The "Fraction Sense" Emerges from a Deep Convolutional Neural Network

Yun-Shiuan Chuang (yunshiuan.chuang@wisc.edu)
Department of Psychology, University of Wisconsin-Madison, 1202 W. Johnson Street
Madison, WI 53706 USA

Edward M. Hubbard (emhubbard@wisc.edu)
Department of Educational Psychology, University of Wisconsin-Madison, 1025 W. Johnson Street
Madison, WI 53706 USA

Joseph L. Austerweil (austerweil@wisc.edu)
Department of Psychology, University of Wisconsin-Madison, 1202 W. Johnson Street
Madison, WI 53706 USA

Abstract

Fractions are a critical building block for the development of human mathematical cognition, but the origins of this concept are not well-understood. Recent work has found that a whole number sense is present in deep convolutional neural networks (DCNNs) pre-trained for object recognition and uses them as a model for investigating human numerical cognition. Do DCNNs also have a fraction sense? If so, is it dependent or independent of whole number processing? We investigated the neural sensitivity of a pretrained DCNN to both whole numbers and fractions. We replicated and extended previous research that the sense of whole number emerges in a different DCNN architecture. Further, we showed that DCNN is also sensitive to fraction value, i.e., the ratio of numerosities. Testing this model, our results suggest that the fraction sense relies on the whole number sense.

Keywords: deep convolutional neural network; emergent sense of number; ratio-processing system; approximate number system

Introduction

Understanding the origins of human mathematical abilities is critical for both theoretical and applied reasons. How people develop increasingly sophisticated mathematical capabilities is not well understood. One crucial building block to achieving a fully developed mathematical cognition is using and understanding the concept of a fraction appropriately (Siegler et al., 2012; Matthews, Lewis, & Hubbard, 2016). There are two main theories for how fraction understanding develops. Some have argued that fraction understanding is bootstrapped from the whole number system using logical rules (i.e., divide two whole numbers; Dehaene, 2011; Feigenson, Dehaene, & Spelke, 2004). Others have proposed that fraction cognition emerges in a bottom-up manner based on the primitive non-symbolic ratio-processing system (RPS; Lewis, Matthews, Hubbard, & Matthews, 2015). The RPS is analogous to the approximate number system, but instead of estimating a single numerosity, it estimates relative numerosities. Its existence has been supported by studies with infants (McCrink & Wynn, 2007) and non-human primates (Vallentin & Nieder, 2008). In this paper, we build on this work by conducting a computational investigation of whether state-of-the-art object recognition models have a fraction sense, and if they do, whether it is built on a whole number sense.

Over the last decade, machine learning models of object recognition have advanced to human-like performance. The standard model for state-of-the-art object recognition is a deep convolutional neural network (DCNN; Krizhevsky, Sutskever, & Hinton, 2012). Their performance is so striking that some researchers use them as a tool for investigating human object recognition (Cadieu et al., 2014). Recently, mathematical cognition researchers have found that a "whole number sense" emerges within these DCNNs, despite being trained purely for object recognition (DeWind, 2019; Nasr, Viswanathan, & Nieder, 2019). In fact, a whole number sense has been found even in an untrained DCNN with random weights (Kim, Jang, Baek, Song, & Paik, 2019). In this paper, we build on this work by analyzing whether these models have a fraction sense. To the best of our knowledge, our work is the first work to do so.

To investigate whether a whole number sense can emerge from an object recognition system, previous work examined how sensitive a DCNN is to whole numbers, presented as dot arrays (DeWind, 2019; Kim et al., 2019; Nasr et al., 2019). They found that a whole number sense can emerge within a DCNN, in a manner that is invariant to various confounds (e.g., area, density). Further, the activation patterns of the whole-number-selective neurons in a DCNN (the artificial neurons) are similar to the neural activity of humans and primates when they use their whole number sense. These studies indicate that number sense is an emergent property of DCNNs. This could due to statistical regularities in visual input, visual recognition training and/or the architecture of a DCNN. Here, we asked whether a DCNN also has a fraction sense.

In this paper, we use DCNNs to investigate the relation between whole number and fraction processing. First, we show that the whole number sense is robust to the particular DCNN used (previous work only used one DCNN – AlexNet; Krizhevsky et al., 2012). Next, we examine whether DCNNs also have a fraction sense. Then, we explore whether fractions are processed in a manner that is independent from whole number processing. We concluded with a discussion of the limitations and implications of our work.

Methods

Deep Convolutional Neural Network

To test the robustness of number sense results to DCNNs other than AlexNet, we used the VGG16 model (Simonyan & Zisserman, 2014) with weights trained to classify the 1000 object classes in ImageNet (Russakovsky et al., 2015). Note that the model has never been explicitly trained to learn the concept of numerosity. Any sensitivity to numerosity is an emergent property of learning to recognize objects, visual statistics and/or the model architecture itself.

Stimuli

We used dot arrays as stimuli (See Fig. 1), in which the number of dots defined a stimulus' numerosity. For fractions, we used two presentation formats to ensure the robustness of the results: side-by-side and intermixed (See Figs. 1bc). For side-by-side fractions, the stimulus was composed of a dot array in the top-left (the numerator) and in the bottom-right (the denominator) of the image. This format representation avoided confounds between the fraction's value and the ratio of size and spacing, but leaves correlations with the ratio of convex hull (see below for definitions of size, spacing, convex hull, field areas etc.). To address the potential confound of convex hulls, we also used the intermixed fraction format. For intermixed fractions, there was a single dot array and the numerator and denominator were represented by different colors. This design ensured that numerator and denominator arrays share the same convex hull and thus avoids this confound. However, this format has a different limitation: the fractions value correlates with the ratio of sparsity. The intermixed fraction format mirrors the one used in infant and non-human primate work (McCrink & Wynn, 2007; Vallentin & Nieder, 2008). Using two fraction stimulus formats help ensure that our results are robust to these confounds.

Whole Number Stimulus Set The whole number stimuli were generated with the scripts used by DeWind (2019). The stimuli were evenly sampled from a three-dimensional stimulus space (Fig. 1d) that includes three axes: numerosity, size, and spacing. For each stimulus, numerosity was the number of dots, size was the area of each dot multiplied by total area of all dots, and spacing was the field area (area surrounded by the convex hull) multiplied by the sparsity. Following best practices from previous work (DeWind, Adams, Platt, & Brannon, 2015), we used these three axes to minimize confounds with simple image features.

For this study, numerosity had 5 levels, ranging from 1 to 5 (Fig. 1a). Size and spacing also had 5 levels with equal intervals between levels. All combinations between numerosity, size, and spacing were used to generate stimuli (with random dot positions). This resulted in 125 unique whole-number stimuli.

Side-by-side Fraction Stimulus Set The side-by-side fraction stimuli were generated by inserting a dot array in the top-left corner (the numerator) and a dot array in the bottom-

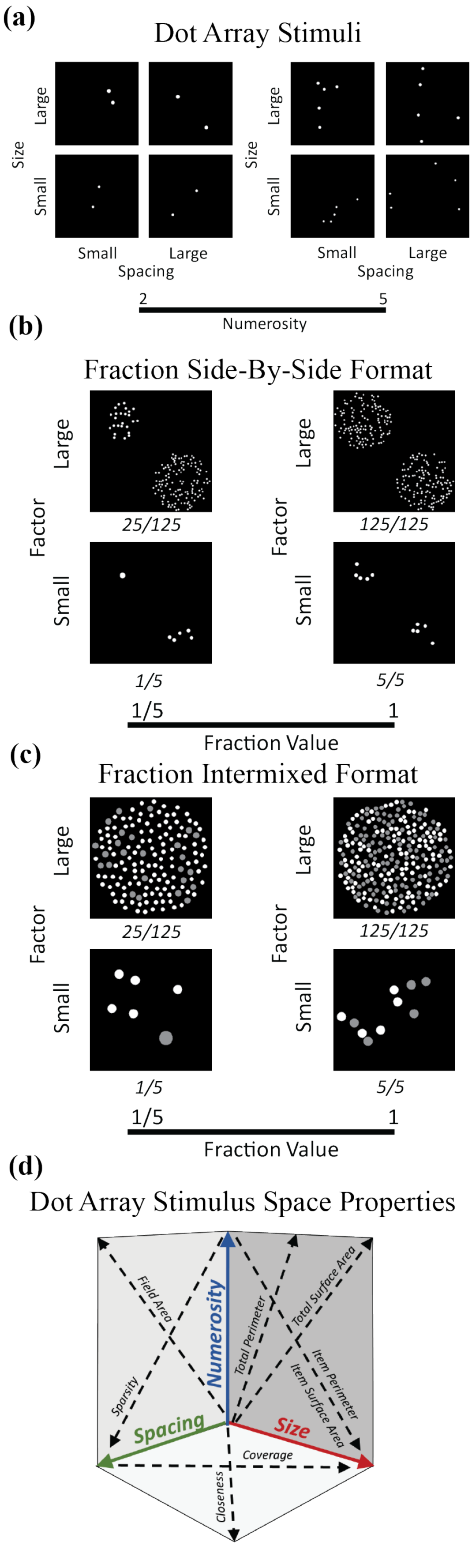

Figure 1: Stimuli. (a) Whole number stimuli (modified from DeWind, 2019 with consent). (b) Side-by-side (top-left numerator, bottom-right denominator) and (c) intermixed fraction stimuli (gray numerator, white denominator). (d) Dot-array stimulus space. 3 principle axes are numerosity, size, and spacing. Dotted lines represent other stimulus features. See text for details regarding the properties of this stimulus space (modified from DeWind et al., 2015 with consent).

right corner (the denominator) of the same image (Fig. 1b). The fraction value had 5 levels, ranging from one-fifth to one in equal intervals. For each fraction value, there were 25 factors to encode the number of dots per numerator or denominator value. For example, 25 images were included where $5, 10, 15, \ldots, 125$ dots were used for the numerator and denominator for the fraction value of $5/5$. With 5 fraction values and 25 factor levels, there were 125 fraction stimuli. Critically, the fraction value was not perfectly correlated with either the numerator or denominator. Therefore, the DCNN could not determine the fraction value solely from one of the dot arrays and must use both the numerator and denominator to solve it correctly. To avoid the DCNN from deciding the fraction value simply by the ratio of size or spacing between the numerator dot array and denominator dot array, we fixed the size ratio and spacing ratio to constant across all 125 stimuli.

Intermixed Fraction Stimulus Set As shown in Fig. 1c, intermixed fraction stimuli were composed of dots of two colors, gray (numerator) and white (denominator). As with the side-by-side format, we fixed the size ratio to be constant. However, unlike that format, we also fixed the field area ratio to be constant. We did this because the two dot arrays were intermixed and have roughly the same convex hull. This means that the ratio of spacing between the numerator and denominator corresponded perfectly to the fraction value, i.e., ratio of spacing = ratio of (field area $\times$ sparsity) = ratio of (field area^2/numerosity) = ratio of (field area^2)/ratio of numerosity) = 1/fraction value. We used the same fraction values and number of dots per numerator and denominator as we did with the side-by-side format.

Activations of Artificial Neurons

To measure the sensitivity of artificial neurons to numerosity and other features, we first computed the intermediate activation of the artificial neurons given a stimulus as input. We only focused on the output activation of the 13 convolutional layers of VGG16 in our analysis. We excluded the output activation of the fully connected layers as they were built for classifying the 1000 objects in the ImageNet challenge, rather than as part of the visual processing itself. This resulted in 13,547,520 hidden neurons of interest. For each stimulus set, we constructed an activation matrix $\mathbf{A}$ that was of size, number of stimuli $\times$ number of neurons. The row $\mathbf{a}_s$ was the intermediate activation of all neurons of interest for the s^{th} stimulus.

Neural Sensitivity Measurement

For each neuron, the neural sensitivity to numerosity and other features were measured as the model R^2 of a multiple regression with dummy-coded regressors. The sensitivity for artificial neuron j is

$$\widehat{a}_{j,s} = \sum_{l=1}^{5} \left(b_{l,j} \times I(n_s = l) \right). \quad (1)$$

The dependent variable is the predicted activation of neuron j for stimuli s. l indexes the possible numerosity values (in discrete levels), n_s is the numerosity of stimulus s, and $I(\cdot)$ is the indicator function (value is one when its input is true, and zero otherwise). $b_{l,j}$ is the regression parameter for numerosity level l for neuron j, which was estimated by minimizing sum of squared error. The model R^2 of the model above indicates the neural sensitivity of neuron j to numerosity. The rationale of using this measure is that if a neuron encodes

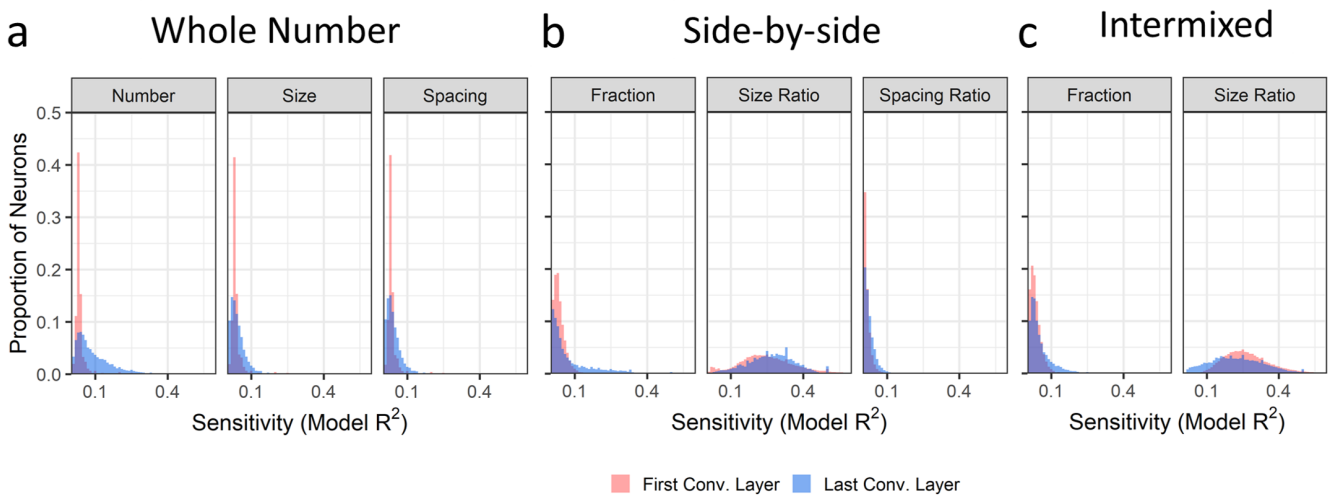

Figure 2: The histograms of neural sensitivity in the first and last convolutional layer for the (a) whole number stimulus set, (b) side-by-side fraction stimulus set, and (c) intermixed fraction stimulus set. The sensitivity is measured by the model R^2 of the regression model for each feature. Spacing was omitted for the intermixed format (see text for details).

numerosity more, its corresponding predictor weight should increase as well, resulting in higher model R^2. This measure was used by DeWind (2019) to investigate the whole number sense in DCNNs. Artificial neuron sensitivity for other features (e.g., size, spacing) was calculated in the same manner.

Artificial Neuron Number Sense Selectivity

Selective neurons were defined as those sensitive to a specific property, while being insensitive to other confounding properties. We examined three types of selectivity. Following DeWind (2019), we defined an artificial neuron as selective to whole-number if its sensitivity to numerosity was above 0.1 and its sensitivities to both size and spacing were lower than 0.001. For the side-by-side format, fraction-selective neurons were defined as having a sensitivity to fraction value above 0.1 and sensitivity to the ratio of size and the ratio of spacing below 0.001. For the intermixed format, the threshold for the ratio of spacing was not applied due to the issues discussed above. There were insignificant variation in some features due to image pixelation and rounding.

We defined an artificial neuron as *fraction-dissociable* if its sensitivity to the fraction value was above 0.1 and its sensitivity to the numerosity of the numerator and denominator were both below 0.001. Fraction-dissociable neurons are those that encode the fraction values while not encoding the numerator or denominator themselves.

Monte Carlo Permutation Test on Selective Neurons

We performed a Monte Carlo permutation test (Dwass, 1957) on the selective neurons in order to ensure they did not emerge just by chance. For each type of selective neurons, we constructed a permutation distribution by sampling 3,000 random permutations of which numerosity labels correspond to which stimuli, and then recalculated the regression analyses with the permuted data. For instance, for whole-number selective neurons, at each random permutation, we shuffled the labels of numerosity randomly, estimated the neural sensitivity to numerosity, and computed the proportion of whole-number selective neurons. After 3,000 random permutations, we then derived the p-value by counting how many permutations have proportion of whole-number selective neurons greater than or equal to the observed proportion. In addition, we calculated the 99% confidence interval (CI) of the p-value based on the Binomial distribution. We performed the permutation test on fraction-selective neurons with the same process, except that we shuffled the labels of fraction value rather than numerosity.

Results

The DCNN Whole Number Sense

Focusing on the whole number stimuli, we found neurons that are sensitive to numerosity, size, and spacing (Fig. 2a). For all three features, the neural sensitivity increases from the first to the last convolutional layer.

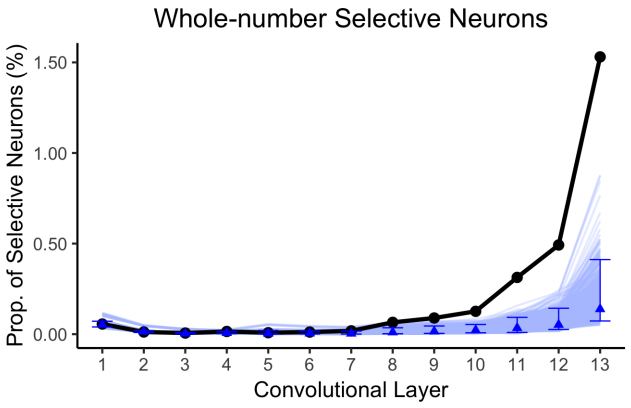

Figure 3: The proportion of whole-number selective neurons (black line) and the proportion of such neurons based on Monte Carlo permutations (light blue lines) across 13 convolutional layers. For each layer, the blue triangle labels the median of the permutations, along with the error bars labelling the 1% and 99% quantiles.

Do numerosity-sensitive neurons really detect numerosity or some confound, such as size or spacing that might be correlated with numerosity? As shown in Fig. 3, there are whole-number selective neurons, and there is a greater proportion of number-selective neurons in deeper than shallower layers. For the first convolutional layer, the proportion of whole-number selective neurons is 0.056%, which is not reliably different from random the permutations, p = .334, 99% CI of p = [.314, .354]. In contrast, in the last convolutional layer, the proportion increases to 1.53%, which is reliably different from random permutations, p = .0003, 99% CI of p = [.000, .001].

The DCNN Fraction Sense

Fraction selective neurons increases in proportion with depth (Figs 4ab). Note that there are fewer fraction selective neurons than whole-number selective neurons, e.g., at the last convolutional layer, 1.54% of whole-number selective neurons vs. 0.082% (side-by-side) or 0.11% (intermixed) fraction selective neurons. However, the permutation test suggests that only the proportion of side-by-side fraction selective neurons is reliably different from random permutations, but this is not true for the proportion of intermixed fraction selective neurons. For instance, at the last convolutional layer, the proportion of side-by-side fraction selective neurons is reliably different from random permutations, p = .002, 99% CI of p = [.000, .004], but not for intermixed fraction selective neurons, p = .123, 99% CI of p = [.102, .143]. Nonetheless, this supports the argument that a fraction sense (at least for side-by-side fractions) can emerge from a DCNN merely trained to classify objects.

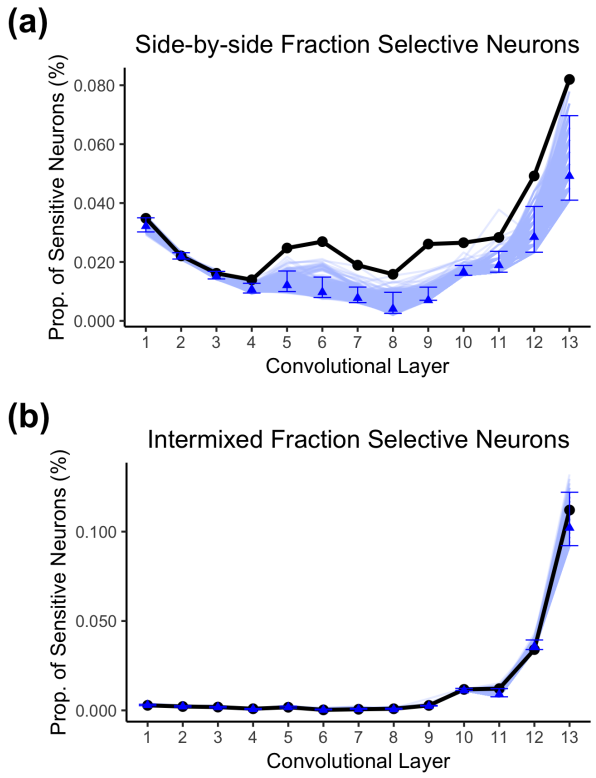

Figure 4: The proportion of fraction selective neurons across 13 convolutional layers for (a) the side-by-side (b) intermixed fraction formats. The proportion of fraction selective neurons (black line) and the proportion of such neurons based on Monte Carlo permutations (light blue lines) across 13 convolutional layers. For each layer, the blue triangle labels the median of the permutations, along with the error bars labelling the 1% and 99% quantiles.

Dissociability of the Sense of Fraction and the Sense of Whole Number

Interestingly, we found no fraction dissociable neurons for either fraction format. To check if this may be due to thresholds being too strict, Figure 5 visualizes the relations between sensitivity to fraction value and the numerator or denominator. For both fraction formats, there are essentially no neurons that are sensitive to the fraction value without being sensitive to the numerator's value. To a lesser extent, this is also true for the denominator. The lack of fraction dissociable neurons supports the dependence of the fraction sense on the whole number sense.

Discussion

In this paper, we investigated the extent to which DCNNs trained to recognize objects contain a whole-number or fraction sense. We found that the whole-number sense emerges in VGG16, replicating and extending previous work that found it emerges in a different DCNN architecture. This provides robustness across particular models to the emergence of a whole-number sense in DCNNs.

We also demonstrated that the DCNN has a fraction sense, but this is limited to the side-by-side fraction format. One explanation for why fraction sense does not emerge for intermixed fractions is that evaluating side-by-side dots and intermixed dots involve distinct perceptual processes (e.g.,Norris & Castronovo, 2016). For humans, evaluating the fraction value of an intermixed format stimulus requires an attentional shift from the dark dots (numerator) to the bright dots (denominator). In contrast, the side-by-side fraction format does not necessarily requires attentional shifts because distinguishing numerator from denominator can be accomplished simply by the spatial location on the retina. Given that the VGG16 is a feedforward neural network without top-down attentional modulation, it is unclear whether it is capable of evaluating displays of fraction that humans need attentional shifts to accomplish. Alternatively, the difference in the result between two fraction formats might be due to differences in how the displays control for convex hull (Gebuis, Cohen Kadosh, & Gevers, 2016). Another difference between the side-by-side fraction format and the intermixed fraction format is that the latter one fixes the ratio of convex hulls to one. To fully understand under what condition a fraction sense can emerge from a DCNN, it will be important to test it with additional fraction formats. For example, a fraction value could be represented with two juxtaposed bars with different lengths or pairs of circles (Jacob, Vallentin, & Nieder, 2012; Matthews et al., 2016).

Furthermore, we did not find artificial neurons sensitive to fraction values that were not sensitive to the numerator or the denominator. This provides support for fractions being bootstrapped from the whole-number system, rather than from the RPS. However, the threshold analyses we conducted have limitations.There are alternative explanations. First, the absence of fraction dissociable neurons might have to do with prior statistical structure. In the real world (e.g, images, textbook), larger fractions tend to have larger numerators. This is also true for our fraction stimuli set where the numerator value is partially correlated with the fraction value. Second, the fact that we fixed the ratio of other visual aspects (e.g., the ratio of size, the ratio of spacing) to one might force the fraction sense to rely on the ratio of numerosity. Third, our analyses focused on analyzing a neural network's internal representation, rather than its behavior. To be certain that the network encodes numerosity, we would also need to demonstrate that the network can also produce behaviors consistent with a proper numerosity understanding (e.g., classify images by their fraction numerosity). To deal with some concerns with our threshold analyses, we plan to test the interdependency of processing of different features (e.g., fraction and whole number) using neural network perturbation analysis (Khakzar, Baselizadeh, Khanduja, Kim, & Navab, 2019). This would examine whether the DCNN could perform the fraction comparison task with the whole-number selective neurons ablated.

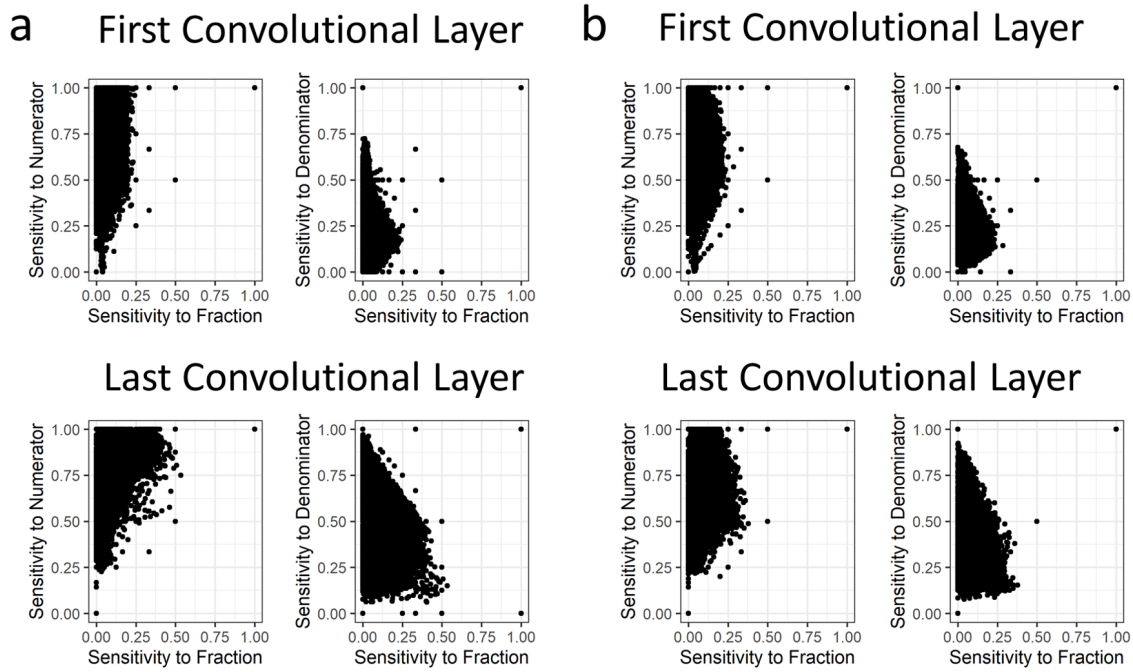

Figure 5: The relations between fraction sensitivity, numerator sensitivity, denominator sensitivity in (a) the side-by-side fraction format and (b) the intermixed fraction format for the first and the last convolutional layers.

One essential question is why DCNNs have a number sense at all. Previous research suggested that the number sense is being bootstrapped from the visual processing system, as the DCNNs were pretrained to recognize visual objects (DeWind, 2019; Nasr et al., 2019). However, an alternative account is that the structure of a convolutional neural network itself results in numerosity sensitive neurons (Kim et al., 2019). Kim et al. (2019) showed that an untrained AlexNet with random weights is still able to detect whole numbers. This implies that the number sense may be due to architectural constraints, rather than a by-product of object recognition. Nonetheless, their results focused purely on whole number and it is unknown whether an untrained network would have a fraction sense.

Another limitation of our work is that a DCNN is by no means a perfect model for visual processing. But, it is arguably the best available model that we can analyze feasibly (Cadieu et al., 2014). There are certainly differences in global shape processing (Doerig, Bornet, Choung, & Herzog, 2020). However, it is unclear to what extent a whole-number or fraction sense depends on this type of processing. Ideally, our findings should be corroborated with fine-grained neural measurements, such as eCOG.

Another interesting issue is how different mechanisms for numerosity estimation relate to how a DCNN detects numerosity. Previous studies have shown that when the density of dots is very high, or the total number of dots is very low, perceptual mechanisms such as texture sensitivity or subitizing comes into play (Anobile, Cicchini, & Burr, 2013, 2015; Anobile, Turi, Cicchini, & Burr, 2015). In the current fraction stimulus set, the fractions with a large number of dots in both the numerator and denominator have a large density of dots (number of dots per square degree). For human visual perception, when the density of dots in an array is greater than 5 dots per square degree, the involved mechanism switches from number estimation to texture sensitivity. On the other hand, for very small number of dots (less than about 4), people tend to rely on subitizing (i.e., the ability to exactly enumerate without counting). DCNNs also differ from human vision in terms of the visual resolution across the retina. It is unclear how these different numerosity mechanisms relate to DCNN numerosity detection and their number and fraction sense more broadly.

There are also constraints related to using the dot array format for the fraction stimulus sets. Specifically, the design of whole number stimulus set ensures numerosity is not confounded with other stimulus features (e.g., dot size, sparsity) because stimuli were sampled equally from a deliberately designed stimulus space (DeWind et al., 2015). This is not the case for both of the fraction stimulus sets. For example, in the intermixed representation of fraction, the denominator and numerator share the same convex hull and thus result in a perfect correspondence between the fraction value and some potential confounding features (e.g., the ratio of spacing, the ratio of sparsity). As mentioned before, one way to ensure the robustness of our results about a fraction sense is to test its robustness to more fraction formats. We will investigate these issues in our future work.

References

Anobile, G., Cicchini, G. M., & Burr, D. C. (2013). Separate Mechanisms for Perception of Numerosity and Density. *Psychological Science*, *25*(1), 265–270.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2015). Number As a Primary Perceptual Attribute: A Review. *Perception*, *45*(1-2), 5–31.

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision*, *15*(5), 1–12.

Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., ... DiCarlo, J. J. (2014). Deep Neural Networks Rival the Representation of Primate IT Cortex for Core Visual Object Recognition. *PLOS Computational Biology*, *10*(12), e1003963.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Second ed.). NY: Oxford University Press.

DeWind, N. K. (2019). The number sense is an emergent property of a deep convolutional neural network trained for object recognition. *bioRxiv*.

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265.

Doerig, A., Bornet, A., Choung, O., & Herzog, M. (2020). Crowding reveals fundamental differences in local vs. global processing in humans and machines. *Vision Research*, *167*, 39–45.

Dwass, M. (1957). Modified Randomization Tests for Nonparametric Hypotheses. *Ann. Math. Statist.*, *28*(1), 181–187.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314.

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35.

Jacob, S. N., Vallentin, D., & Nieder, A. (2012). Relating magnitudes: The brain's code for proportions. *Trends in Cognitive Sciences*, *16*(3), 157–166.

Khakzar, A., Baselizadeh, S., Khanduja, S., Kim, S. T., & Navab, N. (2019). Explaining Neural Networks via Perturbing Important Learned Features. *arXiv:1911.11081*.

Kim, G., Jang, J., Baek, S., Song, M., & Paik, S.-B. (2019). Spontaneous generation of innate number sense in untrained deep neural networks. *bioRxiv*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in NeurIPS* (pp. 1097–1105).

Lewis, M. R., Matthews, P. G., Hubbard, E. M., & Matthews, P. G. (2015). Neurocognitive architectures and the nonsymbolic foundations of fractions understanding. In *Development of Mathematical Cognition* (pp. 141–164).

Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science*, *27*(2), 191–202.

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, *18*(8), 740–745.

Nasr, K., Viswanathan, P., & Nieder, A. (2019). Number detectors spontaneously emerge in a deep neural network designed for visual object recognition. *Science Advances*, *5*(5), eaav7903.

Norris, J. E., & Castronovo, J. (2016). Dot Display Affects Approximate Number System Acuity and Relationships with Mathematical Achievement and Inhibitory Control. *PloS one*, *11*(5), e0155543-e0155543.

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.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... Chen, M. (2012). Early Predictors of High School Mathematics Achievement. *Psychological Science*, *23*(7), 691–697.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*.

Vallentin, D., & Nieder, A. (2008). Behavioral and Prefrontal Representation of Spatial Proportions in the Monkey. *Current Biology*, *18*(18), 1420–1425.