

Neurocinematics: An Empirical and Theoretical Investigation of Visual  
Attention and Event Segmentation during Dynamic Scene Viewing

An Undergraduate Thesis, Presented to  
The Faculty of the Program of Neuroscience  
In Partial Fulfillment of the Requirements for the  
Degree of Bachelor of Arts  
Bates College  
Lewiston, Maine

Jane Mayer  
William Seeley  
January 6, 2014



## TABLE OF CONTENTS

---

Title Page	1
Background	3
Introduction	4
Chapter One: The Idea	9
1.0 <i>An Introduction to the Psychology of Cinematics: A Brief History</i>	9
1.1 <i>Movie Realism: If it looks like it, it must be it</i>	11
1.2 <i>Cognitivism: An Introduction to its Importance</i>	13
1.3 <i>Cognitivism: Movie Perception ≠ Ordinary Perception</i>	14
1.4 <i>Cognitivism: The Ordinary Perceptual Processing Story of Movies</i>	16
1.5 <i>Movies: From the Discontinuous to the Continuous</i>	20
1.6 <i>Natural Vision: From the Discontinuous to the Continuous</i>	23
1.7 <i>So, What?</i>	24
1.8 <i>Where We Are, Where We Came From</i>	26
1.9 <i>Where We Are Going</i>	29
Chapter Two: The Commonality of Viewer Perception, A Look at Eyes, Brains, and Event Perception	30
2.0 <i>Introduction: An Outro to Theory and an Intro to Evidence</i>	30
2.1 <i>Watching Eyes Watch Movies</i>	31
2.2 <i>Watching Brains Watch Movies</i>	35
2.3 <i>The Events We All Perceive</i>	40
2.4 <i>Event Segmentation Theory (EST): The Framework</i>	45
2.5 <i>Neural Correlates of EST's Computational Model</i>	48
2.6 <i>Event Segmentation Theory: Perspectives and Caveats</i>	56
Chapter Three: A Necessarily Emotional Story	58
3.0 <i>Introduction</i>	58
3.1 <i>The Amygdala</i>	59
3.2 <i>The Amygdala and its Not So Necessarily Emotional Story</i>	64
3.3 <i>The Prefrontal Cortex</i>	66
3.4 <i>The PFC and Affective Processing</i>	68
3.5 <i>Perceptual Processing in the Dorsal and Ventral Pathways: An Overview</i>	71
3.6 <i>The Dorsal/Ventral Relationship during Affective Perceptual Processing</i>	74
3.7 <i>Motivation to Find the Puzzle Key: The influence of motivation on affective visual processing</i>	77
3.8 <i>To the Dual Competition Model, and Beyond!</i>	81
3.9 <i>Why is Affective Processing Important Within the Context of this Paper?</i>	84
Conclusion	85
References	89
Acknowledgements	99

## BACKGROUND

---

The overwhelming majority of what we know about human cognition stems from ‘unrealistic’ empirical research. Speaking loosely, this is due to studies’ use of ‘unrealistic’ stimuli. Of course, the word ‘unrealistic’ seems rather antagonistic, but it speaks to the fundamental difference between cognition in the context of a laboratory and cognition in the context of everyday. Within the context of a laboratory, the dynamic, perceptually replete environment is reduced into a well-controlled abstraction of the world. For example, visual attention researchers may show participants a picture of an angry man’s face to simulate the perception of an actual angry man. Indubitably, the reduction of a real angry man to a static 2D image eliminates a multitude of perceptual facets and thus provides only a partial visual processing story. Although this ‘abstract away confounds’ approach serves as a pragmatic means to sidestep the methodological constraints imposed by instrumentation that does not allow for real-time data collection within a contextually rich environment, it inherently limits our understanding of how human cognition actually functions; it provides us with an ecologically invalid understanding of how brains’ process information; and it is a roadblock towards our understanding of human cognition. Yet, the inherent unreality of current methodologies is not news to any cognitive researcher. Rather, it is an age-old problem that everyone knows about, but no one knows what to do about. And for the last few decades, it is a problem that has been easier to sidestep than confront. Yet, as we learn more about the brain’s cytoarchitecture and connectivity, it is becoming increasingly pressing that researchers find a more ecologically valid methodology to study real-time cognition within a contextually rich environment. And if neuroscientists need inspiration for this paradigm shift, they should probably start watching some movies.

Movies possess perceptually gripping and attention grabbing qualities that pull viewers in. These qualities, however, are not merely the result of a camera capturing movement and time. Movies utilize various editing techniques to intentionally *create* the viewers' experience. This, in turn, provides movies with the ability to sequentially trigger various perceptual, emotional and cognitive processes, as well as, capture and direct our attention (Konigsberg, 2007). And while researchers have remained relatively uninterested in examining the perceptual facets of movie viewing over the last 50 years, the last decade has brought a renewed interest in studying the psychological underpinnings of movies. Known as neurocinematics, this field of research aims to empirically analyze the experience of movie viewing—and what it shows us is that just because researchers cannot currently bring the laboratory into the dynamic, contextually rich world outside, doesn't mean that they can't bring the dynamic, contextually rich outside world into the laboratory.

## INTRODUCTION

---

According to the American Bureau of Labor Statistics (2012), we spend an average of 2.58 hours a day watching some form of edited moving images—movies, TV, etc. In other words, we spend more time watching edited moving images than we do reading, thinking, relaxing, playing computer games, exercising or travelling—or any other leisure activity for that matter. But *why* do we spend one-fifth of our waking lives in front of an illuminated 2D screen? What is it about these edited moving images that cause us to devote so much of our time to them?

The reason for this is because movies are powerful stimuli. Movies capture and maintain viewers' attention, and thus guide and control viewers' experience. (Although the use of the

word ‘control’ here may strike some as contentious, it is difficult to deny movies’ grasp on viewers—have you really never waited until a commercial break to go to the bathroom?) The perceptual power of movies is easily reducible to movies’ grasp on viewers’ visual attention. Through the capture of our attention, movies exert control over our perceptual experience; and because movies are strictly limited by their medium, it is solely through their perceptual qualities that movies attain their power to guide us through an evolving narrative and sequentially trigger various perceptual, emotional and cognitive processes within the viewer. This, folks, answers our question about why we watch movies: movies are compelling because of their perceptually salient, attention-grabbing qualities. And while this simple answer may satisfy the curiosity of a layperson, it marks just the beginning of this story—a story that until recently, scientists devoted little attention to.

With the advent of new technology, researchers are beginning to see the use of movie stimuli as an untapped methodology for studying perception within a more naturalistic setting—a contextually rich setting that allows for real-time measurements of perceptual processing. Armed with tools such as fMRI, eye-tracking, and computer vision, empirical investigations into movie perception represent a new frontier—a frontier of interest and importance to experts whose research is entirely unrelated to the movie industry. Yet, as the utilization of movie stimuli in neuroscientific research gains momentum, there appears to be a growing division between those who use movie stimuli and those who understand the current and potential value of movie stimuli—a division that cannot and must not continue. This uncoupling—the uncoupling of use and knowledge—will not only lead to erroneous utilizations of movie stimuli and misconceptions about their value, but it will also hinder researchers from cashing in on the true methodological potential of movie stimuli. The true value of movie stimuli isn’t that they merely

provide another—although more naturalistic—means to sidestep the field of neuroscience’s methodological impasse. The true value of movie stimuli is in their ability to illustrate that the solution to studying human cognition in the context of reality doesn’t necessarily require developing more mobile instrumentation to bring the laboratory outside. The lesson we learn from movie stimuli is that the brain uses a much simpler, reduced physics than the physics of our environment, and if we can identify the limits of its physics, we can construct viable methods to bring reality into the laboratory.

Created through the lens of human perception, the realism of movies is due to the relationship between how we perceive our everyday environment and how we perceive movies. And while there are clearly some vast differences between movie perception and everyday perception—we don’t call 911 or run out of the theater after we witness a murder—there are also fundamental parallels between these two modes of perception—since, we do cry when our favorite character dies. There is something real within movies. There is something within these intentionally constructed artifacts that tricks our brains. Of course, as noted earlier, this deception is limited to some extent, since no matter what happens on-screen, we’re not going to call 911. But, through an understanding of how the perceptual processing of movies and the perceptual processing of everyday coincide—through an understanding of how movies trick our brains and the extent to which they do so—we learn something invaluable about how our brains process perceptual information: we don’t need the reality of our everyday environment to experience it. Movies show us that all we need to perceive something as *real* are the necessary percepts constructed in a way that accommodates the ecological design of our perceptual processing system. The ideological basis of this type of approach towards perception isn’t a new idea though, it stems from research on the neuroscience of art; a field that acknowledges the dissociation between the

objective, physical reality of the world and our subjective perception of it. And the application of such an approach to understand movie perception isn't a new idea either—this is the type of account cognitivist film theory gives on movie perception.

Cognitivism is an empirical investigation of movie viewers' perception that differs from other psychological film theories in that it acknowledges the difference between movie perception and everyday perception. Cognitivism, rather than just haphazardly accepting our ability to perceive movies, seeks to provide an empirical understanding of the perceptual processes that account for our ability to perceive an illuminated 2D screen as something other than that. The naturalistic account of movie perception that cognitivism provides is framed in terms of movies' psychological realism—the idea that movie realism is the product of real-time cognitive processing. According to this account, what fundamentally differentiates our perception of movies from reality is that the phenomenological experience of movies is not a mimesis of real-life phenomenological experience. While the phenomenological unreality of the movie experience may seem like a drawback for the methodological value of movie stimuli, this review argues that such a limitation does not degrade the current value of using movie stimuli to study real-life cognitive processing, and that such a limitation is possible to overcome through the development of more advanced simulation equipment. Although no well-articulated argument for the latter occurs within the context of this paper, we suggest that the phenomenological unreality of the movie experience results from movies' inability to elicit the crossmodal visual enhancement effects that occur within everyday perception (Dieter, Hu, Knill, Blake & Knill & Tadin, 2014; Ramos-Estebanez et al., 2007); and that, although there is currently no simulation technology capable of producing the kinesthesia-induced visual sensations that we experience

within the context of the real-world, the development of simulation technology capable of eliciting such visual enhancements is possible within the future.

In support of cognitivism's account of movie perception—specifically the cognitivist account purported by Carroll and Seeley (2013)—the current review first delineates the theoretical and empirical viability of the cognitivist account. Then, using the cognitivist account to frame recent neuroscientific research on movie perception, this review argues that movies are artifacts of human perception; this not only provides an ecologically viable explanation for the perceptual power and widespread accessibility of movies, but also points to the fact that the human visual systems uses a much more reduced physics than the physics of the environment. This review concludes that movies are not only artifacts of human perception, but that they are more specifically artifacts of human visual consciousness.

The global purpose of this review is to advocate that the most viable solution to overcome the field of neuroscience's methodological impasse—an impasse that continues to hinder our scientific understanding of human cognition—is to adopt the ideologies that disciplines such as the neuroscience of art purport. Through the widespread adoption of such ideologies—ideologies that promote the dissociation of the reality perceived by the human brain and the objective reality of the environment—it is clear that the solution to field of neuroscience's current impasse is not necessarily the development of instrumentation that allows for mobile, real-time data collection within the context of the real-life environment. Rather, movie perception reveals that it is possible to artificially recreate the reality perceived by the human brain.



## CHAPTER ONE:

### THE IDEA

---

#### *1.0 An Introduction to the Psychology of Cinematics: A Brief History*

Inquiry into movie perception began in the early 1900s. The focus of such early investigations lay on the nature and impact of movie viewing on culture and society (Lindsay, 1915; Munsterberg, 1916). The foundation of these investigations, although not rooted in empirical inquiry, provided the backbone from which scientific investigations into movie perception spurred. And the substantial impact of these early thinkers on the development of psychocinematics<sup>1</sup> is evident from insights such as Hugo Munsterberg's:

whatever in nature or in social life interests the human understanding or human curiosity comes to the mind of the spectator with an incomparable intensity when not a lifeless photograph but a moving picture brings it to the screen (Munsterberg, 1916).

Insight such as this points to the type of ubiquitous qualitative understanding we seem to have about movies; these dynamic visual stimuli appear to contain some figment of reality. And while this qualitative understanding of movies speaks to their perceptual power, it was not until the late 20<sup>th</sup> Century that researchers began to quantify this power. During this time, the field of psychocinematics began to emerge as researchers made more empirical inquiries into the psychology of cinematics. Marking the transition away from a strictly theoretical analysis of movie perception towards a more empirical one, psychocinematics grounded itself in a scientific analysis of a viewer's aesthetic response to movies (Shimamura, 2013); simply put, an aesthetic response refers to "a hedonic response" or "a preference or liking judgment" (Shimamura, 2013). Thus, the advent of psychocinematic marked the first time that science and theory came together

---

<sup>1</sup> Although there is not an established field called "psychocinematics," this word is used throughout this paper to denote the interdisciplinary conglomerate of those that study movie perception; it is necessary to note that this term comes from Arthur Shimamura's (2013) textbook, "Psychocinematics."

to elucidate the unique and powerful quality of the film medium. Today, the study of psychocinematics:

considers the psychological and biological underpinnings of our movie experience, addresses the aesthetic nature of movies from an empirical approach, develops movies for psychological research as a means of understanding cognitive processes as they unfold dynamically in time, [and] encourages a multidisciplinary approach that includes psychology, cognitive science, neuroscience, philosophy, film studies, filmmaking, art history, sociology and other related endeavors” (Shimamura, 2013).

As such, what contemporary psychocinematics shows us is that our qualitative understanding of movies as containing some figment of reality is right—there really is something inherently and quantifiably realistic about movies. And as the empirical study of movie perception proliferates, it is becoming increasingly obvious that the field of neuroscience needs to start paying attention.

To prove this, this chapter will explore the relationship between movies and reality. Beginning with a look at why movie realism—the perspective that movies are a mimesis of reality—fails to account for the perceptual power of movies, this chapter will argue that cognitivism provides a more viable, ecologically based and empirically backed account of the reality within movies. The cognitivist account advocates that movies are experienced as psychologically, but not phenomenologically, real and that as a result of movies’ psychological reality, the human brain cognitively processes movies in the same way it cognitively processes reality. Through an understanding of movies in this way, it will become clear that movies are more than just dynamic visual stimuli—movies are artifacts of human perception; and because of this, the utility of movies within neuroscience extends beyond their value as visual stimuli. The aim of this chapter is to show that perceptible, human-made objects are of just as much value to an understanding of the human brain as direct probes of the brain. Thus, let us begin to decompose the relationship between movie and brain.

### 1.1 *Movie Realism: If it looks like it, it must be it*

According to movie realism, movies are thought to mimic reality. As a mimesis of reality, movies' ability to elicit psycho-physiological responses in viewers is explained in terms of the fact that they are experienced as illusions of reality. And in a folk intuition type of way, this makes sense (in fact, it is difficult to explain how else a 2D illuminated screen could make us cry).

However, this perspective tells a rather incomplete story about the cognitive and emotional grip movies have on viewers. This is apparent when we see that the adoption of such a view—a view that movies are experienced as illusions of reality—entails *all* the facets of real-life experiences. This means that the experience of a movie must parallel the psychological, physiological and behavioral aspects of a real-life experience—and this simply just isn't how we experience movies. The surface implausibility of this type of experiential mimesis between movies and real-life is easy to see when we think about how we respond to movies. I mean, do you really call 911 after witnessing someone get murdered in a movie? Or run out of the theater when you see someone on-screen walking towards you with a knife? No, our phones remain in our pocket and our butts in our seat. Thus, there is a clear difference in our behavioral responses to movie events and to real-life events.

But, does the difference in our behavioral responses to movie events and real-life events mean that the illusion thesis is totally wack? Not necessarily. According to Carroll and Seeley (2013), the illusion thesis—the view that we experience movies as illusions of reality—is reducible to two types of illusions, cognitive illusions and perceptual illusions. To accept that movies are experienced as cognitive illusions—illusions in which “we believe what we see is true”—

requires the behavioral responses that we have already illuminated as absent during movie viewing (Carroll & Seeley, 2013). If viewers believed that movie events were real, there would be chaos ensuing in movie theaters around the globe. Clearly, this is not the case. The view that we experience movies as perceptual illusions, however, holds more weight. Unlike cognitive illusions, perceptual illusions remain unaffected by cognitive processes. That is, the experience of a perceptual illusion remains unchanged whether the viewer believes that they are in a movie theater or in the middle of a bloody shooting. In this way, perceptual illusions are cognitively impenetrable (Carroll & Seeley, 2013). The perceptual illusion component of the illusion thesis, however, still falls short in explaining viewer perception.

The shortcoming of trying to explain the movie-viewing experience in terms of perceptual illusions is the realist hypothesis itself. This is because, although there does appear to be something perceptually illusive about the movie-viewing experience, the realist hypothesis is that we experience perceptual illusions in the same way that we experience reality. This would entail that we are actively experiencing movie events along with the unfolding narrative. And again, this just isn't the case. During the movie-viewing experience, the viewer adopts a perspective that is drastically different than their everyday, egocentric perspective. This means that, while we egocentrically orient ourselves to the objects and events of our everyday experience, we don't orient ourselves to the narrative's unfolding events; we don't bodily treat these two experiences as equal and therefore, as Carroll and Seeley (2013) put it, "even if in some sense our eyes are fooled, our bodies aren't."

Yet, as just stated, there really does seem to be something perceptually illusive about the movie-viewing experience. In this way, the shortcoming of the realists' hypothesis is more qualitative in

nature than anything else; its illusion postulate goes beyond movies' actual illusionary aspects—illusions of movement, objects and dynamic events—and in doing so, nullifies its viability. Just because movie realism is too ambitious in its illusion postulate, however, does not mean that the realistic qualities of movies cannot be understood in terms of illusions. Rather, what the failure of movie realism shows us is that there *are* fundamental differences between movie perception and ordinary perception. And because of this, any viable explanation of the reality within movie perception must necessarily acknowledge the unreality within movie perception. This means that we don't need to drop the idea that movies' possess illusive qualities. This means that a viable explanation of movies' illusive qualities must necessarily acknowledge that there are key differences between our perception of movies and our perception of everyday. The reason that the realist account falters in comparison to the cognitivist account of movie perception is exactly this. In many ways, the cognitivist account is actually built off of trying to understand how these two modes of perception—movie perception and ordinary perception—overlap and how this overlap lends an understanding to human perception in general. As a result of this approach, cognitivism offers a realistic account of the reality within movie perception and for this reason is the focus of the rest of this chapter.

## 1.2 *Cognitivism: An Introduction to its Importance*

The reason cognitivism is vital to this paper is that provides us with a holistic, ecologically viable, empirically based account of the relationship between movie perception and ordinary perception. And although a coherent argument for the utility of movie stimuli within the field of neuroscience could be made without the cognitivist account all together, it is through the lens of cognitivism that the utility of movie stimuli can be seen as more than just a means to yield better results—it is through the lens of cognitivism that we are able to understand that movie stimuli

represent more than just a means to sidestep the methodological impasse facing brain studies. Rather, what the cognitivist account shows us is that through an in-depth theoretical and empirical understanding of movie perception, movie stimuli don't just ameliorate the gap between the laboratory and the reality outside of it, movie stimuli provide a means to bridge this gap. To understand how cognitivism lends itself to such conclusions about the utility of movie stimuli, the following discussion delineates i) how movie perception diverges from ordinary perception, ii) how, regardless of these divergences, movie perception is easily explained by current models of ordinary perceptual processing and iii) what this shows us about ordinary perception.

### *1.3 Cognitivism: Movie Perception $\neq$ Ordinary Perception*

Cognitivism, as opposed to film realism, is an empirical investigation of movie perception that acknowledges the differences between movie perception and ordinary perception, and seeks to reconcile these differences in terms of the human brain's natural perceptual capacities. As a naturalistic approach to movie perception, the focus of cognitivism is to understand movie perception in terms of neuronal processing.

Arguably the most important aspect of a cognitivists' approach is the recognition that our perception of movies is critically different than our perception of everyday. Unlike ordinary perception, movies:

- contain only audiovisuo information; hence, are perceptually sparse
- represent space as discontinuous
- contain no egocentric orienting information
- utilize a point of view that defies the spatiotemporal limitations of real life

Markedly, it is the concession of differences between movies and the real world that instills cognitivism with a sort of explanatory drive—for it is the cognitivists' job to explain how movies, whose salient information differs in form than that of ordinary perception, are processed both similarly and differently than the processing of real-live information.

The cognitivist model purported by Carroll and Seeley (2013) is a good starting point for our present discussion because it is not only explicit in its claim about the reality within movie perception, but also the reason for this. The general claim of this model is:

Movies function as attentional engines intentionally designed to focus perception on those aspects of the depictive scaffolding of shots and scenes diagnostic for their narrative content and meaning. The information structure of the movie will, therefore, perhaps counterintuitively, closely match the information structure of real-time experience (Carroll & Seeley, 2013).

The power of this claim is in the empirical models of ordinary perceptual processing Carroll and Seeley (2013) use to ground it. These three well-established perceptual processing models of everyday perception provide us with a naturalistic perceptual processing story about how movies are perceived; which is of utmost importance because this allows for the contextualization of movie perception within everyday perception, as well as the contextualization of everyday perception within movie perception. Through this we are able to see that the two types of perception function symbiotically in elucidating the general reality of human perception; the reality that movie perception and ordinary perception are more similar than most can imagine.

In order to see the power of cognitivism's account of movie perception and its value for the study of neuroscience, the following section will delineate the perceptual processing models that ground the account's claims: diagnostic object recognition model, a biased competition model for selective attention, and a model of everyday visual routines. By providing a three-part

processing story about ordinary perception, the model elucidates that movies obtain their realistic, experiential power through the manipulation of naturalistic perceptual processes. Through this manipulation, the mimesis between movie perception and ordinary perception shows us that certain aspects of reality—in this case, visual perception—are able to be intentionally constructed; and while we still see that the mimesis between movie perception and ordinary perception is currently limited in its realism to certain aspects of visual perception, this is likely a medium-specific constraint that is possible to overcome. Why this is of importance to the field of neuroscience is because it points to the theoretical possibility of constructing reality within the laboratory—not a simulation that is like or close to reality, but reality itself—and finally being able to study the human brain within the dynamic reality it evolved to cope with.

#### 1.4 Cognitivism: *The Ordinary Perceptual Processing Story of Movies*

The object recognition story Carroll and Seeley (2013) tell begins with the subunits of movies, camera shots. According to Carroll and Seeley (2013), these basic building blocks—camera shots—are ‘recognitional prompts.’ What this means is that the camera shots are stimuli with a structure that is sufficient to allow a viewer to identify their descriptive content. The viewer identifies a recognitional prompt’s descriptive content by comparing declarative knowledge about the stimulus’ structure and function with its formal and compositional content (Carroll & Seeley, 2013). In this way, when we think about the function of camera shots within a movie, it is clear that camera shots do indeed function as recognitional prompts.

As such, camera shots are pictorial representations tailored to communicate *something* to the viewer. What that something is, depends on the form and amount of visual information within



the shot; both of which can vary tremendously between different shots, but each of which contains visual content that is specifically tailored to its communicative purpose. And since movies' pictorial representations are intentionally created, their image structure is representative of the choices—about detail, phenomenal content, etc.—that went into its creation. This means that each camera shot, each pictorial representation, in a movie is the product of an innumerable number of choices made by the filmmaker in order to produce an image structure that is tailored to visually convey the shot's descriptive content.

The one way that the filmmaker is limited in choice about how to convey information to the viewer is through the limitations imposed by the film medium itself. Although it is rather obvious that a filmmaker cannot exceed the limits of the medium, this medium-imposed limit on how information can be communicated to the viewer is what relates movie perception to human perception. Like the film medium, the human brain sets a limit on how information can be conveyed to us. What Carroll and Seeley (2013) postulate is that when these two sets of restrictions are considered in tandem, "the pragmatics of image structure in pictorial representations" is seen to "mimic the pragmatics of ordinary perception" (Carroll & Seeley, 2013). In other words, the limitations that the film medium presupposes, parallel the limitations that human cognition presupposes. The basis of this claim stems from the role that selectivity plays in human perception.

Although often unaware of this, we are creatures that only selectively perceive our environment. In other words, we see what we want to see and what we need to see. Selectivity is key to humans' ability to navigate effectively and efficiently through the environment. This is because there is an infinite amount of available information within our dynamic environment just calling

for us to perceptually process it and a finite amount that we are able to process—in other words, there is a capacity to the brain’s processing power. And because of this, we need to be selective. We need to make decisions about what is and is not of important within our environment and allocate our cognitive resources accordingly. Because of this, our capacity to filter out irrelevant stimuli from relevant stimuli is paramount to our ability to cope with our environment—our ability to selectively attend to stimuli endows us with the ability to locate, identify and respond appropriately to environmental stimuli relative to our goals.

This type of selectivity is also seen within movies’ pictorial representations. Here, an image structure need only provide information sufficient enough for viewers to “recognize the intended content of a scene and to mold their affective, perceptual, and semantic engagement with the characters and events depicted” (Carroll & Seeley, 2013). This is accomplished through stereotyped formal devices and editing techniques in which movies’ recognitional prompts are sequenced and ordered to convey a sufficient amount of information to the viewer. It is the pragmatic mimesis of movies’ pictorial representations and the pictorial representations of ordinary experience from which we can begin to understand the role object recognition plays within both ordinary perception and movie perception.

The *diagnostic recognition framework* for object recognition is grounded in the concept of diagnostic features. Diagnostic features are “sets of sensory features sufficient to enable an organism to perceptually recognize the identity, shape, location and affordances of objects and events in the environment” (Carroll & Seeley, 2013). What determines whether a feature is diagnostic or not is the task at hand, or the target task. And for object recognition to occur in this manner successfully requires the ability to flexibly attend to the information necessary for the

task at hand. The ability to flexibly orient is often due to the perceptual salience of some features within the environment—that is, perceptually salient features exogenously capture our attention. But since situations arise in which the diagnostic features necessary for the task at hand are not the most perceptually salient features in the environment, selective attention plays a key role within the diagnostic recognition framework.

The ability to escape from the exogenous control of perceptually salient features is explained by the *biased competition models* for selective attention. What this means is that, when a conflict arises between perceptually salient information and task-relevant diagnostic salient information, top-down processes help us sort out which information to attend to—that is, top-down processes bias our perception. This, in turn, allows us to dynamically alter and tune sensory systems in a task-relevant manner. These top-down biases prime lower-order sensory processing systems to expect—to in some sense, be on the look out—for the perceptually salient features diagnostic of the target task. Empirically, this has been shown to occur through the effect of top-down feedback on the bottom-up perceptual systems that encode perceptual salience; for example, prefrontal cortex feedback to premotor and parietal areas' has been shown to prepare populations of neurons to encode for potential targets through increasing their firing rates, while simultaneously inhibiting distractions (Kaster, 2004). Interestingly, and of importance to the present discussion, feedback from areas such as the amygdala, orbitofrontal cortex, and ventromedial prefrontal cortex have been shown to bias perception towards emotionally salient features within the environment (Pessoa & Adolphs, 2010).

Emerging as a means to accomplish goals, the success of these top-down perceptual biases in helping us accomplish target tasks is evident from stereotyped behaviors. Seeing as we live in a

relatively stable environment, stereotyped behaviors allow attention to be directed at a task's diagnostic features only when needed. The ability to 'attend when needed' drastically reduces the cognitive load of ordinary perception—a cognitive load would otherwise be too energetically expensive and ineffective to maintain. Thus by offloading the processing of unnecessary perceptual information, attention can be directed on the basis of diagnostic salience rather than perceptual salience.

Selectivity of a similar kind is also seen in movies. From choices about the image structure of recognitional prompts to choices about camera cuts and sequencing, movies are produced through a tedious, selective process. This process begins with movies' most basic unit, shots. Shots, and the sequences they are arranged into, are dynamic pictorial representations. In and of themselves shots are perceptually discontinuous, both spatially and temporally. However, from these perceptually stagnant spatial and temporal moments emerge narratologically coherent sequences constructed to transcend shots' temporal and spatial limitations—ultimately endowing the narrative with its own distinct spatiotemporal features. Accomplished through the use of formal film conventions and editing techniques that have evolved over time, filmmakers are able to intentionally manipulate perceptually deplete pictorial representations and thus transform them into perceptually meaningful artifacts.

Selectivity is thus the bridge between movie perception and ordinary perception from which we must attempt to understand the relationship between ordinary perception and movie perception. Because of this, the editing techniques and formal conventions through which movies are constructed from require careful examination.

### *1.5 Movies: From the Discontinuous to the Continuous*

Through an extensive editing process, innumerable camera shots are transformed into a movie. It is through this editing process that spatiotemporally discontinuous shots are constructed into a continuous movie narrative. Thus, a focus on the film editing techniques used to accomplish this is therefore paramount to the present discussion.

A shot is characterized as a single, static camera recording that contains no perceptible temporal or spatial discontinuities (Carroll & Seeley, 2013; Smith, 2006; Smith, Levin, Cutting, 2012). When shot A has exhausted its narrative intention, a ‘cut’ is made and the current shot is replaced by the next shot, shot B for example (Smith, 2005). This is all film editing really is: “[the] coordination of one shot with the next” (Bordwell & Thompson, 2001). And while its definition is rather simple, the actual editing process is not.

On average contemporary movies contain around 1,000 to 2,000 edits. This means that edits occur every 2.7-5.4 seconds on average (Smith, 2006). As a movie viewer, this information is counterintuitive to our intuitions about movies; in fact, the thought of new shots replacing old shots every 2.7-5.4 seconds is almost unfathomable. This is because it doesn’t *seem* like this many edits occur during a movie—we don’t *feel* like movies spend their entire duration jumping between spatiotemporally discontinuous shots. And the reason for this is due to the success of a film editing technique called continuity editing. An editing technique that developed with the purpose of making the discontinuous appear continuous.

Continuity editing (CE), also known as Hollywood style editing, is considered the ‘standard’ editing technique in the film industry. The utilization of this technique by any filmmaker merely requires adhering to a set of rules known as continuity editing rules (CE rules). These rules provide filmmakers with formal conventions on how to stage, film and edit scenes in such a way

that “the viewer can comprehend the scene with minimum effort” (Smith, 2006); the effect of which makes edits “invisible” and the narrative appear continuous (Smith, 2006; Smith, 2012). CE rules are best thought of as a set of heuristic guidelines on how to sparse together discontinuous scenes in order to create the effect of continuity.

Just to provide an idea of what type of guidelines these are, let us quickly look at the CE guidelines for how to spatially establish a scene. As part of CE rules, the 180° System is a set of camera placement and editing guidelines on how to spatially establish, and then analytically break down, scenes in order to effectively tell a story (Smith, 2006). The goal of the filmmaker here is to establish a “clearly identifiable line in the 3D space of the scene down which all action, character movement, glances, and dialogued occur” called an “axis-of-action” (Smith, 2006). To establish a scene’s axis-of-action, the first shot of a scene is filmed perpendicular to the axis-of-action and shows the scene’s location and relative position of characters within it (Smith, 2006)—this is called an *establishing shot* (Smith, 2006; Smith, 2012). The analytical break-down of the scene after the axis-of-action is established merely involves maintaining the spatial coherency of the scene; which is accomplished by ensuring that all subsequent shots are filmed on the same side of the axis as the establishing shot (Smith, 2006).

To enhance the articulate a scene’s content—characters, action and objects of interest—a different technique known as variable framing is used. This involves “enhancing the perceptual salience of elements within depicted scenes” through the indexing, bracketing, and scaling of diagnostic information and is accomplished through the use of discontinuous camera movement, continuous camera movement and lens movement (Carroll & Seeley, 2013). Indexing, bracketing and scaling are summarized as follows:

- Indexing is a means of pointing out information with the camera
- Bracketing is the framing and reframing of objects, characters and events through the narrowing or widening a scene's field of view. It is the natural effect of camera movement
- Scaling is the change in relative size, orientation and context of the elements within a scene. It is the natural effect of camera movement (Carroll & Seeley, 2013)

The effect of indexing, bracketing and scaling is that it provides viewers with the sufficient information needed to comprehend scenes within the context of the narrative and thus help the viewer in the identification and spatiotemporally orientation of scenes within narrative sequences. Through the use of variable framing, and CE rules as a whole, movies' discontinuous shots are transformed into a perceptibly continuous narrative. Interestingly, this is also what occurs during everyday perception: discontinuous visual information is transformed into continuous visual representations.

#### 1.6 *Natural Vision: From the Discontinuous to the Continuous*

Everyday perception also involves stitching together discontinuous pieces of visual information into continuous representations. Just like how movies are limited by their film medium, we are limited by our physiologically capacities.

Natural vision allows us to see about  $120^\circ$  of a scene, over 90% of which is in very low resolution due to the retina's distribution of photosensitive cells (Smith, 2011). In other words, natural vision is limited in i) how much visual information it is privy to at any one time and ii) the acuity of the visual information it has access to. As a result, we cannot encode an entire visual field simultaneously. Instead we must move our eyes—through saccadic eye movements—to specific details within our visual field (Smith, Levin, Cutting, 2012; Smith, 2011). Saccadic eye movements, as they are termed, consist of saccades—eye movements—and

fixations—stationary moments—and allow for a visual target to be projected onto the retina’s most sensitive part, the fovea (Smith, Levin, Cutting, 2012; Smith, 2011).

And what are we saccading too? Research on saccadic eye movements has shown that the details we attend to within our visual field are, of course, not arbitrary. This is where our previous story about object recognition and selective attention come in; our visual fixations are tuned to our actions to that we can successfully cope with our perceptually replete environment. In fact, this coupling of perception and behavior is so intense that that less than 5% of fixations focus on task-irrelevant objects (Carroll & Seeley, 2013). Thus, the visual information privy to neural processing is not only extremely small compared to all the available information within our field of vision, but it is also the necessary portion; the perceptual information we need to accomplish target-tasks.

What this means is that we’re continually—around 2-3 times a second—making saccadic eye movements in order to intentionally hone in diagnostic information within the environment (Smith, 2013). And when we do this, there is a 20-50ms time frame in which no visual information is inputted into our perceptual processing systems—this is called saccadic suppression (Smith, 2013). But it sure doesn’t seem like that; just like it doesn’t seem like cuts are made so often while watching a movie. Rather, we tend to think of our visual stream as continuous. Just like we tend to think of movies as a continuous narrative rather than a bunch of rapidly changing segments. So, what? How does this relate to movies?

### 1.7 *So, What?*

There are clearly some remarkable parallels in the structure of visual information within a movie and the structure of visual information during natural visual routines. For one thing, the manner



in which specific details within a visual scene are selected and honed in on within everyday visual routines mirrors the way in which the camera selects and hones in on specific visual details within a scene (Smith, Levin & Cutting 2012). The camera, just like our eyes, tends to the narratologically important—the diagnostically salient—visual features of a scene such as faces, areas of high motion, etc. And what about the continual camera cuts? Reminiscent of saccadic eye movements, right? While, it turns out that CE rules us specific attention cues to yoke the timing of these cuts to the timing of our saccades such as conversational turns, sudden onsets of motion, shifts in character gaze, etc (Smith, Levin & Cutting 2012). The idea here—the idea that we’re about to adopt and then run with as far as we can—is that movies are artifacts of human visual perception. Here is a good example of what I mean. In real-life, we might hear a loud sound to our left, our eyes saccade on over to check it out (who knows what is over there!), there is a 20-50ms period in which no visual information is inputted, our eyes fixate on the diagnostically salient visual stimulus (in this case, a forklift; since we’re currently touring a micro-brewery in Denver and the tour guide keeps mentioning that we need to watch out for forklifts; so due to top-down feedback, our sensory areas have primed to look for visual stimuli that look like forklifts). In the movie house, we might hear a loud sound to our left, our eyes saccade on over to check it out (who knows what is over there!), there is a 20-50ms period in which no visual information is inputted and during which the camera has also cut to the left to ‘see’ what’s over there, and then our eyes fixate on the diagnostically salient visual stimulus (which is again a forklift; since the present sequence of shots has been intentionally constructed to let us know that we’re in a micro-brewery in Denver and the camera keeps zooming-in on the tour-guide every time he mentions to watch out for a forklift and it is getting really annoying; so,

due to top-down feedback, our sensory areas have primed to look for visual stimuli that look like forklifts).

Thus, the construction of movies mimics the construction of our visual consciousness! Of course, this seems like a rather poignant and unfounded claim—and that’s because it is! While, at least, we most definitely can’t make this claim from the previous discussion. All we can conclude from the previous discussion is that, on the surface, there appears to be some marked similarities between the way filmmakers structure visual information within a movie and the way the human brain structures visual information within the context of everyday. In Section 2 of this paper, however, we are going to make the aforementioned conclusion; but before I divulge how we’re going to do this, let us quickly look back at the last section to see where we came from and how we got here.

### 1.8 *Where We Are, Where We Came From*

In regards to our discussion as a whole, we’re currently at the point in which it is beginning to become clear that there is some sort of weird relationship between the perceptual processing of movies and the perceptual processing of everyday. Yet, we’re also at the point where it is difficult to see how we’re supposed to make sense of these similarities. Let’s use the last section’s micro-brewery example to show how we can make sense of all this (see Section 1.7). Here, it is rather obvious that this qualitative example was intentionally phrased in order to highlight the similarities between our real-life experience and movie experience at a micro-brewery—so let’s work through it in a more honest way.

First off, if we’re in a micro-brewery in real-life, our current experience involves more than just the sights and sounds of the micro-brewery (who knows, maybe we’re drinking a beer and

smelling all those great micro-brewery smells: hops, yeast, hippies, etc.). Secondly, our eyes don't zoom-in on the tour guide's face every time he annoyingly mentions the forklift. And thirdly, when we hear the sound to our left, we don't just move our eyes; we move our body as well (the orienting response). These three aspects of the real-life micro-brewery experience return to the distinguishing features between a real-life experience and a movie experience mentioned in Section 1.3. These features are summarized in below in Table 1.

Table 1. Basic comparison of real-life micro-brewery experience and movie micro-brewery experience.

Real-Life Experience	Movie Experience
Visual, auditory, olfactory, gustatory, tactile information	Visual, auditory information → perceptually sparse
Cannot defy limitations of space or time (i.e., no close-ups of tour guide's face)	Can defy limitations of space and time (i.e., close-up of tour guide's face) → utilize P.O.V. that defies spatiotemporal limitations of the real world
Egocentric perspective (i.e., orienting response)	Non-egocentric perspective (i.e., no orienting response) → contain no orienting information

OK, now let's work through what the sections after Section 1.3 can tell us about how, despite these differences, there are aspects of these two brewery experiences that are exactly the same in terms of being visual experiences. First off, while we might be standing in the dynamic perceptually replete environment during our real-life experience, that doesn't mean we perceive the totality of this environment. Rather, the perceptual capacity of our brain limits the perceptual replete-ness of this experience—so, let's get rid of movies' perceptual sparseness as a distinguishing feature of movie perception. As a result of this perceptual capacity, we are selectively attending to specific details of brewery scene at large that are important either because of their exogenous perceptual salience (motion, light, unique colors, etc.) or because of their diagnostic perceptual salience (the forklift!). And as we do this, there is a continual and constant (remember, around 2-3 times a second!) interruption in the input of visual information

to perceptual processing areas within our brain—so, let's get rid of movies' spatiotemporal discontinuity as a distinguishing feature of movie perception too.

And in consideration of the mimesis between the editing process and ordinary perceptual processing models, the claim of this section is that movies are artifacts of human visual perception. Movies are created by humans, and intended for humans. Therefore, the development of editing techniques that mimic ordinary perceptual processing routines shouldn't be a surprise to anyone. These techniques allow filmmakers to recreate a real-life visual experience; and through the manipulation of the brain's natural perceptual processes, filmmakers are able to capture the eyes and minds of audiences. If this claim is true—if movies are artifacts of visual perception—then a viewers' visual perception of movies is the same as their visual perception of the environment.

But then, how can we explain the discrepancy between how viewers respond behaviorally in the context of the real world and in the context of the movie house? One possible way to rationalize this is to consider the two aspects of movie perception that we have not been able to 'explain away' in terms of natural perceptual processing: the P.O.V. that defies the spatiotemporal limitations of the real world, and the viewer's non-egocentric perspective that results as a the lack of orienting information. Since both of these features relate specifically to perspective, it is likely that they the 'unrealistic' elements within movies; the features of movies that allow viewers to dissociate their perspective from the perspective of the movie. In this way, a reasonable hypothesis is that these features are responsible for the uncoupling of how a viewer behaviorally responds to movie events and how a viewer behaviorally responds to real-life events. What is reasonable about such a claim is that it is ecologically logical; there is no

ecological reason for the human brain to have developed a means to processes a dissociated external perspective. And while some reader may think this last statement is an antithesis of the previous preceding discussion, then they've missed the point. The viewer clearly does not adopt the camera's perspective. Rather, the camera functions as a set of, in some way, objective human eyes through which the filmmaker uses to unpack the narrative in an ecologically coherent manner; the entailment of which is that the unfolding of narratologically diagnostic visual information the way in which the natural perceptual processes select and attend to salience within the environment. In this, the use of visual cues that naturalistically prompt shifts in visual attention towards salience, such as eye gaze, allow camera shifts and eye gaze shifts to coincide in such a way that the viewer's encoding of perceptual information is no difference during a movie than it is during everyday.

This section therefore claims that movies are artifacts of visual perception. In this way, movies naturalistically grip viewers' visual attention due to a mimesis in the presentation of visual information. As such, the perceptual processing of movies is identical to the perceptual processing of everyday. And while we do not experience movies in the exact same way as everyday, this does not mean that there are not components of the movie viewing experience that are experienced in the same way as real-life. Thus, given the mimesis in visual perception, the theoretical postulate of this paper is that: movies, as artifacts of visual perception, are naturalistically processed by the human brain; as a result of movies' naturalistic perceptual processing, there is no difference between the psychological experience of movies and the psychologically experience of everyday.

### 1.9 *Where We Are Going*

If there is no difference between the psychological experience of movie viewing and the psychological experience of everyday visual perception, we must prove this. But, how? The most viable way is to conduct a full-fledged empirical examination of movie perception. If the CE guidelines used by Hollywood movies do indeed evoke naturalistic perceptual processes, we should find empirical that movies utilizing these techniques are perceptually processed similarly across viewers. The reason for this is because natural perceptual processes within the human brain evolved similarly in order to cope with the environment. Although it is difficult to predict the extent to which subjective differences will influence the perceptual processing of movies, we should at least be able to see a difference in the perceptual processing of movies utilizing CE guidelines and movies that do not—if CE guidelines mimic ordinary perceptual processing models, there should be more across-viewer similarity in the perceptual processing of movies utilizes such editing techniques. Such empirical inquiries into movie perception will not only allow us to determine the validity of the claims made within this section.

## CHAPTER TWO:

### THE COMMONALITY OF VIEWER PERCEPTION, A LOOK AT EYES, BRAINS, AND EVENT PERCEPTION

---

#### *2.0 Introduction: An Outro to Theory and an Intro to Evidence*

This chapter seeks to address the general question, how do we perceive movies? If we are to investigate this question from a scientific standpoint, the most obvious starting point is an investigation of effects; in this regard, the question is best considered in terms of movies' effects on viewers. Based on the cognitivist grounded claims about the relationship between movie perception and ordinary perception, we expect to see a high degree of across-viewer similarity in

the perceptual processing of Hollywood movies compared to alternatively edited (and unedited) dynamic visual scenes.

Through the following examination of empirical data, we will see that Hollywood movies do indeed affect viewers by-and-large very similarly. Evidence for this stems from studies that have found a high level of coordination across viewers' perceptual processing of movie stimuli that utilize CE techniques compared to less structured (edited) dynamic visual stimuli; although viewers are unaware, there is a high level of synchrony across-viewers' eye-movements, brain activity, and perception of events during the viewing of Hollywood movie stimuli (Cutting, Brunick & Candan, 2012; Hasson et al., 2008; Smith, 2006; Smith et al., 2011; Smith et al., 2012; Smith, 2013). To determine the exact relationship between CE techniques and such across-viewer synchronization, following discussion examines and delineates the features responsible this synchronization.

### 2.1 *Watching Eyes Watch Movies*

First things first: if the CE rules utilized by Hollywood movies do indeed function to focus and direct viewers towards narratologically diagnostic visual features within scenes—through the use of techniques such as variable framing, as discussed in Section 1.5—viewers should all be looking at the same places at the same times. And, this is exactly the case. Through the use of eye tracking technology, studies have shown that there is a high level of coordination—both spatially and temporally—in the fixation of viewers' gaze while watching movies (Hasson et al., 2008; Smith & Henderson, 2008). This clustering of viewers' gaze at one location during a movie—often referred to as attentional synchrony—is considered to be an integral feature of movies that allows for their global coherency among audiences (Hasson et al., 2008; Mital et al.,

2011; Smith, 2006; Smith et al., 2011; Smith et al., 2012; Smith & Mital, 2013; Smith, 2013).

But, what visual features are responsible for this eye gaze synchrony? This is exactly the question Parag Mital and his colleagues (2011) set out to answer.

In 2011, Mital and his colleagues examined the visual features responsible for the attentional synchrony effects observed among movie viewers. Since visual attention is controlled both involuntarily and voluntarily, Mital et al., (2011) postulated that it is the involuntary control of viewers' visual attention that accounts for the synchronization of eye gaze across-viewers. This involuntary, or exogenous, control describes the capture of attention by low-level visual information such as color, edges, motion, and luminescence (Mital et al., 2011; Smith, 2012; Smith, 2013); and out of these features, motion and the appearance of novel objects are the only two visual features that have been invariantly shown to exogenously capture attention, regardless of the viewers' task (Mital et al., 2011; Smith, 2013). In contrast, the voluntary, or endogenous, control of visual attention results from the internal allocation of attention to cognitively relevant stimuli; cognitively relevant factors include subjective preferences, desires, viewing task and recognition of object and scene semantics (Mital et al., 2011; Smith, 2012; Smith, 2013). The purpose of Mital et al.'s study (2011) was to determine which visual features contribute most to the clustering of viewers' gaze during dynamic scene viewing.

Mital et al. (2011) identified motion and the appearance of novel objects as the visual features that contribute most to viewer gaze clustering during dynamic scene viewing. Of these two features, however, motion was found to be the most predictive of viewers' eye-gaze—motion predicted viewers' eye gaze 68-71% of the time Mital et al.'s (2011). This finding supports the view that eye gaze is under exogenous control and suggests that movies' visual features—motion



and the appearance of objects—captures viewers’ attention involuntarily. However, even Mital et al. (2011) acknowledges that such a conclusion cannot be made since it is also likely that top-down (endogenous) processes may direct viewers’ gaze.

To examine whether or not eye gaze is influenced by endogenous factors, Mital et al. (2011) conducted a post-hoc survey of the visual areas where viewers’ gaze clustered. The data showed that frames’ “cluster areas” also contained semantically rich visual information: objects such as eyes, faces, hands, vehicles and text. As Tim Smith puts it, “motion may simply coincide with the features we are actually interested in...motion may predict where we look, but we may look there because we are interested in people and their actions, not the motion itself” (Smith, 2013).

To determine whether or not an endogenous factor such as task affects viewers’ eye gaze, Taya et al. (2012) asked participants to watch a clip of a tennis match and determine which player won the point. As a control, other participants were asked to free-view the clip without a task. The data demonstrated that the viewers’ task had no affect on their eye movement behavior. No significant difference between the groups’ saccadic amplitudes, inter-saccadic intervals or gaze coherence was found. This experiment thus supports the idea that eye gaze is under exogenous, rather than endogenous, control.

However, a year later Smith and Mital (2013) produced data that refuted Taya et al.’s (2012) results. Smith and Mital (2013) suggested that Taya et al.’s (2012) results are not indicative of how exogenous and endogenous factors influence visual attention in our natural environment; this is because Taya et al. (2012) asked participants’ to complete a task in which visual sensory and semantic salience overlapped. In this task, the movement of the tennis ball revealed information relevant to the viewing task (figuring out who won the point). The influence of

sensory salience and semantic salience therefore cannot be distinguished within Taya et al.'s (2012) experiment. To dissociate sensory salience and semantic salience, Smith and Mital (2013) assigned participants a viewing task that required them to attend to clips' static visual features while ignoring dynamic visual features. Referred to as 'spot-the-location,' the viewing task Smith and Mital (2013) used asked viewers to identify the location of clips. Compared to those asked to free-view the scenes, there was less synchronization among viewers' eye gaze during the 'spot-the-location' condition, as indicated by a higher number of gaze clusters. This was due to the fact that the 'spot-the-location' required viewers to visit more peripheral areas of the scenes.

The picture we get from these eye gaze studies suggests that although sensory salience—and motion in particular—is predictive of eye gaze, semantic salience also plays an influential in directing viewers' eye gaze. From these studies alone, the role of sensory salience and semantic salience in directing viewers' eye gaze in a synchronized fashion is rather ambiguous as it appears that while sensory salience is capable of directing eye gaze, it is nonetheless influenced in a top-down manner by task (an endogenous factor).

Speaking specifically in terms of highly structured movie stimuli, such as Hollywood movies, Hasson et al. (2008) have shown that the synchronization of eye gaze is due to the structure of these movies. Comparing viewers' eye gaze while watching a segment of *The Good the Bad and the Ugly* to viewers' eye gaze while watching a fixed-shot video taken in Washington Square park, Hasson et al. (2008) showed that there is invariance among viewers' eye gaze, as well as brain activity, while watching highly structured film compared to watching unstructured film. Since the filmmaker must therefore direct viewers' eye gaze in some way through the structuring

of the medium, this provides direct evidence that movies' structure affects their perceptual processing—an idea that we will come back to later within the paper.

Although these studies suggest that the synchronization of viewers' eye gaze is necessary in order to obtain brain activity synchronization, it is not sufficient in producing brain activity synchronization. This was shown from a study that found viewers' eye gaze to be highly correlated while watching a silent film in both the forward and backward direction; but found that a high amount of variance in both viewers' comprehension of the narrative as well as brain activity (Hasson, Yang et al., 2008; for discussion of results refer to Section 2.2).

Let us now examine viewer synchronization in terms of the synchronization of viewers' neural activity.

## 2.2 *Watching Brains Watch Movies*

The synchrony of movie viewers' eye movement and brain activity was also investigated in an fMRI study conducted by Hasson et al. (2004). Using fMRI to measure participants' brain activity, Hasson et al. (2004) analyzed their data using a novel method referred to as the inter-subject correlation (ISC) analysis. This method allows researchers to identify similarities in viewers' brain activity over time by measuring and comparing the time course of across viewer voxel-activity (Hasson et al., 2004; Hasson et al., 2008). Additionally, Hasson et al. (2004) conducted a reverse-correlation analysis; this analysis works by looking at which stimuli were present during moments of certain brain regions' activity (Hasson et al., 2004). In this experiment, fMRI data was collected from five participants while free-viewing a 30-minute clip from *The Good, the Bad and the Ugly*.

Inter-subject correlation (ISC) analysis of data revealed a high degree of similarity across viewers' neocortex activity. Across viewers, Hasson et al. (2004) found that approximately 45% of neocortex activity followed a similar time course during movie viewing and was highly correlated ( $p < 0.001$ ). The correlated regions included most auditory and visual cortices—including early retinotopic areas, the occipitotemporal cortex and the intraparietal cortex—and extended to the superior temporal sulcus (STS), the lateral sulcus (LT), retrosplenial gyrus, secondary somatosensory regions in the postcentral sulcus, inferior frontal gyrus and limbic system regions in the cingulate gyrus (Hasson et al., 2004). Hasson et al., (2004) additionally identified two types of correlation among viewers: i) a spatially nonselective activation wave and ii) specific activation of functionally distinct brain areas (Hasson et al., (2004).

In the reverse-correlation analysis of fMRI data, Hasson et al. (2004) dissociated the stimulus events that evoked the nonselective correlation and the functionally selective correlation; here, the 'nonselective' activity wave refers to spatially nonselective, inter-area cortical responses, whereas the 'selective' activity wave denotes spatially selective responses that occurred in functionally distinct cortical areas. To find out what evoked such responses, researchers looked back at the on-screen stimuli events that occurred during moments of correlated activity across viewers.

Through an examination of the on-screen stimuli that evoked the nonselective activity wave, researchers found that emotional or surprising stimuli (loud noises, narrative twists) evoked the highest peak activations among viewers. Researchers then subtracted this nonselective activity from the dataset to examine the stimuli responsible for evoking the functionally specific correlated brain activity across viewers. It is noteworthy that even when the nonselective activity

wave was subtracted from the correlation data, the level of across viewer correlation remained relatively high ( $24\% \pm 8.5$ ). Researchers identified that this functionally specific correlated activity occurred in the fusiform face area (FFA), the parahippocampal place area (PPA) and the middle postcentral sulcus (PCS). Movie shots containing close-ups of faces evoked activity in the FFA; shots containing images of indoor/outdoor scenes, buildings and open fields evoked PPA activity; and shots in which delicate hand movements were performed evoked PCS activity. Unlike the latter two regions of interest (ROIs), which only further replicated past findings, the link between delicate hand movements and activity in the middle postcentral sulcus was a novel finding. Hasson et al. (2004) speculated that PCP activation—a region that is not part of the brain's well-known sensory areas—reflects the visuo-somato-motor system found in macaque monkeys. The visuo-somato-motor system is thought to link behavior and perception and is also known as the “mirror” system.

To test the affect of movies' temporal dynamics on the synchronization of viewers' brain activity, Hasson, Yang et al. (2008) compared viewers' ISC while watching a clip from *The Adventurer* and *City Lights*—two silent films—in either the forward direction or backwards direction. A ‘piecewise scrambled’ condition was created by varying the length of shot sequences and used to probe the effect of the movie narratives' temporal coherency on viewers' brain activity; in this condition, the length of shot sequences within the clip were varied: short ( $4 \pm 1$  s), intermediate ( $12 \pm 3$  s), and long ( $36 \pm 4$  s).

Hasson, Yang et al.'s (2008) results showed that the activation of different brain regions during movie viewing is either driven by movies' sensory input itself or movies' content (temporal coherency). Regions in which activity was driven by movies' sensory information were visual

processing areas such as MT+, V1, lateral occipital (LO), and the PPA. Among these areas, the primary visual cortex and the motion-sensitive MT+ displayed the highest reliability in responses regardless of the movie clips' temporal structure (forward/backward). However, the reliability of responses within higher order cortical areas known to be involved in perceptual processing—specifically the precuneus, posterior lateral sulcus (LS; or Wernicke's area), frontal eye field (FEF), superior temporal sulcus (STS) posterior superior temporal sulcus (STS), intraparietal sulcus (IPS), posterior lateral sulcus (LS; or Wernicke's area), and temporal parietal junction (TPJ)—depended on the movie clips' temporal structure; activity in these regions was only reliable during the forward viewing of the clips.

The piecewise scrambled condition was used to ensure that the reliability of responses in these higher order areas depends on the clips' temporal structure. Corroborating the finding that the reliability of early sensory areas did not depend on the clips' temporal structure, researchers found that regardless of the shot sequences' length, activity in early visual processing areas (MT+, V1, LO, PPA) was highly reliable. For high order areas such as the STS, precuneus, FEF, LS, and TPJ, the reliability of responses for sequences with short and intermediate timescales. The time-dependent responsiveness of these higher order areas makes sense since these areas are known to be involved in the integration of multimodal sensory information and cognitive processes which likely requires an accumulation of a certain amount of sensory information (see review Hasson et al., 2008). These results indicate that these regions are dependent on movies' temporal coherency, and thus movie content, rather than movies' sensory information.

To further investigate the influence of movies' narrative structure on the ISC of viewers, Hasson et al. (2008) compared viewers' ISC while watching four different types of movies (here, 'types'

refers to the level of aesthetic control, or structure). A comparison between ISCs obtained from a TV episode directed by Alfred Hitchcock, a movie directed by Sergio Leone, a TV episode directed by Larry David, a fixed-shot video of Washington Square Park (unstructured) revealed that movies' narrative structure significantly affects the degree to which viewers' brain activity is synchronized: Alfred Hitchcock evoked 65% of across viewer cortex synchronization, Sergio Leone evoked 45%, Larry David 18% and the fixed-shot of Washing Square park less than 5%. Such results speak to level of audiences' collective engagement with the movie stimuli and speak to the impact of movies' structure on audiences' collective engagement. And since each of these clips varying in their structure—truly, how edited they are—this indicates that movie structure not only affects movies' perceptual processing in terms of viewers' eye gaze, but also in terms of viewers' brain activity. While the high level of brain activity synchronization evoked by Hitchcock reflects his ability to globally manipulate and direct viewers' minds through the visual structuring of the clip, Hasson et al. (2008) interestingly found high levels of ISC across viewers' auditory cortex; suggesting that Hitchcock's choice of music was also successful in collectively engages audiences. The low ISC evoked by the Washington Square Park clips corroborates the finding of more variance in the cluster of viewers' eye gaze while watching this clip as well; without a high level of structure (editing), viewers' are left to direct their own eye gaze which for obvious reasons—since viewers are processing different visual information within the scene—results in variance across viewers' brain activity (Hasson et al., 2008).

In summary, the studies reviewed in this section tell us that something aside from perceptual salience drives brain activity synchronization across viewers. Thus, from this discussion emerges the first two integral pieces to our puzzle:

1. visual attention is necessary, but not sufficient in explaining the synchronization of viewers' brain activity
2. there is a division in the cortex between cortical regions that respond stereotypically to sensory information and higher order cortical regions that respond stereotypically to narrative content; additionally, and most obviously, there is a division between cortical regions that respond stereotypically and those that do not (since no level of structure has shown to elicit 100% brain activity synchronization)

With this in mind, it is clear that most logical next step is to understand the way in which narrative structure elicits stereotypical responses within higher order cortical areas.

### *2.3 The Events We All Perceive*

As we've seen, visual attention is necessary but not sufficient to account for across viewer synchronization. This leaves us with a question: what else explains across viewer synchronization? To answer this we must begin by thinking about what besides visual attention is necessary for everyday perceptual processing. One of the most obvious answers is temporal structure. Although it is difficult to conceptualize how movies' temporal cues are perceptually opaque, we must follow empirical data rather than introspective intuitions to address the dissociative effects of visual cues and temporal cues on viewer synchronization. And since movies defy the spatiotemporal limitations of the real-world, this seems all the more reason that event segmentation plays a vital role in movie perception; because of this, it appears that viewers' must be able to segment movie narratives into discrete temporal structures in order to understand how movies' multiple and various temporal parts coherently relate to the narrative as a whole. More specifically, we now turn to event segmentation not because it must necessarily occur, but because we know it occurs during movie viewing. Before we discuss event segmentation in regard to movie viewing, however, let us first turn to event segmentation within the context of everyday.



In the context of ordinary life, temporal structure is essential to the process of human understanding. To most people, this probably seems rather strange—‘like really, how did we just jump from the importance of VISION to *temporal structure*!?’ But this reaction is only rational. Like most perceptual processes, we are hopelessly unaware of the things we desperately need. Temporal structure is one of those things.

The environment is replete of information available to human perception—in fact, we’re bombarded by it. And while this information is dynamic, continuous and complex, this is not how we understand it to be. Instead, the continuous information is segmented into temporal parts. These temporal parts are not only meaningful, but they are also correlated with ecologically relevant features within the environment (Zacks, Braver, Sheridan et al. 2001). In turn, knowledge about event structure allows us to flawlessly navigate through our everyday activities. This is evident from the fact that when someone is making a peanut butter sandwich it is by no means mysterious that they spread the peanut butter on a piece of bread with a knife. This type of information is vital in how we understand everyday activities. Knowledge about event structure is fundamental to effectively cope with the vast amount of information continuously surrounding us. It is this type of knowledge that allows us to fill in, relate and make predictions about everyday activities.

Because temporal structure plays such a fundamental role within our ability to understand everyday events, it is rather unsurprising that observers ubiquitously segment ongoing events into temporal parts that are reliable and meaningful indicators of the action within the event. And while the ubiquitous segmentation of events into temporal parts was first shown through behavioral studies, in 2001 Jeffrey Zacks—one of the most prominent event segmentation

researchers—and his colleagues investigated whether this segmentation of activity is a normal aspect of neural information processing.

To do this, Zacks, Braver, Sheridan et al. (2001) collected fMRI data as participants watched video clips of everyday activities; these video clips consisted of one actor performing four everyday activities: ‘making a bed’ (316s), ‘doing the dishes’ (258s), ‘fertilizing a houseplant’ (120s), and ‘ironing a shirt’ (298s). It should be noted, that all video clips were filmed from a fixed location and contained no cinematic effects (cuts, zooms, pans) in order to avoid the influence of external event segmentation. The researchers hypothesized that if transient changes in neural activity correlate with event boundaries, then event segmentation is an ongoing component of perceptual processing within the brain. And what they found was just that. As participants watched the clips, researchers observed a network of brain regions displayed transient changes in neural activity at event boundaries.

Through a post-hoc event-related analysis of fMRI data, Zacks, Braver, Sheridan et al. (2001) were able to specifically identify a network of cortical areas that exhibited time-locked transient changes in activity correlated with perceptual segmentation. These areas included areas of the bilateral posterior cortex and the right frontal cortex. In the posterior cortex, the most highly correlated region was identified as bilateral Brodmann’s areas 19/37 in the occipital/temporal junction. Peak activity in Brodmann’s areas 19/37 was found to correspond with the human MT (or V5) complex; an area known to be responsible for the processing and understanding of biological motion and action. In the right frontal cortex, the precentral sulcus (BA 6) was the region highly correlated with perceptual segmentation. This area is known to be responsible for eye movements and active shifts in attention and is also thought to be a human analog for

monkey's frontal eye field (FEF). The supposed relationship between the precentral sulcus and the frontal eye field is interesting to mention here because research on monkeys' FEF has allowed for the subdivision of this area into two functionally distinct anatomical areas. In Zacks, Braver, Sheridan et al.'s (2001) study, the observed activity in the precentral sulcus corresponds to the anterior subregion of the monkey FEF; this region is known to be involved in the active shift of attention, whereas the posterior subregion is associated with eye movements.

Statistical analyses revealed that the time-course of cortical activity differed for coarse-grain and fine-grain segmentation (coarse-grain refers to a large temporal unit while fine-grain refers to a small temporal unit). Zacks, Braver, Sheridan et al. (2001) observed that some transient changes in activity occurred before an event boundary. Analyzing the most focal regions of activity (previously discussed), researchers found that pre-boundary neural activity was task and time-dependent. In the active segmentation trial—when participants pushed a button in response to coarse or fine-grained segmentation during clip viewing—pre-boundary activity increased in bilateral posterior regions for coarse-grain segmentation (left,  $F_{5,75} = 5.3$ ,  $p < 0.001$ ; right,  $F_{5,75} = 12.2$ ,  $p < 0.001$ ) as well as fine-grain segmentation (left,  $F_{5,75} = 3.94$ ,  $p = 0.003$ ; right,  $F_{5,75} = 4.98$ ,  $p = 0.001$ ). Although both posterior regions displayed an increase in pre-boundary activity in the passive segmentation trial as well—when participants free-viewed the clips—this activity was only found to be reliable for coarse-grain segmentation in the right posterior region (left,  $F_{5,75} = 2.00$ ,  $p = 0.088$ ; right,  $F_{5,75} = 4.63$ ,  $p = 0.001$ ). In comparison to posterior regions, the frontal region only displayed reliable pre-boundary activity increases during active, coarse-grain segmentation ( $F_{5,75} = 2.729$ ,  $p = 0.028$ ).

Additionally, behavioral data analyses provided further support for the existence of a hierarchical structure within event segmentation. This was shown by first calculating the i) mean distance between fine-grained event boundaries and all points within the video clip and then the ii) mean distance between coarse-grained event boundaries and all points within the video clip. A significant correlation between coarse-grain boundaries and fine-grain boundaries was found ( $t_{53} = 9.04, p < 0.001$ ) and thus indicated that fine and coarse-grain segmentation boundaries do not function independently of one another.

In conjunction, Zacks, Braver, Sheridan et al.'s (2001) results strongly suggest that the human brain contains functionally distinct networks involved in the perceptual processing of temporal information. Furthermore, these neural networks appear to respond to temporal structures with some level of specificity. While posterior brain regions' displayed reliable pre-boundary activity increases at least to some extent in both active and passive segmentation trials, pre-boundary activity increases in the precentral sulcus were more selective and only associated with active segmentation of coarse-grain temporal segmentation. The difference in selectivity between posterior and frontal region activity during ongoing perceptual processing suggests that top-down and bottom-up influences play a dissociative role in the perception of temporal structure. This, in light of Zacks, Braver, Sheridan et al.'s (2001) behavioral data, supports the view that there is some type of perceptually specific functionality between top-down and bottom-up processing networks. But, does this actually relate to movie viewing?

Aside from the fact that it is difficult to understand how viewer comprehension could occur without successfully segmentation movie narratives' temporal structure, research demonstrates that event segmentation is a rather ubiquitous aspect of the movie viewing experience. Evidence

of this stems from findings of around 87% agreement among movie viewers whether a shot within an edited movie stimuli is part of the same event as the preceding shot or not (Cutting, Brunick & Candan, 2012). What is most important to our present discussion, however, is that the dissociative role of top-down and bottom-up processes found in the aforementioned studies have also been found in studies not only specifically probing edited movie stimuli, but specifically probing movie stimuli that utilize CE; and what is even more outstanding is that in this study, researchers found that the largest neural responses evoked by continuity edits that marked event boundaries were in visual processing areas (Magliano & Zacks, 2011). This robust increase in activity within visual processing areas, although counterintuitive, actually directly supports the success of continuity editing in directing viewers' visual attention. With this in mind, let's look at how event segmentation provides us with a model to understand the similarity in viewers' top-down and bottom-up perceptual processing of movies. To do this, let us look at Jeffrey Zacks and his colleagues' (2007) theory on the role of top-down and bottom-up perceptual processes within ongoing event segmentation.

#### *2.4 Event Segmentation Theory (EST): The Framework*

In 2007, Zacks, Speer, Swallow, Braver & Reynolds (2007) put forth a theory called *Event Segmentation Theory* (EST). EST is a culmination of event structure research and aims to computationally and neurophysiologically account for event structure perception. It should be noted, since several definitions of events exist, that this theory operates under a broad and categorical definition of events. Thus the events EST references are "events that involve goal-directed human activity and are of modest duration (seconds to tens of minutes)" (Zacks, Speer, Swallow, Braver & Reynolds (2007).

The three key features of events are their hierarchical structure, recurrence and cyclicity. Here, hierarchical structure refers more to the processing of information; in this way, the perceptual networks involved in event segmentation are thought to be hierarchical in that sensory information inputs is transformed and used to generate *perceptual predictions*, or representations of future states of the world. Perceptual predictions and thus events' hierarchical structure is vital to the anticipation and planning of appropriate, future actions—without them, humans would react rather than respond to incoming environmental stimuli. The recurrent nature of events refers to the modulation of earlier processing stages by later processing stages. And lastly, the cyclical nature of events references the ability to monitor incoming sensory information (what is happening now) with perceptual predications through continuously comparing the actual and with the expected.

Zacks, Speer, Swallow, Braver & Reynolds (2007) postulate generally describes event segmentation processing as the transformation of sensory information input (by perceptual processing) into semantically replete representations. An important aspect of this information processing system is that the perceptual processing of incoming sensory information is bias by *event models*, or representations of 'what is happening now.' Often equated to working memory representations, event models integrate multimodality sensory information to form a relatively stable model of the current event. Aside from sensory information, event models are influenced by *event schemata*, or representations of events previously encountered that exhibit shared features with the current event; unlike the inherent transient nature of event models, event schemata exist permanently within the system. Accordingly, the system utilizes bottom-up and top-down processes. In this way, bottom-up processing involve sensory information input and top-down processing involve event schemata.

To ensure that the system's perceptual predictions do not diverge from the system's event models, the system contains an error detection module that continuously compares the system's perceptual predictions to its event model. If perceptual predictions are no longer accurate—or in other words, activity in the environment is now less predictable than before—the transient error predication signal will eventually cause a gating mechanism to open which forces the system to update its current event model. For an event model to be updated, the system must reset the event model and increase the influence of sensory information on event model. A postulate of particular importance to our current discussion is that the content of event models is resolved by both top-down and bottom-up influences; here, event schemata are responsible for the top-down influences while sensory and perceptual representations of the current state are responsible for the bottom-up influences on the construction of event models.

What exercises executive control in this event segmentation network, anything? The answer is nothing, according to Zacks, Speer, Swallow, Braver & Reynolds (2007). In this theory, event segmentation is not a subcomponent controlled by some cognitive module within the system. Rather, cognition is subservient to event segmentation. The postulate here is that event segmentation i) controls the allocation of cognitive resources and ii) updates working memory. The former involves the transient prediction error signal, which the theory equates with attention. Attuned to the system's processing resource needs, the error prediction signal allows the system to dynamically adapt to the system's needs based off its current event models. In this way, event segmentation contains its own mechanism for attention and functions without the help of external attentional resources.

The system described above is the computational model of EST. Based on empirical evidence, Zacks, Speer, Swallow, Braver & Reynolds (2007) hypothesize that this computational model is representative of how event segmentation is implemented neurally within the human brain. The next section examines the neural correlates upon which this computational model is based.

### 2.5 Neural Correlates of EST's Computational Model

The architecture of EST's computational model represents a culmination of empirical studies on the perceptual processing of event segmentation in the human brain. This section presents empirical evidence that supports the computational model of EST put forth by Zacks, Speer, Swallow, Braver & Reynolds (2007). To aid in conceptualize, a schematic drawing for the neural implementation of EST has been provided below:

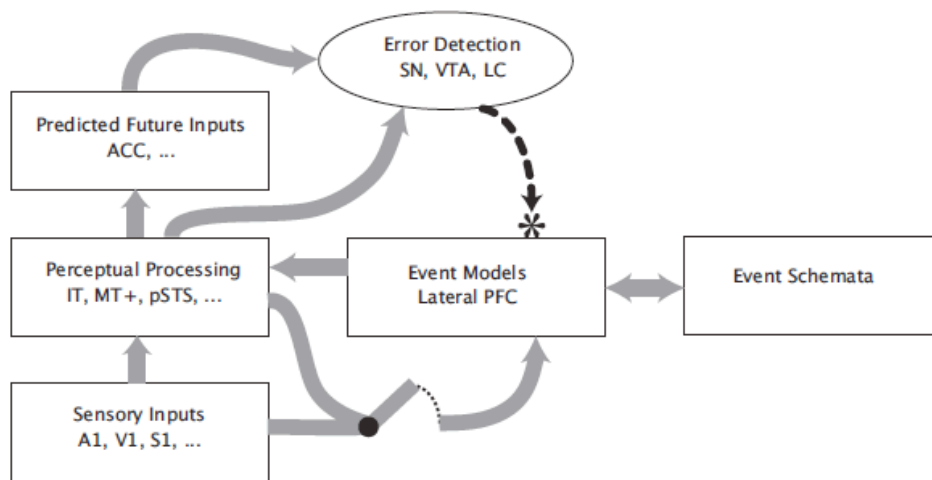


Figure 1. Schematic depiction of the model, with hypotheses about the neurophysiological structures corresponding to the different components of the model. Thin gray arrows indicate the flow of information between processing areas, which are proposed to be due to long-range excitatory projections. Dashed lines indicate projections that lead to the resetting of event models. PFC = prefrontal cortex; IT = inferotemporal cortex; MT+ = human MT complex; pSTS = posterior superior temporal sulcus, ACC = anterior cingulate cortex, SN = substantia nigra, VTA = ventral tegmental area, LC = locus ceruleus; A1 = primary auditory cortex; S1 = primary somatosensory cortex; V1 = primary visual cortex (Zacks, Speer, Swallow, Braver & Reynolds 2007).



### *Sensory Inputs*

The cortical processing of sensory information begins first in the primary sensory areas such as, the primary visual cortex (V1), primary auditory cortex (A1) and primary somatosensory cortex (S1). Perception research often focuses on the perceptual processing of visual information since V1 and its projections downstream to extrastriate cortical areas are relatively well-known.

In V1, low-level sensory representations are inputted and processed in a feedforward manner. Information processing in this area involves the extraction of complex information from inputted sensory representations, such as orientation, direction and color. The capability of V1 to organize sensory representations in this way is a result of its own organization. V1 is composed of functionally distinct, topologically mapped areas with a columnar organization (Mountcastle, 1997). A consequence of V1's hierarchical organization of these functional maps is that the each subsequent area's response properties are more complex and selective than the preceding area. That is, as early sensory representations are feedforward through V1 to extrastriate cortical areas, their organization becomes increasingly complex; these representations continue to increase in complexity after V1 as well.

### *Perceptual Processing*

Subsequent visual processing is segregated into two functionally distinct perceptual processing pathways that specialize processing certain types of information. In general, the occipitotemporal pathway (or 'ventral stream') specializes in the identification of objects, while the occipitoparietal pathway (or 'dorsal stream'), specializes in the processing visual spatial and motion information (Felleman & Van Essen, 1991; Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1978). Although considered separate processing pathways, findings indicate that

communication does occur between the two processing pathways (Felleman & Van Essen, 1991; Ungerleider & Mishkin, 1978).

Similar to V1, areas within the ventral and dorsal streams are selectively activated by particular visual features of the distal visual environment. Until recently, evidence of cortical selectivity stemmed from studies that used static or simplistic stimuli—such as static pictures or black and white moving dots. Although suggestive, these studies provided a more modular perspective of cortical selectivity—a perspective detached from everyday perceptual processing. It was not until the use of movie stimuli, however, that these pathways were investigated within a more dynamic, realistic context (Bartels & Zeki, 2004; Hasson, Nir, Levy, Fuhrmann & Malach, 2004). It is these latter studies that provide the most reputable support for EST’s postulate that “perceptual processing of ongoing activity into discrete events is a spontaneous concomitant of ongoing perception” (Zacks, Speer, Swallow, Braver, & Reynolds, 2007).

Neurophysiological evidence that cortical selectivity occurs during the ongoing perceptual processing of activity is provided by two studies. In one of these studies, Bartels & Zeki (2004) had participants free-view a 22-min segment of James Bond’s *Tomorrow Never Dies* during fMRI recording. Data analysis on fMRI data showed that participants’ BOLD signals for feature-selective cortical regions were correlated with the presentation of those features within the movie. Additionally, the brain regions exhibiting cortical feature-selectivity corroborated with previously identified cortically selective areas. Face selective cortical activity, for instance, was observed in the lateral fusiform gyrus (FG) and superior temporal (ST) gyrus and sulcus (ST activity extended bilaterally along the middle temporal gyrus and was more prominent in the right hemisphere). Whereas, activity in the extrastriate body area (EBA) and anterior lateral FG

was shown to be selective for human bodies. Similarly, Hasson, Nir, Levy, Fuhrmann & Malach's (2004) fMRI study that was discussed previously also identified feature-selective cortical regions; these regions were also identified through a reverse-correlation analysis of fMRI data. Corroborating findings that the FG is face-selective, analyses also discovered that activity in the parahippocampal region—which was previously implicated as building and space selective—was correlated with spatially descriptive movie shots (establishing shots or wide-angle shots). Together, these two studies support the view that cortical selectivity occurs during ongoing perceptual processing.

#### *Changes in the environment and event segmentation*

According to EST, event segmentation relies on changes in the environment. The importance of change within EST is evident when we remember the role change plays within EST's computational model. Here, the function of the prediction error mechanism relies on changes in the environment to transiently increase the prediction error signal, eventually leading to the detection of an event boundary and resetting of the event model. The following subsection will delineate empirical evidence that supports EST's view on the relationship between change and event segmentation.

The strongest empirical support for the relationship between change and event segmentation comes from studies utilizing fMRI to track the neural correlates of event segmentations. These studies show that specific brain areas exhibit selective, time-locked transient responses to event boundaries during ongoing perceptual processing (Speer, Swallow & Zacks, 2003; Zacks, Braver, Sheridan & Donaldson, 2001; Zacks, Swallow, Vettel & McAvoy, 2006). Across the board, transient increases in activity were identified in the MT+ complex at event boundaries

within the movie stimuli; in Zacks, Braver, Sheridan & Donaldson (2001), as previously discussed, BAs 19/37 were identified as the most highly correlated areas involved in observed MT+ activity. In addition to regions within the posterior cortex, the front eye field (FEF) area displayed transient increases in activity correlated with event boundaries (Speer, Swallow & Zacks, 2003; Zacks, Braver, Sheridan & Donaldson 2001). Furthermore, the posterior superior temporal sulcus (pSTS) displayed selective activation during event boundaries (Zacks, Swallow, Vettel & McAvoy, 2006). Taken all together, this data suggests a relationship between changes in the environment and the segmentation of events during ongoing perceptual processing.

#### *A Neural Mechanism for Error Detection*

EST hypothesizes that the error detection mechanism is neurally represented in the anterior cingulate cortex (ACC), the substantia nigra, ventral tegmental area, and locus ceruleus. Here, the ACC is thought to be the region in which future prediction inputs are represented. Whereas the three-midbrain nuclei represent the brain region in which the error detection function is implemented.

By and large, research on the neural implementation of error detection stems from Eugene Sokolov's theory on the *orienting response*. The orienting response is characterized as the mechanism by which current neuronal models of external stimuli are reset when they fail to match the current external stimuli. Accordingly, a mismatch between the external stimulus and current neuronal model triggers the orienting response; and this response is thought to reset the neuronal model and transiently increase the influence of sensory input (Sokolov, 1990).

Empirical evidence implicates that the ACC, substantia nigra, ventral tegmental area, and the

locus ceruleus are the brain regions involved in the processing and implementation of the orienting response.

The resetting of this neuronal model appears to be implemented by the three midbrain regions previously mentioned—the substanti nigra, the ventral tegmental area, and the locus ceruleus. These three regions function as neuromodulatory systems within the midbrain and are known to send error signals throughout the cortex; the widespread nature of these error signals is a result of the systems' divergent and extensive projections throughout the cortex (Schultz & Dickinson, 2000). A notable quality of these three neuromodulatory systems is that they have been found to compute different types of error signals. While the dopaminergic neurons of the substanti nigra and ventral tegmental areas have been shown to selectively respond differences in actual and predicted reward, norepinephrine neurons in the locus ceruleus have been shown to respond to errors regarding external stimuli (Schultz & Dickinson 2000; Usher, Cohen, Servan-Schreiber, Rajkowski & Aston-Jones. 1999).

The ACC, on the other hand, is known to be involved in the regulation of behavioral responses to error detection rather than the detection of error itself (Botvinick, Cohen & Carter 2004; Gehring & Knight 2000). The role of the ACC as a modulator of behavioral responses to error is supported by studies in which ACC activation was found to be strongest during response override tasks—or in other words, tasks that required participants to overcome irrelevant interferences (see review Botvinick, Braver, Barch, Carter & Cohen 1999). The way in which this was determined was through the use of measuring a special event-related potential (ERP) called the event-related negativity potential (ERN); this type of potential occurs during moments

of error in cognitive tasks and is used as a measure of error during ongoing cognitive activity (Botvinick, Cohen & Carter 2004; Gehring and Knight 2000).

In addition to the ACC, studies such as Gehring and Knight's (2000) have identified the prefrontal cortex (PFC) as playing a role in the modulation of action during error detection as well. In Gehring and Knight's (2000) study, activity in both the lateral prefrontal cortex and the ACC was associated with action modulation. A possible explanation for Gehring and Knight's (2000) data, one that they support themselves, is that responses to the detection of error are controlled by two pathways and mediated by neuromodulatory systems in the basal nuclei (Cohen, Botvinick & Carter, 2000). This view hypothesizes that both control pathways are mediated by neuromodulatory systems in the basal nuclei and respond to error detection; with extensive projections to both the PFC and ACC, neuromodulatory systems' are hypothesized as having a role in implementing the necessary corrective action through its connections with the motor system (Cohen, Botvinick & Carter, 2000; Gehring & Knight, 2000). What is different about the two supposed control pathways is that one is more selective and is subject to further mediation by the PFC (Cohen, Botvinick & Carter, 2000; Gehring & Knight, 2000).

Furthermore, an ERP study by Sitnikova, Kuperberg, & Holcomb (2003) supported the view that the neural mechanisms for error detection are utilized during event segmentation. In this study, participants were asked to watch video clips of everyday activities, such as someone brushing their teeth. Half of the participants watched 'incongruent' clips in which at the end of the clip the actor performed the task using an inappropriate object; the other half watched 'congruent' clips in which the actor used an object conventionally used to perform the task. ERP data collected during this viewing task demonstrated that the presentation of incongruent objects elicited a

larger ERN potential than congruent objects. Given the similarity between this ERN potential and the *N200*—an ERP associated with the processing of subsiding predictability (Kutas & Hillyard, 1980)—Sitnikova, Kuperberg, & Holcomb (2003) hypothesized that this response reflected the mismatch of an object’s semantic features. In a study a year later, it was shown that the earliest neural correlate for ERN responses to semantic incongruently occur in the ACC, and after which spreads to prefrontal areas (Frishkoff, Tucker, Davey & Scherg, 2004).

When taken all together, the empirical evidence presented within this section supports the view that the brain employs a neural error detection mechanism to detect inconsistencies between its current state and incoming sensory information.

#### *The neural representation of events*

Event models and event schemas, according to EST, are implemented within the anterior lateral PFC. Here, the PFC is thought of in terms of managerial knowledge, or information representation; rather than in terms of action planning, as it is typically discussed (Sirigu et al. 1996). Accordingly, the information in the PFC is considered to represent “goal hierarchies, temporal order of events, causal links between actions, rules, etc.” (Sirigu et al., 1996). This representational perspective is supported by data that implicates the importance of the PFC in the representation of events.

Often story scripts are used to study the relationship between event representation and the PFC since narrative structure provides an accessible means to delineate subject’s comprehension of sequence and temporal order. These studies have shown that patients with PFC damage are impaired in temporally ordering script narratives and recalling components of the narrative (Sirigu et al., 1996; Zalla, Phipps, Grafman, 2002). In Zalla, Phipps & Grafman’s (2002) study,

patients with PFC damage as showed impairments in understanding how events within the narrative might be interrelated. This lack of narrative comprehension suggests that the PFC plays a role in both processing structural and inferential narrative aspects; a view that supports EST's proposed role of the PFC in event representation during ongoing perceptual processing.

Additionally, this finding was corroborated by a study utilizing movie stimuli (S. J. Hanson, Negishi, & Hanson, 2001). In this study, subjects watched a videotape of actors performing everyday activities and were asked to press a button every time they detected an event boundary. During the viewing, subjects' brain activity was measured concurrently using EEG and fMRI. ERP data showed positive active areas 20 seconds before an event boundary evolve into ERN potentials within prefrontal areas and spread to temporal and visual areas. fMRI data recorded during this activity spread indicated that the activity within the regions evoking ERPs displayed time-locked transient increases in activity; the regions that displayed the most significant activity were the dorsal lateral prefrontal cortex and the anterior cingulate cortex.

All in all, this data provides support for EST's view that the PFC is involved in the representation of events.

## *2.6 Event Segmentation Theory: Perspectives and Caveats*

Through the consideration of EST's theoretical and computational model in relation to the neural correlates that potentially implement this framework, EST provides a plausible account for both the function and implementation of event segmentation during ongoing perceptual processing.

In relation to the previous discussion on attentional synchronization among viewers, the viability of Event Segmentation Theory suggests that event segmentation is the driving force behind



visual attention; possibly the necessary missing puzzle driving viewer synchronization. This, in part, is explained by the cognitive control event segmentation is postulated to exercise over the allocation of perceptual resources during ongoing perceptual processing.

Unfortunately, it appears that Event Segmentation Theory neither includes nor explains an integral component of human perception: emotion. Instead, EST provides an explicitly cognitive perspective on the factors that influence and control perceptual processes within the brain. And in consideration of our current understanding of emotion—that is, that the human brain integrates affective and cognitive processes in an important, influential manner during online processing—the truancy of emotion in EST's proposed model of perceptual processing is detrimental to its fidelity. Of course, this is not said to invalidate EST in its entirety; rather this is said to present the next puzzle piece that bears attention.

Through the following investigation of how emotion and cognition are neurally integrated, it will become clear that a theory on human perception without emotion is necessarily fallacious from the get-go. In the next chapter, the *necessity* of emotion within this unfolding puzzle will become apparent. It is important to note, however, that the following chapter strays from this paper's explicit discussion of movies. You might therefore ask, why is it necessary to include the following necessarily emotion story? Quiet simply, it is of utmost importance to illuminate the direction future research on movie perception must necessarily trend towards. While EST of course provides us with an important story about how we might possibly model and test the interaction between top-down and bottom-up perceptual processes, it is one of the only remotely viable contemporary models out there. And as researchers continue to test EST's postulates in regard to movie perception, no current research has raised the following concern about where

EST necessarily falls short in its explanation of ongoing perceptual processing: the affect of affective processing. Through the following chapter, it will become obvious that emotion must be considered within any model of perceptual processing. The following chapter is also important to this paper's discussion in that it points to the way in which cooperation between these two disjointed fields of cognitive neuroscience research will be mutually benefit to their progressions; while movie perception researchers must begin to model and test the influence of affective processes on the perceptual processing of movies, emotion researchers must begin to use movie stimuli as a means to obtain the influence of affective processing with a more realistic, contextually rich environment.

### CHAPTER THREE:

#### A NECCESARILY EMOTIONAL STORY

---

##### *3.0 Introduction*

Over the last decade cognitive neuroscience has progressed its understanding of emotion rather rapidly. That is not to say that the breath of our contemporary understanding of emotion has been widely adopted into most cognitivists' accounts, however—take EST for example, which fully ignores emotion all together. It is instead to say, that there has been some resilience among researchers (whether out of habit or intention) to compartmentalize emotion and cognition as two entirely different and separate neural systems, cognition trumping emotion in most cases.

And the reason for this most obviously extends from the pervasive characterization of emotion as a lower-level, automatic, sub-cognitive process. An exemplary portrayal of emotion as such is seen in Robert Zajonc's work, where emotion is portrayed as "inescapable" and subservient to cognitive, and even perceptual processes: "because affective judgments are inescapable, they

cannot be focused as easily as perceptual and cognitive processes...affective reactions are thus less subject to control by attentive processes” (Zajonc, 1980). Unsurprising is such a subservient view, however, when considered in light of the earliest emotion theories, such as James Papez’s, that identified emotion with subcortical areas such as the hypothalamus and amygdala (Papez, 1937).

Connectivity, cytoarchitecture and neuroimaging advances, however, brought a more complex understanding of the postulated ‘subcortical’ emotion-network, as well as the identification of emotion with higher up, more ‘cognitive’ brain areas. It is now known that frontal areas, such as the PFC, are also involved in emotion and that these areas are not only highly connected to subcortical areas such as the amygdala, but also that regions involved in emotion are some of the most highly connected in the brain (see Pessoa, 2008 for review). The subsequent chapter’s discussion will begin with an examination of the amygdala’s role in both cognition and emotion; after which, the role of the prefrontal cortex will be discussed.

### 3.1 *The Amygdala*

Often, the amygdala is considered as in terms of its role in the affective processing of fear. So prevalent is this characterization, in fact, that the amygdala is sometimes even referred to as the ‘fear centre.’ As a result, the amygdala’s processing is typically portrayed as automatic and independent of top-down influences. In reality, however, fear-processing is not the amygdala’s sole function. Instead, research has linked amygdala function to several other processes, some of which are cognitive. Of concern to our discussion is the link between attention and the amygdala. In contrast to the amygdala’s typical characterization as an affective-processor, attention is considered solely as a cognitive process; and the fact that recent research has shown an

influential interaction between these ‘affective’ and ‘cognitive’ processors is of interest to this discussion as it indicates that the compartmentalization of ‘affective’ and ‘cognitive’ neural systems is erroneous.

It is well known that the primary function of attention is to modulate the processing of sensory information. Given the brain’s finite processing capacity and the infinite amount of perceptual information in the environment, the modulation of sensory processing via attention endows us with the ability to selectively process the important, meaningful aspects of our environment—and thus, efficiently navigate our environment. Evidence for the affect of attention of sensory processing stems from findings of increased neural firing in the visual cortex while a stimuli is attended and increases in behavioral performance.

Interestingly, attended stimuli are not the only stimuli known with a processing advantage. Research shows that emotionally-laden stimuli have a processing advantage as well. In a study conducted by Anderson and Phelps (2001), it was shown that attention is not responsible for emotionally-laden stimuli’s privy to processing—instead, this advantage in processing was shown to be amygdala-dependent. Using an attentional-blink paradigm, researchers showed that emotionally-laden stimuli, compared to neutral stimuli, weakened the perceptual deficit in awareness of the subsequent stimuli; supporting the notion of this weakening in perceptual deficit as amygdala-dependent, the study found that this perceptual deficit in awareness was not weakened in patients with bilateral amygdala damage (Anderson & Phelps, 2001).

The above finding of amygdala-dependent processing affects have lead researchers to investigate how, and at what level, the amygdala affects perceptual processing. For obvious reasons, researchers have turned to an investigation of the amygdala’s affect on visual processing—and

these investigations have proven fruitful. In one study investigating the influence of amygdala activation on visual areas, participants were asked to perform a gender-discrimination task while shown gray-scale images of emotional faces that varied in emotional intensity. Given the nature of this task, the study demonstrated a correlation between amygdala activation and the activation of extrastriate areas that was independent of attention. Using PET neuroimaging, researchers found that the certain brain regions (see Appendix, Table 1) were selectively activated by the presentation of prototypical fearful, happy and neutral faces. Additionally, parametric analyses revealed that the activity of brain regions presented in were sensitive to the emotional intensity of the presented faces (see Appendix, Table 2). Furthermore, the researchers' regression analyses revealed a correlation between amygdala activation and the activation of specific brain regions (see Appendix, Table 3), including visual areas within the extrastriate cortex (Morris, Friston, Buchel, Frith, Young, Calder & Dolan, 1998).

In a latter fMRI study, Vuilleumier, Richardson, Armony, Driver & Dolan (2004) not only corroborated Morris et al.'s (1998) finding, but also revealed that the correlation between amygdala activity and distant cortical areas is in fact due to modulation by the amygdala. By studying the effects of amygdala damage on the activation of distant cortical areas during the processing of emotional faces, Vuilleumier et al.'s (2004) study was the first study to show that the amygdala is specifically responsible for the observed enhanced neural responses.

Specifically, Vuilleumier et al.'s (2004) findings revealed that patients with amygdala damage, relative to control groups, displayed significantly weakened neural responses in distant cortical areas that are normally enhanced during the processing of emotional faces. Areas that displayed significantly weaker activation in amygdala damage patients relative to controls included: the left striate, left superior occipital, right superior temporal sulcus, right rostral cingulate, bilateral

posterior fusiform, bilateral medial fusiform, bilateral anterior fusiform, bilateral superior occipital, bilateral lateral occipital and bilateral lingual areas (For ANOVA  $F$ -tests, all  $P \leq 0.05$ ). This study thus showed that the sensory processing of emotional stimuli is modulated by the amygdala through ipsilateral projections to cortical visual processing areas.

The next study of interest is based on the aforementioned findings that i) the amygdala interacts with attention and that ii) the amygdala affects early visual processing. Following up on these findings, Phelps, Ling and Carrasco (2006) corroborated previous neuroimaging studies' findings with behavioral data. In this study, researchers examined the effect of emotion and attention on contrast sensitivity—specifically, the sole effect of emotion on contrast sensitivity and conjunctive effect of emotion and attention on contrast sensitivity. In the first part of the experiment, the effect of emotion, irrespective of attention, was tested using an orientation discrimination task. Here, researchers found that the presentation of emotional faces (in this case fearful) significantly increased participants' contrast sensitivity at threshold; this showed the presentation of emotional-laden stimuli alone affects the processing of visual information. In the second part of the experiment, participants again performed an orientation discrimination task but this time the researchers used an attentional pre-cue paradigm to present the emotional faces—that is, the target stimuli for the discrimination task was pre-cued by the presentation face-stimuli (neutral or fearful) either in the peripheral of the subsequent target stimuli or distributed randomly within the display. Results indicated that contrast sensitivity was highest for peripheral, compared to distributed, pre-cues, as well as, when a fearful face pre-cued the target stimuli. Additionally, the effect of emotion on contrast sensitivity was found to be higher for peripheral pre-cues than distributed. These findings thus indicate that although both attention

and emotion increase contrast sensitivity, the conjunctive effects of emotion and attention increase contrast sensitivity the most.

It is therefore evident that the modulation of visual processing is influenced by both emotion and attention—with emotion potentiating attentional affects. What is unclear, however, is the manner in which emotion and attention interact to influence perceptual processing. To address this question, Pessoa, Padmala and Morland (2005) investigated the modulation of amygdala responses by attention and cognition during the processing of emotion-laden stimuli. In this experiment, participants were asked to complete varying discrimination tasks that involved different attentional demands. To examine the effect of attentional load on amygdala responses, Pessoa, Padmala and Morland (2005) asked participants to sustain central fixation while either neutral or fearful faces were concurrently presented with the target stimuli. In one of these ‘attentional’ conditions, the task required participants to discriminate the sex of the face (male or female) after presentation. In the other ‘attentional’ condition, the task required participants to indicate if the orientation of the two peripheral (left/right) bars that were also always presented along with the face. For the latter condition, the difficulty—or in other words, the attentional load required to complete the task—was varied (easy, medium, hard) through the manipulation of the angular difference between the two bars. A comparison between these two attentional conditions was used as a means to measure the affects of attentional load on evoked responses to unattended emotional faces. To study the effect of cognition on emotionally evoked responses, researchers used a bar-only discrimination condition—no faces were presented during this condition and task difficulty was equivalent to the “hard” attentional bar-discrimination task. Through a comparison of participants’ amygdala responses elicited by these conditions, researchers found that both attention and cognition modulate amygdala activity. The results

revealed that increases in attentional load decrease amygdala responses; evident of this was the observed reduction of amygdala responses during the ‘attentional’ bar-discrimination tasks compared to the sex-discrimination task. While the presentation of neutral or fearful faces elicited differential amygdala responses in the “easy” bar-discrimination task, these differential responses were not observed for the more difficult (medium, hard) bar-discrimination tasks. Additionally, a reduction in amygdala activity was observed during the bar-only (cognitive-task) relative to the sex-discrimination task; this indicates that amygdala activity is modulated also by cognition and provides support for the idea that emotion and cognition interaction in a reciprocal and mutually suppressive manner (Drevets & Raichle, 1998).

### *3.2 The Amygdala and its Not So Necessarily Emotional Story*

From the aforementioned studies in section 2.1, we get a story that suggests that enhanced processing of emotion-laden stimuli is necessarily modulated by the amygdala. Although a commonly accepted characterization of the amygdala’s involvement in the modulation of emotional stimuli and its interactions with attention, recent research challenges this conception of the amygdala’s functional role in affective processing. Actually, ‘challenge’ is an understatement since recent research has revealed that the amygdala is not in fact *necessary* for the enhanced processing of emotion-laden stimuli.

In direct opposition to Anderson and Phelps’ (2001) attentional-blink study, a study by Bach, Talmi, Hurlmann, Patin and Dolin (2010) found that patients with selective bilateral amygdala damage displayed an attenuation in attentional blink comparable to the healthy control group after the presentation of emotion-laden stimuli—in this case, aversive words. And this study is not alone in its finding that enhanced processing of emotion-laden stimuli is not modulated, or at



least not *necessarily* modulated, by the amygdala. For instant, in another lesion study, a patient with bilateral amygdala damage showed no impairment in the rapid identification of fearful or threatening stimuli during a visual-search task (Tsuchiya, Moradi, Felsen, Yamazaki & Adolphs, 2009). Similarly, a processing advantage of emotional, relative to neural, stimuli was observed in patients with unilateral amygdala damage during a visual-search task (Piech et al., 2010); the enhanced processing of emotion-laden stimuli in patients with amygdala damage was comparable to controls and was not affected by the lateralization of amygdala damage (Piech et al., 2010).

Furthermore, a later study conducted by Piech et al. (2011) showed that the amygdala is not necessary for the emotional modulation of attention. Using an emotional blink paradigm (EBA)—a paradigm in which emotion-laden stimuli disrupt task-relevant target detection through the disruption of attention—this study observed no comparable difference between the attentional disruption, and subsequent processing deficits, caused by the presentation of emotion-laden stimuli in patients with unilateral (left or right) amygdala damage and controls (Piech et al., 2011).

The truth of the matter is that as of right now, this is the picture of the amygdala that we must live with—at least for now. What we must take away from this story, however, is that the amygdala is most definitely involved in the processing of affective information in some way and that it shares a relationship with cognition, as seen from its interaction with attention (see Hsu & Pessoa, 2007; Pessoa, Padmala & Morland, 2005). In consideration of the interaction between emotion and cognition and the current amygdala impasse, it is clear that a further investigation of

this cognition-emotion interaction is in order. But, where do we start? What bridges these two processes?

The main neural substrate contender for the cognition-emotion interaction is the PFC. Although the next section will discuss a few reasons for this, it is of importance here to mention the reason that bridges the current amygdala discussion with the subsequent PFC discussion: the PFC and amygdala interact to modulate the processing of affective visual information. As a bridge from the amygdala to the PFC, let's quickly look at a study conducted by Lim, Padmala and Pessoa (2009). In this fMRI study, participants were first aversively conditioned to affectively salient visual scenes ( $CS^+$ ). Participants then participated in an attentional blink task with both  $CS^+$  and  $CS^-$  trials. As expected, Lim and colleagues found that participants detected previously conditioned affectively significant stimuli ( $CS^+$ ) better detected than affectively neutral stimuli ( $CS^-$ ). Through an analysis of fMRI data, Lim and colleagues found that the amygdala mediated the processing of affectively significant stimuli through both a direct connection with the visual cortex and an indirect fronto-parietal mediated connected with the visual cortex. The regions identified within the fronto-parietal circuitry were the right inferior parietal lobule (IPL), right superior frontal gyrus, left anterior insula and bilateral middle frontal gyrus (MFG).

It is this cognitive-emotional interaction that we will now turn our discussion to—for it is clearly the next piece in the puzzle.

### 3.3 *The Prefrontal Cortex*

The prefrontal cortex, like the amygdala, is considered one of the basic neural systems involved in affective processing. Anatomically, the PFC is divided into six regions: dorsolateral (DLPFC), ventrolateral (VLPFC), frontopolar (FP), orbitofrontal (OFC), ventromedial (VMPFC) and

dorsomedial (DMPFC) (Ray & Zald, 2010). In general, the PFC is known to have a central role in the control of attention, the maintenance and manipulation of information, and regulation of information flow (see review Pessoa 2008). The most important regions of the PFC involved in emotion, broadly speaking, are the OFC and medial prefrontal cortex (including the anterior cortex, specifically its anterior and subgenual sites); additionally, anterior regions of the insula are often discussed in the context of PFC affective processing due to their relation with the OFC (see review Pessoa, 2008; Pessoa, 2013).

Of vital importance to understanding the PFC's role in affective processing are the following inputs into the PFC from the amygdala and outputs from the PFC to the amygdala. In terms of inputs, the OFC and medial PFC receive a substantial amount of direct amygdalar inputs; whereas, amygdalar inputs into the DLPFC and FP are thought to be indirect and are known to have a much weaker influence on the PFC. It is, also, known that the ACC, OFC, VLPFC, DLPFC and DMPFC have strong reciprocal connections with the amygdala (Banks, Eddy, Angstadt, Nathan & Phan, 2007).

In terms of affective perceptual processing of visual information, the anterior OFC is considered to have an important role in the evaluation of sensory information. The anterior OFC contains two functionally distinct, but related, circuits: i) the medial OFC and ii) the lateral OFC.

Accordingly, the medial OFC is thought to receive low-spatial frequency, also referred to as coarse-grained, visual information through the posterior parietal cortex—this allows the medial OFC to receive information about the 'gist' of the visual scene (Bar, 2003; Barrett & Bar, 2009). The medial OFC also shares strong reciprocal connections to the motion areas (specifically the MST and MT+) within the lateral parietal areas and is involved in discrimination spatial

locations and guiding action. Additionally, the medial OFC is known to influence other prefrontal areas, specifically the VMPFC and ACC. The lateral OFC, on the other hand, is involved in the integration of multimodal sensory information and bodily information. The input of bodily information arises from the VMPFC and anterior insula—the conveyed information via the VMPFC and anterior insula arises is supplied by the medial OFC. The visual sensory information input into the lateral OFC is carried through the lateral OFC's strong reciprocal connections to the TE, TEO and temporal pole areas in the inferior temporal areas. The visual information that the lateral OFC receives is highly processed and rich in detail.

It has long been thought the visual information conveyed to the medial OFC is rapid, compared to the highly processed visual information that the lateral OFC receives. This temporal difference has long been attributed to the fact that the medial OFC receives input from magnocellular pathways. Magnocellular pathways are often thought to have a temporal processing advantage due to their direct connections with lateral geniculate nucleus (LGN)—a region of the thalamus that directs early visual information to V1 or the dorsal stream. Research indicates, however, that the LGN also shares direct connections to ventral stream, specifically the parahippocampal gyrus and amygdala. However, more recent research has been shown that the supposed temporal processing difference between visual information inputs into the medial OFC and lateral OFC does not actually exist (Barrett & Bar, 2009).

### *3.4 The PFC and Affective Processing*

In terms of affective processing, the relationship between the PFC and amygdala is functionally important. The reason for this regards the PFC's connectivity to both cortical and subcortical areas. Among the most known pathways implicated in affective visual processing are the PFC's

connections with the amygdala and visual processing areas. In general, the PFC is known to receive both coarse and fine-grained visual information through either direct or relatively direct connections to both primary and extrastriate visual cortices. Specifically, the VLPFC receives visual input from V4, the DLPFC from V2, and BA 46 via connections with BA 8. Additionally, as previously discussed, the OFC shares strong reciprocal connections to both the dorsal and ventral visual processing pathways. Additionally, the amygdala is also known to receive highly processed visual information through its direct connections with the visual cortex (see review Pessoa, 2008). When we begin to consider all of these sensory inputs in terms of PFC-amygdala connectivity, an affective, modulatory story starts to unfold.

Today, this affective processing story is told in terms of PFC-amygdala connectivity and is often referred to as the cognitive-emotional interaction. Of course there is a lot more to this interaction than just the connectivity between the PFC and amygdala, but it is with this relationship that we will begin our discussion—the most logical reason for this is because it is the relationship between these two regions that sparked early research on the interaction between cognition and emotion.

About two decades ago, researchers began to look beyond mechanisms of bottom-up visual processing and probe the influence of top-down cortical control on the processing of visual information. Although early investigations identified the prefrontal cortex as a possible anatomical substrate for top-down control of visual processing in primates (Barbas, 1995), research remained relatively illusive on how the PFC implemented this cortical control. In 2003, however, Moshe Bar published one of the most substantial and influential postulates on how this cortical control is implemented. Bar's suggested paradigm (2003) worked off of findings such as

Barbas' (1995) that identified direct connections in the magnocellular pathway between early visual areas and the PFC (Barbas, 1995; Bar, 2003). Specifically, direct connections between V2 and the DLPFC, as well as, the ventral area V4 and the VLPFC were identified within the magnocellular pathway (Barbas, 1995). From this Bar (2003) suggested that 'gist-level' visual sensory information is rapidly sent to the PFC, analyzed for its basic representational properties, and then sent back to early sensory areas. Although this analysis was framed in terms of non-emotional visual stimuli, Bar, along with colleague Lisa Barrett, later extended this framework to affective visual processes in 2009.

In Barrett and Bar's (2009) proposed top-down control neural circuitry, the specific prefrontal area responsible for receiving input from visual areas is the OFC. Paralleling Bar's (2003) proposal for the processing of non-emotional visual stimuli, the OFC is thought to receive and analyze 'gist-level' visual input; this information is then fed back to visual processing pathways to influence the further processing of sensory information. Of the two distinct, yet functionally related regions of the OFC—previously discussed in the last section—Barrett and Bar (2009) postulate that the medial OFC receives direct input from visual areas associated with the processing of motion and spatial information (MST, MT+), while the lateral OFC integrates this 'gist-level' information with multimodal sensory information. Accordingly, the low-frequency visual information inputted into the medial OFC is thought to provide motion and spatial information about the current context; in turn, the medial OFC is hypothesized to project feedback to visual areas allowing the medial OFC to guide bodily action, spatially locate visual items, and modify bodily states based off of past knowledge. It is postulated that the medial OFC, through its projections to the VMPFC and ACC, 're-instates' the context in which affective visual information of its sort was previously experienced. The medial OFC is also hypothesized

to convey this spatial information to the lateral OFC, where it is subsequently integrated with the lateral OFC's multimodal sensory inputs. The purpose of this integration is to transform 'gist-level' affective information into "an experience-dependent representation" of the current visual field (Barrett & Bar, 2009).

### *3.5 Perceptual Processing in the Dorsal and Ventral Pathways: an Overview*

The neural circuitry, or neural networks, involved in affective visual attention are divided into two distinct pathways, the dorsal frontoparietal and the ventral frontoparietal. For our purposes, these two pathways will be referred to as either the dorsal or ventral pathways (or streams) from now on.

Both of these largely distributed pathways are charged with different, yet related, roles in attentional control. For instance, the dorsal pathway functions to maintain the locus of attention (even in the midst of distractors), and select attended stimuli based on the goals and prior knowledge of the system. The dorsal pathway is associated with the DLPFC and involves dorsal frontal areas near or at the FEF, intraparietal sulcus (IPS), and the superior parietal lobule (SPL) in the parietal lobe (see review Corbetta & Shulman, 2002a; Corbetta & Shulman, 2002b; Kincade, Abrams, Astafiev, Shulman & Corbetta, 2005; Ochsner 2012; Viviani 2013). The ventral pathway, on the other hand, functions to interrupt and reset attention to behaviorally relevant stimuli. The ventral pathway is associated with the VLPFC, inferior frontal gyrus (IFG) and anterior insula and involves the inferior parietal lobule (IPL) and temporo-parietal junction (TPJ) (see review Corbetta & Shulman, 2002a; Corbetta & Shulman, 2002b; Kincade, Abrams, Astafiev, Shulman & Corbetta, 2005; Ochsner, 2012; Viviani, 2013). The question of import

here is, what happens when goal-directed attention competes with stimulus-driven attention? To answer this, let's look a bit closer at each system.

Studies on the dorsal pathway have shown that it is activated when subjects voluntarily attend to a location—the dorsal pathway is thus considered the endogenous attentional control center (Corbetta & Shulman, 2000; see review Corbetta & Shulman, 2002a; Corbetta & Shulman, 2002b). The sustained activation of dorsal pathway areas—FEF, SPL and IPS—during participants' voluntary control of attention is indicative of the dorsal pathways' role in the maintenance of the current attention set (see review Corbetta & Shulman, 2002a; Corbetta & Shulman, 2002b). In a study specifically probing the relationship between cognition and emotion during the viewing of emotion-laden visual stimuli, researchers found that while activity in the medial PFC increased during a cognitive task—on trials where participants rated the emotional intensity of the stimuli—activity in the medial PFC decreased during trials in which the emotion-laden stimuli were passively viewed (Taylor, Phan, Decker & Liberzon, 2003). Importantly, the pattern of activity in emotion-processing areas such as the anterior insula and basolateral nucleus of the amygdala was the inverse of medial PFC activity (Taylor et al. 2003). This further supports the role of the medial PFC in the control of behaviorally relevant attentional sets during affective visual processing.

On the other hand, activation of the ventral pathway occurs during moments in which attention reorients to unattended stimuli (Corbetta & Shulman, 2000; Corbetta & Shulman, 2002b). And if this attentional reset involves reorienting to unattended, behaviorally relevant stimuli, ventral pathway activation will co-occur with dorsal pathway activation for a brief moment during the reset of attention (Corbetta et al., 2008). Exemplary of dorsal/ventral activation patterns during



ongoing processing, activity in the dorsal pathway is higher during the cue period in target detection tasks, whereas activity in the ventral pathway is higher during the target period; and corroborating the ventral pathway's role in the reorientation of attention, ventral pathway activity is higher when an incongruent target is presented during a target detection task (see review Cabeza, Ciaramelli & Moscovitch, 2012). However, when attention is focused on a task, there is a reduction in ventral pathway activity (Corbetta et al., 2008). Given the ventral pathway's function in controlling attentional resets, it is important to continue on some detail about what drives the reorientation of attention, or in other words, how the ventral pathway defines salience.

Noteworthy, is that the ventral pathway appears to define salience as more than just sensory salience. Although this is obvious from studies, such as Taylor et al.'s (2003)—where behaviorally relevant stimuli were selectively attended to over emotionally-laden stimuli—it appears, still, that the ventral pathway computes salience based off of more than just perceptual and goal-based salience and is also influenced cognitively by more than just goal-directed attentional sets. For instance, in a meta-analysis of 120 fMRI semantic processing studies, Binder, Desai, Graves and Conant (2009) found that areas within the ventral pathway—such as, the IFG, SFG, MFG, VMPFC, and interior parietal lobe areas like that angular gyrus (AG) and supramarginal gyrus (SMG)—were activated during written or spoken word comprehension. Aside from the integration of semantic knowledge, the ventral pathway also exercises endogenous attentional control in response to stimuli previously associated with reward (Anderson, Laurent & Yantis, 2011; Harsay, Cohen, Oosterhof, Forstmann, Mars & Ridderinkhof, 2011; Padmala & Pessoa, 2011). For example, a study Padmala and Pessoa (2011) found that cue-responses to previously conditioned reward stimuli reduced MPFC/ACC activity—two areas implicated in interference/ conflict monitoring. Based off of other studies'

findings, it appears that this reduction in MPFC activity is mediated by a dorsal/ventral pathway interaction. Evidence for this stems from findings that previously rewarded cue-stimuli result in i) an increase in the functional connectivity between reward-processing areas, such as the caudate, and the IPS that results in a decrease in MPFC activity activation and that ii) this is mediated via the IFG; evidence for IFG mediation stems from the fact that researchers found a correlation between the increased activation of the IPS and decreased activation of the IFG in rewarded vs. non-rewarded contexts (Padmala & Pessoa, 2011; Harsay et al., 2011). Thus, the determination of what stimuli are ‘behaviorally salient’ not only involves the integration of perceptual, cognitive and emotional salience, respectively, but is contextually determined. This, in light of the dorsal network’s involvement in the determination of attentional sets, suggests that affective visual processing involves an interaction between the dorsal and ventral pathways. The question is, who is in charge?

### *3.6 The Dorsal/Ventral Relationship during Affective Perceptual Processing*

Seeing as context is an important factor in determining how dorsal and ventral pathways interact during affective visual processing, this relationship will be the focus of this section. To set the scene, let’s first look at what happens when emotional and sensory salience compete for processing.

Luo et al. (2007) examined the relative affects of emotional and sensory salience on affective visual processing by independently varying each of these dimensions during the rapid presentation of word stimuli. In this fMRI study, target words were rapidly displayed and were either preceded by a distracter or no distracter at all. Researchers manipulated sensory salience through the presentation of either subliminal or supraliminal non-target word distracters, while

emotional salience was manipulated through the presentation of either emotional or neutral non-target word distracters. What researchers found is that increases in the sensory salience of non-target words caused SPL activity to increase, whereas, increases in the emotional salience of non-target words caused IPL activity increase.

Similar to Luo et al.'s (2007) results, an fMRI study by Mitchell et al. (2008) manipulated the emotional valence of non-target distracters (presented before and after target stimuli). In this study, however, both SPL and IPL activity was enhanced during task performance. Interestingly, Mitchell et al. (2008) found a relationship between IPL activity and emotional valence; while significant increases in IPL activity were observed during the presentation of emotional v. neutral non-target distracters, these increases in IPL activity were greater in response to positive v. negative stimuli. What is more intriguing about Mitchell et al.'s (2008) results is that the opposite pattern of IPL activity to negative v. positive stimuli was observed upon trials in which participants *only* viewed the emotion-laden stimuli—or in other words, when the emotion-laden images were not presented within the 'distracter-target' paradigm. In a study using a colored-word emotional Stroop task, Compton et al. (2003) also found increases in dorsal pathway activity, specifically the DLPFC, in response to both incongruent colored-words and negative emotion-laden stimuli. Compton et al. (2003) also found that during colored-word trials, and not emotion-laden stimuli, resulted in increases in SPL activity, while increases in IPL activity only occurred during the presentation of negative emotion-laden stimuli. These enhanced dorsal pathway responses were attenuated by high vs. low arousal emotional words (Compton et al., 2003).

Other studies have not only corroborated the selective, context dependent increases in ventral vs. dorsal pathway activity to emotion-laden stimuli, but have also found the relationship between the dorsal/ventral pathways to operate in a negatively suppressive manner. In a study by Dolcos and McCarthy (2006), emotion-laden stimuli were presented as distracters while participants performed a delayed working memory (WM) task. Using fMRI analyses, researchers found that the presentation of emotion-laden distracters during the WM delay enhanced activity in ventral stream areas such as the VLPFC. Additionally, compared to the presentation of neutral and scrambled (a no-distraction perceptual control) distracters, researchers found that emotion-laden distracters not only enhanced activity in ventral stream areas, but also simultaneously deactivated dorsal stream areas such as the DLPFC. A later study by Dolcos, Diaz-Granados, Wang and McCarthy (2008) also corroborated this finding. Through the presentation of ‘confusable’ non-emotional distracters (stimuli similar to the target stimuli, in this case faces) as well as, ‘nonconfusable’ emotion-laden distracters (highly arousing, negative valence pictures) during a delayed WM task, researchers showed that DLPFC activity increased in response to non-emotional distracters and decreased in response to emotion-laden distracters. Here, the decrease in DLPFC activity in response to emotion-laden distracters was correlated with an increase in VLPFC activity and thus further points to the dorsal and ventral stream acting in a negatively suppressive manner.

In light of this, it is clear that the differential activation/deactivation of dorsal and ventral pathways is best considered in terms of the current attentional task set. As has been discussed in passing, the influence of previously rewarded stimuli on dorsal/ventral activation is exemplary of the context-specific manner in which these two perceptual processing pathways interact. Since motivation clearly plays a role in the processing of affective visual information, it is thus the

affect of motivation on affective visual processing that we will now turn to. And through the subsequent section's review of motivation, it will become clear that motivation may be more than just a puzzle piece—motivation may in fact be the *key* to this perceptual puzzle.

### 3.7 Motivation to Find the Puzzle Key: The influence of motivation on affective visual processing

This section probably marks the point where everyone is questioning their *own* motivation to continue reading—I mean really, what am I even getting on about at this point? And in an odd and counterintuitive way, I am empathetic to such an opinion. We've quickly moved from movies, to visual attention, to event segmentation and then to a laborious discussion of the amygdala, PFC and their respective associated roles within affective visual processing networks. And now, with almost no transition—aside from the brief mention of one study—we've moved to *motivation*. Where is this going? Is this just a shoddy amalgamation of what cognitive neuroscience can *generally* tell us about *anything*? Is there even a direction, a point, to this discussion? Or to put it more poetically, is there a *way* to this *will*?

These questions are legitimate and in need of answers. And aside from an assessment of the 'shoddy' nature of this discussion, an assessment of these questions is in order. The problem is that, although close, we still need a delineation of one more puzzle piece before the disjointed can be joined. Understandably frustrating, the discussion must continue on a seemingly aloof path one last time.

So here I offer motivation as the final puzzle piece—the key—from which these seemingly tangential micro-discussions can be purposefully united. To delve in, let us discuss Jimura, Locke and Braver's (2010) study.

*The Basics of Motivation, Attention, and Perception*

In a study conducted by Jimura, Locke and Braver (2010), participants performed a WM task in both rewarded and non-rewarded contexts. The allocation of rewards (in this case monetary) depended upon participants' fast and accurate responses. What Jimura, Locke and Braver (2010) found behaviorally is not very surprising: participants' responses were faster and more accurate in the reward-context than in the non-reward context. Although this finding will not be discussed as a focal matter here, it is important to note that the behavioral affect of motivation provides us with a more encompassing story about how cognitive influences both neural processes and behavior. More pertinent to our discussion, though, is what Jimura, Locke and Braver's (2010) event-related fMRI data revealed.

From fMRI analyses, researchers found context-related activation changes in the dorsal PFC (BA 46/9). In the context of reward, analyses revealed sustained activation of the dorsal PFC. This sustained activation occurred early within reward-context trials and then displayed a transient decrease in activation later within reward-context trials (for similar results see Pessoa & Engelmann, 2010). Given that this observed dorsal PFC activity was unaffected by the magnitude of reward within reward-contexts, researchers concluded that this activity was context-dependent. Additionally, researchers found that participants' individual sensitivity to reward was correlated with the degree of sustained activity in the dorsal PFC. Thus, motivation is not only an influential force in determining how visual attention is allocated, but its affect is also related to individuals' subjective characteristics. And from this, it is clear that motivation has something to offer to our story.

Not only does motivation affect known attention neural substrates, it also affects perceptual processing. The affects of motivation on the processing of sensory information have been observed by many studies (Pleger, Blankenburg, Ruff, Driver & Dolan, 2008; Serences, 2008; Weil et al., 2010). Corroborating animal studies (Shuler & Bear, 2006; Pantoja et al., 2007), studies have found that motivation (in the form of reward) affects the processing of sensory information (Pleger et al., 2008; Serences, 2008; Weil et al., 2010). During rewarded trials of a tactile discrimination task, for instance, ventral striatum activity was enhanced (compared to non-rewarded trials); interestingly, the affect of reward on somatosensory signals was observed after the onset of the stimuli, before the administration of reward (Pleger et al., 2008). In studies that involved the processing of visual stimuli, reward was also seen to affect the processing of sensory information; the affect of reward on the processing of visual information was observed through its modulatory effects on visual areas such as V1 (Serences, 2008; Weil et al., 2010).

#### *A Look at How Tomorrow is Motivated by Today*

In many ways, motivation is kind of like a one-night stand that won't stop calling you. Or at least, that is the characterization we get from numerous studies—of course not in a literal sense, but more in the sense of the residual effects of motivation on attention and perceptual processes.

The residual effect of motivation on attention, and thus the processing of sensory information, is evident from studies such as Libera and Chelazzi's (2009) study. Here, Libera and Chelazzi (2009) tested the effects of motivation (for reward) over the course of a few days. What the researchers found is that the allocation of attention to specific objects differed based off of participants' previous encounters with them in the context of reward. Interestingly, the effect of motivation on subsequent testing corresponded to the level of reward associated with stimuli on

previous training trials. For example, while it was harder for participants to ignore distracter stimuli during subsequent testing when those stimuli previously connoted high-reward, it was easier to ignore the distracter stimuli that previously connoted low-reward. Similar results were obtained in a study conducted by Doallo, Patai and Nobre (2013) that showed that previously rewarded visual stimuli were recognized both faster and more accurately during a visual search task 24hrs after the reward training session.

It turns out that residual motivational effects are not only capable of improving subsequent visual processing, but that they are also capable of reducing the effectiveness of subsequent visual processing. Evidence for this stems from a study conducted by Anderson, Laurent and Yantis (2011). This study showed that stimuli previously associated with reward, slowed participants' performance during a subsequent visual search task.

#### *Motivation in Summary*

From this section's review, a more substantiated picture of motivation's affect on perception develops. The effects of motivation on perceptual processing thus appear to be twofold: i) motivation *enhances* executive functions such as attention, thus increasing or decreasing their effectiveness during the processing of sensory information and ii) motivation *reallocates* the resources of executive functions (Pessoa, 2013 p. 181).

With this in mind, it is time that we consider the *whole* picture. An ample starting point comes from Luiz Pessoa's *dual competition model*. This model, which has saturated Pessoa's work over the last decade, culminates in his new book "The Cognitive-Emotional Brain" and provides logical starting point for the subsequent discussion given that it integrates the concepts of motivation, attention, and perception into one, coherent working neural model.



### 3.8 To the Dual Competition Model, and Beyond!

The dual competition model more or less parallels prior models of attention referred to as *biased competition models* (see Desimone, 1998). In truth, what these models really boil down to is that perceptual processes are *biased* in some way. These biases may bottom-up or top-down and are generated by such things as behavioral goals, perceptual saliency or affective saliency.

In terms of Luiz Pessoa's (2013) model, which specifically framed in terms of visual perception, all of the elements that generate biases converge in what he calls, *priority maps*. Similar to the concept of saliency maps, priority maps are generated from the convergence of four 'biasing' components: i) physical salience, ii) endogenous relevance (current goals), iii) affective significance and iv) motivational significance (Pessoa, 2013). And as we have seen, all four of these components influence the perceptual processing of visual information. Furthermore, the dual competition model provides a neural framework for the interaction between these biasing components.

The neural framework Pessoa (2013) provides for these 'cognitive-emotional interactions' involves both "task networks" and "valuation networks." Here, 'task networks' refer to the neural networks involved in attention, while 'valuation networks' refer to the neural networks involved in reward. Neurally, these networks overlap, in that they both involve regions of the amygdala, hypothalamus, anterior insula, OFC and PFC. For quick reference, the categorization of these cognitive-emotional substrates have been included below:

- (i) direct pathways that link task (attention) and valuation networks, such as the OFC and DLPFC
- (ii) neural hub regions where task and valuation networks intersect
  - (a) DMPFC, a region where executive control and affective information intersect

(b) anterior insula, a region involved in the intersection of sensory information, threat, risk, attention, task-switching, long-term memory, working memory, and various cognitive processes

(iii) neuromodulatory regions, which are influenced by the key substrates involved in cognitive-emotional processes and are capable of globally influencing the brain through their diffuse projections

Within the dual competition model, these overlapping neural networks are dubbed *modes of communication* and serve as the neural substrates for cognitive-emotional and cognitive-motivational interactions.<sup>2</sup> In consideration of these regions' known involvement in the processing of perceptually relevant information, as well as their diffuse neural connectivity, this neural framework is not only wholly plausible, but is also the most likely contemporary account of how multidimensional influences are integrated within the brain.

The reason for this is because from the unearthing of these complex and influential cofactors, we obtain a more holistic picture of how human perception functions in the real world—a picture of human perception that delves deeper than mere the mere sensory saliency of visual information in its explanation of how everyday perception functions. Yet, all the while, the involvement of cognitive, emotional and motivational processes appears to stray even further away from an explanation of viewer synchronization; the subjectivity of these perceptual cofactors seems irreconcilable in regard to explaining movies' highly coordinated effect on viewers. However, it is clear that these seemingly antithetical perceptual influences must be reconciled with. And in order to do so, the most logical perspective to adopt is an ecological one. A perspective in which we can frame the influence of affect on visual perception is framed in terms of its objective ecological value.

---

<sup>2</sup> It should be noted that, from here on, I use 'cognitive-emotional' to denote interactions that involve emotion as well as motivation.

When we adopt an ecological framework of affect, it is not difficult to see stereotypical, synchronized affective influences on visual perception as evolutionarily valuable. If a negatively valenced visual stimulus represents danger, it is logical that all human perceivers respond stereotypically to it. And in terms of social interactions, a degree of automaticity within affective responses makes sense. Supporting the notion that affect is another synchronizing force within visual perception, a recent study explored the influence of emotional valence and arousal on across viewer ISC. This study sheds light on this section's previous discussion and helps us tie together the affect of movies' sensory and semantic salience in terms of the manner in which viewers' stereotypically process perceptual and affective environmental information. For this reason, it is with a discussion of this paper that this chapter will conclude.

### *Seeing Affect as a Synchronizing Force*

Examining the influence of affect on the synchronization of viewers' brain activity, Nummenmaa et al. (2012) were able to identify the influence of specific dimensions of emotions on the synchronization of the cortical pathways discussed thus far: visual sensory processing areas, emotional processing areas, and the dorsal and ventral attention networks. Specifically probing the influence of emotional valence and arousal on viewers' ISC, researchers found a correlation between negative valence and increased ISC in emotional processing areas—the thalamus, ventral striatum, insula, and midbrain—as well as the ventral attentional network—specifically the TPJ, precuneus, and VMPFC. In contrast, researchers found that moments of high emotional arousal within the movie stimuli increased ISC in the visual and somatosensory processing areas, as well as, in the dorsal attentional network—specifically in the IPS and FEF. Additionally, regardless of valence, highly arousing emotional stimuli were found to recruit the dorsal attentional network and enhance the time-locking of activity within this attention network.

In terms of ecological value, stereotypical responses to emotion-laden stimuli can be explained in terms of avoidance of danger and approach of reward, as well as, the ability to understand social cues and behave appropriately within social contexts. Nummenmaa et al.'s (2012) findings suggest that the recruitment of the dorsal attentional network appears to be responsible for the orienting of attention to emotion-laden stimuli, whereas the recruitment of the somatosensory processing areas and the ventral attentional network appears to be responsible for the mapping of others emotional states. The final key to our synchronization puzzle thus appears to be viewers' stereotypical affective responses to visual stimuli. Although given the lack of affective visual processing research that has been conducted using movie stimuli, it is difficult to determine the extent of affective processes within movie perception.

### *3.9 Why is Affective Processing Important Within the Context of this Paper?*

What is clear from the following discussion on the influence of affective processing during ongoing perceptual processing is that affect is a vital component of human perception. And as such, it is clear that future research on movie perception must begin to integrate and test the role affective processing plays within the context of movie perception—or better put, the role it plays within the context of everyday perception.

In terms of the Section One's claim—the claim that the movie experience is a psychologically real experience—the preceding discussion on the synchronizing force of affective visual processing pointedly supports the psychological reality of the movie experience. The reason for this is because the empirical evidence presented within this chapter shows that the visual experience of movie viewing extends an entire spectrum of psychological experiences. The fact that the neural processing of movies involves, and is biased by, affective processes indicates that

these stimuli are processed as more than just visual stimuli. Through the following examination of the effect of affective processes on movie stimuli, it is evident that the human brains' processing of these visual stimuli extends beyond their visual properties; the affective processing of movie stimuli shows that the human brain, to some extent at least, processes these stimuli (their narrative, characters, etc.) as no different than it would in the context of everyday.

Clearly, there is something to be said about the psychological reality of the visual movie viewing experience if our affect processes kick-in in response to what is presented to us in the form of 2D illuminated visual images. In and of itself, this not only points to the value of movie stimuli within all disciplines of cognitive research, but also to the viability of this paper's claim: the movie experience is a psychologically real experience.

## CONCLUSION

---

*We have lost sight of the fact that what we produce is what the brain produces—the body is at best a tool for the mind to create.*

What is most striking about movie perception is how easily we take it for granted—how easily we just accept the perceptual power of these dynamic visual images without thinking twice. Instead we simply get home from work, click on the TV and allow hours to pass by. We rarely, if ever, notice the fact that we are sitting alone in a room watching an illuminated 2D screen rapidly flash a series of spatially and temporally discontinuous images before our eyes. That's because we don't perceive these dynamic visual stimuli as what they are. We don't notice the cuts that occur around every 3 seconds or the fact that the camera is following some set of formal guidelines to introduce us to a scene. In fact, if we did notice all these things, if we did perceive movies as the visual stimuli they are, we'd never watch them. But we do watch them; and we

watch them because the markedly unrealistic qualities tend to go unnoticed as we sit back, put our feet up, and get carried away for 90 minutes.

Amidst the rather dense theoretical and empirical discussion of movie perception, this aspect of the movie viewing experience may seem like the easiest to conceptualize but, in truth, it is likely the most difficult. It is this aspect of the movie viewing experience that underlies the preceding discussion. It is this aspect of the movie viewing experience that contemporary psychocinematic and neurocinematic researchers have spent decades trying to create viable theoretical and empirical accounts of why and how we're able to do this. And yet, we don't think twice about our ability to do this. We are content spending hours sitting in front of a 2D illuminated screen as our favorite couch engulfs our body and our favorite movie engulfs our mind. And while this is not really of any radical importance to the movie viewer, it is to the field of neuroscience.

First off, the preceding discussion reveals that the methodological utility of movie stimuli in the field of neuroscience is—at the very least—a means to obtain real-time data on how the brain processes information within a more contextually rich, dynamic environment. And since movie stimuli provide a more contextually rich, dynamic means to study human cognition, they also provide—at the very least—a more realistic picture of how human cognition functions within the context of everyday. If the field of neuroscience is able to acknowledge the utility of movies, this alone is a step in the right direction; a step towards a more realistic understanding of human cognition. But there is something else, something potentially more important for the field of neuroscience to see within the preceding discussion of movie perception—this concerns deception.

Our ability to watch movies is the result of our ability to perceive something as something else. And this, my friends, involves deception on the part of our brain—deception that is not unique to movie perception. The human brain often deceives us. We don't notice the blind spot in the middle of our field of vision or the fact that we make thousands of saccadic eye movements every hour. Instead, our brain appears to 'fill in' the gaps for us; or at least, it seems like it does. And this deception on the part of the brain is important to the field of neuroscience. These things that go unnoticed tell us something about how our brain processes information; and not how we think it processes information, but how it actually processes information. There is a difference between how things appear to us, and how things really are and understanding this perceptual disconnect is important; and while those within the discipline of the neuroscience of art recognize this—recognize that the perception of illusions tell us about the perceptual capabilities of our brains, tell us that there is a difference between the environment we think we live in and the environment we are capable of perceiving—mainstream human cognition researchers often ignore it. The reason movie stimuli might be of such value to the field of neuroscience is that, because the use of movie stimuli is proliferating within the field, it may be possible for the 'fringe' ideology underlying disciplines such as the neuroscience of art to be widely accepted within the field as of great importance to a general understanding of human perception.

If the aspects of movie viewing that give way to this type of deceptive perception are understood, we can begin to understand the building blocks of perception—the necessary and sufficient pieces of the brain's input of information that form our perception of everyday. As Patrick Cavanagh (2005) discusses in his paper, "The artist as neuroscientist," the scenes we see in art transgress the rules of physics; yet, these transgressions are often unnoticed and do not interfere with our ability to understand the depicted scene. In the case of movie perception, it becomes

clear that not all aspects of our reality are *necessary* for the visual processing and comprehension of the scenes of our everyday life. The ability of movies to transgress the physics of reality without impairing the human brains' ability to perceptually process them "reveal[s] that our visual brain uses simpler, reduced physics to understand the world" (Cavanagh, 2005). The key here is that movies are abstractions of the reality we experience; abstractions that contain the *necessary* and *sufficient* amount of perceptual information needed to for us to unambiguously perceive them. Therefore, what movie perception really tells us is that our perception of the real world as stable and continuous has very little to do with the actual nature of the environment we inhabit, and more to do with the manner in which we perceptual process information. The unnoticed saccadic eye movement, the unnoticed movie cut, these things tell us that the reality we perceive is the reality we are capable of perceiving—not the reality that actually exists. Our everyday reality is thus more of a fictional reality the human brain conjures up in order to make sense of its inputted discontinuous snapshots (i.e. saccadic suppression), to fill in the imperceptible aspects of the environment in a coherent manner. And if we can identify the limits of human perception, it is thus theoretically possible that we will be able to simulate the reality of the human brain within the context of a laboratory. The psychological reality of the visual movie viewing experience thus provides us a theoretically possible solution to the field of neuroscience's current methodological impasse.



## References

- Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, 411(6835), 305-309.
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, 108(25), 10367-10371.  
doi:10.1073/pnas.1104047108; 10.1073/pnas.1104047108
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, 108(25), 10367-10371.  
doi:10.1073/pnas.1104047108; 10.1073/pnas.1104047108
- Bach, D. R., Talmi, D., Hurlmann, R., Patin, A., & Dolan, R. J. (2011). Automatic relevance detection in the absence of a functional amygdala. *Neuropsychologia*, 49(5), 1302-1305.
- Banks, S. J., Eddy, K. T., Angstadt, M., Nathan, P. J., & Phan, K. L. (2007). Amygdala-frontal connectivity during emotion regulation. *Social Cognitive and Affective Neuroscience*, 2(4), 303-312.  
doi:10.1093/scan/nsm029; 10.1093/scan/nsm029
- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, 15(4), 600-609.
- Barbas, H. (1995). Anatomic basis of cognitive-emotional interactions in the primate prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, 19(3), 499-510.
- Barrett, L. F., & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1521), 1325-1334. doi:10.1098/rstb.2008.0312; 10.1098/rstb.2008.0312
- Barrett, L. F., & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1325-1334.  
doi:10.1098/rstb.2008.0312
- Bartels, A., & Zeki, S. (2004). Functional brain mapping during free viewing of natural scenes. *Human Brain Mapping*, 21(2), 75-85.

- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex (New York, N.Y.: 1991)*, 19(12), 2767-2796. doi:10.1093/cercor/bhp055; 10.1093/cercor/bhp055
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex (New York, N.Y.: 1991)*, 19(12), 2767-2796. doi:10.1093/cercor/bhp055; 10.1093/cercor/bhp055
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, 8(12), 539-546.
- Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: An integrative theoretical account. *Trends in Cognitive Sciences*, 16(6), 338-352.
- Carroll, N., & Seeley, W. (2013). Cognitivism, psychology, and neuroscience: Movies as attentional engines. *Psychocinematics: Exploring Cognition at the Movies*, , 53-75.
- Cohen, J. D., Botvinick, M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience*, 3, 421-423.
- Cohen, J. D., Botvinick, M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience*, 3, 421-423.
- Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., Miller, G. A., . . . Heller, W. (2003). Paying attention to emotion. *Cognitive, Affective, & Behavioral Neuroscience*, 3(2), 81-96.
- Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience*, 14(3), 508-523.
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, 58(3), 306-324.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201-215.

- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3(3), 292.
- Cutting, J. E., Brunick, K. L., & Candan, A. (2012). Perceiving event dynamics and parsing hollywood films. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1476.
- Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science*, 20(6), 778-784. doi:10.1111/j.1467-9280.2009.02360.x; 10.1111/j.1467-9280.2009.02360.x
- Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 353(1373), 1245-1255. doi:10.1098/rstb.1998.0280
- Dieter, K. C., Hu, B., Knill, D. C., Blake, R., & Tadin, D. (2014). Kinesthesia can make an invisible hand visible. *Psychological Science*, 25(1), 66-75. doi:10.1177/0956797613497968; 10.1177/0956797613497968
- Doallo, S., Patai, E. Z., & Nobre, A. C. (2013). Reward associations magnify memory-based biases on perception. *Journal of Cognitive Neuroscience*, 25(2), 245-257.
- Dolcos, F., Diaz-Granados, P., Wang, L., & McCarthy, G. (2008). Opposing influences of emotional and non-emotional distracters upon sustained prefrontal cortex activity during a delayed-response working memory task. *Neuropsychologia*, 46(1), 326-335.
- Drevets, W. C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: Implications for interactions between emotion and cognition. *Cognition and Emotion*, 12(3), 353-385.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1), 1-47. doi:10.1093/cercor/1.1.1
- Frishkoff, G. A., Tucker, D. M., Davey, C., & Scherg, M. (2004). Frontal and posterior sources of event-related potentials in semantic comprehension. *Cognitive Brain Research*, 20(3), 329-354.

- Gehring, W. J., & Knight, R. T. (2000). Prefrontal–cingulate interactions in action monitoring. *Nature Neuroscience*, 3(5), 516-520.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, 13(2), 159-166. doi:[http://dx.doi.org/10.1016/S0959-4388\(03\)00040-0](http://dx.doi.org/10.1016/S0959-4388(03)00040-0)
- Hanson, S. J., Negishi, M., & Hanson, C. (2001). Connectionist neuroimaging. *Emergent neural computational architectures based on neuroscience* (pp. 560-576) Springer.
- Harsay, H. A., Cohen, M. X., Oosterhof, N. N., Forstmann, B. U., Mars, R. B., & Ridderinkhof, K. R. (2011). Functional connectivity of the striatum links motivation to action control in humans. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 31(29), 10701-10711. doi:10.1523/JNEUROSCI.5415-10.2011; 10.1523/JNEUROSCI.5415-10.2011
- Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. (2008). Neurocinematics: The neuroscience of film. *Projections*, 2(1), 1. doi:10.3167/proj.2008.020102
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, 303(5664), 1634-1640.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4), 679.
- Hsu, S., & Pessoa, L. (2007). Dissociable effects of bottom-up and top-down factors on the processing of unattended fearful faces. *Neuropsychologia*, 45(13), 3075-3086.
- Jimura, K., Locke, H. S., & Braver, T. S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proceedings of the National Academy of Sciences of the United States of America*, 107(19), 8871-8876. doi:10.1073/pnas.1002007107; 10.1073/pnas.1002007107
- Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., & Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 25(18), 4593-4604. doi:10.1523/JNEUROSCI.0236-05.2005

- Konigsberg, I. (2007). Film studies and the new science. *Projections*, 1(1)  
doi:<http://dx.doi.org/10.3167/proj.2007.010102>
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.
- Lim, S. L., Padmala, S., & Pessoa, L. (2009). Segregating the significant from the mundane on a moment-to-moment basis via direct and indirect amygdala contributions. *Proceedings of the National Academy of Sciences of the United States of America*, 106(39), 16841-16846.  
doi:10.1073/pnas.0904551106; 10.1073/pnas.0904551106
- Lindsay, V. (1970). The art of the moving picture. 1915. *New York: Liveright*.
- Luo, Q., Mitchell, D., Jones, M., Mondillo, K., Vythilingam, M., & Blair, R. J. R. (2007). Common regions of dorsal anterior cingulate and prefrontal–parietal cortices provide attentional control of distracters varying in emotionality and visibility. *Neuroimage*, 38(3), 631-639.
- Magliano, J. P., & Zacks, J. M. (2011). The impact of continuity editing in narrative film on event segmentation. *Cognitive Science*, 35(8), 1489-1517.
- Magliano, J., P., & Zacks, J., M. (2011). The impact of continuity editing in narrative film on event segmentation. *Cognitive Science*, 35, 1489-1517. doi:10.1111/j.1551-6709.2011.01202.x
- Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation*, 3(1), 5-24.
- Mitchell, D., Luo, Q., Mondillo, K., Vythilingam, M., Finger, E., & Blair, R. (2008). The interference of operant task performance by emotional distracters: An antagonistic relationship between the amygdala and frontoparietal cortices. *Neuroimage*, 40(2), 859-868.
- Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain : A Journal of Neurology*, 121 ( Pt 1)(Pt 1), 47-57.
- Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain*, 120(4), 701-722.
- Münsterberg, H. (1916). *The photoplay: A psychological study* D. Appleton and company.

- Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience*, 10(5), 360-372.
- Newton, D., Engquist, G. A., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology*, 35(12), 847-862. doi:10.1037/0022-3514.35.12.847
- Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: A synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, 1251(1), E1-E24.
- Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, 23(11), 3419-3432.
- Pantoja, J., Ribeiro, S., Wiest, M., Soares, E., Gervasoni, D., Lemos, N. A., & Nicolelis, M. A. (2007). Neuronal activity in the primary somatosensory thalamocortical loop is modulated by reward contingency during tactile discrimination. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 27(39), 10608-10620. doi:10.1523/JNEUROSCI.5279-06.2007
- Papez, J. W. (1937). A proposed mechanism of emotion. *Archives of Neurology & Psychiatry*, 38(4), 725-743.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148-158.
- Pessoa, L., Padmala, S., Kenner, A., & Bauer, A. (2012). Interactions between cognition and emotion during response inhibition. *Emotion*, 12(1), 192.
- Pessoa, L., & Engelmann, J. B. (2010). Embedding reward signals into perception and cognition. *Frontiers in Neuroscience*, 4, 10.3389/fnins.2010.00017. eCollection 2010. doi:10.3389/fnins.2010.00017; 10.3389/fnins.2010.00017
- Pessoa, L., Padmala, S., & Morland, T. (2005). Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *Neuroimage*, 28(1), 249-255. doi:<http://dx.doi.org/10.1016/j.neuroimage.2005.05.048>

- Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, 17(4), 292-299.
- Piech, R. M., McHugo, M., Smith, S. D., Dukic, M. S., Van Der Meer, J., Abou-Khalil, B., . . . Zald, D. H. (2011). Attentional capture by emotional stimuli is preserved in patients with amygdala lesions. *Neuropsychologia*, 49(12), 3314-3319.
- Piech, R. M., McHugo, M., Smith, S. D., Dukic, M. S., Van Der Meer, J., Abou-Khalil, B., & Zald, D. H. (2010). Fear-enhanced visual search persists after amygdala lesions. *Neuropsychologia*, 48(12), 3430-3435.
- Pleger, B., Blankenburg, F., Ruff, C. C., Driver, J., & Dolan, R. J. (2008). Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 28(33), 8161-8168. doi:10.1523/JNEUROSCI.1093-08.2008; 10.1523/JNEUROSCI.1093-08.2008
- Ramos-Estebanez, C., Merabet, L. B., Machii, K., Fregni, F., Thut, G., Wagner, T. A., . . . Pascual-Leone, A. (2007). Visual phosphene perception modulated by subthreshold crossmodal sensory stimulation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 27(15), 4178-4181. doi:10.1523/JNEUROSCI.5468-06.2007
- Ray, R. D., & Zald, D. H. (2012). Anatomical insights into the interaction of emotion and cognition in the prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, 36(1), 479-501.
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, 23(1), 473-500.
- Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, 60(6), 1169-1181.
- Shuler, M. G., & Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science (New York, N.Y.)*, 311(5767), 1606-1609. doi:10.1126/science.1123513
- Sirigu, A., Zalla, T., Pillon, B., Grafman, J., Agid, Y., & Dubois, B. (1996). Encoding of sequence and boundaries of scripts following prefrontal lesions. *Cortex*, 32(2), 297-310.

- Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology*, 40(1), 160-164.
- Smith, T. J., & Henderson, J. M. (2008). Edit blindness: The relationship between attention and global change in dynamic scenes. *Journal of Eye Movement Research*, 2, 1-17.
- Smith, T. J. (2006). An attentional theory of continuity editing.
- Smith, T. J. (2012). The attentional theory of cinematic continuity. *Projections*, 6(1), 1-27.
- Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision*, 13(8), 16.
- Smith, T., & Shimamura, A. (2013). Watching you watch movies: Using eye tracking to inform cognitive film theory. *Psychocinematics: Exploring Cognition at the Movies*, , 165-191.
- Smith, T. J., Levin, D., & Cutting, J. E. (2012). A window on reality: Perceiving edited moving images. *Current Directions in Psychological Science*, 21(2), 107-113. doi:10.1177/0963721412437407
- Sokolov, E. (1990). The orienting response, and future directions of its development. *The Pavlovian Journal of Biological Science*, 25, 142-150.
- Speer, N. K., Swallow, K. M., & Zacks, J. M. (2003). Activation of human motion processing areas during event perception. *Cognitive, Affective, & Behavioral Neuroscience*, 3(4), 335-345.
- Taya, S., Windridge, D., & Osman, M. (2012). Looking to score: The dissociation of goal influence on eye movement and meta-attentional allocation in a complex dynamic natural scene. *PloS One*, 7(6), e39060.
- Taylor, S. F., Phan, K. L., Decker, L. R., & Liberzon, I. (2003). Subjective rating of emotionally salient stimuli modulates neural activity. *Neuroimage*, 18(3), 650-659.
- Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., & Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nature Neuroscience*, 12(10), 1224-1225. doi:10.1038/nn.2380
- Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology*, 4(2), 157-165.



- Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science*, 283(5401), 549-554.
- Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science*, 283(5401), 549-554.
- Viviani, R. (2013). Emotion regulation, attention to emotion, and the ventral attentional network. *Frontiers in Human Neuroscience*, 7
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, 7(11), 1271-1278.
- Weil, R. S., Furl, N., Ruff, C. C., Symmonds, M., Flandin, G., Dolan, R. J., . . . Rees, G. (2010). Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex. *Journal of Neurophysiology*, 104(3), 1746-1757. doi:10.1152/jn.00870.2009; 10.1152/jn.00870.2009
- Zacks, J. (2013). Constructing event representations during film comprehension. In A. Shimamura (Ed.), *Psychocinematics: Exploring cognition at the movies* () Oxford University Press. doi:10.1093/acprof:oso/9780199862139.001.0001
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, 133(2), 273.
- Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006). Visual motion and the neural correlates of event perception. *Brain Research*, 1076(1), 150-162.
- Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., . . . Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6), 651.
- Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, 130(1), 29-58. doi:10.1037/0096-3445.130.1.29

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151-175. doi:10.1037/0003-066X.35.2.151

Zalla, T., Phipps, M., & Grafman, J. (2002). Story processing in patients with damage to the prefrontal cortex. *Cortex*, 38(2), 215-231.

DO NOT COPY

## ACKNOWLEDGEMENTS

---

Thank you Professor Seeley. Thank you Bates College. Thank you parents. Thank you friends. But actually, thank you Professor Seeley. You've encouraged me to think critically about the world we see around us and you've helped me realize my capabilities as a critical thinker.

DO NOT COPY

## APPENDIX

**Table 1** Regions selectively activated in the contrasts of (A) fearful versus happy, (B) happy versus fearful and (C) emotional versus neutral conditions

Area	Coordinates (x, y, z)	Z-score
(A) Fearful versus happy contrast		
Left amygdala (including periamygdaloid cortex)	-14, -8, -20	4.26
Left cerebellum	-42, -68, -20	3.57
Right superior frontal gyrus (BA6)	22, 4, 64	3.22
Left cingulate gyrus	-10, 28, 16	3.18
(B) Happy versus fearful contrast		
Right middle temporal gyrus (BA 21)	54, 4, -20	3.55
Right putamen	22, -4, 12	3.48
Left superior parietal lobule (BA 5)	-28, -40, 60	3.41
Left superior parietal lobule (BA 7)	-24, -72, 44	3.12
Left calcarine sulcus	-10, -92, 4	3.06
(C) Emotional versus neutral contrast		
Left occipitotemporal sulcus	-40, -32, -12	3.53
Right orbitofrontal cortex	12, 12, -20	3.35
Left pulvinar	-4, -32, 8	3.10

Coordinates of the maximal points of activation and the associated Z-values are shown. The activations in all regions are significant at  $P < 0.001$  (uncorrected). In (C) the neutral condition includes the 0% and 25% faces. In the amygdala, the only area predicted to show a response, this is equivalent to a significance level of  $P < 0.05$ , corrected for multiple spatial comparisons in a  $2 \times 2 \times 2$  cm search region. The activations in the other brain regions would have had the same corrected level of significance if they had also been predicted *a priori*. All  $P$ -values are one-tailed.

**Table 2** Regions selectively activated with (A) increasing fearful intensity and (B) increasing happy intensity

Area	Coordinates (x, y, z)	Z-score
(A) Increasing fearful intensity		
Left anterior insula	-34, 26, 6	3.91
Left posterior insula	-36, -26, -4	3.58
Left inferior frontal gyrus	-24, 40, 0	3.49
Left precuneus	-16, -52, 48	3.47
Left pulvinar	-8, -30, 8	3.36
Right anterior cingulate	10, -4, 32	3.18
(B) Increasing happy intensity		
Left fusiform gyrus	-42, -66, -20	3.90
Left lingual gyrus	-40, -32, -12	3.72
Right calcarine sulcus	16, -70, -12	3.57
Left calcarine sulcus	-10, -98, 0	3.41
Right fusiform gyrus	44, -34, -16	3.28
Right superior temporal gyrus	56, -40, 8	3.15

The contrasts were performed by weighting each condition according to the proportion of fearful or happy prototype in the faces (see Methods section). Coordinates of the maximal point of activation and the associated Z-values are shown. The activations in all regions are significant at  $P < 0.001$  (uncorrected).

**Table 3** *Regions showing an enhanced contribution from the left amygdala during the presentation of (A) fearful and (B) happy faces*

Area	Coordinates (x, y, z)	Z-score
(A) Fearful versus happy psychophysiological interactions		
Left inferior occipital gyrus	-40, -84, -4	4.27
Right middle temporal gyrus	38, -56, 12	3.82
Right dorsal midbrain	12, -34, -4	3.67
Right cerebellum	16, -62, -20	3.40
Right inferior occipital gyrus	44, -80, 0	3.54
Left hippocampus	-34, -16, -16	3.50
Right fusiform gyrus	40, -38, -24	3.49
Left cerebellum	-12, -58, -24	3.29
(B) Happy versus fearful psychophysiological interactions		
Left inferior frontal gyrus	-36, 38, 12	3.58
Right uncus	8, 4, -28	3.41
Right precentral gyrus	56, 0, 16	3.23
Left inferior temporal gyrus	-58, -16, -16	3.23

A regression analysis employing the rCBF values for the maximal voxel in the left amygdala ( $x = -14$ ,  $y = -8$ ,  $z = -20$ ) was used to identify the regions (see Methods section). All activations are significant at  $P < 0.001$  (uncorrected).