Effects of Augmented Reality Head-up Display Graphics' Perceptual Form on Driver Spatial Knowledge Acquisition

NAYARA DE OLIVEIRA FARIA

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial and Systems Engineering

Joseph L. Gabbard, Chair
Sheila G. Klauer
Martha Irene Smith

December 16, 2019
Blacksburg, VA

Keywords: spatial knowledge; augmented reality; driving; head-up display

Effects of Augmented Reality Head-up Display Graphics' Perceptual Form on Driver Spatial Knowledge Acquisition

NAYARA DE OLIVEIRA FARIA

ABSTRACT

In this study, we investigated whether modifying augmented reality head-up display (AR HUD) graphics' perceptual form influences spatial learning of the environment. We employed a 2x2 between-subjects design in which twenty-four participants were counterbalanced by gender. We used a fixed base, medium-fidelity driving simulator at the COGENT lab at Virginia Tech. Two different navigation cues systems were compared: world-relative and screen-relative. The world-relative condition placed an artificial post sign at the corner of an approaching intersection containing a real landmark. The screen-relative condition displayed turn directions using a screen-fixed traditional arrow located directly ahead of the participant on the right or left side on the HUD. We captured empirical data regarding changes in driving behaviors, glance behaviors, spatial knowledge acquisition (measured in terms of landmark and route knowledge), reported workload, and usability of the interface.

Results showed that both screen-relative and world-relative AR head-up display interfaces have similar impact on the levels of spatial knowledge acquired; suggesting that world-relative AR graphics may be used for navigation with no comparative reduction in spatial knowledge acquisition. Even though our initial assumption that the conformal AR HUD interface would draw drivers' attention to a specific part of the display was correct, this type of interface was not helpful to increase spatial knowledge acquisition. This finding contrasts a common perspective in the AR community that conformal, world-relative graphics are inherently more effective than screen-relative graphics. We suggest that simple, screen-fixed designs may indeed be effective in certain contexts.

Finally, eye-tracking analyses showed fundamental differences in the way participants visually interacted with different AR HUD interfaces; with conformal-graphics demanding more visual attention from drivers. We showed that the distribution of visual attention allocation was that the world-relative condition was typically associated with fewer glances in total, but glances of longer duration.

Effects of Augmented Reality Head-up Display Graphics' Perceptual Form on Driver Spatial Knowledge Acquisition

NAYARA DE OLIVEIRA FARIA

GENERAL AUDIENCE ABSTRACT

As humans, we develop mental representations of our surroundings as we move through and learn about our environment. When navigating via car, developing robust mental representations (spatial knowledge) of the environment is crucial in situations where technology fails, or we need to find locations not included in a navigation system's database. Over-reliance on traditional in-vehicle navigation devices has been shown to negatively impact our ability to navigate based on our own internal knowledge. Recently, the automotive industry has been developing new in-vehicle devices that have the potential to promote more active navigation and potentially enhance spatial knowledge acquisition. Vehicles with augmented reality (AR) graphics delivered via head-up displays (HUDs) present navigation information directly within drivers' forward field of view, allowing drivers to gather information needed without looking away from the road. While this AR navigation technology is promising, the nuances of interface design and its impacts on drivers must be further understood before AR can be widely and safely incorporated into vehicles. In this work, we present a user study that examines how screen-relative and world-relative AR HUD interface designs affect drivers' spatial knowledge acquisition.

Results showed that both screen-relative and world-relative AR head-up display interfaces have similar impact on the levels of spatial knowledge acquired; suggesting that world-relative AR graphics may be used for navigation with no comparative reduction in spatial knowledge acquisition. However, eye-tracking analyses showed fundamental differences in the way participants visually interacted with different AR HUD interfaces; with conformal-graphics demanding more visual attention from drivers

DEDICATION

To the Almighty God for giving me the miracle of life and for opening my path to opportunities that allowed me to be here today. Thank you for your guidance, protection and strength in all the moments I felt alone far from my beloved ones.

To my mother Maria Aparecida de Oliveira de Faria and to my father Donizete Aparecido Ribeiro de Faria for their unconditional love and patience. You are the best parents one could ever ask for.

To my brother Daniel de Oliveira Faria and to my sister Poliana Aparecida de Oliveira Faria for always being present in my life.

To Malu Faria for being the little human being that brings joy to all in our family. I am also grateful for Scooby, Sheik and my first little one, Bidu Faria.

To Jadiel Hernandez Rivera for being my ideal partner during my graduate school journey. I really love, respect, and appreciate you. Thank you for your unconditional support throughout my graduate life's challenges and for always being here when I need you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

1. INTRODUCTION

1.1 MOTIVATION

Imagine getting invited to a friend's house after work. You have never been to his house but have his address saved on your phone, which is equipped with Google Maps. You get in your car and confidently follow the suggested route until Google Maps says you reach the destination. When you arrive, you realize that you had the wrong address and your phone battery is now dead. On top of that, you are in a quiet area where this is no one near you can ask for directions. Would you be able to recall your path back to work? In other words, have you acquired enough levels of knowledge of your surroundings (spatial knowledge) so that you can go back to work?

Finding our way through new driving environments has considerably changed in the past decades. We have evolved from navigating by using a collection of fold-out paper maps to using in-vehicle satellite navigation systems. These satellite navigation systems (commonly termed *SatNav*), became increasingly popular and are currently being complemented by highly accurate and real-time smartphone applications such as Google Maps and Waze. SatNav use changes the way we interact with the environment around us. By following simple turn-by-turn navigation instructions, we are able to offload both tactical aspects (route following) and strategic aspects (route planning) of the navigation task to the navigation device. This method of finding our way reduces the requirements necessary to focus on aspects of the environment that are relevant to the navigation task (e.g., locating and identifying key landmarks, developing an internal spatial map of the environment, etc.), and as a result, drivers' cognitive and attentional resources required for navigation are reduced and thus, can be reallocated to perform other tasks unrelated to navigation [1].

Regardless of the potential opportunities that SatNavs provide, their use can also be problematic. For example, research suggests that drivers using SatNav do not develop as much robust environmental spatial knowledge as drivers using paper maps [2, 3]. In other studies, the use of SatNav resulted in cognitive underload, which can be viewed as a strength in some cases (as drivers can perform other tasks unrelated to navigation) but has also been found to result in decreased driver engagement and situation awareness [1]. These effects were also noted by Webber [4] who found that participants using SatNav devices were not only less aware of their environment or surroundings, but also less engaged with the navigation task itself (i.e., passive navigation), and as a result, acquired lower levels of spatial knowledge.

Augmented reality (AR) head-up displays (HUDs) provide new opportunities to display navigation information directly on the windshield within drivers' forward field of view, allowing drivers to gather information needed to navigate without looking away from the road [5] . While AR HUDs are promising, the nuances of visual interface design and its impacts on drivers must be further understood before AR can be widely and safely incorporated into vehicles. Specifically, an impact that warrants investigation is the role of AR HUDs in spatial knowledge acquisition while driving.

1.2 RESEARCH PROBLEM

To date, the vast majority of driving studies investigating spatial knowledge acquisition have employed head-down displays, typically located in the center of the vehicle dashboard. However, AR HUDs allow designers to overlay navigation information directly onto real-world landmarks, thus allowing drivers to potentially acquire greater spatial knowledge as compared to head-down display navigation systems [6].

A very common user interface design approach in the augmented reality community is to incorporate information and 3D graphics directly into the environment, such as placing a virtual couch in your living room, or virtual control widgets on the wall. AR graphics which are geo-referenced, and perceptually placed at fixed locations in the world we term *conformal*, or *world-relative* graphics. Conformal graphics are often assumed to be inherently superior to 2D, screen-relative graphics for navigation due to the natural perceptual form of the conformal graphic and its perceived (and geo-relevant) location in the world. For example, a virtual AR arrow "painted" on the forward roadway can be considered a naturally intuitive design approach since road markings are already done this way. And while one may intuit this to be true, it is not clear whether the perceptual form of the AR graphic has an impact specifically on spatial knowledge acquisition. To the best of our knowledge, no research has yet been done examining the effectiveness of using 3D AR graphics presented via AR HUD to convey driving-related spatial knowledge, where the AR graphics are visually perceived to be spatially integrated into the driving scene (i.e., next to, or part of actual real-world landmarks). Additionally, most current HUD technologies have limited field of view (6-15 degrees), and a single fixed focal plane. Even though the human foveal vision makes up about 1° of the visual field, we depend heavily on peripheral vision for visual driving tasks, and thus, a limited AR HUD FOV can make driving more difficult, because important road elements and environmental landmarks might lie outside the AR interface (which makes presenting conformal graphics in these locations impossible without improvements in AR HUD technologies).

Given these points, it is important to further delve into the extent to which AR HUD interfaces can be leveraged for navigation without demanding unfeasible requirements from vehicle manufacturers. Specifically, we want to understand whether providing conformal AR navigational cues improves spatial knowledge acquisition to the extent that investment in generating larger FOV AR HUDs with potentially multiple focal planes is justifiable.

1.3 RESEARCH OBJECTIVES

1.3.1 MAIN OBJECTIVE

The main objective of this work is to understand how screen-relative and world-relative AR HUD perceptual forms affect the acquisition of spatial knowledge while driving. Therefore, the purpose of this thesis is to address the following key question:

- *R1: How does HUD graphics' perceptual form (world- relative vs. screen-relative) impact drivers' acquisition of spatial knowledge?*
  - Hypothesis 1.A (H1A): Participants will develop higher levels of landmark knowledge using world-relative graphics.
  - Hypothesis 1.B (H1B): Participants will develop higher levels of route knowledge using world-relative graphics.

1.3.2 SECONDARY OBJECTIVES

Additionally, we want to understand the following additional questions:

- *R2: How does HUD graphics' perceptual form (world-relative vs. screen-relative) impact driving measures?* We include this question, since when employing AR to support a primary task (such as driving), we must not lose sight of potential negative effects of AR interface design on primary task performance.
  - Hypothesis 2.A (H2A): Participants using world-relative graphics will have more variability on driving behavior measures (i.e. standard deviation of speed and

standard deviation of lane position) compared to participants using screen-relative graphics.

    - Hypothesis 2.B (H2B): Participants using world-relative graphics will have more missed turns compared to those using screen-relative graphics.

- R3: *How does HUD graphics' perceptual form (world-relative vs. screen-relative) impact drivers' glance behaviors?*
    - Hypothesis 3 (H3): Participants using world-relative graphics will glance longer to the HUD navigation cue compared to those using screen-relative graphics.

2. BACKGROUND AND RELATED WORK

2.1 BACKGROUND AND DEFINITIONS

2.1.1 AUGMENTED REALITY

Virtual reality (VR) can be understood as a “computer-generated environment that gives a user the experience of being in a particular location, different from where the user actually is” [7, p. 150]. In other terms, VR is described as the use of a virtual environment (VE), also known as computer-generated 3D environment, in which people can both navigate and interact [8]. Navigation and interaction can be accomplished by using multiple sensory channels, such as visual, auditory and haptic. VR is known to be the technological predecessor of augmented reality technologies. *Augmented reality* (AR) is defined as a display in which computer-generated graphics are superimposed upon real-world physical objects [9]. VR and AR are closely related to each other, but they differ in terms of immersion levels in which digital and physical objects co-exist in an environment [10]. *Immersion* deals with what the technology is able to deliver from an objective point of view. It refers to the level to which the person is isolated from the real world [8]. The environment is said to be more immersive when the system increases both the use of sensory modalities and fidelity with the real world [11]. Immersive environments give users a sense of presence and a sense of being in the virtual environment [12]. In this context, *presence* is simply a reaction to immersion. As humans, we experience different levels of presence given the same immersive system [11].

Augmented Reality (AR) was first introduced in the 1960s with the work of Sutherland [13] entitled “A Head-Mounted Three-Dimensional Display” that used a see-through head-mounted display to present 3D graphics. Two decades after Sutherland’s publication, AR was established as a research field when Azuma [9] published a well-known augmented reality survey [14]. AR can be thought of as the "middle ground" between completely synthetic and completely real environments [9], in which the AR system is composed of three main properties: 1) combines both virtual and real objects in a real environment, 2) real-time interactive, 3) registered in 3D, so that real and virtual objects can be aligned to each other [14].

Milgram & Kishino [15] defined the reality-virtuality continuum that is mostly used to classify the technical nature and system composition of AR and related systems (see Figure 1). In this scheme, while in augmented reality the surrounding environment is real, in both augmented virtuality and virtual reality the surrounding environment is virtual [14].

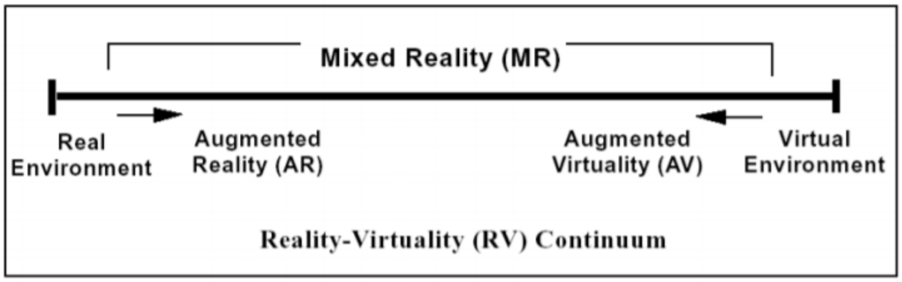


Figure 1- Milgram's Reality-Virtuality Continuum. [15]

According to Wickens et al. [7, p. 155], from an attentional perspective, AR adheres to the proximity compatibility principle (PCP), "facilitating the integration of information between the display and the real world beyond, a display proximity fostered both by co-location (overlay) and the common fate of shared movement when the head rotate". In some AR applications, synthetic imagery is presented through the transparent eye-wear of a head-mounted display (HMD) in which the user concurrently sees the real world and perform a manual task. This is not different from a head-up display in vehicles, except that in this case imagery moves with the head rotation rather than the vehicle rotation [7]. The next section will further discuss the AR head-up display concepts that are applicable to a better understanding of this research.

2.1.2 THE AUGMENTED REALITY HEAD-UP DISPLAY (AR HUD)

Secondary information, such as navigational aids, can be portrayed to drivers in two major display types: head-down displays and head-up displays. A *head-down display* (HDD) is typically located in the center of the dashboard of a vehicle (Figure 2). HDDs deliver information to drivers below their forward-looking line of sight, thus limiting the driver's peripheral vision and ability to concurrently attend to the road and gather information from the display. A *head-up display* (HUD) on the contrary, is a display in which navigation information is displayed directly on the windshield, within drivers' forward field of view (FOV), allowing them to use peripheral vision and to gather information needed to navigate without looking away from the road [5, 16] (see Figure 3).

Figure 2- An example of a head-down display located in the center of the dashboard of a vehicle


Figure 3- An example of a head-up display in a vehicle in which information is displayed within drivers' forward field of view

Smith et al. [17, p. 185] point out two main reasons HUDs might be beneficial to drivers when compared to HDD: (1) visual attention remains in the direction of the road while in use, and (2) display integration with the world can provide more rich information and remove task uncertainty. Additionally, because information is presented directly within the FOV, drivers are no longer required to continuously re-accommodate their focal length between the distant road scene and nearby in-vehicle display [18]. This is important especially for older drivers that usually experience accommodation problems to near objects. Moreover, research have shown that people using HUDs might outperform themselves when using HDDs in other aspects, such as, better performance of vehicle control [19, 20, 17], faster response time to an urgent event [19] and lower levels of mental workload [19, 21].

During the few last years, HUDs have attracted significant interest within the automotive industry and research community. Some car models, such as Audi, Mercedes and BMW already deploy HUDs as part of the vehicle, and each year more manufacturers are offering HUDs as an optional item for comfort and safety. Besides devices already installed (Toyota Prius, Mazda 3, Jaguar XE, Volvo XC90, Lexus RX, Hyundai Genesis, Rolls Royce, Mercedes-Benz Class C, and Class S Sedan and Coupe, Audi and General Motor’s SUVs among others), some other forms of HUDs can currently be found in the market, such as, smartphones apps (Navier HUD, HUDWAY, Sygic), third party devices or self -mounting HUD devices (Garmin HUD +, Kshioe 5.5” Q7 Universal GPS HUD, Generic X6 3” Universal Multi-Function), windshield displays and hologram systems and helmets [22].

In the automotive domain, task-related information is presented on the HUD in two main forms: via *traditional screens* and via *optical see-through displays*. This thesis’ work used an optical see-through HUD. This class of display is characterized by the ability to see the world surrounding the observer through the display, thereby achieving both the maximal possible extent of presence and the ultimate degree of "real space imaging" [10, p. 284]. Optical see-through displays can convey information using different graphic’s perceptual forms, which will be discussed in the following section.

2.1.2.1 AR HUD PERCEPTUAL FORMS

Perceptual form is defined as the visual design of the information conveys, which includes (but is not limited to) the graphics’ shape, color, size, contrast, behavior, and location in the scene. Gabbard at al. [23] helped to conceptualize the design space for AR HUD automotive interfaces regarding two classes of perceptual form graphics related to the driver’s view perspective: screen-relative and world-relative. *Screen-relative* AR graphics are tied to a fixed position on the HUD independent of vehicle position or pose (e.g., a 2D turn arrow indicating turn ahead, Figure 4). On the other hand, *world-relative* AR graphics are perceptually tied to a spatial position in the real world (e.g., a virtual turn arrow “painted” in the lane, Figure 4). World-relative graphics use real-time geolocation, pose-estimation, 3D rendering software, and attract attention to a real-world object that falls within the HUD field of view [24, 23].

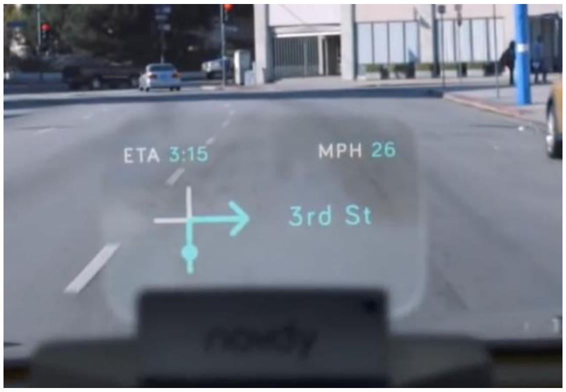
Screen-Relative	World-Relative
	

Figure 4- Examples of screen-relative (left) and world-relative (right) AR HUD perceptual forms

2.1.3 AR HUD CHALLENGES

In contrast to the potential benefits of AR HUDs (see section 2.1.2 for more details) in the automotive field, there are still obstacles to ensuring that this technology is designed both to satisfy human needs and to ensure that that the display interface is compatible with task demands and driver's mental representations. In this section, some of these challenges are discussed.

2.1.3.1 AR HUD PERCEPTUAL FORMS

Current models of spatial/visual attention assume that in the first preattentive stage of visual information processing, complex scenes are visually grouped based on Gestalt principles. In the second attended stage, these perceptual groups are used to control the distribution of spatial attention across the visual field [7, 25]. It has been assumed that by placing information on the HUD within the driver's FOV, one is able of simultaneously using information from both the environment and the device. However, because visual/spatial attention is a limited pool of resources, there are problems with the parallel processing assumption; thus, one cannot focus at two different perceptual groups at once [26]. In fact, early research has suggested that humans cannot simultaneously process HUD instrument symbology and real-world-referenced information [25]. In other words, when individuals focus attention in the world, objects on the HUD are excluded from processing and vice and versa. Additionally, because the preattentive grouping phase directly controls the direction of focused attention, switching from processing information from one perceptual group to another requires a change of spatial attention [26]. This phenomenon creates a switching cost estimated to be 150 milliseconds [25].

Because HUDs don't appear to eliminate the transition cost between the instrument processing and real-world processing, HUDs graphics should be designed in a way that cues segregating HUDs from the world are reduced. McCann, Foyle, & Johnston [25] suggested that an approach for this problem would be to design HUD graphics as conformal as possible to the world-scene and incorporate these graphics into the world itself. For instance, in the aviation field, Wickens & Long [27] reported that conformal HUD graphics resulted in a 30% decrease in flightpath deviation when compared to non-conformal graphics. According to the authors, conformability reduced the transition time between HUD graphics and the environment.

However, there are still technological challenges in placing virtual imagery onto the world itself. For instance, to put the virtual objects at the right place in the real world when driving, the relative position between the camera and the real world must be known. This challenge is also referred to as registration. It is crucial that the algorithm used for registration is accurate and real-time to make drivers ignore the difference between virtual reality and actual reality

2.1.3.2 AR HUD LIMITED FIELD OF VIEW

AR HUDs combine the interface graphics with the driving environment, which can be seen as an advantage in many situations but has also been reported to pose other obstacles. One of these challenges is related to the small field of view presently between 10°-15° (as shown in Figure 5). Although human foveal vision makes up about 1 ° of the visual field, we are heavily dependent on peripheral vision for visual tasks, and therefore limited FOV can make driving more difficult, because important road elements may lie outside the interface

Figure 5- Example of current AR HUD field of view

2.1.3.3 INATTENTIONAL BLINDNESS

Previous research has found that without attention, people fail to perceive important visual features of the environment they are looking at [28]. This phenomenon is known as *inattentional blindness* that, in simple words means looking without seeing. This concept of selective seeing was first reported by Neisser [29]. In this experiment, participants watched a video in which two groups were playing a passing-pall game; one of the groups wore black uniforms and the other wore white uniforms. During the video, a woman unexpectedly appeared carrying an umbrella. Neisser asked participants to count the number of passes only between one group. Surprisingly, after the ball-passing counting task, only 21% of participants reported seeing the woman with the umbrella in the video.

In the aviation domain, inattentional blindness has also been investigated. For instance, Haines [30] conducted a study in which pilots operated an aircraft in a simulated environment in which another aircraft was explicitly blocking the runaway. Surprisingly, some experienced pilots proceeded with the landing process unaware (not seeing) of an airplane blocking the runaway, and thus, they could not avoid the collision between the airplane they were flying and the blocking airplane.

2.1.3.4 COGNITIVE TUNNELING

In AR HUD, *cognitive tunneling* is understood as a phenomenon in which people involuntary fix mental resources on one aspect of the interface at the expense of other sources of information [31]. This phenomenon was first noted by Fischer, Haines, & Price [32] in a study in which they showed that pilots using a head-down display were more likely to observe an aircraft taxiing on a runway than pilots using a HUD. In fact, cognitive tunneling has been shown to decrease pilots' response times to unexpected events when using a HUD [27]. In the same way, McCann, Foyle, & Johnston [25] noticed that HUD's held pilots' more attention than HDD and so, pilots tunneled out the real world when using a HUD. Additionally, Ververs & Wickens [31] investigated how cognitive tunneling is influenced by both conformal and partially conformal HUD graphics. In this case, conformal means that HUD graphics conform to features of the outside world. Based on this study's results, the authors argue that "only truly conformal graphics can combat the effects of cognitive tunneling" [31, p. 3].

2.1.3.5 OCCLUSION

Occlusion means the visual blocking of objects. The main problem with occlusion is the inaccurate separation between the background and foreground that leads to an incorrect depth ordering between objects, and so, objects might look like they are not part of the environment [33]. In this sense, when objects in the real world are completely occluded, under normal circumstances they cannot be seen anymore. Thus, overlaying AR HUDs navigation cues onto the road could potentially overlap graphics with important real-world road objects, that could interfere with driving.

2.1.4 DRIVING BEHAVIOR

Driving behavior can be defined as actions drivers take to "maintain lateral and longitudinal control of the vehicle to safely move the occupants of a vehicle from one point to another" [16, p. 7]. Behavior is linked to people's choices and actions regarding goals, priorities, and trade-offs while driving and it is influenced by the driver's goals, needs, and motivations [34]. *Driving performance* reflects the driver's capability (cognitive, perceptual and motor control) in specific environments [34]. Performance is related to the final product if what a driver is capable of doing given both human limitations and vehicle and environmental constraints [35]. It is noteworthy that while not similar, performance and behavior are closely related, and so, the way drivers behave might influence their driving performance [16]. Performance is related to what the driver can do in a given situation, and behavior is related to what a driver tends to do in a given situation within its limits of performance [35].

Lateral vehicle control, also referred to as lateral behavior, can be defined is a measure of drivers' ability to keep the car in the correct lane [16]. Measures of lateral behavior include but not limited to standard deviation of lane position, mean lane position and standard deviation of steering degrees.

Longitudinal vehicle control, also referred to as longitudinal behavior, is a measure of driver's ability to respond to events ahead and maintain sufficient speed and distance from the vehicle ahead [16, pp. 7,8]. Measures of longitudinal behavior include but not limited to standard deviation of vehicle speed, standard deviation of time headway, standard deviation of distance headway, mean time headway and mean distance headway.

This work analyzed two measures of driving behavior: standard deviation of lane position (derived from lateral vehicle control) and standard deviation of speed (derived from longitudinal vehicle control).

2.1.5 GLANCE BEHAVIOR

Glance behavior is defined as the way drivers “allocate glances and visual attention between the driving task, and the head-up display supported secondary task” [16, p. 6]. A single glance is calculated by summing the saccade leading to an area of interest (AOI) to a series of saccades and fixations within that same AOI. An AOI is defined as a physical location where specific task-related information can be found [7, p. 50]. AOis must be large enough so that eye-tracking glasses can reliably track and discriminate a fixation from one AOI to another. Here, a fixation is defined as the period of time (usually between 200 to 300ms) in which the eyes temporally remain still in a fixed point in space; a saccade is the fasted movement the human body can produce (usually between 30-80ms) and it is defined as the rapid movement of the eyes from one fixation point to another [36].

2.1.6 SPATIAL KNOWLEDGE

As humans, we develop mental representations of our surroundings as we move through and learn about the environment. These spatial representations, also known as cognitive maps, are understood as long-term memory structures that are formed in the hippocampus [37]. Tolman [38], the first author to introduce the concept of cognitive map, proposed that after a satisfactory period of environmental learning, both animals and humans develop mental representations of the space analogue to a real map. It is important to emphasize that the use of the term 'cognitive map' does not imply that there is some cartographic map inside one's head. Instead, the term cognitive map is a way of summarizing the information encoded in a person's cognitive representation of the environment.

Maintaining and developing correct cognitive maps is crucial for navigation tasks. Individuals who have well-formed cognitive maps are capable of navigating based on their own internal knowledge and require fewer cognitive resources [3]. Even though in many situations (e.g. when driving from home to work) the navigation task might involve automatic processing, unanticipated situations (e.g. traffic accidents, bad weather, heavy traffic, GPS not working) might require the ability of navigating alternative routes, and therefore, a well-formed cognitive map comes into play. Moreover, the ability to develop an accurate and comprehensive cognitive map acts as a social function in which individuals are able to navigate for others, providing verbal directions and sketching direction maps [39]. Additionally, it empowers users to “find locations not included in a navigation system’s database (for example districts or specific buildings)” [40, p. 119].

The cognitive mapping process (commonly referred to as acquisition of environmental spatial knowledge), is mainly categorized into three levels: landmark knowledge, route knowledge, and survey knowledge [41]. These three forms of spatial knowledge give distinct levels of both global and local abstraction of the spatial environment, and acquisition of these levels is a dynamic process that occurs in a variety of forms [42]. In this work, only the first two stages of knowledge acquired were measured. This was decided for two main reasons. First, previous work suggests that both landmark and route knowledge are acquired prior to a more complex form of survey knowledge [41]. Second, due to the extensive length of the route used for this experimental study, we expect that a little survey knowledge would have been developed.

1. *Landmark knowledge* – “memory for distinctive objects and/or views within the environment” [3, p. 1]. This level is characterized by visual depiction of salient landmarks [7] in which individual locations are known but not spatially related to one another [43].
2. *Route knowledge* – *“*memory for procedural linking of landmarks, including order, inter-landmark distances and required actions” [3, p. 1]. This level is characterized by the ability of linking together information regarding the relative position of landmarks [7], however, individuals lack an overall understanding of the spatial organization [43]. At this point, individuals can provide specific navigational instructions, such as “turn left at the gas station”, but they are not able of describing the entire path from memory.
3. *Survey knowledge* – *“*memory in which landmark and route knowledge is integrated into a configurational, map-like whole” [3, p. 1]. At this level, a mental map of the environment is reconstructed and individuals can understand the links between locations and answer questions regarding spatial relations [7, 43].

2.1.6.1 MEASURING LANDMARK AND ROUTE KNOWLEDGE

The literature reports several methods to measure levels of spatial knowledge acquisition. Due to the ease and simplicity of use, this work employed two specific methods to measure landmark and route knowledge: iconic recognition task and scene ordering task, respectively.

Iconic recognition tasks involve the correct identification of a target image, including specific landmarks that occurred in the driving scene. For instance, Oliver and Burnett [40] presented participants with both target and distractor images and asked them to identify which images they had encountered in the driving simulation environment. Other studies have employed the same method in both real [44, 4] and virtual [3, 45] environments.

Scene ordering tasks require participants to sort a set of images into the correct order they appeared in the driving scene. In the driving context, several studies have employed this method [4, 46, 40, 3] as a complement of the use of iconic recognition tasks for the landmark knowledge.

2.2 RELATED WORK

2.2.1 NAVIGATION STRATEGIES

When navigating in an environment, humans spontaneously can adopt two strategies: either spatial strategy or response strategy. The spatial strategy is subserved by the hippocampus and involves navigating the environment and forming associations between landmarks (stimulus-stimulus association). In this strategy, that is also a form of explicit memory, one absolves knowledge of the environment and forms a cognitive map during the process [47]. Eventually, this cognitive map allows us to navigate between two points in an area taking a particular route that we have never taken before. On the other hand, response strategy is subserved by the striatum (which includes the caudate nucleus) and involves learning pattern associations such as “turn left or right” from a given starting position (stimulus-response associations) [48, 47]. This latter strategy is kind of an autopilot mode in which after a series of repetitions, you are able to go from home to work using the same route, without necessarily knowing which landmarks or environmental aspects you have seen along the way.

Different strategies are related to different brain size development. Several studies have shown that spatial learning is positively correlated to hippocampus grey matter [48, 49, 50] and response learning is positively correlated to caudate nucleus grey matter [47, 50]. Additionally, it is suggested that it is not only the creation of spatial maps that produces a larger hippocampus but also that people who already have higher hippocampus have a higher likelihood to create spatial maps [51]. Further, grey matter in the caudate nucleus is shown to be negatively correlated to grey matter in the hippocampus, suggesting a competitive interaction between them [47].

Understanding how different navigation strategies (and consequently, wayfinding devices) have the potential to cause physical changes in the brain, it is an important topic to be understood when designing new interfaces. For instance, low grey matter in the hippocampus has been shown to be a risk factor for Alzheimer's, and therefore, using spatial-memory strategies for navigation has the potential to help offset cognitive impairments. Additionally, a reduced risk of dementia is also shown by people who employ spatial strategies to navigate [52].

Because most of the research involving brain-imaging has people navigating in a virtual environment and a virtual maze, there is no direct evidence that wayfinding devices lead to hippocampus atrophy and less activity. However, there are studies linking the hippocampus and navigation while driving without the aid of technology. These researches are presented in the next section.

2.2.2 SPATIAL KNOWLEDGE AND THE HIPPOCAMPUS

The hippocampus is a section in the vertebrate brain in which one of the roles is to facilitate spatial memory in the form of navigation and long-term memory. The most famous study linking spatial navigation and the hippocampus was conducted by Maguire et al. [53]. In this study, the authors investigated how the human hippocampus size is related to spatial navigation abilities. They analyzed structural MRIs of licensed taxi drivers who have extensive navigation experience and compared results with a control group composed of people who were not taxi drivers. Results from this research showed that the posterior size of the hippocampus was significantly larger in taxi drivers than in the control group. Moreover, the right hippocampus volume was shown to be positively correlated to the driving experience as a taxi driver.

After the publication of this study, it was hypothesized that the difference in the hippocampus volume could be associated with the innate navigation expertise of drivers, and not to the driving experience per se. Maguire at al. [54] investigated this possibility. In this study, a total of 26 individuals who were not taxi drivers were analyzed. Using virtual reality, subjects had to navigate between 10 different locations in the town that was complex enough to afford more than one possible route. Then participants were presented with a pair of scenes, and they had to choose which scene was present during the study. Additionally, three other tests were conducted: a map drawing test in which participants had to draw a map of the virtual city, a standardized topographical memory test, and a standardized verbal memory test. Results from this study suggested that there is no association between innate navigation expertise and hippocampal grey matter volume.

In a more recent longitudinal study, Woollett and Maguire [55] investigated for a period of four years the hippocampi growth of 79 aspiring London taxi drivers. A control group of 31 people who were not taxi drivers but were of similar age, education, and intelligence was also included in this study. Initially, all subjects had similar hippocampus size and performed about the same in tests of both long-term memory and working memory. At the end of the study, only 39 aspiring taxi drivers earned their licenses. However, the remaining 20 participants who failed to complete their trainee program agreed to continue in the research. Results from this study showed that participants who earned their license performed much better in some memory tests relative the ones who had failed their license exams. Moreover, MRIs results showed that licensed taxi drivers showed growth in hippocampus size in the period of four years. The authors argued that cognitive exercises and training have the ability to cause physical changes in the brain. In fact, this study proved that the caused the growth in the brain showed by earlier research [53] was in fact due to the training process of navigating.

Although this section provided evidence that driving relying on internal knowledge is related to increased hippocampus activity and volume, currently, there is no direct evidence that driving using navigation systems contributes to less development and atrophy of the hippocampus. However, there are many studies showing that people using navigation systems are not creating the same sort of mental maps compared to people relying on their own internal knowledge. The most relevant researches are presented in the next section.

2.2.3 NAVIGATION SYSTEMS & SPATIAL KNOWLEDGE

Regardless of the potential benefits and opportunities that navigation systems provide, concerns have been raised in the literature about their negative effects on spatial knowledge acquisition caused by their extensive use. Research suggests over-reliance on navigation systems can cause drivers to be mindless of the environment and do not develop as much environmental spatial knowledge as drivers using paper maps [2, 56].

Burnett and Lee (2005) showed that navigation systems negatively impact the development of spatial knowledge. In this study, they compared the use of traditional audio turn-by-turn guidance and paper maps in the formation of a driver’s cognitive maps. They found that drivers who used paper maps remembered more scenes, were more accurate in ordering scenes seen along routes and drew more complex maps with a greater number of landmarks when compared to those using traditional audio turn-by-turn guidance. To explain this phenomenon, the authors argued that drivers using paper-map guidance process information at a deeper level because they have a greater requirement to evaluate the worth of the navigation information they are using or intend to use. In contrast, decisions made by drivers using traditional turn-by-turn instructions are simpler, and thus, require few mental resources than those using paper map guidance. In another study, Ishikawa & Takashshi [57] found that drivers using navigation systems showed poor scene-recognition memory and simply followed the direct route given by the application. On the other hand, drivers using paper-maps remembered more scenes, were able to take alternative routes and were more active in the navigation process.

The influence of navigation systems on spatial knowledge acquisition has been investigated not only for driving but also for pedestrian navigation. For instance, Munzer et al. [58] compared spatial knowledge acquisition and navigation performance for pedestrians using paper maps and navigation systems. In this within-subject design research, participants were exposed to four different conditions, three of them using navigation devices. In two conditions, participants were presented with a speech command together with the view-based picture of the intersection. In another condition, directions were visually indicated with a red line included in the picture of the intersection. Finally, in the last conditions participants used a paper-map to navigate to their destination. The researchers found that pedestrians using navigation aids showed worse spatial knowledge acquisition and orientation, but better navigation performance (e.g. smaller duration to destination) compared to when they used paper-maps. In another study, Willis et al. [59] compared spatial knowledge acquisition for pedestrians using mobile-maps and paper-maps when navigating a large external building environment. In this study, spatial knowledge was measured in terms orientation, Euclidean and route distance estimates. Results showed that pedestrians using mobile-maps performed worse in all measures of spatial knowledge than those using paper-maps. Ishikawa et al. [60] investigated how pedestrian's wayfinding behavior and acquired spatial knowledge are affected by three conditions: direct experience, mobile navigation and paper maps. In this between-subject design study, participants completed six routes using one of the above navigational aids. Better performance (i.e., shorter travel distance, faster travel speed, fewer stops made, smaller direction error, sketch maps with better topological accuracy, and wayfinding tasks rated as less difficult) was achieved in the direct experience condition compared to both mobile navigation system and paper-maps. Also, pedestrians using GPS showed worse performance in all categories compared to both direct experience and paper-maps. Based on these results, the authors argue that the use of mobile navigation aids is less effective than the other two methods to afford a smooth navigation.

2.2.3.1 UNDERLOAD

As it has been mentioned, the use of SatNav frees up more cognitive and attentional resources, resulting in cognitive underload, which can be viewed as a strength in some cases [2]. However, underload has also been found to result in decreased driver engagement and situation awareness [1].

Endsley [61] defines *situation awareness* (SA) as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” [61, p. 36]. SA involves three hierarchical phases that express acquired levels of environmental awareness: perception of the elements in the environment, comprehension of the current situation, projection of future status [61]. In driving, the use of technology to navigate automates aspects of the wayfinding behavior and thus, loss of situation awareness has been found to be a safety-related concern. Besides that, the impact of decreased SA has been investigated on environmental engagement and levels of spatial knowledge acquisition. For instance, Webber [4] found that participants using wayfinding devices were not only less aware of the environment surrounding them, but also less engaged with the navigation task itself (passive navigation) , and as a result, acquired lower levels of spatial knowledge.

2.2.3.2 ENCOURAGING ACTIVE AWARENESS AND ENGAGEMENT

In order to mitigate negative impacts of the extensive use of navigation systems on spatial knowledge acquisition, new research has focused on developing user-interfaces that promote engagement and active awareness with both the driving environment and the navigation process. Particularly, a special focus has been given on “encouraging environmental awareness and learning through changing the information relayed from device to user” [4, p. 31].

Ritcher et al. [62] demonstrated that spatial knowledge acquisition can be positively impacted by encouraging user interaction. In this study, user interaction with the device and the task was promoted by forcing participants to pan and zoom to progress route level information.

Results showed that better survey knowledge performance was yielded when participants interacted with the task and device, compared to participants using a common GPS device with no interaction during task completion.

Parush et al. [56] proposed an unusual approach in which drivers were "kept in the loop" to encourage active environmental and task awareness while navigating using wayfinding aids. In this study, the authors employed two main strategies. First, continuous availability of spatial information was eliminated, so that drivers could access it only when needed. The authors hypothesized that by doing so, drivers would be "more active and "mindful" to the cues in the environment and would not rely continuously on the navigation system" [56, p. 241]. Second, orientation quizzes were periodically introduced during the driving. They hypothesized that by asking participants to indicate their positions in the map, they would employ a "keep the user in the loop" strategy, forcing drivers to be "more "mindful" of the environment, monitor their position, compare and update their cognitive map" [56, p. 241]. Results from this study demonstrated that both strategies of "keeping the user in the loop" during navigation had a positive impact on the levels of spatial knowledge acquired. Further, when these strategies were applied together, better results were achieved.

To date, the most common approach to improve learning and engagement with the environment is landmark-based navigation. This approach has been positively accepted due to its ability to support natural strategies humans use to navigate and its ability to enhance both driver's and pedestrians' spatial knowledge. By providing external reference points that are easily remembered and recognized, drivers no longer need to refer to a navigation display to locate a navigation decision point. Burnett [63] argues that other benefits can be achieved by the use of landmark-based systems, such as, less navigation errors, fewer glances towards an in-vehicle display, lower self-reported workload and higher confidence on navigation performance. In line with this, several studies report that providing turn-by-turn guidance with landmark directions, increases route knowledge of an environment (compared to traditional basic guidance systems) within pedestrian [64] and vehicle [65, 40, 66, 44, 67] navigation aids.

Gramann et al. [65] showed that providing salient turn-by-turn descriptions of landmarks (i.e. "turn right in front of McDonalds) yielded to better results of driving performance and spatial learning as compared to plain directional commands (i.e. "turn right in 150 meters"). In another study, Oliver & Burnett [40] developed a novel learning-oriented navigation system that included audio provision of landmark directions (i.e. "In 50 yards turn right, at the church"). The authors compared driving performance and spatial knowledge acquired between the learning-oriented navigation system and the basic guidance navigation system (which provided basic guidance, i.e. "In 50 yards turn right"). Drivers using the landmark-based system showed better route recall, driving performance, and drew more complex stretch maps than drivers using traditional navigation systems.

2.2.3.3 HUD STUDIES ON SPATIAL KNOWLEDGE

To date, the vast majority of driving studies present landmark-based navigation systems using head-down displays, typically located in the center of the dashboard of the vehicle. With recent advances in AR HUD technology, there are new opportunities to display navigation information directly on the windshield within drivers' forward field of view, allowing drivers to gather information needed to navigate without looking away from the road [5]. Conceptually, AR HUDs allow designers to overlay navigation information directly onto real-world landmarks, thus allowing drivers to potentially acquire greater spatial knowledge as compared to head-down display navigation systems [6] .

For instance, Wu et al. [67] developed a full-windshield head-up display (FWD) prototype to provide landmark-based navigation. This prototype displayed navigation information to users by highlighting key landmarks, such as road signs, monuments, store signs or traffic lights. Although the cost for a commercial prototype is still expensive, the authors argue that this navigation system has the potential to help complex driving situations.

In another study conducted by Large, Burnett & Bolton [18], the authors investigate the effect that four different augmented reality navigation systems (conventional, AR arrows, landmark arrow, landmark box) presented on a HUD have on driving performance. Different HUD elements composed each of the conditions. The conventional condition consisted of a fixed arrow located on the right or left of the HUD display with distance to turn information. The arrow condition consisted of a fixed centered arrow that pointed out to the required turning. Finally, landmark arrow and landmark box conditions highlighted turning landmarks using either a box or an arrow. Results showed that turning response time was lower for landmark-based navigation systems (both arrow and box), with drivers being able to distinguish the correct turning 50 meters sooner than the conventional distance to turn navigation systems. Additionally, when compared to conventional systems, landmark-based conditions showed a higher success rate and smaller perceived workload according to NASA-TLX. Finally, the authors showed that augmented reality graphics presented via HUD has the potential to enhance real-world landmarks that are not easily communicable and identifiable [18].

3. METHODS

3.1 PARTICIPANTS

3.1.1 PARTICIPANTS RECRUITMENT

For this study, Institutional Review Board IRB #17-239 (Appendix A) was approved for data collection with human participants. Participants were students at Virginia Tech who volunteered by answering advertisement flyers (Appendix A) posted around the Grado Department of Industrial and Systems Engineering.

The prerequisites to the participation in the study were:

a) Active driver with a valid US driver's license;

b) Drive a minimum of 3,000 miles per year;

c) Not an employee of any vehicle manufacturers;

d) Age 18+;

e) Have normal or corrected-to-normal vision (after corrected using contact lenses).

3.1.2 PARTICIPANTS DEMOGRAPHICS

We recruited twenty-four participants (twelve males and twelve females) aged between 18-40 years for this study. All participants had a valid driver's license and had a normal or corrected-to-normal vision. Twenty-three participants had a valid US driver's license for at least 6 months (mean 5.89 years; minimum: 6 months, maximum: 15 years). Additionally, five participants had driving experience with a non-US driver's license for at least four years (mean 9 years, minimum: 4 years, maximum: 16 years). Overall, all twenty-four participants had a mean of 7.44 years of driving experience, with a minimum of 6 months and a maximum of 16 years. Further details can be seen in Table 1.

Table 1- Demographic characteristics of participants

Characteristic	Category	n	%
Gender	Male	12	50%
	Female	12	50%
Years with License	3	5	21%
	4	3	13%
	5	3	13%
	6+	13	54%

3.2 EQUIPMENT

A fixed-base, medium-fidelity driving simulator was used for this study. This simulator is composed of the front half of a 2014 Mini Cooper cab fitted with a curved projection with 94 degrees of view displaying a simulated road scene and contains both side and rear-view mirrors that allow participants to view their surrounding environment (Figure 6). The simulator also contains a 7" Lilliput USB monitor mounted directly behind the steering wheel to convey vehicle speed information. Additionally, the simulator is equipped with a Pioneer Cyber Navi HUD with conformal AR graphics capabilities. The area displayed on HUD is 780x260 pixels, FOV is 15 degrees and the virtual image position is approximately 3m away from the eyepoint. The driving simulator software was integrated with customized software, developed using X3D and Python, so that the AR HUD can provide real-time 3D AR graphics perceptually overlaid into the dynamic CG-generated driving scene. That is, unlike other studies that render AR directly into a simulated environment (e.g. using virtual reality), our testbed renders AR graphics onto an aftermarket HUD, calibrated to a projected road scene to produce a more ecologically valid driver experience.

During the study, participants wore SensoMotoric Instruments (SMI) eye-tracking glasses equipped with audio and video recording (see Figure 7). We used iView Eye Tracking Glasses 2.6 software to collect binocular eye gaze data sampled at 60 Hz.

Figure 6- COGENT Lab's AR HUD can be conformal to the simulated world (top left). Top view (right) and side view (bottom left) of COGENT's driving simulator

Figure 7- SensoMotoric Instruments (SMI) eye-tracking glasses used within this study

3.3 DRIVING SCENARIO

This study employed a single, carefully constructed route that contained a series of turns, straightway paths and city intersections. To properly counterbalance and minimize order effects, a driving scenario composed of the same amount of right and left turns without any ordered pattern was constructed. The road scene was created using the National Advanced Driving Simulator miniSim Software, containing four maneuvers to the left, four maneuvers to the right, three straight intersections and a roundabout (see Figure 8). This consideration was important because if a series of four left turns are presented to the driver, one may infer that the next turn will be a right one for example. To better represent a real driving environment, we added ambient traffic that included moving cars, buses, and trucks. A total of twelve decision points was presented to drivers in the following order:

1. Go Straight #1
2. Maneuver #1 to the Right
3. Maneuver #2 to the Left
4. Maneuver #3 to the Right
5. Maneuver #4 to the Right
6. Roundabout #1 Go Right
7. Go Straight #2
8. Maneuver #5 to the Left
9. Maneuver #6 to the Left
10. Go Straight #3
11. Maneuver #7 to the Right
12. Maneuver #8 to the Left

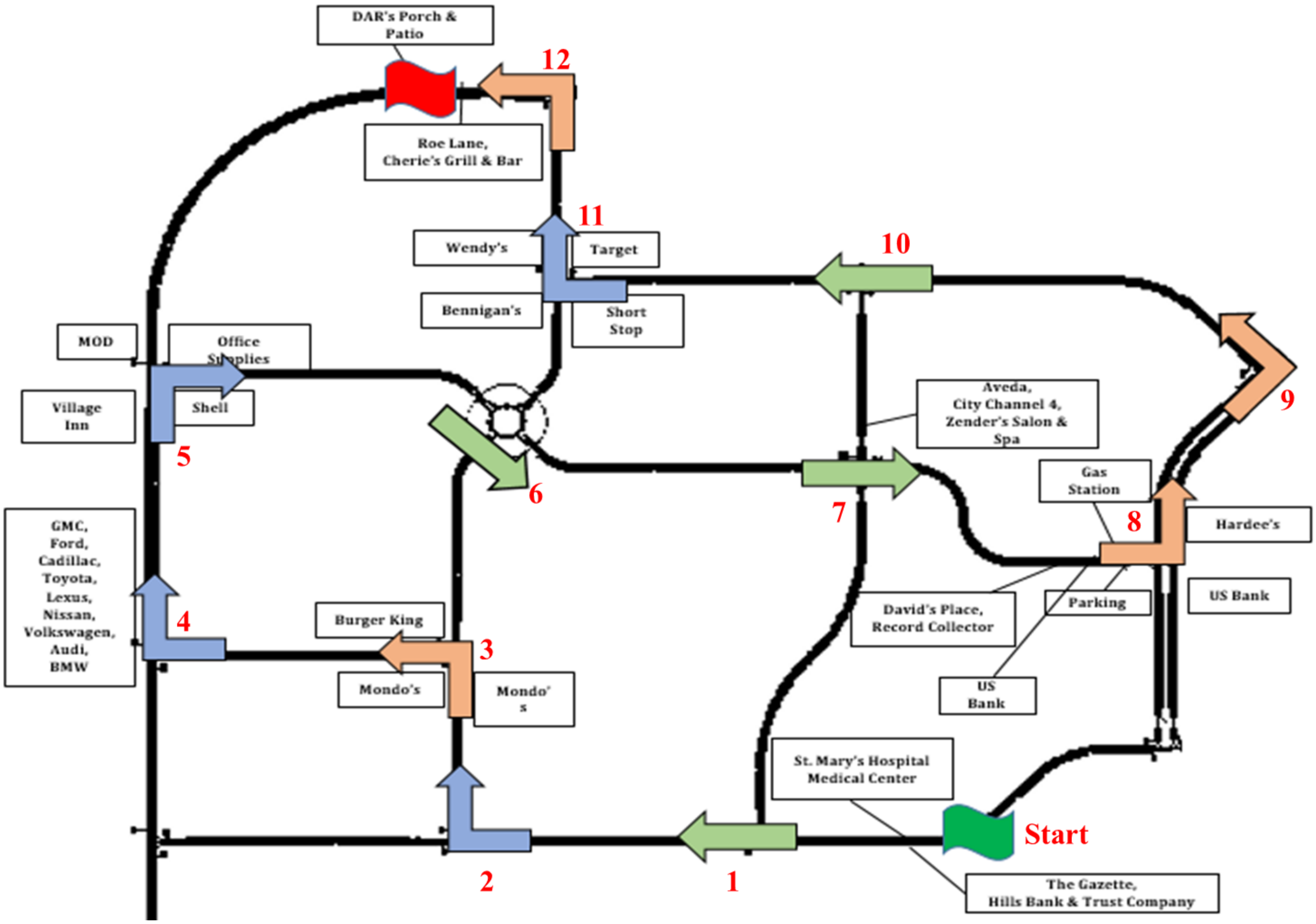


Figure 8- Driving scenario used within this study

3.4 EXPERIMENTAL DESIGN

We employed a 2x2 between-subject experimental design in this study. Table 2 shows an overview of the design used.

Table 2- Overview of Research

Main Objective	Understand how differing AR perceptual forms affect driver's spatial knowledge acquisition.
Research Questions	R1: How does HUD graphics' perceptual form (world-relative vs. screen-relative) impact drivers' acquisition of spatial knowledge (specifically landmark and route knowledge)? R2: How does HUD graphics' perceptual form (world-relative vs. screen-relative) impact driving measures (specifically driving behavior and missed turns)? R3: How does HUD graphics' perceptual form (world-relative vs. screen-relative) impact drivers' glance behaviors?
Independent Variables	• AR HUD graphics' perceptual form (World-relative post sign and Screen-relative traditional arrow) • Gender (male and female)
Dependent Variables	Survey data • Workload • Situation Awareness • Usability Driving data (Behavior) • Speed standard deviation • Lane position standard deviation • Missed turns Gaze behavior • Percentage of time looking at the HUD graphic • Percentage of time looking around the HUD graphic • Mean HUD graphic glance duration • Maximum HUD graphic glance duration • Total number of glances at the HUD graphic • Total number of glances around the HUD graphic Spatial knowledge • Landmark knowledge • Route knowledge
Task	Navigate through turns
Experimental Design	n=24 2x2 factors 1 replication Gender (2): between-subject design, counterbalanced AR HUD graphics' perceptual form (2): between-subject design, counterbalanced

3.4.1 INDEPENDENT VARIABLES

3.4.1.1 WORLD-RELATIVE POST SIGN

The world-relative condition consisted of an artificially created “Post Sign” placed at the corner of an approaching intersection containing a real landmark. This world-relative graphic resembled a real post sign, containing a square base, four tall sides, and two-directional arrow-shaped features that were presented either on the left or right side of the landmark to indicate the direction of turning (see Figure 9).

Figure 9- World-relative post sign interface indicating both left turn (left) and right turn (right)

In the case of straightaways, the arrow-shaped feature was attached at the backside of the post (relative to driver’s approach), which was not visible at a distance but became salient as the driver approached the landmark, as it can be seen in Figure 10.

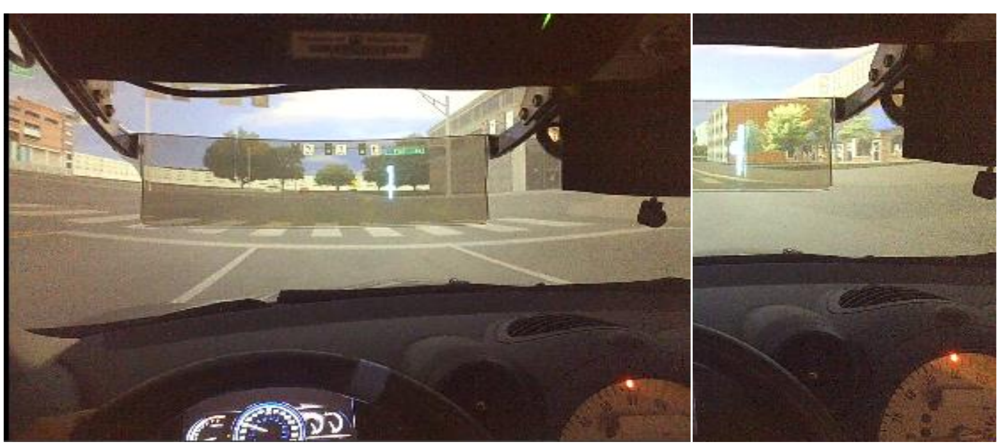


Figure 10- World-relative post sign interface at a distance of a “keep going straight” point drive (left) and approaching said “keep going straight” point (right)

All conformal post signs were positioned on the same side of the road as the driver (right in the US) to increase the likelihood that graphics would stay inside the HUD's field of view. Additionally, post signs were also placed on the far side of each intersection, so that the driver could turn just before reaching the landmark and they appeared 150 meters from intersections.

3.4.1.2 SCREEN-RELATIVE TRADITIONAL ARROW

The screen-relative condition displayed turn directions by using a screen-fixed traditional arrow located directly ahead of the participant on the HUD (see Figure 11). This condition presented a 2D turn arrow (oriented straight, left, or right as appropriate) with no numeric information and appeared on the center of the HUD. To be consistent with the world-relative condition, screen-relative graphics appeared when drivers were 150 meters from a turn and stayed present until the intersection was reached.

Figure 11- Screen-relative traditional arrow as drivers approached the turn

3.4.2 DEPENDENT MEASURES

- Survey Data
    - **Workload:** Workload was subjectively measured using the NASA-TLX questionnaire, which includes six items ranging from very low/perfect (-10) to very high/failure (10). The items included mental, physical and temporal demand, performance, effort, and frustration.
    - **Situation Awareness:** Situation Awareness was subjectively measured using the Situational Awareness Rating Technique (SART) questionnaire, which includes

nine items ranging from 1 (Low) to 7 (High). The items included Familiarity of the situation, focusing of attention, information quantity, information quality, instability of the situation, concentration of attention, complexity of the situation, variability of the situation, arousal, and spare mental capacity.

    - **Usability:** Participants completed an acceptance evaluation of HMI at the end of their session using a usability questionnaire. They ranked the interface in the following subscales: ease of navigation, ease of viewing, trust, effect on driving, effect on distraction, etc. Additionally, there was an opportunity to provide open-ended feedback about the HMI design.

- Driving Data:
    - **Behavior:** Measures of driving behavior included speed standard deviation and lane position standard deviation. Table 3 describes driving behavior variables used in this study.

Table 3- Driving behavior measures collected

Variable	**Unit**	**Description**
Vehicle speed	Miles per hour (mph)	Speed in which vehicle travels
Lane Position	Feet (ft)	Distance from the center of the lane (+ to participants left, - to participants right)

    - **Missed turns:** Refers to the number of turns that participants missed during the drive.

- **Gaze Behavior:** Visual attention was analyzed by using eye movements captured from SMI eye-tracking glasses. The measures used were:
    - Percentage of time looking at the HUD graphic
    - Percentage of time looking around the HUD graphic
    - Mean HUD graphic glance duration
    - Maximum HUD graphic glance duration
    - Total number of glances at the HUD graphic
    - Total number of glances around the HUD graphic

- **Spatial Knowledge**
    - **Landmark Knowledge**: Landmark knowledge was assessed by an iconic recognition task, which involves the correct identification of a target image including specific landmarks that occurred on the experimental traveled route.
    - **Route Knowledge:** Route knowledge was assessed by a scene ordering task, which required participants to sort a set of images into the order they appeared on the experimental traveled route.

3.5 PROCEDURE

Upon arriving to the lab, participants completed a short demographic survey and consented to the research using the IRB approved Informed Consent Form (Appendix A2). To start the study, participants sat in the driving simulator, adjusted the seat to a comfortable driving position, and then we calibrated the HUD vertically and horizontally to ensure that all conformal AR graphics were perceived to be in the same location, independent of participants' height and seat position. Figure 12 shows an example of a calibrated HUD in which AR graphics were overlaid onto the calibration scenario.

Figure 12- AR HUD graphics overlaid onto calibration scenario

We instructed participants to perform a familiarization drive to get comfortable with both car dynamics and driving simulator functionalities. The familiarization drive lasted for a minimum of five minutes and ended when we verified that the participant was comfortable and able to maintain vehicle control while starting, stopping, turning, and, driving straight. Participants were advised to drive at speeds of approximately 30 mph and to obey all traffic rules and norms, including traffic signals. If participants exceeded the speed limit by more than 10%, an audible siren sound was triggered indicating that they needed to slow down.

Participants were not exposed to either of the SatNav conditions prior to the data collection drive in order to mimic a worst-case scenario in which drivers must use a navigation system without prior experience, for example when renting a car. Participants were randomly assigned either the world-fixed or screen-fixed condition and were unaware of which condition was being applied to them during the study.

Each data collection drive lasted between 10 and 15 minutes. During this drive, participants drove through the route and were encouraged to verbally share feedback as they proceeded. The researchers did not provide any driving feedback unless participants missed a turn. After the drive, individuals filled out an open feedback survey, the NASA-TLX questionnaire to record their perceived workload and the SART questionnaire to asses situation awareness.

If any individual exhibited any sign of simulator sickness, they were advised to take a break and asked if they wished to continue the study later. If a participant had to halt their session due to sickness, they were excluded from quantitative analysis, but their feedback was used for qualitative analysis.

After the drive and completion of questionnaires, participants were led to a new room in which they were asked to complete a landmark and route knowledge test. For the landmark test, participants were presented with sixteen random scenes. Of the sixteen scenes, eight were scenes from the driving scenario (on-route) and eight images were scenes that drivers had not seen in the traveled scenario (off-route). Each on-route scene had its corresponding off-route scene, with a slight variation (Appendix B). Participants were asked to sort images into two piles: a pile with images they remembered seeing in the driving scenario and images they did not remember seeing in the driving scenario. After collecting information regarding both image piles, we asked participants to complete a landmark heat map task using VT Qualtrics. In this task, participants were randomly presented with images they indicated they had remembered and were asked to select up to ten points which they believed helped them to remember the scene.

Next, participants were given a route knowledge test. For this, they were given a set of eleven images from the traveled route and were asked to sort them in chronological order and put them in the numbered spaces as seen in Figure 13.

Figure 13- Lab space used for route knowledge test

4. DATA ANALYSIS

For all data, we conducted one between-subject ANOVA for both world-relative and screen-relative conditions. Minitab 18 software was used to carry out all statistical analyses. To ensure that results obtained in this study were reliable, assumptions of independence, normality and equal variances of residuals were checked (Appendix D, E). When assumptions of parametric testing were not met, logarithmic transformation was applied to raw data. In case this strategy did not correct the parametric assumption, we used nonparametric methods. We used the conservative Tukey correction for all post-hoc comparisons and accepted statistical significance at $p<0.05$. This step was important because we wanted to make sure that reliable and correct conclusions were obtained from the study. Also, we wanted to be confident as to whether observed differences in the data samples were simply due to chance, or if the populations were indeed different. Figure 14 shows an example of the output for an individual analysis that was performed, and its plots are described below.

Normality assumption: On the top-left graph, the normal probability plot of residuals is presented and each residual was plotted against its expected value under the normality assumption. When this type of plot is nearly linear, it suggests agreement with normality. In this case, it can be observed that the data followed a straight line and there is no indication of large departure from the normality assumption. The same conclusion can be obtained by the histogram in the bottom left part of the picture. As it can be observed, the data followed a bell shape curve and did not show any strong evidence against the normality assumption. Because each dataset of the analysis was small and had only 24 data points, most of the time the bell shape curve is not perfect. However, this is not strong evidence against the normality assumption. Therefore, to be sure of the normality assumption results, Anderson–Darling goodness of fit test was also performed, as shown in Figure 15. All analyses used a 95% of confidence and a significance level is 0.05. If a p-value greater than 0.05 was found, the null hypothesis failed to be rejected and therefore, we concluded that there is enough evidence to support the hypothesis that the residuals follow a normal distribution with mean 0 and variance α.

Homogeneity of variances assumption: On the top right graph, the plot of residuals against the fitted values is presented. For the data to be considered to have equal variance, the variance of the error terms must be constant (i.e. it must not present any pattern), and they must have a mean of zero. In this case, we observed that residuals form a horizontal band around the 0 line and no pattern is presented. There were no residuals that stand out from the basic random pattern of residuals, indicating that no outliers were presented. Therefore, we concluded that the assumption of equal variances of errors terms was met.

Independence of samples: On the bottom left graph, the residuals against observation order was plotted. For this assumption, it was desired to check whether the data sample came from a randomized sample design and the rows in the dataset do not influence one another. As we observed in this plot, residuals exhibited a normal random noise and bounced randomly around the 0 line. This indicated that there was no serial correlation in the dataset.

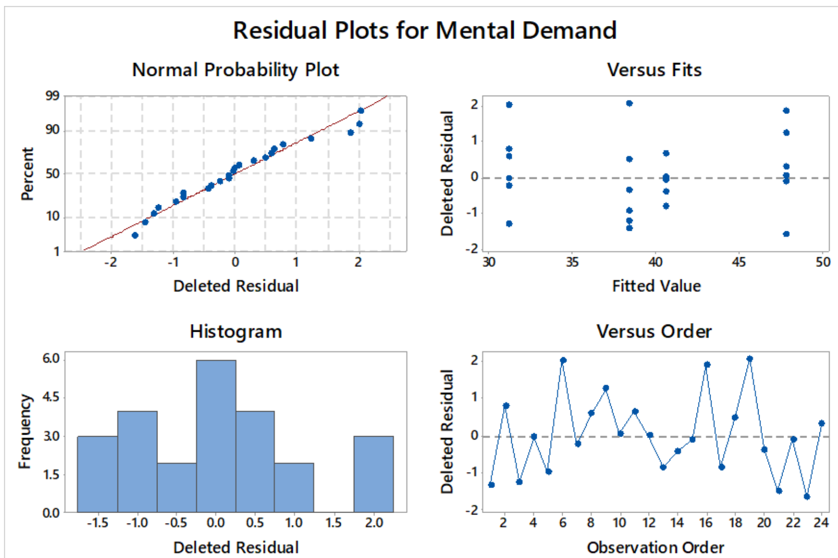


Figure 14- Example of Residuals Assumptions Analysis

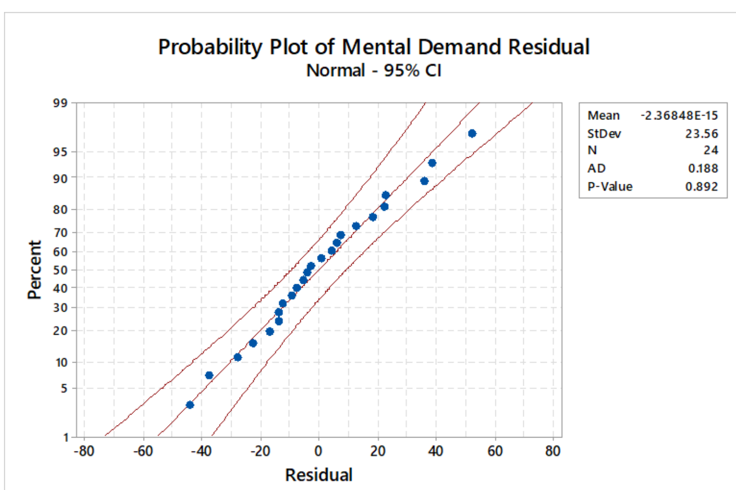


Figure 15- Example of Probability Plot of Residuals with AD Test

4.1 SURVEY DATA

In this study, each participant answered questionnaires regarding perceived workload, situation awareness and interface acceptance using VT Qualtrics. Data was downloaded into Excel and imported to Minitab18 where descriptive statistics of the population and further statistical analysis were carried out.

4.1.1 WORKLOAD

NASA-TLX subscales were individually analyzed between conditions and genders. Overall NASA-TLX workload scores were calculated by averaging raw data from mental, physical and temporal demand, performance, effort, and frustration subscales. For this analysis, no weighting technique was used.

4.1.2 SITUATION AWARENESS

SART is a subjective measure of situation awareness (SA) that uses the following constructs to measure participants' SA (see Table 4).

Table 4- SART Construct Definitions. Source: [68, p. 105]

Construct	Definition
Instability of the situation	Likeliness of situation to change suddenly
Variability of the situation	Number of variables which require one's attention
Complexity of the situation	Degree of complication (number of closely connected parts) of situation
Arousal	Degree to which one is ready for activity (sensory excitability)
Spare of mental capacity	Amount of mental ability available to apply to new variables
Concentration	Degree to which one's thoughts are brought to bear on the situation
Division of attention	Amount of division of attention in the situation
Information quantity	Amount of knowledge received and understood
Familiarity	Degree of acquaintance with situation experience

In this technique, participants rated a series of scales to the degree they perceive (1) demand on resources, (2) supply on resources and (3) understanding of the situation. [69]. Then, these scales were combined to provide an overall SART score for a given situation or system. Equation 1 shows how the SART score was calculated for this study.

$$SA = U - (D - S)$$

Equation 1- Overall SA Calculation using SART

Where:

- **Demand**= instability of the situation + complexity of situation + variability of situation
- **Supply** = arousal + concentration of attention + division of attention + spare mental capacity
- **Understanding** = information quantity + familiarity with the situation

4.1.3 USABILITY QUESTIONS

Participants were asked to fill out an usability questionnaire in which they were asked how much they agree or disagree with five statements about their drive. In each statement, they had to rank their level of agreement from 0 (strongly agree) to 100 (strongly disagree). Data was analyzed for each usability subscale between conditions and genders.

4.2 QUALITATIVE FEEDBACK

We asked participants to write down their thoughts and impressions about the HUD interface after the drive. Feedback from both world-relative and screen-relative conditions was examined and sorted into positive and negative comments. In this section, we did not carry out any further analysis.

4.3 DRIVING MEASURES

We only analyzed driving measures (i.e. standard deviation of speed and standard deviation of lane position) for the period participants were using the navigation cue. Therefore, we gathered driving data from the time the navigation cue appeared (492 ft before the turning) up to the time participants started to cross the intersection (see Figure 16).

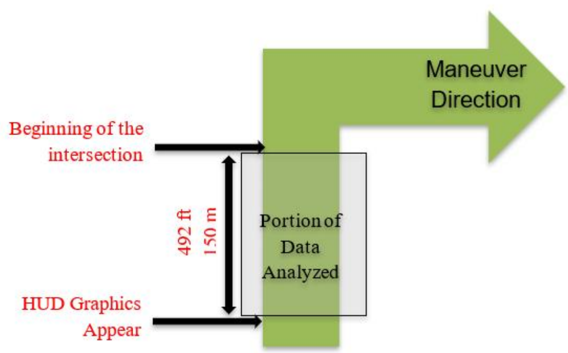
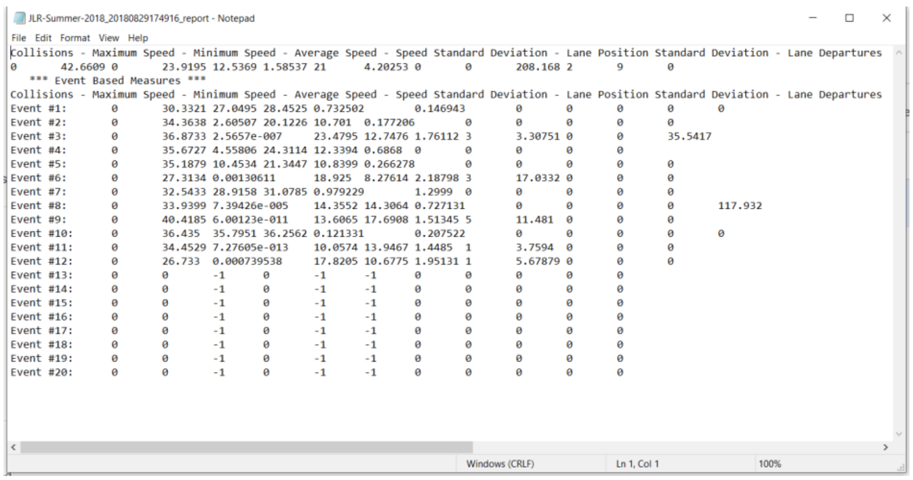


Figure 16-Graphic representation of driving data analyzed

We configured MiniSim software to collect driving measures corresponding to 492 ft before each maneuver. Output was given in .txt format (see Figure 17) and raw data was downloaded into Excel where it was organized using VBA macros. Then, we imported data to Minitab18 where descriptive statistics of the population and further statistical analysis were carried out. Each event from driving measures output (see Figure 17) corresponded to a maneuver decision point, as shown in Table 5



```
JLR-Summer-2018_20180829174916_report - Notepad
File Edit Format View Help
Collisions - Maximum Speed - Minimum Speed - Average Speed - Speed Standard Deviation - Lane Position Standard Deviation - Lane Departures
0       42.6609 0       23.9195 12.5369 1.58537 21      4.20253 0       0       208.168 2       9       0
   *** Event Based Measures ***
Collisions - Maximum Speed - Minimum Speed - Average Speed - Speed Standard Deviation - Lane Position Standard Deviation - Lane Departures
Event #1:       0       30.3321 27.0495 28.4525 0.732502        0.146943        0       0       0       0       0
Event #2:       0       34.3638 2.60507 20.1226 10.701  0.177206        0       0       0       0       0
Event #3:       0       36.8733 2.5657e-007     23.4795 12.7476 1.76112 3       3.30751 0       0       35.5417
Event #4:       0       35.6727 4.55806 24.3114 12.3394 0.6868  0       0       0       0       0
Event #5:       0       35.1879 10.4534 21.3447 10.8399 0.266278        0       0       0       0       0
Event #6:       0       27.3134 0.00130611      18.925  8.27614 2.18798 3       17.0332 0       0       0
Event #7:       0       32.5433 28.9158 31.0785 0.979229        1.2999  0       0       0       0       0
Event #8:       0       33.9399 7.39426e-005    14.3552 14.3064 0.727131        0       0       0       0       117.932
Event #9:       0       40.4185 6.00123e-011    13.6065 17.6908 1.51345 5       11.481  0       0       0
Event #10:      0       36.435  35.7951 36.2562 0.121331        0.207522        0       0       0       0       0
Event #11:      0       34.4529 7.27605e-013    10.0574 13.9467 1.4485  1       3.7594  0       0       0
Event #12:      0       26.733  0.000739538     17.8205 10.6775 1.95131 1       5.67879 0       0       0
Event #13:      0       0       -1      0       -1      -1      0       0       0       0       0
Event #14:      0       0       -1      0       -1      -1      0       0       0       0       0
Event #15:      0       0       -1      0       -1      -1      0       0       0       0       0
Event #16:      0       0       -1      0       -1      -1      0       0       0       0       0
Event #17:      0       0       -1      0       -1      -1      0       0       0       0       0
Event #18:      0       0       -1      0       -1      -1      0       0       0       0       0
Event #19:      0       0       -1      0       -1      -1      0       0       0       0       0
Event #20:      0       0       -1      0       -1      -1      0       0       0       0       0
Windows (CRLF)    Ln 1, Col 1    100%
```

Figure 17- Driving measures txt output example

Table 5- Driving measures events description

Event Number	Maneuver Description
Event #1	Go Straight #1
Event #2	Maneuver #1 to the Right
Event #3	Maneuver #2 to the Left
Event #4	Maneuver #3 to the Right
Event #5	Maneuver #4 to the Right
Event #6	Roundabout #1 Go Right
Event #7	Go Straight #2
Event #8	Maneuver #5 to the Left
Event #9	Maneuver #6 to the Left
Event #10	Go Straight #3
Event #11	Maneuver #7 to the Right
Event #12	Maneuver #8 to the Left

4.4 SPATIAL KNOWLEDGE DATA

4.4.1 BRIEF OVERVIEW OF SIGNAL DETECTION THEORY

Developed based on visual perception research conducted by Wilson Tanner, John Swets, and David Green; signal detection theory (SDT) is a framework mostly used to understand decision-making process under uncertain conditions [70, 71]. In this research, we considered the landmark knowledge test as a decision-making process in which participants had to answer a series of “yes/no” questions (Do you remember seeing this scene during your drive?). Therefore, we could classify participants’ answers into four types of responses that composed the confusion matrix in signal detection theory (see Table 6).

1. A *hit* is the correct identification of a scene as an “on-route scene”.
2. A *false alarm* (also known as type I error) occurs when the participant identifies a scene as part of the traveled route where the presented scene is actually an off-route scene.
3. A *miss* (also known as Type II error) occurs when we present a scene to the participant that was part of the traveled route, and he fails to identify it correctly.
4. A *correct rejection* is the correct identification of a scene as an “off-route scene”.

Table 6- SDT Confusion Matrix

Participants’ response	On-route scene	Off-route scene
“yes”	Hit	False Alarm
“no”	Miss	Correct Rejection

Applying signal detection theory affords the ability to quantify participants’ likelihood of a specific answer (bias) independent from the difficulty (sensitivity) of the decision task. Here, sensitivity (also referred to as discriminability) simply refers to the ability of a person to distinguish between signal and noise (i.e., a scene that was present and a scene that was not present). Due to the fact that SDT experiments can only be used to measure which decision a person has made, rather than to measure the levels of discriminability and bias of participants, these SDT parameters need to be estimated. We estimated both parameters from the *hit rate* (HR) and *false alarm rate* (FAR) equations (see Equation 2 and Equation 3, [72]). HR can be defined as the rate of true positives (also known as sensitivity of the experiment), in other words, HR is a measure of the number of times a participant response was "yes" relative to the number of times the signal was present. On the other hand, FAR can be defined as the rate of false positives of the experiment, which is a measure of the number of times a participant response was "yes" relative the number of times the signal was not present (noise).

$$Hit\ Rate\ = \frac{Number\ of\ "yes"\ responses\ to\ signals}{Number\ of\ signal\ trials}$$

Equation 2 - Signal Detection Theory Hit Rate

$$False\ Alarm\ Rate = \frac{Number\ of\ "yes"\ responses\ to\ noise}{Number\ of\ noise\ trials}$$

Equation 3 - Signal Detection Theory False Alarm Rate

It is important to note that after calculating HR and FAR, we can also calculate the miss rate (MS) and correct rejection rate (CRR).

$$Miss\ Rate = 1 - HR$$

Equation 4 - Signal Detection Theory Miss Rate

$$Correct\ Rejection\ Rate\ = 1 - FAR$$

Equation 5 - Signal Detection Theory Correct Rejection Rate

In SDT, the decision space is assumed to the normally distributed, and therefore, participants' responses can be graphically represented, as shown in Figure 18. The y-axis represents the probability that the scene presented to the participant is either an off-route-scene/noise (top distribution) or an on-route-scene/signal. As can be observed, while the noise distribution is assumed to be random and centered at zero, the signal distribution is shifted to the right. This happens because we assumed that signals produce a stronger internal response than noise. Figure 18 also shows the *decision criterion* along the x-axis. The decision criterion can be defined as the particular level of internal response to the scene, which can change between participants or between a different scene context for the same participant. When a participant produces an internal response greater than the decision criterion, the response will be that the presented scene is an on-route scene, or a signal. On the other hand, when a participant produces an internal response that is smaller than the decision criterion, the response will be that the presented scene is an off-route scene, or a noise.

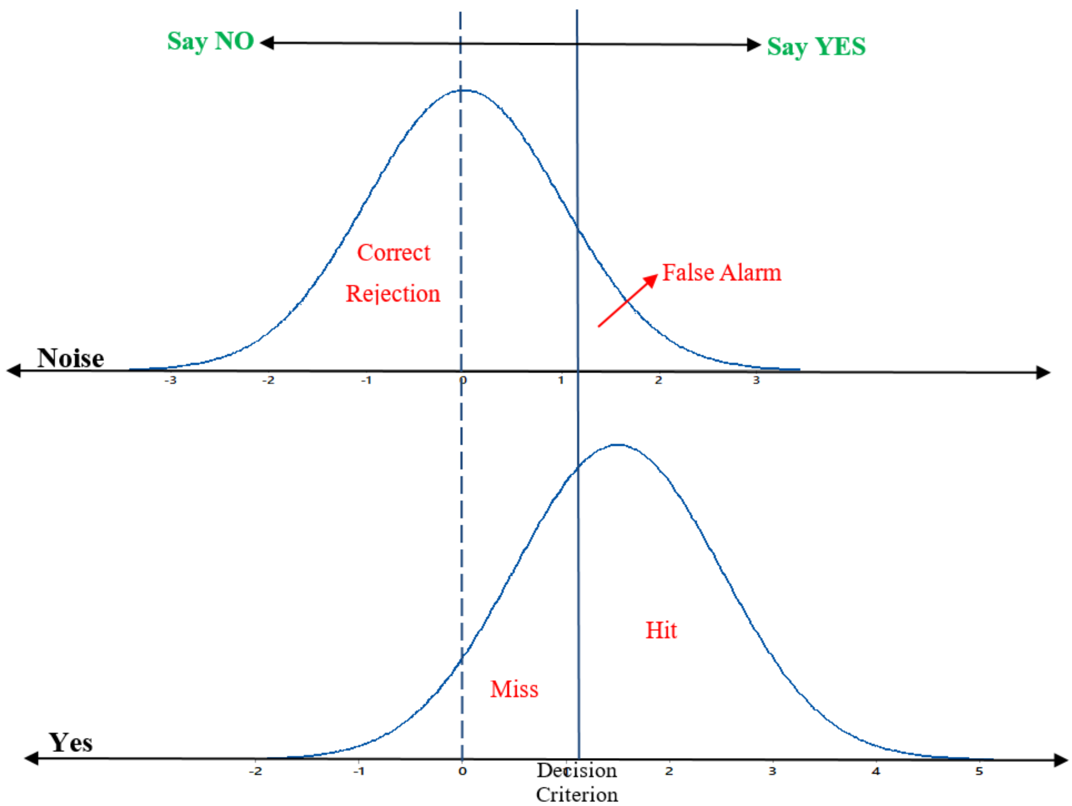


Figure 18- Representation of the decision space for SDT

It is important to emphasize that FAR and HR are influenced by both the decision criterion and the amount of overlap between signal and noise curves. When a participant changes its decision criterion to the left, both HR and FAR will increase. In the same way, a decrease in FAR will also decrease HR. When signal and noise curves don't overlap as much, changing HR will not affect as much the FAR. On the other side, curves that have too much overlap, a change in HR has a significant impact on FAR. Therefore, it is important to understand participants' ability to distinguish between a signal and noise (sensitivity). Figure 19 shows an example of two characteristics that influence sensitivity: a) the amount of separation between the two curves (i.e. mean); b) curves variance (i.e. spread). When signal/noise have means that are far apart from each other, a participant has higher sensitivity than when curves means are close to one another. In the same way, higher sensitivity is achieved when the amount of variance between curves is low.

In the same way than FAR and HR, sensitivity can also be estimated using the *d'* equation (see [72] for more details), which is estimated using z-scores from HR and FAR Gaussian models (see Equation 6).

$$d' = \frac{separation\ between\ curves}{spread\ of\ cuves} = z(HR) - Z(FAR)$$

Equation 6 - Signal Detection Theory False Sensitivity

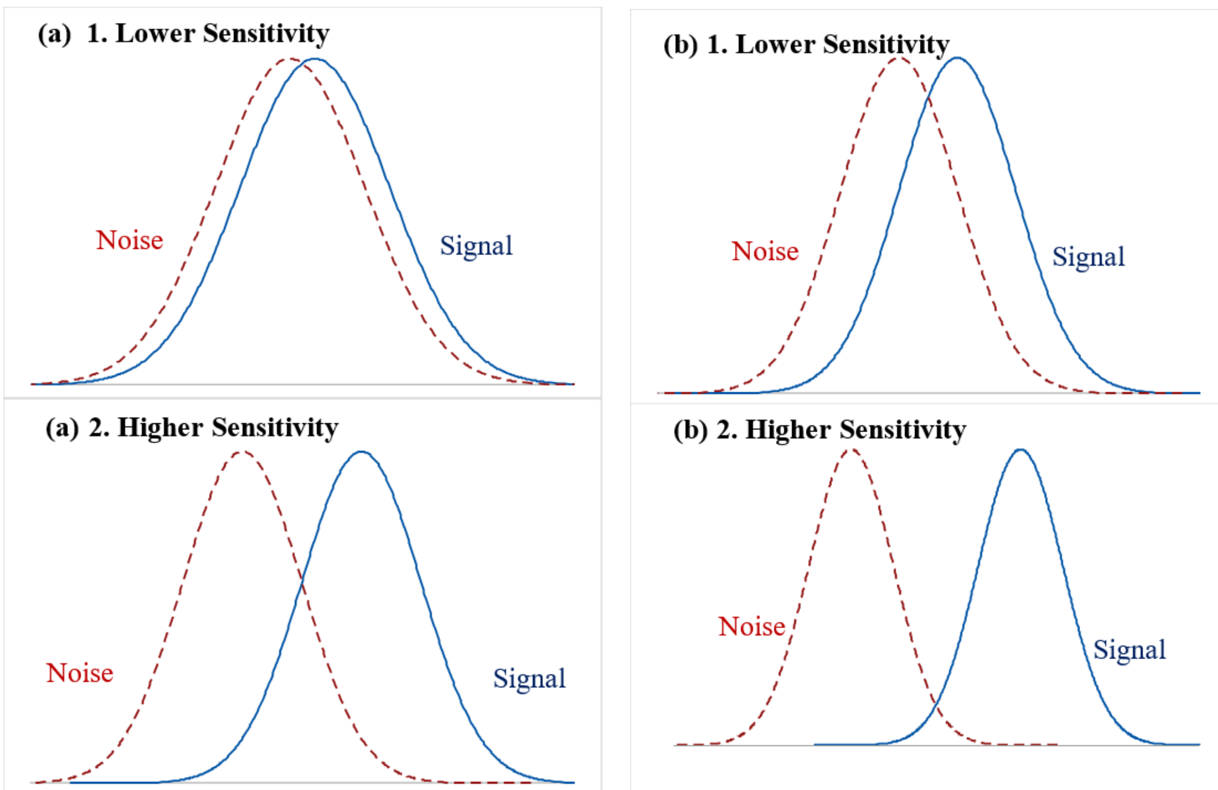


Figure 19- Examples of signal and noise probabilities that are discriminable because of (a) separation between the distribution or (b) low variance of the distributions

Even though the decision criterion is an important measure in SDT, it does not take into account how hard is to detect a signal in a specific task. For example, if a participant's sensitivity is small, the decision criterion might present a bias towards the "no" response. On the other hand, if a participant's sensitivity is large, the decision criterion might present a bias towards the "yes" response (see Figure 20).

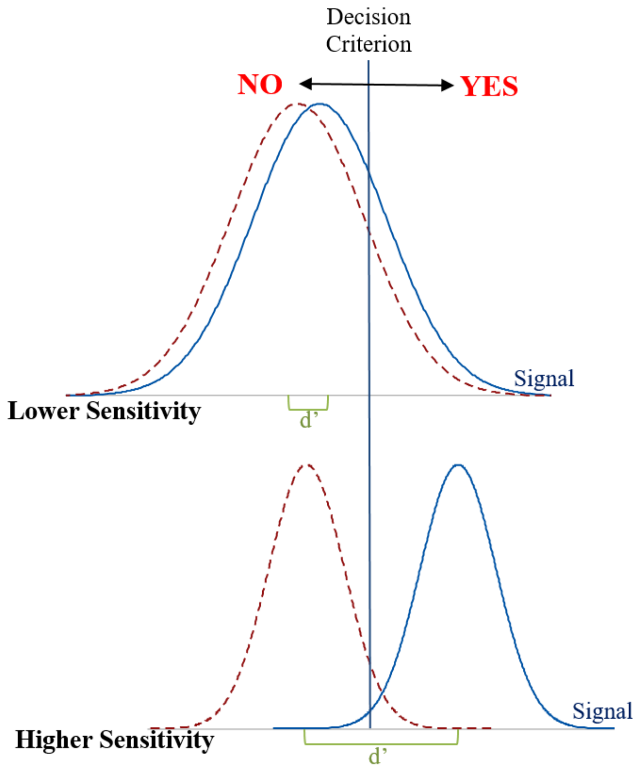


Figure 20- A given decision criterion can either present bias towards saying "yes"(bottom) or "no" (top).

In simple words, bias (c) is a response preference and it is defined as a measure of how far the decision criterion falls from the point at which the signal and noise probability distributions cross. Bias is estimated using the following equation:

$$Bias = e^{d' \, x \, (-\frac{1}{2}(\mathrm{z(FAR)} + \mathrm{z(HR)}))}$$

Equation 7 - Signal Detection Theory Bias

4.4.2 LANDMARK KNOWLEDGE TEST / SDT DATA CODIFICATION

We downloaded raw data from the landmark knowledge test into Excel where it was organized and coded. As we previously mentioned, participants were presented with a set of images (stimulus) in which they had to sort into two piles. Therefore, for the SDT analysis, the participants' response is the dependent variable. If a participant correctly placed an image into the correct pile, the response was a hit, and therefore, it received a codification of 1. If a participant did not correctly place an image into the correct pile, the response was a false alarm, and therefore it received a codification of 0. Hit rate, false alarm rate, miss rate, and correct rejection rate were calculated using Equation 2, Equation 3, Equation 4 and Equation 5 respectively; and its codification scheme can be seen in Figure 21.

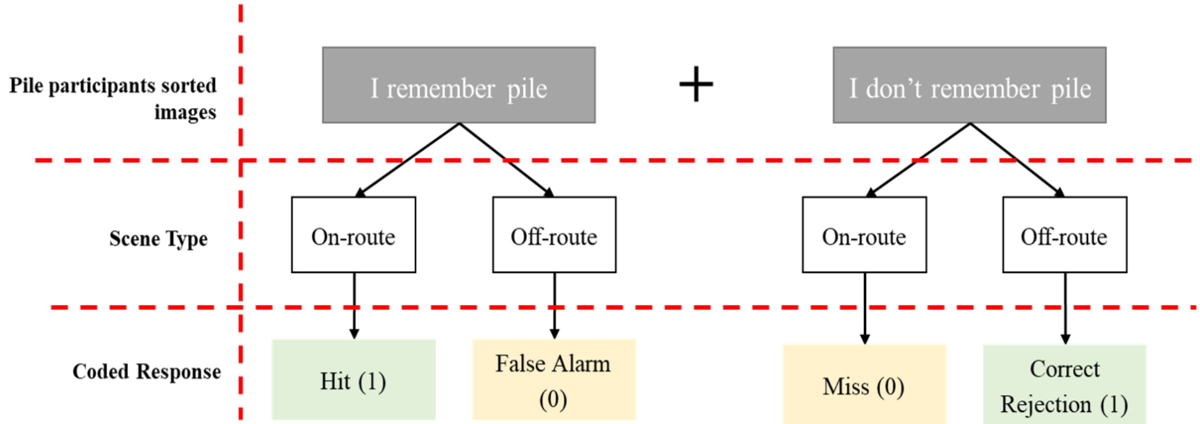


Figure 21- SDT response codification scheme

Table 7 presents an example of dataset codification for three participants in one specific scene. It is important to note that each participant response is recorded for each specific on-route and off-route scenes (n=16 images). In this sense, we decided to codify scene names according to the direction of the maneuver. The complete dataset can be seen in the appendix of this document.

Table 7- Example of response codification for three participants and one Image ID

Participant	Condition	Gender	Image ID	Direction	Pile	SDT	Binary Response
4001	Screen-relative	Male	ST	Straight	I remember	HIT	1
4002A	Screen-relative	Male	ST	Straight	I remember	HIT	1
4003	Screen-relative	Male	ST	Straight	I remember	HIT	1

The landmark test dataset was composed of 4 right turns, 4 left turns, 4 straight turns and 4 special conditions (i.e. 2 roundabout scenes and 2 start of simulation environment scenes). A total of 16*24*2 = 768 (number of stimuli * number of participants*piles) data points were recorded for each experimental condition. No data point was removed; thus, the final dataset contained a total of 768 recorded values for analysis. Minitab and Microsoft Excel software were used to carry out analysis and plot graphs.

4.4.3 LANDMARK KNOWLEDGE ANALYSIS

After coding the dataset, as explained in the previous section, we calculated sensitivity using Equation 6 and bias using Equation 7. Additionally, we were interested in investigating participants' overall performance/chance of success in the landmark test task. Thus, we defined overall performance as the correct placement of scenes in the correct pile. For this measure, we used the following equation:

$$Overall\ Performance = \frac{Hit\ score + Correct\ Rejection\ \text{score}}{2}$$

Equation 8 - Signal Detection Theory Overall Performance

Finally, we analyzed landmark knowledge by calculating the proportion of scenes recognized as having been seen before.

4.4.4 LANDMARK KNOWLEDGE REGRESSION ANALYSIS

For the signal detection theory problem, the response is binary (i.e. it has two possible outcomes), therefore, a model that models probabilities or odds needs to be fitted and so, an ordinary ANOVA analysis is not the best method for the problem. Thus, we decided to use a logistic regression approach that provided a simple way of estimating and testing the signal detection parameters. Table 8 presents the logistic regression design overview used for this work.

Table 8- Logistic Regression Overview

Method	Logistic Regression for Signal Detection Theory Parameters
Independent Variables	• Condition ○ World-relative post sign ○ Screen-relative traditional arrow • Gender ○ Male ○ Female • Direction ○ Start ○ Straight ○ Left ○ Right ○ Roundabout
Dependent Variables	Participant overall performance in the landmark test

4.4.5 LANDMARK KNOWLEDGE HEAT MAP

Figure 22 shows an example of the output from the heat map landmark test. For each of the 16 scenes presented to participants, we pre-defined areas of interest that we believe participants would select during the experiment. Table 9 shows an example of the heat map quantitative output given by Qualtrics. Complete results from both graphical and quantitative heat map output can be seen in Appendix G of this document. Because of the experimental setup, we were unable to sort data among conditions. Therefore, results from this section included data from both world-relative and screen-relative conditions altogether.

Table 9- Examples of areas of interest click count generated from heat maps

Area of Interest	%	Count
St. Mary's Hospital sign	36.36%	8
Building in the left corner	18.18%	4
Advertisement sign	4.55%	1
White building	27.27%	6
Other	13.64%	3
Total	100%	22

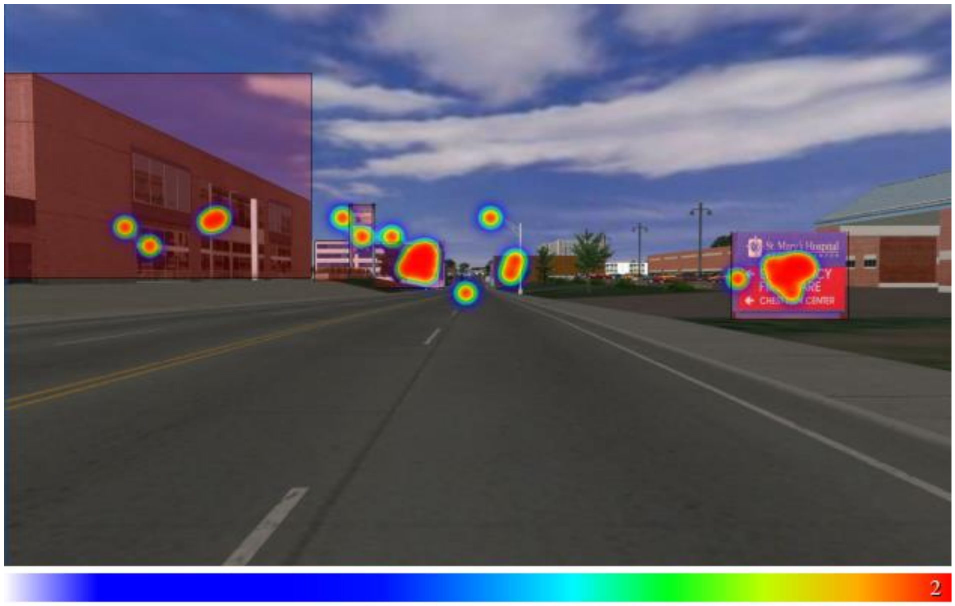


Figure 22- Heat map example from landmark knowledge test

Quantitative heat map raw data was uploaded to Excel where it was arranged and the areas of interest were manually categorized into eleven categories as follows:

1. Bank
2. Buildings
3. Fast Food
4. Fence
5. Gas Station
6. Roundabout
7. Sign (included traffic signs, advisement signs)
8. Store
9. Street Name
10. Traffic Light
11. Vehicles

4.5 GAZE BEHAVIOR

We used SMI eye-tracking glasses to collect participants' gaze behavior. Fixations, saccades, and blinks were tracked over time for each participant using SMI's iView ETG software. We imported data into SMI BeGaze version 3.7 where we used semantic gaze analysis to distill measures for two different areas of interest (AOIs): "HUD Graphic" (see Figure 23) and "Others". Glances coded as "others" included all glances participants had during the data collection besides the HUD navigation instruction (i.e. road, vehicles, landmarks, etc.).

Figure 23- HUD Graphics AOI

After performing semantic gaze analysis for some participants, we realized that in some intersections participants did not have any fixations directly at the HUD graphic. Yet, they were able to follow navigation instructions. After checking eye-tracking calibration for these cases, we concluded that calibration was not a problem, and therefore we decided to introduce a new AOI: "around HUD graphic". In this sense, the final scheme for analysis consisted of three AOIs:

A. **HUD graphic:** included all fixations in which the driver was looking directly at the AR HUD graphic.
B. **Around HUD graphic:** SMI BeGaze version 3.7 provides the option of visualizing the useful field of view of fixations. Therefore, this AOI included all fixations within the driver's field of view in which the driver was looking at locations adjacent to the HUD graphic and the HUD graphic was visible in the FOV.
C. **Others:** included any other fixation.

4.5.1 GAZE BEHAVIOR DATA PROCESSING

During the fixation coding process, data regarding 6 participants (3 from the world-relative condition and 3 from the screen-relative condition) had to be discarded. For one participant, eye-tracking data was not recorded, and the other five participants had problems with the calibration of eye-tracking glasses. In this way, none of this data could be used. Therefore, the final dataset of sematic gaze analysis was composed of only 18 participants in total.

After the manual fixation coding process in SMI BeGaze was complete, we exported data for all 18 participants into an excel file composed of 98,750 rows total of saccades, fixations, and blinks. Then, we developed a python script to calculate each glance with participant number, condition, AOI name and maneuver direction. Glance from the raw data was calculated by summing the duration of a saccade leading to an area of interest (e.g. HUD Graphics) with consecutive saccades and fixations within that same area of interest. In this analysis, blinks were excluded from the dataset.

5. RESULTS

5.1 SURVEY DATA

5.1.1 WORKLOAD NASA-TLX

5.1.1.1 MENTAL DEMAND

We found no main or interaction effects on mental demand (see Table 10). Based on the interval plots, there is no significant difference between the response of mental demand between genders and conditions (see Figure 24 and Figure 25).

Table 10- ANOVA: NASA- TLX of mental demand

Source	F Ratio	p-value
Gender	F (1,20) = 0.49	0.490
Condition	F (1,20) =0.83	0.372
Gender*Condition	F (1,20) = 1.61	0.219

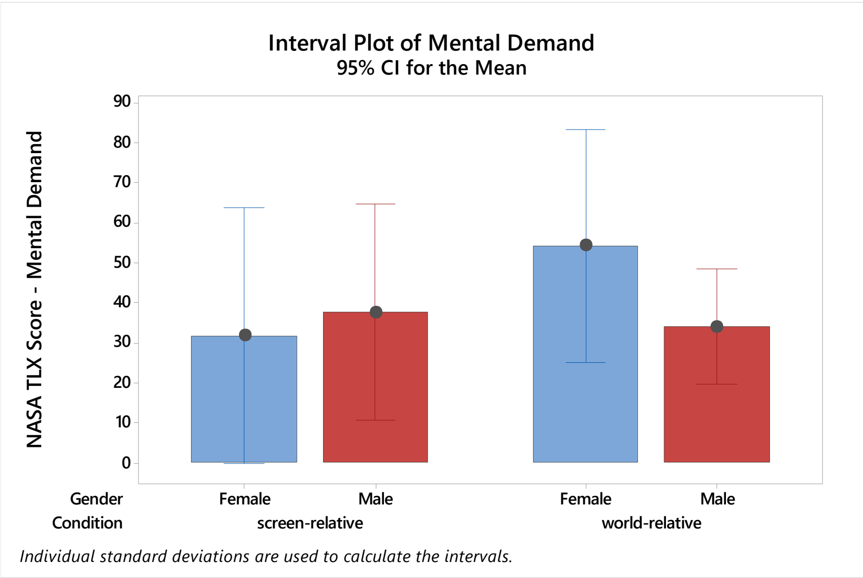


Figure 24- Interval plot of mental demand by gender and condition.

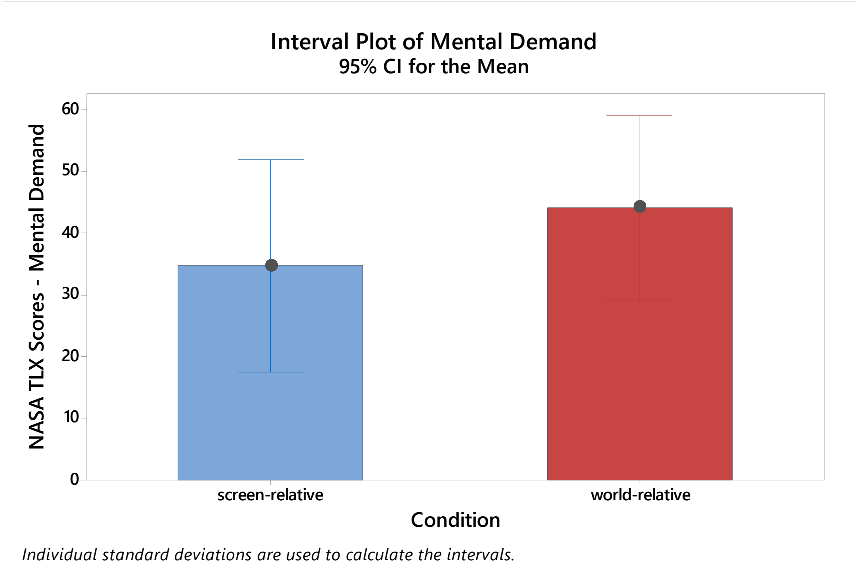


Figure 25- Interval plot of mental demand by condition.

5.1.1.2 PHYSICAL DEMAND

Checking ANOVA assumptions for the physical demand category, we observed that the data did not follow a normal distribution. After performing a box-cox transformation (see Table 11), physical demand data was found to be normal, with a relatively constant variance within the groups.

Table 11- ANOVA Box-Cox transformation of physical demand

Rounded λ	0.5
Estimated λ	0.397404
95% CI for λ	(0.0579043; 0.766904)

We found no main or interaction effects on physical demand (see Table 12). Additionally, based on the interval plots, there is no significant difference between the response of physical demand between genders and conditions (see Figure 26 and Figure 27).

Table 12- ANOVA: NASA- TLX of physical demand

Source	F Ratio	p-value
Gender	F (1,20) = 1.51	0.200
Condition	F (1,20) =0.53	0.734
Gender*Condition	F (1,20) = 0.13	0.769

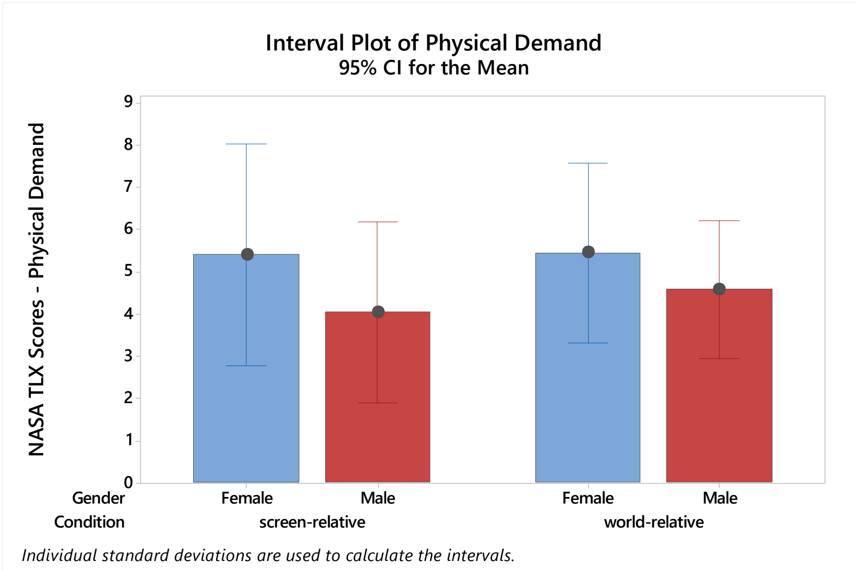


Figure 26- Interval plot of physical demand by gender and condition

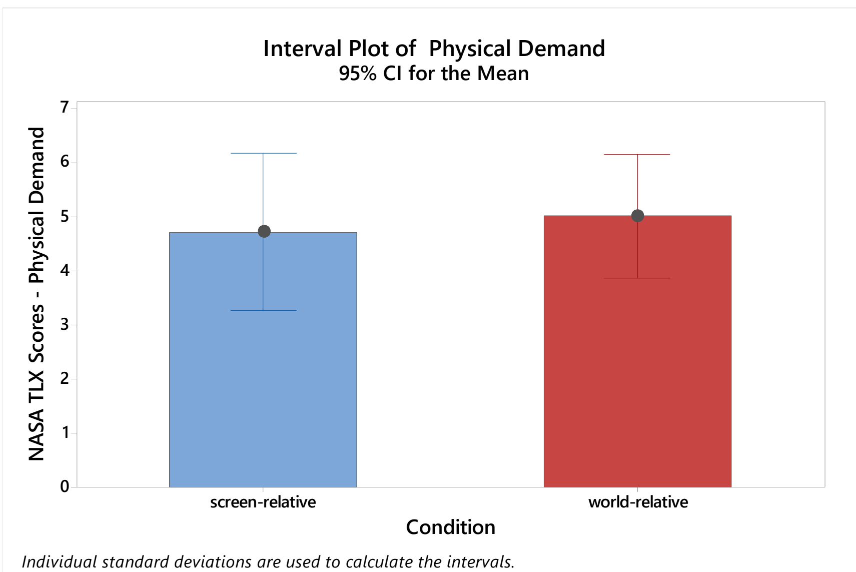


Figure 27- Interval plot of physical demand by condition

5.1.1.3 TEMPORAL DEMAND

There were no main or interaction effects on temporal demand (see Table 13). Based on the interval plots, there is no significant difference between the response of temporal demand between the genders or conditions (see Figure 28 and Figure 29).

Table 13- ANOVA: NASA- TLX of temporal demand

Source	F Ratio	p-value
Gender	F (1,20) = 0.27	0.610
Condition	F (1,20) =0.49	0.493
Gender*Condition	F (1,20) = 0.20	0.657

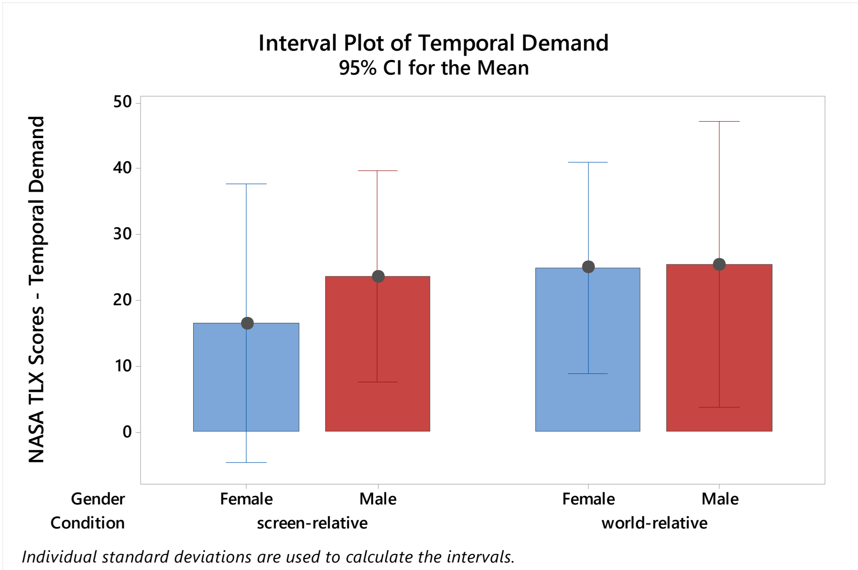


Figure 28- Interval plot of temporal demand by condition and gender

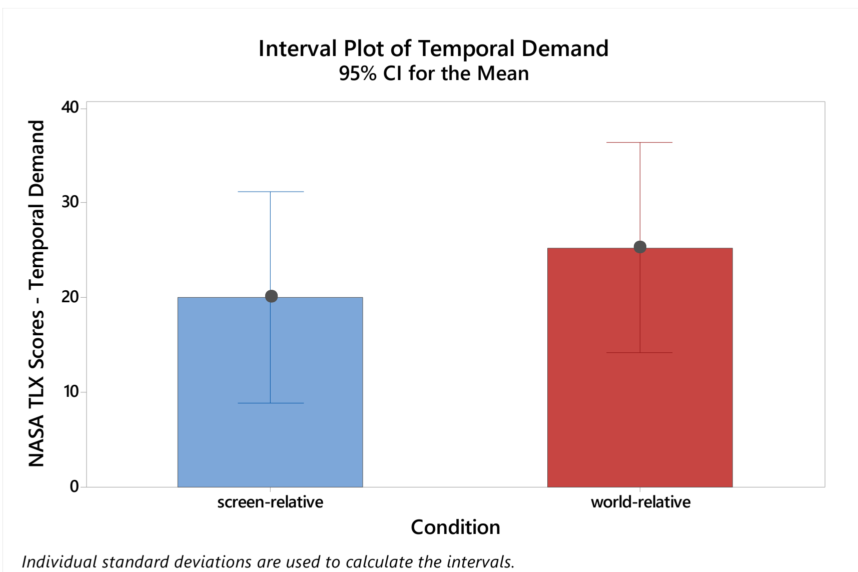


Figure 29- Interval plot of temporal demand by condition

5.1.1.4 EFFORT

There were no main or interaction effects on effort (see Table 14). Based on the interval plots, there is no significant difference between the effort required between the genders for either of the conditions (see Figure 30 and Figure 31).

Table 14 - ANOVA: NASA- TLX of effort

Source	F Ratio	p-value
Gender	F (1,20) = 0.21	0.649
Condition	F (1,20) =0.24	0.628
Gender*Condition	F (1,20) = 0.17	0.681

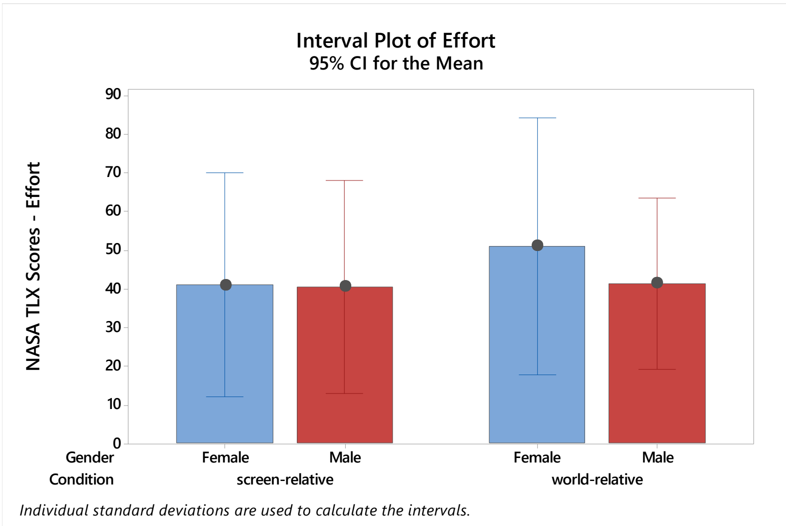


Figure 30- Interval plot of effort by condition and gender

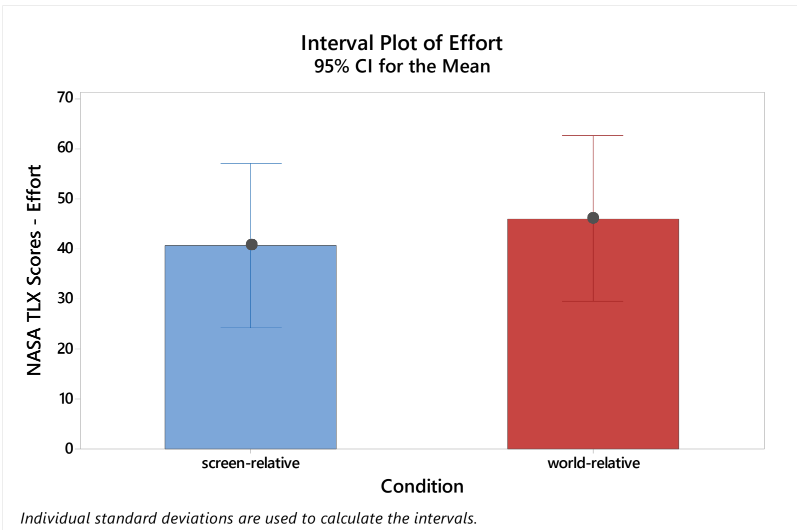


Figure 31- Interval plot of effort by condition

5.1.1.5 FRUSTRATION

There were no main or interaction effects on frustration (see Table 15). There is no significant difference in perceived frustration between the genders or conditions (see Figure 32 and Figure 33).

Table 15- ANOVA: NASA- TLX of frustration

Source	F Ratio	p-value
Gender	F (1,20) = 1.16	0.295
Condition	F (1,20) =0.30	0.590
Gender*Condition	F (1,20) = 0.26	0.618

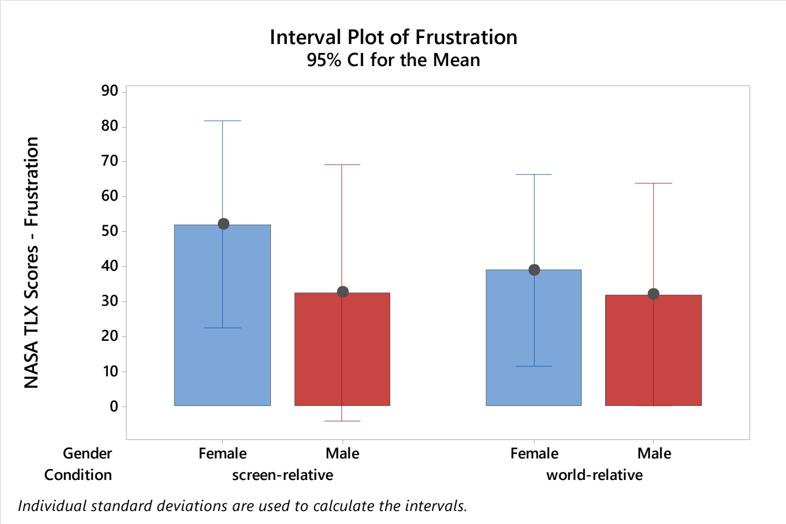


Figure 32- Interval plot of frustration by condition and gender

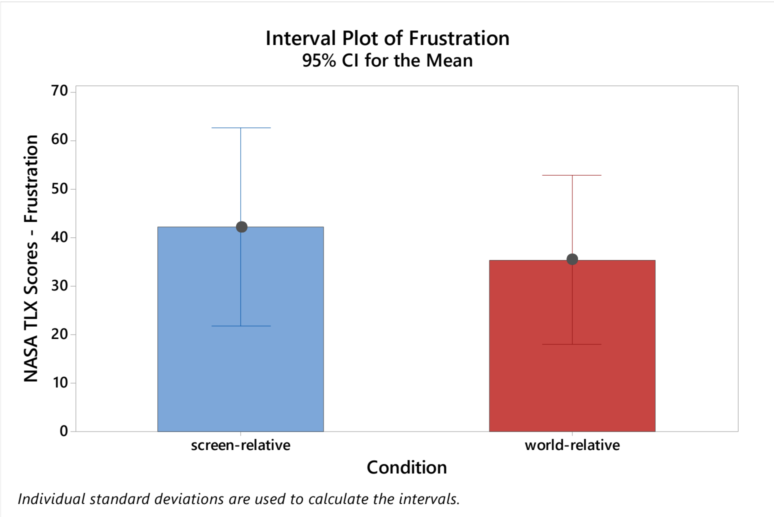


Figure 33- Interval plot of frustration by condition

5.1.1.6 PERFORMANCE

There were no main or interaction effects on performance. The performance between the genders and conditions is not significant based on ANOVA results (Table 16) and interval plots. The p-value is high compared to an alpha level of 0.05, and the plots are not different at a 95% confidence level (see Figure 34 and Figure 35).

Table 16- ANOVA: NASA- TLX of performance

Source	F Ratio	p-value
Gender	F (1,20) = 1.74	0.203
Condition	F (1,20) =0.43	0.520
Gender*Condition	F (1,20) = 0.548	0.548

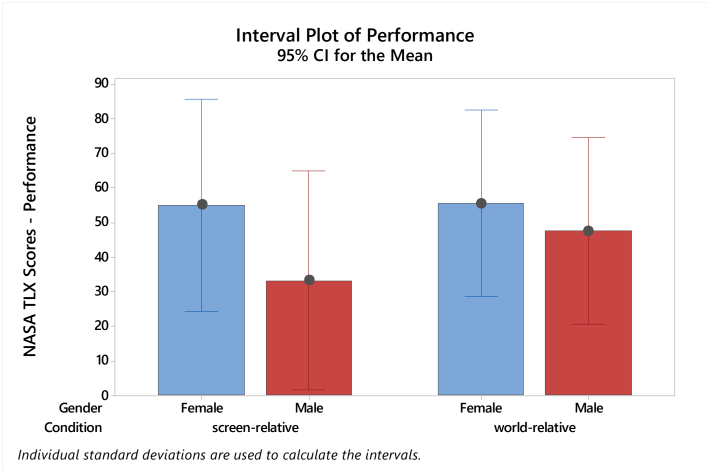


Figure 34- Interval Plot of performance by condition and gender

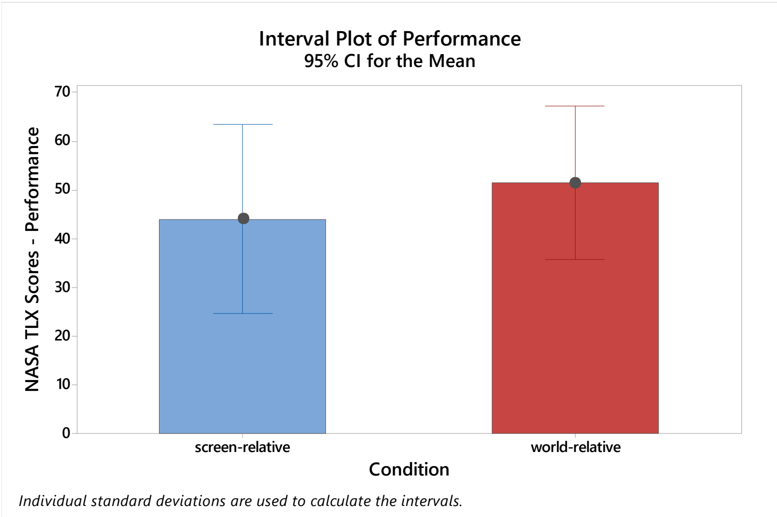


Figure 35- Interval plot of performance by condition

5.1.1.7 OVERALL WORKLOAD

Table 17 shows descriptive statistics of raw scores used to calculated average workload. Conditions that presented the highest workload score in each subscale were highlighted in blue. However, as shown in the previous section, these differences are not statistically significant.

Table 17- Descriptive statistics of raw NASA-TLX scores

Variable	Condition	Mean	StDev	Minimum	Maximum
Mental Demand	Screen-relative	34.75	27.06	0.00	84.00
	World-relative	44.17	23.46	10.00	90.00
Physical Demand	Screen-relative	27.17	22.27	1.00	79.00
	World-relative	28.08	21.64	5.00	90.00
Temporal Demand	Screen-relative	20.08	17.54	0.00	55.00
	World-relative	25.25	17.40	0.00	50.00
Effort	Screen-relative	40.75	25.74	1.00	81.00
	World-relative	46.17	26.10	10.00	93.00
Frustration	Screen-relative	42.25	32.04	1.00	86.00
	World-relative	35.50	27.34	5.00	89.00
Performance	Screen-relative	44.08	30.52	7.00	82.00
	World-relative	51.50	24.79	20.00	98.00

Table 18 shows descriptive statistics of the overall NASA-TLX workload. Again, the condition that presents the highest overall workload score is highlighted in blue.

Table 18- Descriptive statistics of NASA-TLX overall workload scores

Variable	Condition	Mean	StDev	Minimum	Maximum
Average Workload	Screen-relative	34.85	18.05	3.83	65.33
	World-relative	38.44	17.65	13.33	79.67

There were no main or interaction effects of condition and gender on overall workload (see Table 19) and therefore, the average workload between the genders and conditions is statistically similar. Additionally, p-values are high compared to an alpha level of 0.05, and the plots are not different for a 95% confidence interval (see Figure 36 and Figure 37).

Table 19- ANOVA: NASA-TLX of average workload

Source	F Ratio	p-value
Gender	F (1,20) = 0.24	0.603
Condition	F (1,20) =1.22	0.283
Gender*Condition	F (1,20) = 0.02	0.901

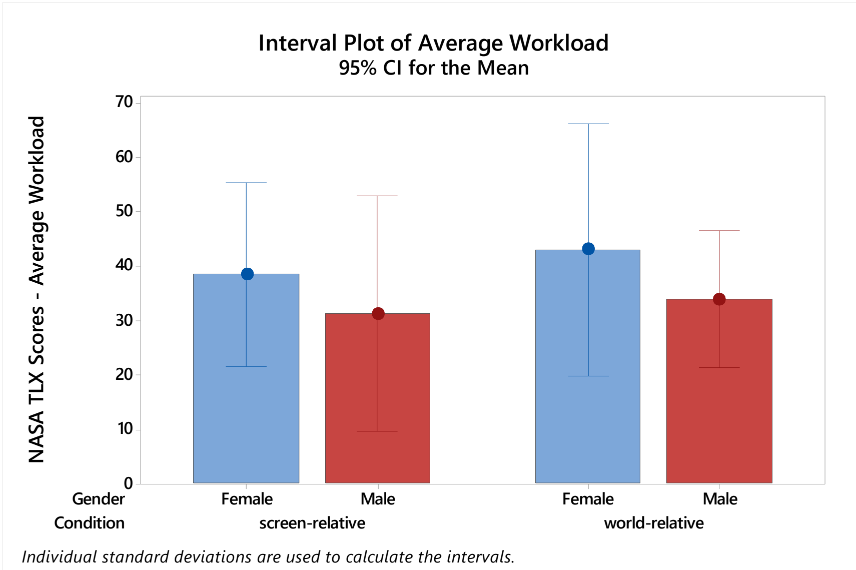


Figure 36- Interval plot of average workload by condition and gender

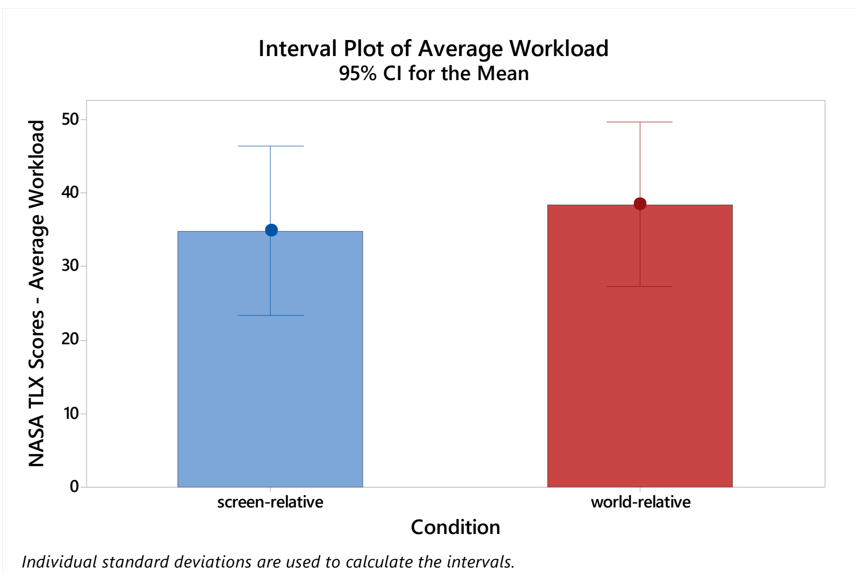


Figure 37- Interval plot of average workload by condition

5.1.2 SITUATION AWARENESS

Table 20 presents descriptive statistics for calculated demand, supply and understanding categories of the SART. Table 21 presents descriptive statistics for each of the nine SART constructs in which conditions that present the best result in each subscale are highlighted in blue. Figure 38 presents the interval plot of these categories. As we can observe, it seems that participants believe that both interfaces provide high levels of supply for attentional resources and low levels of demand for attentional resources. In fact, it seems that supply levels are statistically different from understanding and demand levels. A one paired t-test of differences between supply and understanding categories (t=16.14, $p<0.000$), between demand and understanding (t=0.48, $p<0.636$) and between supply and understanding (t=-11.71, $p<0.000$) were conducted. Results showed that in fact, the supply category is different from both demand and understanding categories.

Table 20- Descriptive statistics of SART demand, supply and understanding categories

Variable	**Condition**	**Mean**	**StDev**	**Minimum**	**Maximum**
Demand	screen-relative	8.667	3.085	5.000	14.000
	world-relative	8.083	3.423	3.000	14.000
Supply	screen-relative	18.500	3.090	13.000	23.000
	world-relative	17.833	3.010	12.000	23.000
Understanding	screen-relative	8.667	1.923	6.000	12.000
	world-relative	7.333	2.146	4.000	10.000

Table 21- SART subscales descriptive statistics

	Variable	Condition	Mean	StDev	Minimum	Maximum
Lower is Better	Instability of the Situation	Screen-relative	3.250	1.485	1.000	5.000
		World-relative	2.833	1.467	1.000	5.000
	Complexity of Situation	Screen-relative	2.167	1.115	1.000	4.000
		World-relative	2.667	1.371	1.000	5.000
	Variability of Situation	Screen-relative	3.250	1.288	2.000	5.000
		World-relative	2.583	1.311	1.000	5.000
Higher is better	Arousal	Screen-relative	4.833	1.267	2.000	7.000
		World-relative	4.833	1.403	3.000	7.000
	Concentration of Attention	Screen-relative	5.167	1.030	4.000	7.000
		World-relative	5.167	1.337	3.000	7.000
	Division of Attention	Screen-relative	4.167	1.193	2.000	5.000
		World-relative	3.750	1.138	2.000	5.000
	Spare Mental Capacity	Screen-relative	4.333	1.303	3.000	7.000
		World-relative	4.083	1.379	1.000	6.000
	Information Quantity	Screen-relative	4.333	1.303	3.000	6.000
		World-relative	4.000	1.206	2.000	6.000
	Familiarity with the Situation	Screen-relative	4.333	1.371	2.000	7.000
		World-relative	3.333	1.371	1.000	5.000

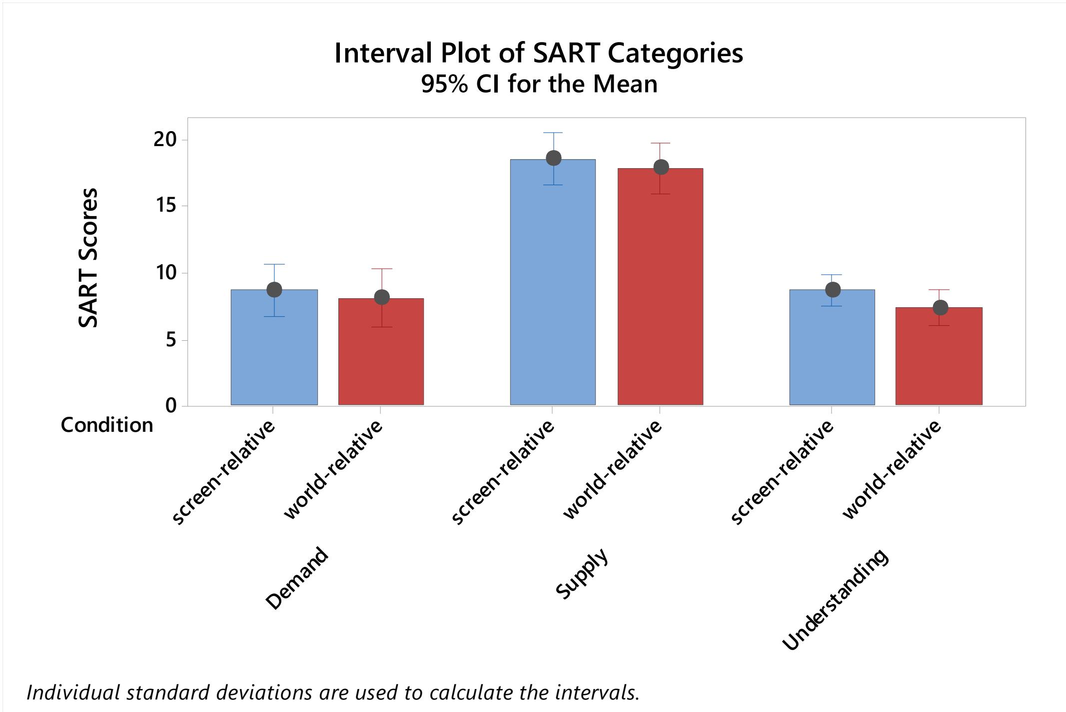


Figure 38- Interval plot of SART categories by condition

Table 22 presents descriptive statistics for overall situation awareness calculation. Again, the condition that present the best result is highlighted in blue.

Table 22- Overall SART descriptive statistics

Condition	n	Mean	St. Dev	Minimum	Q1	Median	Q3	Maximum
Screen-relative	12	18.50	4.25	10.00	15.25	18.50	22.75	23.00
World-relative	12	17.08	5.78	7.00	12.75	16.00	23.00	25.00

ANOVA results showed no main effects of driving condition or gender on the overall situational awareness, as shown in Table 23 and Figure 39).

Table 23- Overall situation awareness ANOVA

Source	F Ratio	p-value
Gender	F (1,20) = 1.83	0.190
Condition	F (1,20) = 0.49	0.493
Gender*Condition	F (1,20) = 0.13	0.722

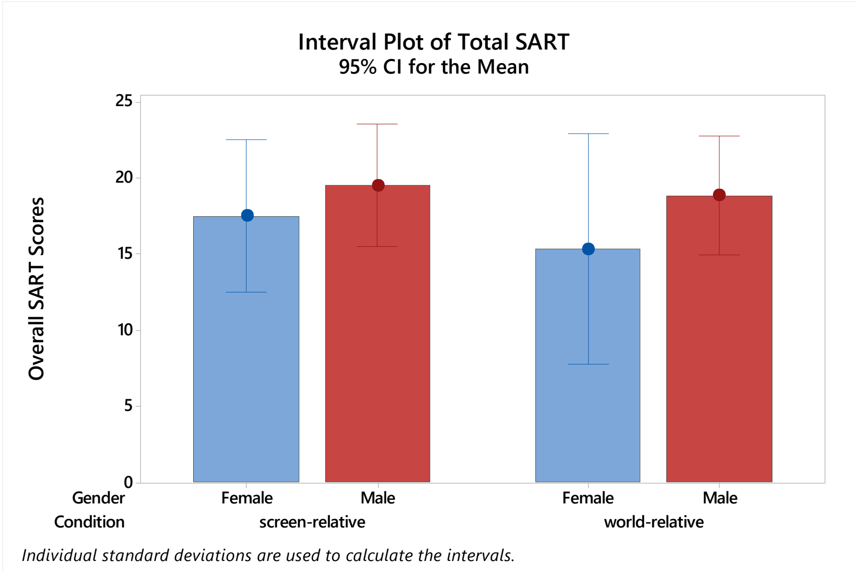


Figure 39- Interval plot of overall situation awareness

5.1.3 USABILITY QUESTIONS

5.1.3.1 EASY TO NAVIGATE

"It was easy to navigate while using this interface". As we can observed in Table 24 the mean rating of easy to navigate subscale for the world-relative condition is higher than the screen-relative condition, which means that participants perceived world-relative post sign graphics as harder to navigate compared to screen-relative traditional arrows.

Table 24- Descriptive statistics of "easy to navigate" subscale

Condition	Total Count	Mean	StDev	Minimum	Maximum
Screen-relative	12	25.83	23.92	0.00	70.00
World-Relative	12	45.00	27.47	10.00	90.00

There were no main or interaction effects between the genders and conditions on easy to navigate scores (Table 25). Although mean scores are higher for the world-relative condition, no statistical difference is found (Figure 40).

Table 25- ANOVA results of "easy to navigate" subscale

Source	F Ratio	p-value
Condition	F (1,20) = 0.02	0.090
Gender	F (1,20) =0.18	0.326
Gender*Condition	F (1,20) = 1.26	0.939

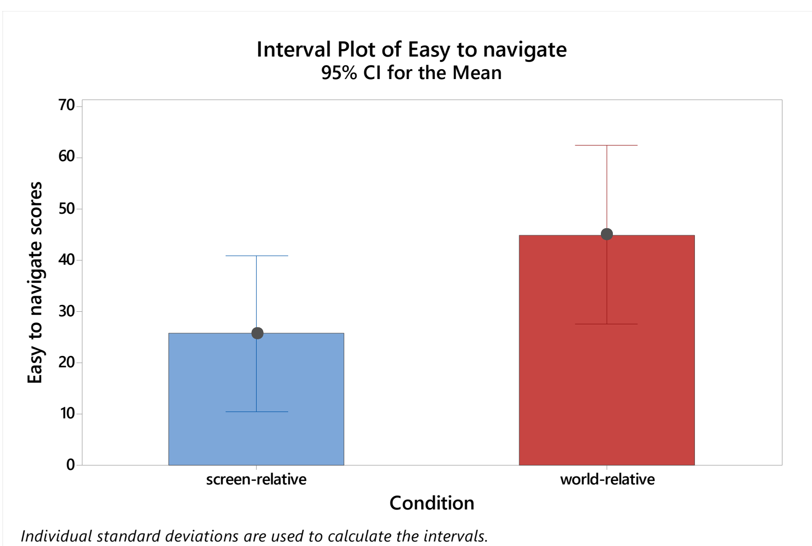


Figure 40- Interval plot of easy to navigate subscale

3.1.3.2 EASY TO VIEW

"The interface was easy to view". As we can observe in Table 26 the mean rating of easy to view scores for both conditions is about the same.

Table 26- Descriptive statistics of "easy to view" subscale

Condition	Total Count	Mean	StDev	Minimum	Maximum
Screen-relative	12	27.50	29.27	0.00	80.00
World-relative	12	28.33	16.97	0.00	60.00

There were no main effects between the genders and conditions on easy to view scores. However, there is an interaction effect of gender*condition (Table 27).

Table 27- ANOVA results of "easy to navigate" subscale

Source	F Ratio	p-value
Condition	F (1,20) = 0.02	0.0925
Gender	F (1,20) =0.18	0.304
Gender*Condition	F (1,20) = 1.26	0.017

Figure 41 shows the interaction plot for the interaction effect found. As we can observe, males perceive the screen-relative interface as easier to view compared to the world-relative interface. On the other hand, females perceive the world-relative interface as easier to view compared to the screen-relative interface.

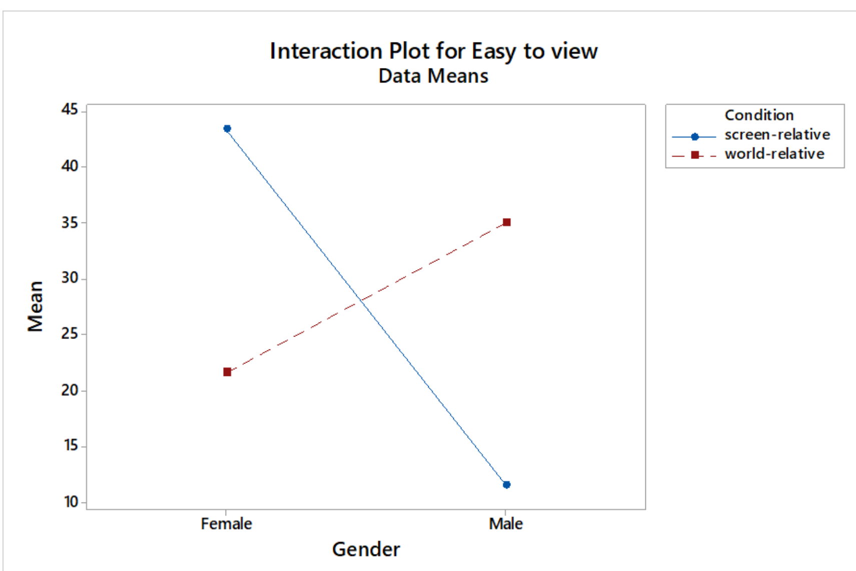


Figure 41- Interaction plot for "easy to view" subscale

3.1.3.3 TRUSTWORTHY INFORMATION

"I trusted the information that the interface was telling me". As we can observe in Table 28, the mean rating for trustworthy information scores for the world-relative condition is smaller than the screen-relative condition. This means that participants trusted more the information conveyed by the world-relative post sign interface.

Table 28- Descriptive statistics of "trustworthy information" subscale

Condition	Total Count	Mean	St. Dev	Minimum	Maximum
Screen-relative	12	21.67	26.91	0.00	70.00
World-Relative	12	12.50	17.12	0.00	50.00

Assumptions of parametric testing were not met, and logarithmic transformations were applied to raw data. However, this strategy did not correct the parametric assumption, and therefore we used the nonparametric method of Kruskal-Wallis. Because p-value is greater than the significance level of 95%, we not have enough evidence to reject the null hypothesis that the group medians are all equal (see Table 29, Table 30, Table 31 and Table 32).

Table 29- Kruskal-Wallis of "trustworthy information" subscale by condition

Method	DF	H-Value	P-value
Not adjusted for ties	1	0.48	0.488
Adjusted for ties	1	0.53	0.465

Table 30- Descriptive statistics of Kruskal-Wallis "trustworthy information" subscale by condition

Condition	N	Median	Mean Rank	Z-Value
screen-relative	12	10	13.5	0.69
world-relative	12	5	11.5	-0.69
Overall	24		12.5	

Table 31- Kruskal-Wallis of "trustworthy information" subscale by gender

Method	DF	H-Value	P-value
Not adjusted for ties	1	0.02	0.885
Adjusted for ties	1	0.02	0.879

Table 32- Descriptive statistics of Kruskal-Wallis "trustworthy information" subscale by gender

Gender	N	Median	Mean Rank	Z-Value
Female	12	5	12.3	-0.14
Male	12	10	12.7	0.14
Overall	24		12.5	

3.1.3.4 POSITIVE IMPACT ON DRIVING

"Using this interface had a positive impact on my driving". There were no main or interaction effects between the genders and conditions on the positive impact on driving scores (Table 33). As we can observe in Table 34 and Figure 42, the mean rating of positive impact on driving for both conditions is about the same (no statistical difference between them as shown in the ANOVA table) and participants to some extent agree that the interface had a positive impact on their driving.

Table 33- ANOVA results of "positive impact on driving" subscale

Source	F Ratio	p-value
Condition	F (1,20) = 0.02	0.890
Gender	F (1,20) =0.18	0.678
Gender*Condition	F (1,20) = 1.26	0.275

Table 34- Descriptive statistics of "positive impact on driving" subscale

Condition	Total Count	Mean	St. Dev	Minimum	Maximum
Screen-relative	12	42.50	28.32	0.00	80.00
World-Relative	12	40.83	29.06	0.00	90.00

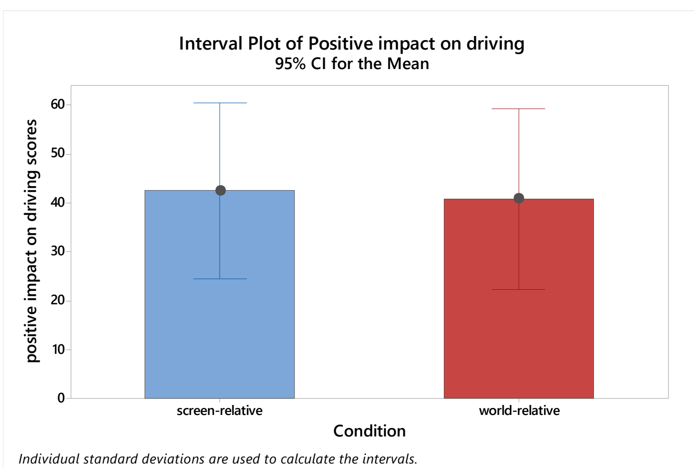


Figure 42- Interval plot of "positive impact on driving" subscale

3.1.3.5 DISTRACTION

"I did not find this interface distracting". As we can observe in Table 35, the mean rating of interface distraction for both conditions is about the same.

Table 35- Descriptive statistics of "distraction" subscale

Condition	Total Count	Mean	St. Dev	Minimum	Maximum
Screen-relative	12	42.5	37.7	0.0	100.0
World-Relative	12	37.50	30.19	0.00	90.00

ANOVA results show that there were no main or interaction effects between the genders and conditions on easy to navigate scores (Table 36). Although mean scores for this subscale is higher for the screen-relative condition (participants perceive this interface as less distracting), no statistical difference is found (Figure 43).

Table 36- ANOVA results of "distraction" subscale

Source	F Ratio	p-value
Condition	F (1,20) = 0.02	0.723
Gender	F (1,20) =0.18	0.244
Gender*Condition	F (1,20) = 1.26	0.411

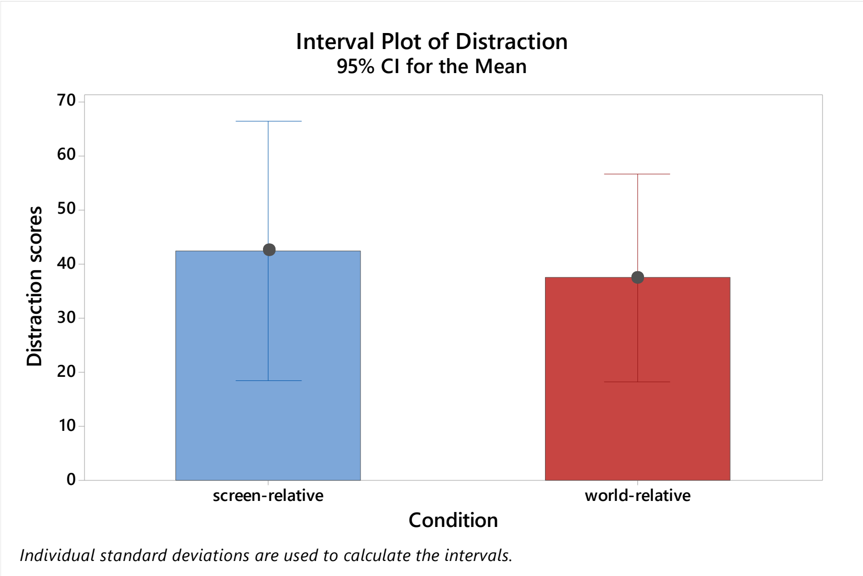


Figure 43- Interval Plot of "distraction" subscale

5.2 QUALITATIVE FEEDBACK

We sorted qualitative feedback from participants into two categories: positive and negative comments. In this section, we did not perform any further analysis.

5.2.1 SCREEN-RELATIVE QUALITATIVE FEEDBACK

Table 37- Screen-relative qualitative feedback

POSITIVE	NEGATIVE
• Nice, very descriptive. • A voice or audio indicator will be a good complement for the indicator system. Compared with usual GPS, I think this system is better because I did have to keep seeing to another place rather than the road. Which I think is safer.	• Signs are a little big. So, sometimes it was difficult to see if the traffic light was green or red.
• It seems to give basic directions similar to the callouts a GPS makes, but visual. • I think it could be helpful to have some kind of constant speed indicator in the upper right-hand corner of the display. • It feels significantly less distracting than having to look over at a phone or a GPS. • With regard to the symbols, most of them made sense but with two exceptions. • Overall though, I do like having the directions overlaid in my vision.	• The arrow to go straight makes sense at the beginning, but after that, I normally would assume you should keep driving straight to get to your destination. This is something that most GPS apps assume as well. • The second thing that seemed a little unintuitive was the symbol to tell which exit to take on the roundabout. I recognized it and knew what it meant because they use it in Waze, but if I had not seen it before while having a GPS tell me what the symbol meant as it appeared, I don't think I would have put the meaning together.
• It was cool though how the directions popped up directly in front of you in real time.	• It was a bit confusing at first because I had to think about what it is trying to tell me and look ahead down the road to understand where it wanted me to turn. • It seemed like it would get in the way most of the time when driving because it stayed there for so long, but if it flashed once and then once again right before the turn I think that'd be more helpful.
• The arrows indicating the direction to go at an intersection were certainly visible and easy to read	• The arrows seemed like they would be a bit in the way of seeing other cars if there were any in traffic going the same direction as me. • Additionally, it seemed like they popped up somewhat later than would be needed in real driving since you would need to know sooner if you have to plan a merge before making the indicated turn. • I was still able to follow the directions and make all the turns, but on some of the longer straight-aways or when approaching an intersection still a ways off found

	myself waiting to see where I'd be going or if I needed to change lanes.
	• The screen did not appear to be the most useful as it seemed like the turn could be further up than it actually was. • My suggestion would be for the arrow to shorten as you are nearing the turns.
• Simple directions	• hard to turn if you are in the wrong lane. More advanced notice would be helpful
	• a head-up cue is overlapped with the traffic light so both of them could not be seen well.
• I liked having the heads-up display directions	• I think they need to appear much sooner so that I know which lane to get into before arriving at the intersection.
	• I found that the HUD could potentially block the view of some of the traffic especially as I approached an intersection. • In terms of the actual display, I felt that the indications lasted longer than was necessary to the point of almost being distracting and making me want to look around them as if they were a physical object obscuring part of the windshield and the traffic circle indicator was confusing.
• I like that I don't need to look at a separate navigation screen for directions. • It also makes it difficult to miss turns!	• I didn't like the size/thickness of the arrow (especially the straight arrow) - it blocked the road in front of me. • I also didn't know what the traffic circle signal meant, but I figured it out
• The navigation was easy to use. • Most signals were intuitive.	• The signal for the roundabout and which turn to take was confusing. It could be redone or explained at the beginning.

5.2.2 WORLD-RELATIVE QUALITATIVE FEEDBACK

Table 38- World-relative qualitative feedback

POSITIVE	NEGATIVE
• I thought it was good, once I was able to understand it. • I did start to like the turn signals towards the end because I could understand where they were directing me, and it gave a good amount of time for me to change lanes.	• The positioning of the system was not the best. • I found it a little weird because the bottom of the glass was right in the middle of my view, which was somewhat distracting because the darkness of the glass and the bright light where the display ended was awkward to drive with. • In the beginning I was not fully aware of what the turn icon was, I thought it was telling me there was a stoplight ahead and I ended up going straight instead of turning. • It was also a little interesting at the roundabout because I saw the turn sign through the center of the roundabout and it was telling me to go straight. I later understood it when I started to go around and saw the part of the turn signal telling me where to exit.
• The ones that were telling me to turn were easy enough to understand once I realized that's what it was	• It was a little hard to understand at first. • It wasn't obvious that it was the navigation system or that it had arrows pointing where to go. • when I was to go straight it wasn't easy to pick that up, other than just that there wasn't an arrow.
• The placement of the navigation system was at a convenient viewing point.	• Understanding what the different arrows meant took a while to adjust to
• The set-up of the heads-up navigation system is pretty cool looking.	• It was initially a little hard to figure out what direction it was giving (left vs right vs straight). • I was expecting a more pronounced left or right area and the directional arrow combined with the straight road was difficult to interpret.
	• I thought it was tiring to try to use the system and drive the car at the same time. • it took a lot of concentration not to crash. • I also was not sure at first what the symbols meant the straight arrow was not as obvious of a graphic. • it took me a second to interpret what each signal left straight and right meant. all while trying to drive.
	• I was kind of confused at first, I did not know the blue markings were the navigation points, but once explained to me I understood how the system worked. • I think having a sound for the navigation signals would help overall with the system.

<table>
<tr><td></td><td>With the navigation system I was confused at first as to the directions given.
I may have missed my first turn because I didn't recognize what the blue symbol meant.
I almost asked why there was a blue F but then I realized "oh, that's probably my symbol to turn right".
Additionally, I think there were some straight directions given but as I approached an intersection I noticed there were tabs pointing in one direction or another, but it was too late to make a turn.</td></tr>
<tr><td></td><td>The HUD display is a bit smaller in area and it is interfering with the clear visibility of the road. If it is wider, it probably would cover all the vision that the driver needs while driving. Or If the display is integrated with the windshield glass, that would be aesthetically good, and the driver would be less distracted with two forms of display-- one through the windshield and the other through the HUD display.
The location of the direction displayed in the HUD display is interfering with what I have to see in the road. It might interfere less if it is moved to the left corner of the HUD display after displaying in the center for a moment.
The shape of the direction symbols is quite new and may be confusing.</td></tr>
<tr><td>Overall, I think it is a really cool idea to have directions presented to you on a screen you look through</td><td>At first, I wasn't really sure about what I was looking for and what the symbols meant, but I caught on.
I am not really a fan of the color that was used to show the turn signals and I think that the symbols were confusing with two arrows pointing in the direction rather than just one.
At first, I thought it was a pole on the side of the road or something.
Sometimes I could not tell if it was telling me to turn or to go straight and I didn't like how the symbol went away before I even reached the intersection sometimes.
Also, on the roundabout, I had no idea which way it was telling me to go. these directions and the symbology were a little bit confusing.</td></tr>
</table>

5.3 DRIVING MEASURES

For this study, we did not observe unusual incidents. There were few collisions during turns, two in total, and usually the navigation system was followed correctly. There have been some occasions when participants missed one or more turns. In this chapter, these examples will be better explained.

5.3.1 MISSED TURNS

A total of eleven participants missed at least one turn. From this total, six participants were males (55%), and five participants were females (45%). More details can be found in Table 39.

Table 39- General statistics of missed turns

Characteristic	Category	n	%
Gender	Male	6	55%
	Female	5	45%
Number of Missed Turns	1	7	64%
	2	1	9%
	3	2	18%
	4	1	9%

The driving scenario for this study had a total of twelve turns. The roundabout event presented the highest number of missed turns – seven participants missed the roundabout, encompassing 36.84% of the total. Further details can be seen in Table 4.

When sorting missed turns by the direction of the maneuver for all participants, we found that straight turns represent 26.32%, right turns represent 21.05%, and left turns represent 15.79% of the total (see Table 41). For the roundabout decision point, we understand that this point is a special case of a right turn and therefore, we decided not to merge statistics from this point with others right turns.

Table 40- Number of missed turns identified by turning point

NAME ID#	DESCRIPTION	Number Missed Turns	%
S1	Go Straight #1	1	5.26%
M1_R	Maneuver #1 to the Right	4	21.05%
M2_L	Maneuver #2 to the Left	2	10.53%
M3_R	Maneuver #3 to the Right	0	0.00%
M4_R	Maneuver #4 to the Right	0	0.00%
R1_R	Roundabout #1 go Right	7	36.84%
S2	Go Straight #2	3	15.79%
M5_L	Maneuver #5 to the Left	1	5.26%
M6_L	Maneuver #6 to the Left	0	0.00%
S3	Go Straight #3	1	5.26%
M7_R	Maneuver #7 to the Right	0	0.00%
M8_L	Maneuver #8 to the Left	0	0.00%

Table 41- Number of missed turns identified by the maneuver direction

Direction	Number of Missed Turns	%
Straight	5	26.32%
Left	3	15.79%
Right	4	21.05%
Roundabout	7	36.84%

Finally, the world-relative condition had the worst performance in terms of the number of missed turns. In fact, 78.95% of the missed turns were made during this condition (see Table 42).

Table 42- Number of missed turns by condition

Condition	Number of Missed Turns	%
Screen-relative	4	21.05%
World-relative	15	78.95%

5.3.2 STANDARD DEVIATION OF LANE POSITION (SDLP)

ANOVA results revealed that there is no significant difference between the two conditions and genders and their standard deviation of lane position (SDLP) during the events. Also, no interaction effects were found (Table 43). There is, however, a significant difference in the SDLP between the different driving maneuvers as shown in Table 43. As we can see in Figure 44, standard deviation of lane position is the highest for the roundabout maneuver.

Table 43- ANOVA of SD lane position due to gender, condition and maneuver direction

Source	F Ratio	p-value
Gender	F = 2.18	0.141
Condition	F =0.17	0.681
Maneuver Direction	**F=71.65**	**0.000**
Gender*Condition	F = 1.99	0.160
Gender*Maneuver Direction	F=1.20	0.312
Condition*Maneuver Direction	F=0.50	0.682

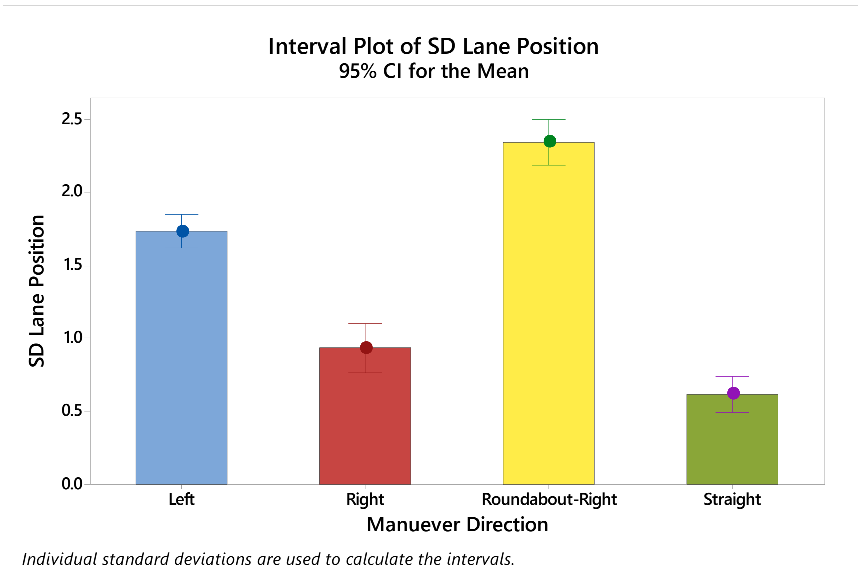


Figure 44- 95% confidence interval plot of SD lane position by maneuver direction

Post hoc Tukey showed that all SDLP from maneuvers directions were significantly different from each other (Table 44).

Table 44- Post-hoc tukey of lane position deviation

Maneuver Direction	N	Mean	Grouping			
Roundabout-Right	24	2.34248	A			
Left	96	1.73451		B		
Right	96	0.93569			C	
Straight	72	0.61933				D

5.3.3 STANDARD DEVIATION OF SPEED

ANOVA results revealed that there is no significant difference between the two conditions and genders and their speed standard deviations during the events. Also, no interaction effects were found (Table 45). There is, however, a significant difference in the speed deviation between the different driving maneuvers. As shown in the 95% confidence interval (Figure 45) standard deviation of speed is the highest for the left maneuver.

Table 45- ANOVA of SD speed due to gender, condition and maneuver direction

Source	F Ratio	p-value
Gender	F = 0.00	0.983
Condition	F =0.85	0.356
Maneuver Direction	**F=26.84**	**0.000**
Gender*Condition	F = 0.06	0.811
Gender*Maneuver Direction	F=0.21	0.889
Condition*Maneuver Direction	F=0.29	0.835

Post hoc Tukey showed that right and left maneuvers directions have statistical similar speed deviations. In the same way, the roundabout and the straight maneuver are statically similar. Finally, left/right maneuvers are statistically different from roundabout/straight maneuvers (Table 46).

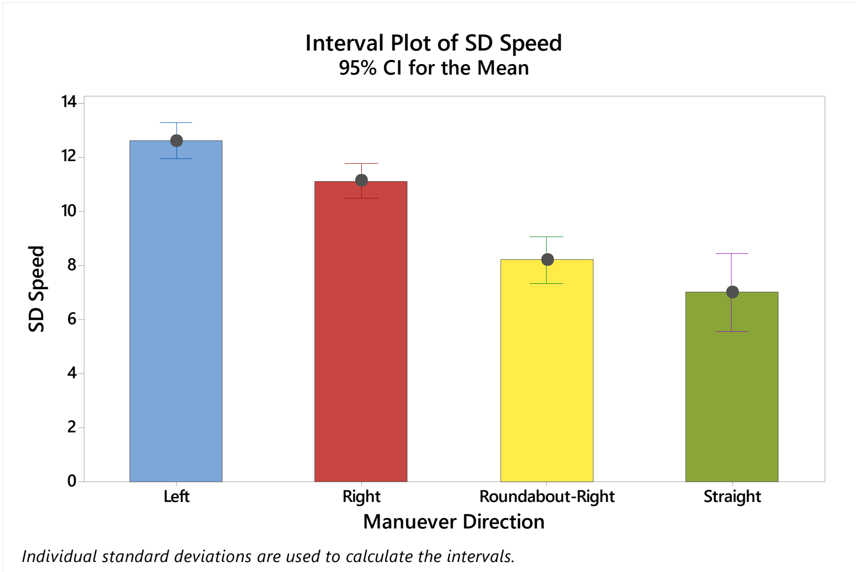


Figure 45- 95% confidence interval plot of SD speed by maneuver direction

Table 46- Post-hoc Tukey of speed deviation

Maneuver Direction	N	Mean	Grouping	
Left	96	12.6000	A	
Right	96	11.1106	A	
Roundabout-Right	24	8.1985		B
Straight	72	6.9999		B

5.4 SPATIAL KNOWLEDGE

5.4.1 LANDMARK KNOWLEDGE

5.4.1.1 INITIAL ANALYSIS

The number of scenes that were correctly placed in both “I remember pile” and “I do not remember pile”, provided a measure of the extent participants were able to recall scenes they have seen, and discard ones they have not seen before. Table 47 provides descriptive statistics for this measure. As we can observe, the mean number of correct image placement is slightly greater for the screen-relative condition.

Table 47- Descriptive Statistics of correct number of image placement in the landmark knowledge test

Variable	Condition	Mean	StDev	Minimum	Maximum
Number of scenes correctly placed	screen-relative	7.917	1.881	5.000	10.000
	world-relative	7.667	1.371	6.000	10.000

A two-sample un-paired t-test was conducted, and results showed that there is no significant difference between world-relative and screen-relative conditions (see Figure 46). This means that participants using both methods can correctly sort the same number of scenes (t (20) =-0.37; $p<0.714$).

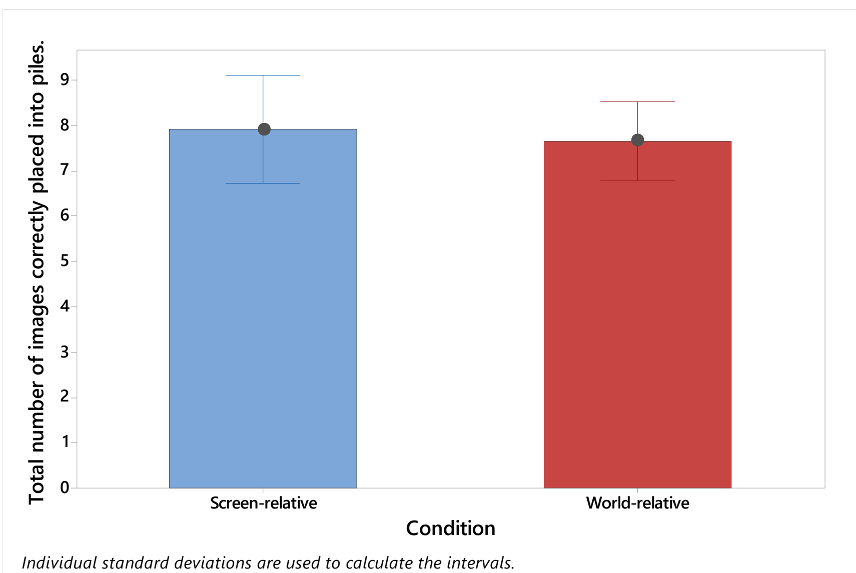


Figure 46-Interval plot bar of the total number of images correctly placed into piles. 95% CI for the mean

Table 48 shows descriptive statistics of the landmark acquisition rate. As we can observe, participants under the screen-relative condition presented a mean of 49.49% landmark acquisition rate and participants under the world-relative condition presented a mean of 47.92% landmark acquisition rate. A two-sample un-paired t-test was conducted, and results showed that there is no significant difference between conditions (t (20) =0.37; $p<0.714$).

Table 48- Descriptive statistics of landmark acquisition rate

Variable	Condition	Mean	StDev	Minimum	Maximum
Landmark Acquisition %	screen-relative	0.4948	0.1176	0.3125	0.6250
	world-relative	0.4792	0.0857	0.3750	0.6250

Table 49 shows descriptive statistics for only on-route scenes that were correctly recognized as having been seen before and Table 50 shows descriptive statistics for only off-route scenes that were correctly recognized as not seen before. As we can observe, the mean number of correctly remembered and correctly discarded scenes are slightly greater for the screen-relative condition.

Table 49- Descriptive statistics of number of correctly recognized on-route scenes

Condition	Mean	StDev	Minimum	Maximum
screen-relative	3.833	0.937	2.000	5.000
world-relative	3.667	1.073	2.000	5.000

Table 50- Descriptive statistics of number of correctly recognized off-route scenes

Condition	Mean	StDev	Minimum	Maximum
screen-relative	4.083	1.240	2.000	6.000
world-relative	4.000	1.128	2.000	6.000

A two-sample un-paired t-test was conducted, and results show that there is no significant difference between world-relative and screen-relative conditions. This means that participants using both methods remembered the same number of on-route of scenes (t (21) =0.41; $p<0.689$) and the same number of off-route of scenes (t (21) =0.17; $p<0.865$). Figure 47 provides a graphical representation of these results.

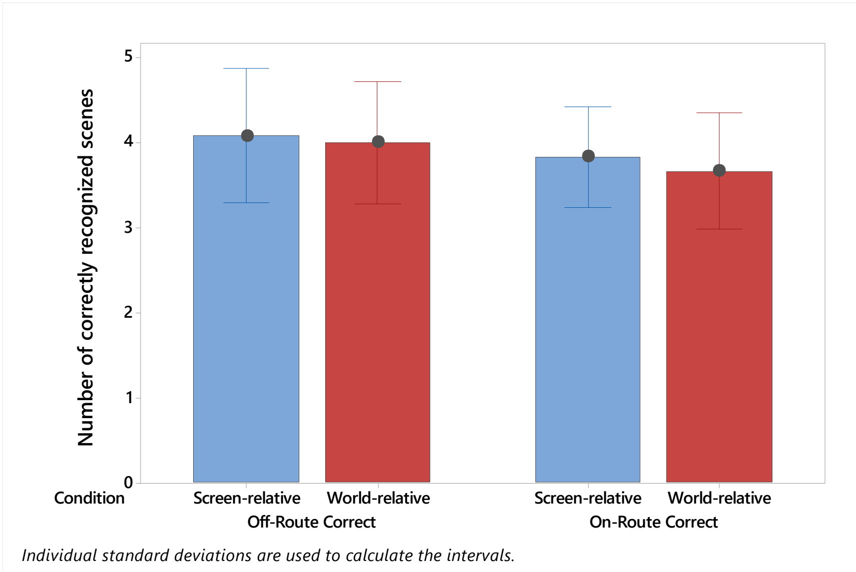


Figure 47- Interval Plot of the number of correctly recognized scenes by condition and on/off routes scenes. 95% ci for the mean

5.4.1.2 SIGNAL DETECTION THEORY

Table 51 presents the final hit, false alarm, miss, and correct rejection scores for each participant. Figure 48 presents the interval plot for calculated signal detection theory parameter rates for all participants by condition.

Table 51- Hit, False alarm, Miss and Correct Rejection scores for each participant

Participant	Correct Rejection	False Alarm	Hit	Miss
4001	5	3	5	3
4003	3	5	2	6
4004	3	5	3	5
4005	2	6	3	5
4006	6	2	4	4
4007	3	5	4	4
4008	3	5	4	4
4009	5	3	4	4
4011	5	3	3	5
4012	5	3	5	3
4013	6	2	4	4
4014	5	3	3	5
4015	4	4	3	5
4016	5	3	5	3
4017	4	4	4	4
4018	4	4	5	3
4019	3	5	4	4
4021	3	5	3	5
4022	2	6	5	3
4023	3	5	4	4
4024	5	3	2	6
4002A	5	3	4	4
4010A	5	3	5	3
4020B	4	4	2	6
Total	**98**	**94**	**90**	**102**

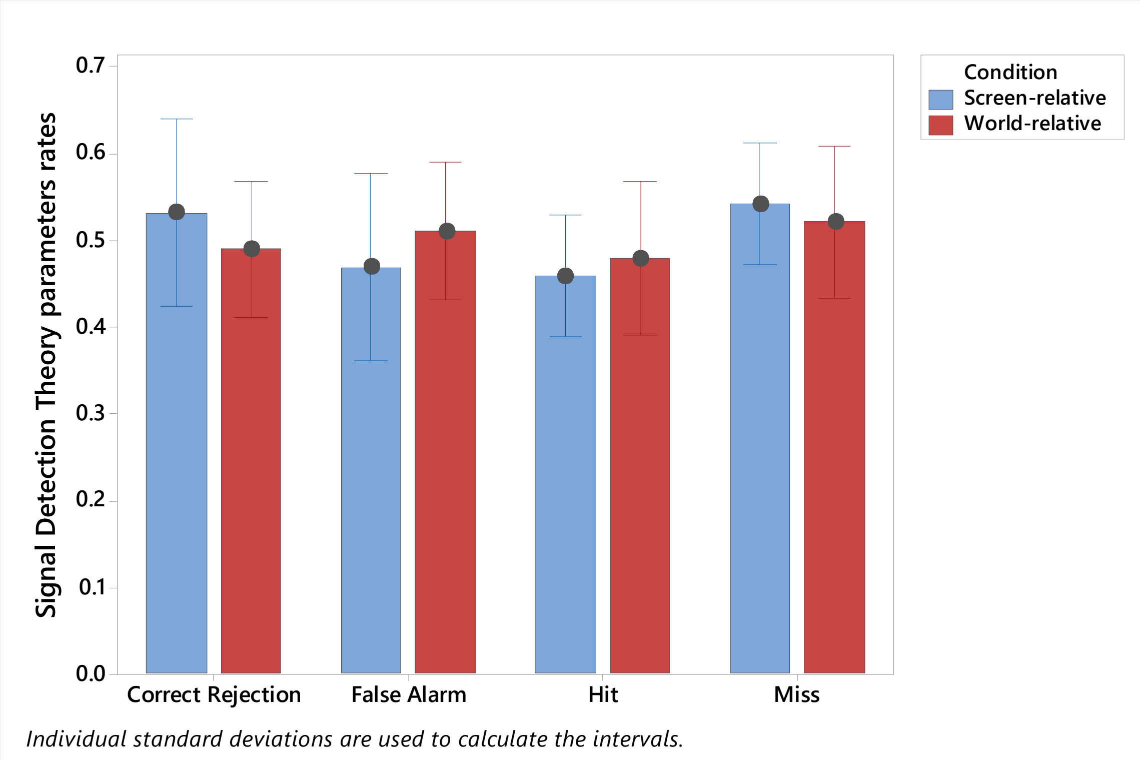


Figure 48- Interval Plot of SDT parameter rates for all participants by condition. 95% ci for the mean

Table 52 presents descriptive statistics of overall signal detection theory parameters of the whole population in both conditions.

Table 52- Descriptive statistics of signal detection theory parameters

Variable	Mean	StDev	Minimum	Maximum
Correct Rejection	0.5104	0.1471	0.2500	0.7500
False Alarm	0.4896	0.1471	0.2500	0.7500
Hit	0.4688	0.1236	0.2500	0.6250
Miss	0.5313	0.1236	0.3750	0.7500

A logistic regression model was fitted for all independent variables. Deviance and Hosmer-Lemeshow Goodness of Fit tests for Chi-Square were carried out and results are shown in Table 53.

Table 53- Goodness of Fit test for Chi-Square for model

Test	DF	Chi-Square	P-Value
Deviance	377	383.97	0.391
Hosmer-Lemeshow	2	1.94	0.925

At a significance level of 95%, $\alpha = 0.05$. Since the p-value for all two tests is greater than the significance level (0.05), the null hypothesis is failed to be rejected and therefore it is concluded that there is no significant difference between the observed and expected value. In other words, the predicted probabilities do not deviate from the observed probabilities, and so the logistic function is a good fit. Table 54 shows the results for the Deviance Analysis, also known as Binomial ANOVA. The output from this ANOVA is similar to the parametric option, however rather than the sum-of-squares and means sum-of-squares displayed, now the resulting deviance from each of the parameters are shown. For the deviance analysis, p-values were calculated using the chi-square distribution. At a significance level of 95%, α=0.05. Since the p-value for maneuver direction is smaller than the significance level (0.05), the null hypothesis is rejected. Therefore, it is concluded that the maneuver direction has a significant effect on the probability of achieving success (hit and correct rejection).

Table 54- Deviance Analysis Results

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	6	10.034	1.6724	10.03	0.123
Condition	1	0.175	0.1746	0.17	0.676
Gender	1	0.385	0.3852	0.39	0.535
Direction	**4**	**9.492**	**2.3729**	**9.49**	**0.050**
Error	377	522.136	1.3850		
Total	383	532.170			

Figure 49 shows the main effect plot for maneuver. As we can observe, when the direction of the turning if left, the overall performance (i.e. the number of hits and number of correct rejections) was worst. The best performance was achieved when the turning direction was straight, followed by the start of the simulation scenario.

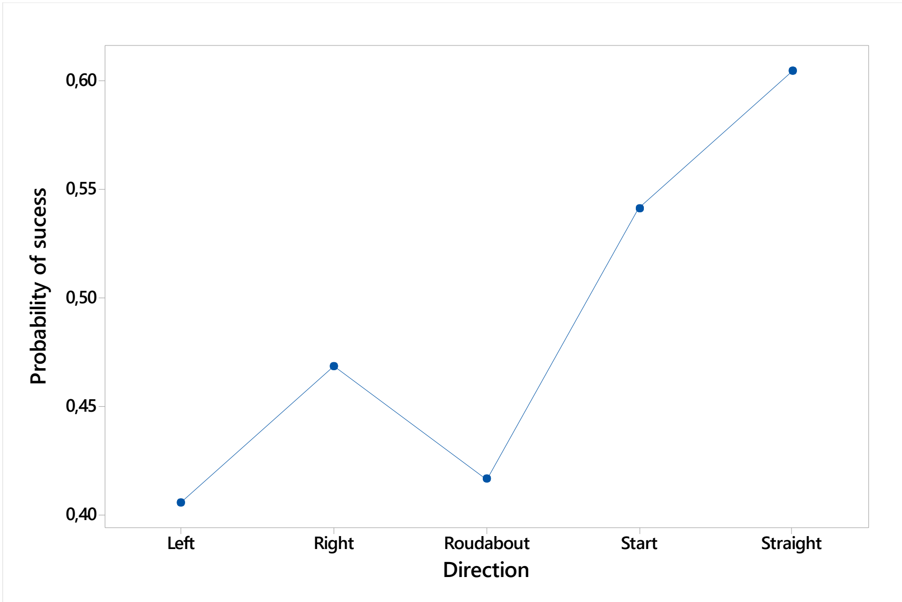


Figure 49- Main effects plot for hit by maneuver direction

Same results were achieved when calculating SDT matrix rates for the direction of turning. As we can observe in Table 55 success rate is the highest for the straight turning direction. As we prove by the logistic regression analysis, this difference is statistically significant.

Table 55- HIT, False Alarm, Misses and Correct Rejection Rates for each turning direction

	Correct Rejection	False Alarm	HIT	Miss	Overall Performance Rate
Left	33%	67%	48%	52%	40.63%
Right	73%	27%	21%	79%	46.88%
Roundabout	13%	88%	71%	29%	41.67%
Start	75%	25%	33%	67%	54.17%
Straight	**54%**	**46%**	**67%**	**33%**	**60.42%**

Table 56 shows the odd ratios for categorical predictors used for the logistic regression analysis. This analysis compared the odds of an event happening at two different levels of the predictor. In this case, level B is the reference level for the factor. When odd ratios are greater than 1, it indicates that an event (in this case, achieving success in the landmark knowledge test) is less likely at level B. On the other hand, odds ratios that are less than 1 indicate that an event is more likely at level B. When comparing experimental conditions, we can observe that the screen-relative condition is more likely to achieve success in the landmark knowledge test, as previously shown in section 5.4.1.1. However, based on the 95% CI, this difference is not statistically significant.

When comparing maneuver direction, we can observe that achieving success in the landmark knowledge is more likely in the straight maneuver compared to the roundabout and left maneuver. Based on the 95% CI this result is statistically significant.

Table 56- Odds Ratios for Categorical Predictors

Level A	Level B	Odds Ratio	95% CI
Condition			
World-relative	Screen-relative	0.9181	(0.6122, 1.3768)
Direction			
Right	Left	1.2897	(0.7282, 2.2843)
Roundabout	Left	1.0440	(0.5164, 2.1104)
Start	Left	1.7277	(0.8589, 3.4754)
Straight	**Left**	**2.2316**	**(1.2527, 3.9753)**
Roundabout	Right	0.8094	(0.4019, 1.6302)
Start	Right	1.3396	(0.6684, 2.6846)
Straight	Right	1.7303	(0.9757, 3.0683)
Start	Roundabout	1.6549	(0.7381, 3.7105)
Straight	**Roundabout**	**2.1376**	**(1.0563, 4.3255)**
Straight	Start	1.2916	(0.6415, 2.6009)

Odds ratio for level A relative to level B

Sensitivity and bias parameters were calculated using equations 6 and 7, respectively. Table 57 presents descriptive statistics for these parameters.

Table 57- Descriptive statistics for SDT parameters

Variable	Condition	Mean	StDev	Minimum	Maximum
Sensitivity	screen-relative	-0.166	0.568	-0.993	0.674
	world-relative	0.050	0.517	-0.674	0.674
Bias	screen-relative	1.0321	0.1192	0.8380	1.2554
	world-relative	1.0165	0.1161	0.7965	1.2554

ANOVA results showed that there is an interaction effect of condition*gender on sensitivity (see Table 63). Interaction plot (Figure 50) revealed that female participants present better sensitivity in the screen-relative condition. On the other side, male participants present better sensitivity in the world-relative condition.

Table 58- ANOVA results for SDT sensitivity

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Condition	1	0.2787	0.2787	1.10	0.306
Gender	1	0.2950	0.2950	1.17	0.293
Condition*Gender	**1**	**1.1309**	**1.1309**	**4.47**	**0.047**
Error	20	5.0550	0.2527		
Total	23	6.7596			

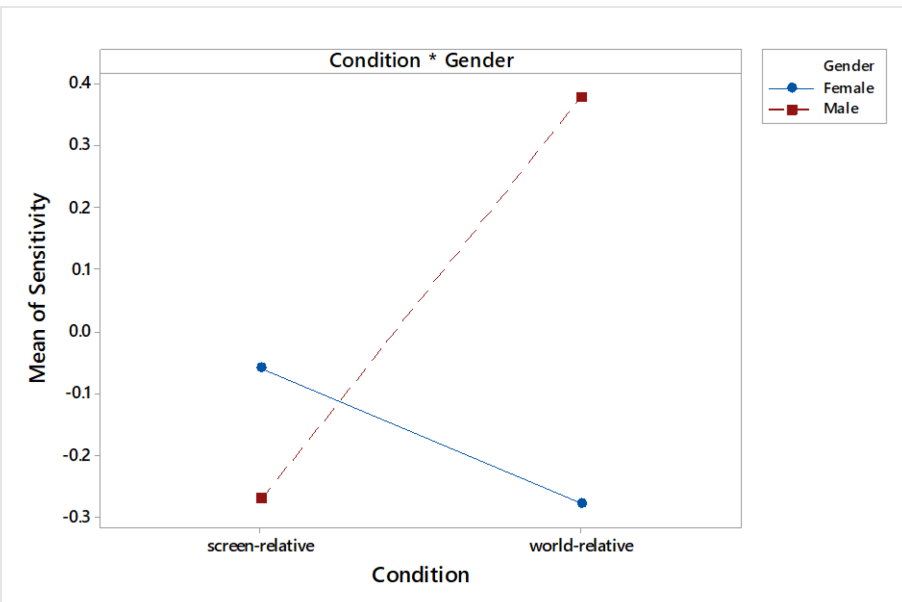


Figure 50- Interaction plot for sensitivity

ANOVA results showed no main or interaction effects of condition and gender on SDT bias (see Table 59).

Table 59- ANOVA results for SDT bias

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Condition	1	0.001454	0.001454	0.10	0.756
Gender	1	0.010667	0.010667	0.73	0.404
Condition*Gender	1	0.000260	0.000260	0.02	0.896
Error	20	0.293656	0.014683		
Total	23	0.306037			

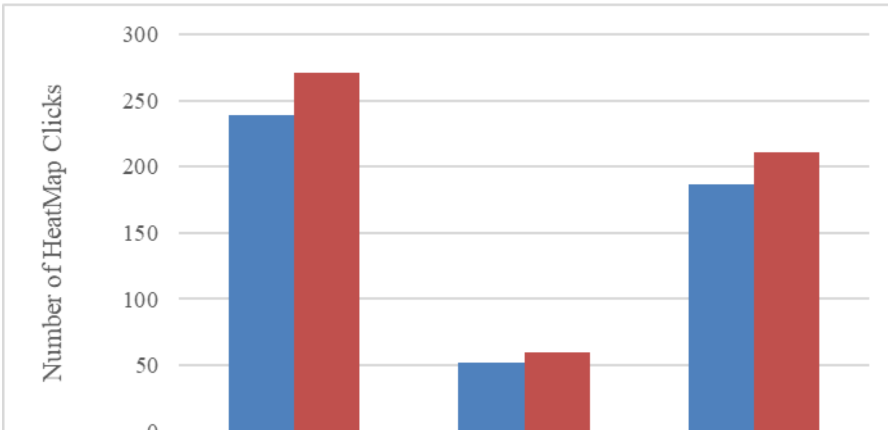
Additionally, sensitivity and bias parameters for maneuver direction were calculated and results are shown in Table 60.

Table 60- Sensitivity and Bias for each turning direction

	Sensitivity	Bias
Left	-0.2086073	1.250532833
Right	-0.1319023	1.228310469
Roundabout	-0.1199313	1.16729642
Start	0.13106747	0.790132825
Straight	0.2350867	0.774869371

5.4.1.3 HEAT MAPS

This phase consisted of eight on-route and eight off-route scenes. Participants were allowed to click up to ten times on scenes they had sorted into the "I remember pile". There was a total of 510 clicks for both conditions. 239 of these clicks were in on-routes scenes and 271 clicks were in off-routes scenes (Figure 51). From this total, 79% of clicks were made in pre-selected areas of interest (landmarks) and 22% were made in other locations (Figure 52).



	Total	Other	Landmark
On-Route	239	52	187
Off-Route	271	60	211

Figure 51-Bar plot of number of clicks

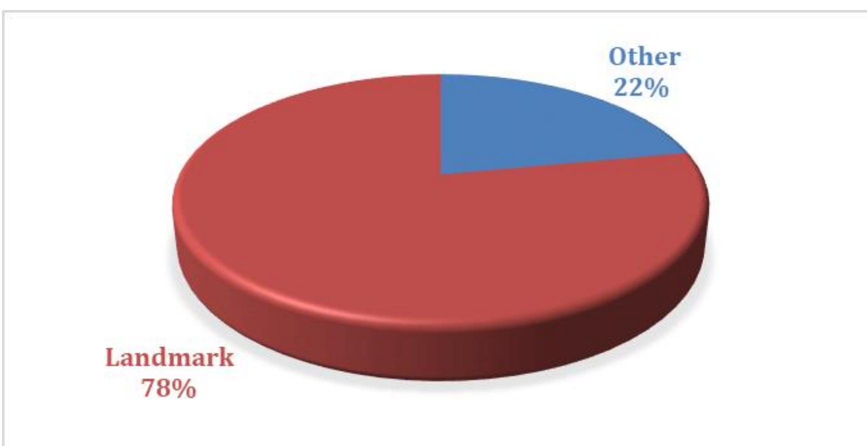


Figure 52- Pie chart of the number of click's by selected area

Table 61 shows the total number of clicks and their corresponding percentages for each landmark category. Figure 53 shows a graphical representation of this data.

Table 61- Clicked landmark count and percentage

Category	Total Clicks	%
Bank	12	3.02%
Building	182	45.73%
Fast Food	17	4.27%
Fence	3	0.75%
Gas Station	16	4.02%
Roundabout	26	6.53%
Sign	92	23.12%
Store	18	4.52%
Street Name	18	4.52%
Traffic Light	11	2.76%
Vehicle	3	0.75%
Grand Total	**398**	

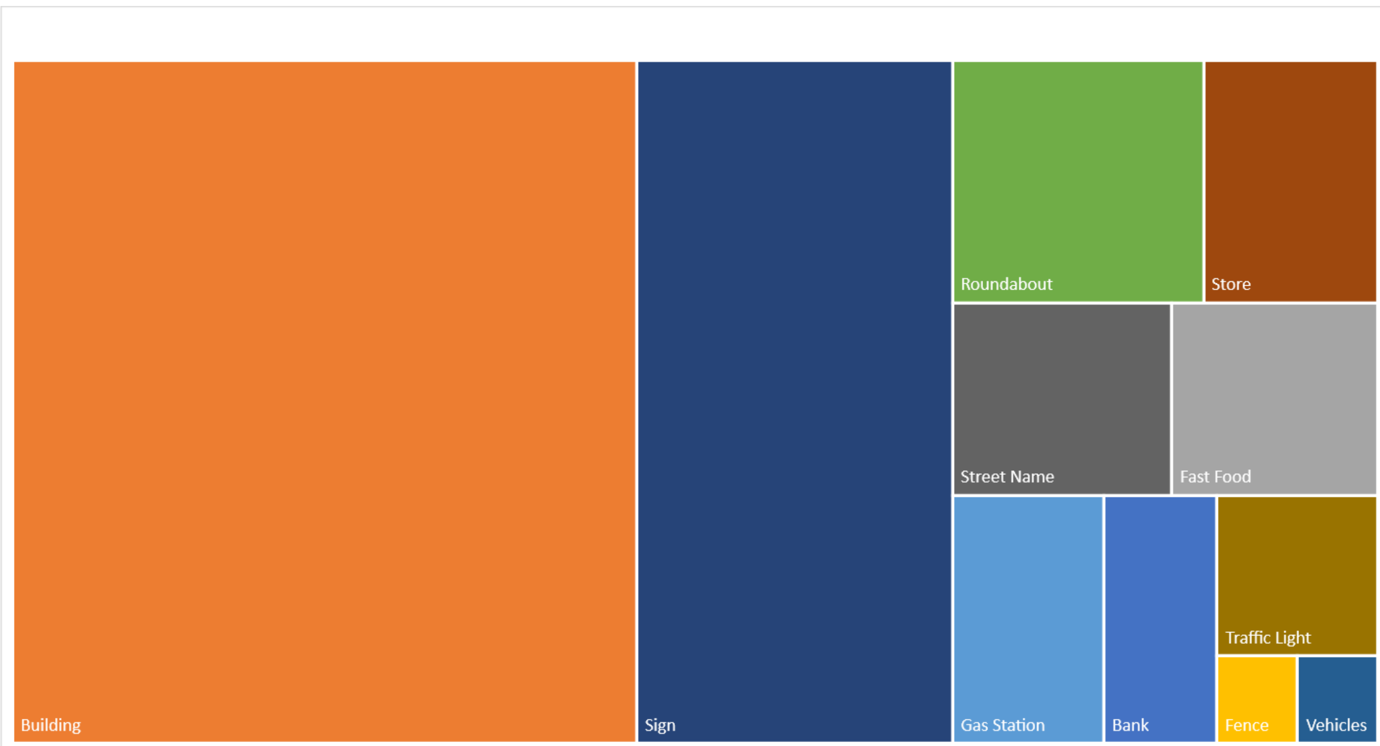


Figure 53- Tree map of clicked landmarks

5.4.2 ROUTE KNOWLEDGE

Table 62 compares the proportion of on-route scenes that were correctly sequenced (route knowledge test) and the proportion of scenes correctly recognized (landmark knowledge test) by each participant. Clearly, participants were generally poor at the route knowledge test. In fact, a paired t-test showed that landmark knowledge acquisition proportion is statistically different from route knowledge acquisition proportion (t= -10.01, $p<0.000$). Same results can be graphically seen on Figure 54.

Regarding the route knowledge test, a two-sample un-paired t-test revealed that subjects using both screen-relative and world-relative methods remembered the same proportion of scenes in the correct order location (t (21) =-1.17; $p<0.256$).

Table 62- Descriptive statistics of landmark/route knowledge comparison

Variable	Condition	Mean	StDev	Min	Max
Landmark Testing	Screen-relative	0.4948	0.1176	0.3125	0.6250
	World-relative	0.4792	0.0857	0.3750	0.6250
Route Testing	Screen-relative	0.1515	0.1050	0.0000	0.2727
	World-relative	0.2045	0.1171	0.0000	0.4545

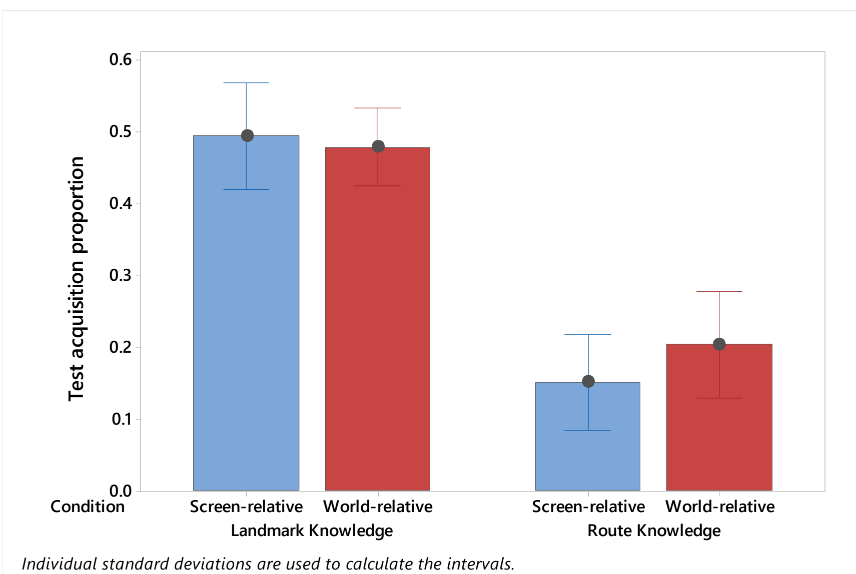


Figure 54- Interval plot of proportion of landmark and route knowledge acquisition by condition

As we mentioned before, there were 11 images for the route knowledge test. In this sense, if a participant puts image A in position 10 and its correct position is 7, the deviation from optimal position would be -3. Table 63 presents the descriptive statistics of this measure. As we can observe, performance is slightly better for the world-relative condition. However, a two-sample un-paired t-test revealed that this difference is not significant (t (261) =0.64; *p<0.523*).

*Table **63**- Descriptive statistics of deviation from the optimal position of route knowledge*

Variable	Condition	Mean	StDev	Minimum	Maximum
Deviation from the optimal position	Screen-relative	3.091	2.564	0.000	9.000
	World-relative	2.894	2.441	0.000	8.000

Figure 55 shows the interval plot of deviation of optimal position by condition and gender.

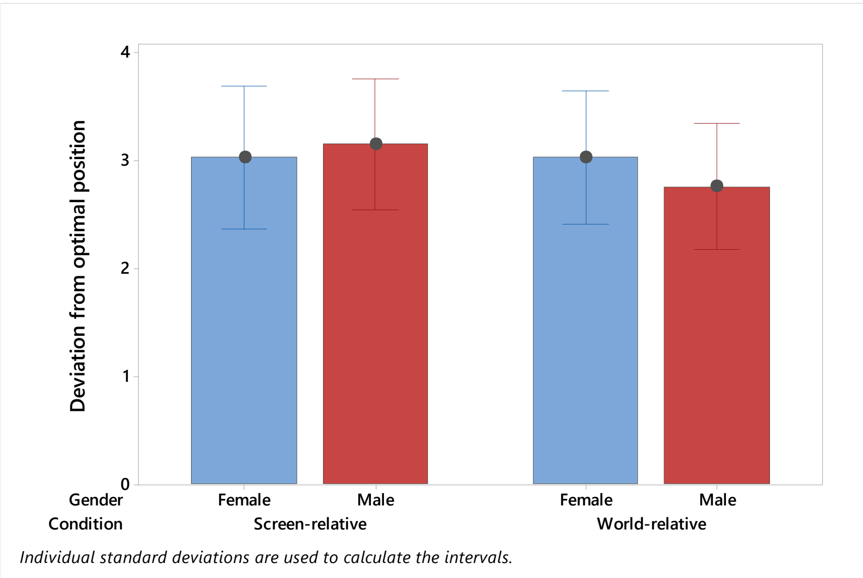


Figure 55- Interval plot deviation from optimal position of route knowledge by condition and gender

5.5 GAZE BEHAVIOR

5.5.1 MAXIMUM HUD GRAPHIC GLANCE DURATION

The maximum HUD Graphic glance duration did not meet ANOVA normality assumption, and therefore box-cox transformation was performed (Estimated λ = -0.0927922). There were no main or interaction effects of condition and gender on the maximum HUD graphic glance duration (Table 64).

Table 64- ANOVA: maximum HUD graphic glance duration

Source	**F-Value**	**P-Value**
Condition	4.37	0.055
Gender	3.64	0.077
Condition*Gender	0.06	0.818

Next, we analyzed which simulation event the maximum HUD glance duration of each participant originated from. Table 65 shows count for this measure with event 4 and 10 presenting the highest counts (four counts each).

Table 65- Count of maximum HUD graphic glance duration per event and condition

		Condition		
		Screen-relative	World-relative	**Total**
Event	1	1	0	**1**
	2	0	1	**1**
	3	1	0	**1**
	4	3	1	**4**
	5	0	0	**0**
	6	1	1	**2**
	7	0	2	**2**
	8	0	1	**1**
	9	0	0	**0**
	10	3	1	**4**
	11	0	0	**0**
	12	0	2	**2**

5.5.2 AVERAGE HUD GRAPHIC GLANCE DURATION

The average HUD Graphic glance duration did not meet ANOVA normality assumption, and therefore box-cox transformation was performed (Estimated λ = -0.495805). There is a main effect of condition on the mean HUD graphic glance duration (Table 66).

Table 66- ANOVA: Average HUD graphic glance duration

Source	F-Value	P-Value
Condition	**7.34**	**0.017**
Gender	4.15	0.061
Condition*Gender	0.49	0.497

Tukey post-hoc test showed that the mean HUD glance duration is higher for the world-relative condition and that this difference is statistically different from the screen-relative condition (Table 67).

Table 67- Tukey Post hoc test : Average HUD graphic glance duration

Condition	N	Mean	Grouping	
World-relative	9	0.705710	A	
Screen-relative	9	0.519909		B

Means that do not share a letter are significantly different.

Based on we can observe that average HUD glance duration tended to decrease for both conditions as the driving simulation time passed. Screen–relative condition showed steadier decrease in average duration compared to the world-relative condition.

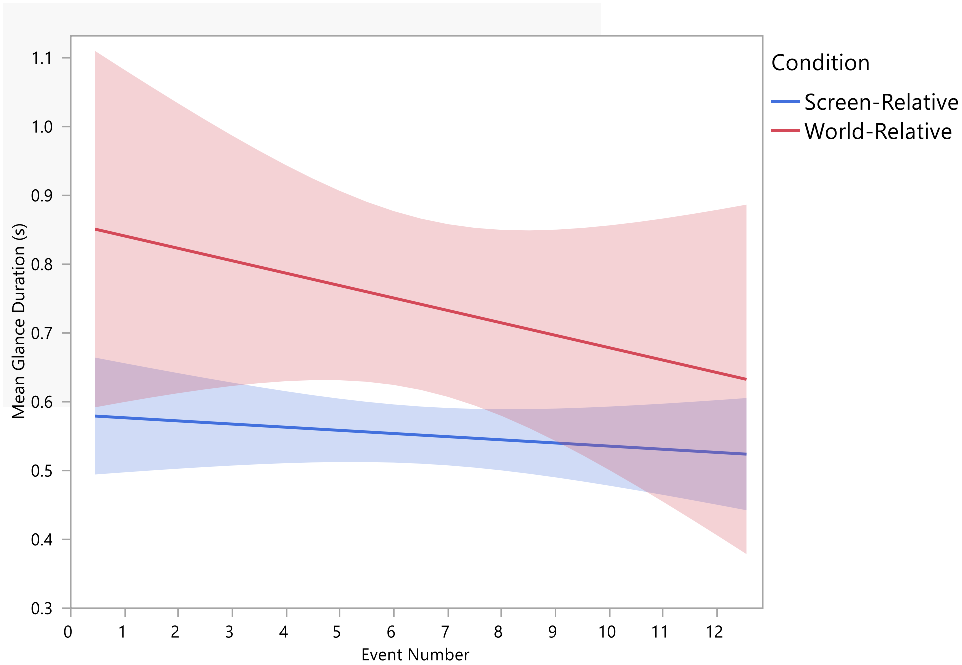


Figure 56- Line chart of average HUD glance duration by condition and event

5.5.3 PERCENTAGE OF TIME LOOKING AT HUD GRAPHIC

We found no main or interaction effects of condition and gender on the percentage of time looking directly at the HUD graphic (Table 68).

Table 68- ANOVA: Percentage of time looking directly at the HUD graphic

Source	F-Value	P-Value
Condition	0.10	0.757
Gender	0.01	0.934
Condition*Gender	0.05	0.829

Descriptive statistics showed that overall, participants looked directly at the HUD graphics at a higher percentage for the screen-relative condition (Table 69). However, as was mentioned, this difference is not statistically significant.

Table 69- Descriptive statistics: Percentage of time looking directly at the HUD graphic

Variable	Condition	Mean	StDev	Minimum	Maximum
% HUD Graphic	Screen-relative	0.1216	0.0602	0.0056	0.1889
	World-relative	0.1133	0.0524	0.0419	0.2083

5.5.4 PERCENTAGE OF TIME LOOKING AROUND HUD GRAPHIC

There were no main or interaction effects of condition and gender on the percentage of time looking around the HUD graphic (Table 70). Additionally, descriptive statistics showed that overall, participants looked around the HUD graphics at a higher percentage for the screen-relative condition (Table 69). However, as was mentioned, this difference is not statistically significant (Table 71).

Table 70- ANOVA: Percentage of time looking around the HUD graphic

Source	F-Value	P-Value
Condition	1.88	0.192
Gender	0.09	0.764
Condition*Gender	1.18	0.295

Table 71- Descriptive statistics: Percentage of time looking around the HUD graphic

Variable	Condition	Mean	StDev	Minimum	Maximum
% Around HUD Graphic	Screen-relative	0.1582	0.0602	0.0327	0.2185
	World-relative	0.1215	0.0386	0.0279	0.1615

5.5.5 PERCENTAGE OF TIME LOOKING AROUND HUD GRAPHIC VS HUD GRAPHIC

Descriptive statistics showed that overall, participants looked around the HUD graphic at a higher percentage compared to directly at the HUD graphic (Table 72). However, a paired t-test revealed that this difference is not significant, meaning that participants looked around and directly the HUD graphic at the same proportion ($t(18) = -1.57$, $p < 0.136$).

Table 72- Descriptive statistics: Percentage of time looking around the HUD graphic

Sample	N	Mean	StDev	SE Mean
% HUD Graphic	18	0.1175	0.0549	0.0129
% Around HUD Graphic	18	0.1399	0.0526	0.0124

5.5.6 TOTAL NUMBER OF GLANCES AT THE HUD GRAPHIC

Descriptive statistics showed that the total number of glances directly at the HUD graphic is higher for the screen-relative condition (Table 73 and Table 72). However, a two-sample t-test revealed the difference between the number of glances at the HUD graphic for both conditions is not significant (t(18) = 1.64, $p<0.129$).

Table 73- Descriptive statistics: Total number of glances at the HUD graphic

Sample	N	Mean	StDev	SE Mean
Screen-relative	9	53.6	29.6	9.9
World-relative	9	35.8	13.4	4.5

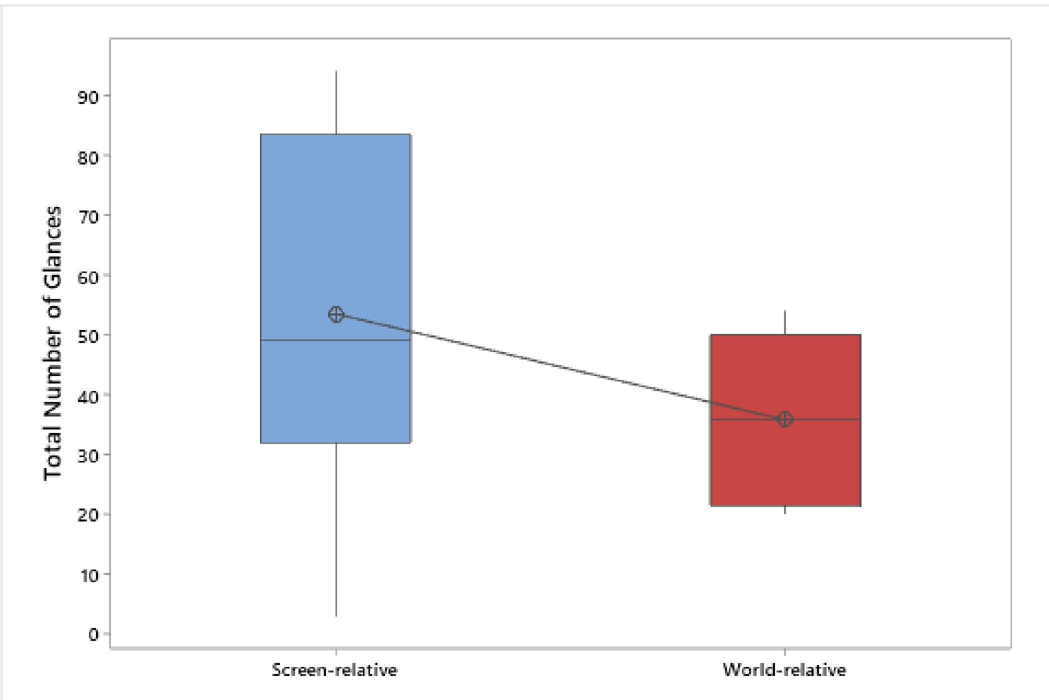


Figure 57- Boxplot of number of glances at the HUD graphic by condition.

5.5.7 TOTAL NUMBER OF GLANCES AROUND THE HUD GRAPHIC

Descriptive statistics showed that the total number of glances around the HUD is higher for the screen-relative condition (Table 74, Table 73 and Table 72). A two-sample t-test revealed the difference between the number of glances around the HUD graphic for both conditions is significant ($t(18) = 2.77$, $p<0.018$).

Table 74- Descriptive statistics: Total number of glances around the HUD graphic

Sample	N	Mean	StDev	SE Mean
Screen-relative	9	69.1	27.7	9.2
World-relative	9	41.0	12.7	4.2

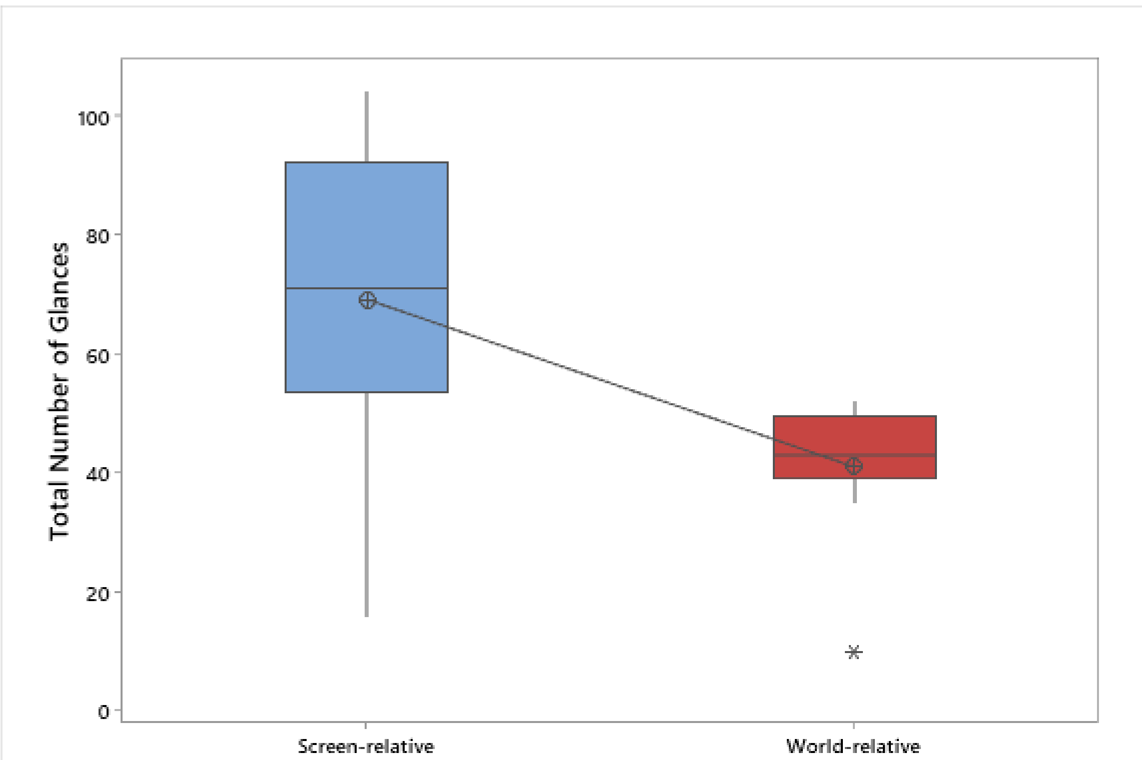


Figure 58- Boxplot of number of glances around the HUD graphic by condition.

6. DISCUSSION

6.1 SURVEYS

NASA-TLX subscales results (Section 5.1.1) showed no significant differences in mental demand, physical demand, temporal demand, effort, frustration, or performance between conditions or genders. However, there were some trends for lower workload in the screen-relative condition. Additionally, screen-relative participants reported being more dissatisfied and less successful with the system than world-relative participants.

The SART survey (Section 5.1.2) showed no statistical difference in awareness between the two conditions. Participants under the world-relative condition reported higher levels of demands on attentional resources than participants under the screen-relative condition. On the other hand, screen-relative drivers reported that the interface provided more elevated levels of supply of attentional resources and more levels of understanding of the situation. Overall, the screen-relative condition provided higher levels of situation awareness. As mentioned by Burnett and Lee [3], the key research question in automotive Human Machine Interfaces (HMIs) is how to develop interfaces that leverage workload (demands) while at the same time promotes active awareness with the environment and navigation. In this line of thought, the screen-relative condition provided higher levels of awareness and lower levels of perceived workload. However, differences between HMI conditions were not statistically significant, and thus, conclusions cannot be drawn.

For the usability survey (Section 5.1.3), no significant differences in any of the usability subscales for genders or conditions were found. Overall, participants perceived the world-relative HMI as harder to navigate and less distracting compared to the screen-relative HMI. For the easy to view subscale, there was an interaction effect, but no main effect of gender and condition. In this case, males suggested that they saw the screen-relative HMI as easier to view while females saw the world-relative HMI as easier to view.

6.2 QUALITATIVE FEEDBACK

Qualitative feedback (Section 5.2) provided some interesting insights about both interfaces. Participants from the world-relative condition showed more difficulties interpreting the navigation interface, which led to uncertainty and missed turns. Some people stated not knowing that the "blue markings" were the navigation cues, or they thought there was a pole or a blue F on the side of the road. Overall performance for left and right turns improved after participants got familiar with the interface. For the straight maneuvers of the world-relative condition, participants did not fully understand the concept of the interface and usually turned whenever they approached the conformal graphic (since the arrow on the back of the pole became more salient). For the straight turns of screen-relative condition, participants believed that there was no real need to show navigation cues for a straight path (as current GPS services). For the roundabout turning of the screen-relative condition, participants commented that the symbol was not too difficult to understand, and the concept is somewhat like the current system used by the navigation application Waze. However, as we observed during the data collection phase, participants were always talking out loud during this turning: “What does the number 2 mean?”, “Do I need to take the second exit?”, “Do I take the roundabout twice?”, “Why is there a number 2 in the traffic circle?”. Although not all participants missed the traffic circle exit, it can be seen that it was challenging to interpret the meaning of the interface. In fact, the roundabout resulted in the most significant deviation in lane position and a lower mean speed compared to left and right turns. This might mean that they were unsure where they should turn and needed to go slower to think about the information the interface was transmitting.

For both conditions, participants liked the idea of presenting navigation cues on a HUD, but they believe that in a real environment with heavy traffic, 150 meters is not enough to react and change lanes when it is needed.

6.3 DRIVING MEASURES

Results showed that AR HUD perceptual forms do not differ in terms of the standard deviation of lane position (Section 5.3.2) and standard deviation of speed (Section 5.3.3) as participants approached the desired turn. Based on these results, we reject H2A: “*Participants using world-relative graphics will have more variability on driving behavior measures (i.e. standard deviation of speed and standard deviation of lane position) compared to participants using screen-relative graphics”.*

For the standard deviation of lane position measure, a main effect of maneuver direction was found. However, this effect is not related to the type of perceptual form provided by AR HUD graphics. The highest lane position deviation was found when participants were turning on the roundabout (mean = 2.34248). We believe that the roundabout was the most challenging maneuver to participants due to three main reasons: a) ability to understand AR HUD graphic meaning for the screen-relative condition (see Figure 59); b) ability to view world-relative graphics on time to make the correct maneuver due to the HUD small field of view (see Figure 60); c) ability to maintain control over the simulator when turning in circle.

Figure 59- Screen-relative roundabout concept used within this work. The number two in the center means that participants should take the second exit.

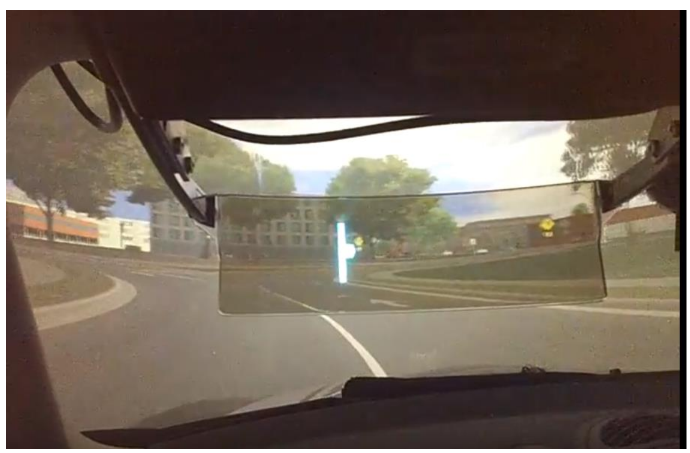


Figure 60- World-relative signpost during the roundabout maneuver. As we can observe, participants could not either see the navigation cue 150 meters from the intersection or see the complete navigation cue.

In fact, when analyzing other driving measures, we found that the roundabout presented the highest number of missed turns (36.84%). Given these points, we believe it is plausible that the roundabout showed the highest lane position deviation. The right maneuver showed a smaller standard deviation of lane position when compared to the left maneuver. This result is expected because usually drivers drive on the right side of the road, and thus, turning right will result in a smaller lane position deviation than turning left. However, when looking at missed turns, the right maneuvers resulted in 21.05% of missed turns, and left maneuvers resulted in only 15.79%. We believe these results might be due to the unfamiliarity of participants with the concept of the conformal graphics. In fact, from all four right turns of the simulated environment, all missed turns happened in the first right maneuver. Further, 36.84% of all missed turns happened within the three first maneuvers (straight, right, and left) and 36.84% of all missed turns happened at the roundabout, resulting in 73.68%.

For the standard deviation of speed measure, a main effect of maneuver direction was found. However, this effect is not related to the type of perceptual form provided by AR HUD graphics. The highest vehicle speed deviation was found when participants were turning left (mean = 12.600) and right (mean =11.1106), respectively. Tukey post-hoc analysis found that these variations are not statistically significant. The roundabout resulted in 8.1985 SD vehicle speed, and the straight tmaneuver resulted in 6.9999 SD vehicle speed. Tukey post-hoc analysis revealed these means are statistically similar. We believe that these last two maneuvers presented smaller standard deviations due to two main reasons: a) when entering a roundabout, drivers tend to reduce vehicle speed and keep it constant to maintain control of the vehicle; b) straight maneuver do not require an actual vehicle turning, and thus, vehicle speed is usually kept constant.

Additionally, results showed that AR HUD perceptual forms differ in terms of the number of missed turns (Section 5.3.1). In fact. world-relative graphics were associated with 78.95% of all missed turns. Based on these results, we fail to reject H2B: "*Participants using word-relative graphics will have more missed turns compared to those using screen-relative graphics*". However, we do believe these results might be due to the unfamiliarity of participants with the concept of conformal graphics as previously explained in the lane position deviation measure.

Because graphic type had little impact on participants' driving behaviors in this study, we can suggest that HUD graphic type may not always influence driver behaviors. When designing future vehicle displays, we should make thoughtful design choices, but understand that any decision may not have a significant impact on driver behavior when approaching a turn. Of course, we need to bear in mind that we provided only navigation directions (therefore, only one category of information) to participants utilizing two types of graphics–one conformal and one screen-relative. Other types of information combined with different kinds of graphics could have a different impact on how drivers perform while using the interface.

6.4 SPATIAL KNOWLEDGE

6.4.1 LANDMARK KNOWLEDGE

Initial analysis (Section 5.4.1.1) showed that drivers under the screen-relative condition were able to correctly place more images in both "I remember" and "I do not remember" piles. However, this difference is not statistically significant. Participants using the world-relative HMI presented a mean of 47.92% landmark acquisition rate, and participants using the screen-relative HMI presented a mean of 49.48%. For this measure, no differences between genders or conditions were found, and therefore we reject H1A: "*Participants will develop higher levels of landmark knowledge using world-relative graphics"*.

When analyzing landmark knowledge using signal detection theory (Section 5.4.1.2), we found that overall, participants under both conditions achieved hit rate of 46.88%, false alarm rate of 48.96%, miss rate of 53.13% and correct rejection rate of 51.04%. In this work, overall performance in the landmark test was defined as the sum of hit and correct rejection rates. Deviance analysis showed no main effect of condition or gender on overall performance. However, we found a main effect of maneuver direction on overall performance probability ($p<0.05$). The main effect plot demonstrated that the straight maneuver has the highest probability of success (successfully recognizing on-route scenes or rejecting off-route scenes) and the left maneuver has the lowest probability of success. Odd rations showed that the screen-relative condition is more likely to achieve success in the landmark knowledge test (in fact, screen-relative condition presented highest landmark knowledge acquisition rate), but based on the 95% CI this difference is not statistically significant. When comparing maneuver direction using odd rations, we observed that achieving success in the landmark knowledge test is more likely in the straight maneuver compared to both roundabout and left turn. Based on the 95% CI this difference is statistically significant.

One of the main strengths of SDT applied to landmark knowledge acquisition problems is that it allows a more thorough characterization of the ability of drivers to recognize features of the environment. For instance, it was possible to use d' to quantify how challenging for participants it was to distinguish between signal and noise between different conditions and maneuvers. Also, we can assess whether participants' performance has been affected by any systematic response preference (i.e. bias). In this work, SDT analysis showed that participants had low sensitivity to the differences between scenes shown during the drive and scenes that were not included in the driving scenario. This could have happened due to the fact we used a medium-fidelity simulator in which landmarks have low salience, and each on-route scene had its corresponding off-route scene that shared some type of similarity between them (see Figuer 61). In this context, we believe that participants did not acquire enough discriminable information about the environment so that signal and noise could be distinguished from each other, and so our results are due to pure chance.

Figure 61- On-route scene (right) and off-route scene (left) example used for the landmark knowledge test

Of some relevance to the variance in decision-making processes in the landmark test, there was some evidence (through participants' open feedback comments) that drivers were may not have been able to pay much attention to the driving environment due to the unfamiliarity with the navigation cue system. Several participants stated that they could not capture enough characteristics of the environment because their attention was focused on understanding the HUD interface instead. Also, no payoffs were offered to participants in this study, and they did not understand the "value" of a correct rejection or a false alarm for example. In this case, they might not have made a great deal of effort to complete this task and recall the features of the driving scene in order to position the scenes in the right pile properly.

Finally, heat map findings suggested that in fact, participants used landmarks as a strategy to navigate (78% of heat map clicks were in landmarks). However, due to study design limitations, we could evaluate whether different conditions presented different clicks patterns in this task.

6.4.2 ROUTE KNOWLEDGE

Route knowledge results (Section 5.4.2) revealed that generally, participants were more deficient at route knowledge test compared to the landmark knowledge test. While landmark knowledge acquisition rate was 49.48% for screen-relative and 47.92% for world-relative, route knowledge was 15.15% for screen-relative and 20.45% for world-relative condition. A two-sample un-paired t-test revealed that participants using both screen-relative and world-relative methods remembered the same proportion of scenes in the correct location. Therefore, we reject H1B: "*Participants will develop higher levels of route knowledge using world-relative graphics".*

6.5 GAZE BEHAVIOR

Concerning glance behavior, no statistical difference was found among conditions for maximum HUD glance duration (Section 5.5.1), percentage of time looking at HUD (Section 5.5.3), percentage of time looking around the HUD (Section 5.5.4), and the number of glances at the HUD (Section 5.5.6). We found that overall, participants under the world-relative condition had higher HUD glance duration than participants under the screen-relative condition. Therefore, based on these results, we fail to reject H3: *"Participants using world-relative graphics will glance longer to the HUD navigation cue compared to those using screen-relative graphics".* Also, we noted that the average HUD glance duration decreased in both conditions as participants were more familiar with the interface.

It is important to emphasize that participants using the screen-relative HMI glanced more times at the HUD graphic, but these glances had shorter durations compared to the world-relative HMI. We believe that participants had more glances at the HUD graphic under the screen-relative condition because, in this case, the navigation arrow graphic appeared in front of the driver that needs to be constantly looking forward to driving safely. On the other hand, because world-relative graphics appeared on the right side of the HUD, drivers don't need to constantly look at that side of the display, and therefore, a lower number of glances towards the HUD graphic in the world-relative condition is plausible. According to Smith [16], "an optimal graphic would require few glances of short duration in the direction of the graphic and would increase the amount of visual attention available for drivers to allocate to other areas with potential hazards and other driving-relevant information". However, in our study, neither of the HMI investigated satisfied this requirement.

We wanted to better understand visual attention allocation to four other important areas of interest: close left landmark, close right landmark, far-left landmark, and far-right landmark. Therefore, we extended the scope of our analysis to include these areas. Although participants glanced fewer times to the world-relative graphics compared to the screen-relative graphics, as we can observe in Figure 62, participants glanced more at the right landmarks in the world-relative condition. Left landmarks presented a similar number of glances for both conditions. It is important to emphasize that world-relative graphics were positioned in front of the far right landmark of the intersection, and as we can observe this AOI was the one that presented the most number of glances. Because of this difference in glance number allocation, we would expect that participants under world-relative condition would acquire more visual information from the environment and thus, acquire higher levels of spatial knowledge. However, looking at an AOI does not necessarily mean that drivers were absorbing and processing valuable information about it.

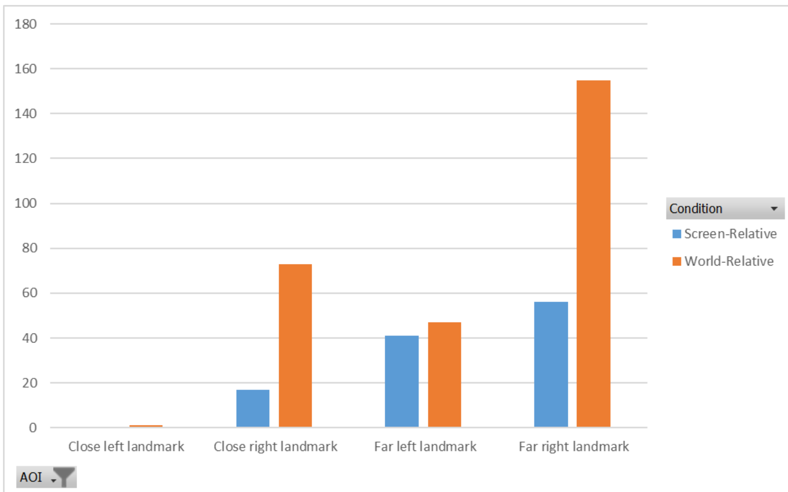


Figure 62- Number of glances towards four different AOIs

Table 75 shows descriptive statistics of each AOI and its respective t-test p-value comparing world-relative (green) and screen-relative conditions (blue). We could not compare results for the close left landmark AOI because under the screen-relative condition because there were no glances towards this AOI.

Finally, when looking at the percentage of time allocated in each AOI, we can observe that participants allocated less time to the HUD graphic and around the HUD graphic under the world-relative condition (Figure 63). Instead, in this case, participants allocated more time to landmarks (especially right landmarks) than the screen-relative condition. Ultimately, time allocated to other AOIs was reasonably similar to both conditions.

Table 75- Descriptive statistics and t-tests for glance behavior analysis

AOI	Number of Glances	Mean Glance Duration (in sec)
Far right landmark	$\mu = 6.22;\ \sigma = 4.02$ $\mu = 11.33;\ \sigma = 8.67$ ***p*<0.137**	$\mu = 0.391;\ \sigma = 0.118$ $\mu = 0.507;\ \sigma = 0.223$ ***p*<0.190**
Far left landmark	$\mu = 5.86;\ \sigma = 3.18$ $\mu = 4.89;\ \sigma = 4.08$ ***p*<0.240**	$\mu = 0.453;\ \sigma = 0.280$ $\mu = 0.471;\ \sigma = 0.197$ ***p*<0.891**
Close right landmark	$\mu = 2.83;\ \sigma = 1.47$ $\mu = 4.89;\ \sigma = 4.08$ ***p*<0.197**	$\mu = 0.299;\ \sigma = 0.280$ $\mu = 0.657;\ \sigma = 0.395$ ***p*<0.030**
Close left landmark	**Information not available for the world-relative condition**	

Screen-relative World-relative

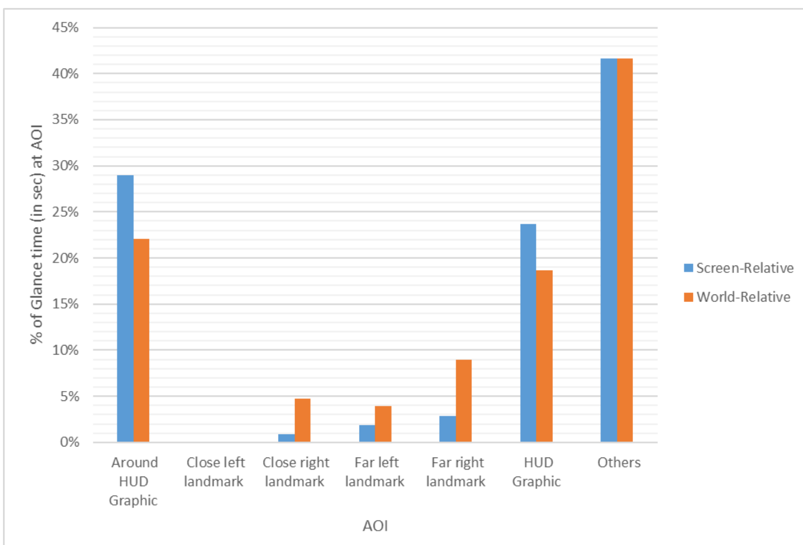


Figure 63- Use two types of graphics' perceptual forms resulted in different allocations of glance time across the seven areas of interest.

7. CONCLUSIONS

In this study, we investigated whether modifying HUD graphics' perceptual form influences spatial learning of the environment. Two different navigation cues systems were compared: world-relative and screen-relative. The world-relative condition placed an artificial post sign at the corner of an approaching intersection containing a real landmark. The screen-relative condition displayed turn directions using a screen-fixed traditional arrow located directly ahead of the participant on the right or left side on the HUD.

Participants using both world-relative and screen-relative conditions were presented with turn-by-turn guidance with different visual user interface designs. We believe that in both cases, participants may voluntarily cease to explicitly engage in effortful knowledge acquisition because of the use of AR HUD for the decision-making process to navigate. The same phenomenon has been noted in other studies using traditional GPS devices [4, 73]. A strategy that has the potential to improve spatial knowledge of the environment while driving is to direct attention towards relevant environmental features, specifically landmarks. We attempted to do this by placing a visual post-sign at the corner of an approaching intersection containing a real landmark. We expected that placing conformal AR cues on the right side of the road would draw attention to the right side features of the environment and therefore, participants would remember seeing these landmarks along the drive. In fact, glance behavior results showed that the distribution of visual attention allocation was that the world-relative condition was associated with glances of long duration directed to the right side of the HUD. However, in this study participants using both AR HUDs' perceptual form interfaces presented similar levels of spatial knowledge acquisition. One of the reasons that this may have happened is due to the fact that the world-relative condition AR graphic was not compatible with drivers' mental model of what a navigation cue looks like. Many participants stated not knowing that the "blue markings" were the navigation cues, or they thought there was a pole or a blue 'F' on the side of the road. Therefore, participants glanced longer to the right side of the interface as they were trying to understand what information the system was transmitting to them. As participants learned how the system works, the average glance duration decreased, and they were more comfortable using navigation cues.

Even though our initial assumption that the conformal AR HUD interface would draw drivers' attention to a specific part of the display was correct, this type of interface was not helpful to increase spatial knowledge acquisition. This finding contrasts a common perspective in the AR community that conformal, world-relative graphics are inherently more effective than screen-relative graphics. We suggest that simple, screen-fixed designs may indeed be effective in certain contexts. Also, we showed that the distribution of visual attention allocation was that the world-relative condition was typically associated with fewer glances in total, but glances of longer duration. Optimal AR graphics would require few glances of short duration in the direction of the graphic and would increase the amount of visual attention available for drivers to allocate to other areas with potential hazards and other driving-relevant information. And as mentioned, this finding alone warrants further investigation since changes in visual attention in more dynamic and dangerous settings can have significant differences in primary task performance (e.g., driving).

8. LIMITATIONS AND FUTURE WORK

Some limitations of the present study should be pointed out. First of all, we did not have a baseline condition with no AR graphic (e.g., a voice-only navigation system or a head down display) and therefore, we could not draw any conclusions regarding the use of a HUD to increase spatial knowledge acquisition while driving compared to other navigation cues. Second, our graphics were static in the sense that they were not animated in the world or across the screen, and therefore we did not fully explore all possibilities of different HUD graphics' perceptual forms and their impact on drivers' gaze and/or spatial knowledge acquisition. Third, this study used a medium-fidelity driving simulator in which characteristics of landmarks were not easily discriminable as they may be in the real world. Therefore, we believe that participants were limited in their ability to acquire critically discernible visual information about the environment in which they were driving. Despite obvious benefits of using a simulator to address AR driving related research questions, it is evident that, like much of outdoor AR research, in-field user-studies are necessary to fully understand the relationship between interface design and spatial knowledge acquisition. Further, even though this research presented traffic in the scene so as to simulate a realistic driving environment for participants, this study was not designed so that surprise events involving hazards, traffic or pedestrians would be present in participants 'driving lane. Therefore, this point should certainly be addressed with future work so that issues such as occlusion or driver's ability to respond to hazards can be addressed. In these cases, perhaps greater visual allocation to the world-fixed post sign could be problematic, despite differences in spatial knowledge acquisition. Also, the sample population for this study was composed only of healthy college students. Future studies should also include a broader variety of population, accounting for differences in age range, especially because it is well-known that spatial knowledge abilities decrease for the older population [74] and therefore, HUDs could have the potential to unequally help (or hurt) this age range of the population.

Also, this study was not designed to evaluate participants' spatial learning versus recognition. Because no other tests were conducted following the day of data collection, our results cannot evaluate long-term memory effects which are the core of many spatial knowledge approaches. Future studies should investigate delays between learning and testing to understand whether long-term spatial learning took place using different AR interface designs presented via HUD.

Ultimately, drivers did not make strategic decisions regarding what route to take as they were only involved with tactical aspects of the navigation task. In this sense, they had only to confirm instead of choosing strategic aspects of the overall navigation task. Even though HUDs have the potential to guide drivers' visual attention to important characteristics of the environment, in this study participants received navigation instructions similar to traditional navigation systems. Therefore, we believe that future studies should investigate the role of "active drivers" who also have choice decisions in strategic aspects of the navigation using a HUD. We hypothesize that drivers in this latter case will develop more refined cognitive maps than drivers simply following turn-by-turn instructions without active engagement

Finally, in a broader level, the consensus in the literature is that normal activities have the potential to induce changes in the grey matter of the brain, and as a consequence, it has obvious implications in rehabilitation methods for those people who have suffered a brain disease or injury [53]. In fact, the hippocampus seems to be crucial to the storage and use of mental maps of our environment, but research have suggested that by using head-down navigation devices, such as the GPS, hippocampal volume and activity tend to decrease. When searching articles for general audience, many of these studies negatively linking navigation devices to brain size can be found, showing that this matter has gone out of university walls: "Ditch the GPS. It's ruining your brain", "Cache cab: taxi drivers' brain grows to navigate London streets, "the knowledge enlarges your brain", "Technology: use or lose our navigation skills", "is GPS ruining our ability to navigate for ourselves?", "Meet your brain GPS: the hippocampus", Are GPS apps messing with our brains?".

However, all of these articles only talk about GPS and no discussion of brain activity and development changes when using HUDs compared to HDD. We believe that the key point is to use wayfinding devices intelligibly, and HUDs could have the potential to help drivers to absolve more knowledge about their environment, and therefore, better exercise their hippocampus. Thus, future work could analyze the link between the use of HUDs, brain activity and spatial knowledge, and compare it to traditional HDDs.

REFERENCES

[1] G. Burnett, "On-The-Move And In Your Car: An Overview Of HCI Issues For In-Car Computing," *International Journal Of Mobile Human Computer Interaction,* Pp. 60-78, 2009.

[2] W. Waters And S. Winter, "A Wayfinding Aid To Increase Navigator Independence," *Journal Of Spatial Information Science,* Pp. 103-122, 2011.

[3] G. Burnett And K. Lee, "The Effect Of Vehicle Navigation Systems On The Formation Of Cognitive Maps," In *Traffic And Transport Psychology: Theory And Application*, Elsevier, 2005, Pp. 407-418.

[4] E. Webber, *Strategy Differences In The Use Of Mobile Devices For Navigation,* University Of Nottigham , 2014.

[5] D. Beck And W. Park , "Perceived Importance Of Automotive HUD Information Items: A Study With Experienced HUD Users," *EEE Access,* Vol. 6, Pp. 21901-21909, 2018.

[6] A. Bolton, G. Burnett And D. Large, "An Investigation Of Augmented Reality Presentations Of Landmark-Based Navigation Using Head-Up Display," In *Proceedings Of The 7th International Conference On Automotive User Interfaces And Interactive Vehicular Applications*, Nottingham, United Kingdom, 2015.

[7] C. D. Wickens, J. G. Hollands, S. Banbury And R. Parasuraman, Engineering Psychology And Human Performance, Pearson Education, 2013.

[8] D. A. Guttentag, "Virtual Reality: Applications And Implications For Tourism," *Tourism Management,* Vol. 31, No. 5, Pp. 637-651, 2010.

[9] R. Azuma, "A Survey Of Augmented Reality," *Presence: Teleoperators And Virtual Environments,* Pp. 355-385, 1997.

[10] P. Milgram, H. Takemura, A. Utsumi And F. Kishino, "Augmented Reality: A Class Of Displays On The Reality-Virtuality Continuum," *Proc. SPIE 2351, Telemanipulator And Telepresence Technologies,* Pp. 282-292, 1995.

[11] M. Slater, "A Note On Presence Terminology," *Presence Connect,* Vol. 3, 2003.

[12] I. Tussyadiah, D. Wang And C. (. Jia, "Virtual Reality And Attitudes Toward Tourism Destinations," In *Information And Communication Technologies In Tourism 2017* , 2017, Pp. 229-239.

[13] I. Sutherland, "A Head-Mounted Three-Dimensional Display," In *Fall Joint Computer Conf., Am. Federation Of Information Processing Soc. (AFIPS),* Washington, D.C. , 1968.

[14] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier And B. Macintyre, "Recent Advances In Augmented Reality," *IEEE Computer Graphics And Applications,* Pp. 34-47, 2001.

[15] P. Milgram And F. Kishino, "A Taxonomy Of Mixed Reality," *IEICE Trans. Information System,* Pp. 1321-1329., 1994.

[16] M. I. Smith, "Informing Design Of In-Vehicle Augmented Reality Head-Up Displays," (Doctoral Dissertation), Blacksburg, VA, 2018.

[17] M. Smith, J. L. Gabbard And C. Conley, "Head-Up Vs. Head-Down Displays: Examining Traditional Methods Of Display Assessment While Driving," In *Automotive'UI 16 Proceedings Of The 8th International Conference On Automotive User Interfaces And Interactive Vehicular Applications*, Ann Arbor, MI, USA, 2016.

[18] D. R. Large, G. Burnett And A. Bolton, "Augmenting Landmarks During The Head-Up Provision Of In-Vehicle Navigation Advice," *International Journal Of Mobile Human Computer Interaction,* Vol. 9, No. 2, Pp. 18-38, 2017.

[19] Y. Liu And M. Wen , "Comparison Of Head-Up Display (HUD) Vs. Head-Down Display (HDD): Driving Performance Of Commercial Vehicle Operators In Taiwan," *International Journal Of Human-Computer Studies*, Vol. 61, No. 5, Pp. 679-697, 2004.

[20] Z. Medenica, A. L. Kun , T. Paek And O. Palinko , "Augmented Reality Vs. Street Views: A Driving Simulator Study Comparing Two Emerging Navigation Aids," *Mobile HCI,* 2011.

[21] M. Ablajimeier, T. Poitschke, F. Wallhoff, K. Bengler And G. Rigoll, "Eye Gaze Studies Comparing Head-Up And Head-Down Displays In Vehicles," *Multimedia And Expo, 2007 IEEE International Conference On Multimedia And Expo,* 2007.

[22] M. Maroto, E. Caño, P. González And D. Villegas, "Head-Up Displays (HUD) In Driving," *Human-Computer Interaction,* 2018.

[23] J. L. Gabbard, G. M. Fitch And H. Kim, "Behind The Glass: Driver Challenges And Opportunities For AR Automotive Applications," In *Proceedings Of The IEEE*, 2014.

[24] C. Merenda, H. Kin, K. Tanous And J. L. Gabbard, "Augmented Reality Interface Design Approaches For Goal-Directed And Stimulus-Driven Driving Tasks," *IEEE Transactions On Visualization And Computer Graphics*, Vol. 24, No. 11, Pp. 2875-2885, November 2018.

[25] R. S. Mccann, D. C. Foyle And J. C. Johnston, "ATTENTIONAL LIMITATIONS WITH HEAD-UP DISPLAYS," In *Proceedings Of The Seventh International Symposium On*, Columbus , 1993.

[26] A. Treisman, "Perceptual Grouping And Attention In Visual Search For Features And For Objects," *Journal Of Experimental Psychology: Human Perception And Performance,* Vol. 8, No. 2, Pp. 194-214, 1982.

[27] C. D. Wickens And J. Long, "Object Versus Space-Based Models Of Visual Attention: Implications For The Design Of Head-Up Displays.," *Journal Of Experimental Psychology: Applied,* Vol. 1, No. 3, Pp. 179-193, 1995.

[28] D. J. Simons And C. F. Chabris, "Gorillas In Our Midst: Sustained Inattentional Blindness For Dynamic Events," *Perception,* Vol. 28, Pp. 1059 - 1074, 1999.

[29] U. Neisser, "The Control Of Information Pickup In Selective Looking.," In *Perception And Its Development: A Tribute To Eleanor Gibson*, Hillsdale, NJ, Erlbaum, 1979, Pp. 201-219.

[30] R. F. Haines, "A Breakdown In Simultaneous Information Processing.," In *Presbyopia Research*, New York, Plenum Press., 1991, Pp. 171-175.

[31] P. Ververs And C. Wickens, "DESIGNING HEAD-UP DISPLAYS (Huds) TO SUPPORT FLIGHT PATH GUIDANCE WHILE MINIMIZING EFFECTS OF COGNITIVE TUNNELING," In *IEA 2000/HFES 2000 Congress*, 2000.

[32] E. Fischer, R. F. Haines And T. Price, "Cognitive Issues In Head-Up Display (NASA Technical Paper 1711)," Moffett Field, CA, NASA Ames Research Center, 1980.

[33] E. Kruijff, J. Swan And S. Feiner, "Perceptual Issues In Augmented Reality Revisited," In *2010 IEEE International Symposium On Mixed And Augmented Reality*, 2010.

[34] J. D. Lee, "Fifty Years Of Driving Safety Research," *HUMAN FACTORS*, Vol. 50, No. 3, P. 521–528, June 2008.

[35] D. Shinar, "Theories And Models Of Driver Behavior," *Traffic Safety And Human Behavior*, Pp. 53-90, 2016.

[36] K. Holmqvist, M. Nystrom, R. Anderson, R. Dewhurst, H. Jarodza And J. V. D. Weijer, Eye Tracking: A Comprehensive Guide To Methods And Measures, Oxford Press, 2011.

[37] C. R. Bennett And N. A. Giudice, "Evaluating Age-Related Cognitive Map Decay Using A Novel Time-Delayed Testing Paradigm," In *Spatial Cognition*, Springer International Publishing AG 2017, 2017, P. 69–85.

[38] E. Tolman, "Cognitive Maps In Mice And Men," *Psychological Review*, Vol. 55, Pp. 189-208, 1948.

[39] M. Hill, "Asking Directions And Pedestrian Wayfinding," *Man-Environment Systems*, Vol. 17, No. 4, Pp. 113-1120, 1987.

[40] K. J. Oliver And G. E. Burnett , "Learning-Oriented Vehicle Navigation Systems: A Preliminary Investigation In A Driving Simulator," In *Proceedings Of The 10th International Conference On Human Computer Interaction With Mobile Devices And Services*, 2008.

[41] A. W. Siegel And S. H. White, "The Development Of Spatial Representatives Of Large-Scale Environments," In *Advances In Child Development And Behaviour,* San Diego, CA, Academic Press, 1975, Pp. 9-55.

[42] A. Klippel, S. Hirtle And C. Davies, "You-Are-Here Maps: Creating Spatial Awareness Through Map-Like Representations. Tation, 10(2-3),.," *Spatial Cognition & Compu,* Vol. 10, No. (2-3), P. 83–93, 2010.

[43] E. Stern And D. Leiser, "Levels Of Spatial Knowledge And Urban Travel Modeling," *Geographical Analysis,* Vol. 20, No. 2, Pp. 140-155, 1988.

[44] V. Antrobus, G. Burnett And L. Skrypchuk, "“TURN LEFT AT THE FAIRHAM PUB” USING NAVIGATIONAL GUIDANCE TO RECONNECT DRIVERS WITH THEIR ENVIRONMENT," In *Proceedings Of The 8th International Conference On Automotive User Interfaces And Interactive Vehicular Applications,* Ann Arbor, MI, USA, 2016.

[45] E. P. Fenech, F. Drews And J. Z. Bakdash, "The Effects Of Acoustic Turn-By-Turn Navigation On Wayfinding," *Proceedings Of The Human Factors And Ergonomics Society Annual Meeting,* Vol. 54, No. 23, Pp. 1926-1930, 2010.

[46] R. Golledge, A. Ruggles, J. Pellegrino And N. Gale, "INTEGRATING ROUTE KNOWLEDGE IN AN UNFAMILIAR NEIGHBORHOOD: ALONG AND ACROSS ROUTE EXPERIMENTS," *Journal Of Environmental Psychology,* Vol. 13, Pp. 293-307, 1993.

[47] V. D. Bohbot, L. Jason, T. Brook, G. Iaria And A. P. Zijdenbos, "Gray Matter Differences Correlate With Spontaneous Strategies In A Human Virtual Navigation Task," *The Journal Of Neuroscience,* Vol. 27, No. 38, P. 10078 –10083, 2007.

[48] K. Konishi And V. D. Bohbot, "Spatial Navigational Strategies Correlate With Gray Matter In The Hippocampus Of Healthy Older Adults Tested In A Virtual Maze," *Frontiers In Aging Neuroscience,* Vol. 5, Pp. 1-8, 2013.

[49] K. Chen, L. Chuah, S. Sim And M. Chee, "Hippocampal Region-Specific Contributions To Memory Performance In Normal Elderly," *Brain And Cognition,* Pp. 400-407, 2010.

[50] D. Head And M. Isom, "Age Effects On Wayfinding And Route Learning Skills," *Behavioural Brain Research,* Vol. 209, No. 17, Pp. 49-58, 2010.

[51] G. Iaria, M. Petrides, A. Dagher, B. Pike And V. D. Bohbot, "Cognitive Strategies Dependent On The Hippocampus And Caudate Nucleus In Human Navigation: Variability And Change With Practice," *The Journal Of Neuroscience,* Vol. 23, No. 13, Pp. 5945-5952, 2003.

[52] V. Bohbot, "Use Spatial Memory To Reduce Risks Of Dementia," 04 January 2011. [Online]. Available: Http://Www.Douglas.Qc.Ca/News/1072/File_En/Press-Release-Bohbot-Dementia.Pdf.

[53] E. A. Maguire, D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. J. Frackowiak And C. D. Frith, "Navigation-Related Structural Change In The Hippocampi Of Taxi Drivers," *Proceedings Of The National Academy Of Sciences Of The United States Of America,* Vol. 97, No. 8, Pp. 4398-4403, 2000.

[54] E. A. Maguire, H. J. Spiers, C. D. G. Good, T. Hartley, R. S. Frackowiak And N. Burgess, "Navigation Expertise And The Human Hippocampus: A Structural Brain Imaging Analysis," *Hippocampus,* Vol. 13, Pp. 208-217, 2003.

[55] K. Woollett And E. A. Maguire, "Acquiring "The Knowledge" Of London's Layout Drives Structural Brain Changes," *Current Biology,* Pp. 2109-20114, 2011.

[56] A. Parush, S. Ahuvia And I. Erev, "Degradation In Spatial Knowledge Acquisition When Using Automatic Navigation Systems," In *8th International Conference, COSIT 2007* , Melbourne, Australia, 2007.

[57] T. Ishikawa And K. Takahashi, "Relationships Between Methods For Presenting Information On Navigation Tools And Users' Wayfinding Behavior," *Cartographic Perspectives,* Vol. 0, No. 75, Pp. 17-28, 2014.

[58] S. Münzer, H. D. Zimmer, M. Schwalm, J. Baus And I. Aslan, "Computer-Assisted Navigation And The Acquisition Of Route And Survey Knowledge," *Journal Of Environmental Psychology,* Vol. 26, No. 4, Pp. 300-308, 2006.

[59] K. S. Willis, C. Hölscher, G. Wilbertz And C. Li, "A Comparison Of Spatial Knowledge Acquisition With Maps And Mobile Maps," *Computers, Environment And Urban Systems,* Vol. 33, No. 2, Pp. 100-110, 2009.

[60] T. Ishikawa, F. Hiromichi, I. Osamu And O. Atsuyuki, "Wayfinding With A GPS-Based Mobile Navigation System: A Comparison With Maps And Direct Experience," *Journal Of Environmental Psychology,* Vol. 28, No. 1, Pp. 74-82, 2008.

[61] M. R. Endsley, "Toward A Theory Of Situation Awareness In Dynamic Systems," *HUMAN FACTORS*, Vol. 37, No. 1, Pp. 32-64, 1995.

[62] K.-F. Richter, D. Dara-Abrams And M. Raubal, "Navigating And Learning With Location Based Services: A User-Centric Design," In *The 7th International Symposium On LBS & Telecartography*, Guangzhou, 2010.

[63] G. E. Burnett, "Usable Vehicle Navigation Systems : Are We There Yet ? Usable Vehicle Navigation Systems : Are We There Yet ?," In *Vehicle Electronic Systems 2000 - European Conference And Exhibition,* 2000.

[64] A. H. J. Oomes, M. Bojic And G. Bazen, "Supporting Cognitive Collage Creation For Pedestrian Navigation," In *Engineering Psychology And Cognitive Ergonomics. EPCE 2009. Lecture Notes In Computer Science*, Berlin, Heidelberg, Springer, 2009, Pp. 111-119.

[65] K. Gramann, P. Hoepner And K. Karrer-Gauss, "Modified Navigation Instructions For Spatial Navigation Assistance Systems Lead To Incidental Spatial Learning," In *Frontiers In Psychology.*, Frontiers Research Foundation, 2017, P. 193.

[66] I. Reagan And C. L. Baldwin, "Facilitating Route Memory With Auditory Route Guidance Systems," *Journal Of Environmental Psychology,* Pp. 146-155, 2006.

[67] W. Wu, F. Blaicher, J. Yang, T. Seder And D. Cui, "A Prototype Of Landmark-Based Car Navigation Using A Full-Windshield Head-Up Display System," In *AMC '09 Proceedings Of The 2009 Workshop On Ambient Media Computing*, Beijing, China , 2009.

[68] M. R. Endsley And D. J. Garland, Situation Awareness Analysis And Measurement, London: Lawrence Erlbaum Associates, 2000.

[69] M. R. Endsley , S. J. Selcon, T. D. Ha&Man And D. G. Croft, "A COMPARATIVE ANALYSIS OF SAGAT AND SART FOR EVALUATIONS OF SITUATION AWARENESS," In *Proceedings Of The Human Factors And Ergonomics Society 42nd Annual Meeting*, 1998.

[70] W. P. Tanner And J. A. Swets, "A Decision-Making Theory Of Visual Detection," *Pyschological Review,* Vol. 61, Pp. 401-409, 1954.

[71] D. M. Green And J. A. Swets, Signal Detection Theory And Psychophysics, New York: Krieger, 1966.

[72] T. D. Wickens, Elementary Signal Detection Theory, New York: Oxford University Press, 2002.

[73] J. Z. Bakdash, S. A. Linkenauger And D. Proffitt, "Comparing Decision-Making And Control For Learning A Virtual Environment: Backseat Drivers Learn Where They Are Going," In *Proceedings Of The Human Factors And Ergonomics Society Annual Meeting*, 2008.

[74] P. Jansen, A. Schmelter And M. Heil, "Spatial Knowledge Acquisition In Younger And Elderly Adults: A Study In A Virtual Environment," *Exp Psychol,* Vol. 57, No. 1, Pp. 54-60, 2010.

[75] S. Kim And A. K. Dey, "Simulated Augmented Reality Windshield Display As A Cognitive Mapping Aid For Elder Driver Navigation," In *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems*, Boston, MA, 2009.

[76] V. Nwakacha, A. Crabtree And G. Burnett, "Evaluating Distraction And Disengagement Of Attention From The Road," In *Virtual, Augmented And Mixed Reality. Systems And Applications*, Springer-Verlag Berlin Heidelberg , 2013, Pp. 261-270.

[77] . R. PANKO, "The Popularity Of Google Maps: Trends In Navigation Apps In 2018," 10 July 2018. [Online]. Available: Https://Themanifest.Com/App-Development/Popularity-Google-Maps-Trends-Navigation-Apps-2018.

[79] G. Leshed, T. Velden, O. Rieger, B. Kot And P. Sengers, "In-Car Gps Navigation: Engagement With And Disengagement From The Environment," In *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems*, 2008.

[80] P. M. Ververs And C. D. Wickens, "Conformal Flight Path Symbology For Head-Up Displays: Defining The Distribution Of Visual Attention In Three-Dimensional Space," Aviation Research Lab Institute Of Aviation, Illinois, 1998.

[81] H. Abdi, "Signal Detection Theory (SDT)," 02 October 2018. [Online]. Available: Https://Www.Utdallas.Edu/~Herve/Abdi-SDT_2009.Pdf.

[82] A.-H. Javadi, B. Emo, L. R. Howard, F. E. Zisch, Y. Yu, R. Knight, J. P. Silva And H. J. Spiers, "Hippocampal And Prefrontal Processing Of Network Topology To Simulate The Future," *Nature Communications,* Vol. 8, 2017.

[83] V. D. Bohbot, J. Lerch, B. Thorndycraft, G. Iaria And A. P. Zijdenbos, "Gray Matter Differences Correlate With Spontaneous Strategies In A Human Virtual Navigation Task," *The Journal Of Neuroscience,* Vol. 27, No. 38, Pp. 10078-10083, 2007.

APPENDIX

A. DOCUMENTS

A1. IRB APPROVAL

VIRGINIA TECH.

Office of Research Compliance
Institutional Review Board
North End Center, Suite 4120
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732 Fax 540/231-0959
email irb@vt.edu
website http://www.irb.vt.edu

MEMORANDUM

DATE:	June 1, 2018
TO:	Joseph L Gabbard Jr, Martha Irene Smith, Kiran Bagalkotkar, Rishikesh Nittala, Sarah Y Lee, Kevin Key, Richard Leslie Greatbatch, Nayara De Oliveira Faria, PRANAY SHAH Shah
FROM:	Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)
PROTOCOL TITLE:	AR Navigation Cues in a Driving Simulator
IRB NUMBER:	**17-239**

Effective June 1, 2018, the Virginia Tech Institution Review Board (IRB) approved the Amendment request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As:	**Expedited, under 45 CFR 46.110 category(ies) 4,6,7**
Protocol Approval Date:	**September 13, 2017**
Protocol Expiration Date:	**September 12, 2018**
Continuing Review Due Date*:	**August 29, 2018**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

Date*	OSP Number	Sponsor	Grant Comparison Conducted?
07/06/2017	P7I7HRM5	University Of Nottingham (Title: Augmented Reality-based HMI Design for Future Vehicles)	Not required (Not federally funded)

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the IRB office (irbadmin@vt.edu) immediately.

A2. INFORMED CONSENT FORM

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
Informed Consent for Participants
in Research Projects Involving Human Subjects

Title of Project: AR Navigation Cues in a Driving Simulator

Investigator(s):	Name	E-mail / Phone number
	Missie Smith, Virginia Tech	mis16@vt.edu
	Dr. Joe Gabbard, Virginia Tech	jgabbard@vt.edu
	Kiran Bagalkotkar, Virginia Tech	kiranb@vt.edu
	Rishi Nittala, Virginia Tech	rishin1@vt.edu
	Sarah Lee, Virginia Tech	Sylee95@vt.edu
	Kevin Key, Virginia Tech	kkey13@vt.edu
	Richard Greatbatch, Virginia Tech	rlg1990@vt.edu
	Nayara de Oliveira Faria, Virginia Tech	nfaria@vt.edu

I. Purpose of this Research Project

The purpose of this project is to understand the difference in individual behavioral response between different navigation cues presented via a head-up display (HUD) while in simulated driving scenarios. The research will be shared with the funding institution, and may be published at a later time.

II. Procedures

You will be asked to drive a simulator vehicle, following the instructions that will be provided to you prior to testing. While driving, you will be asked to perform a visual search task repeatedly. You will be asked to continue driving as normally as possible while performing these tasks. You will be asked to provide verbal responses to these tasks.

You will drive the vehicle in a series of short trips while wearing eye-tracking glasses. During each short drive, eye-tracking data (includes audio recording of the room, forward-facing video, and video recording of your eyes) and video data of the room will be collected. The only identifying audio and video recorded will be your eyes and your voice. Upon completion of each trip, you will be asked to fill out a short survey. The total process is expected to take no more than 2 hours to complete.

III. Risks

You should be aware that there is some risk of induced motion sickness when using a driving simulator – if you are particularly susceptible to motion sickness, migraines, epilepsy, dizziness or blurred vision then you should not take part. If you do take part, you are free to stop the evaluation at any time for any reason, without penalty. In the event of your withdrawal, any data collected will be securely stored along with data from other participants. We advise that you wait for 30 minutes before driving your own car after the end of the study. Any expenses accrued for seeking or receiving medical or mental health treatment will be the responsibility of the subject and not that of the research project, research team, or Virginia Tech.

IV. Benefits

The data collected in this research can be used to help researchers better understand inherent differences between AR head-up navigation cues, and the knowledge gained can be used to design new driving interfaces that are safer than those that are on the market now.

No promise or guarantee of benefits has been made to encourage you to participate.

V. Extent of Anonymity and Confidentiality

No personal information will be included in publications and presentations. Please note that these records will be held by a state entity and therefore are subject to disclosure if required by law.

The Virginia Tech (VT) Institutional Review Board (IRB) may view the study's data for auditing purposes. The IRB is responsible for the oversight of the protection of human subjects involved in research.

Note: in some situations, it may be necessary for an investigator to break confidentiality. If a researcher has reason to suspect that a child is abused or neglected, or that a person poses a threat of harm to others or him/herself, the researcher is required by Virginia State law to notify the appropriate authorities. If applicable to this study, the conditions under which the investigator must break confidentiality must be described.

VI. Compensation

You will be compensated with $10 upon completion of your participation.

VII. Freedom to Withdraw

It is important for you to know that you are free to withdraw from this study at any time without penalty. You are free not to answer any questions that you choose or respond to what is being asked of you without penalty.

Please note that there may be circumstances under which the investigator may determine that a subject should not continue as a subject.

Should you withdraw or otherwise discontinue participation, you will be compensated for the portion of the project completed in accordance with the Compensation section of this document.

VIII. Questions or Concerns

Should you have any questions about this study, you may contact one of the research investigators whose contact information is included at the beginning of this document.

Should you have any questions or concerns about the study's conduct or your rights as a research subject, or need to report a research-related injury or event, you may contact the

Virginia Tech Institutional Review Board at irb@vt.edu or (540) 231-3732.

IX. Subject's Consent

I have read the Consent Form and conditions of this project. I have had all my questions answered. Further,

- I agree to take part in the above study and am willing to:
  - Participate in a car driving simulation activity
  - Be recorded in audio and/or video whilst engaged in the study

- I understand that my information will be held in line with Virginia Tech privacy policies and used for the following purposes:
  - Investigation into vehicle display systems.
  - Dissemination of findings through academic publications.

- I am aware of the potential risks associated with using the driving simulator and confirm that I do not suffer from any of the aforementioned conditions.

- I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without being penalized or disadvantaged in any way.

I hereby acknowledge the above and give my voluntary consent:

____ Date__________
Subject signature

____
Subject printed name

A3. STUDY ADVERTISEMENT FLYER

PAID PARTICIPANTS NEEDED
AR DRIVING SIMULATOR STUDY

SIGN UP HERE

OR SEND AN EMAIL TO: NFARIA@VT.EDU

The COGENT Lab at Virginia Tech (https://cogent.ise.vt.edu/) needs volunteers for a research study comparing augmented reality vehicle displays and how they impact drivers. During the study, you will be asked to drive a mini cooper in a driving simulator wearing eye-tracking glasses. After your driving, you will be asked to answer a questionnaire and rate your experience.

You are eligible to participate if:

- Active driver with valid US driver's license
- Drive a minimum of 3,000 miles per year
- Not an employee of any vehicle manufacturers
- Age 18+
- Have normal or corrected-to-normal vision (after corrected using contact lenses)

Be aware!

Please be aware that there is a risk of induced motion sickness when using a driving simulator. If you are susceptible to motion sickness, migraines, epilepsy, dizziness or blurred vision, or are pregnant, then you should not volunteer.

The study should last no more than 1 hour, and you will receive $10 compensation. Participation is voluntary and confidential. If you are interested in participating, sign up the Doodle form using the QR code above or contact **Nayara Faria at nfaria@vt.edu.**

AR Driving Simulator Study Contact: Nayara Faria nfaria@vt.edu	**AR Driving Simulator Study** Contact: Nayara Faria nfaria@vt.edu	**AR Driving Simulator Study** Contact: Nayara Faria nfaria@vt.edu	**AR Driving Simulator Study** Contact: Nayara Faria nfaria@vt.edu	**AR Driving Simulator Study** Contact: Nayara Faria nfaria@vt.edu	**AR Driving Simulator Study** Contact: Nayara Faria nfaria@vt.edu	**AR Driving Simulator Study** Contact: Nayara Faria nfaria@vt.edu	**AR Driving Simulator Study** Contact: Nayara Faria nfaria@vt.edu

B. SURVEYS

B1. DEMOGRAPHICS SURVEY

Gender:

- Male
- Female
- Other

How old are you?

____

Are you predominantly left- or right-handed?

- left handed
- right handed

What is your nationality?

____

How many years have you held a full U.S. driver's license?

____

How many years have you held a non-U.S. driver's license?

____

Approximately how far (in **miles**) do you weekly drive?

____

What is the make, model and year of the car you normally drive?

____

Does the car you normally drive have a center console large screen display?

- ◯ Yes, factory-fitted
- ◯ Yes, portable device (could be a Garmin or cell phone for navigation)
- ◯ No

Does the car you normally drive have a touchscreen device installed?

- ◯ Yes, factory-fitted (1)
- ◯ Yes, portable device (2)
- ◯ No (3)

Does the car you normally drive have a head-up display (HUD) device installed? A HUD provides information very near the field of view of the driver, allowing the driver to quickly gather information without looking away from the road.

- ◯ Yes, factory-fitted
- ◯ Yes, portable device
- ◯ No

Do you have a smart phone? If yes, please list the 2 most common ways that you use it.

Example: texting, GPS, calling, social media

- ◯ Yes (1) ____
- ◯ No (2)

B2. NASA-TASK LOAD INDEX (NASA-TLX)

	Very Low/Easy Very High/Difficult 0 10 20 30 40 50 60 70 80 90 100
Mental Demand: How mentally demanding was the task? ()	
Physical Demand: How physically demanding was the task? ()	
Temporal Demand: How hurried or rushed was the pace of work? ()	
Effort: How hard did you have to work to accomplish your level of performance? ()	
Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you? ()	

	Perfect Failure 0 10 20 30 40 50 60 70 80 90 100
Performance: How successful were you in accomplishing what you were asked to do? ()	

B3. USABILITY QUESTIONNAIRE

How much do you agree or disagree with the following statements about your most recent drive?

	Strongly Agree	Somewhat Agree	Slightly Agree	Slightly Disagree	Somewhat Disagree	Strongly Disagree

0 10 20 30 40 50 60 70 80 90 100

It was easy to navigate while using this interface. ()	
The interface was easy to view. ()	
I trusted the information that the interface was telling me. ()	
Using this interface had a positive impact on my driving. ()	
I did not find this interface distracting. ()	

B4. SITUATION AWARENESS RATING TECHNIQUE - SART

How changeable is the situation? Is the situation highly unstable and likely to change suddenly (High) or is it very stable and straightforward (Low)?

1 2 3 4 4 5 6 7

Instability of Situation ()	

How complicated is the situation? Is it complex with many interrelated components (High) or is it simple and straightforward (Low)?

1 2 3 4 4 5 6 7

Complexity of Situation ()	

How many variables are changing within the situation? Are there a large number of factors varying (High) or are there very few variables changing (Low)?

1 2 3 4 4 5 6 7

Variability of Situation ()	

How aroused are you in this situation? Are you alert and ready for activity (High) or do you have a low degree of alertness (Low)?

1 2 3 4 4 5 6 7

Arousal ()	

How much are you concentrating on the situation? Are you concentrating on many aspects of the situation (High) or focused on only one (Low)?

1 2 3 4 4 5 6 7

Concentration of Attention ()	

How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (High) or focused on only one (Low)?

	1	2	3	4	4	5	6	7
Division of Attention ()								

How much mental capacity to you have to spare in the situation? Do you have sufficient to attend to many variables (High) or nothing to spare at all (Low)?

	1	2	3	4	4	5	6	7
Spare Mental Capacity ()								

How much information have you gained about the situation? Have you received and understood a great deal of knowledge (High) or very little (Low)?

	1	2	3	4	4	5	6	7
Information Quantity ()								

How familiar are you with the situation? Do you have a great deal of relevant experience (High) or is it a new situation (Low)?

	1	2	3	4	4	5	6	7
Familiarity with the Situation ()								

B.5 LANDMARK KNOWLEDGE SURVEY IMAGES AND AOI

NAME ID	O_ST	ST
#1: START OF DRIVING TRIAL		
	• St. Mary's Hospital sign • Brown Building on the right corner • Brown Building on the left corner • White Building on the center • Street Light Post	• Building in the left corner • Advertisement sign • St. Mary's Hospital sign • White building

NAME ID	O_S1	S1
#2: GO STRAIGHT #1		
	• Brown Bulding on the left corner • Traffic sign on the right corner • Grey Bulding on the right far corner • Street Name on the left corner • Traffic Light • Light Brwon Bulding on the right corner	• Blue Building on the left corner • Traffic Light • Brown Building on the Right Corner • Street Name • Brown Building on the left far corner • Brown/ Blue Building on the right far corner

<table>
<tr><th>NAME ID</th><th>O_M2_L</th><th>M2_L</th></tr>
<tr><td rowspan="2">#3: MANEUVER #2 TO THE LEFT</td><td></td><td></td></tr>
<tr><td>• Salon Spa Bulding on the left cor
• Brown Bulding on the right far corner
• Traffic Lights
• Street Name right corner
• Light Post
• Street Name left corner</td><td>• Street Name
• Burger King
• Traffic Light
• Brown building on the Right Far Corner</td></tr>
</table>

<table>
<tr><th>NAME ID</th><th>O_M4_R</th><th>M4_R</th></tr>
<tr><td rowspan="2">#4: MANEUVER #4 TO THE RIGHT</td><td></td><td></td></tr>
<tr><td>• Brown Bulding on the left corner
• Gas Station
• Street Name
• Advertisement Sign on the left Corner
• Traffic Lights
• US Bank Bulding
• US Bank Sign</td><td>• Village Inn
• McDonalds
• Traffic Light
• Street Name
• Shell Advertisement sign
• Shell Gas Station</td></tr>
</table>

<table>
<tr><th>NAME ID</th><th>O_R1_R</th><th>R1_R</th></tr>
<tr><td rowspan="2">#5: ROUDABOUNT #1</td><td>
</td><td></td></tr>
<tr><td>Roundabout SIgn
Speed Limit Sign
Roundabout
Pedestrian Crossing Sign
Truck</td><td>Crossing sign
Roundabout</td></tr>
</table>

<table>
<tr><th>NAME ID</th><th>O_S2</th><th>S2</th></tr>
<tr><td rowspan="2">#6: GO STRAIGHT #2</td><td>
</td><td>
</td></tr>
<tr><td>Wheelock Welding Bulding
Wheelock Welding Sign
Street Name
Traffic Light
Fence on the left far corner
Brown Bulding on the right far corner
Grey Bulding on the right corner</td><td>Brown building on the left corner
Grey building on the right corner
White building on the back
Street Name
Traffic Light</td></tr>
</table>

NAME ID	O_ M5_L	M5_L
#7: MANEUVER #5 TO THE LEFT		
	• Brown Bulding on the right corner • Grey Bulding in the center • Brown Bulding on the left corner • Must turn traffic sign	• Gas Station • Brown building on the right corner • Bank • Grey building on the left far corner • Grey building on the center • Fast food building on the far corner • Grey building on the right far corner • Parking sign

NAME ID	O_ M7_R	M7_R
#8: MANEUVER #7 TO THE RIGHT		
	• Store on te left corner • Gas Station • Grey bulding on the left far corner • Street Name • Bank Bulding • Bulding on the right corner • Traffic Light	• Short Stop Store • Advertisement sign on the left • Gas Station • Bennigan's store • Grey building on the center • Store on the Right far corner • Target • Store before target

C. DATA COLLECTION: RAW DATA

C1. NASA TLX RAW DATA

Participant #	Condition	Gender	Mental Demand	Physical Demand	Temporal Demand	Effort	Frustration	Performance
4001	screen-relative	Male	0	1	1	1	1	19
4003	screen-relative	Male	30	34	30	40	22	70
4004	screen-relative	Male	76	34	46	75	66	73
4005	screen-relative	Male	25	5	20	40	5	20
4006	screen-relative	Male	45	15	15	25	15	10
4008	screen-relative	Female	9	30	0	22	71	77
4009	screen-relative	Female	15	14	55	15	56	13
4011	screen-relative	Female	84	79	18	81	82	48
4012	screen-relative	Female	29	30	17	52	40	80
4014	world-relative	Male	56	30	25	60	5	20
4015	world-relative	Male	20	5	0	20	10	85
4016	world-relative	Male	40	30	10	23	24	20
4017	world-relative	Male	38	24	18	32	24	60
4018	world-relative	Male	30	39	50	73	89	60
4019	world-relative	Female	77	25	40	64	51	50
4021	world-relative	Female	49	29	7	71	32	49
4022	world-relative	Female	90	90	45	93	62	98
4023	world-relative	Female	45	20	26	20	9	66
4024	world-relative	Female	10	20	10	10	10	20
4002A	screen-relative	Male	50	30	30	62	86	7
4007A	screen-relative	Female	4	3	3	15	2	82
4010A	screen-relative	Female	50	51	6	61	61	30
4013A	world-relative	Male	20	10	50	40	40	40
4020B	world-relative	Female	55	15	22	48	70	50

C2. SART RAW DATA

Participant # :	Condition	Gender:	Instability of Situation	Complexity of Situation	Variability of Situation	Arousal	Concentration of Attention	Division of Attention	Spare Mental Capacity	Information Quantity	Familiarity with the Situation
4001	screen-relative	Male	1	1	3	2	7	5	3	6	5
4003	screen-relative	Male	4	2	4	6	6	5	5	5	5
4004	screen-relative	Male	4	4	5	6	5	5	4	5	7
4005	screen-relative	Male	3	1	2	4	5	5	7	6	2
4006	screen-relative	Male	1	3	2	4	4	2	3	4	5
4008	screen-relative	Female	3	1	2	5	5	4	3	3	4
4009	screen-relative	Female	5	2	2	5	4	3	5	3	4
4011	screen-relative	Female	5	4	5	5	5	5	6	6	4
4012	screen-relative	Female	5	2	3	4	5	2	3	3	3
4014	world-relative	Male	2	3	3	6	7	4	6	4	4
4015	world-relative	Male	1	2	2	6	6	4	4	5	4
4016	world-relative	Male	4	2	4	4	6	4	4	4	4
4017	world-relative	Male	4	4	2	3	6	5	5	4	3
4018	world-relative	Male	2	3	3	5	5	5	5	2	4
4019	world-relative	Female	4	2	4	3	4	2	3	5	5
4021	world-relative	Female	1	1	1	4	4	4	6	5	3
4022	world-relative	Female	4	5	5	7	7	2	1	3	1
4023	world-relative	Female	2	1	2	5	6	5	4	6	4
4024	world-relative	Female	1	1	1	4	4	2	4	2	2
4002A	screen-relative	Male	4	3	4	5	5	5	4	3	3
4007A	screen-relative	Female	2	1	2	5	4	4	5	3	6
4010A	screen-relative	Female	2	2	5	7	7	5	4	5	4
4013A	world-relative	Male	4	4	3	7	4	4	3	4	5
4020B	world-relative	Female	5	4	1	4	3	4	4	4	1

C3. USABILITY RAW DATA

Participant # :	Condition	Gender	easy to navigate	easy to view	Trustworthy information	positive impact on driving	distraction
4001	screen-relative	Male	10	0	0	10	0
4003	screen-relative	Male	10	20	10	20	20
4004	screen-relative	Male	40	30	30	60	70
4005	screen-relative	Male	10	0	0	0	10
4006	screen-relative	Male	0	0	0	30	70
4008	screen-relative	Female	60	70	70	80	100
4009	screen-relative	Female	20	20	60	40	40
4011	screen-relative	Female	70	70	20	70	70
4012	screen-relative	Female	30	80	10	60	0
4014	world-relative	Male	30	30	10	40	30
4015	world-relative	Male	10	20	0	30	80
4016	world-relative	Male	50	60	20	60	40
4017	world-relative	Male	70	50	20	50	40
4018	world-relative	Male	60	10	40	80	70
4019	world-relative	Female	40	20	0	90	0
4021	world-relative	Female	60	20	0	0	20
4022	world-relative	Female	80	40	0	70	90
4023	world-relative	Female	10	0	10	20	20
4024	world-relative	Female	20	20	0	20	10
4002A	screen-relative	Male	50	20	60	80	100
4007A	screen-relative	Female	0	10	0	10	10
4010A	screen-relative	Female	10	10	0	50	20
4013A	world-relative	Male	20	40	0	10	50
4020B	world-relative	Female	90	30	50	20	0

C4. LANDMARK KNOWLEDGE RAW DATA

I DO NOT REMEMBER PILE													
Participant #	Condition	Gender	Images Selected										
4001	Gen2	Male	ST	S1	M5_L	M7_R	O_ST	O_S1	O_M2_L	O_M4_R			
4002A	Gen2	Male	ST	M4_R	R1_R	M7_R	O_S1	O_M2_L	O_M4_R	O_M5_L	O_M7_R		
4003	Gen2	Male	ST	S1	M2_L	M4_R	R1_R	M7_R	O_S1	O_M4_R	O_S2		
4004	Gen2	Male	ST	M2_L	M4_R	M5_L	M7_R	O_S1	O_M4_R	O_M7_R			
4005	Gen2	Male	ST	S1	M4_R	M5_L	M7_R	O_S2	O_M7_R				
4006	Gen2	Male	ST	S1	M2_L	M4_R	O_ST	O_S1	O_M2_L	O_M4_R	O_S2	O_M7_R	
4007	Gen2	Female	ST	M4_R	M5_L	M7_R	O_ST	O_M2_L	O_M7_R				
4008	Gen2	Female	M4_R	R1_R	M5_L	M7_R	O_ST	O_M4_R	O_S2				
4009	Gen2	Female	M4_R	R1_R	M5_L	M7_R	O_ST	O_M2_L	O_M4_R	O_S2	O_M7_R		
4010A	Gen2	Female	ST	S2	M7_R	O_ST	O_S1	O_M4_R	O_M5_L	O_M7_R			
4011	Gen2	Female	ST	M2_L	M4_R	R1_R	M7_R	O_ST	O_M2_L	O_M4_R	O_S2	O_M7_R	
4012	Gen2	Female	S2	M5_L	M7_R	O_ST	O_S1	O_M2_L	O_R1_R	O_S2			
4013	Post Sign	Male	M2_L	M4_R	S2	M7_R	O_ST	O_S1	O_M2_L	O_M4_R	O_S2	O_M7_R	
4014	Post Sign	Male	ST	M2_L	M4_R	M5_L	M7_R	O_ST	O_M4_R	O_R1_R	O_S2	O_M7_R	
4015	Post Sign	Male	ST	S1	M2_L	R1_R	O_ST	O_S1	O_M4_R	O_M7_R	M4_R		
4016	Post Sign	Male	M2_L	M4_R	M7_R	O_ST	O_S1	O_M2_L	O_M4_R	O_M7_R			
4017	Post Sign	Male	S1	M4_R	S2	M7_R	O_ST	O_S1	O_M4_R	O_M5_L			
4018	Post Sign	Male	M2_L	M4_R	M7_R	O_ST	O_M4_R	O_M5_L	O_R1_R				
4019	Post Sign	Female	ST	S1	M4_R	S2	O_ST	O_S1	O_S2				
4020B	Post Sign	Female	S1	M2_L	M4_R	S2	M5_L	M7_R	O_ST	O_M2_L	O_M4_R	O_M7_R	
4021	Post Sign	Female	ST	M2_L	R1_R	M5_L	M7_R	O_ST	O_M2_L	O_M4_R			
4022	Post Sign	Female	ST	M2_L	M4_R	O_S1	O_M4_R						
4023	Post Sign	Female	ST	S1	M2_L	M7_R	O_S1	O_M4_R	O_S2				
4024	Post Sign	Female	ST	S1	M2_L	M4_R	M5_L	M7_R	O_ST	O_M4_R	O_S2	O_M5_L	O_M7_R

I REMEMBER PILE													
Participant #	**Condition**	**Gender**	**Images**										
4001	Gen2	Male	M2_L	M4_R	R1_R	S2	M7_R	O_R1_R	O_S2	O_M5_L			
4002A	Gen2	Male	S1	M2_L	S2	M5_L	O_ST	O_R1_R	O_S2				
4003	Gen2	Male	S2	M5_L	O_ST	O_M2_L	O_R1_R	O_M5_L	O_M7_R				
4004	Gen2	Male	S1	R1_R	S2	O_ST	O_M2_L	O_R1_R	O_S2	O_M5_L			
4005	Gen2	Male	M2_L	R1_R	S2	O_ST	O_S1	O_M2_L	O_M4_R	O_R1_R	O_M5_L		
4006	Gen2	Male	R1_R	S2	M5_L	M7_R	O_R1_R	O_M5_L					
4007	Gen2	Female	S1	M2_L	R1_R	S2	O_S1	O_M4_R	O_R1_R	O_S2	O_M5_L		
4008	Gen2	Female	ST	S1	M2_L	S2	O_S1	O_M2_L	O_R1_R	O_M5_L	O_M7_R		
4009	Gen2	Female	ST	S1	M2_L	S2	O_S1	O_R1_R	O_M5_L				
4010A	Gen2	Female	S1	M2_L	M4_R	R1_R	M5_L	O_M2_L	O_R1_R	O_S2			
4011	Gen2	Female	S1	S2	M5_L	O_S1	O_R1_R	O_M5_L					
4012	Gen2	Female	ST	S1	M2_L	M4_R	R1_R	O_M4_R	O_M5_L	O_M7_R			
4013	Post Sign	Male	ST	S1	R1_R	M5_L	O_R1_R	O_M5_L					
4014	Post Sign	Male	S1	R1_R	S2	O_S1	O_M2_L	O_M5_L					
4015	Post Sign	Male	S2	M5_L	M7_R	O_M2_L	O_R1_R	O_S2	O_M5_L				
4016	Post Sign	Male	ST	S1	R1_R	S2	M5_L	O_R1_R	O_S2	O_M5_L			
4017	Post Sign	Male	ST	M2_L	R1_R	M5_L	O_M2_L	O_R1_R	O_S2	O_M7_R			
4018	Post Sign	Male	ST	S1	R1_R	S2	M5_L	O_S1	O_M2_L	O_S2	O_M7_R		
4019	Post Sign	Female	M2_L	R1_R	M5_L	M7_R	O_M2_L	O_M4_R	O_R1_R	O_M5_L	O_M7_R		
4020B	Post Sign	Female	ST	R1_R	O_S1	O_R1_R	O_S2	O_M5_L					
4021	Post Sign	Female	S1	M4_R	S2	O_S1	O_R1_R	O_S2	O_M5_L	O_M7_R			
4022	Post Sign	Female	S1	R1_R	S2	M5_L	M7_R	O_ST	O_M2_L	O_R1_R	O_S2	O_M5_L	O_M7_R
4023	Post Sign	Female	M4_R	R1_R	S2	M5_L	O_ST	O_M2_L	O_R1_R	O_M5_L	O_M7_R		
4024	Post Sign	Female	R1_R	S2	O_S1	O_M2_L	O_R1_R						

C5. ROUTE KNOWLEDGE RAW DATA

			Image/ Order										
Participant #	Condition	Gender	ST	S1	M2_L	M3_R	M4_R	R1_R	M5_L	M6_L	S3	M7_R	ED
4001	Gen2	Male	8	1	4	11	6	5	9	10	2	7	3
4002A	Gen2	Male	3	4	7	1	5	8	11	9	6	2	10
4003	Gen2	Male	7	1	4	11	9	6	3	10	5	2	8
4004	Gen2	Male	1	8	5	6	3	10	4	7	2	9	11
4005	Gen2	Male	2	3	11	6	7	1	10	9	5	8	4
4006	Gen2	Male	5	9	4	6	1	7	3	10	8	2	11
4007	Gen2	Female	1	9	7	4	5	8	3	11	10	6	2
4008	Gen2	Female	3	2	9	4	11	7	10	8	1	5	6
4009	Gen2	Female	7	8	3	2	6	5	10	4	1	9	11
4010A	Gen2	Female	10	5	4	7	2	8	6	9	1	3	11
4011	Gen2	Female	1	4	10	5	9	6	3	7	2	8	11
4012	Gen2	Female	5	4	3	7	9	6	1	11	2	10	8
4013	Post Sign	Male	3	1	7	10	4	8	6	9	2	5	11
4014	Post Sign	Male	1	6	11	4	8	7	2	10	5	9	3
4015	Post Sign	Male	1	2	10	9	4	7	5	11	3	8	6
4016	Post Sign	Male	3	8	10	4	5	9	2	1	7	6	11
4017	Post Sign	Male	1	3	2	8	10	6	4	7	11	9	5
4018	Post Sign	Male	1	8	6	4	3	5	2	11	7	10	9
4019	Post Sign	Female	2	1	7	11	9	5	4	10	8	3	6
4020B	Post Sign	Female	2	3	4	10	5	9	8	6	1	7	11
4021	Post Sign	Female	1	9	3	8	2	7	10	4	5	6	11
4022	Post Sign	Female	9	7	6	2	8	4	5	10	1	3	11
4023	Post Sign	Female	1	4	9	10	2	6	7	11	5	3	8
4024	Post Sign	Female	1	2	10	4	11	9	7	8	3	6	5

C6. DRIVING MEASURES

SD Lane Position and SD Speed by Participant, Gender, Event and Condition						
Standard Deviation Speed	**Standard Deviation Lane Position**	**Participant#**	**Gender**	**Event**	**Condition**	**Manuever Direction**
1.15476	0.7346	4001	male	1	Screen-Relative	Straight
5.84063	0.325827	4001	male	2	Screen-Relative	Right
8.31021	2.45865	4001	male	3	Screen-Relative	Left
14.2454	0.292445	4001	male	4	Screen-Relative	Right
13.2581	1.0666	4001	male	5	Screen-Relative	Right
11.3839	2.64486	4001	male	6	Screen-Relative	Roundabout-Right
11.9445	0.304396	4001	male	7	Screen-Relative	Straight
13.6339	0.232049	4001	male	8	Screen-Relative	Left
2.45612	2.58796	4001	male	9	Screen-Relative	Left
15.903	0.543757	4001	male	10	Screen-Relative	Straight
14.3809	1.05394	4001	male	11	Screen-Relative	Right
15.5036	2.03288	4001	male	12	Screen-Relative	Left
1.93188	0.346127	4002	male	1	Screen-Relative	Straight
11.207	0.381485	4002	male	2	Screen-Relative	Right
14.1555	1.44619	4002	male	3	Screen-Relative	Left
8.08401	1.77945	4002	male	4	Screen-Relative	Right
15.1864	0.289314	4002	male	5	Screen-Relative	Right
5.78654	2.02768	4002	male	6	Screen-Relative	Roundabout-Right
14.3535	0.569936	4002	male	7	Screen-Relative	Straight
13.0847	1.49655	4002	male	8	Screen-Relative	Left
10.8798	2.72258	4002	male	9	Screen-Relative	Left
13.3099	0.63146	4002	male	10	Screen-Relative	Straight
6.11704	1.9381	4002	male	11	Screen-Relative	Right
14.4466	1.58754	4002	male	12	Screen-Relative	Left
1.42517	0.354971	4003	male	1	Screen-Relative	Straight
12.9079	0.257956	4003	male	2	Screen-Relative	Right
10.2788	1.32738	4003	male	3	Screen-Relative	Left
12.1351	2.54787	4003	male	4	Screen-Relative	Right
8.8144	0.852787	4003	male	5	Screen-Relative	Right
9.90229	2.78024	4003	male	6	Screen-Relative	Roundabout-Right
1.22652	0.601937	4003	male	7	Screen-Relative	Straight
17.8955	1.86307	4003	male	8	Screen-Relative	Left
15.0028	1.92455	4003	male	9	Screen-Relative	Left
0.769662	0.426699	4003	male	10	Screen-Relative	Straight
11.9273	0.824461	4003	male	11	Screen-Relative	Right
16.8355	2.34733	4003	male	12	Screen-Relative	Left
2.95415	0.420098	4004	male	1	Screen-Relative	Straight

10.6112	0.380729	4004	male	2	Screen-Relative	Right
9.55005	1.56741	4004	male	3	Screen-Relative	Left
8.28957	0.769117	4004	male	4	Screen-Relative	Right
10.3443	0.597037	4004	male	5	Screen-Relative	Right
5.08218	2.06082	4004	male	6	Screen-Relative	Roundabout-Right
5.73365	0.769828	4004	male	7	Screen-Relative	Straight
14.1427	1.01655	4004	male	8	Screen-Relative	Left
12.1081	2.04641	4004	male	9	Screen-Relative	Left
13.8434	0.600347	4004	male	10	Screen-Relative	Straight
13.5004	0.143601	4004	male	11	Screen-Relative	Right
8.24903	1.99942	4004	male	12	Screen-Relative	Left
2.88307	0.237322	4005	male	1	Screen-Relative	Straight
7.83284	0.337223	4005	male	2	Screen-Relative	Right
12.444	1.81042	4005	male	3	Screen-Relative	Left
5.12696	2.1242	4005	male	4	Screen-Relative	Right
11.1869	0.168733	4005	male	5	Screen-Relative	Right
6.20645	2.23504	4005	male	6	Screen-Relative	Roundabout-Right
9.07647	0.73962	4005	male	7	Screen-Relative	Straight
14.5036	1.60496	4005	male	8	Screen-Relative	Left
15.4403	1.23291	4005	male	9	Screen-Relative	Left
13.7422	0.324126	4005	male	10	Screen-Relative	Straight
8.49492	0.270242	4005	male	11	Screen-Relative	Right
10.9192	1.78973	4005	male	12	Screen-Relative	Left
0.866705	0.166871	4006	male	1	Screen-Relative	Straight
7.66411	0.262027	4006	male	2	Screen-Relative	Right
13.798	1.4172	4006	male	3	Screen-Relative	Left
4.01829	2.15998	4006	male	4	Screen-Relative	Right
6.1689	0.677957	4006	male	5	Screen-Relative	Right
8.87938	1.58546	4006	male	6	Screen-Relative	Roundabout-Right
11.246	0.339435	4006	male	7	Screen-Relative	Straight
14.2768	1.39646	4006	male	8	Screen-Relative	Left
15.7743	0.958077	4006	male	9	Screen-Relative	Left
2.03206	0.531674	4006	male	10	Screen-Relative	Straight
14.8407	0.280587	4006	male	11	Screen-Relative	Right
9.33684	2.03457	4006	male	12	Screen-Relative	Left
1.44416	0.530252	4007	female	1	Screen-Relative	Straight
10.6784	0.337109	4007	female	2	Screen-Relative	Right
11.6702	1.78245	4007	female	3	Screen-Relative	Left
10.6204	0.245126	4007	female	4	Screen-Relative	Right
11.6301	0.387827	4007	female	5	Screen-Relative	Right
7.57992	1.88652	4007	female	6	Screen-Relative	Roundabout-Right
15.7166	0.276806	4007	female	7	Screen-Relative	Straight
16.9516	1.75002	4007	female	8	Screen-Relative	Left

9.39569	1.87732	4007	female	9	Screen-Relative	Left
12.8003	0.545076	4007	female	10	Screen-Relative	Straight
12.6088	0.441865	4007	female	11	Screen-Relative	Right
11.201	2.30856	4007	female	12	Screen-Relative	Left
0.697507	0.250242	4008	female	1	Screen-Relative	Straight
13.2189	0.675982	4008	female	2	Screen-Relative	Right
14.2173	1.47472	4008	female	3	Screen-Relative	Left
12.9161	0.299161	4008	female	4	Screen-Relative	Right
13.5826	0.0578304	4008	female	5	Screen-Relative	Right
4.76949	2.32485	4008	female	6	Screen-Relative	Roundabout-Right
12.7762	0.263904	4008	female	7	Screen-Relative	Straight
13.7983	1.70709	4008	female	8	Screen-Relative	Left
17.5685	1.54661	4008	female	9	Screen-Relative	Left
14.2098	0.550402	4008	female	10	Screen-Relative	Straight
14.9958	0.480889	4008	female	11	Screen-Relative	Right
11.1907	2.07812	4008	female	12	Screen-Relative	Left
1.23787	0.753356	4009	female	1	Screen-Relative	Straight
6.43779	0.74049	4009	female	2	Screen-Relative	Right
10.7796	2.07322	4009	female	3	Screen-Relative	Left
6.31342	2.44919	4009	female	4	Screen-Relative	Right
3.16921	0.782018	4009	female	5	Screen-Relative	Right
6.19964	2.67493	4009	female	6	Screen-Relative	Roundabout-Right
14.2023	0.275627	4009	female	7	Screen-Relative	Straight
14.3357	2.13763	4009	female	8	Screen-Relative	Left
14.6191	1.56253	4009	female	9	Screen-Relative	Left
7.78782	0.35886	4009	female	10	Screen-Relative	Straight
10.2677	0.362799	4009	female	11	Screen-Relative	Right
5.62369	2.61994	4009	female	12	Screen-Relative	Left
3.32024	0.403263	4010	female	1	Screen-Relative	Straight
11.5887	0.566637	4010	female	2	Screen-Relative	Right
16.6114	1.87069	4010	female	3	Screen-Relative	Left
12.8557	2.53921	4010	female	4	Screen-Relative	Right
10.7046	2.21799	4010	female	5	Screen-Relative	Right
9.41063	2.88763	4010	female	6	Screen-Relative	Roundabout-Right
13.8074	1.60855	4010	female	7	Screen-Relative	Straight
16.4819	0.735692	4010	female	8	Screen-Relative	Left
13.5217	2.4743	4010	female	9	Screen-Relative	Left
0.570111	1.0385	4010	female	10	Screen-Relative	Straight
14.732	2.36281	4010	female	11	Screen-Relative	Right
9.80688	3.00163	4010	female	12	Screen-Relative	Left
1.53055	0.257439	4011	female	1	Screen-Relative	Straight
10.4856	0.662763	4011	female	2	Screen-Relative	Right
8.3232	2.4598	4011	female	3	Screen-Relative	Left

8.60719	2.56013	4011	female	4	Screen-Relative	Right
11.4484	0.686793	4011	female	5	Screen-Relative	Right
8.16988	2.42046	4011	female	6	Screen-Relative	Roundabout-Right
1.08418	1.17943	4011	female	7	Screen-Relative	Straight
10.1797	0.470875	4011	female	8	Screen-Relative	Left
11.6886	1.16422	4011	female	9	Screen-Relative	Left
1.63101	0.356684	4011	female	10	Screen-Relative	Straight
12.7097	0.312459	4011	female	11	Screen-Relative	Right
10.546	2.35306	4011	female	12	Screen-Relative	Left
3.32024	0.403263	4012	female	1	Screen-Relative	Straight
11.5887	0.566637	4012	female	2	Screen-Relative	Right
16.6114	1.87069	4012	female	3	Screen-Relative	Left
12.8557	2.53921	4012	female	4	Screen-Relative	Right
10.7046	2.21799	4012	female	5	Screen-Relative	Right
9.41063	2.88763	4012	female	6	Screen-Relative	Roundabout-Right
13.8074	1.60855	4012	female	7	Screen-Relative	Straight
16.4819	0.735692	4012	female	8	Screen-Relative	Left
13.5217	2.4743	4012	female	9	Screen-Relative	Left
0.570111	1.0385	4012	female	10	Screen-Relative	Straight
14.732	2.36281	4012	female	11	Screen-Relative	Right
9.80688	3.00163	4012	female	12	Screen-Relative	Left
1.12439	0.365403	4013	male	1	World-Relative	Straight
8.94878	0.337182	4013	male	2	World-Relative	Right
8.51324	1.87701	4013	male	3	World-Relative	Left
5.22017	2.59864	4013	male	4	World-Relative	Right
4.48263	1.08133	4013	male	5	World-Relative	Right
10.099	2.61937	4013	male	6	World-Relative	Roundabout-Right
1.35778	2.04658	4013	male	7	World-Relative	Straight
14.5519	0.408829	4013	male	8	World-Relative	Left
12.6267	0.506311	4013	male	9	World-Relative	Left
16.3327	0.458447	4013	male	10	World-Relative	Straight
14.9146	0.81418	4013	male	11	World-Relative	Right
9.10161	1.88866	4013	male	12	World-Relative	Left
0.818573	0.322552	4014	male	1	World-Relative	Straight
14.0039	0.317817	4014	male	2	World-Relative	Right
13.9254	1.43996	4014	male	3	World-Relative	Left
9.4089	0.649341	4014	male	4	World-Relative	Right
13.5716	0.331545	4014	male	5	World-Relative	Right
7.10497	2.52511	4014	male	6	World-Relative	Roundabout-Right
14.3774	0.582295	4014	male	7	World-Relative	Straight
16.3935	1.53896	4014	male	8	World-Relative	Left
15.5685	1.34034	4014	male	9	World-Relative	Left
0.0377661	0.182635	4014	male	10	World-Relative	Straight

14.6353	0.601921	4014	male	11	World-Relative	Right
6.56452	2.03096	4014	male	12	World-Relative	Left
1.03499	0.359609	4015	male	1	World-Relative	Straight
15.3042	0.279089	4015	male	2	World-Relative	Right
12.31	1.45432	4015	male	3	World-Relative	Left
6.80308	1.95913	4015	male	4	World-Relative	Right
14.3856	0.444118	4015	male	5	World-Relative	Right
6.53831	1.7677	4015	male	6	World-Relative	Roundabout-Right
13.4825	0.610565	4015	male	7	World-Relative	Straight
14.6805	1.39094	4015	male	8	World-Relative	Left
11.4989	1.08135	4015	male	9	World-Relative	Left
14.79	0.110468	4015	male	10	World-Relative	Straight
14.6509	0.430564	4015	male	11	World-Relative	Right
1.25203	2.26967	4015	male	12	World-Relative	Left
6.24071	0.932439	4016	male	1	World-Relative	Straight
10.7457	1.02645	4016	male	2	World-Relative	Right
11.8948	1.65444	4016	male	3	World-Relative	Left
11.8072	0.454018	4016	male	4	World-Relative	Right
14.1931	0.122416	4016	male	5	World-Relative	Right
12.0216	1.91148	4016	male	6	World-Relative	Roundabout-Right
14.3538	0.450903	4016	male	7	World-Relative	Straight
15.9248	1.67135	4016	male	8	World-Relative	Left
13.7298	1.33237	4016	male	9	World-Relative	Left
0.164214	0.18849	4016	male	10	World-Relative	Straight
15.0889	0.377945	4016	male	11	World-Relative	Right
15.9272	2.06827	4016	male	12	World-Relative	Left
2.08004	2.84899	4017	male	1	World-Relative	Straight
13.9354	0.353833	4017	male	2	World-Relative	Right
15.1509	0.580493	4017	male	3	World-Relative	Left
11.13	0.490734	4017	male	4	World-Relative	Right
15.4148	0.644453	4017	male	5	World-Relative	Right
11.5889	1.79867	4017	male	6	World-Relative	Roundabout-Right
11.5423	0.942814	4017	male	7	World-Relative	Straight
15.9164	1.99372	4017	male	8	World-Relative	Left
9.61658	1.59221	4017	male	9	World-Relative	Left
16.7096	0.4879	4017	male	10	World-Relative	Straight
14.1527	1.11389	4017	male	11	World-Relative	Right
11.3611	2.6033	4017	male	12	World-Relative	Left
1.05801	0.475213	4018	male	1	World-Relative	Straight
8.28263	0.533968	4018	male	2	World-Relative	Right
6.14188	2.17299	4018	male	3	World-Relative	Left
11.1968	0.948555	4018	male	4	World-Relative	Right
10.2401	0.956093	4018	male	5	World-Relative	Right

9.32956	2.72148	4018	male	6	World-Relative	Roundabout-Right
10.7573	2.64229	4018	male	7	World-Relative	Straight
13.8329	2.19379	4018	male	8	World-Relative	Left
13.9177	1.85051	4018	male	9	World-Relative	Left
0.729106	1.03727	4018	male	10	World-Relative	Straight
6.74108	1.39466	4018	male	11	World-Relative	Right
13.0507	3.52973	4018	male	12	World-Relative	Left
9.03765	0.161644	4019	female	1	World-Relative	Straight
9.25887	0.497808	4019	female	2	World-Relative	Right
8.5223	1.93839	4019	female	3	World-Relative	Left
10.8278	2.91395	4019	female	4	World-Relative	Right
13.0268	0.575716	4019	female	5	World-Relative	Right
7.75247	2.41104	4019	female	6	World-Relative	Roundabout-Right
13.2727	0.538678	4019	female	7	World-Relative	Straight
14.2867	1.56195	4019	female	8	World-Relative	Left
11.3464	0.978373	4019	female	9	World-Relative	Left
15.6355	0.388849	4019	female	10	World-Relative	Straight
13.0813	0.507577	4019	female	11	World-Relative	Right
15.1853	1.59255	4019	female	12	World-Relative	Left
0.437354	0.583732	4020	female	1	World-Relative	Straight
16.675	0.541018	4020	female	2	World-Relative	Right
12.8721	1.70368	4020	female	3	World-Relative	Left
7.98931	2.402	4020	female	4	World-Relative	Right
11.0096	0.393127	4020	female	5	World-Relative	Right
8.95366	2.58051	4020	female	6	World-Relative	Roundabout-Right
13.9823	0.942069	4020	female	7	World-Relative	Straight
14.7496	1.56756	4020	female	8	World-Relative	Left
15.0256	1.54754	4020	female	9	World-Relative	Left
13.8203	0.0402823	4020	female	10	World-Relative	Straight
11.0248	0.959823	4020	female	11	World-Relative	Right
16.5331	1.75521	4020	female	12	World-Relative	Left
1.04821	0.658986	4021	female	1	World-Relative	Straight
12.4236	0.303266	4021	female	2	World-Relative	Right
12.1141	1.71808	4021	female	3	World-Relative	Left
11.6952	2.32661	4021	female	4	World-Relative	Right
13.184	0.629304	4021	female	5	World-Relative	Right
7.27684	2.28349	4021	female	6	World-Relative	Roundabout-Right
12.4912	1.51866	4021	female	7	World-Relative	Straight
17.2071	1.90974	4021	female	8	World-Relative	Left
12.8693	1.46335	4021	female	9	World-Relative	Left
1.52641	0.664594	4021	female	10	World-Relative	Straight
13.1173	0.288969	4021	female	11	World-Relative	Right
11.4996	2.3397	4021	female	12	World-Relative	Left

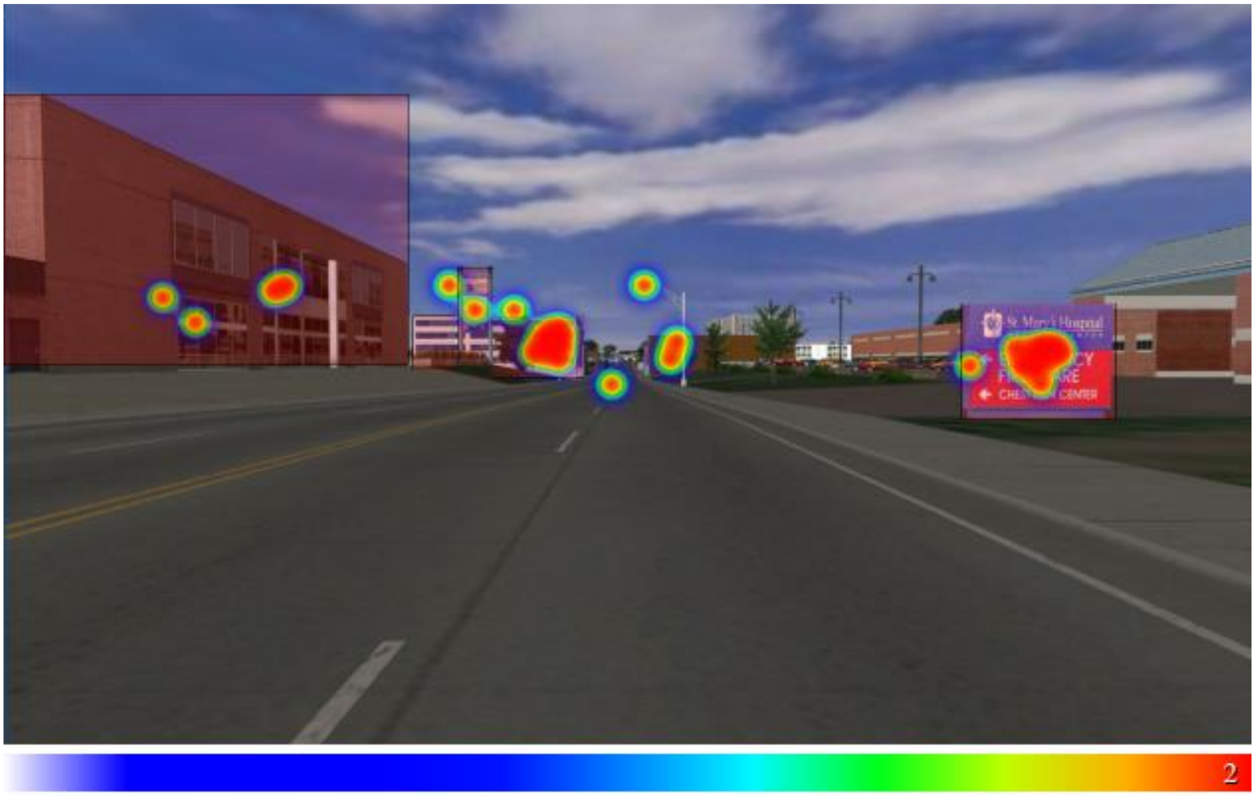
1.18788	0.521539	4022	female	1	World-Relative	Straight
13.6084	0.819037	4022	female	2	World-Relative	Right
13.3619	2.00993	4022	female	3	World-Relative	Left
4.48448	2.8467	4022	female	4	World-Relative	Right
7.84187	2.42004	4022	female	5	World-Relative	Right
5.40624	2.46084	4022	female	6	World-Relative	Roundabout-Right
9.70202	0.387487	4022	female	7	World-Relative	Straight
10.163	2.0416	4022	female	8	World-Relative	Left
8.54641	1.43274	4022	female	9	World-Relative	Left
9.41058	0.424849	4022	female	10	World-Relative	Straight
5.77484	1.16371	4022	female	11	World-Relative	Right
9.51143	2.26276	4022	female	12	World-Relative	Left
1.18171	0.0580206	4023	female	1	World-Relative	Straight
14.2409	0.216173	4023	female	2	World-Relative	Right
14.7776	1.4333	4023	female	3	World-Relative	Left
9.16065	2.19708	4023	female	4	World-Relative	Right
15.4785	0.102413	4023	female	5	World-Relative	Right
9.63445	2.53578	4023	female	6	World-Relative	Roundabout-Right
12.1684	0.161932	4023	female	7	World-Relative	Straight
14.9882	1.48716	4023	female	8	World-Relative	Left
12.4207	0.915509	4023	female	9	World-Relative	Left
15.3552	0.499436	4023	female	10	World-Relative	Straight
13.9948	0.243658	4023	female	11	World-Relative	Right
6.784	1.99887	4023	female	12	World-Relative	Left
0.732502	0.146943	4024	female	1	World-Relative	Straight
10.701	0.177206	4024	female	2	World-Relative	Right
12.7476	1.76112	4024	female	3	World-Relative	Left
12.3394	0.6868	4024	female	4	World-Relative	Right
10.8399	0.266278	4024	female	5	World-Relative	Right
8.27614	2.18798	4024	female	6	World-Relative	Roundabout-Right
0.979229	1.2999	4024	female	7	World-Relative	Straight
14.3064	0.727131	4024	female	8	World-Relative	Left
17.6908	1.51345	4024	female	9	World-Relative	Left
0.121331	0.207522	4024	female	10	World-Relative	Straight
13.9467	1.4485	4024	female	11	World-Relative	Right
10.6775	1.95131	4024	female	12	World-Relative	Left

C7. SDT PARAMETERS

Condition	Gender	Correct Rejection	False Alarm	Hit	Miss	Sensitivity	Bias
screen-relative	Male	0.625	0.38	0.625	0.375	0.6373	1
screen-relative	Male	0.375	0.63	0.25	0.75	-0.9931	0.838
screen-relative	Male	0.375	0.63	0.375	0.625	-0.6373	1
screen-relative	Male	0.25	0.75	0.375	0.625	-0.9931	1.1933
screen-relative	Male	0.75	0.25	0.5	0.5	0.6745	1.2554
screen-relative	Female	0.375	0.63	0.5	0.5	-0.3186	1.0521
screen-relative	Female	0.375	0.63	0.5	0.5	-0.3186	1.0521
screen-relative	Female	0.625	0.38	0.5	0.5	0.3186	1.0521
screen-relative	Female	0.625	0.38	0.375	0.625	0	1
world-relative	Male	0.625	0.38	0.625	0.375	0.6373	1
world-relative	Male	0.75	0.25	0.5	0.5	0.6745	1.2554
world-relative	Male	0.625	0.38	0.375	0.625	0	1
world-relative	Male	0.5	0.5	0.375	0.625	-0.3186	0.9505
world-relative	Male	0.625	0.38	0.625	0.375	0.6373	1
world-relative	Female	0.5	0.5	0.5	0.5	0	1
world-relative	Female	0.5	0.5	0.625	0.375	0.3186	0.9505
world-relative	Female	0.375	0.63	0.5	0.5	-0.3186	1.0521
world-relative	Female	0.375	0.63	0.375	0.625	-0.6373	1
world-relative	Female	0.25	0.75	0.625	0.375	-0.3559	1.1933
screen-relative	Male	0.375	0.63	0.5	0.5	-0.3186	1.0521
screen-relative	Female	0.625	0.38	0.25	0.75	-0.3559	0.838
screen-relative	Female	0.625	0.38	0.5	0.5	0.3186	1.0521
world-relative	Male	0.625	0.38	0.625	0.375	0.6373	1
world-relative	Female	0.5	0.5	0.25	0.75	-0.6745	0.7965

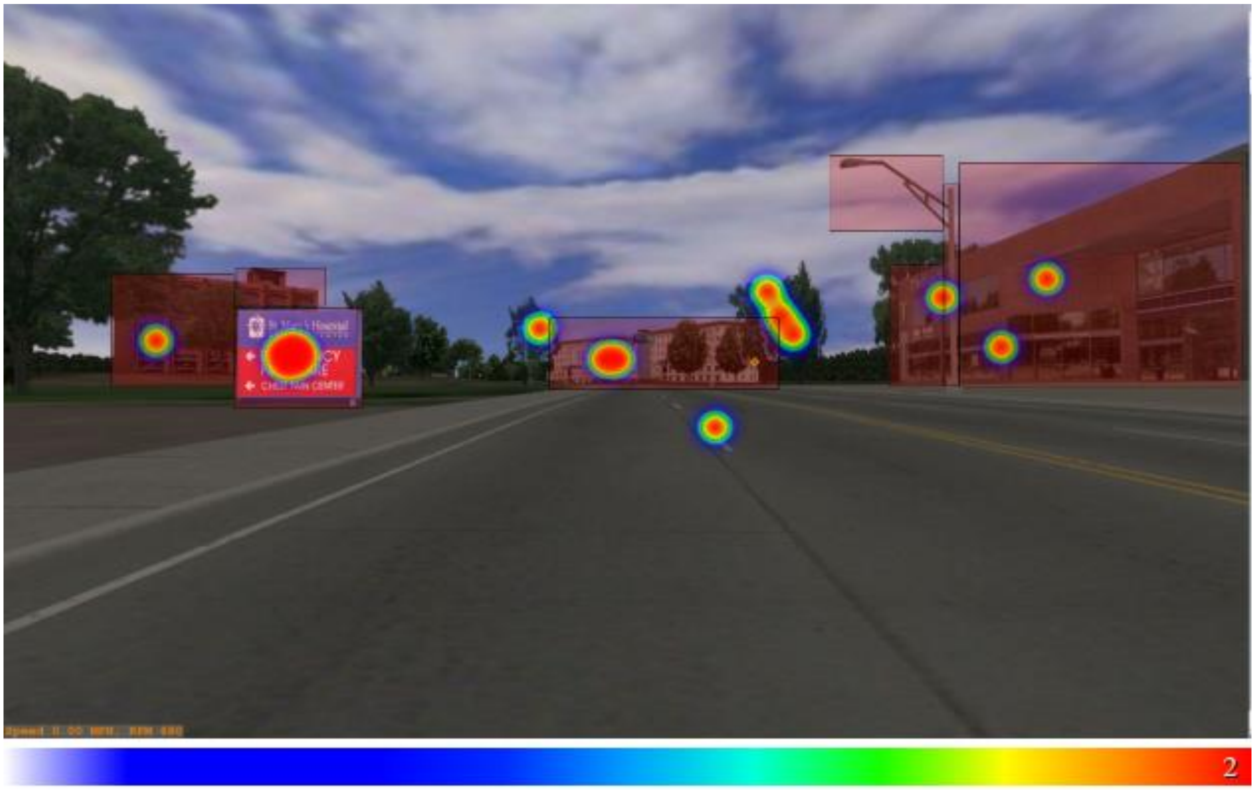
C8. LANDMARK KNOWLEDGE HEAT MAPS

ON ROUTE – ST



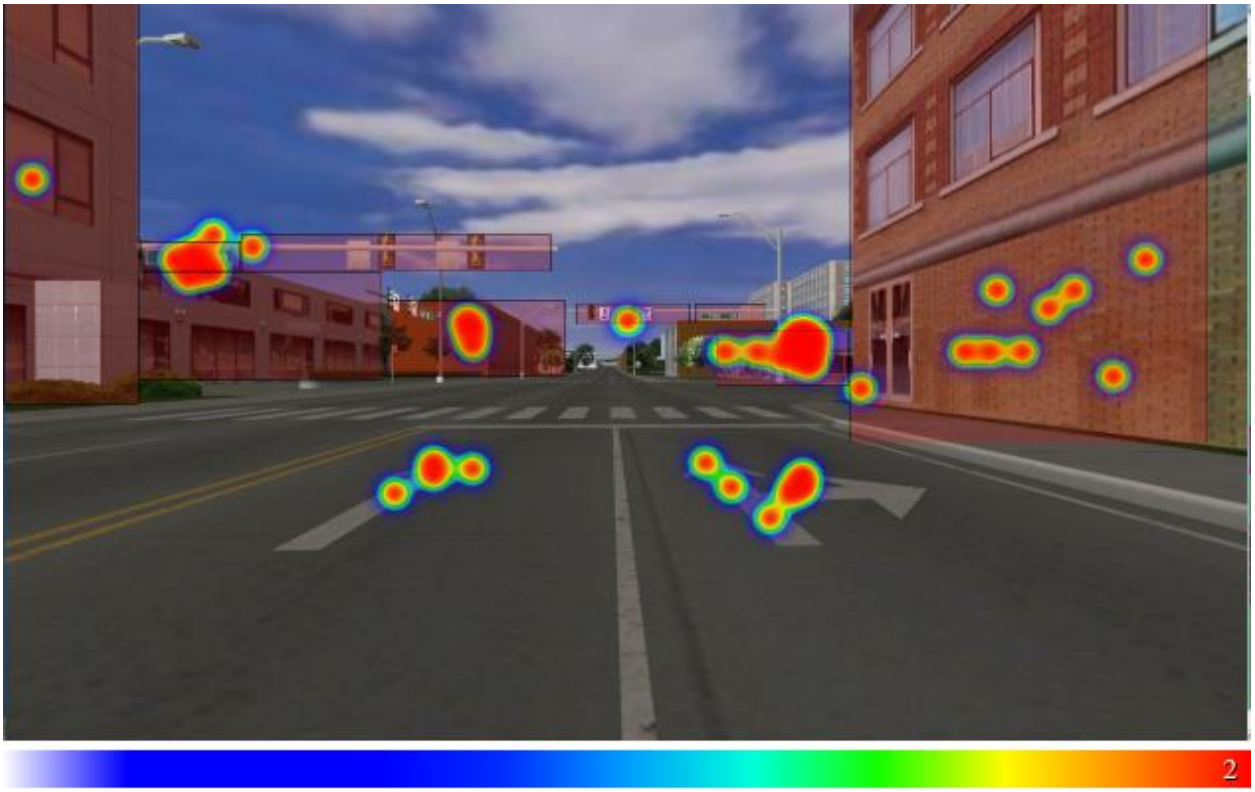
Area of Interest	%	Count
St. Mary's Hospital sign	36.36%	8
Building in the left corner	18.18%	4
Advertisement sign	4.55%	1
White building	27.27%	6
Other	13.64%	3
Total	100%	22

OFF ROUTE – O_ST



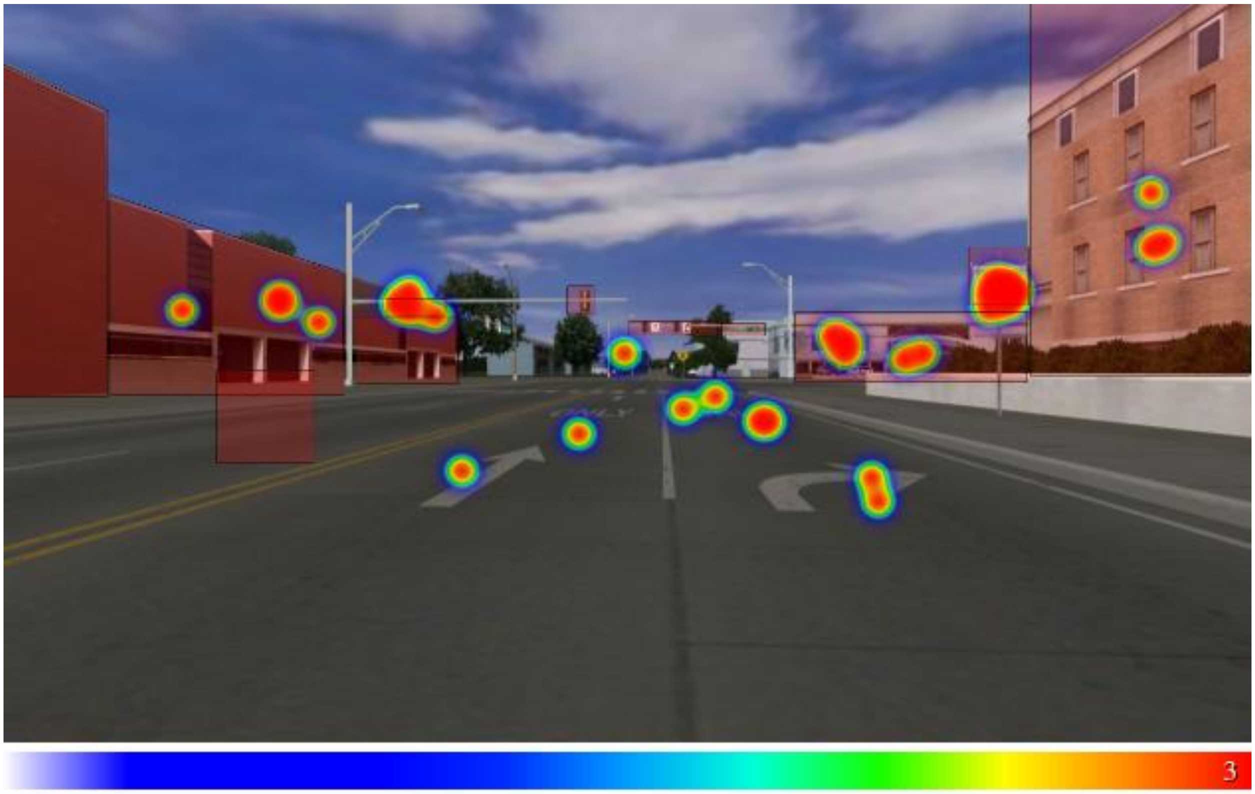
Area of Interest	%	Count
Brown Building on the right corner	20.00%	3
St. Mary's Hospital sign	33.33%	5
Brown Building on the left corner	6.67%	1
White Building on the center	13.33%	2
Street Light Post	0.00%	0
Other	26.67%	4
Total	100%	15

ON ROUTE - S1



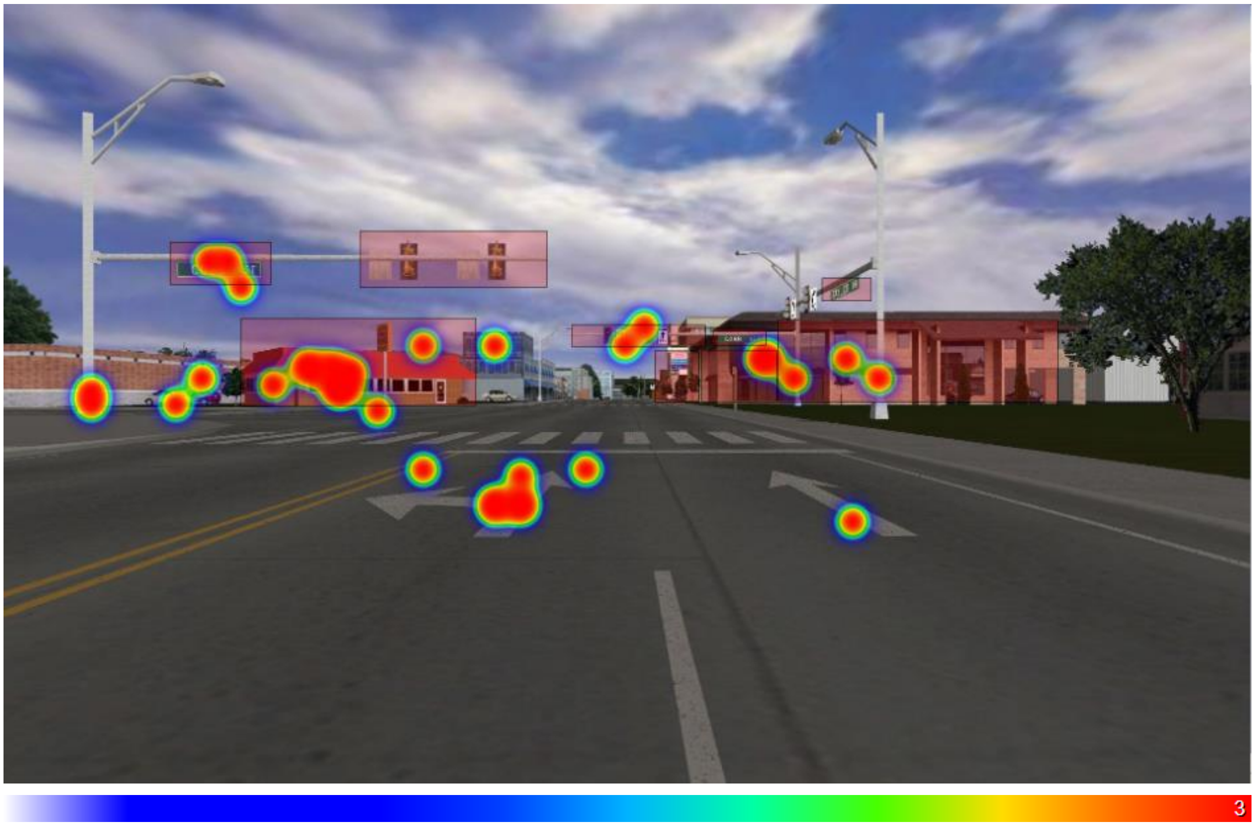
Area of Interest	%	Count
Brown Building on the Right Corner	22.22%	8
Blue Building on the left corner	8.33%	3
Traffic Light	5.56%	2
Street Name	8.33%	3
Brown Building on the left far corner	8.33%	3
Brown/ Blue Building on the right far corner	27.78%	10
Other	19.44%	7
Total	100%	36

OFF ROUTE – O_S1



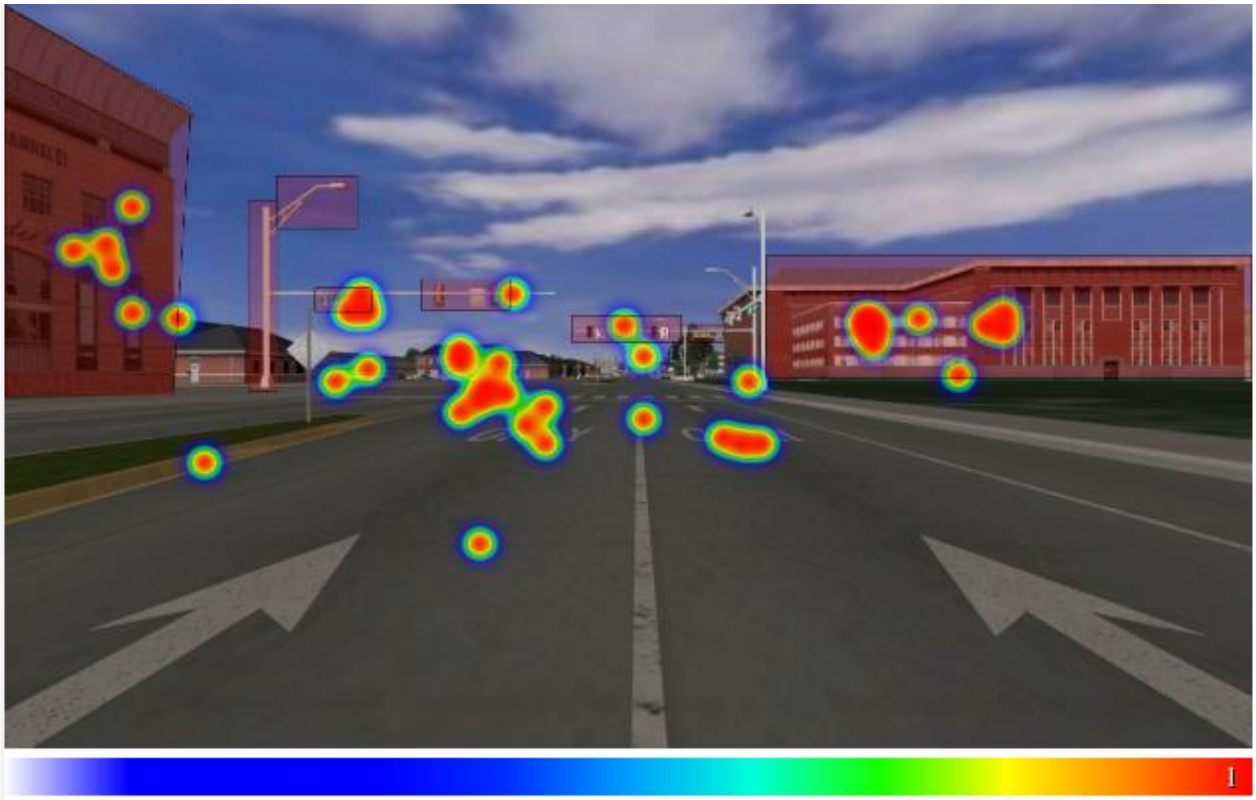
Area of Interest	%	Count
Traffic sign on the right corner	22.22%	6
Grey Building on the right far corner	18.52%	5
Brown Building on the left corner	18.52%	5
Street Name on the left corner	7.41%	2
Traffic Light	0.00%	0
Region #11	0.00%	0
Light Brown Building on the right corner	11.11%	3
Other	22.22%	6
Total	100%	27

ON ROUTE – M2_L



Area of Interest	%	Count
Advertisement sign	0.00%	0
Brown Building on the Right Far Corner	18.52%	5
Burger King	33.33%	9
Other	33.33%	9
Street Name	7.41%	2
Traffic Light	7.41%	2
		27

OFF ROUTE – O_M2_L

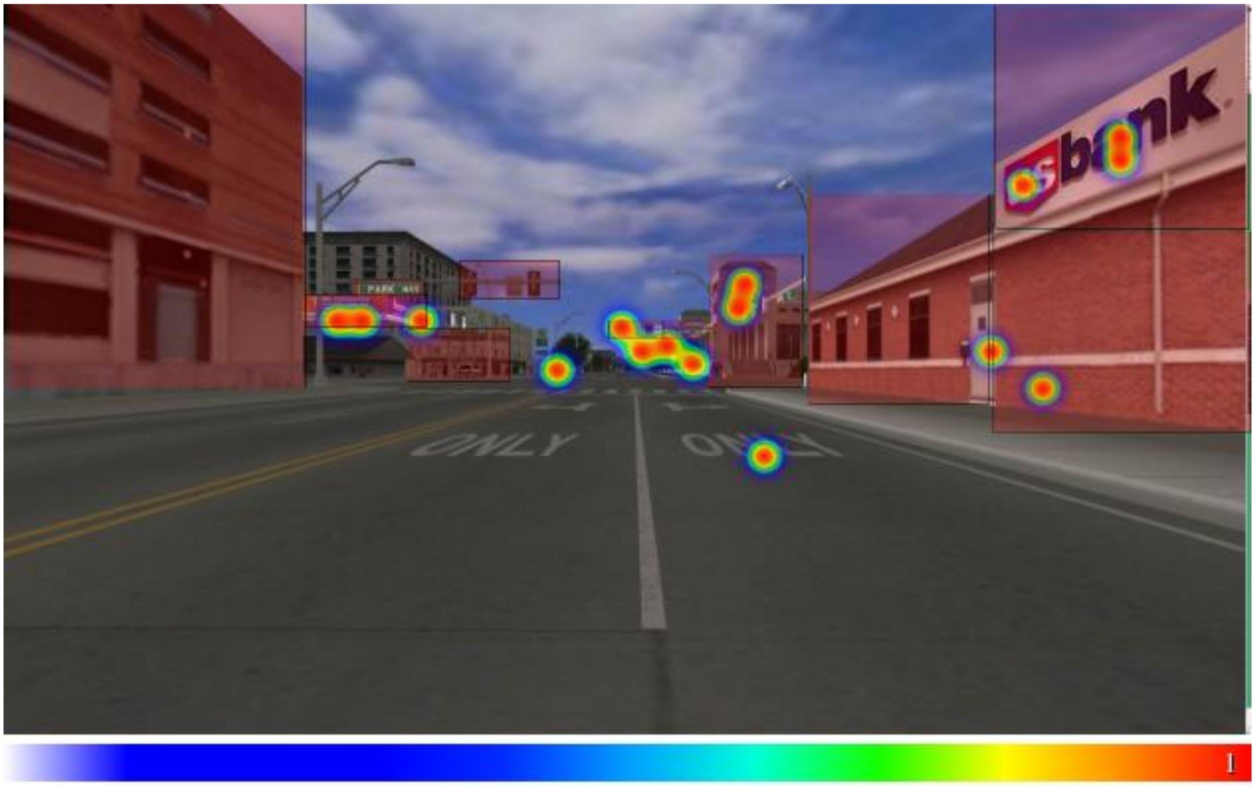


Area of Interest	%	Count
Brown Building on the right far corner	30.00%	9
Traffic Lights	3.33%	1
Salon Spa Building on the left corner	16.67%	5
Light Post	0.00%	0
Street Name right corner	6.67%	2
Street Name left corner	0.00%	0
Other	43.33%	13
Total	100%	30

ON ROUTE – M4_R

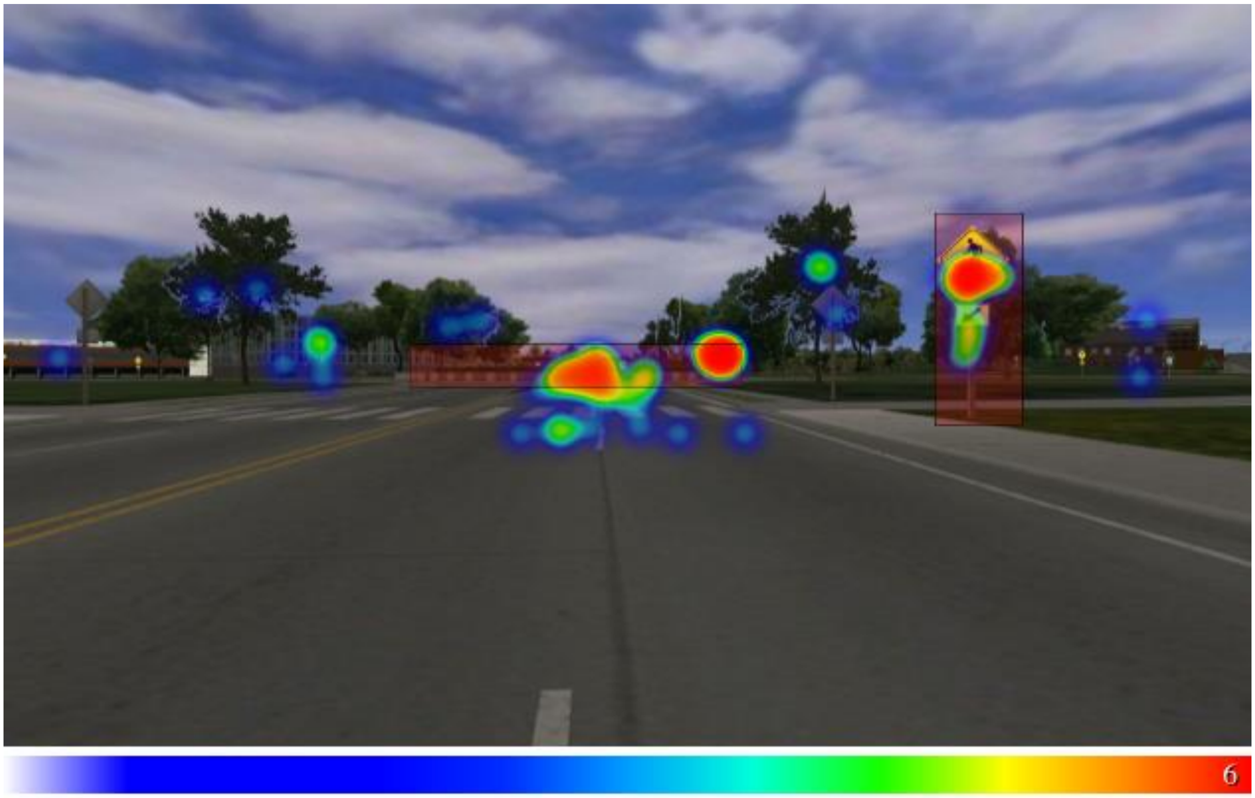
Area of Interest	%	Count
Street Name	7.14%	1
Shell Advertisement sign	21.43%	3
Village Inn	21.43%	3
Traffic Light	0.00%	0
Shell Gas Station	0.00%	0
McDonalds	21.43%	3
Staples	14.29%	2
Other	14.29%	2
Total	100%	14

OFF ROUTE – O_M4_R



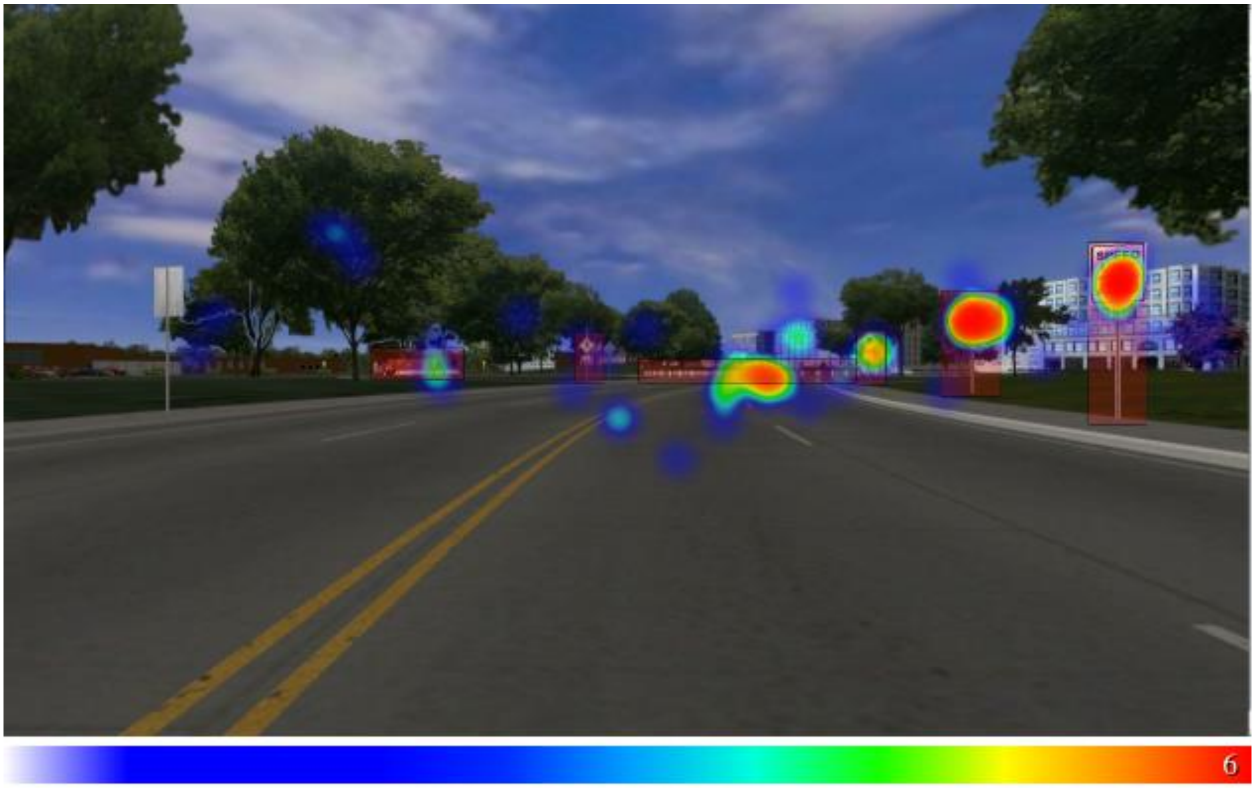
Area of Interest	%	Count
Traffic Lights	7.69%	1
Brown Building on the left corner	0.00%	0
Gas Station	0.00%	0
Street Name	0.00%	0
Advertisement Sign on the left Corner	15.38%	2
US Bank Building	15.38%	2
US Bank Sign	23.08%	3
Grey Building on the far corner	15.38%	2
Other	23.08%	3
Total	100%	13

ON ROUTE – R1_R



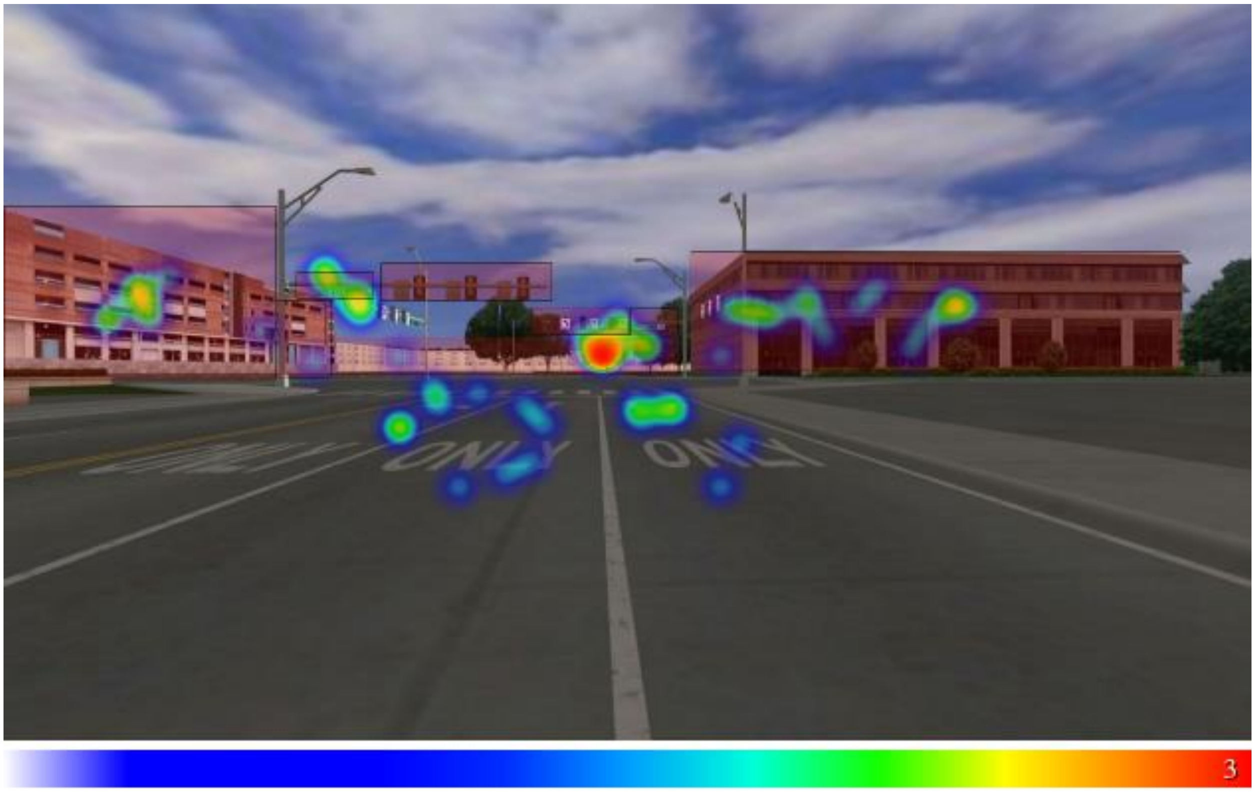
Area of Interest	%	Count
Crossing sign	30.00%	12
Roundabout	37.50%	15
Other	32.50%	13
Total	100%	40

OFF ROUTE – O_R1_R



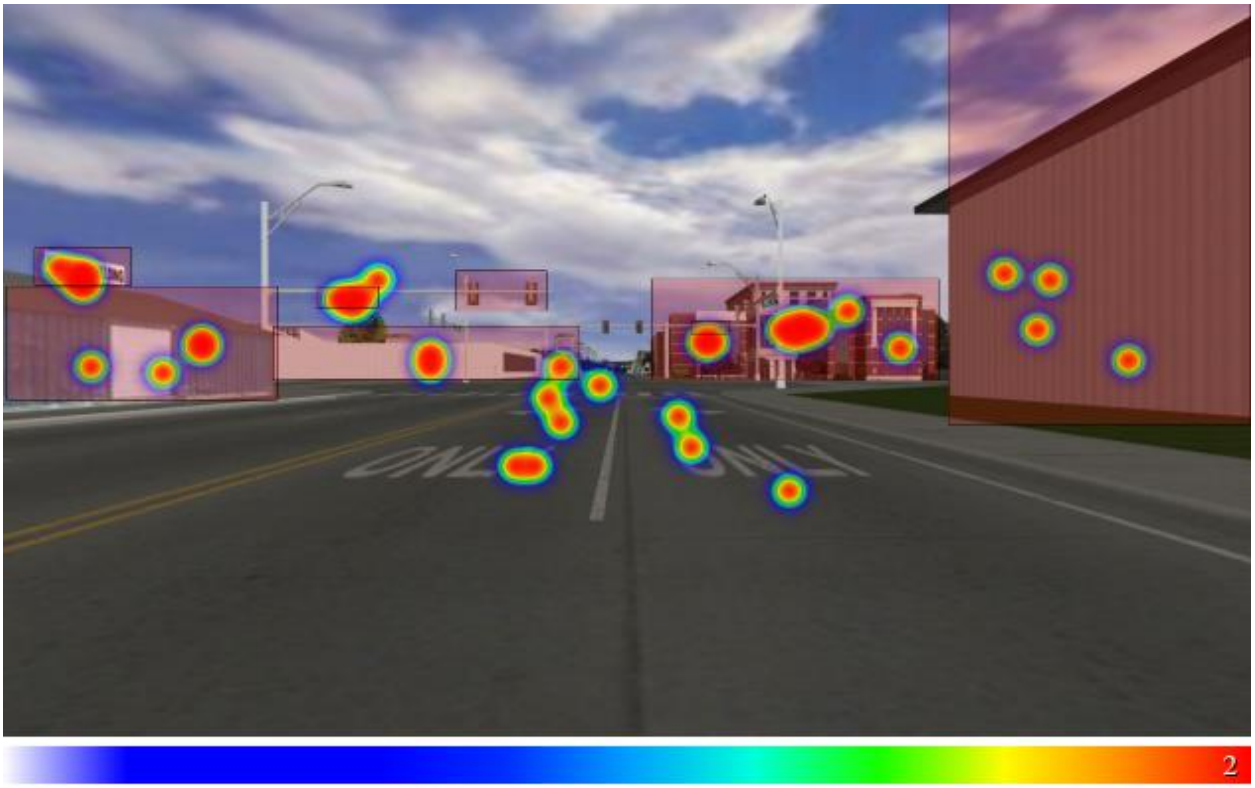
Area of Interest	%	Count
Roundabout Sign	27.14%	19
Speed Limit Sign	24.29%	17
Pedestrian Crossing Sign	8.57%	6
Roundabout	15.71%	11
Truck	4.29%	3
Other	20.00%	14
Total	100%	70

ON ROUTE – S2



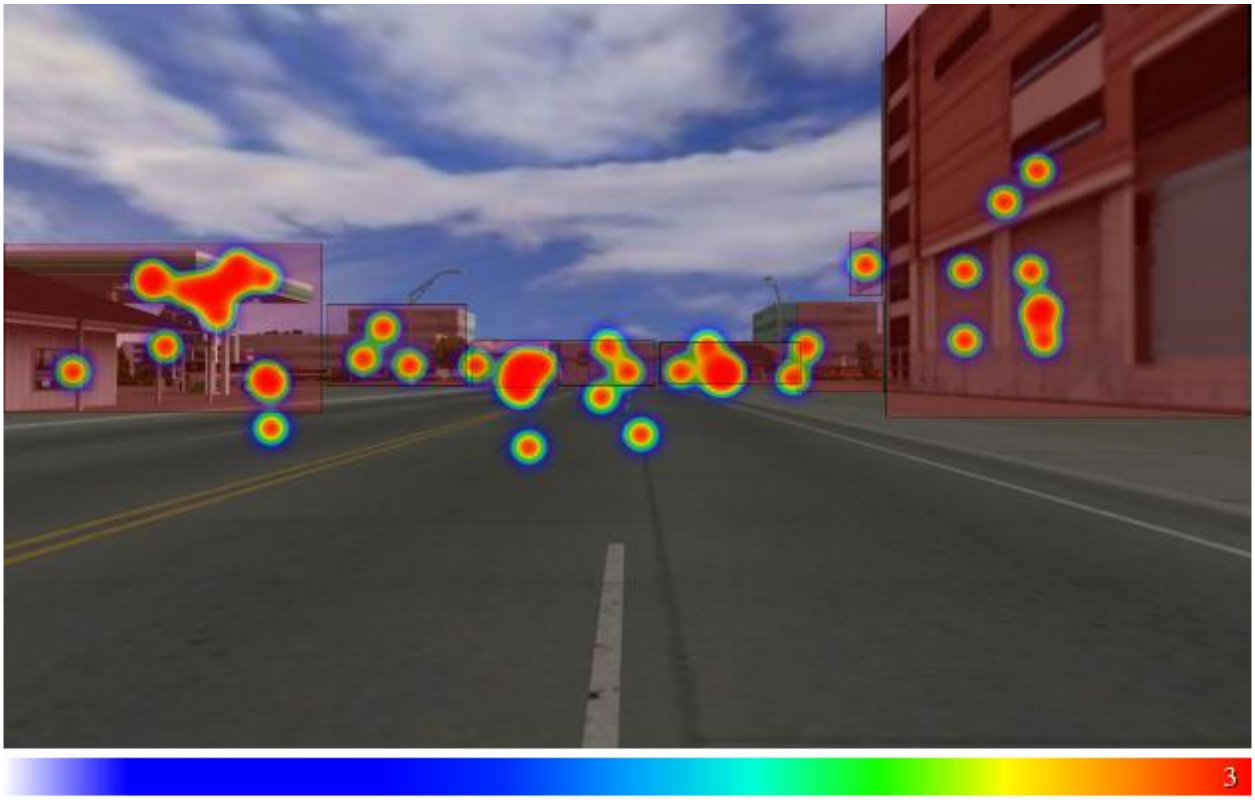
Answer	%	Count
White building on the back	13.04%	6
Brown Building on the left corner	19.57%	9
Grey Building on the right corner	32.61%	15
Street Name	6.52%	3
Traffic Light	4.35%	2
Other	23.91%	11
Total	100%	46

OFF ROUTE – O_S2



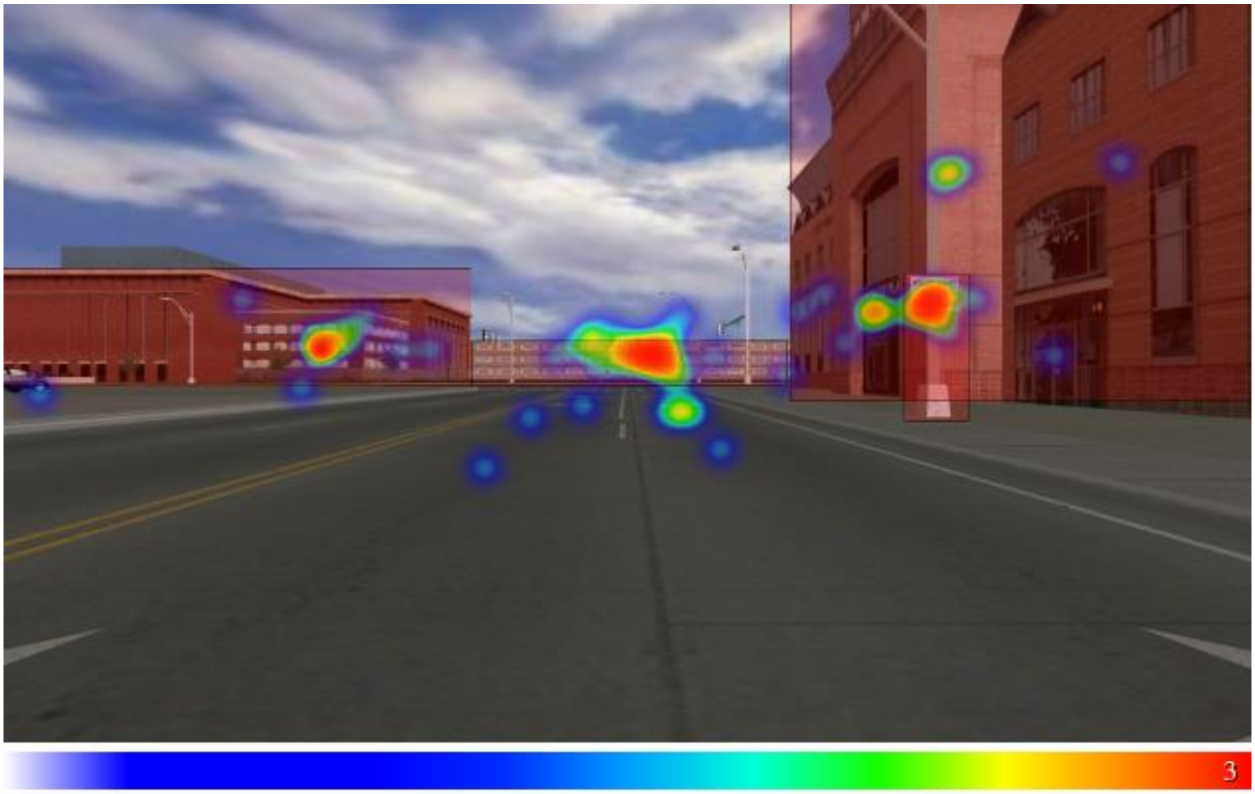
Area of Interest	%	Count
Brown Building on the right far corner	25.81%	8
Street Name	9.68%	3
Wheelock Welding Building	16.13%	5
Wheelock Welding Sign	9.68%	3
Traffic Light	0.00%	0
Fence on the left far corner	9.68%	3
Grey Building on the right corner	12.90%	4
Other	16.13%	5
Total	100%	31

ON ROUTE – M5_L



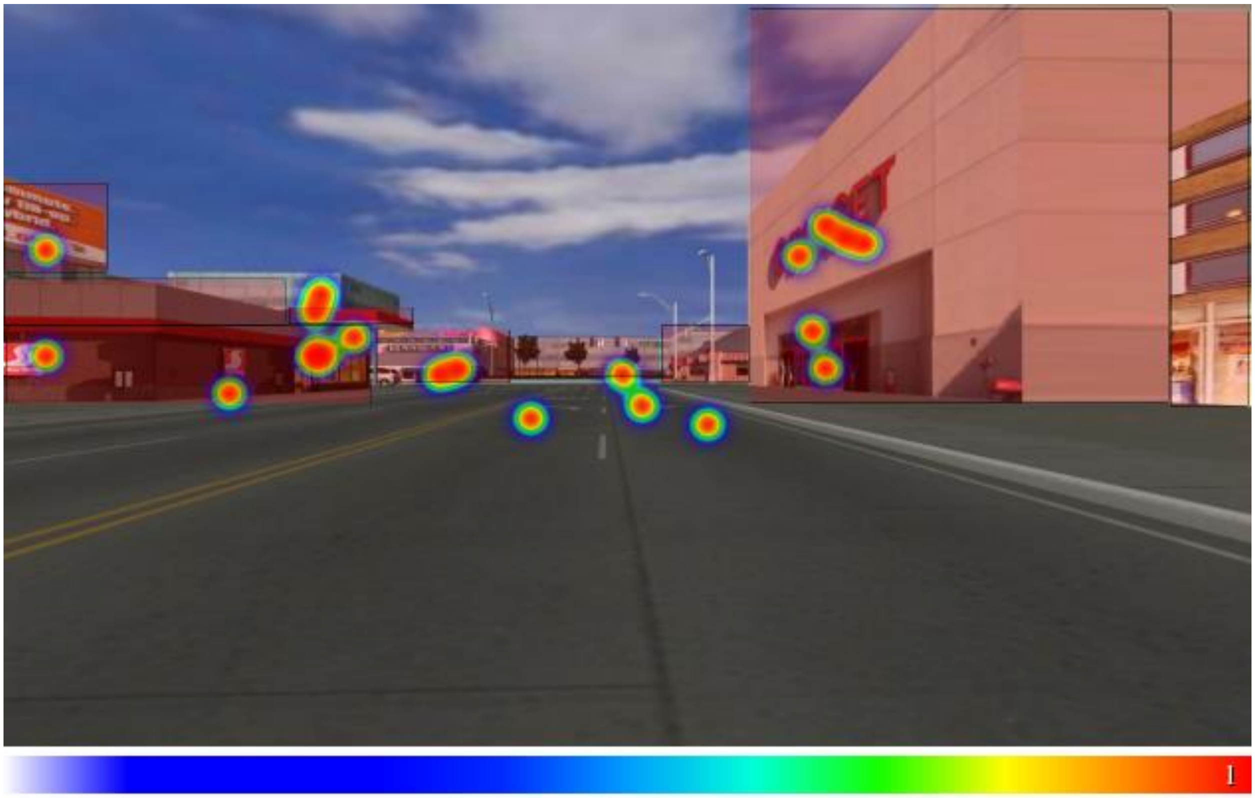
Area of Interest	%	Count
Gas Station	26.32%	10
Brown Building on the right corner	18.42%	7
Bank	13.16%	5
Grey Building on the center	5.26%	2
Grey Building on the left far corner	7.89%	3
Fast food building on the far corner	13.16%	5
Grey Building on the right far corner	2.63%	1
Parking sign	2.63%	1
Other	10.53%	4
Total	100%	38

OFF ROUTE – O_M5_L



Area of Interest	%	Count
Brown Building on the right corner	26.42%	14
Grey Building in the center	26.42%	14
Brown Building on the left corner	18.87%	10
Must turn traffic sign	15.09%	8
Other	13.21%	7
Total	100%	53

ON ROUTE – M7_R



Area of Interest	%	Count
Advertisement sign on the left	6.25%	1
Short Stop Store	18.75%	3
Gas Station	6.25%	1
Store on the Right far corner	0.00%	0
Store before target	0.00%	0
Grey Building on the center	6.25%	1
Bennigan's store	12.50%	2
Target	31.25%	5
Other	18.75%	3
Total	100%	16

OFF ROUTE – O_M7_R

Area of Interest	%	Count
Bank Building	6.25%	2
Store on the left corner	3.13%	1
Gas Station	15.63%	5
Grey building on the left far corner	18.75%	6
Street Name	6.25%	2
Building on the right corner	15.63%	5
Traffic Light	9.38%	3
Other	25.00%	8
Total	100%	32