Title
Three-Dimensional Object Completion in Humans and Computational Models

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Yiu, Eunice
Collins, Jasmine
Gopnik, Alison

Publication Date
2022

Peer reviewed

Three-Dimensional Object Completion in Humans and Computational Models

Eunice Yiu (ey242@berkeley.edu)
Department of Psychology, University of California, Berkeley
Berkeley, CA 94720 USA

Jasmine Collins (jazzie@berkeley.edu)
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley
Berkeley, CA 94720 USA

Alison Gopnik (gopnik@berkeley.edu)
Department of Psychology, University of California, Berkeley
Berkeley, CA 94720 USA

Abstract

Three-dimensional objects pose a challenge for our visual system, since we can only view objects from a single limited perspective at a given moment. Previous work found that given a limited perspective, infants represent 3D objects as complete volumes. Our study replicated this finding in 4- to 7-year-olds and adults, using an explicit prediction measure rather than looking times. We also explored whether humans have a bias to represent visually limited 3D objects as symmetrical rather than asymmetrical across shape, size, texture, and color. Overall, there was an above-chance preference for full volumetric and symmetrical object completion that increased with age. Low-level perceptual similarity of choices did not predict participants' choices. Moreover, we evaluated *ResNet-50* neural networks on the same tasks: they represented objects as complete volumes, but did not show substantial preference for symmetrical 3D representations. This raises the possibility that incorporating human symmetry biases could improve computer vision.

Keywords: perceptual development; three-dimensional object perception; symmetry

Introduction

Humans can extract a general representation of a three-dimensional object from a single perspective or image. This ability helps us to recognize objects, reason about object affordances, interact with objects, and understand scenes. However, it raises a puzzle. Given the limited information from a single perspective, how do we infer what the unrevealed portions of objects will look like?

Humans are remarkably able to conceive general representations of objects, and can make amodal completions of objects even when they are partially occluded (Bruno, Bertamini, & Domini, 1997). Previous work showed that when presented with occluded two-dimensional surfaces (van Lier, Leeuwenberg, & Van der Helm, 1995) or self-occluded three-dimensional objects (van Lier & Wagemans, 1999), adults prefer global as opposed to local completions. Furthermore, young infants represent simple 3D objects as complete and solid volumes instead of incomplete and hollow volumes even when they see a limited perspective that is compatible with either interpretation (Soska & Johnson, 2008; 2013). Soska, Adolph, & Johnson (2010) further suggested that the visuo-manual exploratory skills and self-sitting experience of infants facilitate their ability to complete 3D objects as solid volumes.

All these findings across different paradigms provide evidence that humans hold certain prior expectations about objects whose forms are not fully revealed, and those expectations influence their inferences about the unseen parts of objects. However, it is unclear what other sorts of predictions underpin human 3D object completion, beyond global completions for 2D surfaces and solid volumes for 3D objects. One potential perceptual bias that has not been explored is symmetry.

Objects in our visual world, natural or manmade, commonly exhibit mirror and/or rotational symmetries (Darvas, 2007; Tyler, 1995). It has been argued that much of our understanding of objects is guided by the perception and recognition of repeated or common patterns (Thompson, 1961). In fact, symmetry is a salient cue in human perception from early development. 4-month-old infants can discriminate bilaterally symmetrical patterns from asymmetrical forms (Fisher, Ferdinandsen, & Bornstein, 1981; Pornstein & Krinsky, 1985). Not only can infants process bilaterally symmetrical patterns more immediately than asymmetrical patterns (Bornstein, Ferdinandsen, & Gross, 1981), but they also develop faster and more accurate recognition memory for the former (Bornstein & Stiles-Davis, 1984). There have been theories suggesting that this ability to appreciate symmetry confers cognitive and evolutionary advantages. American computer scientist Alan Perlis posits that symmetry is a complexity-reducing concept. Treating the other half of a bilaterally symmetrical object as the same thing dramatically reduces information processing load for recognition (Gross & Bornstein, 1978). Symmetry in physical ornaments and motor patterns of living organisms also serves as an indicator of fitness in mate selection (Enquist & Arak, 1994; Zaidel, Arde, & Baig, 2005). It is thus interesting to explore the role symmetry plays in human 3D object completion and representation. When we approach a novel object from a single viewpoint, do we generally expect it to be symmetrical rather than asymmetrical? Does this expectation vary across development? What kinds of symmetry, such as symmetry in material (color and texture)

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

and geometry (shape and size), do humans incorporate in their object completion? Are they equally favored, or are some kinds of symmetry more preferred than others?

In the present study, we aim to, first, investigate if the finding that infants prefer complete volumetric representations of objects from limited viewpoints (Soska and Johnson, 2008; 2013) applies to older children and adults viewing more diverse novel 3D objects; second, examine preference for symmetrical completions of these novel objects in children and adults. We further compare participants' responses against the predictions made by *ResNet-50*, a highly popular state-of-the-art neural network in computer vision (He et al., 2016). We tested three forms of the neural network: one that is supervised and trained on ImageNet – a large dataset consisting of hundreds of object categories and millions of images (Russakovsky et al., 2015), one that is self-supervised and trained on ImageNet as well, and one that is untrained serving as a random feature baseline.

Understanding human priors that underly the 3D object completion from limited perspectives may also be relevant to computer vision. In neural networks, the ability to recognize objects from multiple viewpoints is acquired through an abundance of 2D images displayed from different viewpoints for each object category. There is often neither a direct transfer of 3D competence across categories, nor extracted or built-in priors about what novel views of objects should look like. Further, state-of-the-art detection and segmentation methods are only capable of recognizing and localizing visible object parts (He et al., 2017; Kirillov et al., 2020).

We hypothesize that older humans, like infants, prefer complete volumes, and may likewise prefer symmetry to asymmetry in their object completion. In particular, they may care more about symmetry in geometry than in material, since many studies have shown that both children and adults have a shape bias. The shape bias refers to the inclination to classify, sort, and name objects on the basis of shape rather than other object elements such as color or texture (Landau, Smith, & Jones, 1988; Smith et al., 2002). In contrast, an ImageNet-trained *ResNet-50* may not necessarily show preferences for complete volumes and geometric symmetry in 3D object completion, but it may show preference for material symmetry: like many standard convolutional neural networks, the training of *ResNet-50* may lead it to be more texture-biased than shape-biased (Geirhos et al., 2018; Ringer et al., 2019).

Experiment 1

In Experiment 1, we expanded Soska and Johnson (2008; 2013)'s infant study in three ways: one, we evaluated preferences for 3D solid volumetric completion in *older children and adults*; two, we tested *more diverse and complex 3D stimuli*; three, we tested *explicit predictions* about the object's appearance, rather than the more implicit infant looking-time measures. In addition to presenting participants with two possible options as in the original study, we introduced a physically impossible distractor option that conflicted with the limited viewpoint to ensure participants were not merely making random guesses.

Methods

Participants. 38 child participants aged between 4 years old and 7 years old (M_{age} = 5.94 years, *SD* = 1.16, 20 females) were recruited and tested at children's museums. More specifically, the sample comprises 10 4-year-olds (M_{age} = 4.53 years, *SD* = .36), 10 5-year-olds (M_{age} = 5.44 years, *SD* = .36), 9 6-year-olds (M_{age} = 6.48 years, *SD* = .32) and 9 7-year-olds (M_{age} = 7.49 years, *SD* = .25). An additional 4-year-old was tested but excluded from the sample analysis due to selecting impossible distractors in 1/3 of the test trials. In addition, 40 adult participants (M_{age} = 28.70, *SD* = 6.88; 20 females) were recruited on Prolific to complete the same task. The same experimental stimuli were also tested on self-supervised, supervised, and untrained *ResNet-50*.

Stimuli and Procedure. The study was performed on a computer screen. Participants were introduced to two virtual characters exploring a toy store, and were told that the toys were located on a shelf that was too high for them to reach, and so these objects could only be viewed from a limited perspective with their back parts being occluded. Participants were asked by the experimenter to help predict what 14 novel and abstract 3D toys would look like if they were taken off the shelf and turned around. The 14 objects were divided between 2 practice trials and 12 test trials. All objects were downloaded from Thingi10K, a large dataset of 3D printing models (Zhou and Jacobson, 2016), and were further edited in Blender (an open-source 3D computer graphics software) for adaptation to the experiment.

Practice Trials. The practice trials were designed to ensure that the participants' understood the basic object completion task. In each of the 2 practice trials, participants saw a novel object on the shelf from a limited viewpoint followed by two 15° pivoting options representing what the two characters respectively thought the object would look like if it was turned around (Figure 1a). Only one of the two options was physically possible and did not conflict with the limited viewpoint in terms of shape, size, texture, and color.

Critically, the experimenter asked the participant, "See this object on the shelf? If you take if off the shelf and turn it around, will it look like [pointing to the two options] this or this?" After they selected one view, participants saw a full 360° rotation video of the object and were told whether their response was correct. All participants went through both practice trials before proceeding to the test trials.

Test Trials. In each of the 12 test trials, participants were shown a novel object on the shelf from a limited viewpoint as in the practice trials. This time, the experimenter asked the same question, but participants had to decide among three instead of two different options, "See this object on the shelf? If you take it off the shelf and turn it around, will it look like 1, 2, or 3?" The choices included a possible complete volumetric option, a possible incomplete volumetric option, and an impossible distractor option that conflicted with the limited viewpoint of the object. They were presented in a randomized, counterbalanced order (Figure 1b).

Similar to the practice trials, these options in the test trials pivoted by 15° to facilitate the perception of the depth and three-dimensionality of the object, but the full rotation was not revealed. Once participants made a choice, they were rewarded with encouragement irrespective of what they chose. The goal was to motivate younger children to continue with the task without shaping their responses.

While all children, adults, and *ResNet-50* were tested on the same stimuli, the task was administered in slightly different formats. Child participants were guided through the experiment by a human experimenter, whereas adult participants finished the task in the form of a Qualtrics survey (the experimenter's questions and instructions were written out in words). Since *ResNet-50* was trained with static 2D images, it was fed with the center frame of each 15-degree pivoting option for evaluation. In each trial, we extracted the feature representation from each neural network for the limited viewpoint image and each option image, and considered the option with the highest cosine similarity with the limited viewpoint as the model's choice.

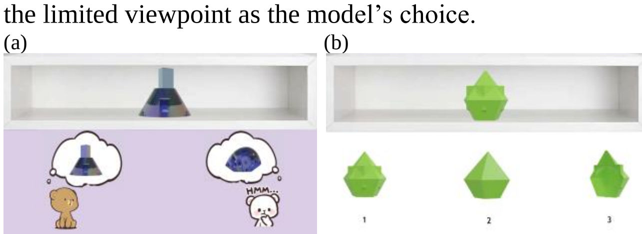

Figure 1: A (a) practice trial involving one possible option and one impossible option) and a (b) test trial involving two possible options (1 and 3), and one impossible option (2).

Results and Discussion

Object Completion Accuracy. First, we determined participants' object completion accuracy to evaluate their ability to understand our task and to complete the rest of each object with a form that was not physically impossible. This was computed as the proportion of test trials in which participants did not choose the distractor options. In other words, we tested whether participants preferred the possible choices to the impossible distractor option. Children scored a mean of 97.4% (*SE* = .89%); age effects were not observed from 4 to 7 years old. Adults scored a mean of 99.2% (*SE* = .50%) (Figure 2a). The high accuracy scores in objection completion confirm that both children and adults could demonstrate object completion and the task was appropriate for both age groups. The Welch's two-sample, two-tailed t-test showed that adults' scores were marginally higher than children's, $t(77) = 1.75$, $p = .084$; Cohen's effect size $d = .40$ suggested small significance. The self-supervised and supervised *ResNet-50* scored 100%; even the untrained *ResNet-50* (random baseline) scored 91.7% (Figure 2b), suggesting that even random pixel-level features capture the information needed to solve the task.

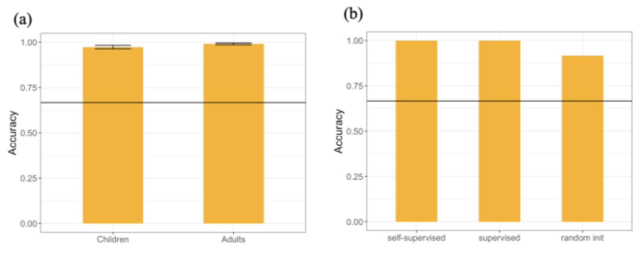

Figure 2: Mean object completion accuracy of (a) humans and (b) *ResNet-50*. Error bars show 1 standard error. Horizontal lines indicate chance-level accuracy (66.7%).

Complete Volumetric Preference. Complete volumetric preference refers to the proportion of test trials in which the complete volumetric option was selected out of the total number of trials in which the distractor option was not chosen. We dropped one child who selected the distractor options in 4 out of the 12 trials. All adults passed this critical requirement and were considered in the analysis.

Both children and adults chose the complete volumetric option significantly above the chance level of 50% (symmetrical option vs. asymmetrical option). Children had a mean complete volumetric preference score of 58.6% (*SE* = 4.27%), while adults had a mean score of 82.7% (*SE* = 3.18%) (Figure 3a). While age effects were not observed from 4 to 7 years old, the Welch's two-sample, two-tailed t-test on children and adults revealed significantly stronger preference in adults, $t(77)$, $p < .001$. Cohen's effect size $d = 1.03$ suggested high practical significance. Thus, our complete volumetric preference continues to strengthen past the age of 7, potentially through increasing exposure to the visual statistics of objects (for instance, noticing there are more objects with complete volumes in the environment).

Like human participants, all three forms of *ResNet-50* also showed preference for complete volumes over incomplete volumes: the self-supervised network showed a complete volumetric preference of 83.3%, the supervised network 91.7%, and the untrained network 81.8%. The incomplete objects possessed convex surfaces which often had darker shadings than the surfaces of the reference object in the shelf. This might have prompted *ResNet-50*, even in its untrained form, to eliminate the incomplete option.

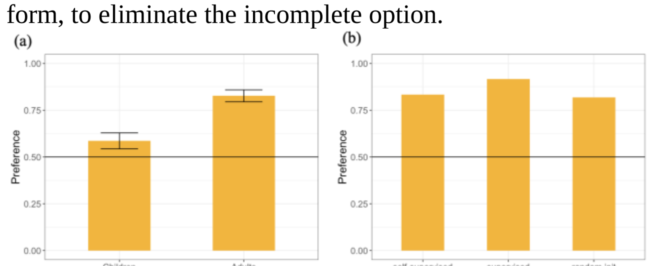

Figure 3: Mean complete volumetric preference of (a) humans and (b) *ResNet-50*. Error bars show 1 standard error. Horizontal lines indicate no preference (50%).

Experiment 2

We adopted the same paradigm in Experiment 1 to study another potential preference in 3D object completion: bilateral symmetry.

Methods

Participants. 82 child participants aged between 4 years old and 7 years old (M_{age} = 5.90 years, SD = 1.13, 40 females) were recruited and tested at children's museums. More specifically, the sample comprised 16 4-year-olds (M_{age} = 4.29 years, SD = .27), 22 5-year-olds (M_{age} = 5.25 years, SD = .29), 22 6-year-olds (M_{age} = 6.31 years, SD = .22) and 22 7-year-olds (M_{age} = 7.29 years, SD = .28). Sixteen additional children were tested but excluded from the sample analysis as they selected the incorrect distractors in at least 1/3 of the test trials. This could be due to inattention, language comprehension issues, or simply inability to understand the task. Again, the same task was also tested on 40 adult participants (M_{age} = 23.65, SD = 4.94; 31 females) who were recruited on Prolific and on *ResNet-50*.

Stimuli and Procedure. The task design and objects used in Experiment 2 were identical to those used in Experiment 1, but the available options were different. In Experiment 2, the choices included a possible *symmetrical* option, a possible *asymmetrical* option, and an impossible distractor option. The asymmetrical option could be asymmetrical in its geometry (shape or size) or material (color or texture). To prevent the inherent differences between objects from confounding with the type of asymmetry presented, each object was edited for asymmetry in all four conditions (shape, size, color, and texture) (Figure 4). These four types of edits of the same object were then allocated to four separate test sets. Participants were randomly assigned to one of the test sets and thus only saw each object once. This enabled preferences for the different types of asymmetry to be assessed and compared across the same set of objects. In other words, the 12 test trials were comprised of 3 trials with asymmetrical shape options, 3 trials with asymmetrical size options, 3 trials with asymmetrical color options, and 3 trials with asymmetrical texture options. The order of options was randomized.

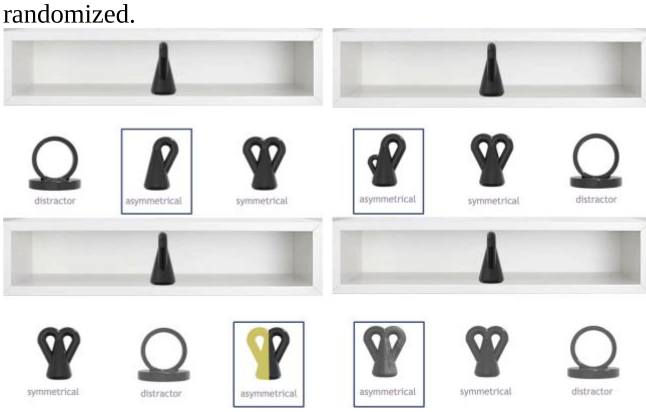

Figure 4: Four types of asymmetrical edits (circled) for the same object (from left to right, top to bottom): shape, size, color, and texture, as presented in the test trials.

Results and Discussion

Object Completion Accuracy. One-sample, two-tailed t-tests on children and adults respectively showed significantly above-chance accuracy (66.7%) across all conditions (p < .001). Children had mean accuracy scores 96.4% in shape (SE = .58%), 96.5% in size (SE = .63%), 96.0% in color (SE = .71%), and 96.4% in texture (SE = .59%). Adults also scored highly above chance for shape (M = 99.4%, SE = .35%), size (M = 100%, SE = 0%), color (M = 100%, SE = 0%), and texture (M = 98.1%, SE = .56%) as well (Figure 5a). In contrast, *ResNet-50* demonstrated poorer performance than humans on the task. The self-supervised *ResNet-50* scored a mean accuracy of 83.3% in shape and color, 100% in size, and 91.7% in texture; the supervised *ResNet-50* scored 83.3% in shape, color, and texture, and 91.7% in size; the untrained *ResNet-50* did not show any capability of completing 3D objects: it scored 33.% in shape, 58.4% in size, 66.7% in color, and 58.4% in texture (Figure 5b).

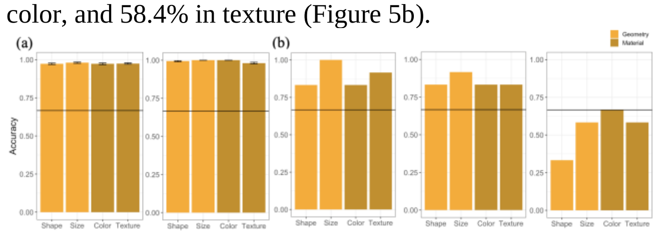

Figure 5: Mean object completion accuracy of (a) (from left to right) children and adults, and (b) (from left to right) self-supervised, supervised, untrained *ResNet-50* in geometry and material conditions. Error bars show 1 standard error. Horizontal lines indicate chance-level accuracy (66.7%).

Symmetry Preference across Conditions. We measured participants' symmetry preference by computing the proportion of test trials in which the symmetrical option was selected out of the total number of trials in which the distractor option was not chosen. We dropped sixteen children who selected the distractor options beyond chance level in at least 4 out of the 12 trials, retaining only those with an accuracy score above 66.7% for our analysis of symmetry preference. Again, all adults passed this critical requirement and were considered in the analysis.

Both children and adults chose the symmetrical option significantly above the chance level of 50% (Figure 6a). Children attained mean symmetry scores of 58.1% in shape (SE = 3.13%), 66.7% in size (SE = 2.82%), 80.1% in color (SE = 3.00%), and 68.3% in texture (SE = 3.49%). Meanwhile, adults had mean symmetry scores of 80.8% in shape (SE = 3.89%), 95.0% in size (SE = 2.25%), 86.7% in color (SE = 2.61%), and 92.9% in texture (SE = 3.32%). Children and adults respectively revealed significantly above-chance symmetry preference across all conditions (p < .001), but there was an exception: children as a group did not show symmetry preference for shape beyond chance level, p = .12. Unlike human participants, *ResNet-50* showed much weaker symmetry preference: the self-supervised *ResNet-50* scored 50.0% in shape and size, 60.0% in color, and 54.5% in texture; the supervised *ResNet-50* scored 50.0%

in shape and texture, 54.5% in size and 70.0% in color; the untrained *ResNet-50* scored 50.0% in shape, 71.4% in size, 37.5% in color, and 28.6% in texture (Figure 6b). Overall, *none of the neural networks* showed a substantial preference for completing objects in a symmetrical fashion as our human participants did. Perhaps understanding symmetry in objects requires 3D reasoning that cannot be acquired from training on ImageNet. Unlike in Experiment 1, the symmetric option cannot be deduced from 2D cues alone.

A within-subjects ANOVA with condition (shape, size, color, texture) as the independent variable and symmetry preference as the dependent variable yielded a main effect of condition in both children, $F(3,324) = 8.82$, $p < .001$, and adults, $F(3,117) = 7.36$, $p < .001$. We used Bonferroni corrections, leading to an adjusted alpha of .0125 in our multiple comparisons. We found that children had greater symmetrical completion preference for color than for shape $t(81) = 5.06$, $p < .001$, and size, $t(81) = 3.26$, $p < .01$. Adults on the other hand, exhibited significantly stronger symmetrical completion preference for size than for shape, $t(39) = 3.15$, $p < .05$. The differences in preferences between the types of symmetries are more flattened out in adulthood than in childhood, as adults preferences approach ceiling.

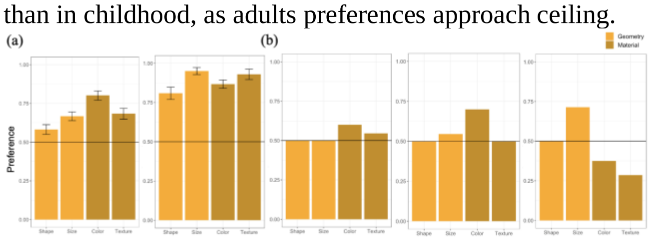

Figure 6: Mean symmetry preference of (a) (from left to right) children and adults, and (b) (from left to right) self-supervised, supervised, untrained *ResNet-50* in geometry and material conditions. Error bars show 1 standard error. Horizontal lines indicate no preference (50%).

Symmetry Preference and Age. Overall, the results of the study suggest a general increase in preference for symmetrical object completions across age. Mean symmetrical preference across conditions increased from a mean of 61.5% (*SE* = 2.92%) among the 4-year-olds (M_{age} = 4.29 years, *SD* = .27) to a mean of 77.3% (*SE* = 3.07%) among the 7-year-olds (M_{age} = 7.29 years, *SD* = .28). In an exploratory analysis, given the few number of trials per condition, we compared symmetry preferences for geometry (shape and size combined) (*M* = 61.7%, *SE* = 2.03%) and material (color and texture combined) (*M* = 73.5%, *SE* = 2.57%). We found that there is a significantly positive correlation between age and symmetry preference for both geometry, $r(81) = .34$, $p < .01$), and material, $r(81) = .35$, $p < .01$, respectively. As indicated in Figure 7, the developmental progression of symmetry preferences in both the material and geometry conditions appear to be similar, with material symmetry preference being stronger than geometry symmetry, $t(81) = 3.58$, $p < .001$, Cohen's effect size $d = .56$ suggested medium significance. When comparing the symmetry preferences in each of these four symmetry conditions between all child participants and all adult participants (with Bonferroni corrections applied, leading to an adjusted alpha of .025), we found that adults showed a higher symmetry preference than children for shape, $t(120) = 4.54$, $p < .001$), size, $t(120) = 7.86$, $p < .001$, and texture, $t(120) = 5.11$, $p < .001$, but not for color, $t(120) = 1.66$, $p = .10$. This suggests that from childhood to adulthood, there is further strengthening of symmetry preferences in object completion with regards to shape, size, and texture. The older we get, the more biased towards symmetry we are when we imagine the occluded surfaces of 3D objects.

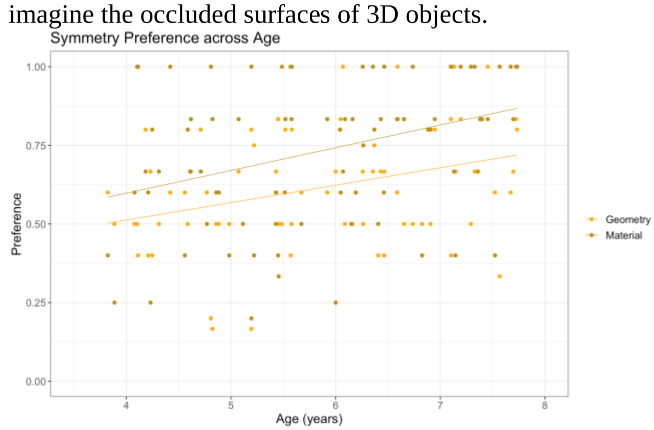

Figure 7: Symmetry preference for geometry (shape & size) and material (color & texture) from 4 to 7 years of age. Each dot represents a child. Data is linearly fitted by solid lines.

General Discussion

The present study provides initial evidence that both children aged 4-7 years old and adults incorporate a preference for complete volumes and symmetry into their completion of novel objects from limited perspectives. Our task asks participants to select what is most plausible to them within a given set of possibilities, so it does not probe what participants might expect outside these options. Nevertheless, this finding complements the existing work on object completion and further demonstrates how priors or expectations concerning the occluded parts of objects may shift across development. The preference for completing novel objects as solid, complete volumes is evident in the looking times of infants (Soska and Johnson, 2008; 2013), and we showed via an explicit prediction measure that this preference continues to strengthen with increasing age. Similarly, completing objects in a symmetrical fashion is observed in early childhood, but adults show an even stronger preference for doing so. This implies that complete volumetric and symmetrical completion preferences are not rigid, built-in priors. Infants' self-sitting experience and visual-manual exploration of objects predicted their looking longer at a volumetrically incomplete object than a volumetrically completion object in 3D object completion (Soska, Adolph, & Johnson, 2010). Preferences may continue to develop through dynamic interactions with objects and scenes in the environment over time. More empirical work is needed to investigate this possibility at a later age.

Some may argue that instead of actually demonstrating complete volumetric and symmetrical object completion preferences, participants were choosing options based on low-level perceptual similarity (e.g., an option was chosen because it looked most perceptually similar to the limited viewpoint); or based on whether or not the 2D limited viewpoint was symmetrical (e.g., the symmetrical option was chosen because the 2D limited viewpoint presented was symmetrical). To test the first possibility, we used "F-score," a metric commonly used in computer vision for multi-view 3D reconstruction (Tatarchenko et al., 2019). This metric evaluates the perceptual similarity of each option relative to the limited viewpoint and relative to the other options in terms of 2D and 3D *geometry*. The F-score evaluates the distance between object surfaces and counts the percentage points that lie within a certain distance on another object under comparison. We also used "SSIM," the structural similarity index measure (Wang et al., 2004), to evaluate perceptual differences in terms of *material*. We did not find any statistically significant correlation between the 2D or 3D F-scores (whether it be relative to the limited viewpoint or relative to other options), or SSIM, and how often an option was chosen by children and adults in both experiments. To test for the second possibility, we evaluated whether the limited viewpoints were symmetrical or asymmetrical along the vertical axis (since we evaluated preferences for bilateral symmetry across the trials). Again, performance was similar whether the limited viewpoint was symmetrical or not.

Our finding of a robust human expectation for symmetry in object completion aligns with the vast literature on human sensitivity to and preference for symmetry (Palmer, Schloss, & Sammartino, 2013; Scheib, Gangestad, & Thornhill, 1999). The increasing tendency to complete objects in a symmetrical manner with age also resonates with existing developmental studies that human symmetry preference increases with age (Huang et al., 2020; Humphrey, 1997).

However, contrary to our initial hypothesis, and in spite of children's well-established shape bias, children exhibited the strongest preference for symmetrical object completion for color and the weakest preference for shape. One possible reason is that bilateral asymmetries in shape (one half of the object is one shape and the other half is a different shape) may occur more frequently than bilateral asymmetries in color (one half of the object is one color and the other half is a different color) in the artefacts in our everyday environments, such as kitchen tools or toys. From infancy, humans have a strong capacity for visual statistical learning (Fiser & Aslin, 2001; Kirkham, Slemmer, & Johnson, 2002), and may have come to expect fewer color asymmetries than shape asymmetries. This could explain the relatively lower shape symmetry preference in both children and adults. Another possibility stems from the fact that young children between 3.5 and 6 years of age also tend to be faster and more accurate at naming the color of an object when the object was presented in an abstract shape (Prevor & Diamond, 2005). The abstractly shaped objects in our study may therefore have increased children's attention to color and driven their preference for color symmetry.

Like children and adults, pre-trained *ResNet-50* networks completed objects at above-chance accuracy in both experiments, while the untrained network failed to do so in Experiment 2, potentially because the distractor options were harder to visually distinguish in that experiment. Critically, in their object completion, the *ResNet-50* networks preferred volumetrically complete objects in Experiment 1, but not symmetrical objects in Experiment 2. Compared to the reference objects on the shelf and the volumetrically complete options, the volumetrically incomplete options had convex surfaces reflected through darker shadings. Darker shadings might contribute to a lower cosine similarity, thereby causing *ResNet-50* to avoid the incomplete options. This also lends some support to the hypothesis that neural networks are texture-biased and attend less to shape (Geirhos et al., 2018; Geirhos et al., 2020). That said, SSIM, which includes luminance and contrast masking terms, did not predict the choices of *ResNet-50*. By contrast, *ResNet-50* cannot use 2D cues such as shading in Experiment 2. Trained or untrained, it demonstrated little to no preference for symmetrical completions, suggesting that 3D symmetrical object completion may not be achieved solely through cosine similarities following training on 2D images. The ImageNet-trained *ResNet-50's* did not necessarily develop 3D understanding of objects the way humans do, and thus may not further develop preferences for symmetry in 3D objects.

From a broader perspective, the present study may shed light on understudied aspects of object completion, an important problem in both human cognition and computer vision. We examined preferences for solid volumes and bilateral symmetry, the developmental trend from childhood to adulthood, as well as the differences between humans and ResNet-50 neural networks in completing occluded parts of an object. Given the small number of trials (n = 12) in these experiments, we hope to scale up the object dataset to generate more reliable findings. Further studies may also explore other variables such as the axis of symmetry and the animacy of objects to further understand this important aspect of perception. Manmade structures tend to possess more segments of straight lines, longer linear lines, and coterminations than animates or natural objects (Iqbal and Aggarwal, 2002). Hence, it is possible that symmetry preferences may differ between these categories. On the computational side, we plan to train neural networks on 3D (as opposed to 2D) object data and test the preferences of single-view 3D reconstruction networks.

Acknowledgements

This work was funded by the DARPA [047498-002] Co-PI Machine Common Sense Grant. We are grateful to the members of the Cognitive Development and Learning Lab at UC Berkeley, especially Iran Torres Aleman, Megan Lui, and Yuki Bian, for their help with data collection. We also thank the Bay Area Discovery Museum, Children's Creativity Museum, Lawrence Hall of Science, as well as parents, children, and adults who supported this research.

References

Bornstein, M. H., Ferdinandsen, K., & Gross, C. G. (1981). Perception of symmetry in infancy. *Developmental psychology*, *17*(1), 82.

Bornstein, M. H., & Stiles-Davis, J. (1984). Discrimination and memory for symmetry in young children. *Developmental Psychology*, *20*(4), 637.

Bruno, N., Bertamini, M., & Domini, F. (1997). Amodal completion of partly occluded surfaces: Is there a mosaic stage?. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(5), 1412.

Darvas, G. (2007). *Symmetry: Cultural-historical and ontological aspects of science-arts relations; the natural and man-made world in an interdisciplinary approach*. Springer Science & Business Media.

Enquist, M., & Arak, A. (1994). Symmetry, beauty and evolution. *Nature*, *372*(6502), 169-172.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, *99*(24), 15822-15826.

Fisher, C. B., Ferdinandsen, K., & Bornstein, M. H. (1981). The role of symmetry in infant form discrimination. *Child development*, 457-462.

Geirhos, R., Narayanappa, K., Mitzkus, B., Bethge, M., Wichmann, F. A., & Brendel, W. (2020). On the surprising similarities between supervised and self-supervised models. *arXiv preprint arXiv:2010.08377*.

Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., & Brendel, W. (2018). ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. *arXiv preprint arXiv:1811.12231*.

Gross, C. G., & Bornstein, M. H. (1978). Left and right in science and art. *Leonardo*, *11*(1), 29-38.

Hayakawa, S., Kawai, N., & Masataka, N. (2011). The influence of color on snake detection in visual search in human children. *Scientific reports*, *1*(1), 1-4.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, 2961-2969.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Huang, Y., Lyu, J., Xue, X., & Peng, K. (2020). Cognitive basis for the development of aesthetic preference: Findings from symmetry preference. *Plos one*, *15*(10), e0239973.

Huang, Y., Xue, X., Spelke, E., Huang, L., Zheng, W., & Peng, K. (2018). The aesthetic preference for symmetry dissociates from early-emerging attention to symmetry. *Scientific reports*, *8*(1), 1-8.

Humphrey, D. (1997). Preferences in symmetries and symmetries in drawings: Asymmetries between ages and sexes. *Empirical Studies of the Arts*, *15*(1), 41-60.

Iqbal, Q., & Aggarwal, J. K. (2002). Retrieval by classification of images containing large manmade objects using perceptual grouping. *Pattern recognition*, *35*(7), 1463-1479.

Kirillov, A., Wu, Y., He, K., & Girshick, R. (2020). Pointrend: Image segmentation as rendering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9799-9808).

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35-B42.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development*, *3*(3), 299-321.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual review of psychology*, *64*, 77-107.

Pornstein, M. H., & Krinsky, S. J. (1985). Perception of symmetry in infancy: The salience of vertical symmetry and the perception of pattern wholes. *Journal of experimental child psychology*, *39*(1), 1-19.

Prevor, M. B., & Diamond, A. (2005). Color–object interference in young children: A Stroop effect in children 3½–6½ years old. *Cognitive development*, *20*(2), 256-278.

Ringer, S., Williams, W., Ash, T., Francis, R., & MacLeod, D. (2019). Texture bias of CNNs limits few-shot classification performance. *arXiv preprint arXiv:1910.08519*.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, *115*(3), 211-252.

Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1431), 1913-1917.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological science*, *13*(1), 13-19.

Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: motor skill acquisition facilitates three-dimensional object completion. *Developmental psychology*, *46*(1), 129.

Soska, K. C., & Johnson, S. P. (2008). Development of three-dimensional object completion
in infancy. *Child development*, *79*(5), 1230-1236.

Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: motor skill acquisition
facilitates three-dimensional object completion. *Developmental psychology*, *46*(1), 129.

Soska, K. C., & Johnson, S. P. (2013). Development of three-dimensional completion of complex objects. *Infancy*, *18*(3), 325-344.

Tatarchenko, M., Richter, S. R., Ranftl, R., Li, Z., Koltun, V., & Brox, T. (2019). What do single-view 3d reconstruction networks learn?. In *Proceedings of the IEEE/CVF*

Conference on Computer Vision and Pattern Recognition (pp. 3405-3414).
Thompson, D. (1961). On Growth and Form (JT Bonner, ed. ab.).
Tyler, C. W. (1995). Empirical aspects of symmetry perception. *Spatial Vision*, *9*(1), 1-8.
van Lier, R. J., Leeuwenberg, E. L., & van der Helm, P. A. (1995). Multiple completions primed by occlusion patterns. *Perception*, *24*(7), 727-740.
van Lier, R., & Wagemans, J. (1999). From images to objects: Global and local completions of self-occluded parts. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1721.
Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, *13*(4), 600-612.
Zaidel, D. W., Aarde, S. M., & Baig, K. (2005). Appearance of symmetry, beauty, and health in human faces. *Brain and cognition*, *57*(3), 261-263.
Zhou, Q., & Jacobson, A. (2016). Thingi10k: A dataset of 10,000 3d-printing models. *arXiv preprint arXiv: 1605.04797*