The Nature of Visual Attention

Wesley Sauret

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy.

Chapel Hill
2010

Approved by:

Ram Neta

William Lycan

Jesse Prinz

ABSTRACT

Wesley Sauret: The Nature of Visual Attention
(Under the direction of Ram Neta)

The central question dealt with in this paper is 'what is the nature of visual attention in primates?' After examining the scientific literature on the neuroscience of attention, it seems that attention in primates has two key features. First, its causal upshot is the modulation of a mechanism implementing divisive normalization. Second, this modulation has its causal origin in the frontal eye fields and the lateral intraparietal cortex. This is a causal description of attention, but those that find multiple realizability compelling are welcome to consider attention as having a more general function of strengthening representations. The causal description is then seen as describing the specific implementation of the function of strengthening visual representations in primates.

ACKNOWLEDGMENTS

I want to thank my advisor Ram Neta for patiently directing me through the sea of ideas I had about attention, to the topic that became the focus of this paper. I would also like to thank Nate Sharadin for reading a draft of this paper, and my readers, Bill Lycan and Jesse Prinz, for their contributions to this paper.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

1. Introduction

This paper will consider two answers to the question, 'what is the nature of visual attention in primates?'[1] The first answer highlights certain causal facts about attention; the second focuses on functional facts about attention.

We begin by examining the causal upshot thesis, which says that attention just is any process that has a certain causal upshot (or effect). If this thesis wants to avoid being overly liberal, attention needs to have a particular causal upshot that is uniquely identifiable as the upshot of attention.[2] This requirement is difficult to meet, so we will consider a natural objection to the view. In light of the objection, we will move on to a revised version of the thesis, called the causal process thesis. The causal process thesis says that attention is any process that not only has the right sort of causal upshot, but also has the proper causal origins as well. The causal process thesis thus claims that attention is a 'causal kind' like sunburns and footprints are. Finally, we will compare the causal process thesis to a functionalist thesis about the nature of attention.

These views about attention, as I understand them, rely on empirical research to determine what the specific causal upshot, causal origin, or function of attention actually is. In their current form, they make very general claims about what it is for a process to count

---

[1] It is worth pointing out that I included that qualifier "in primates" for a reason. I intend to provide a chauvinist account of visual attention in primates, without regard for what it might be for some alien species (or machine) to have some form of attention. Of course, the true aim is to give an account of attention in *humans*, but none of the neuroscientific data that I will discuss in this paper has been gathered from humans (that would be unethical). So while I am primarily interested in what attention is in humans, this account more accurately describes what attention is in *primates*, and possibly even mammals in general (depending on whether the brain regions important to attentional phenomena in primates have homologs in other mammals).

[2] I suspect that many neuroscientists investigating the neural correlates of attention hold a view like this.

as attention. But, in order to evaluate whether those general claims are right, we will need to evaluate them in the context of what our best empirical theories say about the causal and function aspects of attention. In the end, we will see that the causal process thesis (insofar as it is distinct from the functionalist thesis in primates) best fits the empirical research on the neural underpinnings of attention, and thus constitutes our best answer to the question of what the nature of attention is.

Since our first task is to evaluate the causal upshot thesis, we will first need to see what the empirical data say about the causal upshots of attention. In the section 2. we will review the past research on visual attention, and in section 3. we will look at some recent findings by Lee & Maunsell (2010).

2. A Brief Review of Past Results

Experiments on spatial attention in the visual cortex can be divided into two kinds: those that use a single stimulus and those that use a pair of stimuli. We'll start with the studies that have presented a single stimulus in a recorded neuron's receptive field. Results using this experimental paradigm indicate that, while the specific magnitude of attentional modulation varies with stimulus contrast (Reynolds et al. (2000), Williford & Maunsell (2006)), the attentional enhancement found at a particular stimulus contrast tends to be proportional (i.e., multiplicative) across the tuning curve, leaving the shape of the neuron's tuning curve unchanged (Treue & Martinez-Trujillo (1999), McAdams & Maunsell (1999)). The models of these effects are often called gain modulation models.[3]

Studies using a pair of stimuli in a neuron's receptive field have yielded more complicated results. When a preferred stimulus and a non-preferred stimulus both appear in

[3] Gain is an electrical term that describes the ability of a circuit to amplify the power or amplitude of a signal. Gain *modulation*, then, describes the ability of attention to amplify (or, in some cases, suppress) the firing rate of a neuron (or population of neurons).

the receptive field of a neuron, the neuron's response is a weighted average of its response to each stimulus individually. When attention is then directed toward the preferred stimulus, the response is enhanced. But when attention is directed toward the non-preferred stimulus we find a suppression of the neuron's response (Reynolds et al. (1999), Ghose & Maunsell (2008)). The net effect of attention (enhancement or suppression) in the paired stimuli condition depends on the selectivity of the neuron. Attention here does not simply enhance responses across the board, as is found in single stimulus experiments. Models used to explain the effects of paired stimuli in the receptive field typically offer an explanation in terms of a biased competition mechanism.[4]

Two recent papers have aimed to provide a computational model that can reconcile the results from these two types of experiments. In short, these papers suggest that the different effects found with one and two stimuli in the receptive field can be explained if the attentional modulation is the result of a response *normalization* mechanism (Reynolds & Heeger (2009), Lee & Maunsell (2010)). Normalization is a form of gain control in which neurons' responses are inhibited in proportion to the activity of neighboring neurons. Because normalization has a divisive effect on all of a neuron's responses, it can scale neural responses without altering the neuron's stimulus preference or selectivity. Since it can increase sensory gain without affecting tuning curves, it provides a *pure* form of gain modulation (Lee & Maunsell (2009)). The Normalization Model (as I will call it) aims to explain the nonlinear aspects of neuronal activity (e.g., the suppression found with two

[4] Biased competition has two parts, the biasing and the competition. The competitive aspect describes how when two stimuli are present in a receptive field, they compete with each other, thereby leading the neuron to respond at a rate that is the average of its response to each individually. The biasing describes the ability of attention to bias the competition in favor of one of the stimuli, thereby enhancing or suppressing the firing rate (depending on which stimulus is attended).

stimuli) via a mechanism that increases the inhibition a neuron receives as the strength of its input increases.

3. Recent Results

Lee & Maunsell (2010) wanted to evaluate the ability of the Normalization Model to account for the neurophysiological data on these differing effects of attention. But, since it was well known that increased task difficulty increases the strength of attentional modulation (Spitzer et al. (1988), Chen et al. (2008)), they performed a new set of experiments that used the same stimuli for both the single stimulus and paired stimuli paradigms and an experimental design that forced the monkeys to use same attentional strategy in both paradigms. They hoped that these additional controls would rule out any changes in task difficulty between the single stimulus and paired stimuli conditions.

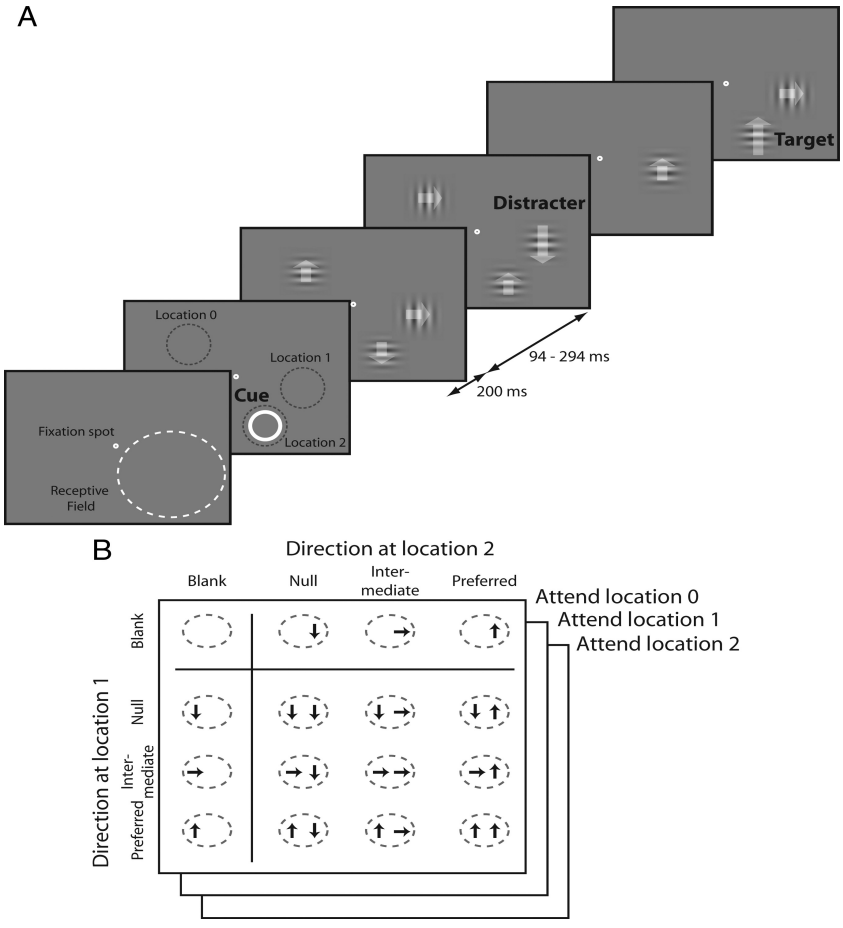

Figure 1. The experimental design. **A** shows the sequence of frames in a single trial. **B** shows the stimulus and attention conditions that were collected and analyzed. 48 conditions were collected for each MT neuron (16 stimulus configurations x 3 attention locations).

Lee & Maunsell recorded from neurons in macaque area MT that were sensitive to direction of motion. Figure 1 A shows the full task layout. The basic task setup had three locations where stimuli could appear. Two of the locations were within the recorded neuron's receptive field, and the third location was in the opposite hemifield. The monkey was presented with a succession of frames in each trial. Each frame lasted for 200ms and was followed by a blank period of a random duration (94-294ms). Within each frame, the specific number of Gabors[5] and the direction of each Gabor were randomly selected from a predefined set of possible combinations.[6] At the beginning of a trial the monkey was presented with a yellow circle in one of the three locations. This cued the monkey to attend to the stimuli that appeared in that location in each frame. The monkey's task was to respond when a Gabor with a different drift *speed* appeared at the cued location (the target), while ignoring any speed changes at uncued locations (distractors). On average, 1500+ frames were presented to the monkey for each neuron they recorded from.

---

[5] A Gabor is another commonly used stimulus in vision research. It is a sinusoidal grating that has been windowed by a 2D bell-curve function. The stimuli in Figure 1 A show what this looks like. In this experiment the sinusoidal grating component of the stimulus drifted in a specified direction.

[6] Since the number of Gabors presented on each frame was random, this made it unlikely that monkey was adjusting its attentional strategy each time a pair of stimuli appeared inside the receptive field (as opposed to the appearance of a single stimulus). The task design thus helped prevent the monkey from increasing their attentional effort when they were presented with two stimuli (which would be bad, since increased effort leads to increased attentional modulation).

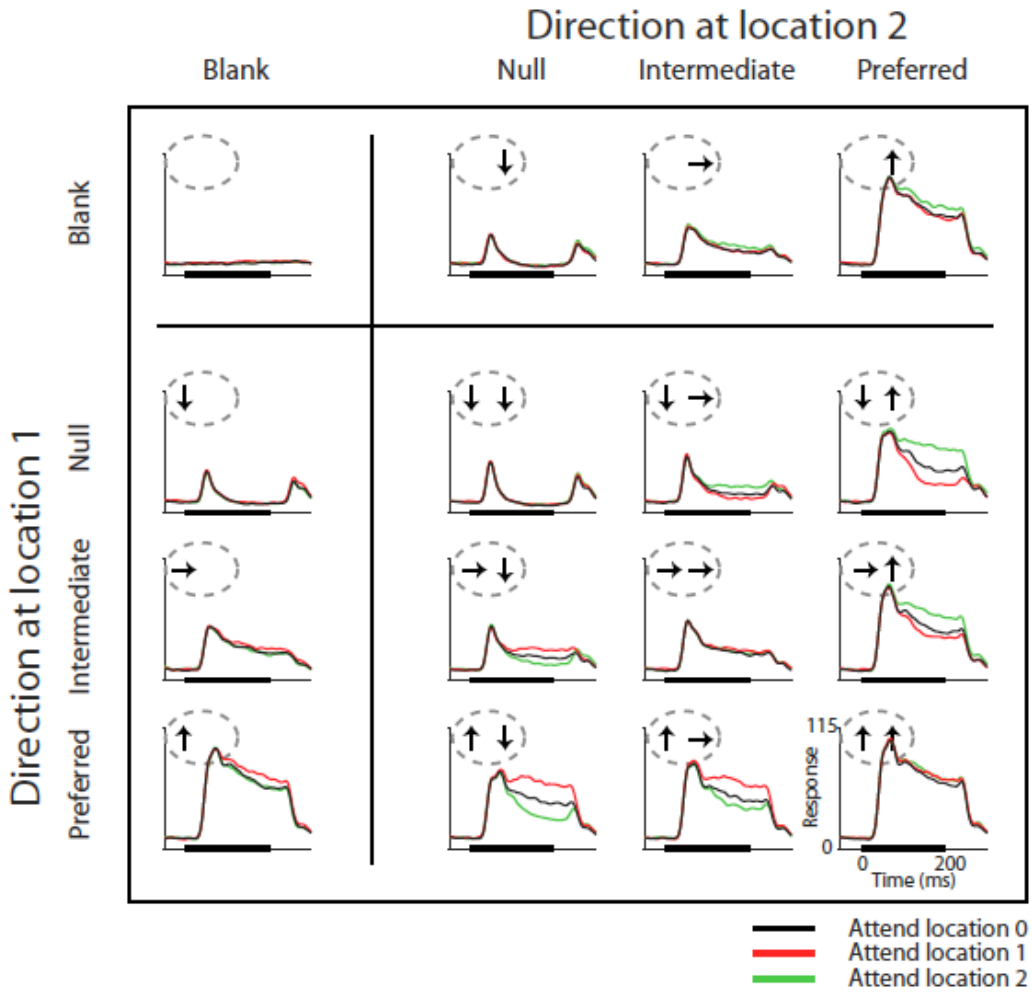

Figure 2. Average responses of 70 MT neurons. The black bars at the bottom of the each histogram indicate the presence of the stimulus (200ms).

Lee & Maunsell's findings are summarized in Figure 2. Each histogram shows the average responses to one stimulus condition (in the same arrangement as Figure 1 B). There are three colored lines plotted on each histogram, one for each of the three locations that attention was directed towards (see Figure 1 A for the three locations).

When a single stimulus appeared inside the receptive field (top row and left column), the neuron's response varied depending on the stimulus' direction of drift. Strong sustained responses were obtained in response to the preferred direction and transient responses at the onset and offset of the null direction. This response pattern is typical of direction selective

neurons with a bell shaped tuning curve. The red and green lines show the effect of directing attention toward each of the locations within the receptive field. Lee & Maunsell found that in the single stimulus condition, attention to the location of the stimulus caused a modest increase in response (red lines in the left column; green lines in the top row; ~9% response increase for the preferred direction).

When a pair of stimuli appeared inside the receptive field, the neuron's response was typically an average of the response to each stimulus alone (indicated by the black lines). This confirms the results obtained by Reynolds et al. (1999). Very strong attentional modulation was found when attention was shifted between the locations occupied by the pairs of stimuli in the receptive field. The strongest modulation occurred in the condition where one stimulus moved in the preferred direction and the other moved in the null direction. Attending the preferred direction increased the response by ~59%, relative to attending to the null direction. Attention to either stimulus (preferred or null) moved the response of the neuron *toward* the response it would have had if the attended stimulus appeared alone (as it does in the single stimulus conditions). But, attending to the stimulus in the paired condition did not move the response *fully* in that direction. Lee & Maunsell calculated that attentional modulation had less than half the effect of removing the unattended stimulus entirely. Ghose & Maunsell (2008) described a similar result in V4. This is an important result because it means that attention does not act to simply *filter* out the unattended stimulus. If attention filtered out unattended stimulus, we would predict it would have roughly the same effect as removing the unattended stimulus. Instead, Lee & Maunsell found it only has ~44% of the effect of removing the stimulus.

The results obtained here are largely in agreement with the results reported in earlier papers that used varying tasks and methodologies (see Section 2). At least two are worth

noting in particular. First, a key finding that Lee & Maunsell replicated is the moderate modulation of single stimuli that was found by McAdams & Maunsell (1999). Second, they demonstrated a very strong attentional enhancement of firing rates when attention is shifted between a preferred stimulus and a null stimulus in the paired condition, which was reported by Reynolds et al. (1999).

4. The Normalization Model

We now return to the issue of whether or not the Normalization Model can account for Lee & Maunsell's finding strong attentional modulation in the paired stimuli condition, but only moderate modulation with a single stimulus in the receptive field. Both Lee & Maunsell (2010) and Reynolds & Heeger (2009) have proposed similar normalization models that seem to be equally successful. Indeed, as Lee & Maunsell puts it,

> "Recently, others have proposed that attention may act through normalization mechanisms, and have presented detailed normalization models (Ghose and Maunsell, 2008; Boynton, 2009; Ghose, 2009; *Reynolds and Heeger, 2009*). These models are more elaborate than the one considered here, including, for example, terms that allow for variation in the spatial extent of attention. Because those models have similar form and more free parameters, they undoubtedly would perform at least as well in fitting the data we report here" (p. 3064, italics added).

Our discussion of the Normalization Model here will concern its general features, so we will be looking at the central features shared by both models.

The core idea of the Normalization Model is that attention reshapes the distribution of activity across the population of neurons, shifting the balance between excitation and suppression. Consider the case described in the last section: Lee & Maunsell presented two stimuli within the receptive field of a recorded neuron, one of the stimuli moving in the preferred direction, the other in the null direction. In that condition, they found that the response of the neuron was the average of its response to each stimulus individually (e.g., high for preferred, low for null, medium for the pair). In the Normalization Model, *only* the

preferred stimulus contributes to the excitatory drive of the neuron, but *both* stimuli contribute to the suppressive drive[7]. This creates a stronger suppressive input to the neuron, which is why the neuron's response to the pair of stimuli is lower than its response to the preferred direction (e.g., medium vs. high). Now consider what happens when attention is shifted into the receptive field and onto the preferred stimulus. Attention acts to multiply the excitatory drive to the neuron from the preferred stimulus, thereby acting to lessen the suppressive impact of the suppressive drive. According to Lee & Maunsell, this should lessen the net contribution of the null stimulus to the suppressive drive by ~44%. When attention is shifted to the null stimulus, attention acts to multiply the contribution of the null stimulus only to the suppressive drive (since it can't contribute to the excitatory drive). This results in a stronger suppressive drive, which leads to a reduction in the neuron's firing rate (Reynolds & Heeger (2009)).

Although our discussion of the behavior of the Normalization Model was at a purely descriptive level, the Normalization Model is first and foremost a computational model, powered by equations describing the firing rate of neurons in various conditions. Since descriptive models of past data run the risk of being mere ad hoc explanations of these findings, it would be nice to back up this description of its expected behavior with some demonstrative data. Luckily, the equations allowed Lee & Maunsell to simulate the activity of neurons in the various experimental conditions, which we can compare against the activity of recorded neurons *in vivo*. This gives them a powerful way of testing whether the model truly fits the data. Although a detailed description of the equations themselves is outside the scope

[7] In the Normalization Model, the suppressive drive is divisive, which is why it is able to suppress neuronal responses without affecting their selectivity. This contrasts with the multiplicative effect of attention, which can multiply both the contribution of a stimulus to the excitatory drive and its contribution to the suppressive drive.

of this paper, we will want to examine the results of Lee & Maunsell's simulations to determine how well their simple Normalization Model fit the experimental data.

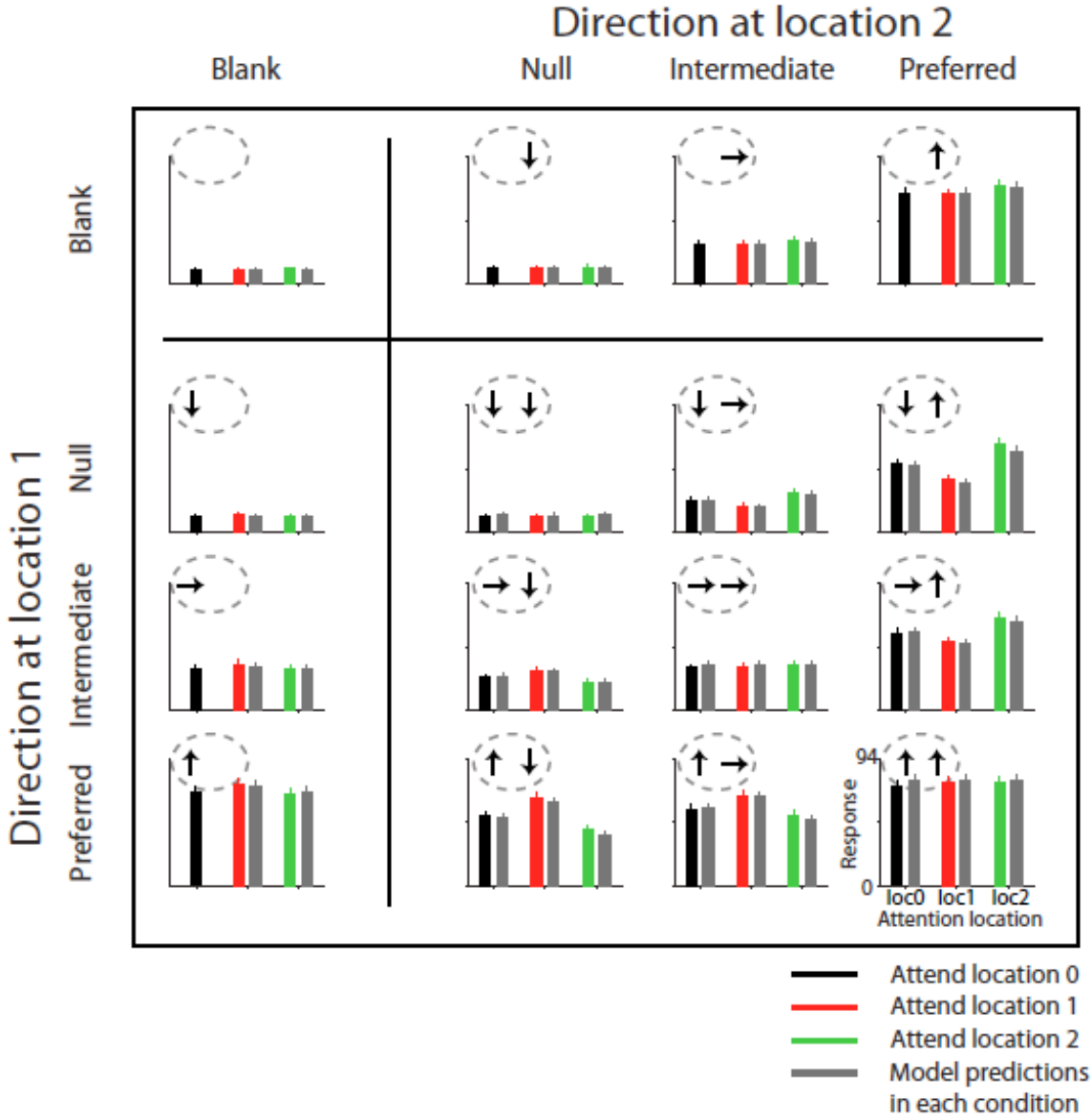

Figure 3. The Model's Fit. This shows a comparison of the outputs of the Normalization Model against the population average of 70 MT neurons. On this graph, the histograms from Figure 2 have been reduced to simple bars describing the average firing rate. Additionally, next to each bar graph is a grey bar indicating the model's prediction for the average firing rate in that condition.

As we can see from Figure 3, the model does a very good job of predicting the firing rates observed experimentally. Over all, the model predicted ~93.5% of the observed variance. For example, the predicted effect of attention on the response of neuron to a single

preferred stimulus was a 7% enhancement in the firing rate, and experimentally Lee & Maunsell found the enhancement was ~9%. The predicted modulation for shifting attention between the preferred stimulus and the null stimulus when both are in the receptive field was a 51% change, and experimentally they observed a ~59% change. Crucially, this means that the Normalization Model was able to predict the disparity in the strength of attentional modulation that is seen between the single stimulus paradigm (7%) and paired stimuli paradigm (51%). This was the main discrepancy in the data that we had set out to explain at the beginning of this section. It seems, then, that since the Normalization Model is highly successful at predicting the experimental data on the causal upshots of attention, we can say that these causal upshots are best described by gain modulation as implemented by a normalization mechanism.

5. How Common is Normalization?

Returning now to our task of evaluating the causal upshot thesis about the nature of attention, recall that this thesis says that attention just is the process that has a certain causal upshot. I take it that this claim is best understood as a claim about how we decide whether some process is an instance of the process of *attention* (as opposed to, say, ordinary "bottom-up" processing[8]). The causal upshot thesis says that we make this decision by looking to see whether the process has the specified causal endpoint. Now, the casual endpoint of attention that was suggested by our examination of the empirical data was divisive normalization. Thus, this thesis about attention says that any process that ends in normalization is an instance of attention. If the theory wants to avoid being liberal, the hope is that

[8] Throughout the paper, we should understand the term "bottom-up" processing to be compatible with both feedforward and feedback cortical circuitry. "Bottom-up" processing is supposed to contrast with "top-down" influences.

normalization is a unique form of processing that is found when, and only when, attention modulation occurs.

Unfortunately for the causal upshot thesis, normalization is a ubiquitous cortical processing strategy. From the perspective of a neuroscientist, this is a virtue of the Normalization Model, because this means that attention acts via a mechanism already known to be used elsewhere by the brain. The computational model is thus more biologically plausible as it avoids posting a novel mechanism that has never been seen before. From the perspective of the causal upshot theorist, this means that their theory is absurdly liberal, as it will call a vast set of processes *attentional* processes, despite the fact most of them have nothing to do with attention (e.g., the process may be computing how much force your hand should exert when grasping your cup of coffee).

In fact, normalization has been used to explain several phenomena throughout the visual cortex. It was originally used to explain nonlinearities in the responses of V1 cells (Heeger (1992), Carandini et al. (1997)), and later was used to explain the nonlinear response properties of other visual areas, such as MT in the dorsal stream (Simoncelli & Heeger (1998)) and IT in the ventral stream (Zoccolan et al. (2005)). Finally, normalization has been suggested as a mechanism to reduce redundancy in representing natural stimuli (Schwartz & Simoncelli (2001)). According to Reynolds & Heeger (2009), "normalization, therefore, has been proposed as a "canonical" neural computation" (p.159). It seems that the natural interpretation of this is that visual attention does not have a uniquely identifiable causal upshot. Instead, attention co-opts the already present "bottom-up" stimulus processing

mechanisms and, when it acts, it merely adjusts their functioning so as to modulate their gain.[9]

6. A Hypothetical Reply

Since normalization is just a description of a particular kind of computation, someone might argue that, contrary to my general claims above, the causal upshot of attention cannot be normalization *per se*. Instead, what we mean when we say that the casual upshot of attention is normalization is that the causal upshot of attention is the activation of some biophysical mechanism that *implements* the normalization computation. And, since the normalization computation can be implemented in several ways, the causal upshot view might be salvageable if we suppose that *attentional* normalization has a unique implementation of the normalization computation; one that is different than all the other implementations of normalization in the brain. Reynolds & Heeger (2009) mention this possibility,

> "Normalization might not have a single biophysical mechanism. It might instead emerge from a complex combination of a variety of mechanisms (Priebe and Ferster, 2008). Regardless of the mechanism(s), normalization appears to operate at multiple (perhaps all) stages of the visual system" (p.181).

This lends support two of the claims we've discussed: (1) there may be multiple implementations of normalization, and (2) normalization is ubiquitous.

I think that this reply makes an important point; the empirical data at this time does not completely rule out the possibility that the attentional uses a unique implementation of normalization. However, the empirical evidence makes it highly unlikely. The computational

---

[9] This is the important conclusion that I need for my argument against the causal upshot view. The argument succeeds as long as the mechanism that any new theory of attention proposes is common throughout the brain. Past models, such as simple gain modulation and biased competition, also described mechanisms common to many brain regions. Normalization is common, and I suspect that any better theory of attention that comes along will also describe a canonical computational mechanism (See the next section for more on this).

architecture of the cortex is *very* similar from one area to another. The types of neurons, their arrangements and connections, are highly stereotyped, and these stereotypes are repeated throughout the cortex (Mountcastle (1997)). This suggests (but does not prove) that each area is conducting computations of the same form (e.g., linear summation, divisive normalization, and spike threshold) (Reynolds & Heeger (2009)). Along these same lines, the few mechanisms that have been proposed as implementations of divisive normalization appear to be found throughout the brain. Let's look at one example. Tiesinga et al. (2004) have proposed a mechanism for implementing attentional gain modulation that uses interconnected networks of inhibitory interneurons. These fast-spiking inhibitory interneurons used are found throughout the cortex, and have been demonstrated to generate synchronous firing from the excitatory pyramidal neurons that they modulate (Cardin et al. (2009)). If the inhibitory interneurons generate both synchrony and normalization, then we should expect neural synchrony wherever such an implementation is used for normalization. This seems to be exactly what we find: synchronous firing of this sort is both correlated with attention and found throughout the cortex, just like normalization (Cardin et al. (2009)). Since neither synchrony nor normalization have been found to be uniquely correlated with attentional modulation, such implementations will not serve the purposes of the causal upshot theorist.

If it's very unlikely that attention uses a unique implementation of divisive normalization, then this gives us good reason to think that the causal upshot thesis about attention is seriously flawed. But, even though the thesis is flawed, it does seem to pick up on an important aspect of the nature of attention. It seems right to say that the causal upshot of attention (whether described causally or functionally) is a necessary part of what attention is. So, while the causal upshot thesis does do a good job of identifying a necessary condition

for what attention is (i.e., normalization), it does not give a sufficient characterization of the nature of attention.

7. How to Fix the Causal Upshot Thesis

What we would like from an explanation of the nature of attention, if possible, is a set of necessary and sufficient conditions. Our task now is to see if we can find some other factor that can be added to the causal upshot thesis to make it a sufficient characterization. To that end, the question we need to answer in this section is, since there is no uniquely identifiable "mark" of attention on visual processing, what is it that distinguishes the result of attentional modulation from those of ordinary stimulus processing? If we can answer this question, we will have a way to fix the causal upshot thesis.

The answer to this question will become clear by actually going through the process of trying to distinguish two very similar visual representations. Imagine that we have two visual representations, and one of the representations is the product of attentional modulation and the other is solely a product of "bottom-up" processing. The two representations are otherwise identical: they both have been normalized by the same mechanisms, they are both representations of the stimulus, the two representations are equally strong, etc. How are we going to determine which representation has been affected by attentional modulation?

The answer to this question is suggested by the initial description of the case itself. The description said that one is a result of ordinary processing and the other has been modified by attention. We can thus determine which one has been modulated by attention, if we look at the causal *origin* of the various components of the neural activity that constitutes the representation. On one hand, the representation that was modified by attention will have at least two components, one whose input to the normalization mechanism has its causal

origin in the brain regions implicated as the source of attention, and another whose input to the normalization mechanism was caused by the external stimulus itself. On the other hand, the representation that was a product of ordinary processing will have its source solely in the external stimulus. We now have the answer to our question: the way to determine which representation is the causal upshot of attention is to find out whether certain brain regions were the direct cause of a component of the input to the normalization mechanism. If they did causally participate, it has been modulated by attention; if not, it has not.

Since the causal origin of the input matters, this makes attention a causal kind, like sunburns and footprints. We can see the similarity by looking at what it is to be a footprint or sunburn. There are two criteria that need to be met for something, say an impression in the mud, to count as a footprint. First, the impression in the mud must be shaped like the foot of which it is supposed to be a footprint. Second, it must be a causal product of that very same foot (i.e., the foot of which it is supposed to be a footprint). Similarly, there are two criteria that must be met for something to count as an instance of sunburn. First, the damage to the skin must have certain characteristic properties. Second, the sun must have caused that damage to the skin.[10]

If either criterion fails to be met, the thing in question will fail to count as either a sunburn or as a footprint. For example, if we find an impression in the mud that is shaped exactly like a foot, but wasn't caused by a foot (but was instead caused by some mischievous person digging in the mud), then, even though the result is exactly the same as that of a footprint, this impression fails to be a footprint. Similarly, I might pour a chemical concoction on your arm that causes exactly the same sort of damage that overexposure to

[10] Or, if the cause is not our sun, some sort of solar radiation, such as UV rays, must be its cause. You might want to use this more specific claim if you think that an organism could get a sunburn in another galaxy.

the sun causes, but this damage still will not be sunburn since the sun didn't cause it. What this shows is that in order to be the sort of thing they are, these causal kinds must have a certain characteristic structure *and* a specific causal etiology.

Attention also has these two sorts of requirements. First, it must result in gain modulation via divisive normalization (i.e., it requires a characteristic structure). Second, the specific brain areas related to attention must cause it (i.e., it requires a certain etiology). If either of these conditions fails to be met, it isn't attentional modulation.[11] But, in order to fully flesh out this thesis, we now need to determine which brain regions are the causal origins of attention.

8. The Causal Origins of Attention

There are three classes of evidence concerning the causal origins of attentional modulation in visual cortex: correlational, anatomical, and causal. We will examine each of these in turn. To begin with, we have correlational studies, which can be subdivided into two groups. First, there are the brain imaging studies that have found activation in certain regions in humans when performing attentional tasks. Second, we have single cell recording studies in monkeys. The imaging studies have found several areas are correlated with attentional effort outside of the visual cortex, but most of the areas correlated with attention are not consistently found across studies. Two regions that are consistently highlighted as being activated during attentional effort are the frontal eye field (FEF) and the lateral intraparietal area (LIP) (Kastner & Ungerleider (2000)). Next, we will look at a recent recording study done by Gregoriou et al. (2009) in monkeys. They recorded from neurons in the FEF and in

[11] If this view is right, then it implies that "bottom-up attention" is not really attention *per se*, but a sort of processing that is often closely associated with attention, maybe because it often leads to attention capture. I think this is an issue worth investigating in the future.

V4. They found that attention to a stimulus inside of the receptive fields of neurons in both areas lead to enhanced oscillatory coupling between the FEF and V4, particularly at gamma frequencies.[12] This coupling appeared to be *initiated* by the FEF, which changed oscillatory frequencies ~8-13ms before V4 did, across a range of oscillatory frequencies. Considering the known conduction velocities and synaptic delays between these two areas, the time-shifted coupling at gamma frequencies is of the right delay to be able to optimize the postsynaptic impact of spikes from the FEF on V4. This study gives us a clearer picture of how it is that the FEF are correlated with the attentional effects that are found in visual areas like V4.

Next we have the anatomical evidence, which was already hinted at by the previous experimental results. Anatomical studies looking that the interconnectivity of various brain regions have determined that the FEF is both (1) directly connected to most of the visual cortex, and (2) indirectly connected to visual cortex via its connections to the LIP, which has projections to the visual cortex (Kastner & Ungerleider (2000)). These connections mean that it is anatomically possible for the FEF (and LIP) to be the causal origins of attentional modulation.

Finally, we have some direct causal evidence that the FEF can cause changes in V4 processing that are remarkably similar to those occurring with attentional cuing. Moore & Armstrong (2003) performed an experiment in which a microstimulation electrode was placed near FEF neurons that had overlapping receptive fields with neurons in V4 that they were recording from. They found that microstimulation of the FEF neurons enhanced the gain of the stimulus-evoked responses from the V4 neurons. Crucially, this enhanced the

---

[12] As was mentioned elsewhere, synchronous firing, particularly gamma oscillations, has been found to be correlated with attention.

gain of the V4 neurons in the same manner as the presentation of an attentional cue to attend to the stimulus in their receptive field. Together these three sets of evidence strongly suggest that the FEF, and the regions it projects to (such as the LIP), are the causal origins of attentional modulation.

9. Quick Summary of the Causal Process Thesis

The causal process thesis says that attention is a causal kind, like sunburns and footprints. This means that there are two requirements that must be met in order for a process to count as attention: it must have the right causal upshots, and it must have the right causal origins. By examining the available experimental data, we found that the causal upshot of attention seems to be modulations of a mechanistic implementation of divisive normalization, and that the modulations of this mechanism have their causal origin in the FEF and LIP. Thus, the causal process thesis says that attention is the process that begins with activity in the FEF and ends with modulation of the mechanisms implementing divisive normalization.

10. The Functionalist Thesis

Functionalist accounts of the different faculties of the mind are popular alternatives to causal accounts. We have already described a suitable candidate for a causal account of the nature of attention; so we will now see what a thesis that focuses on functional facts might say about attention. The functionalist thesis we are considering makes the very general claim that attention is any process that performs a certain function. One promising way of characterizing the specific function of attention draws on the empirical data that we have discussed in this paper. On this view, attention acts to increase the signal to noise ratio of visual representations in favor of the attended feature. Attention does this by selectively enhancing the responses of neurons selective for the attended feature and suppressing the

activity of neurons selective for other features. Increasing the signal to noise is a way of strengthening the representation of an attended feature. So, the functionalist thesis we will discuss says that the function of attention is to strengthen the representation of the attended feature by increasing the signal/noise ratio.

A natural response to this functionalist thesis is to ask whether it is empirically distinct from the causal process thesis. If it is, then what are some empirical predictions that it makes that are different than those made by the causal process thesis? Questions like these are difficult for the functionalist thesis to answer if it wants to be a viable *competitor* to the causal process thesis, because in answering them they run into a dilemma. On one hand, if the functionalist thesis does not make any empirical predictions that can differentiate it from the causal process thesis, then the functionalist thesis will be empirically identical to the causal process thesis (at least insofar as we are concerned with attention in primates). On the other hand, if the thesis does make empirical predictions that are different than those made by the causal process thesis, then it will end up in trouble. This is because the causal process thesis was specifically designed to reflect the empirical data on visual attention in primates. Any prediction the functionalist thesis might make for how attention works in primates that is substantially different than one made by the causal process thesis will likely end up deviating from the empirical data. Neither horn of this dilemma is particularly satisfying for the functionalist thesis that wants to be a viable alternative to the causal process thesis in primates.

On the first horn of the dilemma, the functionalist thesis basically amounts to the causal process thesis in primates. In a sense, this is a good thing given the tight relation between the empirical data and the causal process thesis. If the causal process thesis is seen as describing the specific *implementation* of the functionalist thesis in primates, then the casual

process thesis ends up being a precisification of the functionalist thesis. In this case, we should have no qualms about accepting the functionalist thesis over the causal process thesis. The only real difference between the two on this construal is that the causal process thesis is narrowly concerned with attention in primates and makes no predictions at all about what attention is like in machines, aliens, etc. The functionalist thesis, on the other hand, predicts that in other organisms attention will still have the function of strengthening representations, and so these organisms will have processes that count as attention if the process functions to strengthen representations. To the extent that we find multiple realizability compelling, we should prefer this version of the functionalist thesis over the narrower causal process thesis.[13]

On the other horn of the dilemma, the functionalist thesis hopes to remain distinct from the causal process thesis in primates. If the function of attention is to strengthen representations by increasing the signal to noise ratio, then, since there are many ways of implementing that function, the functionalist thesis in its general form will make a variety of false predictions unless it is able to specify an implementation for each homologous group of organisms. Presumably, all of the ways of implementing the function of strengthening representations are going to be incorrect except one (for each group).[14] Thus, what the functionalist thesis needs to do, to remain empirically viable, is specify the *correct* implementation of this function for the organism in question. If it does this successfully, then it won't have to worry so much about making false predictions. But once the functionalist is in the business of specifying implementations, they will describe the

---

[13] But, given the biochauvinist scope of this paper, I wish to remain agnostic regarding which of the theories to endorse.

[14] This assumes that attention is a single unified kind and has a single implementation of that kind.

implementation in primates (i.e., the causal process thesis), in which case they are back in the first horn of the dilemma.

11. Conclusion

The question with which we began was 'what is the nature of attention in primates?' After examining the scientific literature on attention we have settled on an answer to the question: The implementation of attention in primates has two key features. First, its causal upshot is the modulation of a mechanism implementing divisive normalization. Second, this modulation has its causal origin in the FEF and LIP. This is a narrow description of what attention is for primates. For those that find multiple realizability compelling, attention has the more general function of strengthening representations, and this describes the specific implementation of it in primates.

REFERNCES

Carandini M., Heeger D.J. & Movshon, J.A. (1997). Linearity and Normalization in Simple Cells of the Macaque Primary Visual Cortex. *Journal of Neuroscience* 17: 8621–8644.

Cardin J.A., Carlen M., Meletis K., Knoblich U., Zhang F., Deisseroth K., Tsai L. & Moore C.I. (2009). Driving Fast-spiking Cells Induces Gamma Rhythm and Controls Sensory Responses. *Nature* 459: 663-667.

Chen Y., Martinez-Conde S., Macknik S.L., Bereshpolova Y., Swadlow H.A. & Alonso J.M. (2008). Task Difficulty Modulates the Activity of Specific Neuronal Populations in Primary Visual Cortex. *Nature Neuroscience* 11: 974–982.

Ghose G.M. & Maunsell, J.H. (2008). Spatial Summation can Explain the Attentional Modulation of Neuronal Responses to Multiple Stimuli in Area V4. *Journal of Neuroscience* 28: 5115–5126.

Gregoriou G.G., Gotts S.J., Zhou H. & Desimone R. (2009). High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. *Science* 324: 1207-1210.

Heeger D.J. (1992). Normalization of Cell Responses in Cat Striate Cortex. *Visual Neuroscience* 9: 181–197.

Kastner S. & Ungerleider L.G. (2000). Mechanisms of Visual Attention in the Human Cortex. *Annual of Review Neuroscience* 23: 315–341.

Lee J. & Maunsell J.H. (2010). Attentional Modulation of MT neurons with Single or Multiple Stimuli in Their Receptive Fields. *Journal of Neuroscience* 30: 3058-3066.

Lee J. & Maunsell J.H. (2009). A Normalization Model of Attentional Modulation of Single Unit Responses. *PloS ONE* 4: e4651.

McAdams C. & Maunsell J.H. (1999). Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. *Journal of Neuroscience* 19: 431-441.

Moore T. & Armstrong K.M. (2003). Selective Gating of Visual Signals by Microstimulation of Frontal Cortex. *Nature* 421: 370–373.

Mountcastle V.B. (1997). The Columnar Organization of the Neocortex. *Brain* 120: 701–722.

Reynolds J., Chelazzi L. & Desimone R. (1999). Competitive Mechanisms Subserve Attention in Macaque Areas V2 and V4. *Journal of Neuroscience* 19: 1736-1753.

Reynolds J., Pasternak T. & Desimone R. (2000). Attention Increases Sensitivity of V4 Neurons. *Neuron* 26: 703-714.

Reynolds J. & Heeger D.J. (2009). The Normalization Model of Attention. *Neuron* 61: 168-185.

Schwartz O. & Simoncelli E.P. (2001). Natural Signal Statistics and Sensory Gain Control. *Nature Neuroscience* 4: 819–825.

Simoncelli E.P. & Heeger D.J. (1998). A Model of Neuronal Responses in Visual Area MT. *Vision Research* 38: 743–761.

Spitzer H., Desimone R. & Moran J. (1988). Increased Attention Enhances Both Behavioral and Neuronal Performance. *Science* 240: 338–340.

Tiesinga P.H., Fellous J.M., Salinas E., Jose J.V. & Sejnowski T.J. (2004). Inhibitory Synchrony as a Mechanism for Attentional Gain Modulation. *Journal of Physiology* 98: 296–314.

Treue S. & Martinez Trujillo J. (1999). Feature-Based Attention Influences Motion Processing Gain in Macaque Visual Cortex. *Nature* 399: 575-579.

Williford T. & Maunsell J.H. (2006). Effects of Spatial Attention on Contrast Response Functions in Macaque Area V4. *Journal of Neurophysiology* 96: 40-54.

Zoccolan D., Cox D.D. & DiCarlo J.J. (2005). Multiple Object Response Normalization in Monkey Inferotemporal Cortex. *Journal of Neuroscience* 25: 8150–8164.