1. First example of the benefit of having parallel processing pathways is when visual signal transmit from retina to LGN, the diameter of axon bundle forms a bottleneck, thus a condensed parallel representation of visual scene gives the advantage of saving space and transmit information more efficiently. Another reason is that visual information is complex and redundant, parallel processing with modular computation offers higher speed. A third reason is that sometimes we want to separate different attribute of visual information and process them separately, as illustrated in the example of dorsal and ventral stream of visual processing(p367).

2. Originally, dorsal pathway is thought to process motion and depth, while ventral pathway is thought to do process object form and color(p366). These pathways are best understood as processing visual information for different behavior goals, in this case action guidance and object recognition. Certain visual feature like binocular disparity is processed in both of these pathways because they are relevant to the computational goal of both of these pathways.

3. Neuro response is enhanced when subject is attending to the preferred stimuli and suppressed when attending to the non-preferred stimuli. In Figure 1 graph a), the response of neuron is smaller when attending to non-preferred stimuli than when attending to a place outside its receptive field(p296). That shows the suppression effect.

4. Narrowing of tuning curve increase selectivity in the sense that the unit response less to neighboring stimulus and (probably) more strongly to the preferred stimulus. Multiplicative increase of response increase selectivity because according to information theoretical derivation information content in a neuron’s response is the highest when the slop of the tuning curve is the highest, multiplicative increase of response don’t change the width of the tuning curve but increase its maximum slope. Indeed a narrowing of tuning curve is consistent with this interpretation of selectivity because narrowing also increase the maximum slope of tuning curve.

5. Because the visual system cannot get rid of attentional influence of previous stages, and each time computation happens there can only be more attentional influence added to the response.

6. Featural attention refers to attentional effect based on visual feature, in this way multiple object in a visual field can have increased response if their feature is attended to. This challenges the “spotlight of attention” idea because spotlight is inherently spacial, and is incompatible to the observation of featural attention.

7. Computation happens in those later stages may include non-linearities. For example rectification, saturation, on-linear summation and so on, so a linear modulation of input can result in a non-linear modulation of output.