane"), as well as a gallery of crowdsourced part labels that corresponded to it. For each stroke in the presented drawing, annotators were prompted to tag it with the part label that described the part of the depicted object that it represented. Annotators were permitted to label a stroke with multiple part labels if they believed a stroke to represent multiple different parts of the depicted object, and were able to write their own custom label if they believed that none of the provided part labels were fitting. They could also label a stroke as unintelligible if they could not discern what it represented. Annotators also completed an "attention-check" trial, consisting of a pre-selected drawing that had been annotated by a researcher and then randomly inserted into the set of drawings. If annotators did not match the researcher's annotation criteria for this drawing, data sessions from these annotators were excluded from subsequent analysis.

Data preprocessing To determine how often annotators agreed on what each stroke of children's drawings represented, we calculated the inter-rater consistency among annotators. Across drawings, annotators agreed on the same part label for 69.9% of strokes and had modest improvement

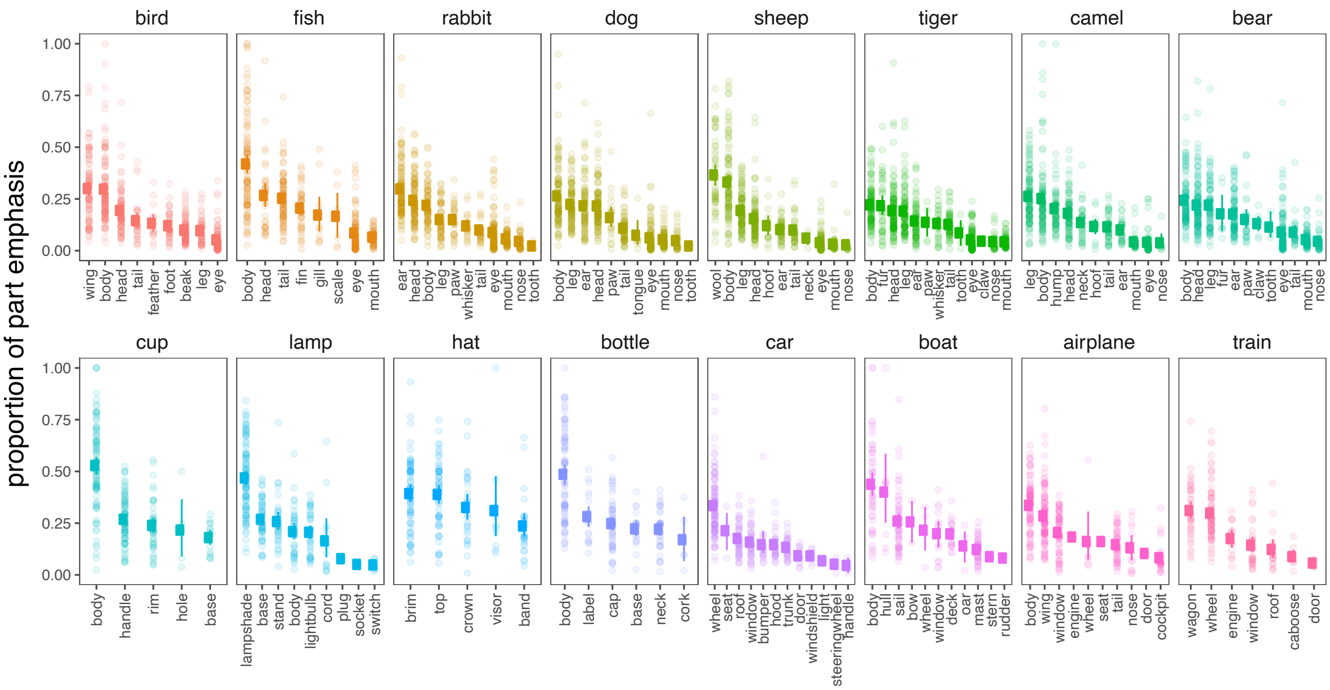

Figure 3: Proportion of part emphasis (i.e., amount of ink allocated to drawn object parts) in each object category. Error bars represent 95% bootstrapped CIs.

with drawings produced by older children (4-year-old drawings = 68.3% mean agreement, 8-year-old drawings = 69.8% mean agreement). We retained strokes that were assigned the same part label by at least two of three annotators. Our resultant dataset therefore contained 14,159 annotated strokes across 2,088 drawings. Additionally, although we found that annotators infrequently wrote custom labels, they only used 68 of the available 82 part labels. Strokes that were labeled as unintelligible were not counted as distinct parts in analyses.

Results

Which parts of objects did children prioritize in their drawings? To investigate which parts were included and visually emphasized in children's drawings, we calculated how often distinct object parts were included in the drawings. We then evaluated their visual emphasis by calculating the proportion of the total length of strokes that were attributed to a particular object part in a drawing (e.g., wing), relative the total length of all strokes in the entire drawing. If strokes were used to represent multiple object parts, we took the total length of the stroke and divided it by the number of parts that it represented.

Overall, children prioritized certain parts of object categories over others, by both including more object parts and devoting a greater amount of strokes to them (see Fig. 3 for the five most frequently included parts per object category). For example, children were more likely to prioritize "ears" in their depictions of rabbits and "wheels" in their depictions of cars. In drawings of animals, children tended to include more generic bodily features such as "body" and "head", although intriguingly some categories tended to emphasize parts that were uniquely characteristic of those animals, such as "wing" for birds and "wool" for sheep. By contrast, other generic facial features (e.g., "eye", "mouth", "nose", "tooth") were less frequently emphasized. In drawings of household objects and vehicles, children also tended to include generic features like "body" and prioritize parts that were uniquely characteristic of those categories (e.g., "lampshade" in lamp, "wagon" on train).

Do older children draw more object parts? Across object categories, we found that the proportion of children's drawings consisting of strokes labeled as unintelligible decreased as children matured (4-year-olds = 29.7% of strokes, 8-year-olds = 12.4% of strokes); this result corroborates general prior findings that older children produce more recognizable drawings (Long et al., 2021). Additionally, we found that older children were more likely to use a single stroke to represent multiple parts of an object (4-year-olds = 0.192% of strokes, 8-year-olds = 0.274% of strokes), suggesting that children gradually become skilled at producing more complex and efficient representations of object parts.

To evaluate how the number of uniquely identifiable parts attributed to each object category varied by age, we fit a linear mixed effects model predicting the number of object parts from age, including random intercepts for each object category. We found a reliable increase in the number of depicted parts across age for all object categories ($B=0.395$, $SE = 0.041$, $df = 2071$, $p<.001$), demonstrating that children gradually produce more detailed depictions by including

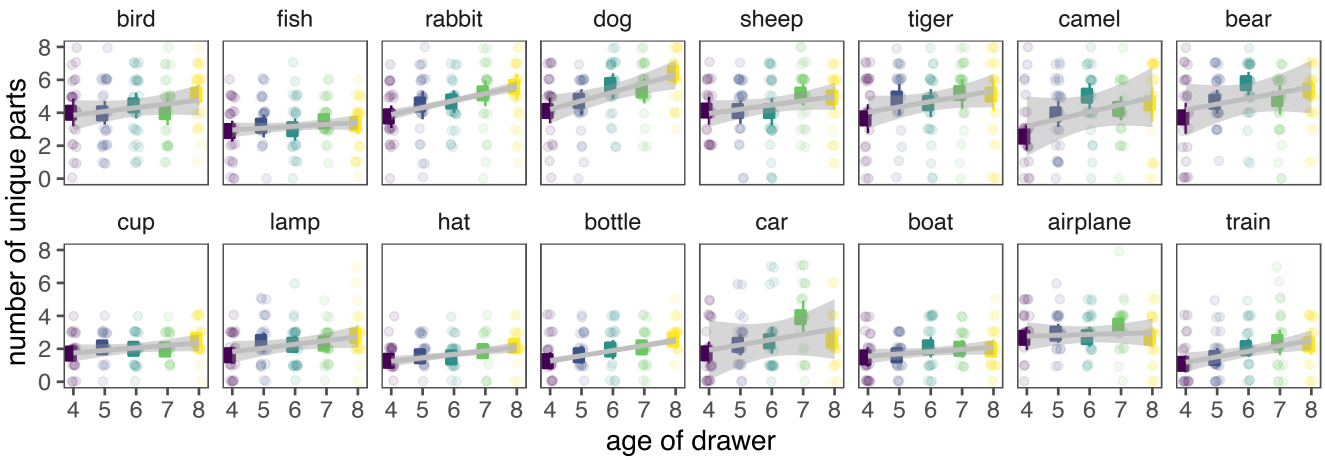

Figure 4: Number of unique parts per each object category across age

more object parts across development. Nonetheless, we observed wide variation in the number of object parts depending on the category of the drawn object, revealing that children tend to draw a greater number of object parts for animals relative to inanimate objects (e.g., household objects, vehicles) (Fig. 4).

Do more recognizable drawings contain more object parts? To test how variability in the number of drawn object parts may predict how well naive viewers can recognize the depicted objects, we leveraged the recognition scores previously generated by a second set of children in Long et al. (2021). We first examined children's recognition of each object category as a function of the number of unique parts included in each drawing (Fig. 5). We found that drawings with both too few and too many object parts tended to be less recognizable to children, relative to drawings with an intermediate number of unique parts. We additionally observed wide item variability both between and within broad object categories (i.e., small animals, large animals, vehicles, household objects), in which some categories tended to have more variation in the number of depicted parts and variation in the relative strength of a quadratic relationship to their recognizability (Fig. 5). For example, the quadratic relationship was strongest for object categories like lamps (see row 3 of Fig. 1 for examples, where more sparse drawings of lamps were less identifiable and more detailed drawings could be mistaken for mushrooms). On the other hand, drawings of object categories like bears that were more enriched with object parts were more recognizable to naive children (see row 1 of Fig. 1 for examples). This variability in recognizability suggests that as children gain more semantic knowledge object categories, they may need to adjust how many parts of object they should choose to include in their drawings in order to produce more recognizable drawings.

We next examined the relationship between the number of unique object parts and age of the child recognizing the drawings and found a similar quadratic relationship between the two variables. In order to assess these trends, we fit a generalized linear mixed effect model to the recognition data, modeling the quadratic interaction between the number of parts in each drawing and the age of the child recognizing the drawing, with the maximal random effects structure possible. We found a significant main effect of unique parts, confirming the observation that drawings with a greater number of parts tend to be better recognized. In addition, we found significant interactions between both terms of the polynomial and the age of the recognizer, suggesting that children's ability to integrate this extra part information during recognition changed across development (see all coefficients below in Table 1).

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	0.05	0.23	0.20	0.84
Parts	129.39	10.04	12.89	<0.001
Parts**2	-34.98	3.24	-10.79	<0.001
Age	0.34	0.02	17.04	<0.001
Parts x Age	12.70	2.71	4.69	<0.001
Parts**2 x Age	-8.95	2.46	-3.64	<0.001

Table 1: All model coefficients from a generalized, linear mixed effect model predicting how well children could recognize drawings of object categories as a function of their own age (Age; recognizer age) and the number of unique parts included in each drawing.

Discussion

How do children transform their semantic knowledge of object categories into recognizable depictions? In the current paper, we investigate the developmental changes that characterize children's increasingly sophisticated ability to include and visually emphasize different parts of object concepts in their drawings, and the impact of those changes on their downstream recognizability. To accomplish this, our work builds on a large-scale dataset previously collected by Long et al. (2021) by generating quantitative estimates of the semantic part information of drawn objects conveyed in children's drawings of 16 common object categories. We first crowdsourced part decompositions of the object categories, and then leveraged the generated lists of part labels to provide detailed annotations of how each pen stroke in children's drawings corresponded to the different parts of the depicted objects.

Across development, we found that children became progressively more efficient at drawing, using less strokes that were labeled as unintelligible and using more multipurpose strokes to represent multiple parts of objects, supporting findings by previous studies documenting the consistent gains in children's representational abilities (Luquet, 1927). Critically, however, our findings additionally offer a novel examination of how children choose to visually express their semantic knowledge about the parts of object concepts. We found that children tended to include and visually emphasize certain parts of objects more frequently than others (e.g., including more strokes representing wings than eyes when drawing birds), with drawings of animals containing more part information than those of inanimate object categories (e.g., vehicles, small household objects). As children grew older, they also tended to include more unique parts of objects in their drawings, suggesting that a marked developmental change in children's drawings is the increased complexity of their part structure.

However, exploratory analyses revealed a non-monotonic relationship between the number of unique parts that children depicted in their drawings and how well other naive children could recognize the drawn object. Although we found that this trend varied across object categories, the majority of drawings in our dataset that contained too few or too many parts were often less well-recognized than those with an intermediate number of object parts, demonstrating that increased richness in part information does not necessarily translate to maximal drawing recognizability. Additionally, we observed developmental differences in children's ability to use such part information when attempting to recognize drawn objects; older children were more sensitive to the presence of extra unique part information during drawing recognition than younger children, suggesting that their ability to integrate visual information across multiple object parts changes across development. These findings therefore offer preliminary evidence that what drives the recognizability of children's drawings is their increasing enrichment of part information that characterizes their semantic knowledge of object concepts, but that such recognition success is tempered by a mismatch between their increasing inclusion of object parts and what is actually most informative to visually prioritize about those objects for other viewers. While this work contributes an important characterization of how children visually express their semantic knowledge about parts, it leaves open additional questions about children's sensitivity to the salience of different parts of objects, as well as how children map the correspondence between the parts of drawn and real-world objects both when producing and interpreting drawings.

A major contribution of our work is the development of a public, large-scale dataset of semantic annotations of children's drawings. By systematically measuring the part information conveyed in these drawings, this dataset enables future endeavors that may examine more fine-grained hypotheses about children's semantic knowledge about parts, as well as the developmental changes that can be investigated through their drawings of object concepts. In particular, a critical direction concerns the spatial organization of such semantic part knowledge. One idea is that children learn to increase the visual salience of certain parts of objects relative to others across development, in order to better represent the conceptual salience of those parts as they gain more understanding about what characterizes an object category. For example, when drawing rabbits, children may learn to exaggerate the relative size of the "ears" compared to its "head" either to better fit their understanding of what characterizes rabbits or to better communicate the salience of these informative features to other viewers. Future research exploring the spatial size and location of how children draw parts of objects will likely provide vital insight on how children learn to modulate the visual salience of certain parts, as well as insights on the visualization strategies that children employ to produce more recognizable drawings.

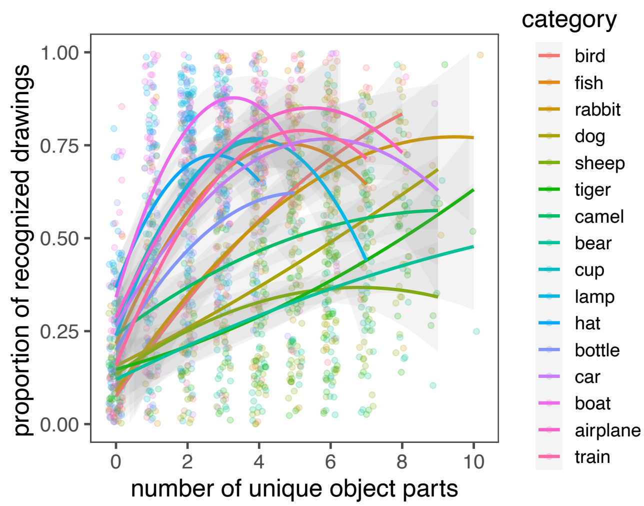

Figure 5: Drawing recognition for each category as a function of the number of unique parts included in each drawing; each individual dot is a unique drawing.

Overall, our paper documents how children's visual expression of their semantic part knowledge changes across development and affects how recognizable their drawings are to others. Further, by publicly releasing our dataset of annotated children's drawings, these data generate opportunities for future studies that seek to understand the relationship between the development of object recognition, semantic knowledge, and object part structure. We propose that systematically examining how children acquire this uniquely human ability to produce visualizations will provide novel insights into how we encode and represent our visual world.

Acknowledgments

Many thanks to the members of the Cognitive Tools Lab at UC San Diego and the Language and Cognition Lab at Stanford University for helpful discussion. We gratefully acknowledge the San Jose Children's Discovery Museum where this dataset was collected. This work was supported by an NSF CAREER Award #2047191 to J.E.F. J.E.F is additionally supported by an ONR Science of Autonomy award and a Stanford Hoffman-Yee grant.

References

Barrett, M., & Light, P. (1976). Symbolism and intellectual realism in children's drawings. *British Journal of Educational Psychology*, *46*(2), 198–202.

Bova, S. M., Fazzi, E., Giovenzana, A., Montomoli, C., Signorini, S. G., Zoppello, M., & Lanzi, G. (2007). The development of visual object recognition in school-age children. *Developmental neuropsychology*, *31*(1), 79–102.

Bremner, J. G., & Moore, S. (1984). Prior visual inspection and object naming: Two factors that enhance hidden feature inclusion in young children's drawings. *British Journal of Developmental Psychology*, *2*(4), 371–376.

Chappell, P. A., & Steitz, J. A. (1993). Young children's human figure drawings and cognitive development. *Perceptual and Motor skills*, *76*(2), 611–617.

Fan, J. E., Yamins, D. L., & Turk-Browne, N. B. (2018). Common object representations for visual production and recognition. *Cognitive science*, *42*(8), 2670–2698.

Fury, G., Carlson, E. A., & Sroufe, A. (1997). Children's representations of attachment relationships in family drawings. *Child development*, *68*(6), 1154–1164.

Huey, H., Walker, C., & Fan, J. (2021). How do the semantic properties of visual explanations guide causal inference?

Juttner, M., Muller, A., & Rentschler, I. (2006). A developmental dissociation of view-dependent and view-invariant object recognition in adolescence. *Behavioural brain research*, *175*(2), 420–424.

Juttner, M., Wakui, E., Petters, D., & Davidoff, J. (2016). Developmental commonalities between object and face recognition in adolescence. *Frontiers in psychology*, *7*.

Karmiloff-Smith, A. (1990). Constraints on representational change: Evidence from children's drawing. *Cognition*, *34*(1), 57–83.

Kellogg, R. (1969). *Analyzing children's art*. National Press Books Palo Alto, CA.

Light, P., & Simmons, B. (1983). The effects of a communication task upon the representation of depth relationships in young children's drawings. *Journal of Experimental Child Psychology*, *35*(1), 81–92.

Long, B., Fan, J., Chai, Z., & Frank, M. C. (2021). Parallel developmental changes in children's drawing and recognition of visual concepts.

Long, B., Fan, J., & Frank, M. C. (2018). Drawings as a window into developmental changes in object representations. In *Proceedings of the 40th annual meeting of the cognitive science society.*

Luquet, G.-H. (1927). Le dessin enfantin.(bibliothèque de psychologie de l" enfant et de pédagogie.).

Minsky, M., & Papert, S. (1972). *Artificial intelligence progress report* (Tech. Rep.). Cambridge, MA, USA.

Mukherjee, K., Hawkins, R. X., & Fan, J. W. (2019). Communicating semantic part information in drawings. In *Proceedings of the 41st annual meeting of the cognitive science society* (pp. 2413–2419).

Natu, V. S., Barnett, M. A., Hartley, J., Gomez, J., Stigliani, A., & Grill-Spector, K. (2016). Development of neural sensitivity to face identity correlates with perceptual discriminability. *Journal of Neuroscience*, *36*(42), 10893–10907.

Piaget, J. (1929). The child's concept of the world. *Londres, Routldge & Kegan Paul*.

Sitton, R., & Light, P. (1992). Drawing to differentiate: Flexibility in young children's human figure drawings. *British Journal of Developmental Psychology*, *10*(1), 25–33.

Tversky, B. (1985). Development of taxonomic organization of named and pictured categories. *Developmental Psychology*, *21*(6), 1111.

Tversky, B. (1989). Parts, partonomies, and taxonomies. *Developmental Psychology*, *25*(6), 983.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of experimental psychology: General*, *113*(2), 169.

Yamins, D., Hong, H., Cadieu, C., Solomon, E., Seibert, D., & DiCarlo, J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, *111*(23), 8619–8624.