6.1 Summary and Conclusion

We make approximately 3 eye movements per second in order to sample the world around us. Eye movements are necessary for vision because visual acuity drops off sharply away from the fovea. The focus of this thesis is why we move our eyes where we do. Specifically, the work presented here examines eye movement control in natural scenes.

There are two main answers to the question of why we look where we do in a picture. The first is that eye movement control is a product of our cognitive knowledge structures and we look where we do in service of our ongoing task or goals. The second is that eye movement control is a product of the environment and we look where we do in order to react to sudden events or because we are attracted to particularly salient regions. In terms of natural scene viewing, these two answers form the basis of a debate that has been going on for years.

Many researchers believe that eye movement control is largely under cognitive control (i.e., top-down control) because there is overwhelming evidence to suggest that things such as our expertise, knowledge and task goals can account for most of the eye movements we make in a scene. However, this perspective tends to overshadow a very critical part of the equation – the scene itself. Salient regions of the environment can also influence eye movements (i.e., bottom-up control) in natural scenes. One of the challenges of studying eye movement control in natural scenes is that it is difficult to be sure why a certain region was fixated. For example, did you look at a particular region on the table by your door because that’s where you keep your keys, or did you look there because your keychain has a very bright Lego pirate that stands out relative to the otherwise nondescript background of your table?

There is a long established tradition for studying eye movement control that utilizes basic displays (such as lines and dots) with simple manipulations. In situations like this it is absolutely transparent whether eye movements are under top-down or bottom-up control. The work presented in this thesis applies knowledge and techniques from these more basic displays to natural scenes by imposing more control on the scenes themselves. This allows for the
extent to which eye movements are under top-down or bottom-up control to be more precisely measured.

In Chapter 2, we manipulated the contrast gradient across a natural scene such that one side of the image was increased in contrast relative to the other. We showed that early and short-latency eye movements are biased toward the higher contrast side of the image both when observers view the scene for a later memory test and when they are searching for a target within the scene. This confirms that eye movements early on in scene viewing are under bottom-up control (Foulsham & Underwood, 2008; Parkhurst et al., 2002) and aligns with research that utilizes more basic displays (e.g., van Zoest et al., 2004). This finding also supports the idea that salience influences eye movements early in time, while top-down control influences eye movements later in time—an idea that has been gaining traction in work utilizing more basic displays (Siebold et al., 2011; van Zoest & Donk, 2008).

Chapter 3 investigated whether scene gist (the finding that a scene can almost instantly be recognized; e.g., Oliva, 2005) could override the early effects of salience reported in Chapter 2. We showed participants a brief scene preview followed by a contrast manipulated image. Even though the scene preview led to an increase in the first fixation latency into the scene and the size of the first saccade, it was not enough to override the early effects of salience. Initial, short-latency eye movements were biased to land on the higher contrast side of the image. This suggests that even though scene gist can guide eye movements, it is not enough to completely override the early influence of salience.

The research in Chapter 4 looks beyond the initial moments of scene perception to investigate the role of salience in the presence of sudden object changes within the scene. An individual object in the scene was increased or reduced in salience. Critically this change occurred either during a fixation (which has long been known to capture attention) or during a saccade (which is essentially invisible to an observer as the change is masked by saccadic suppression). When the change occurred during a fixation it indeed captured the eyes, regardless of whether it was an increase or a reduction in salience.
When it occurred during a saccade, it only captured the eyes when it was increase in salience. This latter finding suggests that low-level salience can influence change detection, a phenomenon that is largely attributed to top-down control (Brockmole & Henderson, 2005a).

It has recently been discovered that when looking at natural scenes, eye movements tend to land in the center of objects (Nuthmann & Henderson, 2010). This is largely attributed to top-down factors (i.e., that eye movement control is object, rather than feature-based). Additional analyses in Chapter 4 and the experiments in Chapter 5 demonstrated a new role for salience that had yet to be discovered: it can influence within-object landing positions in natural scenes. In Chapter 4, when the object changed during a fixation (which meant that it occurred along with a very salient transient signal), within object landing positions were more central than when no change occurred or when the change occurred during a saccade. In Chapter 5, observers were shown scenes where an object had been manipulated such that it increased in luminance from left to right, or from right to left. In both a memorization and visual search task, fixation landing positions within these objects were biased toward the brighter side of the object, compared to landing positions within objects that were unmanipulated. The findings in Chapter 4 and 5 suggest that object-based targeting might work in a similar way to saccadic averaging in more basic displays (where eye movements tend to land in between two objects; see Van der Stigchel & Nijboer, 2011 for a comprehensive review).

Taken together, the findings of the experiments presented in this thesis make the role of salience in natural scene viewing explicit. By manipulating the salience of natural scenes and the objects within them, the constraints of top-down and bottom-up control in natural scenes can be more clearly defined. Eye movements in natural scenes are not just under top-down control (despite the enormity of evidence to suggest this) but can also be controlled by stimulus salience. Salience influences eye movements early on in scene viewing and affects within-object landing positions.