Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Brain Research

Research Report

# A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network

Rakesh Sengupta[a], Bapi Raju Surampudi[a,c,*], David Melcher[b]

[a]Center for Neural and Cognitive Sciences (CNCS), University of Hyderabad, Hyderabad 500046, India
[b]Center for Mind/Brain Sciences (CiMEC), University of Trento, Trento 38068, Italy
[c]School of Computer and Information Sciences, University of Hyderabad, Gachibowli, Hyderabad 500046, India

ARTICLE INFO

*Article history:*
Accepted 9 March 2014
Available online 6 August 2014

*Keywords:*
Visual sense of numbers
Computational model
On-center off-surround
Neural network
Enumeration
Spatial attention
Individuation
Numerical cognition

ABSTRACT

It has been proposed that the ability of humans to quickly perceive numerosity involves a visual sense of number. Different paradigms of enumeration and numerosity comparison have produced a gamut of behavioral and neuroimaging data, but there has been no unified conceptual framework that can explain results across the entire range of numerosity. The current work tries to address the ongoing debate concerning whether the same mechanism operates for enumeration of small and large numbers, through a computational approach. We describe the workings of a single-layered, fully connected network characterized by self-excitation and recurrent inhibition that operates at both subitizing and estimation ranges. We show that such a network can account for classic numerical cognition effects (the distance effect, Fechner's law, Weber fraction for numerosity comparison) through the network steady state activation response across different recurrent inhibition values. The model also accounts for fMRI data previously reported for different enumeration related tasks. The model also allows us to generate an estimate of the pattern of reaction times in enumeration tasks. Overall, these findings suggest that a single network architecture can account for both small and large number processing.



## 1. Introduction

According to the theory of a visual sense of number (Burr and Ross, 2008a, 2008b), the ability to rapidly estimate the numerosity of a set of items reflects a basic, perceptual process. Developmental studies have shown that infants show an ability to distinguish between different numerosities at a young age (Xu and Spelke, 2000). Studies of non-human animals, such as cotton-top tamarins (Hauser et al., 2003), point towards the possible evolutionary origins of the visual

*Correspondence to: Current address: Cognitive Science Lab, International Institute of Information Technology (IIIT), Gachibowli, Hyderabad 500032, India.
*E-mail addresses:* qg.rakesh@gmail.com (R. Sengupta), bapiks@yahoo.co.in (B.R. Surampudi), david.melcher@unitn.it (D. Melcher).

http://dx.doi.org/10.1016/j.brainres.2014.03.014
0006-8993/© 2014 Published by Elsevier B.V.

number sense. It has also been shown that making decisions based on numerosity is possible even in societies where language does not have words for distinguishing larger numbers (Gordon, 2004). The visual number sense also seems to distinguish between a range of smaller numbers (the subitizing range: Kaufman et al., 1949) and the range of larger numbers. Numbers in the subitizing range are characterized by their rapid and confident enumeration with very high degree of accuracy. For larger numbers humans use one of the two strategies – (1) precise sequential counting, or (2) estimation, a rapid but inexact process of enumeration. Counting is precise but is much slower than either subitizing or estimation. Estimation, although faster than counting, is much less precise (Whalen et al., 1999). An important question, then, is whether subitizing and estimation share a common mechanism and thus possibly a common substrate.

To even begin to understand how to disambiguate the answer we have to address certain complexities. In the visual domain, it has been argued that enumeration can be linked to object individuation, a visuo-spatial mechanism that allows us to locate and track a limited set size of objects (Piazza et al., 2011; Melcher and Piazza, 2011). In contrast, another approach has been to characterize subitizing as estimation mechanism operating at small numbers (Gallistel and Gelman, 1992; Dehaene and Changeux, 1993). However, the different Weber fractions[1] over the two ranges suggest a fundamental distinction based on small or large numerosities (Revkin et al., 2008). Trick and Pylyshyn (1994) have suggested subitizing, unlike estimation, might employ pre-attentive mechanisms that index potential objects. Burr et al. (2010) have systematically manipulated both spatial and temporal attention to show that such manipulations indeed affect subitizing performance but not estimation, which led them to suggest that in addition to a possible pre-attentive mechanism that is active across all numerosities there is an additional attentive mechanisms necessary for enumeration within subitizing range.

In order to explore possible computational strategies for enumeration Dehaene and Changeux (1993) developed a model using a reinforcement-based supervised learning approach (with a proposed extension towards self-organization) to explain possible learning mechanisms in infants. However this model restricts enumeration only up to five items. As another possible solution Stoianov and Zorzi (2012) have trained 'deep' networks to use pixel by pixel information of images through unsupervised learning. They show that numerosity detectors can emerge in the highest level of the generative network. Their model could account satisfactorily for numerosity comparison task data in monkeys and human adults in the larger number (estimation, not subitizing) range. These numerosity detectors can be compared to the 'number neurons' reported by Nieder et al. (2002) and Roitman et al. (2007). Both the models mentioned above have complex structures that allow for learning of numerosity detectors. Another interesting model was developed by Grossberg and Repin (2003), based on an on-center off-surround architecture. This model involved an interaction between a spatial number map and semantic categories to explain error rates and reaction times in human numerosity comparison data. In the present work we wanted to investigate what properties a network of numerosity detectors should have in order to account for enumeration performance across both small and large numbers of items.

Building on work from Roggeman et al. (2010), we constructed a recurrent on-center off-surround network that receives a normalized pre-processed input and the output is the mean steady state activity of the network. This kind of network has been used to describe different kinds of phenomena in the domain of vision and working memory (Grossberg, 1973; Usher and Cohen, 1999). However any computational account of numerosity has to pass a few crucial tests: (a) the model simulations should qualitatively demonstrate how the different regimes of numerosity (subitization and estimation) emerge through internal network dynamics; (b) the model should be able to account for empirical findings like numerosity comparison data in human adults; and (c) it should be possible to make testable predictions from the model. In the following sections we will try to show how the proposed model fares under these conditions.

One advantage of our approach is that it does not start with dedicated number neurons that activate to specific number estimates, but attempts instead to see whether such capacities might emerge from a recurrent on-center off-surround network used more generally in perception of objects and scenes. A second aspect of our approach is that, rather than trying to fit the data with a model based on a large number of parameters, the proposed network has fewer number of parameters and the same network has been used to fit disparate data sets. Finally, we are able to generate novel testable predictions that emerged from the model itself regarding the link between sensory processing and numerical cognition.

## 2. Results

### 2.1. *Model*

We began by modeling critical features of a recurrent on-center off-surround network reported in Usher and Cohen (1999) and used in Roggeman et al. (2010). It is essentially a saliency map model based on the nonlinear leaky competing accumulation models (LCA) that have been used to account for performance in multiple-alternative choice paradigms (Bogacz et al., 2007). These models capture the recurrent on-center and off-surround nature of activity observed in neural systems (Grossberg, 1973; von der Malsburg and Buhmann, 1992) particularly in the visual modality.

The network consists of a single layer of completely interconnected nodes (Fig. 1). Each node corresponds to a neuronal assembly encoding an object or location of an object (or particular features) depending upon the cognitive phenomenon being modeled. The three main parameters that define the type of network are $\alpha$ (strength of self-excitation

[1]Weber fraction or Weber ratio refers to the minimum relative change in stimulus intensity in order for the stimulus level to be perceived as different from a reference stimulus intensity.