Chapter7.

Closing Thoughts.

7.1 The processing of natural stimuli is favored by the brain

The goal of this thesis was to understand the role of attention in the processing of natural stimuli at both the behavioral and neuronal levels. As we mentioned at the beginning, much of our current understanding of attention has been shaped by how the visual system processes simple geometric shapes. However, given that most of these stimuli are rarely encountered on a daily basis, it is important to expand our investigations to the realm of natural stimuli. The experiments discussed in this thesis have shown that in contrast to the artificial geometric shapes on which many theories of attention are founded, natural stimuli are treated specially by the visual system.

We have observed that natural stimuli can be processed efficiently even when focal attention is not fully available to them. Following the original demonstration by Li et al. that animals or vehicles could be detected in natural scenes in the near-absence of attention (Li, VanRullen et al., 2002), we have shown that this processing ability holds for even finer discriminations of natural stimuli (in particular face stimuli), right down to the level of the identification of individuals. These results are surprising from a computational point of view since identifying a face constitutes an enormous challenge for machine vision systems. However, given that we engage in face identification countless times a day, it shouldn’t come as a surprise that the human visual system is well tuned for such tasks. Furthermore, it is advantageous for the visual system to be able to perform these routine tasks (such as face identification, gender discrimination, or animal or vehicle categorizations) with the least effort possible, in order to free up resources for the processing of unexpected and possibly threatening events and objects.

7.2 Underlying neuronal substrates for the efficient processing of natural stimuli

In contrast to the processing of natural stimuli, computationally simpler tasks (such as telling a bisected red-green disk apart from its mirror image) are nearly impossible for observers unless attention is engaged. This fact, that under identical conditions (i.e. dual-task, where the peripheral stimuli are presented in isolation) some tasks suffer while others are effortless, could reflect important aspects of the architecture underlying human vision. What about the structure of the visual system could support the processing of natural stimuli that we observe?

In Chapter 5, we presented evidence that demonstrates that natural and familiar stimuli are represented explicitly within single neurons in high level areas. The representations we observed were of an abstract nature, in that they corresponded to specific high-level, semantic concepts rather than a particular visual property. It is possible that the hardwired selectivity of neurons such as these, established through familiarity with the stimuli, could underlie the efficient processing we have observed in the dual-task experiments. A similar hypothesis was proposed by VanRullen and colleagues (VanRullen, Reddy et al., 2004). In other words, the activity of a neuron that specifically encodes all elements from the category “animal” could be sufficiently strong, even in the near-absence of attention, to perform an animal/non-animal discrimination task. Although we did not control for attention in the experiments described in Chapter 5, this premise is supported by several early monkey experiments, which have observed category specific responses in IT (e.g. to faces) (Gross, Rocha-Miranda et al., 1972). These studies were performed in anesthetized animals where attention was unlikely to have played a significant role in generating the responses. Thus, engaging the attentional resource could be superfluous to achieving good performance, in the case of single isolated stimuli.

According to this idea that natural scenes (presented in isolation) are exempt from attentional requirements because they have explicit neuronal representations, human vision should also be efficient for other stimulus features that are represented explicitly (VanRullen, Reddy et al., 2004). And indeed, there is evidence that provides support for this premise. For instance, simple orientation, color, or motion detection tasks are known to occur pre-attentively, and adequate neuronal representations for these features are found in early visual areas such as V1 and V5/MT (Baker, Petersen, Newsome, & Allman, 1981; Maunsell & VanEssen, 1983; Zeki, 1983). On the other hand, for stimuli that do not correspond to explicit neuronal representations, performance suffers tremendously unless attention is fully available. In these cases it is likely that attention could play the critical role of binding different features together (as proposed by the Feature Integration Theory, (FIT)) at the neuronal level to create the relevant representations.

Compatible with these theories, the data presented in Chapter 4 lend support to the idea that the near-absence of attention does not necessarily weaken the representation of isolated stimuli in their “essential nodes” (the specific regions in the brain that support awareness of particular facets of vision; see Chapter 5). In conditions when the stimuli were relevant to the observer, high levels of BOLD activity were observed in the relevant brain area (FFA), even when attention was not focused on the stimuli. Under the influence of top-down factors such as behavioral relevance, these high levels of activity could form a basis for the efficient processing we observe at the behavioral level.

These results have important implications because they challenge some aspects of the influential FIT. One of the strong premises of the FIT is that while the pre-attentive processing of low-level features can be directly supported by the visual system, complex object recognition tasks and their corresponding neuronal representations must rely on attention. Our behavioral data (along with previous results (Li, VanRullen et al., 2002; Rousselet, Fabre-Thorpe et al., 2002)) clearly contradicts this assertion. Additionally, at the cortical level, our fMRI study shows that strong cortical representations can be maintained in areas known to support high-level tasks even in the near-absence of spatial attention. Taken together, in contrast to the claims of the FIT, our behavioral and fMRI findings suggest that high-level complex stimuli, presented in isolation, do not pose a major problem for the visual system.

For everyday life, these data thus suggest that certain stimulus types, when relevant or meaningful to the observer, can automatically trigger their corresponding neuronal populations without engaging the relatively laborious focal attention mechanism. Given the limitations of the attentional resource, this is indeed a desirable property.

7.3 Why then, does processing fail on occasion when attention is unavailable?

The discussion so far has focused on the idea that natural stimuli can be processed outside the focus of attention without suffering behavioral deficits and without decreases in the associated neural activities. However, as we mentioned in the introduction to this thesis, in several situations, observers can be oblivious to objects (including natural ones) unless attention is directly involved. This phenomenon was also demonstrated in our change-blindness (CB) experiments (Chapter 6), where we observed that both behavioral performance and the neuronal signal were lower on incorrect trials when attention was less likely to be engaged. Thus, we are faced with an apparent contradiction: the data from the first portion of this thesis suggests that natural stimuli can escape the attentional requirement, but the last chapter seems to demonstrate just the opposite. What factors could explain this conflict?

7.3.1 Top-down behavioral relevance (or the lack thereof) during CB

This discrepancy is reminiscent of the apparent inconsistency between the findings of our fMRI study and many previous reports: we demonstrate that BOLD activity does not necessarily decrease significantly in the near-absence of attention, but other reports have observed substantial drops. As we discussed in Chapter 4, this contradiction can be explained by taking into account the top-down influence of behavioral relevance. Perhaps, the same factor can help address the differences between our dual-task and change-blindness results.

During the dual-task condition of the experiments described in Chapters 2 and 3, subjects always expected a stimulus in the periphery and knew which aspects of it were relevant to their task. For instance, in the face-gender discrimination experiment, subjects' expectations included the knowledge that the stimulus would be a face in the periphery, and that the gender of the face was the key factor. During CB, however, subjects' expectations were not similarly constrained—neither by prior knowledge of the location of the relevant picture, nor by the type of stimuli that would be presented, nor by which aspects could change (e.g. subjects could not know beforehand that an apple would be presented, or whether its color (or some other feature) or the entire apple itself could change). Consequently the top-down signal was likely to have been a much weaker influence in the change-blindness experiments. In the absence of this signal, we have shown in Chapter 4 that the lack of focal attention results in a significant decrease in cortical activity. Thus, similarly during change-blindness, neuronal activity could have been lower on the incorrect trials (when attention was less likely to be engaged) due to this lack of top-down expectation.

7.3.2. The effects of cluttered environments

Another factor that could explain the discrepancy in the effects of attention between the dual-task and change-blindness paradigms could be related to the manner in which the stimuli are presented in these experiments. In the DT experiment, the face stimuli were presented in isolation in the periphery. Monkey electrophysiology studies have shown that moving the focus of attention onto or away from (well-contrasted) isolated stimuli has a small effect on neuronal responses (Reynolds, Pasternak et al., 2000). Thus, during the DT experiment, the fact that focal attention had no significant effect on the representation of the face stimuli is compatible with these observations.

However, in the CB experiment, all stimuli were presented in cluttered environments. In these conditions, competitive interactions arise between stimuli, and it is known that focal attention significantly affects neuronal responses by biasing competition. As we discussed in the introduction to this thesis, presenting a “good” and a “bad” stimulus together introduces competition between the stimuli that weakens the influence of the “good” stimulus on the cell’s response. In the CB experiment, we presented each “preferred” stimulus along with 3 stimuli that we knew drove the cell weakly. Thus competition would be especially strong in this situation. Therefore, the differences in neuronal activity we observe between attended and unattended conditions during the CB paradigm could be a result of the stronger influence of attention in cluttered scenes.

7.4 Focal attention and behavioral relevance revisited

The differential effects of focal attention and behavioral relevance were encountered in the fMRI dual-task experiment in Chapter 4. It was remarked there that when stimuli are processed in the near-absence of focal attention (i.e. the central and dual-task conditions) top-down influences such as behavioral relevance can significantly impact the BOLD signal. In particular, we observed that even in the near-absence of focal attention, the top-down influence of behavioral relevance was sufficient for sustaining strong cortical representations of the stimuli.

However, the dual-task paradigm did not allow us to discern whether focal attention and behavioral relevance can have equivalent effects on neuronal activity. In other words, we could not investigate the reverse interaction between these two factors, namely if in the near-absence of *behavioral relevance*, focal attention could similarly enhance the activity of underlying neuronal populations. The change-blindness paradigm however affords us a look at this question.

As we have recently argued, during change-blindness situations, subjects' expectations of potential stimuli and events are relatively unconstrained and consequently it could be maintained that these stimuli are processed in the near-absence of behavioral relevance. Under these conditions therefore, the effects of manipulating focal attention can be investigated relatively independently of top-down influences. Accordingly, as we have demonstrated in Chapter 6, during change-blindness situations shifting the focus of attention to particular stimuli does indeed increase their underlying neuronal response.

Thus, resembling the effects of behavioral relevance in the near-absence of focal attention, focal attention appears to be able to strengthen neuronal representations in the near-absence of behavioral relevance. In other words, these two factors could work independently of each other, although they appear to have equivalent effects on neuronal representations. Engaging either one could be sufficient to strengthen and maintain representations of objects at the neuronal level, but in the absence of both factors, neuronal representations would tend to suffer.

7.5 A proposed framework for attention and the processing of natural scenes

Information from scenes rich in complex objects enters our eyes constantly to be processed by the visual system. In some cases, attention plays a critical role in this processing. In others, it is not fully necessary. In all cases, however, the relevant objects and events must be represented at the neuronal level so that accurate percepts can be generated. How does the visual system accomplish this task?

Based on the experiments described in this thesis, I propose the following framework for natural stimulus processing. This framework, which is summarized in Figure 7.1, is of course a simplified interpretation of our data, but I hope it will serve as a useful tool for summing up our conclusions. The assertions of the framework are not likely to apply to unnatural and unfamiliar stimuli that are not explicitly represented at the neuronal level (Chapter 5).

In the following framework, we consider two extreme situations in which an observer could encounter natural stimuli: in isolation, or as a part of a cluttered scene. (Since objects are normally encountered in a variety of conditions, these two extremes are used here only as a first approximation.) In these situations, depending on the various attentional conditions, different outcomes are possible as has been discussed throughout this thesis. These outcomes are illustrated in Figure 7.1 and detailed further below. Each of the following points corresponds to a particular situation (stimulus/attentional condition) as indicated by the corresponding letter in the figure.

A. When an object is present in isolation, and it is both attended to and task relevant or meaningful to the observer, strong cortical representations are created (peripheral-task condition of the dual-task paradigm; Chapter 4). These representations could provide the necessary support for efficient processing of these stimuli as was manifested at the behavioral level in the dual-task paradigm (Chapters 2 and 3).

B. Even if the focus of attention is not on the object, as long as it is behaviorally relevant to the observer, the integrity of the underlying neuronal representations is not necessarily compromised (dual-task condition in Chapter 4). Instead, the neuronal activities triggered by the presentation of the object appear to be strong enough to generate percepts of the object and support its processing (Chapters 2 and 3).

C. However, if the object is not meaningful to the observer, the underlying cortical representation is substantially weaker (central task condition in Chapters 2, 3, and 4). This is not to suggest that cortical representations are not formed for this object. Instead, the object still appears to activate the relevant neurons significantly above baseline. This activity could be maintained online, and in the event that the corresponding information is required by the observer, top-down signals or focal attention could strengthen the underlying neuronal signals (Chapter 4).

D. For objects that are presented in cluttered scenes in the near-absence of both top-down expectations and focal attention (chapter 6; see also point H), the strong competitive interactions between objects significantly weaken the corresponding neuronal signals, and these objects are consequently in danger of falling into oblivion.

E. In these cluttered environments and outside the influence of top-down signals, focal attention can play a critical role in object representation. By selecting out a certain subset of stimuli, focal attention effectively biases the neuronal competition in favor of these objects. Based on our findings in Chapter 6, and also several monkey electrophysiology studies, attention appears to have the effect of filtering out the influence of the competing objects. Thus to all intents and purposes, the influence of focal attention brings us back to the condition of the "isolated stimulus."

F. Since attention is now focused on this stimulus, even though top-down expectations are minimal (which makes this point different from point A), the activity of the relevant neuronal population is strengthened. As was shown in Chapter 6, these higher levels of activity could support the generation of accurate percepts for this stimulus.

G. Presumably, adding top-down signals, such as behavioral relevance to the existing influence of focal attention in cluttered scenes, would continue to bias the competition in favor of the attended and relevant object, again taking us back to the case of the "isolated stimulus" (as shown in A). This possibility is just speculation, however, since our data (Chapter 6) does not provide us with direct evidence to support this claim.

H. The independent effect of behavioral relevance in crowded environments is an open question. Under our experimental conditions, we could not directly examine the influence of this factor on neuronal activity (this factor was either absent altogether (Chapter 6), or present in conjunction with focal attention (Chapters 2, 3, and 4). However, it is plausible that these signals would affect crowded scenes and isolated stimuli similarly and favor the relevant objects thereby strengthening the underlying neuronal activity.

This framework is an attempt to succinctly summarize our findings about how the brain could process naturalistic stimuli. Some aspects of the framework might appear to be grossly simplified for, as we have said earlier, the proposed framework uses two extreme situations as an approximation for the myriad possibilities in real life. However, it is the larger picture that I would like to leave my readers with. The results of this thesis, along with a few other previous studies, seem to strongly suggest that the processes underlying the perception of natural objects are not mere extensions of the mechanisms we are familiar with from experiments on simple geometric shapes. Instead, the visual system seems to be well adapted for these stimuli and can often process them without calling upon additional resources such as focal attention. Hopefully, future studies will elucidate these processes further, and thereby bring us to a deeper understanding of how the brain tackles its constantly changing, rich visual environment on a daily basis.

7.6 Honestly, closing thoughts

It's another glorious day in Southern California. As we walk out of the basement, into an afternoon flooded with light, we are once again taken aback by the nature that surrounds us. Struck by the complexity of all that we (almost) effortlessly see, our thoughts turn to the attentional mechanisms and cortical areas that generate our experience of the world (at least they should, given that we’ve spent our morning mulling over the visual system). We realize with gratitude that we are fortunate these complex systems function normally in our lives — systems so complex that we still don't understand them. And we marvel at the design of it all.

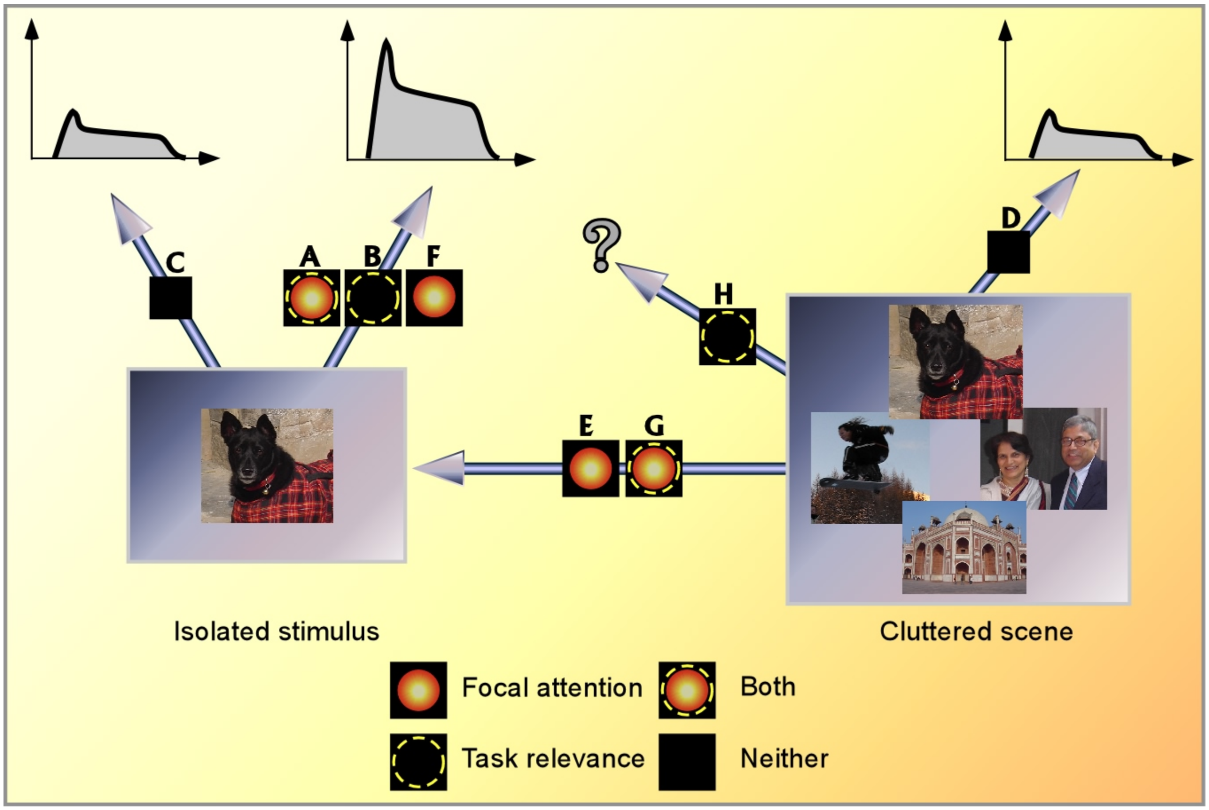

Figure 7.1 A proposed framework for the processing of natural stimuli when presented in isolation or in cluttered environments. Processing of stimuli in isolation is depicted on the left, and processing of stimuli in cluttered environments is shown on the right. Focal attention and top-down task relevance can have different effects on the neuronal representations of natural stimuli. For an isolated stimulus, the influence of focal attention and/or behavioral relevance is sufficient for eliciting high levels of cortical activity (A, B, F). If neither influence is available, neuronal signals are weaker (C). In cluttered scenes, the neuronal signal is weakened as a result of competition between stimuli (D). However, shifting attention to an object filters out the influence of unattended objects. This has virtually the same effect as presenting an object in isolation (E, G). The effect of top-down behavioral relevance to cluttered scenes in the absence of focal attention is an open question, since our experiments do not allow us to speak to this issue (H).