2016

Improving Anomaly Detection through Identification of Physiological Signatures of Unconscious Awareness

Alyssa Marie Piasecki
Wright State University

Improving Anomaly Detection through Identification of Physiological Signatures of Unconscious Awareness

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Biomedical Engineering

By

Alyssa Marie Piasecki
B.S., Wright State University 2015

2016
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

April 20, 2016

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Alyssa Marie Piasecki ENTITLED Improving Anomaly Detection through Identification of Physiological Signatures of Unconscious Awareness BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Biomedical Engineering.

Mary Fendley, Ph.D., Thesis Director

Jaime Ramirez-Vick, Ph.D., Chair
Department of Biomedical, Industrial
and Human Factors Engineering

Committee on
Final Examination

Rik Warren, Ph.D.

Nasser H. Kashou, Ph.D.

Mary Fendley, Ph.D.

Robert E.W. Fyffe, Ph.D.,
Vice President for Research and
Dean of the Graduate School

ABSTRACT

Piasecki, Alyssa Marie. M.S.B.M.E., Department of Biomedical, Industrial and Human Factors Engineering, Wright State University, 2016. *Improving Anomaly Detection through Identification of Physiological Signatures of Unconscious Awareness.*

Missed anomalies have the potential to cause detrimental effects in the Intelligence, Surveillance, and Reconnaissance (ISR) domain. One possible cause of these missed anomalies is that cognitive processing may not reach conscious awareness and may only be perceived by the unconscious mind. Identification of correlates of these unconscious processes could provide an insight into potential missed targets. The present study explored missed anomalies in a visual search task and the possibility of unconscious awareness. Eye metrics were recorded and a "Detection Threshold Model" was created and validated with a nominal logistic regression model, in order to characterize the search patterns and eye metrics of detection, non-detection, and possible unconscious detection. Results indicated that eye metrics of fixation count, fixation duration, mean saccade length, and backtrack rate predicted detections and non-detections with an overall accuracy of about 90%. Additionally, gaze plots of possible unconscious detections revealed signature search patterns of detection.

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

This research was supported by the Dayton Area Graduate Studies Institute (DAGSI) and the Air Force Research Laboratory (AFRL), 711th Human Performance Wing, Human Effectiveness Directorate, Human Analyst Augmentation Branch located at Wright-Patterson AFB. With sincere appreciation, I thank them for their collaboration and support on this effort. I would also like to thank my graduate advisor, Dr. Mary Fendley, for her continuous guidance and support throughout the duration of this research.

1.0 Introduction

1.1 Background

The ever-increasing rate of automation in image capture and the resulting increase in volume of imagery to analyze outpaces the ability of intelligence analysts (IA) to process them. Due to the increase in the amount of data availability, the cognitive limit of IAs is continuously pushed to its limits, and the identification of anomalies within the military image analysis task is becoming increasingly time-critical (Fendley & Narayanan, 2012; Muller & Narayanan, 2009; Duvall, 2005; Maule, 1997). Additionally, the random and unexpected nature of certain anomalies provides an increased difficulty of timely anomaly detection (Warren, Smith, & Cybenko, 2011). Due to the increase in availability of data and a subsequent increase in workload of IAs, time-saving tactics are often employed in an attempt to keep up with the increasing pace of data generation. A common example of a time-saving tactic in the Intelligence, Surveillance, and Reconnaissance (ISR) domain is to selectively scan intelligence feeds for specific regions of interest (ROIs) and ignore others that do not seem of particular importance. This selective attention could, therefore, lead to missed anomalies in unexpected ROIs or outside the ROIs. Additionally, it is well-known and accepted in the field of cognitive psychology that people have no conscious experience of most of what happens in the human mind (Heuer, 1999; Simon, 1957; March, 1978). This research will explore this concept of unconscious awareness through missed anomalies in a visual search task and presents a methodology for determining possible unconscious detections by investigating physiological signatures of detected and un-detected anomalies.

1.2 Research Objective

The objective of this study is to not only show that unconscious processing exists, but that there are physiological eye-tracking signatures of these phenomena that can be detected and used for acknowledgement and mitigation purposes. The present study will investigate physiological signatures of anomaly detection during missed anomalies in a visual search task, create a model of detection versus non-detection, and identify events in which unconscious anomaly detection may be occurring. This paper will first provide relevant background information and then go on to discuss the methodology, analysis, results, and discussion of the present research.

2.0 Literature Review

2.1 Intelligence, Surveillance, and Reconnaissance

Intelligence, Surveillance, and Reconnaissance (ISR) programs are a crucial part of the United States Department of Defense (DoD) that serve as the center for planning, execution, and assessment of issues concerning global situational awareness and national security. In order to fully understand what the ISR community is responsible for, the term can be broken down into its parts. According to Barber (2001), intelligence is defined as "the product of processed information concerning hostile or potentially hostile forces", surveillance is the "systematic observation by technical sensors or human beings [which] implies continuous 24 hours a day / 7 days a week [observation] of areas or forces of interest", and reconnaissance is defined as the "directed mission(s) to obtain specific information". Combining these separate terms, Barber defined ISR as "the capability that integrates command direction, sensors, and processed formation and intelligence with timely dissemination in order to provide decision makers with effective ‘Situational Awareness’”. In other words, it can be said that one of the main goals of the ISR community is information acquisition in many forms such as real-time video feed or text-based data streams, as well as the corresponding analysis of the gathered intelligence in order to provide timely situational awareness and national security. There exists a specific characterization of intelligence depending on the source of how it was obtained, which includes Human Intelligence (HUMINT) from a person observing, Imagery Intelligence (IMINT) from photographs or other imagery, Signals Intelligence (SIGINT) from electronic signals, and Measurement and Signatures Intelligence (MASINT) from measurable aspects of the target (Chizek, 2003). When analyzing these various types of intelligence, there is the potential for real-time analysis by the ISR community or, in other less time-critical situations, imagery captured by the UAVs can be relayed back to IAs, who will analyze the data “offline”, or not in real-time. The analysis process, which could include still-, motion-, or text-based imagery, is a complex task that involves the integration of many cognitive processes. In order to fully understand this process, it is important to know the steps involved in human reasoning and decision making.

2.2 Human Reasoning and Decision Making

The process of surveying and identifying anomalies in a military image analysis task is a cognitively-demanding and high-workload process. In order to ensure that IAs are making timely, effective, and trusted decisions in critical situations, it is first important to understand the process of decision making, the involvement of human reasoning, and the steps leading up to these cognitive tasks.

Before the process of decision making or human reasoning occurs, the IA first needs to gather the intelligence into what is referred to as "working memory". Working memory is a term that refers to the maintenance and storage of information in the short term and can be described as the system that underlies human thought processes. Following the information processing model developed by Wickens (1992) as shown in Figure 1, sensory information enters a short-term sensory store where the information is transformed into an understandable form by the perceptual processes of the brain. After perception, the information is transferred to working memory which interacts with long-term memory in order to grow and develop the individual's perception of the world and determine a reasonable response to the stimuli.

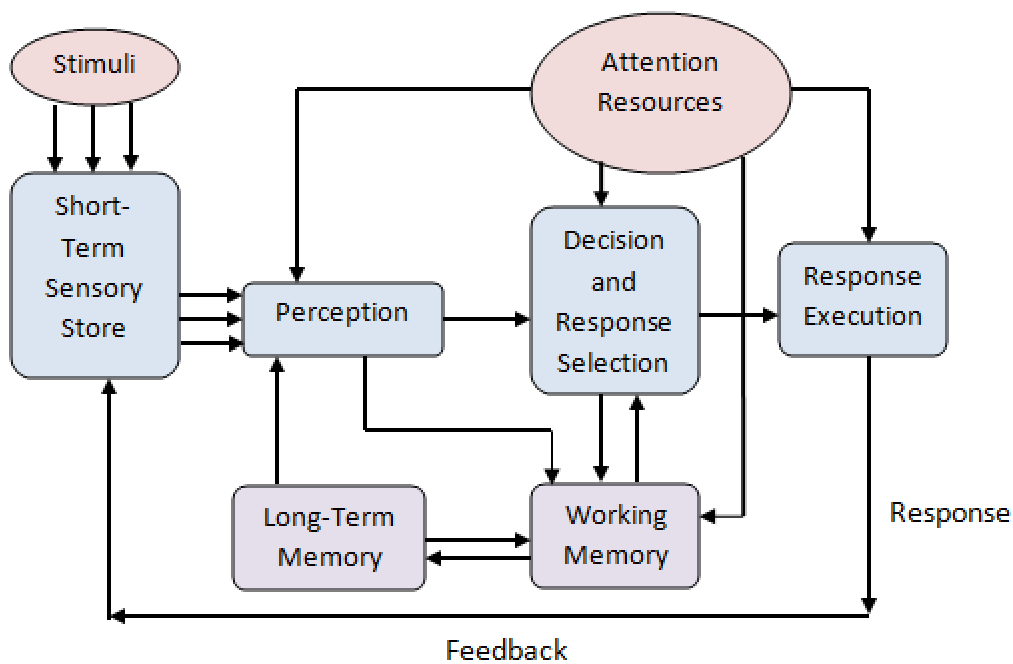


Figure 1: Wickens' Model of Information Processing (Wickens, 1992). A stimulus enters the short-term sensory store where it is transformed into an understandable form in order to be perceived. Working-memory and long-term memory interact to determine an appropriate response that is then executed.

Working memory temporarily maintains this information as a means for providing an interface between perception, long-term memory, and action (Baddeley, 2003). Another view of working memory is that it is made up of three major parts: the central executive, the phonological loop, and the visuospatial sketchpad.

The central executive serves as the control system while the phonological loop and visuospatial sketchpad make up the storage systems of the model. With regards to the central executive function, it is arguably the most important of the three components, however, the least understood. The phonological loop exists to facilitate with the acquisition of verbal skills, such as the ability to learn a new language, whereas the visuospatial sketchpad is responsible for the storage of visual cues. Figure 2 shows the makeup of working memory.

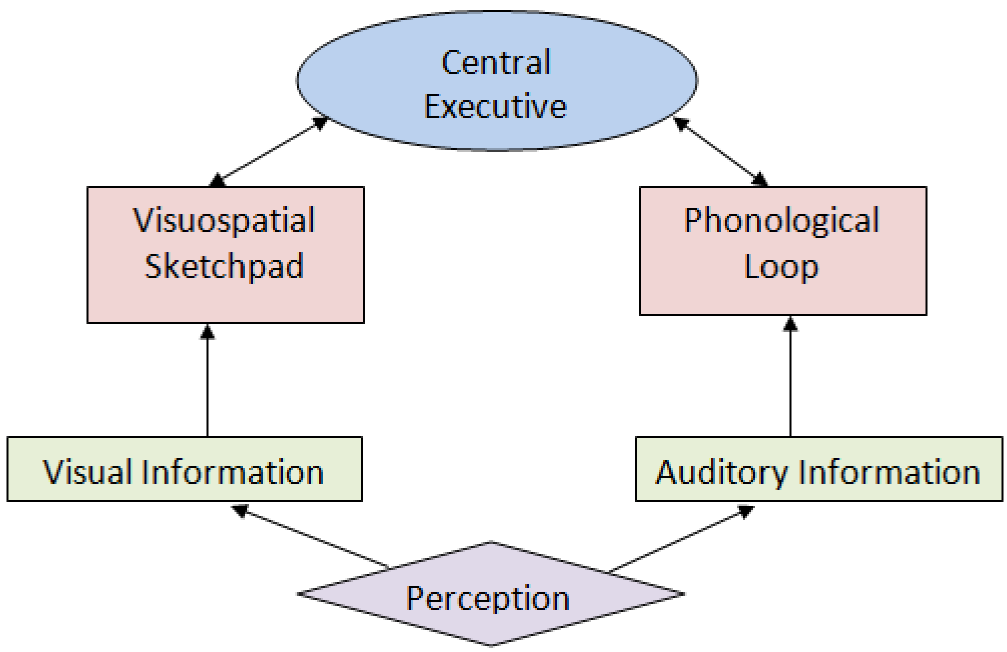


Figure 2: The Baddeley and Hitch (2003) Model of Working Memory. The central executive, or control system, consists of the visuospatial sketchpad, which stores visual information, and the phonological loop, which stores auditory information, in order to achieve perception of the stimulus.

As shown, input perceived by an individual is separated into either visual or auditory information, which is then relayed to the corresponding storage system that ultimately is transferred to the central executive for processing in order to determine if action needs to be taken or if the information needs to be stored in long-term memory, for example.

Once information is gathered in the brain, there are several theories and pathways for that knowledge and intelligence to be further processed and analyzed, such as case-based reasoning, naturalistic decision making, dual-process theory, fuzzy-trace theory, and intuition. Although this is not intended to be an exhaustive list, the following paragraphs will briefly discuss each of these methods in order to gain an understanding of the types of pathways available for decision making and reasoning.

2.2.1 Case-Based Reasoning

Case-based reasoning refers to the problem-solving method in which past experiences and previously stored knowledge about a certain topic are applied to the current situation at hand (Aamodt, 1994). With this approach, a new problem is solved by recalling a specific, similar case from the past and applying that knowledge to the new situation. This means that the strategy is an incremental, sustained learning process, meaning that with each problem solved, another solution is retained in memory for further potential use of future applicable problems. In order to fully understand this concept, consider the following scenario:

> *David, a family doctor, is examining a patient with specific symptoms, and he is reminded of a patient that he had several weeks ago with very similar symptoms. He recalls that he did not think the previous patient's symptoms were serious, and therefore*

just advised the patient to get rest and drink fluids. However, he remembers that the patient came back a couple days later with an even greater decline in health, so he decided to prescribe antibiotics which quickly remedied the symptoms. Using this knowledge, David decides to prescribe antibiotics to the new patient right away.

In this example, David uses a past experience in order to effectively solve the current problem at hand. He recalls that "rest and fluids" was not an effective treatment in a previous patient with similar symptoms, and therefore goes straight to the solution that worked for the previous patient. This method is effective because it is coupled with a "learning" process. The more experiences an individual has, the more knowledge that is stored for use in future problems.

2.2.2 Naturalistic Decision Making

Another method for decision making is called naturalistic decision making. This term refers to decision making in complex real-world settings. Prior to the knowledge of naturalistic decision making, other methods were used by theorists and decision makers, including Multi-Attribute Utility Analysis (MAUA) and Decision Analysis (Klein & Klinger, 2008). The methods of MAUA and Decision Analysis focus on the analytical process of decision making and are theoretically successful with regards to the systematic process of weighing and rating solutions, as well as calculating probabilities. However, the short-coming of these methods is that they fall short when it comes to real-world, time-critical, and high-stress situations. These methods are too time-consuming and require extensive work in order to reach a valid decision. Additionally, it is difficult to factor in ambiguity and dynamic environments. Therefore, naturalistic decision making was designed in order to account for these short-comings and serve as a successful method in complex real-world settings. According to Klein et al. (2008), a complex real-world situation is comprised of ten main features, which are summarized in Table 1.

Table 1: Features of Naturalistic Decision Making (Klein, 2008). These characteristic features are applied towards the ISR domain and specific examples are provided.

	Feature	ISR-Specific Example
1	Ill-defined goals and ill-structured tasks	Temporal and spatial locations of targets or threats are unknown
2	Uncertainty, ambiguity, and missing data	Data feeds may be incomplete or unclear
3	Shifting and competing goals	Multiple monitors, data feeds, and systems to analyze and operate
4	Dynamic and continually changing conditions	
5	Action-feedback loops (real-time reactions to changed conditions)	Time-critical decisions and appropriate actions needed
6	Time stress	
7	High stakes	
8	Multiple players	IAs must work with personnel to receive intelligence, analyze data, make appropriate decisions, and act on those decisions
9	Organizational goals and norms	
10	Experienced decision makers	

As stated by Klein et al. (2008), naturalistic decision making accounts for "dynamic and continually changing conditions, real-time reactions to these changes, ill-defined tasks, time pressure, significant personal consequences for mistakes, and experienced decision makers." This method revolves around making decisions without performing analyses, without an in-depth comparison of options, and rarely without a search for an "optimal choice." Instead the method consists of finding the first solution that is time-effective, cost-effective, and plausible.

2.2.3 Dual-Process Theory and Fuzzy-Trace Theory

Dual-process theory, a concept that attempts to explain the process of thinking and reasoning, states that human reasoning is made up of two distinct systems. These systems consist of an intuitive, autonomous system and an analytical, controlled system (Gawronski & Creighton, 2013; Reyna & Brainerd, 2011; Evans, 2008; Wixted, 2007; De Neys, 2006; Evans, 2003). Branching from this concept, another term attempting to explain the process of human reasoning, known as fuzzy-trace theory, is a derivation from the dual-process theory that was originally used to predict improvement in the ability to reason from childhood to adulthood. Through relevant studies (Reyna & Brainerd, 2011), it was determined that there are two parallel memory representations formed in the mind: verbatim traces and gist traces. Verbatim traces refer to knowledge remembered word-for-word exactly and tend to be more specific while gist traces refer to remembering a general meaning or concept.

2.2.4 Intuition

Another important topic to be discussed in this review concerning decision making is intuition. There is not one exact definition of intuition that is unanimously agreed upon and, as Betsch (2008) states, “There are as many meanings for the term *intuition* as there are people using it.” However, the term tends to refer to reaching an answer or solution or idea without conscious effort or reasoning. Some consider it a source of knowledge, some a process, and some even a structure of the brain (Betsch & Glöckner, 2010; Horstmann, Ahlgrimm, & Glöckner, 2009). Nevertheless, the concept that a thought, solution, or idea can be developed without conscious thought is a view that differs to quite an extent from the other previously discussed decision making strategies, although most closely related to naturalistic decision making. The concept of intuition brings several questions to mind, such as “how do these concepts develop in the brain?”, “by what mechanism do these thoughts reach consciousness?”, and “what neurophysiological biomarkers could exist to track these processes?” Numerous studies have attempted to answer these questions, which will be further investigated later in this paper.

2.3 Signal-Detection Theory

A method for analyzing detections involves a concept known as signal-detection theory. The term signal-detection theory (SDT) refers to a statistical technique in which a signal, target, or object of interest is identified through noise or distraction. As the name suggests, it is a technique used to differentiate between a measured electrical signal, such as an electrocardiograph (EEG), and the associated noise of the system. However, the techniques of SDT can also be applied to the field of psychology, such as with the detection of an anomaly or object of interest (Stanislaw & Todorov, 1999).Taking this application into consideration, there are various methods of applying SDT depending on what is being tested and the type of experiment at hand. The main type of experiment that will be focused on in this paper is the "Yes-No Experiment", in which sensitivity is measured in terms of the ability to distinguish between stimuli (Macmillan & Creelman, 2004). Examples of this type of task include finding abnormalities in X-rays or differentiating between two slightly different images of the same scene. In both of these examples, there are clear anomalies (e.g., a fracture in the X-ray or a fire hydrant in one of the images) that may or may not be present in each of the stimuli (each X-ray or each image of the scene). When searching for these anomalies, there are four possible choices for the outcome, as shown in Table 2. If there is an anomaly present, it can either be detected or not detected. Likewise, if there is no anomaly present, there can be a false detection or a correct rejection.

Table 2: Response Matrix for "Yes-No" Experiment (Macmillan & Creelman, 2004). A binary response is characterized by the reality of whether the target is truly present or not in the stimulus.

	Response	
	"Yes"	**"No"**
Target Present	Hit	Miss
Target Not Present	False alarm	Correct rejection

In this type of experimental setup, it is relatively simple to calculate the hit rate. This would be accomplished by dividing the number of "hits" by the total number of “signal trials”, or trials where anomalies or targets known to be in the stimuli. Similarly, the false alarm rate is calculated by dividing the number of "false alarms" by the total number of “catch trials”, or stimuli without targets (Macmillan & Creelman, 2004). For example, take into consideration the task of differentiating between two slightly different images of the same scene. Specifically, imagine two images of a kitchen scene. In some trials, the two images are exactly the same (catch trials). In other trials, a cooking pot is present on the stove in one but not the other with all other details of the images being exactly the same (signal trials). If the participant was asked to state whether there was a difference between the two images, this would be a typical "Yes-No" experiment in which "hits" would be when the participant correctly states that there is a difference and "false alarms" would be when the participant states there is a difference when in reality the two images are exactly the same, as shown in Table 3. Therefore, the false alarm rate is calculated by dividing the total number of false alarms by the total number of catch trials.

Table 3: Response Matrix for Kitchen Scene. This is an example application of the “Yes-No” Experiment Response Matrix, in which classification of the binomial response is dependent on the reality of whether the kitchen pan is present in the scene.

	Participant’s Response	
	“Yes”	**“No”**
Pan Present	Hit	Miss
Pan Not Present	False alarm	Correct rejection

Alternatively, the false alarm rate is not as simple for other experimental setups. Now consider a task in which participants are asked to find abnormalities in X-rays. The number of "false alarms" is clear: the number of times the participant stated there was an abnormality somewhere in the X-ray when there actually wasn't. However, there is no easy or accurate way to calculate the total number of possible false alarms in the stimuli. There are an infinite number of possibilities for an individual to mistake a normal object as an abnormality; therefore the false alarm rate is not clear. Depending on the experiment, there are possible ways to get around this; however, they tend to be not as accurate or convenient as the hit rate calculation. Therefore, the specific experiment at hand has a strong influence on what results can be accurately measured and calculated.

With regards to differences among experimental setups, they depend partially on the modality of stimulus presentation and partially on the type of anomaly or target being detected. These differences between anomaly types need to be investigated in order to fully understand the possible outcomes of an experiment, which will be covered in the next section.

2.4 Anomaly Detection

An anomaly is commonly understood as an occurrence that deviates from what is normal or expected (Chandola, Banerjee, & Kumar, 2009). Other common terms used to refer to these phenomena are outliers, exceptions, contaminants, or surprises. A simple depiction of an anomaly is shown in Figure 3.

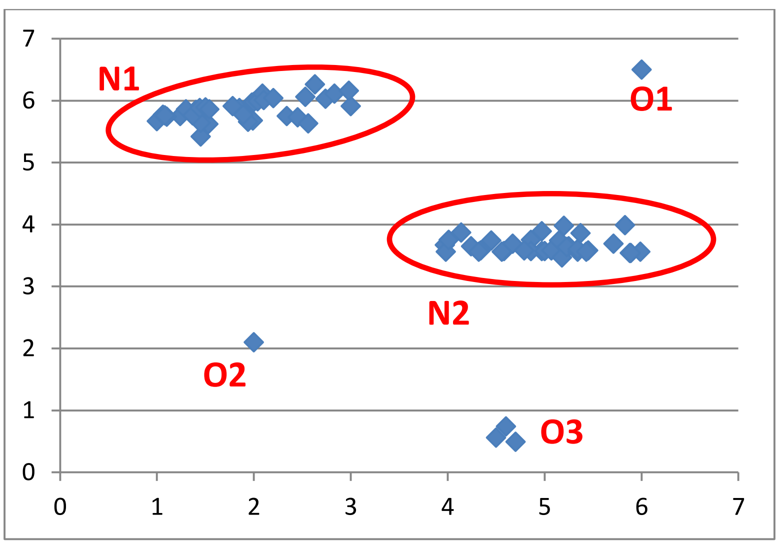


Figure 3: Simple example of anomalies in 2-dimensional data set (Chandola, et. al, 2009). The general trend of the data is depicted by N1 and N2. Points O1, O2, and collection of points O3 are all anomalous.

In this two-dimensional data set, the collection of data points denoted by N1 and N2 shows the expected nature of the data. However, point O1, point O2, and collection of points O3 do not conform to the expected nature of N1 and N2. Therefore, these three points or collection of points are considered anomalous. This, of course, is a very simple example in order to easily portray the definition of an anomaly. Real-world cases are much more complex and involve a more in-depth analysis in order to identify the anomalous points or events.

Applying this definition of an anomaly, it can be inferred that anomaly detection refers to the process of identifying these patterns, occurrences, or behaviors that do not conform to expected behavior. The process of anomaly detection has many possible applications in a wide variety of domains, such as fraud detection, insurance or health care, and military surveillance for enemy activities (Chandola, Banerjee, & Kumar, 2009). The potential types of anomalies within these different domains are numerous, and there are various ways of categorizing them. One way is to separate the anomalies into point anomalies, contextual anomalies, and collective anomalies, as shown in Table 4 (Chandola, Banerjee, & Kumar, 2009). Point anomalies focus on one individual event that is unexpected with respect to the rest of the data, a contextual anomaly is an individual event or outlier that is considered anomalous only in a specific context but not in others, and a collective anomaly is a series of data points or events that are considered anomalous but not necessarily each point individually.

Table 4: Categorization of Anomaly Types (Chandola, et. al, 2009). Point, contextual, and collective anomaly types are described, along with real-world examples.

Type of Anomaly	Definition	Example
Point	One individual data point or instance or event is considered as an outlier with respect to the rest of the data.	A car is going in the wrong direction (against traffic) on a road.
Contextual	One individual data point or instance or event is considered as an outlier in a specific situation (but not otherwise).	The outside temperature in July is 30 degrees Fahrenheit (this would not be anomalous if it were, for example, December).
Collective	The occurrence of a collection of related data points or instances or events is considered as an outlier (but not necessarily individually).	An ECG recording shows a series of flat-line data points.

The present study focuses on point anomalies, which can be separated further into the spatial and temporal domains, as shown in Table 5. Spatial domain point anomalies refer to individual instances that appear anomalous with respect to their position or location. Alternatively, temporal domain point anomalies refer to individual instances that appear anomalous with respect to the time of occurrence (Chandola, Banerjee, & Kumar, 2009).

Table 5: Types of Point Anomalies (Chandola, et. al, 2009). Point anomalies consist of spatial and temporal types. Descriptions, along with common and ISR-specific examples, are described.

Type of Point Anomaly	Definition	Common Example	ISR-Specific Example
Spatial *denoted by <u>underlined text</u>	Focuses on a geographical area or location (a point location or a point area); comparison of multiple still images of a scene	Noticing an image of a gorilla inserted in a computed tomography (CT) lung cancer screening	Aerial still images reveal an individual <u>entering a vacant building</u> at the *same time every day late at night.*
Temporal *denoted by *italicized text*	Includes a time period during which data was collected; comparison of a dynamic scene	Noticing a gorilla walking through players passing a ball	Drone video feed shows an individual <u>digging with a shovel in a deserted area</u> *late at night*.

As shown in Table 5, two common, generic examples, as well as two ISR-specific examples, are provided in order to illustrate the difference between spatial and temporal anomalies. The two generic examples are very well-known studies (see Drew, Vo, & Wolfe, 2013; Simons & Chabris, 1999). The spatial example refers to a study in which naïve observers as well as expert radiologists examined computed tomography (CT) lung cancer screening images for lung nodules, which appeared as small light circles, and failed to detect the presence of a black gorilla 48 times the size of the average nodule (Drew, Vo, & Wolfe, 2013). The temporal example is a study in which a participant watched a video of players passing a ball, with some players wearing black shirts and others wearing white shirts (Simons & Chabris, 1999). The participants were told to count the number of passes that either the white team or the black team had and ultimately failed to detect an individual dressed in a gorilla costume walking directly through the middle of the scene. For the ISR-specific examples, each example contains certain aspects of both spatial and temporal anomalies, where the spatial parts of the anomaly are denoted by underlined text and the temporal components are denoted by italicized text. For the example in the top row, "entering a vacant building" refers to the spatial component due to the fact that it refers to a specific location. The temporal component is "the same time every day late at night" since it refers to a specific point in time. The example in the second row can be explained in a very similar way. The phrase "digging with a shovel in a deserted area" refers to a specific location; therefore it is a spatial anomaly. The temporal component is "late at night" because, once again, it refers to a specific point in time.

During the process of identifying these specific types of anomalies, there are certain phenomena that can occur that ultimately decrease the performance of an IA and cause the potential for a missed anomaly. Two of the more widely studied examples of these phenomena are change blindness and inattentional blindness. Change blindness refers to the phenomenon that occurs when an individual fails to detect changes in a visual scene when the physical changing of the scene is masked, usually by short flickers of the image (Rich & Gillam, 2000). Inattentional blindness is a similar concept, however differs in that it refers to the inability to detect a clearly identifiable, unchanging object in a scene (Gu, Stocker, & Badler, 2005). In Table 5, the common examples from the literature for spatial and temporal anomalies are both examples of inattentional blindness. As the definitions suggest, these two phenomena have the potential to cause detrimental effects in the ISR domain. Research exploring the causes of these happenings, as well as possible ways to detect and mitigate them is needed. The next section will explore these studies, with a focus on physiological monitoring, specifically eye-tracking, as a means of detection.

2.5 Eye-Tracking

The wide area of research devoted to exploring the causes of missed anomalies provides for numerous ways for characterization and detection, including electroencephalography (EEG), electrocardiography (ECG), and eye-tracking methodologies. For the purposes of this study, eye-tracking methods will be used in the experimental setup and, therefore, will be focused on in the remainder of this paper.

As stated previously, the "lung nodule" study as well as the "gorilla costume" study both portrayed well-known examples of inattentional blindness. For the "lung nodule" study (Drew, Vo, & Wolfe, 2013), eye-tracking methodologies were used during experimentation on all participants. The results revealed that twenty out of twenty-four expert radiologists failed to report seeing the gorilla, even though eye tracking confirmed that 12 out of the 20 radiologists that failed to detect the gorilla actually looked directly at the gorilla's location when it was visible in the CT scans (mean dwell time 547 ms). This discovery raises questions, such as the cause for the inattentional blindness, possible identifiers of the occurrence of this phenomenon, and the possible inclusion of unconscious processes during the examination of images.

Several other studies (Droll, Hayhoe, Triesch, & Sullivan, 2005; Simons, Chabris, Schnur, & Levin, 2002; Hollingworth & Henderson, 2002) show similar results, indicating that changes in a scene receive longer fixation durations even though the changes are not consciously recognized. For example, Droll et al. (2005) performed a study to explore how the visual scene or task at hand affects the acquisition of information from that scene. Results from the study revealed that fixation durations on changed objects were longer than other areas of the scene, yet the change still went unnoticed.

2.5.1 Correlates of Unconscious Awareness

The occurrence of unconscious processing, and possibly unconscious awareness, during anomaly detection is the basis of this research effort and is a topic that has been explored by numerous other researchers (Rensink, 2004; Spering, Pomplun, & Carrasco, 2011; Spering & Carrasco, 2015; Galpin, Underwood, & Chapman, 2008; Rothkirch, Stein, Sekutowicz, & Sterzer, 2012; Underwood, Templeman, Lamming, & Foulsham, 2008; Chen & Yeh, 2012). It is evident, from studies focusing on change blindness, that relatively little information from the visual world is internally stored. However, change blindness could be due to other reasons even if a mental representation of the pre-change visual scene is stored. One example of this is the failure to compare the pre-change scene to the post-change scene (Simons, Chabris, Schnur, & Levin, 2002; Hollingworth & Henderson, 2002). Taking into consideration the change blindness paradigm, it has been proposed that, even though a participant does not provide an explicit reporting of the change, it does not mean that the change was not detected at all. It only means that an explicit report is not sensitive enough to measure the change (Hollingworth & Henderson, 2002). In a study performed by Simons et al. (2002), it was shown that participants failed to notice a change initially; however, they were later able to report the exact change when the experimenter provided a clue as to what the change was. This provides an interesting proposition that the participants stored a mental representation of the scene; however it did not reach consciousness until explicitly pointed out to them. This study, therefore, provides evidence that visual information acquisition and mental encoding can occur as unconscious processes. Research by Rensink (2004) further supports these findings, who stated that a visual experience (i.e., consciously seeing or noticing) is not required in order to become aware of an object, event, or the surroundings. In the study, images of a scene and a changed scene were presented to the participant, with the change being either being related to presence (or non-presence), color, or location of an object. Participants viewed the images and were asked to press a key when they first had a “feeling” that a change was occurring and again when they saw explicitly what the change was. Results from the study suggested that visual changes can be sensed without an explicit visual experience. A follow-up to this study was performed by Galpin et al. (2008), who found that sensing did indeed occur in participants without an actual visual experience (as opposed to being random and guess-based) and, furthermore, that sensing and actually visually seeing are two different processes altogether. Other studies have further shown that there is a possibility that visual information processing is distinct and different for perception and for motor action, indicating that eye movements can reflect unconscious visual processing (Spering, Pomplun, & Carrasco, 2011; Spering & Carrasco, 2015).

The goal of the present study is to not only show that unconscious processing exists, but that there are physiological eye-tracking signatures of these phenomena that can be detected and used for acknowledgement and mitigation purposes. There are significantly fewer papers focused specifically on these goals; however, research has been done in an attempt to accomplish these tasks (Rothkirch, Stein, Sekutowicz, & Sterzer, 2012; Jacob & Hochstein, 2009). Rothkirch et al. (2012) performed a study in which participants performed a search task in order to locate a Gabor patch that was made supposedly invisible using continuous flash suppression (CFS) techniques (Tsuchiya & Koch, 2005). According to Rothkirch et al. (2012), “CFS is thought to largely disrupt neural signals from the suppressed eye at early central processing stages, but may leave some subcortical processes and responses in dorsal visual cortical areas relatively preserved”. The participants were asked the location of the Gabor patch, the orientation, and were subjected to a confidence rating. Results of “very unsure” participants showed that location and orientation were at chance level and, therefore, the participants had no subjective or objective awareness of the Gabor patch. However, dwell times of the participants revealed that they were increased by 40% for the Gabor patch area relative to the control areas. These results indicate that participants’ eye movement patterns were affected by the unconscious perception of stimuli.

The present research will explore the concept of unconscious awareness through missed anomalies in a visual search task. A methodology for determining possible unconscious detections is described by investigating physiological signatures of detected and un-detected anomalies.

2.6 Test Objective & Hypothesis

Rothkirch et al. (2012) provided evidence that objects made inherently invisible, and shown to be "unseen" by an individual, can still exhibit various effects in the search patterns and fixations of that individual. This leads to the notion that possible unconscious visual processing can occur during search tasks, or any other image processing experience for that matter. The details of the image may not reach the level of conscious awareness; however, they are processed at some level below this conscious threshold through which effects can still be experienced. The present study aims to apply this phenomenon to an environment more closely related to that experienced by IAs in the ISR domain by using ISR-related images and search tasks. A simulated analyst environment is also accomplished by the ambiguity of the presence of an anomaly. In other words, subjects are not told if or when an anomaly is present in the image, much like a real IA experience. Given the previously explained research, it is hypothesized that the presence of an anomaly changes the search activity of the individual and, more specifically, causes the individual to increase the fixation count, fixation duration, saccade length, and backtrack rate in the area of the anomaly, as follows, where μ is the mean of the respective metric:

$$H_o: \mu_{detections} = \mu_{non-detections}$$

$$H_1: \mu_{detections} > \mu_{non-detections}$$

3.0 Research Methodology

3.1 Participants

A total of 33 Wright State University engineering students participated in this study. The subject pool consisted of 12 female and 21 male subjects with ages ranging from 20–52 years (mean = 23.6 years). All participants had normal or corrected to normal vision. During the analysis phase, three participants were discarded due to a low quality of eye-tracking data (12%, 15%, and 17%), which refers to the percentage of samples collected throughout all trials for that participant. The data quality of the remaining 30 participants ranged from 37%-80% (mean = 64.5%) and were used for the analysis. The variation in these percentages can be attributed to factors such as the participant blinking or looking off-screen. Figure 4 shows a box plot of the data quality percentages for all participants. Note the three discarded participants as outliers in the figure.

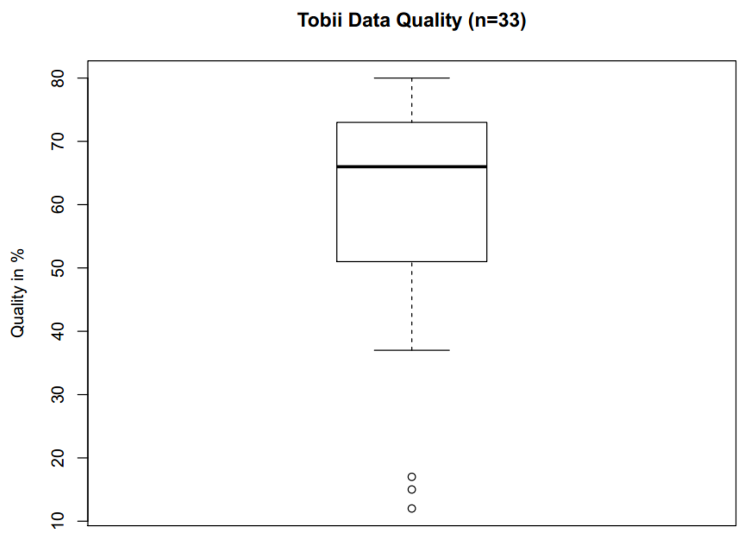


Figure 4: Box plot of Data Quality Percentages across Participants. The data quality refers to the percentage of samples collected from the eye-tracker. Three participants are identified as outliers and were excluded from the analysis.

3.2 Testing Environment & Apparatus

All testing was performed in the Human Performance and Cognition Laboratory in the Neuroscience Engineering Collaboration building at Wright State University. The lighting, ambient temperature, and ambient noise level were held constant for all participants. Testing was performed using a Tobii T120 Eye Tracker monitor with a data collection rate of 60 Hz, screen size of 17", and screen resolution of 1280 x 1024 pixels. Participants were seated approximately 50-70 cm from the Tobii monitor.

3.3 Stimuli

The experimental display consisted of a static background RGB (8 bit unsigned integer) image with static card suit symbols and dynamic signals hidden in the image.

The images used were a series of frames taken from infrared (IR) movies. Each scene had one of each of the four standard playing card suit symbols (heart, spade, diamond, club) hidden in the image. An example of the static background image with hidden suits is shown in Figure 5. Note that the red circles are for indication purposes only and were not present during the actual experiment.

Figure 5: Example of Suits. An example image is depicted with the locations of the four hidden suits indicated by red circles.

The signals were small flashing circles, of which there were two difficulty levels referred to as "easy" and "difficult". Two difficulty levels were established in order to account for differences in visual thresholds among different individuals and provide for a stronger chance of unconscious visual awareness among all participants. According to Carmi et al. (2006), the primary characteristics of visual images that are noticed by an individual are color contrast, motion, and flicker (Rummukainen, Radun, Virtanen, & Pulkki, 2014; Hamel, Houzet, Pellerin, & Guyader, 2015; Wu, Wick, & Pomplun, 2014; Açık, Bartel, & König, 2014). Therefore, these three characteristics were manipulated in order to provide for a clear difference between the easy and difficult signals. Each easy signal was made up of a small circle approximately 50 pixels in diameter that flashed three times (250 ms flash duration with 50 ms between each flash) while “traveling” in a linear fashion and had a mean contrast difference amplitude of 23.2 units. Each difficult signal was made up of a small circle approximately 50 pixels in diameter that flashed three times (250 ms flash duration with 50 ms between each flash) in a stationary position and had a mean contrast difference amplitude of 4.9 units.

An example of an easy signal that appears 12 s into the 35 s trial is shown in Figure 6. Recall that an easy signal flashes three times while moving in a linear fashion. In Figure 6, each of the three windows shows a different position of the signal as it moves. Note that this particular signal is moving up and to the left as it progresses. Once again, the red circles are for indication purposes only and were not present during the actual experiment. The timeline below the images indicates the time that each of the three flashes appeared. As previously discussed, each flash stays on the screen for 250 ms and then disappears, with a 50 ms wait time before the next flash appears. Therefore, the 3 flashes of an easy signal appear at 12 s, 12.3 s, and 12.6 s, assuming that the signal starts at 12 s from the start of the trial.

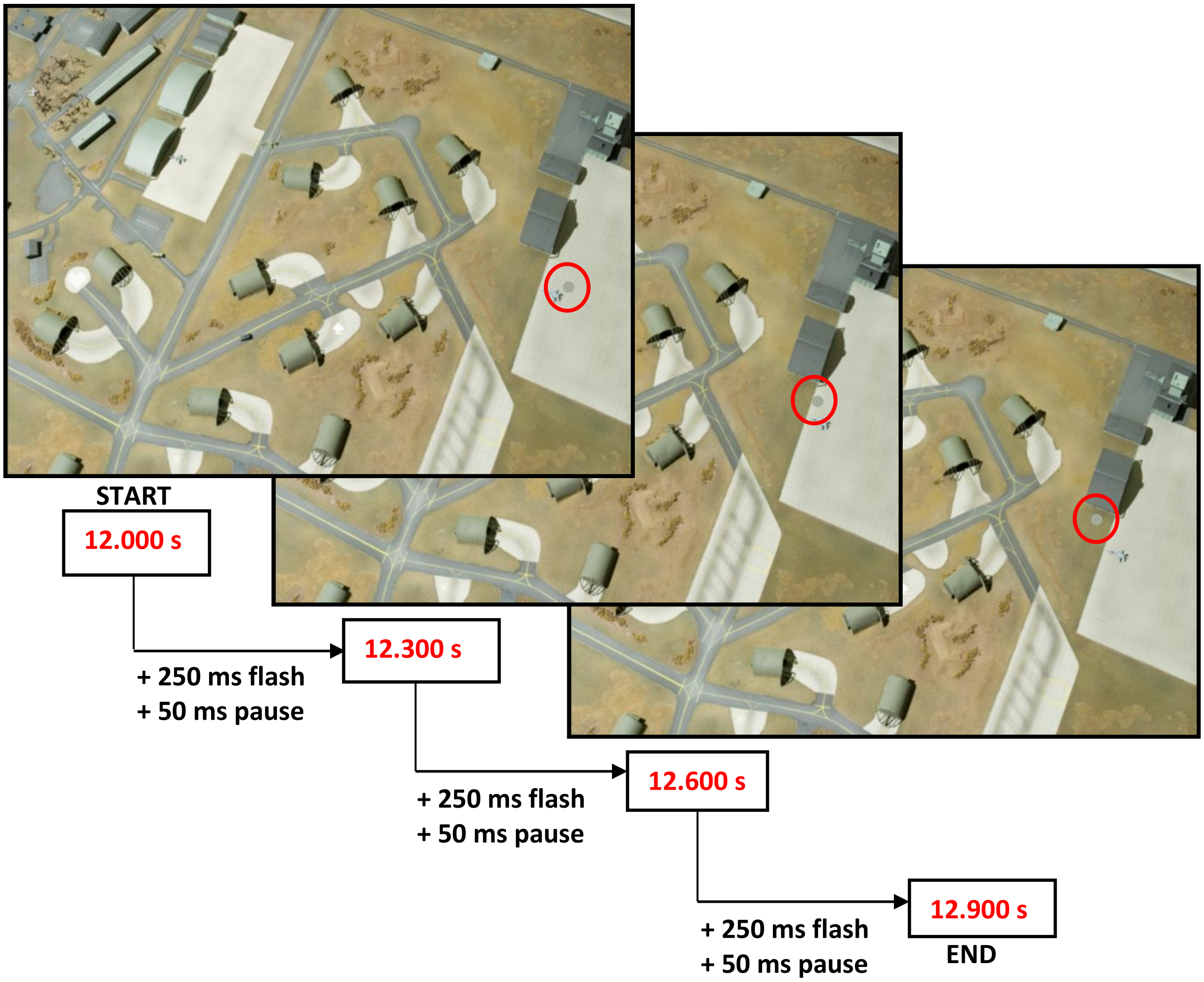


Figure 6: Example of an Easy Signal. In each of the three windows, the location of the signal is depicted by a red circle as it moves. As an example, the signal has an onset time of 12 s. The signal flashes three times, with a 250 ms flash and 50 ms pause between each flash.

The "mean contrast difference" values for the easy and difficult signals were calculated using Matlab. The gray index value at the center of each signal was measured, along with the gray index value of the background surrounding the signal at a distance of 25 pixels from the center of the signal (approximately equal to the radius of the signal). The absolute difference of these two gray index values was then calculated in order to have a relative difference in contrast for each signal. The contrast difference measured was the overall contrast difference of that signal. As the easy signals "travelled" in a linear direction, each of the flash dots was against a different background as it moved. This means that each of the three flashes that made up each signal had different contrast differences. Therefore, in order to obtain an overall contrast difference for each signal, the three contrast differences from the three flash dot positions of each easy signal were averaged in order to have one contrast difference value.

There was the possibility of a combination of 0 or 1 easy signals and 0, 1, or 2 difficult signals in each image. The total number of difficult signals in the experimental setup across all trials was higher than the number of easy signals to provide more opportunities to exhibit unconscious detection. The matrix in Table 6 shows the six possible combinations of these two levels of signals.

Table 6: Possible combinations of Easy and Difficult Signals. There was a possibility of zero, one, or two difficult signals and zero or one easy signals. From these, six possible combinations of signals exist. These combinations were repeated once to give a total of twelve trials.

	D0	D1	D2
E0	E0D0	E0D1	E0D2
E1	E1D0	E1D1	E1D2

As shown in Table 6, there was a possibility of a minimum of zero signals and a maximum of three signals (1 easy and 2 difficult) in any one image. Whether the total number of signals was 0, 1, 2, or 3, the timing of the signals was held constant for all trials (12s, 20s, 25s).

Figures 7 and 8 show the cumulative locations of signals and suits, respectively, for the twelve trials. In Figure 7, each of the six easy signals is shown as a combination of three small red arrows, symbolizing the three flashes, travelling in a linear fashion. The tail of each small red arrow indicates the center of the flash and the head of the arrow indicates the position 25 pixels from the center where the gray index value was measured for the contrast difference calculation. The three red arrows of each easy signal are connected by a larger black arrow. The tail of the black arrow shows the start of the sequence of flashes and the head indicates the end of the sequence. The positions of the twelve difficult signals are indicated by small blue arrows. Once again, the tail of the arrow indicates the center of the signal and the head indicates the position 25 pixels from the center where the gray index value was measured. In Figure 8, the relative positions of the forty-eight card suits (twelve trials with four suits each) for the experiment are shown.

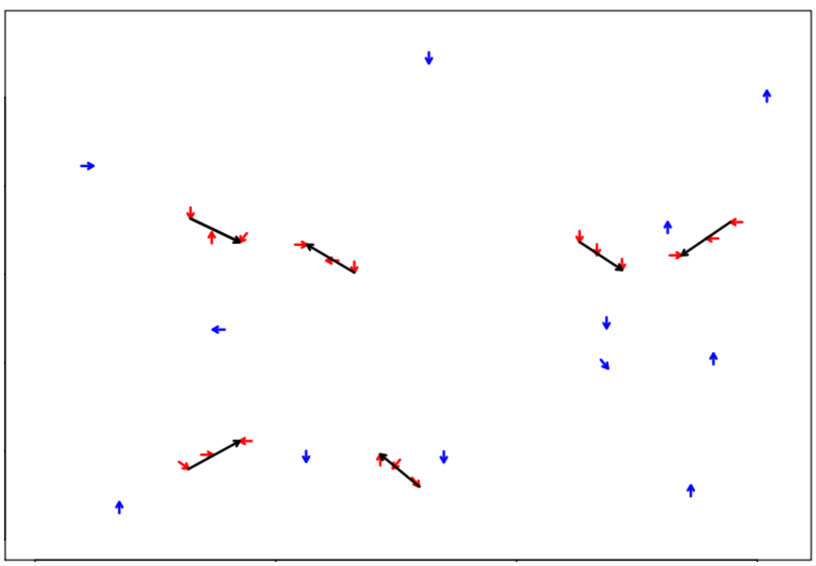


Figure 7: Signal Reference Locations. The six easy signals are depicted by large black arrows, with the tail of the arrow indicating the starting position and the head indicating the end position. The three small red arrows in each large black arrow indicate the locations of the three flashes of each easy signal. The blue arrows indicate locations of the difficult signals. The tail of the red and blue arrows indicates the center of the flash and the head of the arrow indicates the position 25 pixels from the center where the gray index value was measured for the contrast difference calculation.

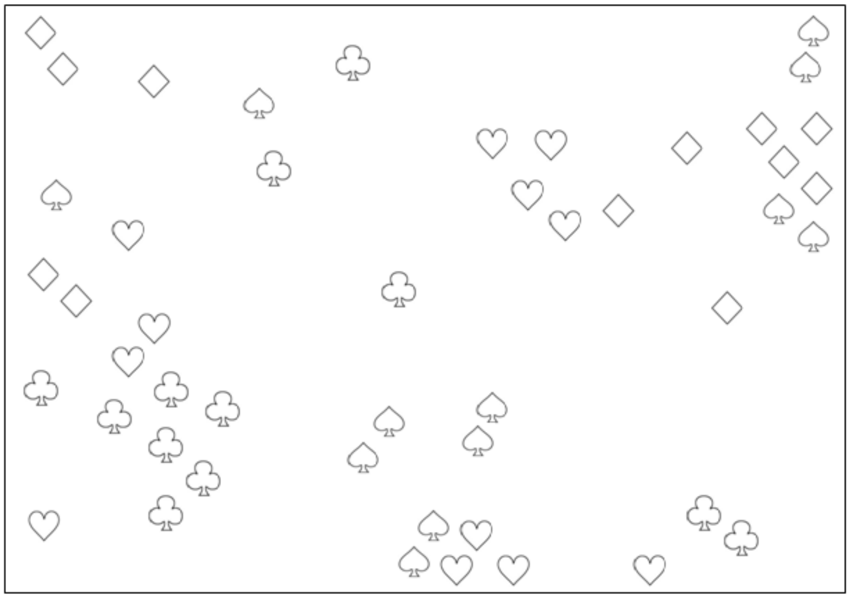


Figure 8: Suit Reference Locations

3.4 Experimental Design

The experiment was a 2x3 within subjects design consisting of two independent variables: number of easy signals (two levels: zero or one) and number of difficult signals (three levels: zero, one, or two). These factors were repeated once to give a total of 12 trials in the experiment. The dependent variable was the detection of the target (two levels: detected or not-detected), where "target" is the all-inclusive term used to refer to suits and signals, collectively. These variables were then used to perform a nominal logistic regression in order to determine if the target metrics of fixation count, fixation duration (s), mean saccade length (pixels), and backtrack rate (/sec) could predict detection.

3.5 Procedure

All participants signed a written consent form prior to experimentation. A pre-test questionnaire was administered to obtain basic demographic information as well as experience with video/computer gaming. Participants were briefed about the study and told that their main objective was to find the card suits hidden in the image and press the space bar when they found each suit. They were also told that there may or may not be small flashing circles somewhere in the image and that they did not have to take any action if/when they saw these signals. An eye calibration of the Tobii monitor was performed, and the participant then completed a practice trial with one easy signal in order to ensure understanding of the testing procedure. Participants then viewed each of the twelve images and answered a short post-clip questionnaire after each trial. This post-clip questionnaire asked the participants to state if they saw any signals, in which quadrant of the computer monitor they noticed the signal(s), and their confidence rating in seeing each signal(s). After completion of the twelve trials, a post-test questionnaire was administered to the participants, asking about their search strategy during the visual search task. Upon completion of the post-test questionnaire, participants were thanked for their time and released.

A flow chart summary of the procedures is shown in Figure 9. The total time for the experiment was approximately one hour.

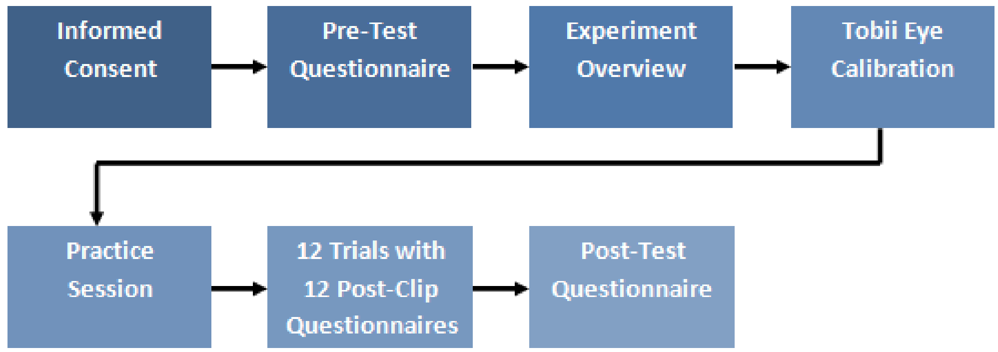


Figure 9: Experimental Timeline

4.0 Results

A preliminary analysis of data was performed in order to determine the total number of detections and false alarms (FA) for suits (Figure 10) and signals (Figure 11). Additionally, hit and FA rates were calculated for each target type (Figure 12-13). In order to calculate the FA rate, the total number of FA opportunities was estimated to be equal to the total number of target detection opportunities (Stanislaw & Todorov, 1999). Therefore, there were a total of 18 signal FA opportunities per participant (6 easy + 12 difficult) and a total of 48 suit FA opportunities per participant (4 suits x 12 trials).

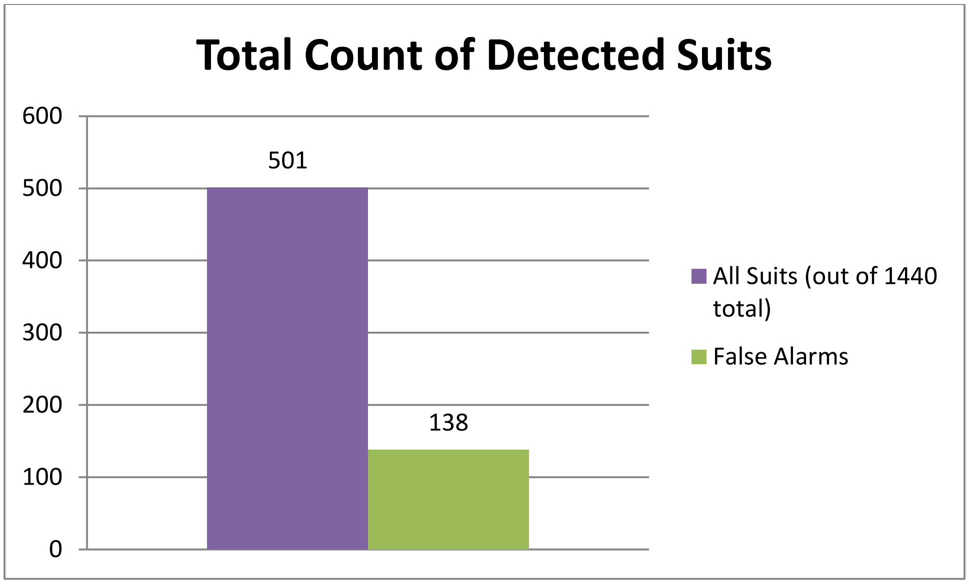


Figure 10: Count of Suit Detections and False Alarms. Out of a total of 1440 suits, 501 suits were detected and there were 138 false alarms.

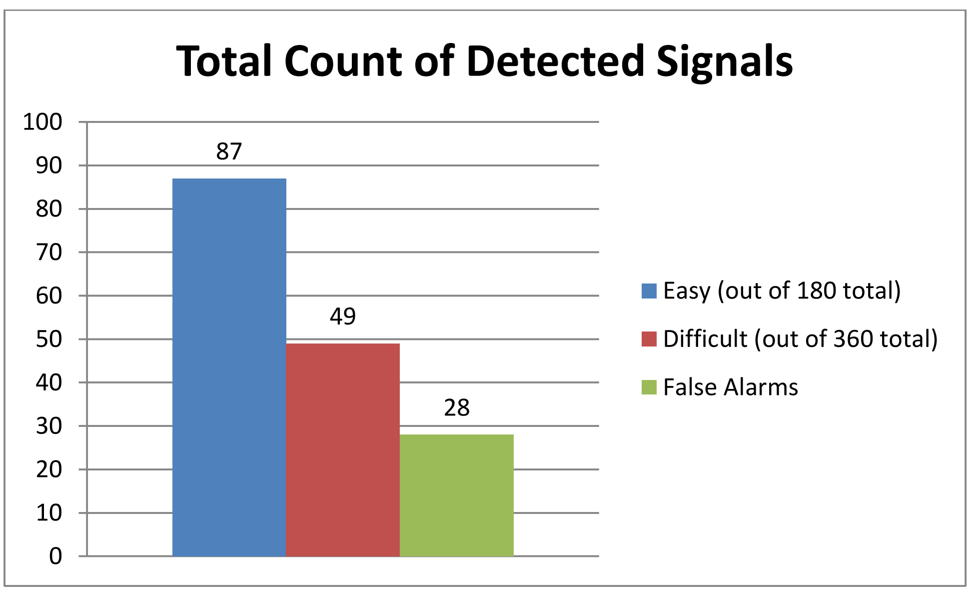


Figure 11: Count of Easy & Difficult Signal Detections and False Alarms. There were 87 out of 180 easy signals detected, 49 out of 360 difficult signals detected, and 28 signal false alarms.

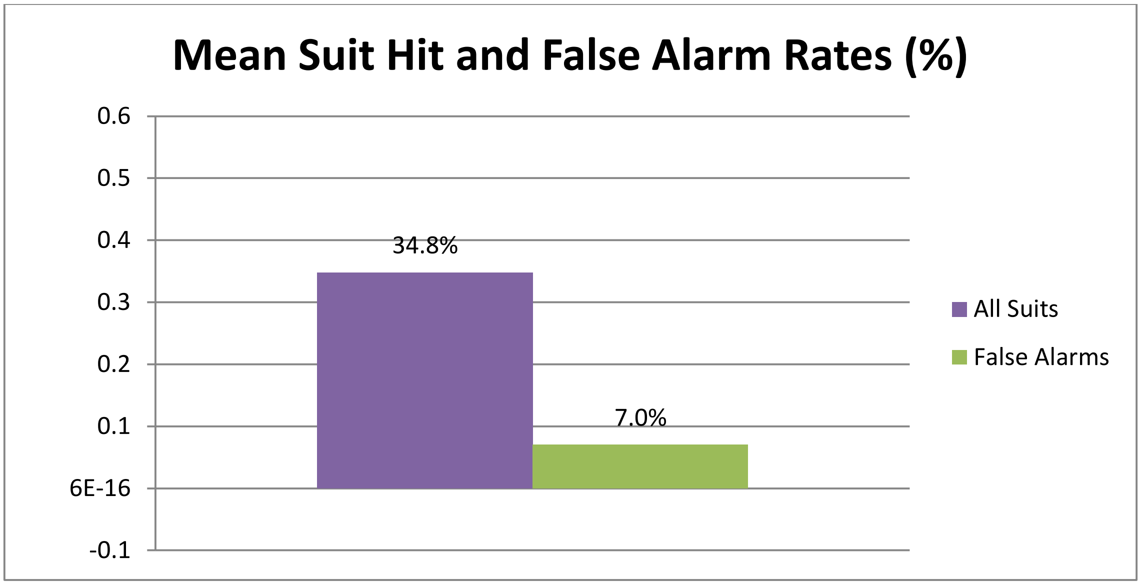


Figure 12: Suit Hit and False Alarm Rates. The suit detection rate was 34.8% and the false alarm rate was 7.0%.

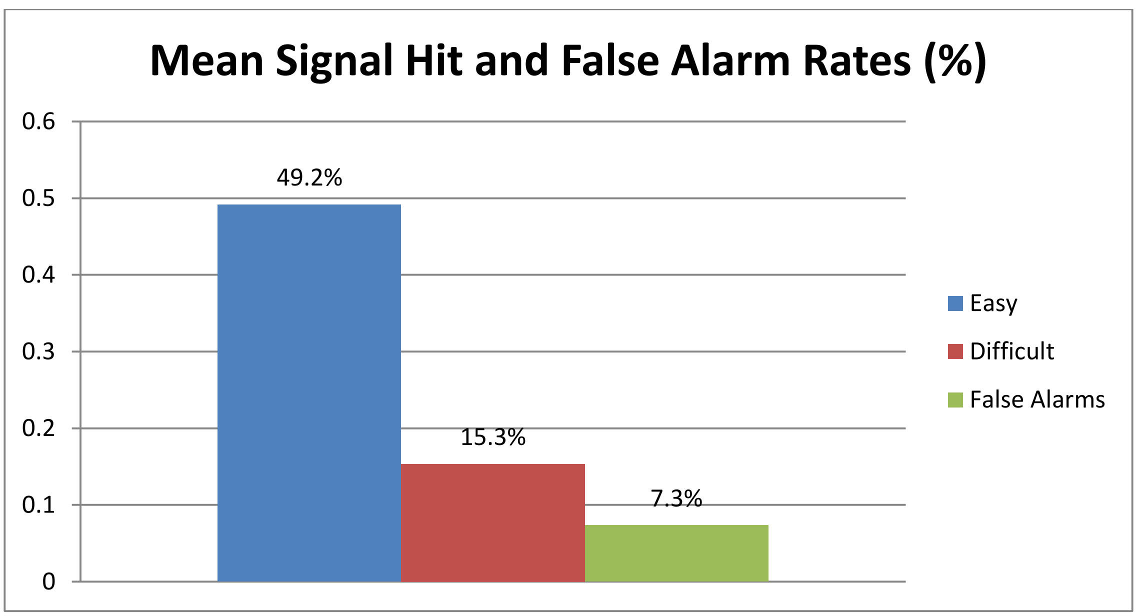


Figure 13: Signal Hit and False Alarm Rates. The easy signal detection rate was 49.2%, the difficult signal detection rate was 15.3%, and the false alarm rate was 7.3%.

4.1 Eye-Tracking Metrics

Four eye-tracking metrics were investigated for correlation with detection/non-detection and, ultimately, unconscious detection. These four metrics included fixation count, fixation duration (s), mean saccade length (px), and backtrack rate (/sec). The measurement of fixation count and fixation duration were extracted from the Tobii Studio software; however, mean saccade length and backtrack rate were not directly available and were calculated using a custom algorithm written in Matlab. Saccade length was defined to be the length in pixels between two sequential fixations (Moacdieh & Sarter, 2015; Goldberg & Kotval, 1999). A backtrack was defined as an angle greater than ninety degrees between two sequential saccades and, therefore, backtrack rate was the number of backtracks per second (Moacdieh & Sarter, 2015; Goldberg & Kotval, 1999).

The time interval used for measurement of each of the four metrics was based on mean choice reaction time and was estimated to be 1000ms (Vaportzis, Georgiou-Karistianis, Churchyard, & Stout, 2015; Ratcliff & Smith, 2004). However, measurement of each of the four metrics used in the model varied, depending on the type of target. For signals, the metrics were measured within the +1000ms interval after the signal appeared on the monitor. For detected suits, the metrics were measured within the ±1000ms interval of detection. For undetected suits, the metrics were measured across the entire 35s trial due to the lack of a timestamp for detection or appearance. In order to account for this variation in measurement times, the metrics were normalized with respect to time.

4.2 Detection Classification

The analysis of results and pathway for identification of possible unconscious detections is depicted in Figure 14. Results of the visual search task were separated into detected and not detected responses. Behavior of the detections was analyzed, and a model was created in order to depict this behavior. This “Detection Threshold Model” was based on the physiological metrics of fixation count, fixation duration, mean saccade length, and backtrack rate. The results from the unpaired-t tests were inspected in order to determine which significant metrics to include in the model. Using this model, the experimental data was applied and, specifically, non-detections were applied in order to determine if the behavior was similar to detections, indicating a potential unconscious detection. In addition, gaze plots and an analysis of transition rate were performed in order to investigate the behavior of each trial as a whole. These analyses are discussed later in this paper.

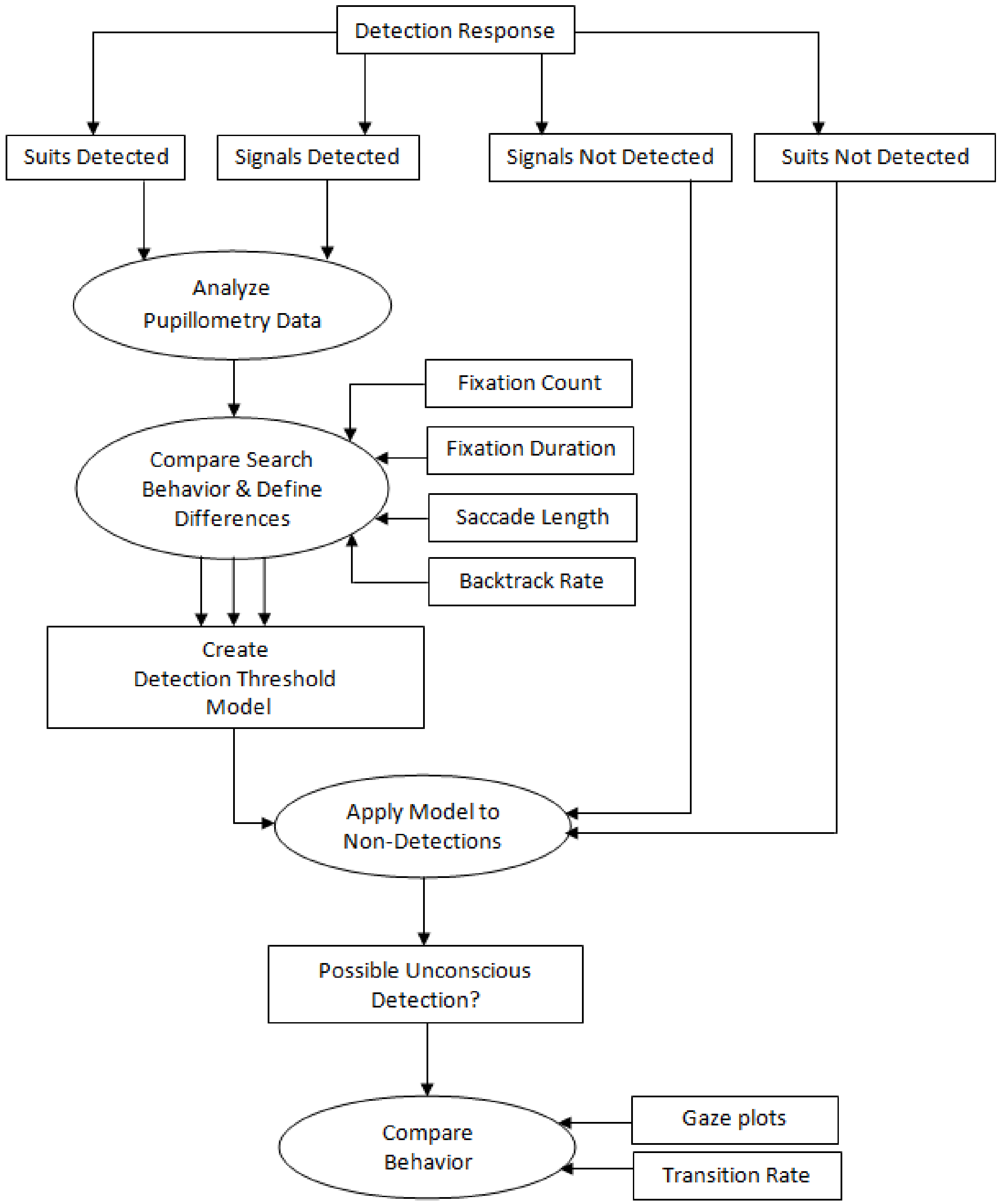


Figure 14: Experimental Analysis Model. Pupillometry data of detections was analyzed in order to create a Detection Threshold Model. Non-detections were applied to this model for classification as an unconscious detection or true non-detection. Gaze plots and transition rate were analyzed for overall search patterns.

Scatterplots with jitter applied, along with box plots, were created to serve as a preliminary investigation of the differences between detection and non-detection for fixation count (Figure 15), fixation duration (Figure 16), mean saccade length (Figure 17), and backtrack rate (Figure 18).

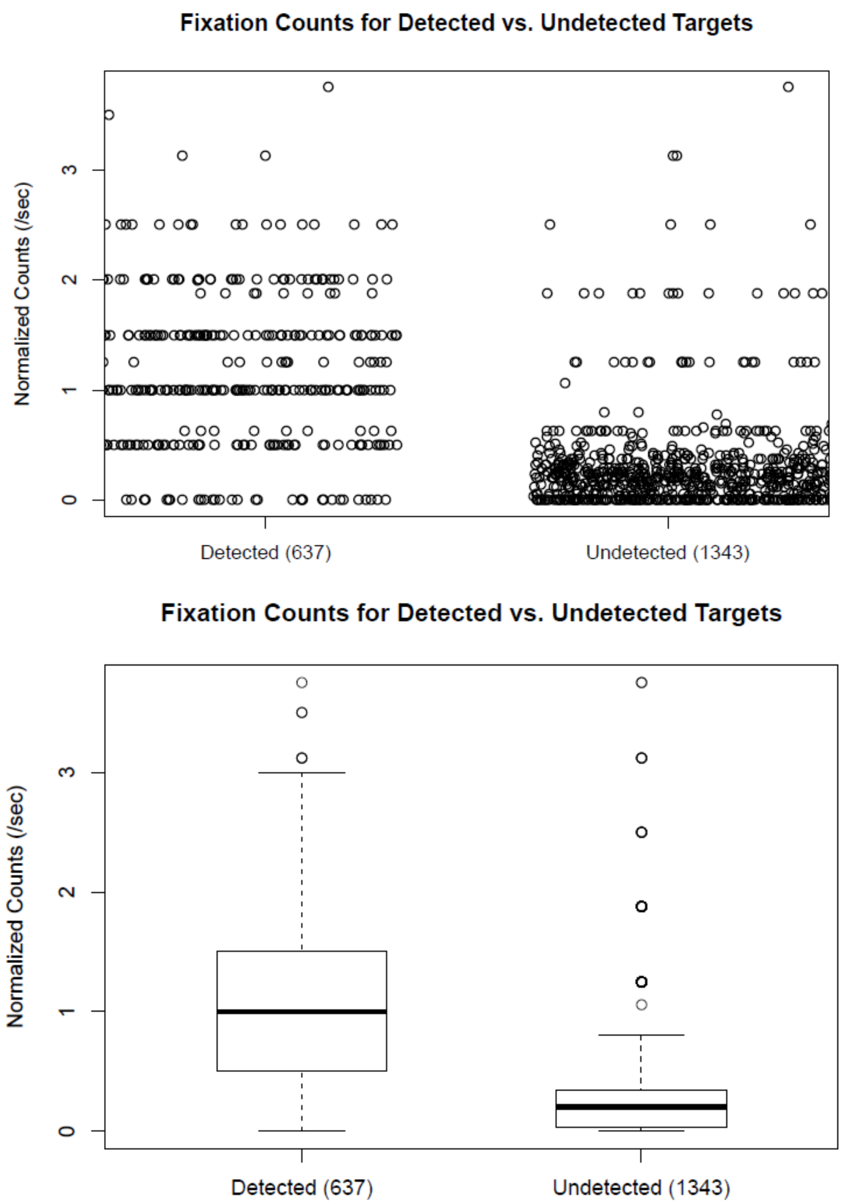


Figure 15: Scatterplot with jitter applied (top) and boxplot (bottom) of fixation count for detections and non-detections.

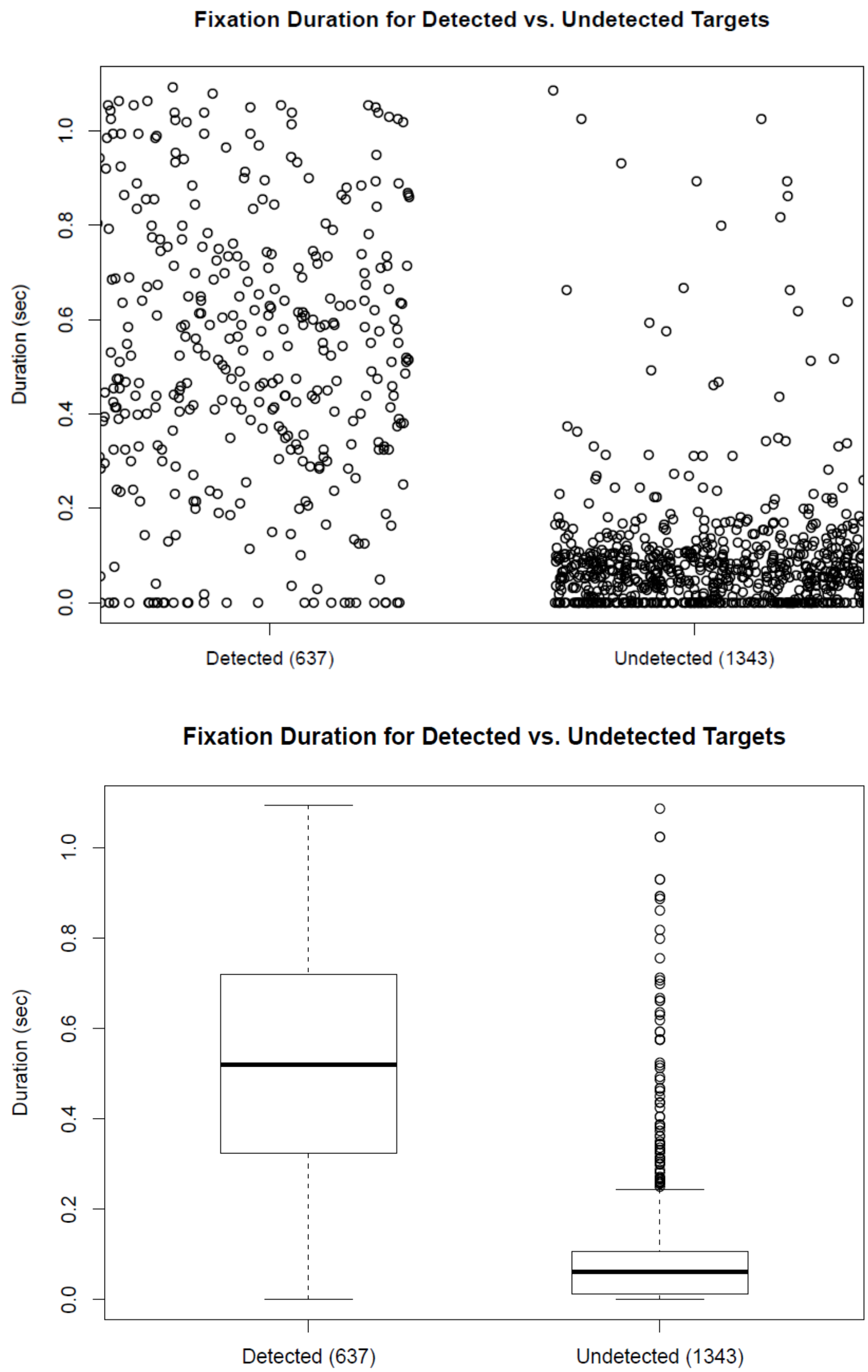


Figure 16: Scatterplot with jitter applied (top) and boxplot (bottom) of fixation duration for detections and non-detections

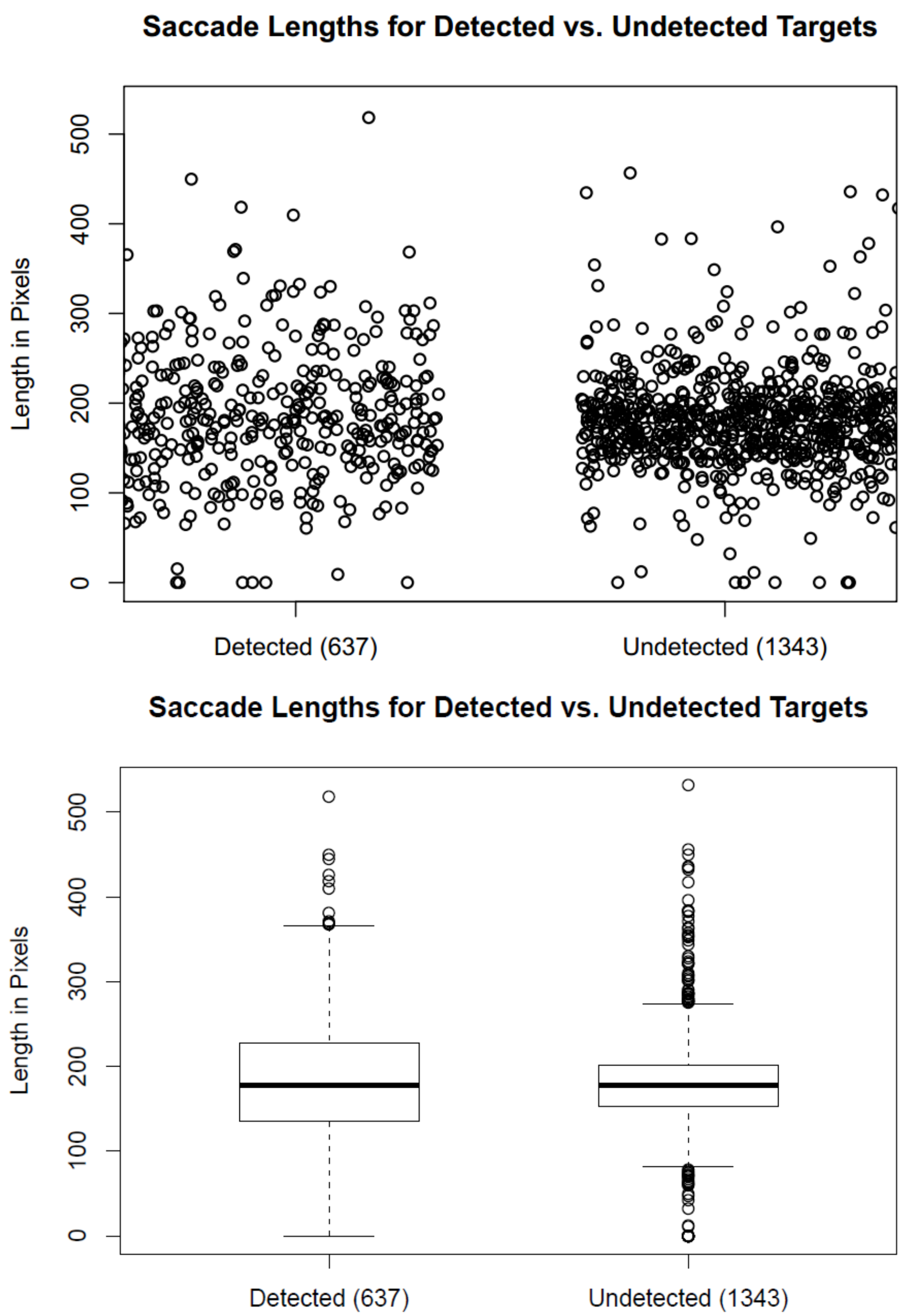


Figure 17: Scatterplot with jitter applied (top) and boxplot (bottom) of mean saccade length for detections and non-detections

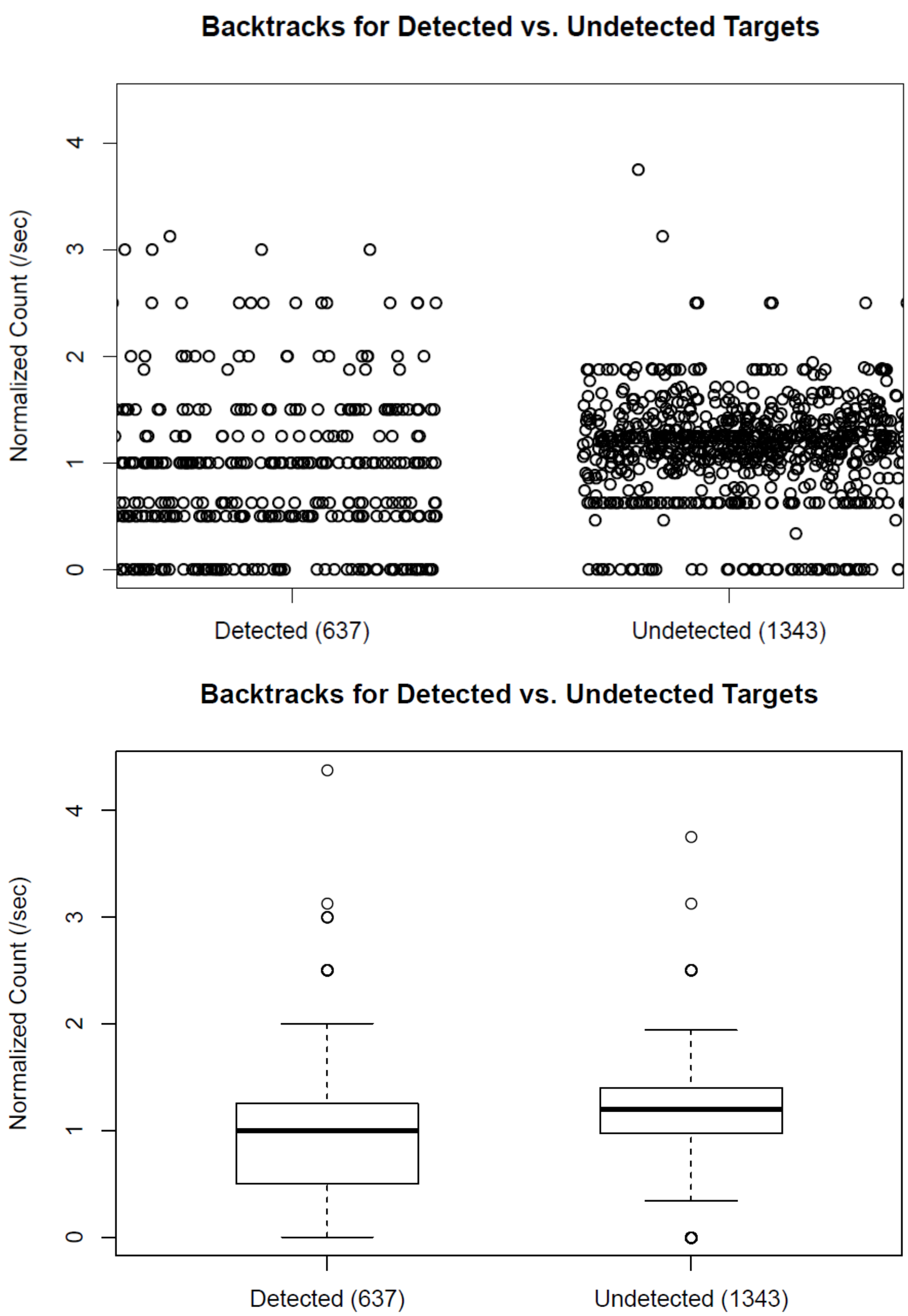


Figure 18: Scatterplot with jitter applied (top) and boxplot (bottom) of backtrack rate for detections and non-detections

A two-tailed unpaired-t test (overall α=0.05, individual α=0.0125) was performed using JMP software for each target type (suits and signals) and each metric (fixation count, fixation duration, mean saccade length, and backtrack rate) to give a total of eight statistical analyses that compared these metrics for detected and non-detected targets. For each of these tests, the mean of the respective metric was calculated for each participant for detected and non-detected targets. One participant had zero signal detections and, therefore, the mean metric value was set equal to zero. These test results are shown in Table 7.

Table 7: Results from Unpaired-t test Statistical Analyses ($\boldsymbol{t_{29,0.0125} = 2.462}$) for Detected and Undetected Targets

Target	Metric	Test Statistic	p-value
Suits	Fixation Count	-28.2896	<0.0001*
	Fixation Duration	-21.1761	<0.0001*
	Mean Saccade Length	-1.29747	0.2047
	Backtrack Rate	1.424373	0.1650
Signals	Fixation Count	-4.22498	0.0002*
	Fixation Duration	-4.52739	<0.0001*
	Mean Saccade Length	1.257681	0.2185
	Backtrack Rate	1.490045	0.1470

4.2.1 Detection Threshold Model

The Detection Threshold Model used to characterize the eye physiology of detections was based on a combination of a prototype model of categorization, which is a type of cognitive model in which the classification of a new target is based on the similarity to each category prototype (Cohen & Basu, 1987; Ashby & Maddox, 1993), and a decision ladder (Rasmussen & Goodstein, 1985). This model is depicted in Figure 19. As shown in the model, possible unconscious detections can be identified by using the metrics shown to be significant through the unpaired-t test analysis. Thresholds in the model were determined based on the statistical analysis of each metric along with the visual separation of detection and non-detection in a scatterplot of the data. The overall classification of a non-detection as an unconscious detection or true non-detection was determined by the number of metrics that were above threshold. If all metrics were above threshold for a given non-detected target, it was classified as an unconscious detection. All targets from the experiment were applied to the Detection Threshold Model, and the results are summarized in the confusion matrix in Table 8.

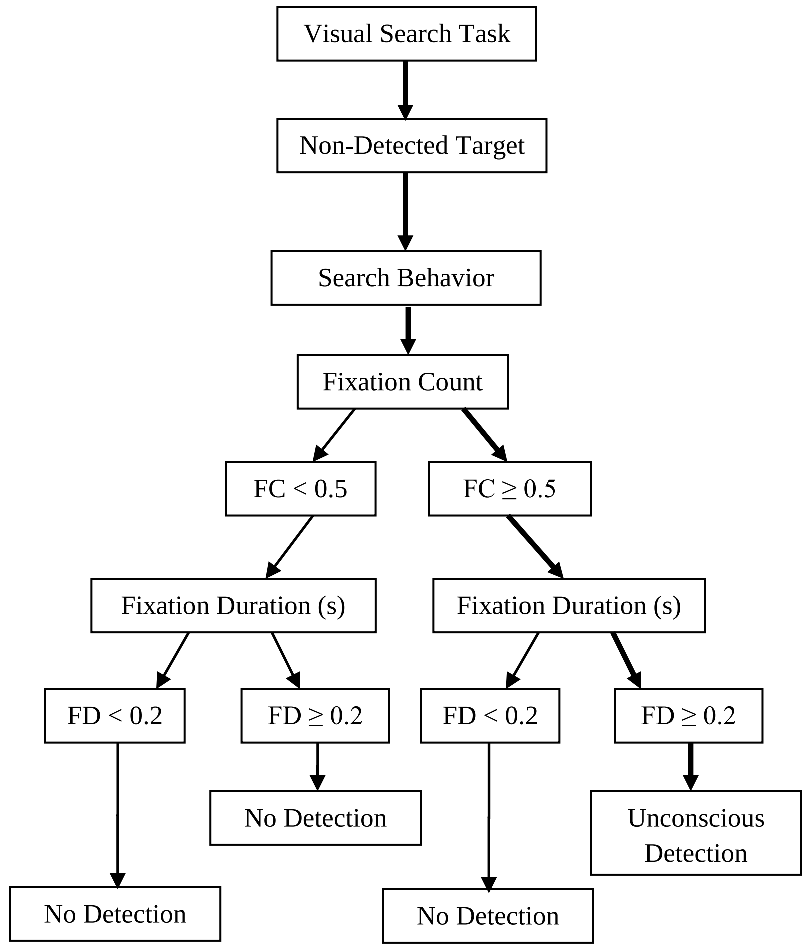


Figure 19: Detection Threshold Model using Fixation Count (FC) and Fixation Duration (FD) as Metrics. Thresholds for the two metrics were determined using scatterplots of the data. A target was classified as an unconscious detection if above threshold for both metrics.

Table 8: Confusion Matrix Results from Detection Threshold Model. Classification, label, and overall accuracies were calculated for the results.

		Predicted		Classification Accuracy
		Detected	**Non-Detected**	
Actual	**Detected**	541	96	84.9%
	Non-Detected	102	1241	92.4%
Label Accuracy		84.1%	92.8%	Overall=90%

4.2.2 Model Validation

In order to validate the classification of data performed by the Detection Threshold Model, a nominal logistic regression was performed on the data using JMP software to compare the metrics of each individual target (N=1980) in order to classify as detected or non-detected. Results revealed that fixation duration, mean saccade length, and backtrack rate were all statistically significant ($p < 0.0001$), and fixation count was marginally significant ($p = 0.0807$) at the 5% confidence level. Therefore, all four metrics were included in the model, and the results are shown in Table 9. Cumulative probability plots were also created for each of the factors in order to examine individual effects on the response. Figure 20 shows these cumulative probability plots for fixation count (top left), fixation duration (top right), mean saccade length (bottom left), and backtrack rate (bottom right).

Table 9: Confusion Matrix Results from Nominal Logistic Regression. Classification, label, and overall accuracies were calculated for the results.

		Predicted		Classification Accuracy
		Detected	**Non-Detected**	
Actual	**Detected**	485	152	76.1%
	Non-Detected	61	1282	95.5%
Label Accuracy		88.8%	89.4%	Overall=89.2%

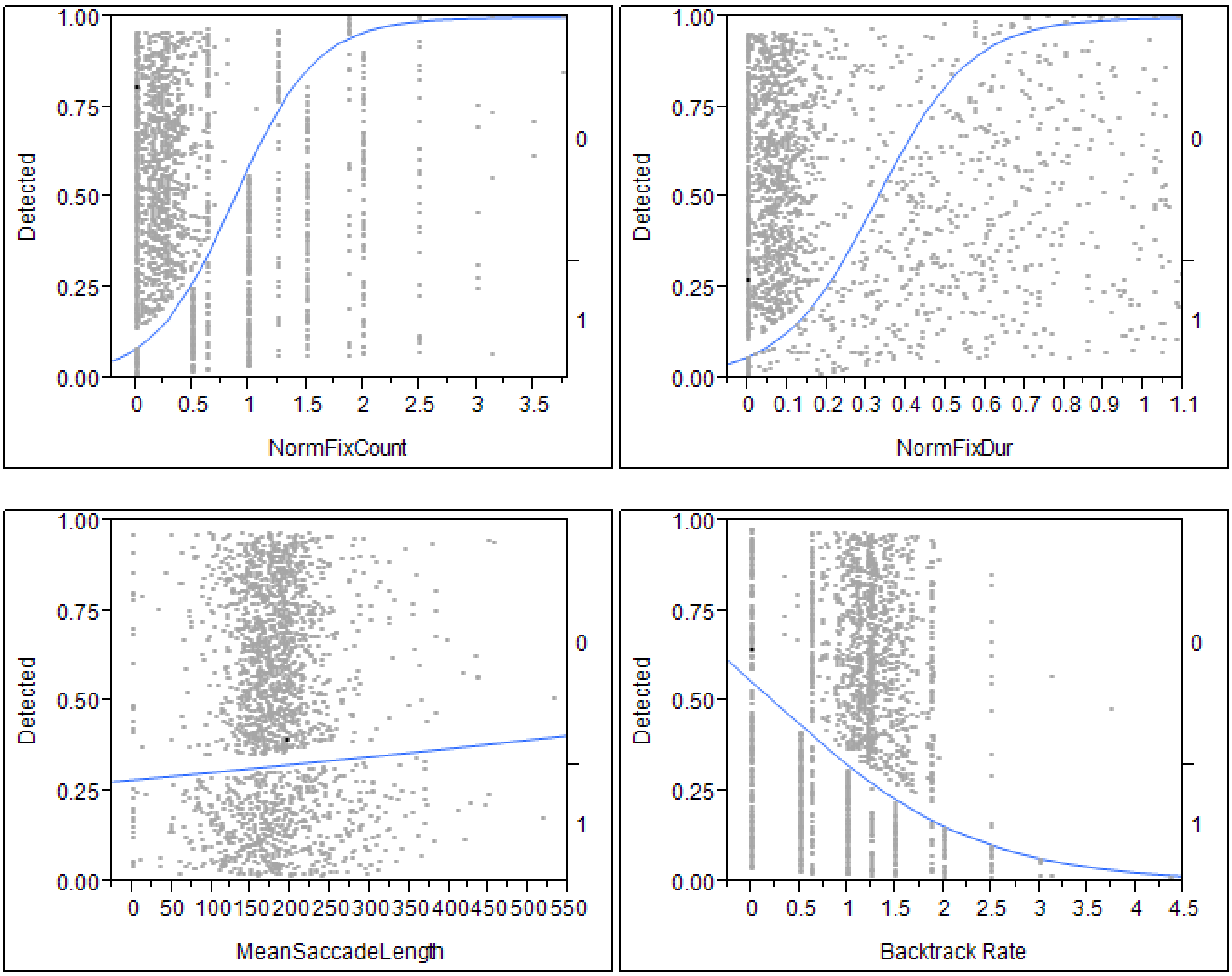


Figure 20: Cumulative Logistic Probability Plots of fixation count (top left), fixation duration (top right), mean saccade length (bottom left), and backtrack rate (bottom right). Fixation count and fixation duration both show a strong correlation with detection and, as the metric increases, the probability of a detection increases. Mean saccade length shows a relatively weak, although positive, correlation with detection. Backtrack rate shows a negative correlation with detection. As backtrack rate increases, the probability of detection decreases.

4.3 Overall Search Pattern Analysis

The Detection Threshold Model and logistic regression analyses classified each target based on individual metric calculations. In addition to this, the overall search strategy of each trial was characterized using gaze plots and the calculation of transition rate in order to determine if these metrics correlated and agreed with the categorization performed by the Detection Threshold Model and logistic regression. In order to calculate transition rate, the experimental images were divided into nine equal subsections. An automated algorithm was designed using Matlab, in which transition rate was equal to the number of times per second that the fixation location moved from one subsection to another (Moacdieh & Sarter, 2015). A linear regression analysis was performed on the mean transition rate of each participant in order to determine its correlation with the mean number of detections for that respective participant. Results from the analysis indicated no statistical significance ($p = 0.4707$).

Gaze plots of all 61 possible unconscious detections that were common between the two models were also created as an additional measure of overall search patterns. For each of these plots, the search pattern was analyzed over the entire thirty-five second trial. Signature search patterns were recognized in some of the plots, and these were further analyzed to determine if these signature patterns occurred at the time onset of the signal. Figures 21 and 22 show gaze plots for possible unconscious detection of easy signals, while figures 23 and 24 show gaze plots for difficult signals. For each plot, only fixation data corresponding to the onset time interval of the target were included for illustration purposes in order to determine any search pattern characteristics or correlates of detection. The fixations are shown as circles, with duration indicated by the circle diameter, and are connected by lines symbolizing the saccades. The position of the possible unconsciously-detected target is indicated by a white circle in each plot. In order to have a standard for comparison of these search patterns, a gaze plot for a conscious detection of an easy signal is depicted in Figure 25.

Figure 21: Gaze Plot of Possible Unconscious Detection of Easy Signal. The search patterns indicate a large saccade towards the signal and increased fixations around the area of the target.

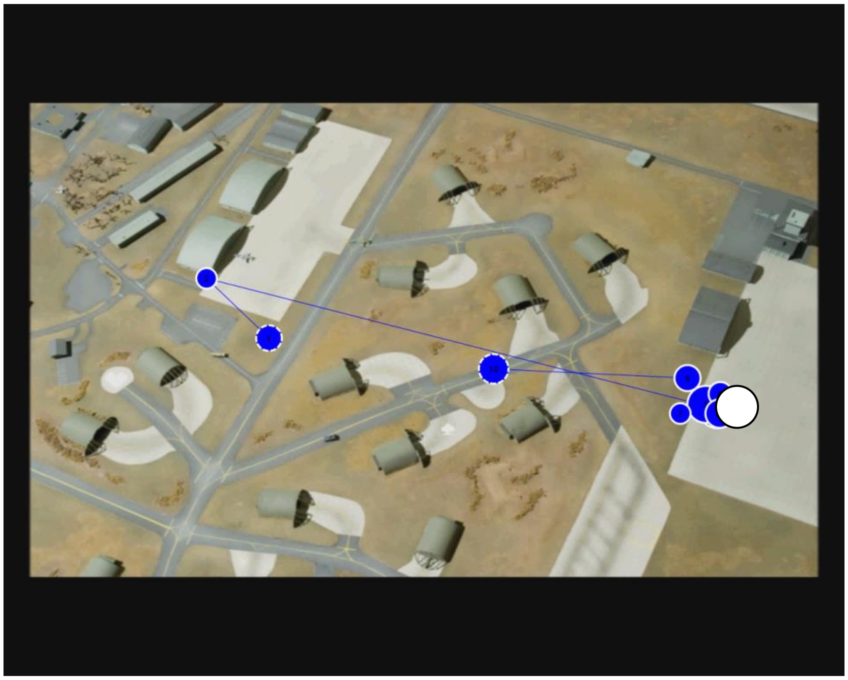


Figure 22: Gaze Plot of Possible Unconscious Detection of Easy Signal. The search patterns indicate a large saccade towards the signal and increased fixations around the area of the target.

Figure 23: Gaze Plot of Possible Unconscious Detection of Difficult Signal. The search patterns indicate a large saccade towards the signal and increased fixations around the area of the target.

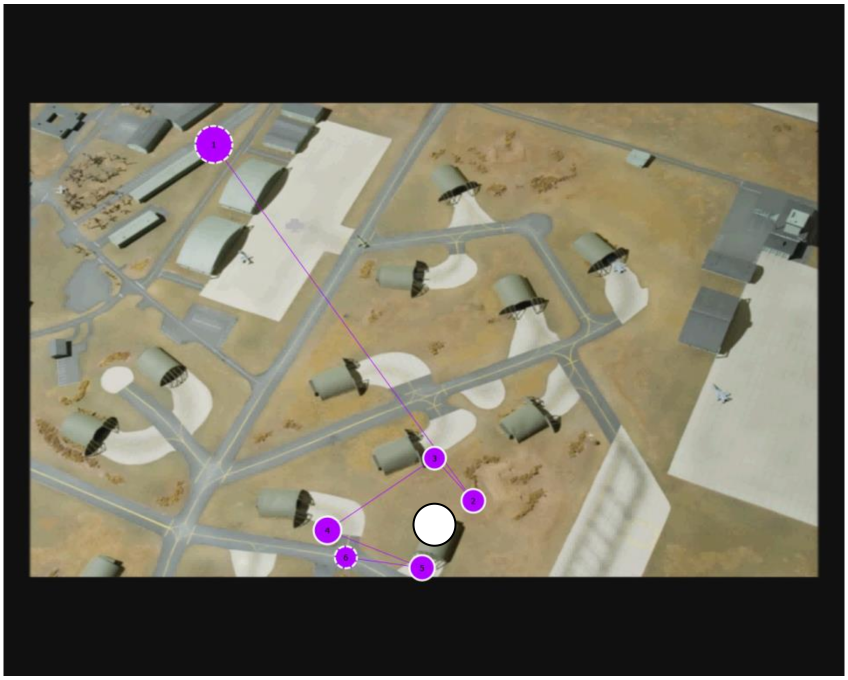


Figure 24: Gaze Plot of Possible Unconscious Detection of Difficult Signal. The search patterns indicate a large saccade towards the signal and increased fixations around the area of the target.

Figure 25: Gaze Plot of Conscious Detection of Easy Signal. The search patterns indicate a large saccade towards the signal and increased fixations around the area of the target.

5.0 Discussion

This experiment used two models, a Detection Threshold Model and a nominal logistic regression model, in order to determine which metrics were correlative of conscious detection, and also possibly unconscious detection. Results from the unpaired-t test statistical analysis revealed that fixation count and fixation duration were the only statistically significant metrics for both suits and signals. Using this information, the Detection Threshold Model was implemented with the inclusion of only these two metrics. The threshold for each of the metrics was determined from the scatterplots and boxplots of the data, in which the thresholds for detection vs. non-detection for normalized fixation count and fixation duration were 0.5 fixations and 0.2 seconds, respectively.

All targets from the experiment were applied to the Detection Threshold Model and, as shown in Table 8, the model correctly classified 84.9% of the detected targets and 92.4% of the non-detected targets. Additionally, 84.1% of labeled detections were correctly predicted and 92.8% of labeled non-detections were correctly predicted. The confusion matrix shows that 102 targets were predicted as “detected” even though the participant did not acknowledge that they saw these. Therefore, each of these 102 targets could potentially be unconscious detections, which could account for the lower label accuracy for detections at 84.1%. Nevertheless, the overall accuracy of the model was 90%, indicating that the model is fairly accurate in differentiating between detections and non-detections in a visual search task.

Out of the 102 possible unconscious detections, there were 33 easy signals, 53 difficult signals, and 16 suits. This indicates that, as expected, the difficult signals provided more opportunities for unconscious detection. Additionally, signals as a whole provided more possibilities for unconscious detection as compared to suits. This could be due to the fact that the suits were stationary, making it easier to consciously recognize once found. However, the signals were only on screen for a short period of time, providing a shorter time window for the unconscious detection to reach conscious awareness.

In order to further validate the Detection Threshold model, results from the logistic regression were analyzed. Fixation duration, mean saccade length, and backtrack rate were all significant, and fixation count was marginally significant, therefore all metrics were included in the model. This model correctly classified 76.1% and 95.5% of detections and non-detections, respectively. Additionally, 88.8% and 89.4% of detections and non-detections, respectively, were labeled correctly. As shown, 61 undetected targets were predicted as "detected", indicating potential unconscious detections. Out of these 61 possible unconscious detections, there were 22 easy, 39 difficult, and no suits. This further justifies the aforementioned conclusions that the harder the target is to see, the higher the possibility of an unconscious detection.

In order to examine the effects of each factor on the detection response for the logistic regression, cumulative probability plots were constructed for each of the factors. As shown in the top left plot of Figure 20, the probability of detection (1) increases as fixation count increases. At a normalized fixation count of 3.75 fixations, almost all of the probability is attributed to detection. The top right plot indicates similar results for fixation duration, and the probability for a detection increases as fixation duration increases. The cumulative probability plot for mean saccade length in the bottom left of

Figure 20 indicates that this factor has a minimal effect on the detection response, as shown by the small slope of the curve. However, there is a slight increase in the probability of detection as the mean saccade length increases. The bottom right plot in Figure 20 indicates that the probability of a detection slightly decreases as backtrack rate increases. At a backtrack rate of zero per second, the probability of a detection is about 65%, however, at a backtrack rate of 4.5 per second, almost all of the probability is attributed to a non-detection. These cumulative probability plots for each individual metric seem to agree with the unpaired-t tests performed for the Detection Threshold Model. Fixation count and fixation duration were statistically significant, while mean saccade length and backtrack rate did not seem to be as accurate of predictors for the response. However, when all factors are examined together in the logistic regression model, mean saccade length and backtrack rate are statistically significant. This could possibly be due to interactions between these factors that are only measured with the logistic regression model, and not with the unpaired-t tests or individual metric cumulative probability plots.

As compared to the Detection Threshold Model, the logistic regression model predicted fewer potential unconscious detections. This could be due to the consideration of interactions between metrics in the logistic regression, while the unpaired-t tests examined each metric separately. The inclusion of these possible interactions may have provided for a more refined, restrictive model, leading to fewer predicted unconscious detections. However, when examining the specific targets for the possible unconscious detections, all 61 predicted unconscious detections from the logistic regression model were also included in the unconscious detections of the Detection Threshold Model. This supports the validation of both models, although the logistic regression model seems to be more restrictive.

In order to analyze the overall search patterns of each trial, gaze plots and a transition rate analysis were performed. The transition rate analysis did not indicate any significant relationship between the mean number of detections per trial and the mean transition rate. However, this could be a result of testing the mean transition rate, as compared to individual transition rates per trial.

Examination of the gaze plots revealed signature search patterns correlative of detection. Specifically, search patterns revealed a long saccade followed by increased fixations around the area of the undetected target. These same search patterns were also revealed in gaze plots of detected targets, as shown in Figure 25. The correlation of the search patterns between conscious detections and possible unconscious detections further validates the performance of both models and suggests that these search patterns are indicative of a detected target, whether conscious or unconscious.

6.0 Conclusion

Results from this research revealed that the Detection Threshold Model and the logistic regression model both predicted detections and non-detections with similar overall accuracies of about 90% and 89.2%, respectively, even though the metrics used to create each model differed. Individually, the metrics of fixation count and fixation duration were strong indicators of whether a target was detected or not detected. However, mean saccade length and backtrack rate also showed significance when analyzed using nominal logistic regression methods, which could indicate that there are significant interactions among these metrics. Additionally, all possible unconscious detections predicated by the logistic regression model were also predicted by the Detection Threshold Model, further validating both models. With the incorporation of interaction effects in the nominal logistic regression, this model indicated to be more restrictive with regards to the classification of potential unconscious detections.

These results were further strengthened by analyzing gaze plots of the predicted unconscious detections that were in common between the two models. These gaze plots indicated signature search patterns of detection, consisting of a long saccade towards the target followed by multiple fixations around the location of the target. These search patterns could potentially be used to identify future possible unconscious detections in a visual search task. Additionally, search patterns were shown to change during the onset of a target, as compared to the search patterns used for the remainder of the trial as stated in the post-test questionnaire. These sudden changes in overall search patterns could further validate the identification of an unconscious detection.

7.0 Future Work

The results of this research provided strong evidence for a reliable model of detection, through which unconscious detections could be identified. In order to strengthen these findings, further work could be performed with a larger sample size in order to provide a further validation of the results. Additionally, research into other possible correlative metrics of unconscious detection could strengthen the Detection Threshold Model and provide for a more in-depth understanding of the cognitive processes involved with visual detection. The Detection Threshold model could also benefit from the inclusion of baseline data in which no targets are present. This data could further validate the model by providing a comparison of true non-detections with the possible unconscious detections in this study. Using the information gathered in this study, further research could be aimed towards developing a real-time analysis and warning of possible unconscious detections. This development could have the potential of directly impacting the ISR domain by improving situational awareness, decreasing missed anomalies, and reducing time and costs.

8.0 Appendix

8.1 Pre-Test Questionnaire

1) What is your gender?
   □ Male
   □ Female
   □ Decline to answer

2) What is your age (in years):

   ______ years old

   □ Decline to answer

3) Please select your Ethnicity origin (or Race):
   □ White
   □ Hispanic or Latino
   □ Black or African American
   □ Native American or American Indian
   □ Asian / Pacific Islander
   □ Other: __________________
   □ Decline to answer

4) What is your primary language?
   □ English
   □ Spanish
   □ Arabic
   □ Other: ________________
   □ Decline to answer

5) Do you have normal or corrected-to-normal vision?
   □ Normal
   □ Contact lenses
   □ Glasses
   □ No, I do not have normal or corrected-to-normal vision.
   □ Decline to answer

6) On average, how often do you play computer and/or video games?

   ____ hours/week

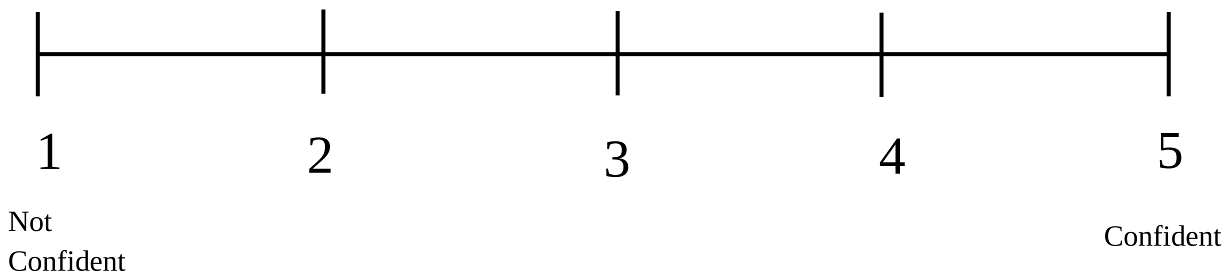
8.2 Post-Clip Questionnaire

Which quadrant of the computer monitor did you see the **“blink”** or **“blinks”**, if any at all? (Mark an “x” in the appropriate quadrant below for every “blink” you saw, if any)

Use the 1-5 scale below to rate your confidence in seeing each **blink**. (Write the rating # next to the “x” you made in the quadrant)

If you did not see any **blinks**, please write “None”.

1-5 Scale:



Quadrants:

II	I
III	IV

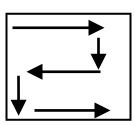
8.3 Post-Test Questionnaire

1) Did you notice a difference between the **blinks** across all images? If yes, what was the difference? If no, what did all the **blinks** look like?

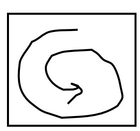
2) What strategy did you use in looking for **targets**?

3) About how far into the 12 trials did you notice the first **blink**? Did this influence your search strategy for the remaining trials? Did you try to locate the **blinks** along with / instead of the assigned **targets**? Please explain in detail.

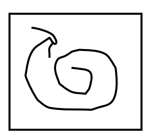
4) Which strategy below do you think is closest to the one you used? Did this change? Why?



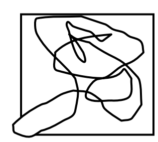
Horizontally



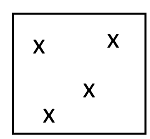
Spiral
Inward



Spiral
Outward



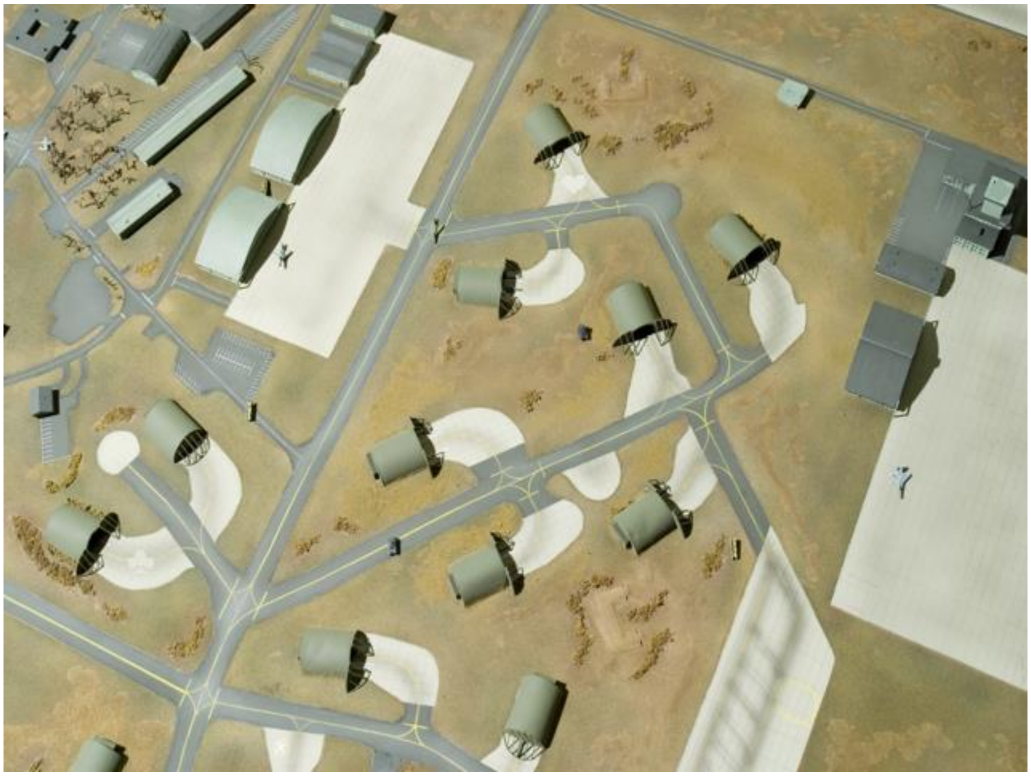
Random

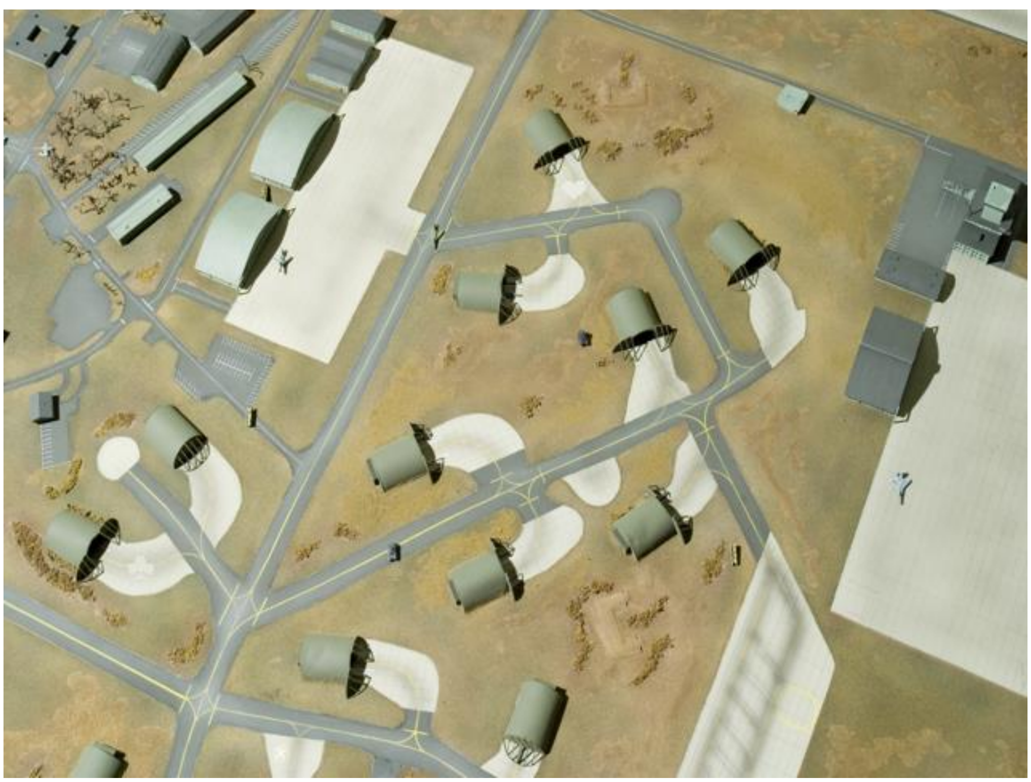


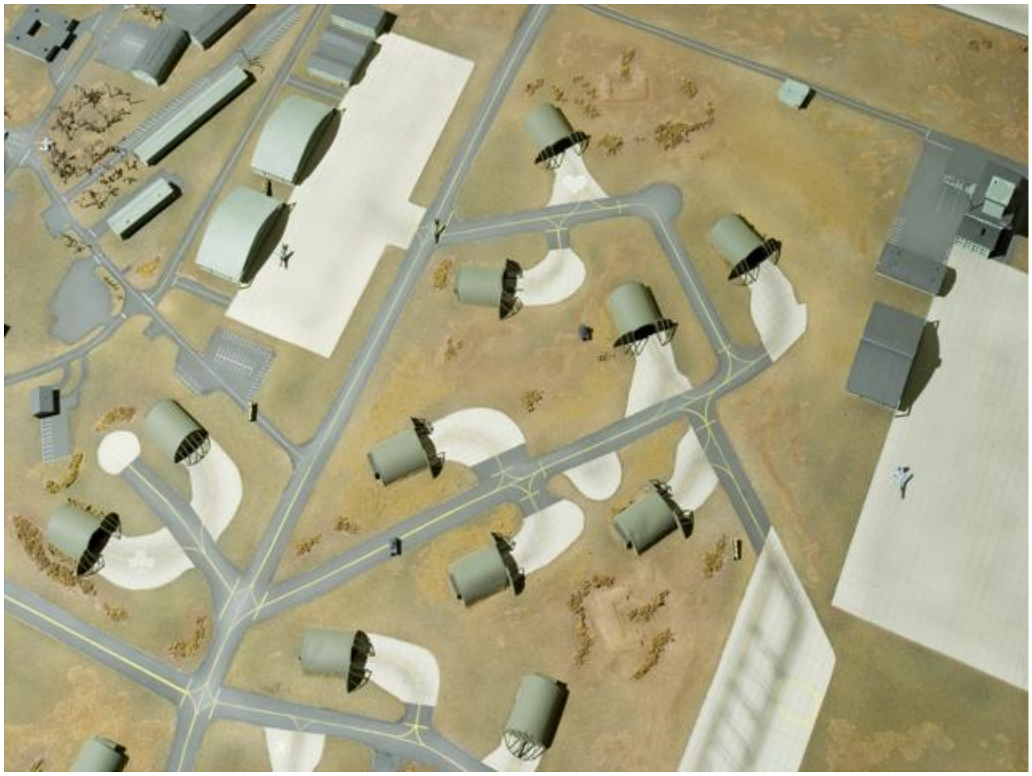
Object
Specific

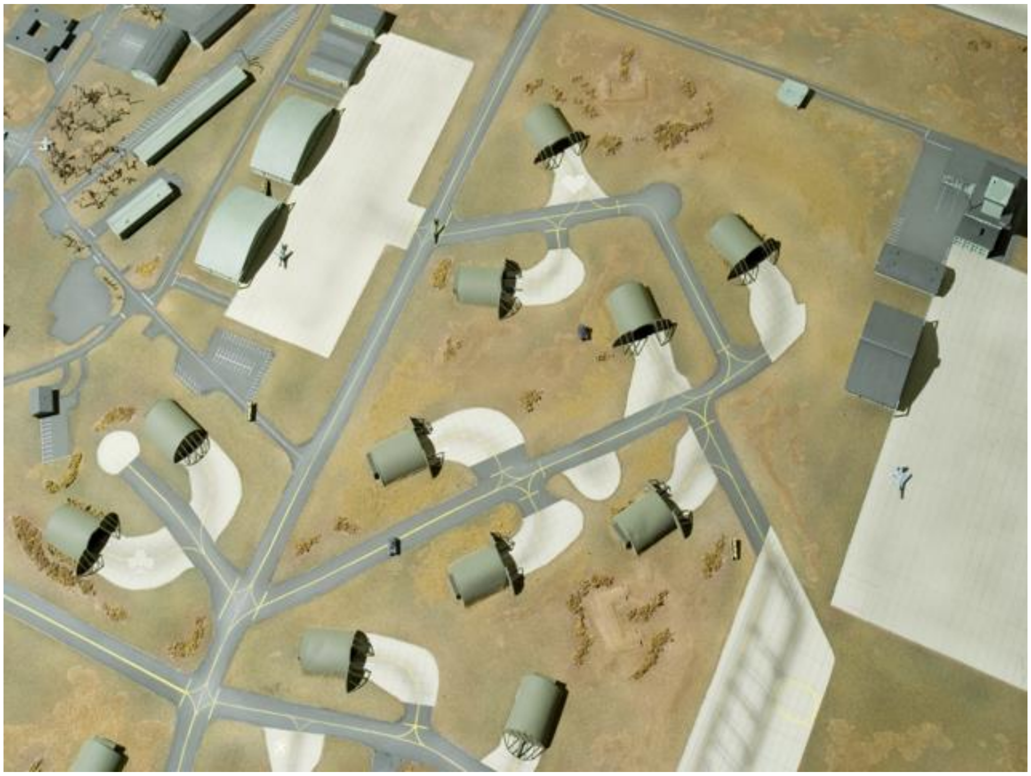
5) If your search strategy changed, which one do you think worked better? Why?

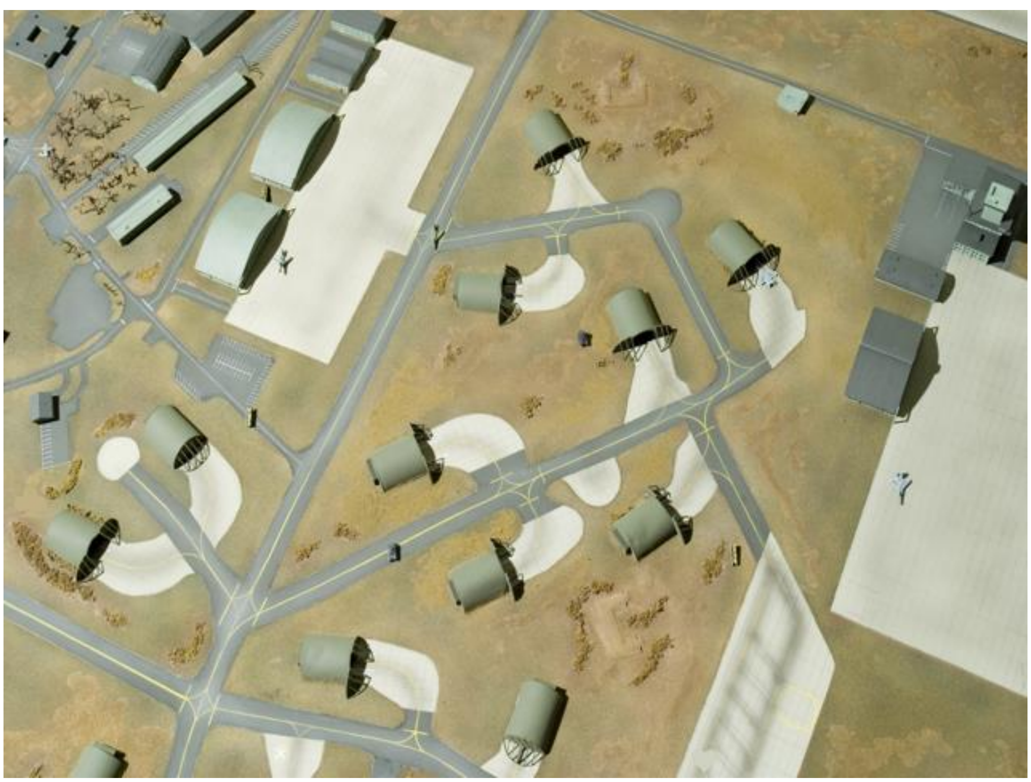
8.4 Stimulus set for twelve trials of experiment











8.5 Nominal Logistic Regression Analysis

Nominal Logistic Fit for Detected

Converged in Gradient, 6 iteration

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Differenc	593.7315	4	1187.463	<.0001 *
Full	650.0195			
Reduced	1243.7509			

RSquare (U)	0.4774
AICc	1310.07
BIC	1337.99
Observations (or Sum Wgts	1980

Measure	Training	Definition
Entropy RSquare	0.4774	1-Loglike(model)/Loglike(0)
Generalized RSquar	0.6306	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.3283	∑ -Log(ρ[j])/n
RMSE	0.2983	√ ∑(y[j]-ρ[j])²/n
Mean Abs Dev	0.1859	∑ |y[j]-ρ[j]|/n
Misclassification Rat	0.1076	∑ (ρ[j]≠ρMax)/n
N	1980	n

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fi	1938	642.74306	1285.486
Saturated	1942	7.27641	**Prob>ChiSq**
Fitted	4	650.01947	1.0000

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-3.0583622	0.2798963	119.39	<.0001 *
NormFixCount	0.38226918	0.2188799	3.05	0.0807
NormFixDur	7.40164464	0.6441938	132.02	<.0001 *
MeanSaccadeLengt	0.00557979	0.0011565	23.28	<.0001 *
Backtrack Rate	-0.7059785	0.1349696	27.36	<.0001 *

For log odds of 1/

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
NormFixCount	1	1	2.98561888	0.0840
NormFixDur	1	1	234.797764	<.0001 *
MeanSaccadeLengt	1	1	23.6274976	<.0001 *
Backtrack Rate	1	1	27.5383237	<.0001 *

Nominal Logistic Fit for Detected

Converged in Gradient, 6 iteration

Odds Ratios

For Detected odds of 1 versus

Tests and confidence intervals on odds ratios are likelihood ratio based.

Unit Odds Ratios

Per unit change in regresso

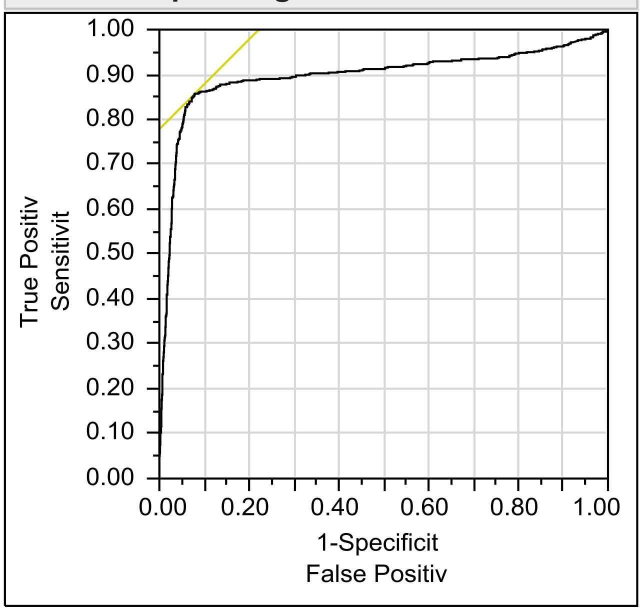
Term	Odds Ratio	Lower 95%	Upper 95%	Reciprocal
NormFixCount	1.465607	0.949307	2.241243	0.6823114
NormFixDur	1638.677	488.1151	6106.624	0.0006102
MeanSaccadeLengt	1.005595	1.003329	1.007894	0.9944358
Backtrack Rate	0.493625	0.37849	0.642691	2.025828

Range Odds Ratios

Per change in regressor over entire rang

Term	Odds Ratio	Lower 95%	Upper 95%	Reciprocal
NormFixCount	4.19339	0.822765	20.62212	0.2384706
NormFixDur	3279.836	872.1099	13826.75	0.0003049
MeanSaccadeLengt	19.45687	5.860136	65.55249	0.0513957
Backtrack Rate	0.045563	0.014256	0.144547	21.947562

Receiver Operating Characteristic



Using Detected='1' to be the positive leve

AUC
0.89725

Confusion Matrix

Actua		Predicte
Training	**1**	**0**
1	485	152
0	61	1282

8.6 Matlab Pseudo Code

8.6.1 Mean Saccade Length Calculation

Identify all fixation data points within AOI time interval of respective trial and participant

Calculate saccade length between all sequential fixation points (e.g. calculation of saccade length for fixation points 1 and 2 below):

$$Saccade\ Length(1,2) = \sqrt{(Fixation.X1 - Fixation.X2)^2 + (Fixation.Y1 - Fixation.Y2)^2}$$

Calculate mean saccade length

8.6.2 Backtrack Rate Calculation

Identify all fixation data points within AOI time interval of respective trial and participant

Calculate saccade length between all groups of three sequential fixation points (e.g. calculation of saccade length for fixation points 1 and 2, 2 and 3, & 1 and 3 below):

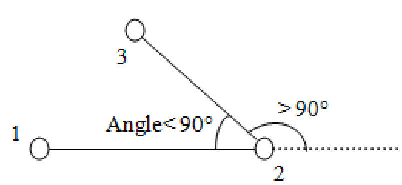
$$Saccade\ Length(1,2) = \sqrt{(Fixation.X1 - Fixation.X2)^2 + (Fixation.Y1 - Fixation.Y2)^2}$$

$$Saccade\ Length(2,3) = \sqrt{(Fixation.X2 - Fixation.X3)^2 + (Fixation.Y2 - Fixation.Y3)^2}$$

$$Saccade\ Length(1,3) = \sqrt{(Fixation.X1 - Fixation.X3)^2 + (Fixation.Y1 - Fixation.Y3)^2}$$

Calculate angle (using the Law of Cosines) between all sequential saccades (e.g. calculation of angle between saccade length (1, 2) and saccade length (2, 3) below):

$$Angle = cos^{-1}\frac{Saccade\ Length\ (1,2)^2 + Saccade\ Length\ (2,3)^2 - Saccade\ Length\ (1,3)^2}{2 * Saccade\ Length\ (1,2) * Saccade\ Length\ (2,3)}$$



If angle is less than 90 degrees (meaning, by definition, angle from first saccade to second saccade is greater than 90 degrees)

Add one backtrack to total number of backtracks for that AOI

Else

No Backtrack

Calculate backtrack rate by dividing by AOI time interval

8.6.3 Transition Rate Calculation

Identify all fixation data points within trial for respective participant

Define boundaries (in pixels for 1280 x 1024 pixels screen size) for nine equal subsections of image

Calculate which subsection each fixation is located in using boundaries

Determine number of times subsection changes from one fixation to the next in order to calculate number of transitions

Calculate transition rate by dividing number of transitions by 35 s trial

7.0 References

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1), 39-59.

Açık, A., Bartel, A., & König, P. (2014). Real and implied motion at the center of gaze. *Journal of Vision*, 14(1), 1-19.

Ashby, F.G., & Maddox, W.T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology,* 37(3), 372-400.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839.

Barber, C. J. (2001). An intelligence, surveillance and reconnaissance (ISR) vision for the Canadian forces. *Canadian Military Journal*, 2(4), 2001-2002.

Betsch, T. (2008). The nature of intuition and its neglect in research on judgment and decision making. *Intuition in Judgment and Decision Making*, 3-22.

Betsch, T., & Glöckner, A. (2010). Intuition in judgment and decision making: Extensive thinking without effort. *Psychological Inquiry*, 21(4), 279-294.

Carmi, R., & Itti, L. (2006). Visual causes versus correlates of attentional selection in dynamic scenes. *Vision Research*, 46(26), 4333-4345.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, 41(3), 1-58.

Chen, Y.C. & Yeh, S.L. (2012). Look into my eyes and I will see you: Unconscious processing of human gaze. *Consciousness and Cognition.* 21, 1703-1710.

Chizek, J. G. (2003). Military transformation: intelligence, surveillance and reconnaissance. Library of Congress, Washington, D.C. Congressional Research Service.

Cohen, J.B., & Basu, K. (1987). Alternative models of categorization: toward a contingent processing framework. *Journal of Consumer Research*, 455-472.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, 17(5), 428-433.

Drew, T.,Vo, M. L., & Wolfe, J. M. (2013). The invisible gorilla strikes again: Sustained inattentional blindness in expert observers. *Psychological Science,* 24(9), 1848-1853.

Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance,* 31(6), 1416-1438.

Duvall, L. M. (2005). Accounting for Human Information Processing: A Critical Element for Future United States Air Force Command and Control Systems. Air University, Maxwell Air Force Base, Alabama.

Evans, J. S. B. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454-459.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255-278.

Fendley, M. & Narayanan, S. (2012). Decision aiding to overcome biases in object identification. *Advances in Human-Computer Interaction, 2012*, 7.

Galpin, A., Underwood, G., & Chapman, P. (2008). Sensing without seeing in comparative visual search. *Consciousness and Cognition, 17*(3), 672-687.

Gawronski, B. & Creighton, L. A. (2013). Dual-process theories. *The Oxford Handbook of Social Cognition*, 282-312.

Goldberg, J.H. & Kotval, X.P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.

Gu, E., Stocker, C., & Badler, N. I. (2005). Do you see what eyes see? Implementing inattentional blindness. Paper presented at the *Intelligent Virtual Agents*, 178-190.

Hamel, S., Houzet, D., Pellerin, D., & Guyader, N. (2015). Does color influence eye movements while exploring videos? *Journal of Eye Movement Research,* 8(1):4, 1-10.

Heuer, R. J. (1999). *Psychology of intelligence analysis.* Washington, D.C.: Center for the Study of Intelligence.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 113-136.

Horstmann, N., Ahlgrimm, A., & Glöckner, A. (2009). How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes. *Judgement and Decision Making,* 4, 335-354.

Jacob, M., & Hochstein, S. (2009). Comparing eye movements to detected vs. undetected target stimuli in an Identity Search Task. *Journal of Vision*, 9(5), 1-16.

Klein, G., & Klinger, D. (2008). Naturalistic decision making. *Human Factors,* 50(3), 456-460.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide. 2nd Edition.* New York, NY: Cambridge University Press.

March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics*, 9, 587-608.

Maule, A. J. (1997). Strategies for adapting to time pressure. In R. Flin, E. Salas, M. Strub, & L. Martin, *Decision making under stress: Emerging themes and applications*, 271-279. Ashgate: Aldershot.

Moacdieh, N., & Sarter, N. (2015). Clutter in electronic medical records: examining its performance and attentional costs using eye tracking. *Human Factors and Ergonomics Society,* 57(4), 591-606.

Muller, A. C., & Narayanan, S. (2009). Cognitively-engineered multisensor image fusion for military applications. *Information Fusion*, 10(2), 137-149.

Rasmussen, J. & Goodstein, L.P. (1985). *Decision support in supervisory control* (Riso Rep. M-2525). Roskilde, Denmark: Riso National Laboratory.

Ratcliff, R. & Smith, P.L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Annual Review of Psychology,* 111(2), 333-367.

Rensink, R. A. (2004). Visual sensing without seeing. *Psychological Science,* 15(1), 27-32.

Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review,* 31(2), 180-206.

Rich, A., & Gillam, B. (2000). Failure to detect changes in color for lines rotating in depth: The effects of grouping and type of color change. *Vision Research,* 40(10), 1377-1384.

Rothkirch, M., Stein, T., Sekutowicz, M., & Sterzer, P. (2012). A direct oculomotor correlate of unconscious visual processing. *Current Biology,* 22(13), R514-R515.

Rummukainen, O., Radun, J., Virtanen, T., & Pulkki, V. (2014). Categorization of natural dynamic audiovisual scenes. *PloS One*, 9(5), e95848.

Simon, H. A. (1957). Models of man: social and rational. New York: John Wiley and Sons, Inc.

Simons, D. J. & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception.* 28, 1059-1074.

Simons, D. J., Chabris, C. F., Schnur, T., & Levin, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness and Cognition*, 11(1), 78-97.

Spering, M., & Carrasco, M. (2015). Acting without seeing: Eye movements reveal visual processing without awareness. *Trends in Neurosciences*, 38(4), 247-258.

Spering, M., Pomplun, M., & Carrasco, M. (2011). Tracking without perceiving: A dissociation between eye movements and motion perception. *Psychological Science*, 22(2), 216-225.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1), 137-149.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, 8(8), 1096-1101.

Underwood, G., Templeman, E., Lamming, L., & Foulsham, T. (2008). Is attention necessary for object identification? Evidence from eye movements during the inspection of real-world scenes. *Consciousness and Cognition*, 17, 159-170.

Vaportzis, E., Georgiou-Karistianis, N., Churchyard, A., & Stout, J.C. (2015). Dual task performance in Huntington's disease: a comparison of choice reaction time tasks. *Neurophyschology*, 29(5), 703-712.

Warren, R., Smith, R. F., & Cybenko, A. K. (2011). *Use of Mahalanobis Distance for Detecting Outliers and Outlier Clusters in Markedly Non-Normal Data: A Vehicular Traffic Example*. Technical Report, DTIC Document. Wright-Patterson AFB, Dayton, Ohio.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114(1), 152-176.

Wu, C. C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. *Frontiers in Psychology, 5*.