Influence of Excitation/Inhibition Imbalance on Local Processing Bias in Autism Spectrum Disorder

Yukie Nagai, Takakazu Moriwaki¹, and Minoru Asada

({yukie, takakazu.moriwaki, asada}@ams.eng.osaka-u.ac.jp)

Graduate School of Engineering, Osaka University

2-1 Yamada-oka, Suita, Osaka 565-0871, Japan

Abstract

People with autism spectrum disorder (ASD) tend to detect local patterns of visual stimuli more quickly than global patterns, which is opposite to the behavior of typically developing people. We hypothesized that the imbalance between excitation and inhibition neurons in the visual cortex causes the local processing bias observed in ASD. Stronger inhibitory connections could diminish the neural activities and thus prevent global feature integration, whereas properly balanced connections would enable the cortex to detect features of any size. We verified our hypothesis by employing a computational neural network called a neocognitron. Our experimental results demonstrated that the network with stronger inhibitory connections exhibited a local processing bias, whereas the network with properly adjusted connections showed a moderate global bias. Moreover, the networks with extremely strong or weak inhibitions revealed no perception bias. These results suggest that an excitation/inhibition imbalance causes multiple types of atypical perception in ASD.

Keywords: autism spectrum disorder; local processing bias; excitation/inhibition balance; neocognitron

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that is characterized by impaired social interaction and communication (Baron-Cohen, 1995). For example, people with ASD show difficulties in making eye contact, establishing joint attention, and reading others' intentionsabilities that are otherwise usually acquired in the first few years of life. Unlike the traditional view of ASD, recent studies in cognitive science, neuroscience, and Tojisha-Kenkyu² suggest that a core problem of ASD lies in the lower perception and action rather than in higher cognition (Frith & Happé, 1994; Happé & Frith, 2006; Ayaya & Kumagaya, 2008). A new hypothesis was proposed called "weak central coherence" (Frith & Happé, 1994; Happé & Frith, 2006), suggesting that a diminished ability to integrate information or an enhanced ability in the lower perception causes social difficulties in ASD. Researchers in Tojisha-Kenkyu has provided further insight into the mechanism. They argue that ASD may be associated with an abnormal sensitivity to prediction error, and thus individuals with ASD acquire a different internal model from typically developing people (Ayaya & Kumagaya, 2008). Difficulties in social communication may therefore be caused by the difference between the internal model of ASD and that of non-ASD rather than by disabilities intrinsic to ASD (Nagai, 2015).



Figure 1: Evaluation of local processing bias in ASD. The left figure shows the visual stimuli consisting of hierarchically compound letters. The right graph plots the reaction time to identify the local or global letter (adapted from (Behrmann, Avidan, et al., 2006)).

Behavioral evidence of a local processing bias in pattern recognition has emerged to support the above hypothesis (Behrmann, Avidan, et al., 2006; Behrmann, Thomas, & Humphreys, 2006; Simmons et al., 2009). People with ASD focus stronger and quicker visual attention to spatially local features than global features, whereas typically developing people exhibit the opposite behavioral pattern. Figure 1 shows the result of a behavioral experiment (Behrmann, Avidan, et al., 2006), in which subjects with or without ASD were asked to respond to a visual stimulus of a large letter (global) composed of smaller letters (local) by naming the local or global letter presented (left: the visual stimuli, right: reaction time). Subjects with ASD exhibited a stronger local bias in recognition, especially when the local and global letters were inconsistent (e.g., a large "S" consisting of small "H"s).

However, an open question remains as to the specific neural mechanisms involved in this atypical perception in ASD. Although a number of studies have revealed behavioral and neural abnormalities in ASD, no study has yet explained the underlying mechanism of these symptoms. We address this issue from a computational point of view. A computational approach has potential to bridge the gap between behavioral findings and neural evidence (Vattikuti & Chow, 2010; Bakhtiari, Mohammadi Sephavand, Nili Ahmadabadi, Nadjar Araabi, & Esteky, 2012). Our hypothesis inspired by the result of neuroscience studies, is that the imbalance between excitatory and inhibitory neurons in the visual cor-

¹T. Moriwaki is with Ishida Co., Ltd. since April 2014.

²A developing research area, in which researchers suffering from ASD investigate their atypical sensorimotor experiences from a first-person perspective.

tex causes a local processing bias in ASD. Previous studies have reported an atypical excitation/inhibition balance (E/I balance) and relevant gamma-band activity in ASD (Sun et al., 2012; Snijders, Milivojevic, & Kemner, 2013). Furthermore, the association between the E/I imbalance and social dysfunction was demonstrated in a mouse experiment (Yizhar et al., 2011). Based on this evidence, we consider that the E/I balance might influence the integration of spatial information. For example, stronger inhibitory connections could diminish the response of succeeding neurons and thus hinder global feature integration. We verify our hypothesis by employing a computational neural network called a neocognitron (Fukushima & Miyake, 1982; Fukushima, 1988, 2003).

Neocognitron: A Hierarchical Neural Network for Pattern Recognition

We introduce a neocognitron (Fukushima & Miyake, 1982; Fukushima, 1988, 2003) as the basis of our computational model. The neocognitron is a hierarchical, multilayered neural network, which is capable of robust visual pattern recognition. Figure 2 (a) shows the typical architecture of the network. Inspired by the structure of the primary visual cortex, the network consists of multiple layers: U_0 is the input layer that receives visual stimuli, U_G is the contrast-extraction layer that detects the edge features of the input, and U_{S1} to U_{C5} are the layers that recognize the patterns presented in the input. Visual stimuli fed into the network are processed through receptive fields (denoted by circles in Figure 2 (a)), which cover local areas in the lower layers and gradually integrate the information into the higher layers.

The network learns to recognize visual inputs by updating the connections between the U_S and U_C layers. Figure 2 (b) illustrates the connections from the (l-1)-th to l-th layers. U_S and U_C are alternately arranged in the network (i.e., $U_{S1} \rightarrow U_{C1} \rightarrow U_{S2} \rightarrow U_{C2} \dots$) and contain three types of cells: Scells, C-cells, and V-cells. S-cells serve as feature-extracting cells, similar to simple cells in the visual cortex. S-cells in U_{Sl} receive excitatory connections (a_{Sl}) from C-cells in the preceding layer and receive inhibitory connections (b_{Sl}) from V-cells in the same layer, and thereby learn to respond selectively to particular features in their receptive fields. The output of S-cells (u_{Sl}) is calculated by

$$u_{Sl} = \frac{\theta_l}{1 - \theta_l} \varphi \left[\frac{1 + \sum^{K_{Cl-1}} \sum^{A_{Sl}} a_{Sl} u_{Cl-1}}{1 + \theta_l b_{Sl} v_l} - 1 \right], \quad (1)$$

where θ_l is a fixed threshold, $\varphi[\]$ is a function defined by $\varphi[x] = \max(x, 0)$, K_{Cl-1} and A_{Sl} are the number of cell planes and the size of the receptive fields, respectively, and u_{Cl-1} and v_l are the outputs of C-cells and V-cells, respectively. Here, the connections a_{Sl} and b_{Sl} are the targets to update through learning and define the E/I balance, which is our point of interest with respect to the atypical perception in ASD. The learning rule of a_{Sl} is defined by

$$\Delta a_{Sl} = q_l c_{Sl} u_{Cl-1}, \tag{2}$$



(a) Typical architecture of the neocognitron



(b) Connections between U_S and U_C layers

Figure 2: Neocognitron

and b_{Sl} is calculated accordingly by

$$b_{Sl} = \sqrt{\sum^{K_{Cl-1}} \sum^{A_{Sl}} \frac{a_{Sl}^2}{c_{Sl}}}$$
(3)

so that u_{Sl} learns to respond selectively to particular visual features. Here, q_l is a positive constant parameter determining the learning speed, and c_{Sl} represents fixed connections from C-cells to V-cells. After learning, S-cells in the lower layers (e.g., U_{S1} and U_{S2}) are able to extract local features such as lines in particular orientations, whereas those in the higher layers (e.g., U_{S4} and U_{S5}) are able to extract global patterns such as letters, numbers, and diagrams consisting of multiple local features.

C-cells resemble complex cells in the visual cortex. That is, they receive fixed excitatory connections (a_{Cl}) from S-cells in the preceding layer, where S-cells are trained to extract the same feature but from slightly different positions. The output of C-cells (u_{Cl}) is calculated by

$$u_{Cl} = \Psi \left[\sum_{l=1}^{A_{Cl}} a_{Cl} u_{Sl} \right], \tag{4}$$

where A_{Cl} is the size of the receptive fields, and $\psi[]$ is a function defined by $\psi[x] = \varphi[x]/(1 + \varphi[x])$. An important point here is that C-cells respond if at least one of the connected

S-cells yields an output. C-cells thus spatially blur the responses of S-cells and make the network robust against position errors in visual stimuli. The mechanism of the neocognitron is described in greater detail in Fukushima and Miyake (1982) and Fukushima (1988, 2003).

Influence of E/I Balance on Local/Global Information Processing

We hypothesize that the E/I balance influences the local/global processing bias in pattern recognition. Figure 3 illustrates how the relative strength of the inhibitory connections b_{Sl} from V-cells affects the output of S-cells u_{Sl} in the neocognitron: (a) proper inhibitory connections, (b) stronger inhibitory connections, and (c) weaker inhibitory connections. We suppose that the neocognitron has been trained with visual stimuli, each of which contains a single number but with various types, sizes, and positions. The network is then presented with hierarchically compound numbers, as shown in the leftmost of Figure 3 (a), where multiple small number "3"s form a large number "2" according to their close positioning.

If the network has properly adjusted connections, it should be able to detect the numbers presented in the input image regardless of their sizes and positions. As illustrated in the rightmost image of Figure 3 (a), the output of S-cells, which receive proper inhibitory connections from V-cells, maintains the features of both "3" and "2" as the integration of multiple "3"s. The extracted features then enable the network to recognize both the local and global patterns in the succeeding layers. In contrast, stronger or weaker inhibitory connections cause local or global processing bias in the recognition. If inhibitory neurons have stronger connections, as shown in Figure 3 (b), they suppress the global response of S-cells so that they extract only the core features of the input. Under this scenario, only the local number "3" becomes visible, but not the global number "2", because of the sparse activation of S-cells. On the other hand, weaker inhibitory connections produce the opposite effect. Figure 3 (c) depicts how weaker inhibitions enhance the responses of S-cells and therefore diminish the local features of the input. In this case, only the global number "2" becomes visible by connecting the local number "3"s, while "3" per se becomes increasingly difficult to recognize.

Taken together, our hypothesis suggests that the atypical perception in ASD as well as the typical perception in non-ASD can be modeled by the E/I balance in the visual cortex. Of particular importance, the use of a unified architecture allows for not only the local processing bias but also another type of symptom in ASD to be modeled. Behavioral studies suggest two types of ASD symptoms: hyperesthesia and hypoesthesia. Hyperesthesia corresponds to the local processing bias because it exhibits increased sensitivity to perceptual stimuli. Hypoesthesia, on the other hand, corresponds to a super-global or no bias in pattern recognition (i.e., difficulty in pattern recognition), because it shows reduced sensitivity



(a) Proper inhibitory connections maintain both local and global features.



(b) Stronger inhibitory connections extract only local features.



(c) Weaker inhibitory connections emphasize global features.

Figure 3: Influence of E/I balance on local/global processing bias

to the stimuli. Our computational model thus reveals the underlying neural mechanism based on a common architecture: stronger inhibitory connections (i.e., a lower E/I balance) may cause hyperesthesia, whereas weaker inhibitory connections (i.e., a higher E/I balance) may cause hypoesthesia.

Pattern Recognition Experiment

Training and Testing Data

To verify our hypothesis, we conducted a pattern recognition experiment using the neocognitron. The visual stimuli used for the training and testing of the network are presented in Figures 4 (a) and (b), respectively. Each stimulus was a black and white image with a size of 130×130 pixels. The training data contained a single number per image, where the number was "0" to "9" drawn in different sizes and positions (for a total of 400 images). The neocognitron was trained in a super-



Figure 4: Training and testing data used in the experiment

vised manner in order to differentiate the numbers regardless of their sizes and positions.

The testing data were hierarchically compound numbers. All possible combinations of "0" to "9" were created (a total of 90 images). Note that the compound numbers were only used for testing, but not for learning. Thus, the network was not explicitly trained toward local or global processing but rather spontaneously acquired a bias to either local or global processing due to the balance between the excitatory and inhibitory connections.

The Neocognitron and its E/I Balance

The neocognitron used in our experiment consisted of an input layer U_0 , a contrast-extraction layer U_G , and five U_S-U_C layers. The number of U_S-U_C layers was determined based on learning performance. U_0 and U_{C5} were given as 130×130 C-cells and 10 C-cells, respectively, so that each cell in U_{C5} represented a single type of the input numbers. The size of the receptive fields was set from 4 to 8 pixels, depending on the layers.

The E/I balance of the neocognitron was manipulated to examine its effect on the local/global processing bias. We modified Eq. (3), which determines the relative strength of the inhibitory connections, as below:

$$b_{Sl}' = \alpha b_{Sl} = \alpha \sqrt{\sum_{l=1}^{K_{Cl-1}} \sum_{l=1}^{A_{Sl}} \frac{a_{Sl}^2}{c_{Sl}}}$$
(5)

so as to control the E/I balance by changing the coefficient α . Note that α was fixed over the learning and testing phases. If $\alpha = 1.0$, which is equivalent to Eq. (3), the network should acquire properly balanced inhibitory connections, as shown in Figure 3 (a). This network was expected to reproduce the behavior of typically developing people. If $\alpha > 1.0$ or $\alpha < 1.0$, the network should acquire relatively stronger or weaker inhibitory connections, as shown in Figures 3 (b) and (c), respectively. These networks were expected to reproduce the two types of ASD symptoms (i.e., hyperesthesia and hypoesthesia). In our current experiment, α ranged from 0.5 to 1.4 to observe the diverse influence of the E/I balance.



Figure 5: Influence of E/I balance on local/global information processing

Experimental Results Local/Global Processing Bias Caused by E/I Imbalance

We first examined the influence of the E/I balance on local/global information processing. Figure 5 shows the relationship between α and the local/global bias evaluated in the testing phase. The vertical axis Δu is defined by

$$\Delta u_{target} = u_{C5}(k_{target}) - \frac{1}{K_{C5}} \sum_{k=0}^{K_{C5}-1} u_{C5}(k), \qquad (6)$$

where $u_{C5}(\cdot)$ is the output of C-cells in U_{C5} , and K_{C5} is the number of C-cells in the same layer. In our experiment, $K_{C5} = 10$ because there were ten numbers to recognize (i.e., "0" to "9"). k_{target} was either the local or global number in the test stimulus. Therefore, Δu represents how strongly the local and global numbers in the compound numbers were recognized compared to the irrelevant numbers.

A two-way ANOVA revealed a significant interaction between α and the local/global bias. The Δu value for the global numbers was significantly higher than that for the local numbers when $\alpha = 0.9$ and 1.0 (p < 0.001), whereas the opposite result was found when $\alpha = 1.1, 1.2, 1.3, \text{ and } 1.4 \ (p < 0.001).$ These results suggest that a change in the E/I balance effectively reproduced the behavioral difference observed between typically developing people and people with ASD. According to Behrmann, Avidan, et al. (2006), and as shown in Figure 1, typically developing people tend to detect global patterns more quickly than local patterns. In our experiment, higher Δu values for the global numbers demonstrated such a global processing bias in typically developing people, where the strength of the output corresponded to a shorter reaction time shown in Figure 1. The properly balanced E/I connections (i.e., $\alpha = 0.9$ to 1.0) enabled the network to recognize the global numbers more strongly than the local numbers. In contrast, higher Δu values for the local numbers demonstrated the local processing bias in ASD. Inhibitory connections enhanced by $\alpha = 1.1$ to 1.4 forced the network to extract the local features more strongly than the global ones, which generated an ASD-like bias toward the local information.

Moreover, our results suggest that our computational model with a variable E/I balance can reproduce multiple types of ASD symptoms. ASD is characterized by both hypoesthesia as well as hyperesthesia in sensory perception. As mentioned above, hyperesthesia might be related to a local processing bias, whereas hypoesthesia might be linked to a super-global or no processing bias. When $\alpha = 0.5$ to 0.7, the neocognitron did not show any bias toward the local or global patterns. Indeed, the network could recognize neither the local nor global numbers. This inability to recognize the patterns also appeared when $\alpha > 1.4$. These results indicate that hypoesthesia in ASD shares a common neural mechanism with hyperesthesia, and that a change in the E/I balance produces the difference between these symptoms.

Internal Representation of the Neocognitron Acquired through Learning

Next, we analyzed the internal representation of the neocognitron to investigate the effect of the E/I balance on the network structure. Figure 6 presents the responses of (a) the whole network when $\alpha = 1.0$ and of (b) only U_{C4} when $\alpha =$ 0.5, 1.0, and 1.2. In this example, the compound numbers "3" (local) and "2" (global) were given as the input to the network.

Figure 6 (a) first demonstrates the hierarchical feature integration acquired in the network, which is a basic function of the neocognitron. Local lines in particular orientations were detected in U_{C1} , while more global patterns such as corners and longer lines were extracted in U_{C2} and U_{C3} . The response of U_{C4} demonstrated that the network recognized both the local and global features. The top three cell planes (denoted by squares) in U_{C4} showed wider activation in the image area, whereas the bottom four showed narrower activation at different positions. The black regions represented the preferred areas of the C-cells, where the C-cells became activated if particular patterns were detected. The existence of both wider and narrower activation in U_{C4} indicates that the network could recognize both the local and global numbers in the visual stimuli.

Figure 6 (b) shows the effect of different α values ($\alpha = 0.5$, 1.0, and 1.2) on the representation of U_{C4} . Comparing these three conditions reveals how the E/I balance affected the local/global information processing in the neocognitron. As described above, the neocognitron with a proper E/I balance (i.e., $\alpha = 1.0$) enabled U_{C4} to extract both the local and global features by creating cell planes responding to features of various sizes. In contrast, when $\alpha = 0.5$, only a single cell plane was created, which covered the whole region of the input image. The network with such a limited internal representation could not differentiate the visual patterns, because U_{C4}



(b) Response of U_{C4} when $\alpha = 0.5$, 1.0, and 1.2

Figure 6: Internal representation of the neocognitron after learning

always responded regardless of the type, size, and position of the input. When $\alpha = 1.2$, the network exhibited the opposite behavior. U_{C4} created many cell planes, each of which responded to a specific local pattern but not to a global pattern. Therefore, the network with stronger inhibitory connections showed a local processing bias similar to that of ASD. These analytical results reveal how the E/I balance affected the internal structure of the network and thus produced the local/global processing bias in pattern recognition.

Discussion and Conclusion

This paper has presented our computational experiments to verify the hypothesis that the E/I imbalance causes the atypical perception observed in ASD. A local processing bias and E/I imbalance are well known features of ASD; however, no previous study has demonstrated their association. Our computational model revealed that the E/I balance affects the internal representation of the visual cortex and thus produces a local/global processing bias in pattern recognition.

One of our most notable experimental results is that the unified neural architecture could reproduce multiple symptoms of ASD as well as typical non-ASD behavior by changing only one parameter (i.e., the E/I balance). A proper E/I balance led to a behavioral pattern representative of typically developing people, whereas a higher or a lower E/I balance resulted in hyperesthesia or hypoesthesia, which is representative of ASD. Only a change in the E/I balance generated the behavioral differences between typical and atypical visual perception. This result provides new insight into ASD and further supports a recent argument that the hyperesthesia and hypoesthesia of ASD are two sides of the same coin; although their behaviors appear different, they may share a common underlying mechanism (Ayaya et al., 2013). Our computational model is therefore highly plausible because of its unified architecture.

There has been another hypothesis about the neural mechanism underlying local processing bias. A magnocellular deficit or its abnormal activity might be a cause of the observed local processing bias in ASD (Sutherland & Crewther, 2010; McCleery, Allman, Carver, & Dobkins, 2007). The human brain has two parallel pathways for visual processing: the magnocellular pathway conveying the global and coarse information of visual input, and the parvocellular pathway conveying the local and fine information of the input. Therefore, a magnocellular deficit in ASD could result in difficulty of recognizing the global feature of a stimulus. Our results do not deny this possibility but instead provide another potential explanation for these observed ASD symptoms. Furthermore, our model has an advantage of reproducing multiple aspects of the ASD symptoms, instead of only the local processing bias, by employing a shared neural architecture. We intend to further investigate the relationship between different neural models for ASD.

Acknowledgements

This work is partially supported by MEXT/JSPS KAK-ENHI (Research Project Numbers: 24119003, 24000012, 25700027).

References

- Ayaya, S., Kawano, T., Mukaiyachi, I., Tojisha-Kenkyukai, N., Ishihara, K., Ikeda, T., & Kumagaya, S. (2013). *Tojisha kenkyu no kenkyu (in japanese)* (K. Ishihara, Ed.). Igakushoin.
- Ayaya, S., & Kumagaya, S. (2008). Hattatsu shougai tojisha kenkyu (in japanese). Igaku-shoin.
- Bakhtiari, R., Mohammadi Sephavand, N., Nili Ahmadabadi, M., Nadjar Araabi, B., & Esteky, H. (2012). Computational model of excitatory/inhibitory ratio imbalance role in attention deficit disorders. *Journal of Computational Neuroscience*, 33(2), 389–404.

Baron-Cohen, S. (1995). Mindblindness. MIT Press.

- Behrmann, M., Avidan, G., Leonard, G. L., Kimchi, R., Luna, B., Humphreys, K., & Minshew, N. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia*, 44(1), 110–129.
- Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: visual processing in autism. *Trends in Cognitive Sciences*, *10*(6), 258–264.
- Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind". *Cognition*, 50, 115–132.
- Fukushima, K. (1988). Neocognitron: A Hierarchical Neural Network Capable of Visual Pattern Recognition. *Neural Networks*, 1, 119–130.
- Fukushima, K. (2003). Neocognitron for handwritten digit recognition. *Neurocomputing*, *51*, 161–180.
- Fukushima, K., & Miyake, S. (1982). Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, *15*(6), 455–469.
- Happé, F., & Frith, U. (2006). The Weak Coherence Account: Detail-focused Cognitive Style in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5–25.
- McCleery, J. P., Allman, E., Carver, L. J., & Dobkins, K. R. (2007). Abnormal Magnocellular Pathway Visual Processing in Infants at Risk for Autism. *Biological Psychiatry*, 62, 1007–1014.
- Nagai, Y. (2015). Mechanism for cognitive development. In H. Ishiguro, M. Osaka, T. Fujikado, & M. Asada (Eds.), *Cognitive neuroscience robotics: A: Synthetic approaches* to human understanding. Springer (in press).
- Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, 49(22), 2705–2739.
- Snijders, T. M., Milivojevic, B., & Kemner, C. (2013). Atypical excitation-inhibition balance in autism captured by the gamma response to contextual modulation. *NeuroImage: Clinical*, *3*, 65–72.
- Sun, L., Grutzner, C., Bolte, S., Wibral, M., Tozman, T., Schlitt, S., ... Uhlhaas, P. J. (2012). Impaired Gamma-Band Activity during Perceptual Organization in Adults with Autism Spectrum Disorders: Evidence for Dysfunctional Network Activity in Frontal-Posterior Cortices. *Neurobiology of Disease*, 32(28), 9563–9573.
- Sutherland, A., & Crewther, D. P. (2010). Magnocellular visual evoked potential delay with high autism spectrum quotient yields a neural mechanism for altered perception. *Brain: A Journal of Neurology*, *133*, 2089–2097.
- Vattikuti, S., & Chow, C. C. (2010). A computational model for cerebral cortical dysfunction in autism spectrum disorders. *Biological Psychiatry*, 67(7), 672–678.
- Yizhar, O., Fenno, L. E., Prigge, M., Schneider, F., Davidson, T. J., O'Shea, D. J., ... Deisseroth, K. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, 477(7363), 171–178.