DRIVING, EYE-TRACKING AND VISUAL ENTROPY: EXPLORATION OF AGE AND TASK EFFECTS

By

Jess Gilland

B.S., University of South Dakota, 2002
M.A., University of South Dakota, 2004

A Dissertation Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

Department of Psychology
The University of South Dakota
May, 2008

The members of the Committee appointed to examine

the Dissertation of Jess Gilland find it

satisfactory and recommend that it be accepted

____

Chairperson - Frank Schieber, Ph. D.

____

Jan Berkhout, Ph. D.

____

Thomas Langworthy, Ph. D.

____

Doug Peterson, Ph. D.

____

XT (Xiao-Tian) Wang 王晓田, Ph.D.

Abstract

Most driving-related eye-movement research completed before this present study has relied upon *qualitative* techniques to evaluate questions. These techniques primarily consist of descriptive statistics on distributional data. This is adequate for a single study, but can not be reliably compared to the results of another study having dissimilar conditions. This has led to a need for the development and testing of a new measurement technique that produces *quantitative* results that are scaleable across differing situations. The quantitative metric of *entropy* – derived from information theory – has been adapted toward this purpose. Visual entropy values were calculated from the eye movements of 28 young (*M* = 24, *SD* = 5.3 yrs) and 14 older (*M* = 75, *SD* = 5.6 yrs) participants who drove an instrumented vehicle along a predefined route while engaged in several functionally different cognitive tasks. Driver age and subsidiary task type (none, verbal, visual-spatial) were systematically varied in order to assess the sensitivity of the entropy metric to discriminate behaviors generated across these situational factors. Resulting entropy values were directly compared to several more common visual metrics such as pupil size, saccadic amplitude, and fixation dwell time. In all metrics, effects were most discernible in the visual-spatial condition which was further exacerbated in the older driver group. The finding that a *visual attention* task inhibited a *visual scanning* behavior the most supports Wickens' multiple resource theory. The finding that older individuals were the most affected in all cases supports the idea that older individuals may have diminished visual-cognitive resources available during driving compared to their younger counterparts. Upon comparison of the different visual metrics it was found that the entropy metric proved to be more sensitive to attentional demands than all alternative visual metrics assessed (e.g. pupil size, dwell times, and saccadic amplitudes). The findings of this study strongly suggest that global measures of eye movement behavior captured by the entropy metric are useful for understanding the correlation between normal adult aging and task–induced cognitive demands within the context of real-world driving.

Acknowledgments

Many people donated *vast* amounts of time and support during this project. I would like to give thanks to these individuals:

- I would like to thank Dr. Frank Schieber. Frank's advice was invaluable during idea development, protocol implementation, and writing evaluation. Without his expertise, this project would not have progressed as efficiently as it did.

- I would like to thank the members of my committee for taking time out of their busy schedules to oversee the project throughout the years.

  Dr. XT Wang
  Dr. Doug Peterson
  Dr. Jan Berkhout
  Dr. Thomas Langworthy

- I would like to thank my wife Rebecca and two boys (Zander & Gavin) for their constant support throughout this project. I know it took longer than anticipated, but everything always works out in the end.

Table of Contents

List of Figures

List of Tables

This study was designed to bridge the gap between the qualitative methods currently used for evaluating vision in driving research and quantitative that are actually desired. It will be demonstrated how Information theory, more specifically *entropy*, can be used to derive these meaningful values on a quantitative scale. The power of the entropy concept is that one can calculate an *exact* minimum and maximum quantity of "information" present. By specifying a range, one can standardize the data before correlating with performance. Standardized data is much more reliable for comparison across multiple studies. The ultimate goal of this research was to develop and validate the entropy methodology using eye movements.

1.0 Review of Eye tracking literature

Eye tracking is a topic that only recently became "everyday" science due to the shear expense of both time and money needed to collect this type of data. Eye movements vary by individual and are highly directed by situation which makes it hard to correlate one study's result with the next.

The next section will provide a framework from which to view eye-tracking theory. This will also serve to facilitate the understanding of what possible situations can arise during its use. What follows is a comprehensive summary of the types of apparatus used, major terms that should be known, and key studies reinforcing what to expect when involved in this type of research.

1.1 Eye-Tracking Instrumentation

In order to do any high fidelity research on visual scan patterns, an "eye-tracker" is needed. An eye-tracker is simply an apparatus used to extrapolate and graphically represent where a person is looking. There are several distinct types of eye-trackers each varying in its theory of operation. The most common systems are dark pupil, bright pupil, and dual purkinje. Each has its own subtleties associated with each.

Bright and dark pupil systems will be explained first. A bright pupil system was used in this study, although both are operationally the same. These types of eye-trackers are regarded as the most basic and use pupil position along with light reflections off the surface of the cornea to triangulate where the eye is “looking” at any given moment. One can think of this as similar to satellites in a GPS system. Given the relative distances to several different satellites one can triangulate nearly any position in space. In an eye-tracking system these "satellites" are the equivalent of the pupil and at least one of several corneal reflections. The "location", is the point of gaze. The pupil is the large open center of the eye that allows light transmission to the retinal tissues. The corneal reflections are simply reflections of light originating from the outside world being reflected off the cornea.

In both these systems, the eye will be illuminated by the use of a light source, usually infrared (IR). Sometimes when the eye is flooded with IR light, it will enter the pupil and be reflected off the retina back through the pupil on nearly the same viewing plane as the recording camera. This will result in a "bright" pupil. This can be visualized by thinking of a camera flash causing the common red eye phenomenon in many pictures. This is the same principle, only using IR-light, which isn't visible without a specialized camera. The pupil is used as the anchor point and assumed that all sight is 90° to the center surface of the pupil projecting outward. By assuming straight vision from the pupil and triangulating that with the reflection of a given point source of light referenced by its reflections off the cornea, one can conclude the direction of gaze. The opposite of a bright pupil system is a "dark" pupil system. As one would expect, this measures the same elements as the bright pupil system only with a dark pupil. This is accomplished by locating the recording camera in such a way that it is not on the same viewing plane as the light reflected off the retina. It relies on a dark center to define the pupil as opposed to bright, but the two systems are operationally the same.

The main precision from either of these two systems comes from the addition of a second set of reflections called corneal reflections. These are simply light reflections off the cornea. These are usually artificially produced by external lighting elements to maintain a brightness that can be easily recognized by a computer program over all other lighting "noise". One can use a single reflection as is often the case with many head mounted units like the Applied Science Laboratories (ASL) model 501, or opt to use a more robust two-light reflection unit like that seen in many car setups like the ASL ETS-PC. A two-reflection setup will rely on two differently shaped light sources to assure discrimination from each other on the cornea. The single reflection model has strength in that it needs less computer processing power making it much more simple setup. It is weak in that it is not completely robust to movement. The double reflection model is good for situation where lots of movements is expected since the system can always default to a single reflection state if the subject turns his or her head and one reflection is lost.

Another lab-based system that has been used for eye tracking is called a Dual-Purkinje Tracker. This is a *very* high precision tracker that interprets movements of the eye with incredible precision (1000 Hz). This system uses a lawfully defined set of four

"Purkinje" reflections that appear where light entering the eye is refracted and reflected outward. The 1st reflection occurs on the immediate outside surface of the cornea. The 2nd appears on the inside surface of the cornea. The 3rd occurs on the front side of the lens and the 4th is on the inside of the lens. Each reflection will refract to a different location and appear as a different intensity when observed from the front of the eye. Although Purkinje reflections occur at 4 locations, reflections #1 and #4 are usually the only ones used with this system.

The main downfalls of the Dual-Purkinje tracking system are that *reflections*: (a) are very hard to see (limited working environment), (b) only visible over a small viewing range (<30° - limited usable visual field), and (c) *very* prone to head movement (reflections move in miniscule increments relative to the visual scene). Due to these three shortcomings, this type of tracker can only be used on subjects with immobilized heads in controlled settings, making a car study pretty much out of the question.

1.2 Important Terminology

Several important terms are associated with this discipline. The first is the concept of a *fixation*. A fixation is defined as the period of time after the eye acquires a new target and ceases movement. It is during this time that visual information is extracted and relayed to the brain. This period of time can range greatly by individual but usually averages around 200-300 milliseconds in duration (Green, 2002; Irwin, 1996; (300ms observed in Zwahlen & Schnell, 1998)). However, it can also range from less than 100ms to well over a second depending on the situation (Harris et al, 1988). Fixation duration can also be influenced by target size as Harris et al. (1988) reported that durations of a fixation tend to decrease as target size is increased. The target size effect was reasoned to be due to fixations being terminated by interference of peripheral "non-foveal" information when a target is large. Essentially there are too many distracters WITHIN the target itself.

It should be noted however that a fixation on an object does not mean it has been processed efficiently enough for recall as demonstrated in Louma (1988). Louma reported as high as a 54% miss rate for a pedestrian sign fixated 100% of the time. They clearly looked but did not "see" *more* than half the time. Nevertheless, eye fixations tend to follow a distinct trend: Given any search task, fixations will tend to cluster to areas and objects containing the highest quantity of visual information rather than other less informative locations (Buswell, 1935; Mackworth and Morandi 1967). However, this too can be influenced as demonstrated by Yarbus (1967) when he observed a shift in gazing behavior dependent on the task. When viewing directions were changed, subjects now began to look at the areas that each felt held the most information in that specific situation. To provide an example of this, say there was a picture of ten people on a football bleacher. One would expect most people to study the faces of the individuals the majority of the time as this would provide the greatest source of information given no other search strategies. However, if the subject was now instructed to identify the *context* of the picture, one would expect a shift of fixations to areas other than the faces as they no longer provide as much information as the peripheral details given the new situation.

Knowing that potential "strategy" confounds are possible, extra caution was taken in the present study to avoid coercive search strategies. By using a defined driving route and special circumstance directions, we believe that the majority of eye-movements observed in this study accurately reflect undirected and naturalistic scanning behavior unbiased by distracters or search goals.

For this field of research, it is necessary to explain a few additional terms that are used to describe eye-movements in the literature. The next important term is called a *saccade*. This is the movement that occurs as the eye moves to the next area of fixation. Very little if any visual information is actually processed *during* a saccade (Henderson & Hollingworth., 1999; Hollingworth & Henderson, 2003; Duren, 1993). Saccades can be very brief and have very high rotational velocities ranging from 275°/s (Green, 2002) to an extreme 900°/s (Carpenter, 1988.) As a side note, all saccades vary in size and will be labeled *differently* depending on these sizes. In some literature, a large saccade that moves to a new viewing area is termed a transition. This term is interchangeable with a *saccade* in principle, but is generally used to indicate *a new target acquisition* rather than additional detail gathering at the same target. One may also hear the term *glance* in reference to several saccades and fixations around the same target. Right now it is just important to note that there is a distinction between the terms, however, the subtle differences between the three will be clarified in more detail later in this section (see Figure 2).

A saccade links fixations together; however there is debate on what role it plays in picture synthesis. In Henderson & Hollingworth (2003b), they observed a *global* change-blindness effect suggesting that point by point representations of visual scenes may not be cumulative. In other words, one doesn't progressively "stitch" a picture together from smaller parts over time. Instead, each fixation acts like an autonomous snapshot of the environment representing one particular point in time. Each fixation in this respect would produce an independent picture of the scene that decays very rapidly from one fixation to the next. This is supported by the idea that changes occurring *during* a saccade are prone to a low detection rate (Henderson & Hollingworth, 2003b.)

This idea of scene linking between fixations was tested in another study by Hollingworth & Henderson (2003) were they concluded that scene linking *does* occur, but it is only observable over a relatively small area surrounding the current fixation point rather than over an entire scene. Changes occurring further from the current fixation point are progressively less detectable since a reduced amount of attention is devoted to updating the “details” in the periphery with each fixation. Having said this, it should also be noted that there are exceptions to this finding. In the Hollingworth & Henderson (2003) study they also observed that even though the "updated" picture sent to the brain seemed to miss many changes that occurred *during* the saccade, these changes could be made much more evident depending on the context of the change. Replacing an object with context consistent objects was much less noticeable than if that same object was replaced with a non-context consistent object for that scene. For example, if viewing a kitchen scene were a refrigerator was replaced with a stove, one may be prone to miss it because it is semantically logical. A stove would be expected in a kitchen. However, if that fridge was replaced with a potted tree, the change would be much more likely to be detected because it no longer fits the context of the scene. A pictorial example of this concept can be seen in Figure 1.

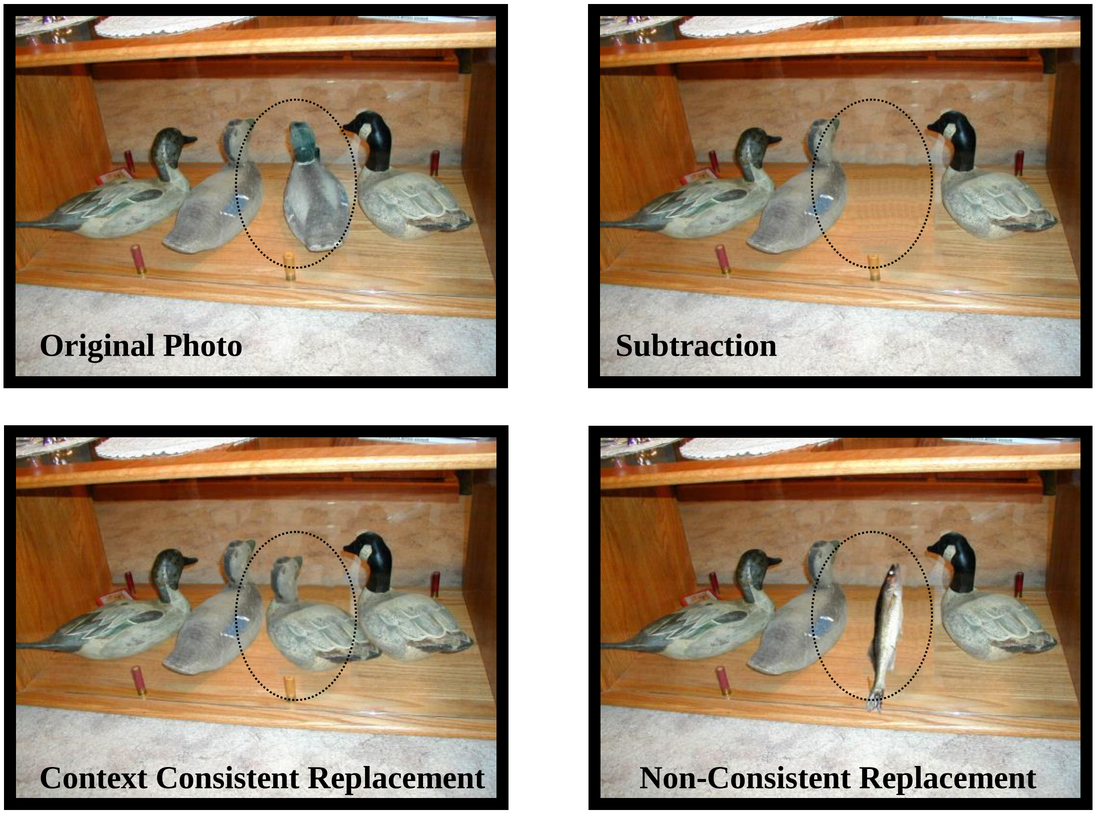


Figure 1. Examples of different picture manipulations. A random object (circled) changes *during* a saccade and then detection rates are assessed. This protocol was used in both Brockmole & Henderson (2004) and Hollingworth & Henderson (2003).

The next important term to understand is a *transition.* A transition is a *new target acquisition* characterized by a MUCH longer saccade to a *new location.* This is thought to be an attention driven mechanism that requires a slight but unconscious attentional shift from the current target to peripheral target acquisition (Henderson, 1993). In other words, a transition is used to locate new targets, whereas a saccade is driven by the need to extract information from the target once located. A pictorial of the relationship between a glance, fixation, transition, and saccade can be seen in Figure 2, adapted from Green (2002).

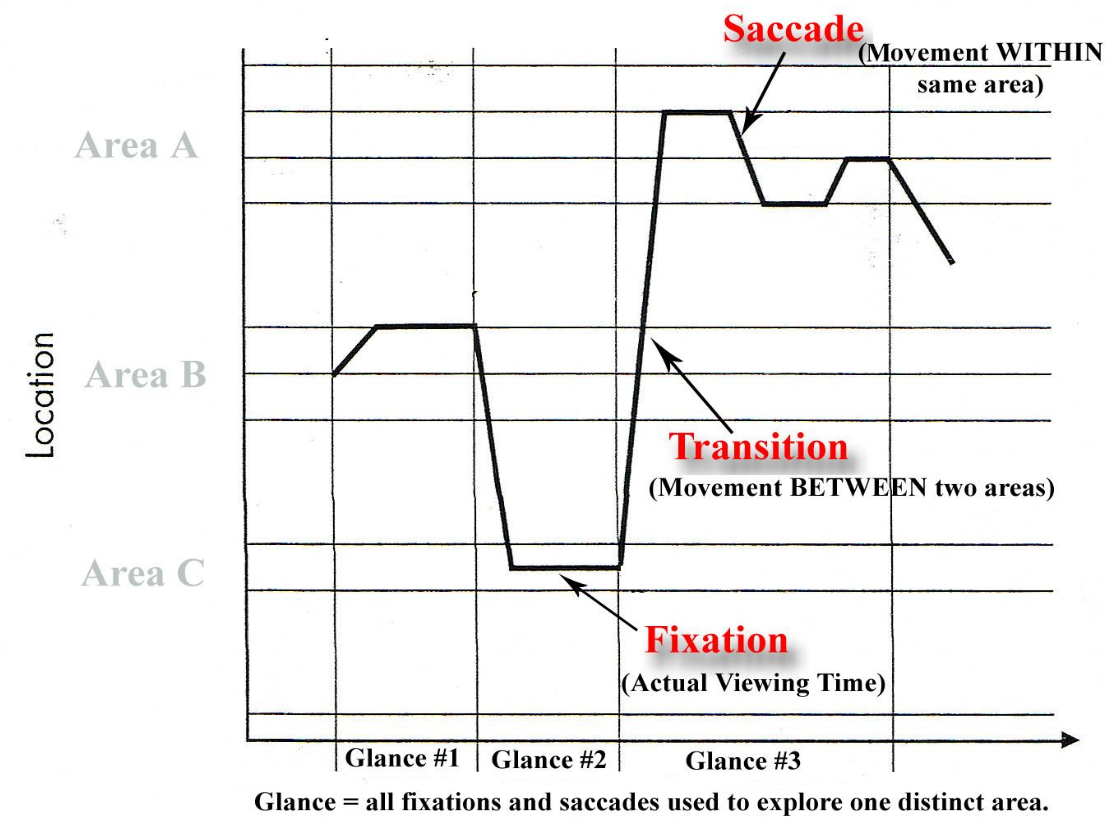


Figure 2. Pictorial representation of the relationship between a fixation, saccade, transition and glance on a given scanning sequence (modified version of Green, 2002.) Note the saccade that connects two different locations is termed a transition. This is because the saccade served to access a new target, *not extract* further peripheral details from a current target.

When one links all fixations, saccades, transitions, and glances together, the result is called a *scan path*. An example of a visual scan path can be seen in Figure 3. It is the scan path that is decomposed into measurable components for the entropy paradigm used in this study.

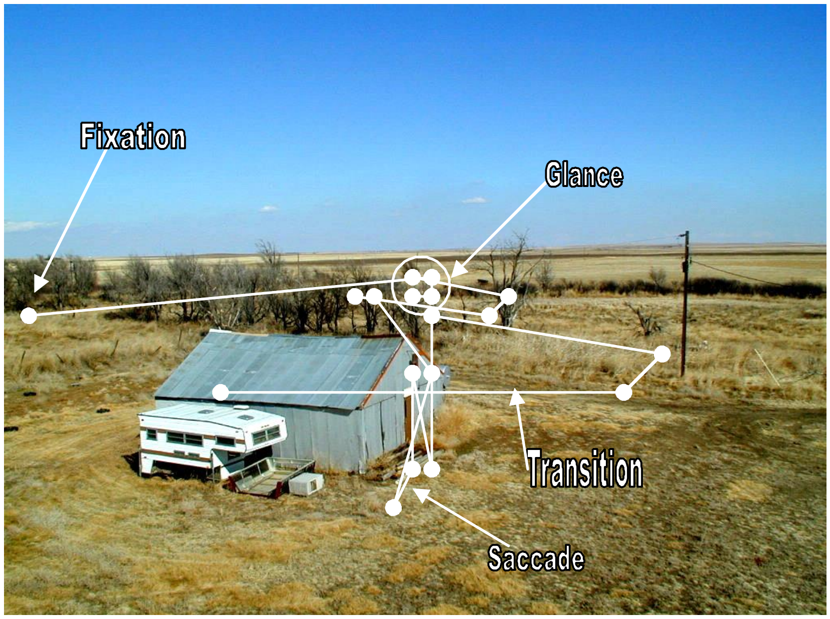


Figure 3. Example of a scene scan path. A scan path is the overall action sequence and includes all categories of descriptors (e.g. saccades, fixations, transitions, and glances.) In the above figure, dots represent points of fixation, whereas lines connect to the "next" fixation in the sequence.

The general concepts of eye-tracking and key terminology have now been reviewed. It is time to understand how visual scan patterns change in respect to driving. This is unique in theory because a new problem is now introduced: dynamic viewing.

1.3 Driving Related Eye Tracking Studies

First, several studies encompassing the most relevant of the "general" eye-movement and driving literature will be explored followed by situation specific variances. To help one understand how everything pieces together in this first section, Figure 4 from Serafin (1994) will be used to contrast straight-road glances observed between several different studies.

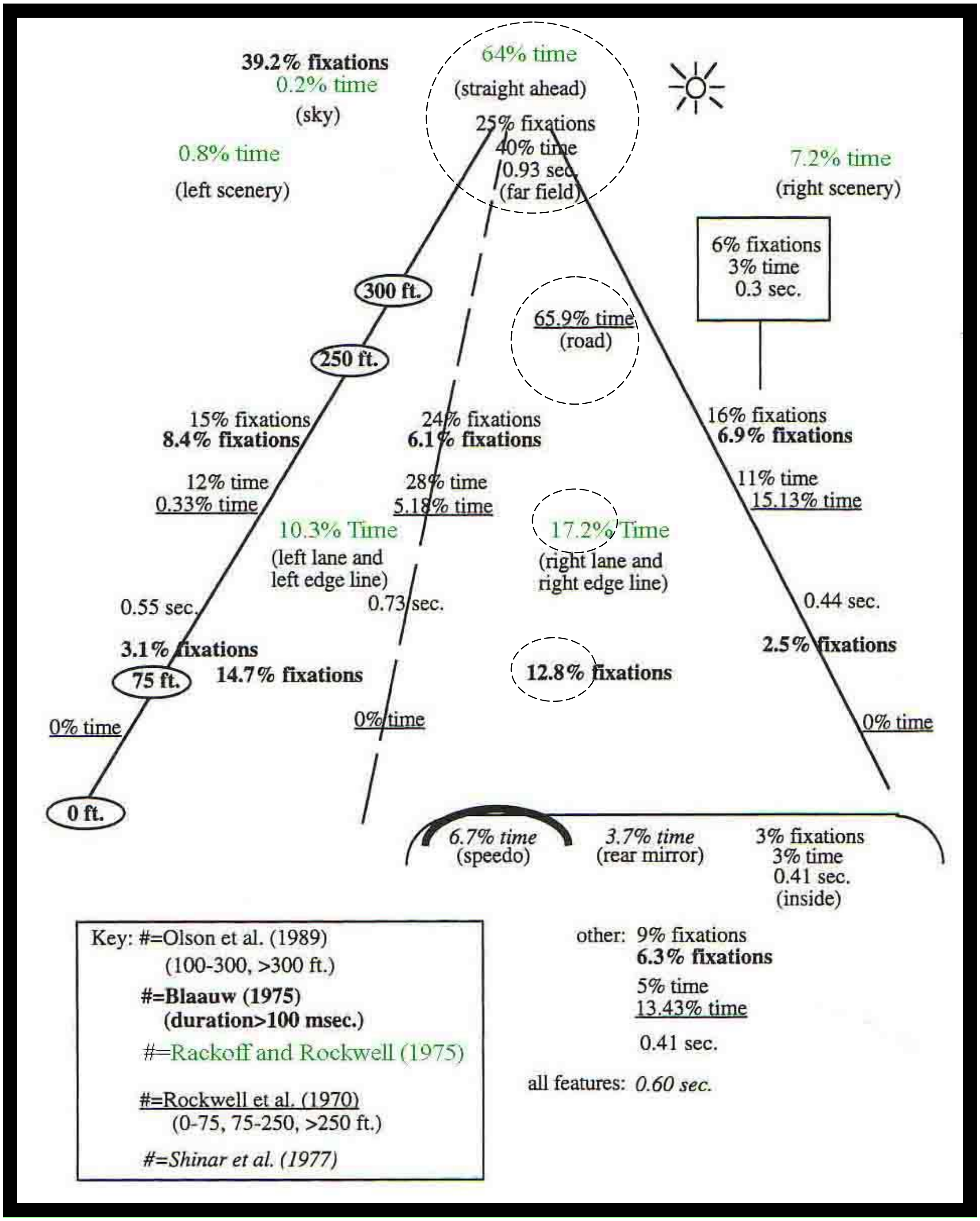


*Figure 4.*Summary of general trends observed in straight lane driving from cited literature as of 1994. (Serafin, 1994).

In figure 4, it is immediately notice that the majority of one's fixations are directed straight ahead (see circles). These can be segregated as looks focusing on the point of expansion (65.9%) or at the road (64%) depending on the study cited. This figure also shows how fixations are biased depending on the distance being viewed. The percentage of fixations seem to be directed more to the left side of the road than the right (3.1% left to 2.5% right at 75ft) when close to the vehicle, whereas a bias to the right is seen further away from the car (.8% left to 7.2% right when greater than 300ft). Edge and lane markings seem to be fixated fairly equally throughout, although the relative amount again varies by study: 15%left, 24% middle, and 16% right reported by Olson (1989) verses 8.4% left, 6.1% middle, 6.9% left reported by Blaauw (1975)(as cited in Serafin, 1994).

Straight Road Driving

Now that a general expectancy has been established, several individual studies from the "straight road driving" literature will be evaluated in more detail. The first study is real-world study done by Mourant and Rockwell (1970). This paper was an exploratory study designed to assess general eye movement patterns while driving. They evaluated eye movements of eight drivers on an expressway using a very early version eye-marker camera system that overlaid fixation position on a film. Subjects were instructed to drive six times through one of two pre-planned routes. Approximately 180 to 200 seconds of data was recorded during each run. The first note they make is that scan patterns seem to change as a function of "experience" with a noticeable "down and left" shift of the entire fixation cluster as the study progressed (see Figure 5). They also reported that this shift typically neared full effect around trial 3 (of 6) and was always skewed towards the right side of the road.

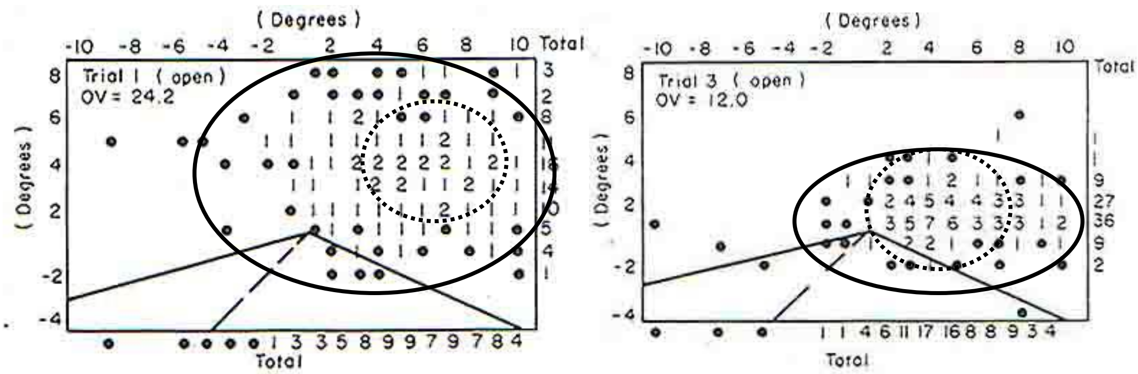


Figure 5. Adapted from Mourant and Rockwell (1970). These two figures demonstrate the effect of learning or “experience” from the 1st trial (left) to the 3rd trial (right). Notice how the fixation cluster’s range truncates and shifts closer to the road.

It was concluded that as the subjects became more comfortable with the road, they decreased the scope of their scanning and began to narrow the focus on areas where they learned the most relevant information was present. The majority of fixations were directed towards areas such as road signs, edge markers, turning lanes, oncoming cars, and parked vehicles.

When performing a car following task, Mourant & Rockwell (1970) observed increased sampling rates of edge markers and increased saccade travel distance from initial fixation. They also observed that the majority of fixations tended to be attracted to lead car as can be seen in Figure 6.

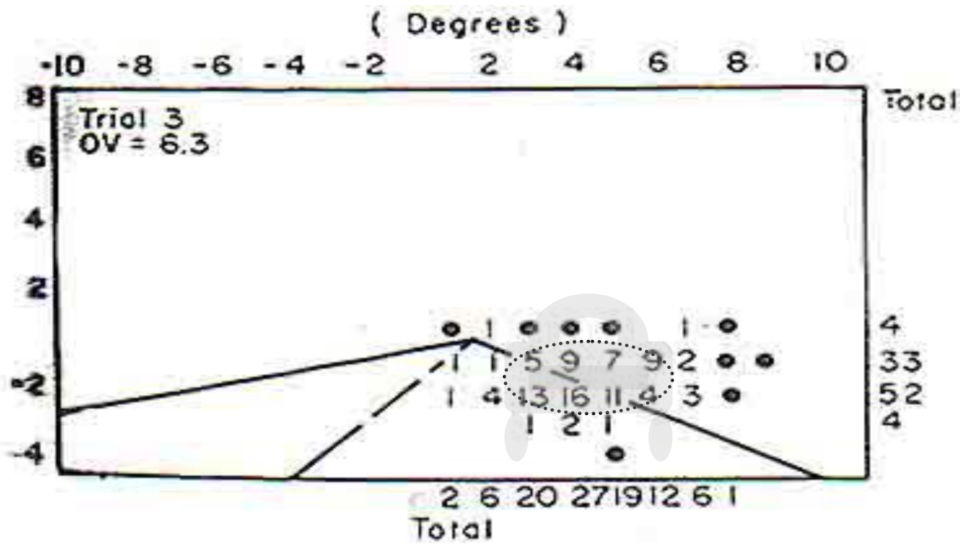


Figure 6. This figure demonstrates how fixations become directed during a car following task (adapted from Mourant and Rockwell, 1970). Notice how tightly the fixations are packed together under this situation. This is a potential confound to any eye-tracking study that has following involved.

In 1973, Bhise and Rockwell wrote a paper for the Ohio Department of Highways discussing some findings on sign reading behaviors derived from 8 separate studies (5 field, 3 laboratory). One strength of this paper was the use of naturalistic measurement techniques *in the field*. They observed that subjects do not concentrate on a sign after the first fixation, but rather begin to time-share with other objects on the road *under normal conditions*. This behavior was determined to be highly dependent on traffic density, instructions to the subject, length of message, familiarity, and relevancy of message. More simply put, scanning behavior is dependent on the cognitive load of the subject. In cases where a subject had a high cognitive load (special instructions, unfamiliar road, high traffic density, inadequate sign instructions, inadequate signing characteristics, etc.) the authors noted an increased dedication of time to concentration, a very low sign legibility distance, and a delayed comprehension of the sign's message.

In another study (Taoka, 1991) spare glances were statistically analyzed to determine how fixation duration interacted with task induced cognitive loading. He observed that people tend to have the longest fixation durations for goals that require the most brain computational power. Examples would include any activity that needs additional interpretation of meaning (signs, temp gauges, etc) or precise physical inputs (radio tuning, etc). The least amount of fixation time was needed for less cognitively intense activities such as simple position changes (flip a switch) or status updates (mirrors, speedometer, etc). This study further reinforces the idea of how variable the information content of a stimulus is and how it can directly influence eye-movements. It was noted from the Henderson lab studies that people tend to fixate more frequently to *informativ*e aspects of a scene in a global sense, but this study helps tie the logic of why one also tend to increase dwell time on these areas.

Now a few of the more *applied* areas of driving related eye-movement study will be examined. These areas will be deemed "situation specific", although they could easily be applied to situations other than just those specified. As a logical starting point, the most prolific bodies of research such as that involving eye-scan behavior modified by curves will be explored first. After that, a review of observations from several other studies involving age, driving experience, lighting, preview time, speed, fatigue, and visual scene complexity will be conducted.

1.3.2 Driving on Curves

Curve navigation is simply how one looks when attempting to navigate a curve. This is by far the most heavily researched area of driving and eye-tracking. Even though this literature is not directly applicable to the present study, it will be discussed in at least some detail due to the magnitude of the literature it accounts for.

In a paper written by Shinar et al. (1977), the fixations of five subjects on straight and curved rural roads were compared. The first finding was that fixations to the left and right side of the road are asymmetrical. There appears to be a larger bias to fixate to the right than to the left. On straight roads, they reported that scanning was very inactive with most of fixations centering mainly on the area of expansion in front of the vehicle. On curves roads, they noticed that fixations tend to follow the geometry of the road. When approaching a curve, they noticed a shift of eye gaze *into the* curve that occurs *several seconds before the curve.* They attributed this gaze shift to the anticipation of the curve. The authors suggest that curve navigation is a more visually intense activity since it requires constant foveal involvement whereas straight lane driving can be maintained using mostly peripheral vision. Since curve navigation is more resource demanding than straight lane driving, it is logical to conclude that this is why fixations need to be more directed (follow geometry of the road) in order to perform that activity.

In another study comparing straight verses curved roads, Olson et al. (1989) reported that eye fixations were uniformly displaced over the road when straight line driving but follow the direction of the curve when turning. They broadened this finding by conducting the same experiment at night where they again found that the same trend except that it was now simply limited to a narrower range mainly confined within the area illuminated by the headlights. This study also observed the same tendency as Mourant and Rockwell (1970) did when their subjects followed a lead car. Olson et al. noticed that when involved in a following task, the majority of fixations are spent looking directly at the lead car and that this effect is even more pronounced at night. Figure 7 adapted from the Olson study, illustrates the differences in how visual information is fixated as a function of day/night, curves, and when following a lead car.

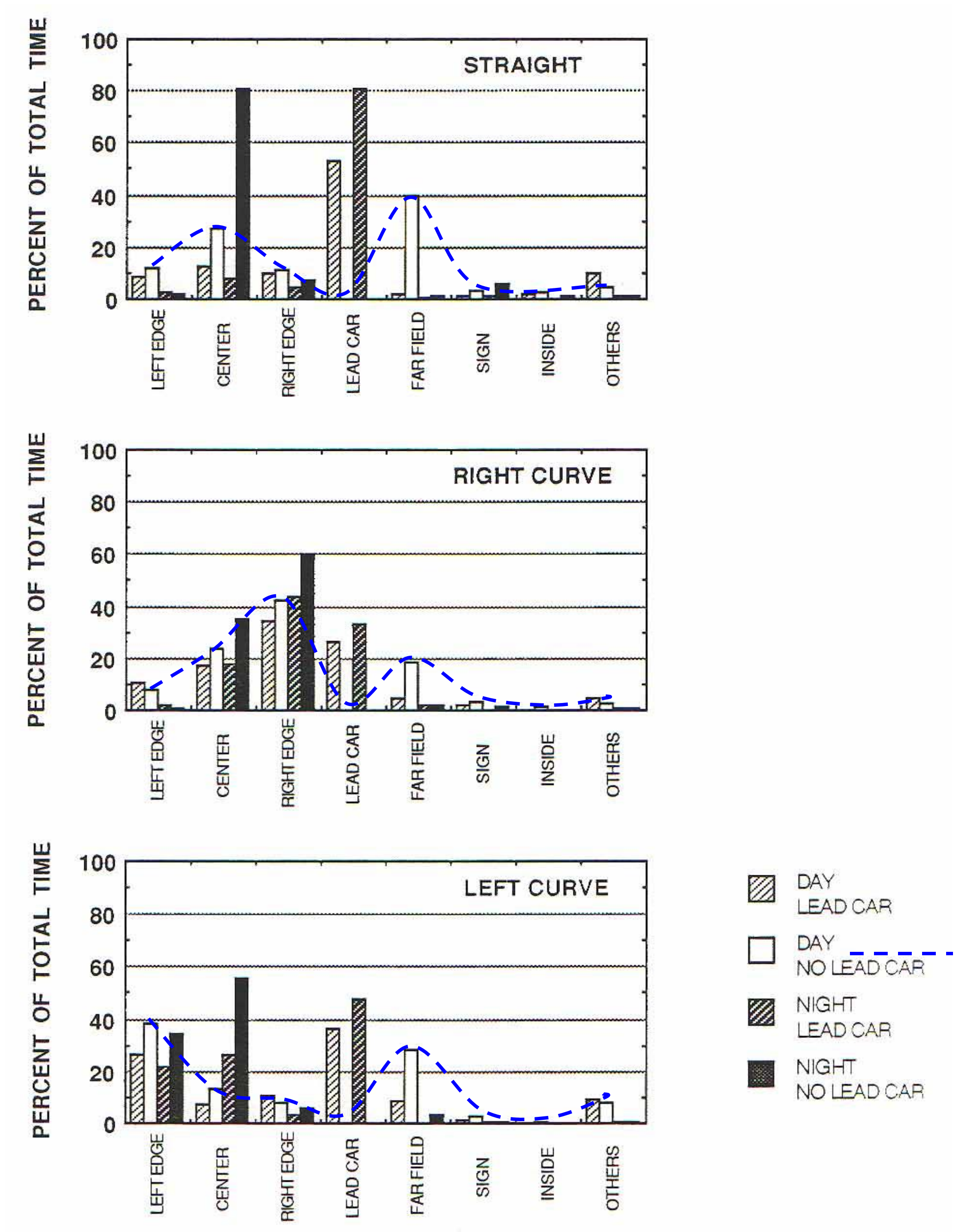


Figure 7. These three charts taken from Olson et al. (1989) show how the monitoring percentages (areas that fixations are dedicated to) change between straight and curved road driving under several unique conditions. Since the "Day – no lead car" scenario is most important to this study, a trend line has been superimposed over the data to help highlight these findings.

Land (1992), Land & Lee (1994) and Land & Horwood (1996) are three more studies that discuss curve fixation phenomena. In the 1992 paper, Land begins to form his hypothesis that head and eye movements may be correlated with each other under certain conditions. In the (1994) paper they compared data from *eye* and *steering* movements during curve navigation to see how well this correlated. Using that data, they ultimately concluded that subjects seek out a tangent point when navigating turns. They also concluded that this point is sought out approximately 1-2 seconds before the curve and periodically re-fixated throughout the curve in order to coordinate with steering. Although taken from the next paper in the series (1996), Figure 8 illustrates this sequence.

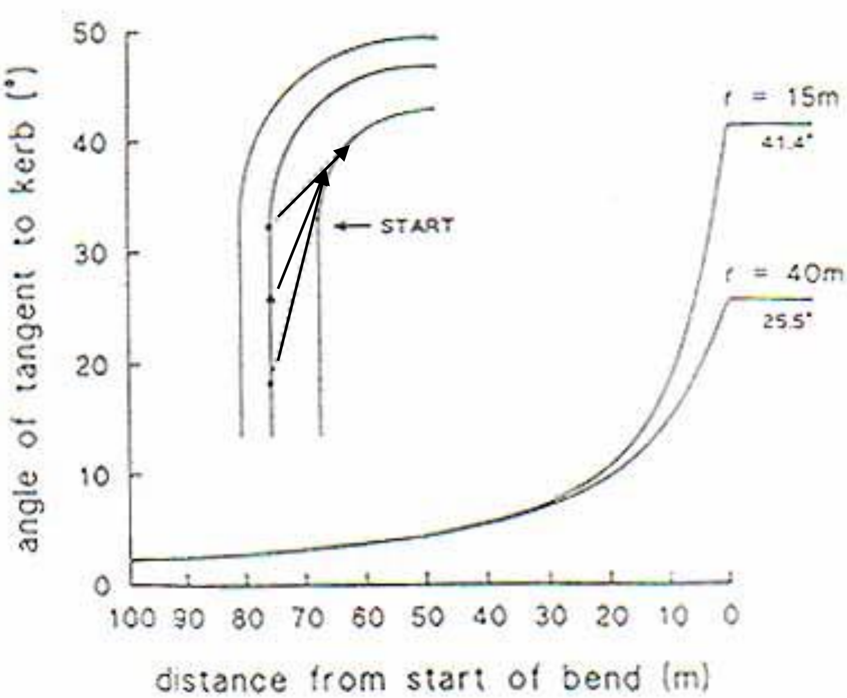


Figure 8. Notice that the tangent point is sought out before the curve, and how it is revisited throughout the turn as an update to the turn. Each point on the middle line represents the current location of the driver in relation to the curve (adapted from Land & Horwood, 1996)

In the 1996 paper Land & Horwood revisited the *eye* and *head* movement concept in a new study using two different curve scenarios and three subjects. It was hypothesized that there are three main phases of gaze behavior for curve navigation. The phases are as follow:

1. Phase 1: The eyes first seek out the curve but no head movement occurs.
2. Phase 2: After noticing that a curve is coming up, the subjects then begins to align the head in a way as to point the eyes into the curve and away from the current vehicle heading.
3. Phase 3: Lastly upon exit of the curve, the eyes lock on a distant target and the head rotates to once again become in alignment with the cars heading again. These phases are illustrated in Figure 9.

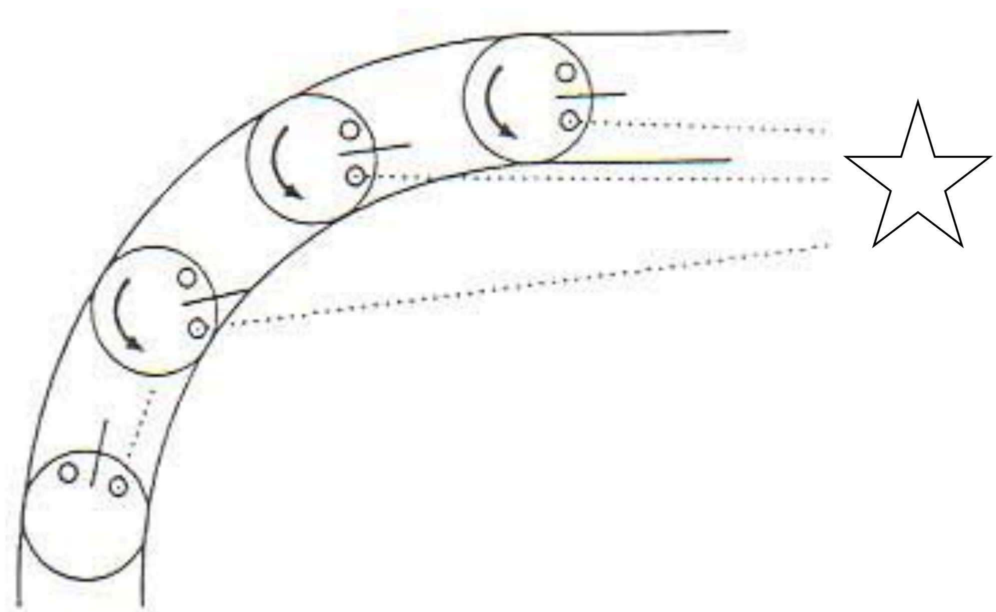


Figure 9. This picture demonstrates the projected path of the head in order to realign with the road when coming out of a curve. After the halfway point, the eye breaks away from the tangent point and searches out a distant target, the driver then maintains this fixation until the exit of the curve. (adapted from Land & Horwood, 1996)

Now from a more applied perspective, three more papers will be discussed that explore the eye tracking and curve navigation relationship when "experience" is added to the mix. Since the main variable is still "curve navigation", these papers will be discussed in this portion of the report, however, there will be more review of this topic in a later section exploring “experience” as the primary variable.

The first of the "experience & curve navigation" papers cited is a paper written by Cohen and Studach, (1977). This paper measured the eye movements of five experienced and four inexperienced drivers while navigating left and right curves. They noted that experienced drivers seem to fixate longer to the left in left turns than to the right in right turns while the inexperienced did not note this trend. An experienced driver's right curve fixation duration was also typically shorter than those observed in the inexperienced group. Overall, it was concluded that eye fixation patterns begin changing before the curve, seem to be directed towards the area of future driving path, and are affected by the driver's experience level.

Dishart and Land (1998) wrote the next "experience & curve navigation" paper to be summarized. This paper again reported that curve navigation strategies are influenced by experience. They observed that initial runs (novice) seemed to elicit an increased usage of tangent points for heading determination as compared to an experienced group when navigating the same curve. The experienced driver tended to use more peripheral cues for heading as they essentially already have a model of what to expect in their mind. However, this observation seemed to peak in the novice group as eventually even the novice will accrue enough “experience” and begin to direct fixations more like the experienced group. The now "experienced" novice will then begin to direct more fixations elsewhere as they have learned to more efficiently optimize their visual search patterns for this specific situation.

To test if these learned scanning patterns were truly attributable to experience, another study was created on arguably the most highly experienced drivers around, racing drivers. This paper written by Land and Tatler (2001) studied the eye-movements of a racing driver during a high-speed practice session. What they found was that the driver spent the majority of time looking near *but not at* the tangent point like that commonly observed in ordinary driving. The altered positioning of the head as to not line up with the expected tangent supports this data. The authors reason that since the driver has such an exaggerated amount of experience, he now is able to use models in his mind to plan future vehicle movements rather than visual cues. He essentially no longer needed to plan routes exclusively using road tangents; he only uses these previous visual cues as confirmation for what he expects. This extra freedom allows the driver to plan other alternative actions such as vehicle passing, etc.

The curve-driving literature was explored simply because it dominates the body of literature dealing with eye tracking while driving. Even now after only reviewing this narrow focus of navigation literature, it is easy to envision how different situations will elicit different scene scanning adaptations. It is now time to investigate other commonly known influences of eye-scanning modification, except this time from the straight-lane driving literature.

1.3.3 Driver Experience (Novice versus Expert)

Chapman & Underwood (1998) sought to simply describe the eye-scanning differences between novice and experienced groups while watching hazardous videotapes. One hundred and twelve subjects participated in this study. Twenty-nine of these were classified as experienced. The main finding from this study was that novice drivers seemed to have much longer fixation durations during hazardous situations as compared to the experienced groups, although, both groups exhibited the same effect to some degree. They reasoned this to be due to the increased cognitive load a novice brings to the table as they have not yet learned how to separate “hazardous” from “non-hazardous” situations, effectively forcing them to process much more of the scene than needed. In other words, experienced drivers are more able to form alternative search strategies that will result in a quicker extraction of relevant information from a scene. Another relevant finding from this study is that they note a global effect of fewer fixations with longer durations in a rural setting as compared to urban. From a logical standpoint, this fixation discrepancy is most likely due to the fact that there is simply less to look at in a rural environment. This study again supports the idea that one's ability to detect a stimulus *can* be influenced by experience.

Crundal et al. (2002) later tested the idea that novice drivers simply "see" less of the world when driving. They evaluated 43 drivers using a lab based visual display showing scenes of potentially hazardous situations. Experience was determined by number of months that the individual had been driving. Subjects were simply required to just step on a footpad when a hazard was noticed. What they concluded was that novice subjects weren’t actually "seeing" less of the peripheral environment (targets), they were actually just using an alternative and much less efficient search strategy due to inexperience on how to search. The more experienced drivers tended to dedicate more attention to peripheral targets in hazardous situations allowing them to extract more detail rather then merely gleaning surface information like the novices. This is to say that they fixated the same information the same number of times as the novices, except they were able to retain more information about what they saw. This was represented by a much higher detection rate in all loading levels (high, low, none) by the experienced drivers. The authors hypothesized that the effect is due more to an inability to effectively deploy attention to peripheral targets, more than missing the targets due to a "tunnel vision" effect where you just simply don’t look to your periphery.

Falkmer and Gregersen (2005) summate what can be expected when comparing experience. This study's sole goal was to confirm many of the previously reported differences between experienced and inexperienced drivers. Overall, this paper provides a nice outline of the “experience” relationship to eye-scanning behavior. Using forty subjects, they confirmed that compared to a more experienced subject, novices’ exhibit:

1. Characteristically more in-vehicle fixations (gauges, dash, etc.)
2. A decreased tendency to scan the horizontal periphery
3. A larger number of fixations directed toward irrelevant traffic cues
4. Increased fixation durations on perceived hazards.

1.3.4 Challenging Visual Conditions

Another potential influence of eye-scanning strategies is one that is often overlooked: lighting. Lighting is an ever-changing variable relevant to the driving world because people drive in situations of decreased light (rainstorms, cloudy days, civil twilight, nighttime, etc) all the time. It is well known that vision is essentially a function of contrast recognition, so it is only logical that anything with the potential to reduce contrast is going to require some form of compensation. In this case that compensation presents itself in the form of alternative eye-scanning strategies. What follows are several papers exploring eye-tracking while immersed in different lighting conditions.

1.3.4.1 Driving in the Rain

In Zwahlen (1980), he explored how eye scanning strategies change as a function of a true "real-world" hazardous situation: wiper failure during rain. Unlike the majority of studies using only simulators or videotape representations, this study was done in the field. Zwahlen used 4 subjects each participating in 16 trial runs during dry, light rain, and heavy rain condition. Wipers were intermittently turned off for 9-18 second intervals randomly by the experimenter. Overall, they found that fixations tended to increase in number and decrease in duration as the rainfall level increased. The main area of fixation placement slowly migrated from the periphery/edge lines in dry conditions to closer to the center in front of the car with increased rainfall as can be seen by the "crosshairs" in Figure 10.

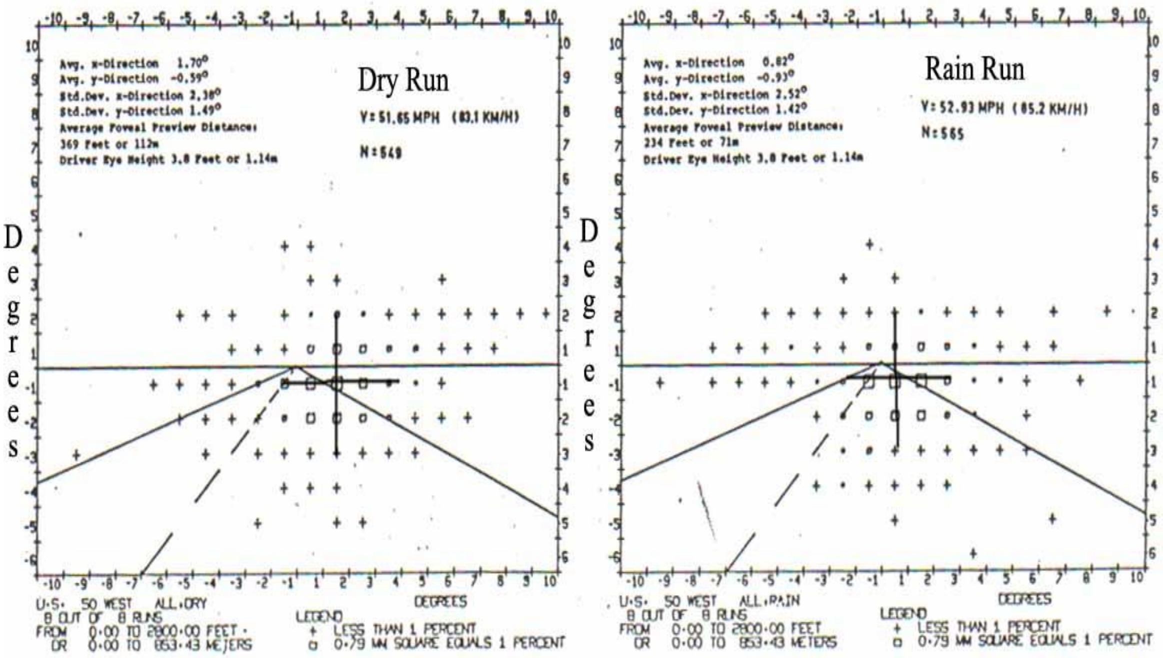


Figure 10. This figure is showing how the fixation cluster remains similar in size, but move as a whole closer to the center during a wiper failure (adapted from Zwahlen, 1980). "Crosshairs" have been added to aid in visualizing the center of the fixation cloud.

Preview distance in front of the car degraded from a fairly long foveal preview of 112 meters in the dry condition to only 62 meters during wiper failure. Eye blinks increased from 1.82 blinks per minute in dry to 6.67 blinks per minute in the rain. Fixations on lane markings also increased significantly as a function of rainfall density. It is reasoned that as peripheral cues become less salient due to the rainfall, one begins to use direct foveal inputs to maintain lane positioning. In other words, since the subject could no longer see anything in the periphery, he simply modified his search to no longer scan it and instead increases the monitoring of what he can see. Another interesting note is that once a shift in scanning occurred, it sometimes persisted for up to 10 seconds after the event ended suggesting that the event was stressful enough and to hinder immediate resource reallocation.

In a later study from the same lab, Zwahlen and Schnell (1999) reported another interesting paper dealing this time with civil twilight where lighting levels are dramatically diminished from daytime levels. They observed that the current federal definition (3.2 lux) may not be an accurate estimate of lighting under twilight conditions. They suggest the actual lighting levels may be closer to (1.7 lux) with clear skies to only (.08 lux) when cloudy. Signs designed to be interpreted at this federally defined level will not be as efficient when observed at these lower lux levels. They found in field trials that most signs exhibited shorter than predicted recognition distances when designed for this exaggerated illumination level. A direct result of "hard to read" signs, is an increase in visual/cognitive loading. Anything that loads the visual system will undoubtedly influence eye-movements into a compensation strategy designed to get the most out of “what you're given”. While this paper does not explicitly report changes in fixation patterns, it does provide a theoretical base that may explain any differences possibly seen due to low light conditions. Lower light will increase your cognitive load as targets need to be processed longer to derive the same meaning. Since driving is a time dependent activity, anything that serves as a "limiter" will force you to prioritize your scan patterns.

1.3.4.2 Reading Signs at Night

Schnell and Zwahlen (1999) was conducted at night on a straight road while using only “low beams”. They reported the trend that brighter pavement markings increased the number of longitudinal eye fixations. In other words, as the markings were made brighter, the subjects began looking further down the road. It should be noted this was somewhat mixed results even though they reported it as significant. They ran nine subjects of which only 6 were used. Of the six, one was much older than the rest (age 53) so there is the potential for data skew. This older subject along with one of the younger subjects did not show the same scanning effect as the other four and seemed to look immediately in front of the car no matter what the condition. This paper does show again how eye-movement patterns can be modified (which is main goal of review), but it may not accurately represent the entire ranges that can be expected to be seen. Do all older subjects see this reverse effect like was observed by the single older subject in this study, or was he just an outlier like the one younger subject appeared to be?

In another sign study by Schieber et al., (2004), sign preview distances were assessed as a function of sign reflectance during nighttime driving conditions. Drawing upon the participation of forty subjects, this study reported significant reductions in detection/recognition distance as sign reflectance was changed from "new" levels (100% reflectance) to the level suggested as the FHWA minimum retro-reflectivity level (15% reflectance). It was also reported that urban situations elicited significantly shorter legibility distances (24% decrease) compared to the rural environments (17% decrease). All recognition distance responses in both rural (32-36ft per inch of letter height) and urban settings (30-37ft per inch of letter height) were observed well under the estimated legibility distance (40ft per inch of letter height) using the federal minimum of 15% reflectance. Another interesting finding is that although no significant age difference existed, there was an overall shift in performance of the older group compared to the younger subjects. It seems that it wasn’t that the older subjects were impacted MORE than younger counterparts due to a decreased reflectance; rather they just performed less optimally throughout ALL conditions suggesting other influences. One aspect of this study that needed more clarity was the influence of speed. This brought about another study in 2005.

The follow-up study designed by Schieber (2005) looked specifically at speed's influence on legibility levels in respect to differing retro-reflectivity levels. In this study, subjects drove at speeds from 5mph to 60mph while observing signs of 15% and 100% luminance. Overall, Schieber stated a 30% drop in legibility distance as a function of speed while only a 12% drop due to reflectance. However it should be noted that the (speed X reflectance) interaction was not significant. In other words, one *will* loose legibility distance due either speed or reflectance independently, but there seems to be no compound effect.

1.3.4.3 Visually Cluttered (i.e., Complex) Backgrounds

Another area of the visual domain that impacts how one scans a scene is simply the complexity of the scene itself. The more visual "information" present, the more effort it takes to extract meaning. This is the idea behind a paper written by Miura (1990) entitled: "Active function of eye movements and useful field of view in a realistic setting". This study looked at peripheral target location performance as a function of the complexity of visual clutter in a viewing scene. Miura observed that as the complexity of the traffic environment increases, both the reaction time for secondary tasks and the number of total eye fixations per unit time will also increase. Another key point is that peripheral target detection DECREASED, regardless of the fact that overall fixation spread seemed to stay similar between conditions. They were still scanning the periphery, just not deriving as much data (also observed in Crundal et al., 2002). In fact, it was noted in this paper that under the highest workload conditions, subjects actually tended to glance to the periphery with an even increased frequency.

Miura noticed an increased latency after each fixation leading him to believe that each fixation must be being processed more deeply even though it wasn't covering as wide an area. This in effect would limit the amount of "information" that could be derived at any given fixation. One can think of this by thinking about laying a paper circle on top of a picture. Instead of a fixation extracting information over a 10° wide circle, they might now only be deriving information from a 5° circle, however still need the same amount of processing resources for both. The increase in visual complexity means that the same amount of information once contained by a 10 °circle is now represented by a 5° and so forth. This study emphasizes two important "visual scanning" behavioral changes associated with increased visual load:

1. One will use an increased number of eye-movements to scan an area.
2. Each fixation will extract its information from a narrower effective angle.

As visual load increases, a subject will increase numbers of fixations, but process so deeply at each fixation that the overlap of fixation peripheries no longer occurs. This gives more opportunity for a target to occur in the newly formed void. This is the foundation for the "tunnel vision" hypothesis proposed by this author.

1.4 Role of Head Movements

Now is a logical time to talk about how eye movements are influenced in respect to visual flow. In the Land et al. (1996) paper it was reported that the positioning of the head plays a roll in determining the range of the field from which the next fixation will occur. The question that arises from this observation is if the head movement also drives heading estimations?

In Cutting et al. (2000) it is proposed the monitoring of visual flow and not head movement that is the main component of one's heading calculations. From this point of view, it isn't that the environment is moving and one just points his head in the direction he thinks it is moving, rather, one actually needs to actively search out objects an calculate heading dynamically. This is important to eye tracking because it is a potential source of scan-path influence since it asserts that *all movements are directed.* One can't drive without knowing heading which means it will be impossible to completely eliminate this search strategy in a driving situation given this hypothesis is true. While this shouldn't be a problem in this study since the route is the exact same for each person in each condition, it is important to be aware of.

In a simulator study conducted by Cutting et al. (2000), the authors attempted to address how people use landmarks (i.e. terrain features) to determine heading. In this study, the main target stimuli were pairs of trees that visually converged of decelerated apart. Cutting noted that subjects had fewer fixations (1.28 near, .98 far) and lower duration of fixation (820msec near, 578msec far) on the furthest objects as compared to the closest. They also observed that object displacement size was judged the majority of the time by looking at the trees themselves (63%) and NOT the gap size between them

(24%). Also, if a subject had fixated the trees immediately *before* making the heading determination, they were much more accurate in a heading assessment than if they had not recently viewed these cues. The authors made the distinction that the peripheral visual cues *most often used to guide* heading are also the cues that can qualitatively defined as "containing the most informative visual information." They ultimately conclude that the best heading determinations need some form of reference object to monitor and that one has to be *actively* involved to remain accurate. Although this study is a lab study, its influences should also be observable in the field since heading determination is an important aspect of driving. It is entirely possible that many of the fixation sequences observed during driving will be purposefully done for this update of one's heading model. If this is the case, it should show up as some sort of regular-interval organized structure to the scanning pattern, which is ultimately what the visual entropy metric seeks to quantify.

1.5 Fatigue effects

Another commonly considered cause of eye-scanning strategy changes is fatigue. One such study exploring this as a potential source of influence is Galley and Andres (1996). In this study the authors looked at the correlations of alcohol, fatigue, saccadic movements, and eye blinks while driving either on the Autobahn or in a city setting. A significant difference between city and Autobahn driving was reported for most measures, however not for the reason first speculated (fatigue). There *was* a significantly higher saccadic amplitude, saccadic velocity, blink velocity, and blink rate in addition to overall lower fixation durations in the city condition compared to the Autobahn. *However,* the fatigue effect noted in both driving situations was "unexpectedly small". Most the variance was accounted for by the changes in visual complexity of the scenes rather then fatigue. This is not entirely a surprise as city environments can easily be reasoned to be more distracting and visually dense, whereas the highway condition should have much less visual clutter.

1.6 Driving Speed Effects

Spijkers's (1992) paper dealt with the concept of driving speed having an impact on how one views the visual world. In this study six subjects were simply instructed to drive along three different road types (urban 4-lane, rural 2-lane, & urban street) at two variations of speed (30-50km/h on street, 50-80km/h others). The main finding was that under increased speed conditions, subjects tended to fixate task relevant items with increased frequency. However, fixation numbers were more dependent on the road type and traffic density than on speed. There was also a noticeable increase in the numbers of fixations dedicated to the road under higher speed conditions, although this was only marginally significant. It was also reported that the number of extreme glances (saccades > 18°) occurred more frequently during lower speeds with an overall narrowing of saccade glance size as speed increases. All rural conditions seemed to represent higher numbers of roadway fixations (83.68%) compared to either of the urban conditions (71.2% 4-lane, 70.56% street). This paper is important because even though it shows that a speed effect did exist, it pointed out that it was very minor. They concluded that the greatest contributor seemed to be the visual complexity of the scene rather than speed. This is good to know in order to prevent a potential validity threat to our experiment. In the present study, speeds were kept constant, but this Spijkers's paper reinforces the idea that even with some speed variance, a significant tolerance could be assumed granted the visual context stays constant.

1.7 Age Effects

In 1978, a study by Shinar et al. was published in Human Factors discussing age related changes in driver spare visual capacity. In this study they instructed 6 young (aged 20-25) and 9 older subjects (aged 63-70) to drive in each of 4 conditions (day/night open road, and day/night following) while keeping their eyes closed as long as possible between update looks. They observed that older drivers' need to keep there eyes open longer than younger drivers for visual data extraction. This finding mirrored the performance of an embedded figure test given before the driving test. If subjects typically needed longer to find embedded figures, they also tended to need longer update times while driving. This study provides a framework for behavioral changes that in later studies will ultimately be reflected by eye-movement scanning pattern adaptations (increased number of fixations, etc.)

The next aging related paper that will be discussed is one written by Rockwell (1988). This study was designed to explore visual-sampling differences needed for different stereo designs. The results are the composite of three studies. They processed 106 subjects each over 1-hour trials on a given stretch of highway. Groups were evenly divided by age. "Aged" in the study was anyone over the age of 45. Due to this small age gap, the results they observed could be reasoned to be no where near as pronounced as if they would have been using 65+ subjects. Nevertheless, they still observed that older subjects tend to have longer average glance durations and needed more glances in order to complete a stereo tuning task as compared to the younger group. They also observed that if you made the interface less intelligible, you could elect longer fixation durations in all groups (20% increase in duration) *as long as* the interface was imperative to the task. However, if it was an optional target, the driver tended used more fixations rather than increasing duration of each to accomplish the task.

The last observation they note is that even though fixation durations can be adjusted by making the interface less intelligible, the impact of the *demands* of the driving task proved to be statistically more reliable at predicting these changes than the visual characteristics of "in-car" targets. In a high traffic condition, the average fixation duration of interior interfaces dropped nearly 20% when compared to low traffic density situations.

In a paper written by McCarley et al., (2001), aging differences in scanning pattern were observed while subjects were immersed in conversation. Twenty-eight subject split evenly into two groups by age (mean = <21 young or >68 older) were instructed to free view traffic scenes that would periodically have targets (streetlights, stop signs, etc.) "erased" briefly then reappear. The goal was to simply detect these changes as they occurred and report what the change was. They used two task scenarios: (a) simple detection and (b) detection with loading task. The loading task consisted of being actively involved in a conversation with a confederate during the main detection task. Older subjects exhibited significantly increased numbers of fixations (mean = 27 per trial in young, 42 old) and decreased fixation durations (mean = 301ms young, 262ms old) as compared to the younger group, although age was not significant with the task interaction. Older subjects also had much slower reaction times for detection recognition than the younger counterparts. Visual changes occurring within 1 degree of a current fixation resulted in 83% detection in younger verses 77% in the older suggesting a UFOV decrease, but again this did not seem to interact with task. Overall, the main conclusion is that age in itself seems to be the main reason for performance decrements, not the task.

Keeping with the aging theme is a paper written by Perryman and Fitten (1996). These authors did not set out to directly measure eye scanning, but it did make one important observation about older drivers nevertheless. They noted that the older drivers in their study made many fewer "eye-movement excursions" than the younger group. This is to say, that older subjects did not fixate as much towards peripheral targets as the younger group. Whether they still saw the targets in peripheral vision but chose not to fixate is undefined, but this is still an interesting finding.

Next is a paper written by Maltz et al., (1999). This two-part study sought to identify trends in eye-scanning behavior contrasted by age. They used ten subjects split evenly into two groups (young <30, 62< Old). In both studies, subjects were simply asked to search out and identify (in order) a set of numbers (1-14) overlaid onto a picture of a traffic image (see Figure 11).

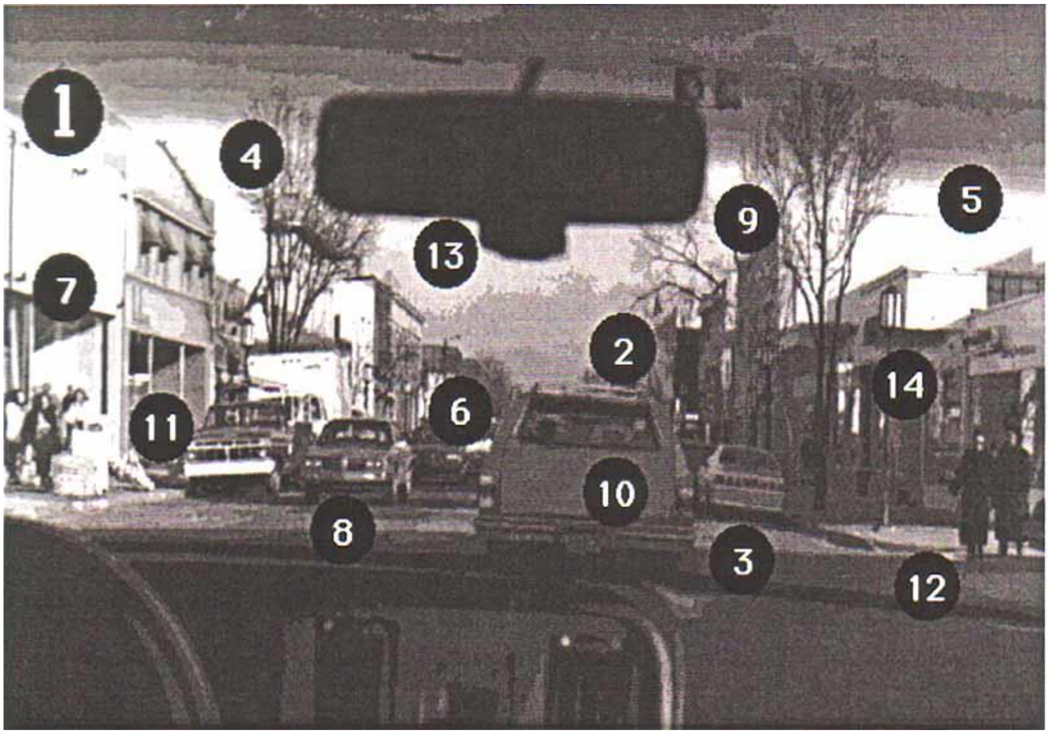


Figure 11. Sample image of driving scene. (taken from Maltz et al., 1999). The subjects would be presented a similar picture and asked order the numbers as fast as possible.

The first of the Maltz et al., (1999) studies used traffic pictures taken *without* a driving reference (not in car), while the second study used pictures cued *from the perspective of the driver* (“in car” like Figure 11). They mainly observed that older individuals: (a) Needed much longer search episodes until completion (27.0 seconds in the old, 16.7 seconds in the young)) and (b) Required more total fixations (87 old, 55.7 young) than the younger group. Older subjects also tended to scan smaller subsets of the picture more, move less distance between fixations (however almost negligible) and even revisit areas previously scanned whereas younger subjects seemed to scan the whole scene more evenly and rarely revisited. It should be noted that even with all these dramatic divisions, the actual durations of fixations *did not* vary by group suggesting that mental processing of each fixation was similar between groups. If this truly is the case, the differences must be due instead to another mechanism varying between groups. Two possible mechanisms have been proposed. The first hypothesis is that an older individual simply uses an inferior search strategy due to the decreased mental efficiency inherently related to age as suggested by Korteling et al., (1990; 1991). The second hypothesis is that older subjects are exhibiting a Useful Field of View reduction (Ball & Owsley, 1992). In the UFOV paradigm, they are processing each point with the same amount of mental effort as the younger group, but the size of the area being processed is more compact.

Another paper reporting aging differences is one measuring hazard detection performance with age. Fildes et al., (2005) ran this study using 40 subjects (young <35, 65<old) in simulator conditions. In the older group, they reported much broader scanning behaviors, slow hazard recognition speeds, less time fixating on hazard once acquired, and a greater number of glances away from the visual scene (mainly toward speedometer) when compared to the younger group. They also reported that older subjects tended to drive slower than the younger subjects. Given the trends observed so far, one would reason this was most likely a coping strategy employed by the older subjects to increase the amount of time available for visual processing and lessen the cognitive load caused by the time constraints of a dynamic environment. Slow down the visual flow and by definition one will slow down the amount of information present to be processed during a given unit of time.

In the same year as Fildes paper, another one written by Underwood et al., (2005) was published. This paper looked again at hazard perception except they derived dissimilar results. They used 24 subjects (young <40, 60<old) who were allowed to free view video clips for potential hazards (cars, pedestrians, etc). This study used a

technique similar to our study by assigning a value to each persons scan path using Markov Matrices. In this study they actually found that there was no statistical difference in speed or difference in visual detection of hazards as compared by age, nor was there any major difference in scan patterns between groups. The only finding they did establish was that older drivers objectively rated the scenes as more visually intense than the younger drivers did. While this study did not observe an effect, it is still important to serve as a contrast to all those that did. It should also be noted that this study did not rely on a driver in the loop during data collection. This in effect may have been able to free up cognitive resources in the older group allowing them to perform at similar levels over a wider range of conditions. They also confined the definition of a "hazard" to ONLY other road users. This could have served to bias glances some as it would allow one to form a search model excluding much of the visual environment. This would further allow one to dedicate attention to a smaller area that when compounded with simple free viewing, might explain the reason no significance was noted.

Serafin (1994) was a study done for UMTRI to ascertain if age results will vary in comparison to the wide body of curve navigation literature that did not account for this variable. This paper starts with a very good summary of general trends (see Figure 4), ends with a follow-up study exploring if there are any deviations from these expectations trends with respect to age. This author measured 32 participants (16 under 35 years of age, 16 over 60) as they *actually drove* around a 2-lane test track. She observed that for all groups there was a larger preference of fixations toward the right scenery on straight roads verses curved, more fixations on higher curved roads than less curved, and all subjects seemed to fixate as far down the road as possible when possible. When compared by age, Serafin found that young drivers tended to have fewer (30) but longer (174ms) fixations than an older subject (36, 145msec), but that the locations of these fixations were not significantly different by age. Again, more fixations are needed as a coping strategy for the older subjects, however, the strategy did not dictate where they choose to look.

In a paper entitled "Visual demand of driving curves as determined by visual occlusion," curve navigation demands were observed by Tsimhoni & Green (1999). They used 12 subjects, 4 in each of 3 groups determined by age (18-24, 24-55, and 55-68). Using a driving simulator that only showed the driving scene when needed (key pressed to update), they determined that the highest visual load is approximately 100-meters before a curve for all groups. Also, the tighter the curve, the more visually loading it was. This mental loading was supported by a subjective questionnaire given at the end of the trials. In respect to age differences, there was more visual loading (as indicated by more key presses) on the older subjects (aged 55-68) of the study compared to the younger (ages18-24) in all situations; however this was only moderately significant. Another interesting finding was that while younger subjects tended to exhibit a slight decrease in visual load due to learning between trials (few update key presses needed), the older subjects did not seem to exhibit this meaning their visual load remained constant.

1.8 Mental Workload and Driver's Eye Movements

Workload is a broad concept encompassing how "hard" or cognitively engaging a task is. It has been applied to *many* situations and has a broad base of literature associated with it. However, even in this broad base of literature, there are only a handful of studies actually correlating workload *with eye-movements.* This is even more truncated if one wants these results to have been measured *while driving*. Many studies report changes in driving related detection rates that they *hypothesize* to be due to scanning, but most do not empirically observe actual fixation patterns.

One of the most widely reported findings in driving related workload is that people will miss targets in the periphery under increasingly stressful situations (Schieber & Gilland, 2005; Miura, 1990,1992; Ball & Owsley, 1992; Harms 1986, 1991; Crundall et al., 2002) however the logic behind why this occurs is still debatable. Two examples of this peripheral degradation can be seen in Figures 12 and 13.

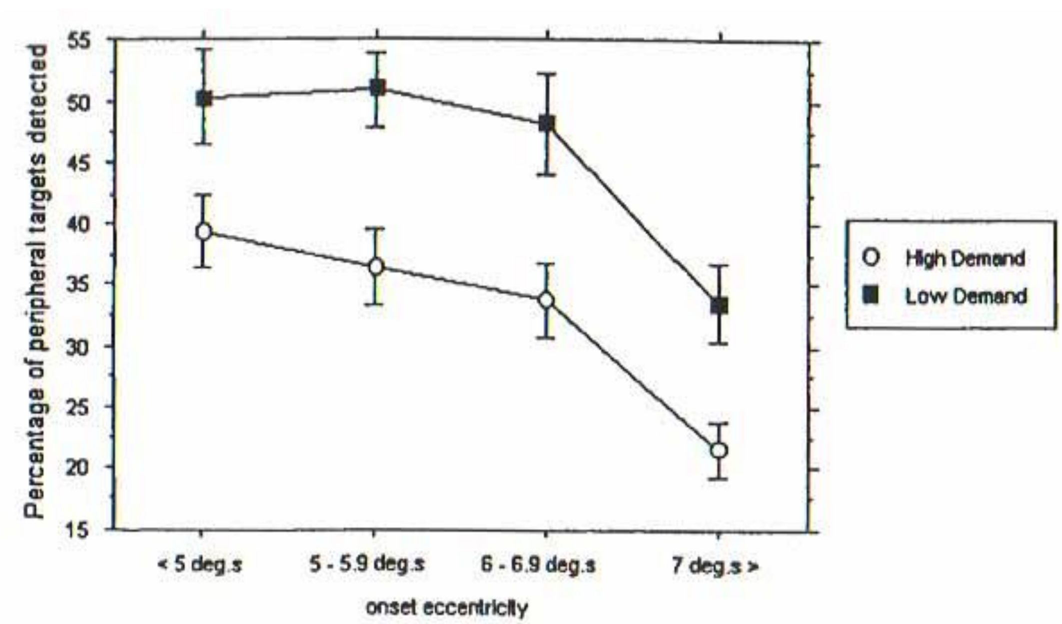


*Figure 12***.** Peripheral target detection as a function of eccentricity and workload level. (Crundall et al., 2002)**.**

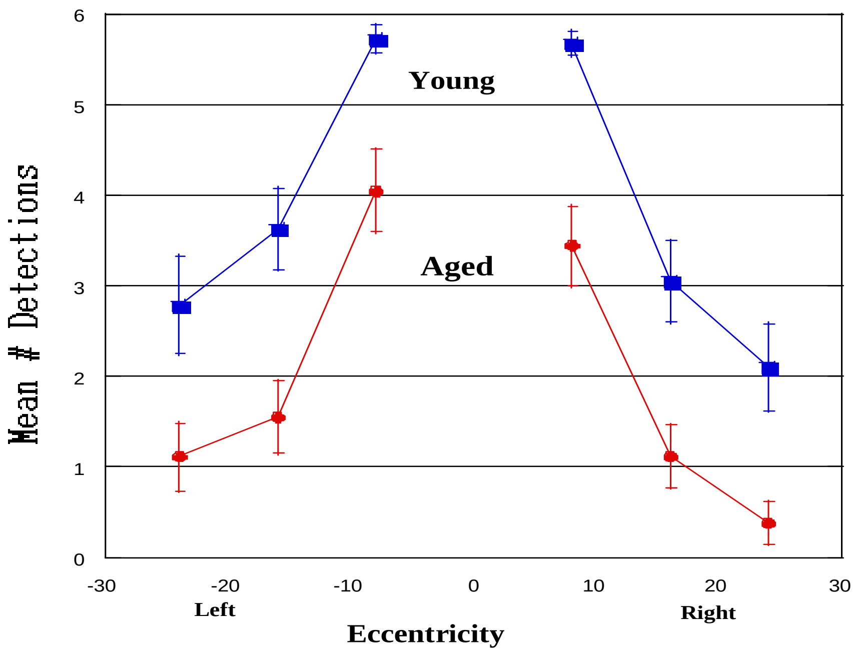


Figure 13. Peripheral target detection as a function of eccentricity and age. It should be noted that this study did not have a "loading" task per se, but a steady cognitive stress was present due to the need to follow a lead car and watch for a central stimuli presentation. A car-following task could be argued to be a very significant source of attentional loading. (Gilland, 2004)

Some studies report that drivers still look at the periphery; they just don't extract as much information from any given glance causing a "miss"(e.g. Serafin, 1994; Miura 1990; Ball & Owsley 1992). This is the basis behind the Useful Field of View and the "tunnel vision" related theories. Other studies report that drivers simply just don't look to the periphery as often as they begin to dedicate more fixations to central vision activities (e.g. Perryman & Fitten, 1996.) A good example of a workload induced central vision bias can be seen in the following (Figure 14) adapted from Harbluk & Noy (2002). Notice the change in how people look and the percentage of time spent looking at a particular area when secondary task difficulty is increased.

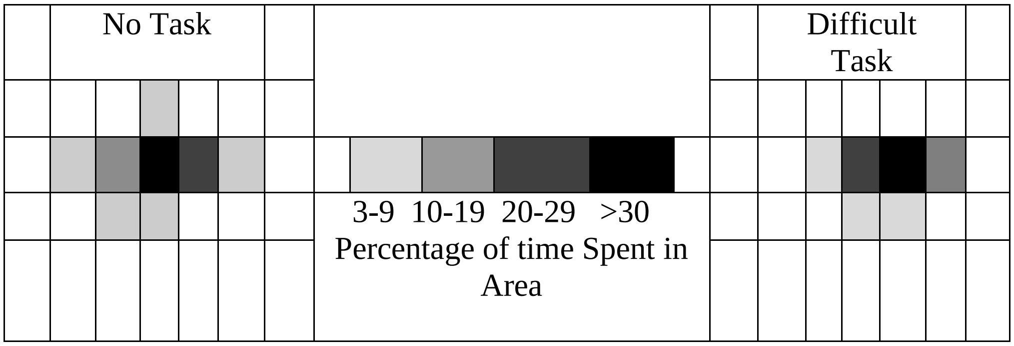


Figure 14. Adapted from Harbluk and Noy (2002). Example of fixation areas observed in one participant under a no task and difficult task condition (mental arithmetic). Note how the fixations are more localized in the difficult task situation.

Some form of secondary task or cognitively loaded environmental factor usually serves as the catalyst to manipulate the "workload" level of these studies. Further manipulation of the subjects' motivational levels and/or the ratio of difficulty/complexity between tasks can also encourage additional mental loading. Alternatively, naturally occurring environmental factors such as traffic density, weather conditions, etc. are also known to bring about changes in task performance. For all cases we are interested in, the primary task will always be driving which dramatically limits the types of secondary tasks one can choose to use. A task has to be sensitive enough to detect a change, diagnostic enough to specify a reason for that change, while at the same time maintain a low level of intrusiveness so as to not interfere with driving itself. As one can see, there is a very fine line that needs to be considered.

A "secondary task" can be anything from dialing a phone to simply repeating numerical sequences while engaged in a primary task. There are a number of secondary tasks that have been used to induce mental loading in a wide range of situations. All loading tasks will produce cognitive interference to some degree, but whether they can be applied to driving and ultimately eye-tracking is of great concern. Luckily, two methods that are practical to driving situations have already been established: mental arithmetic (Harms, 1986; 1991) and spatial imagery (Recarte & Nunes, 2000; 2003).

Nevertheless, since workload has been correlated to eye-scanning changes in other realms, one should logically expect something to occur in the driving domain as well. After all, a technique is a technique. There really is no reason to believe that since a task was difficult in one situation that it will all of a sudden become easy in the next. This logic will be evaluated by the following studies using secondary measures *while* driving to elicit eye-scanning differences.

1.8.1 Loading Tasks

It is now time to shift gears slightly and look directly at several types of *workload-inducing* loading tasks that are believed to be the most attuned with driving. Several of these have already proven to influence eye-scanning behaviors, while others will be assumed so by logic. Several papers in particular were chosen since they deal specifically with loading tasks in driving situations. The first two of these studies will involve eye-scanning pattern changes due to spatial imagery secondary tasks (spatial attention oriented). The final studies will report eye-scanning changes due to a verbal attention oriented tasks.

1.8.1.1 Spatial Imagery Tasks

The first paper is one written by Recarte & Nunes (2000). This experiment entailed twelve subjects performing two types of tasks while driving large and small scale highway roads for approximately eighty minutes. The tasks required either "verbal repeats" or "visual-spatial manipulations". In the verbal repeat task, the subject was asked to rehearse and verbally repeat a set of words beginning with a certain letter specified by the experimenter. The visual-spatial task required the subject to mentally rotate a letter in order to answer the following questions: (a) Are the letters unchanged when rotated horizontally? (b) When rotated vertically? or (c) Are the letters open or closed (e.g. 6, 9, 8 or 0 = closed, whereas 3, 2, 1 or 5 = open)?

Pupillary dilations were used to confirm that the cognitive load was similar for each task and was indeed found to be similar or both. Pupil dilations proved significant when comparing task loading verses no task, but not between tasks or road types. This indicates that similar resource expenditures were being used by each. They observed that when compared to normal driving (no task), fixations were significantly longer during the spatial imagery task than in the verbal task. They also noted smaller saccadic size and a more pronounced decrease in peripheral glances (to mirrors, speedometer, etc) with the spatial imagery task. Another curious finding was a tendency to "freeze" one's fixations during a spatial-imagery task. This suggests a temporary interruption in visual resource is needed to help process the mental variable. Ultimately, they concluded that spatial-imagery produced the greatest degree of effect and that these effects were seen equally across all manipulations suggesting that it is a reliable effect. They suggest that the visual-spatial loading task is most distinguished because it relies on the same "pool" of attention and therefore produces a bottleneck for that type of attention due to the requirement to "time share" with vision needed for driving. Verbal repeat on the other hand, while may share some attention with the visual "pool", is not exclusive to this pool so therefore does not need to time-share as much leaving more attention to be dedicated solely to external visual environment monitoring.

Later in 2003, Recarte and Nunes produced another study to help verify those results observed in 2000. In this study they now used eight variations of mentally engaging tasks to see how they impacted eye-scanning behavior. In this study, each of the twelve subjects engaged in either verbal learning (concrete, abstract), verbal production (concrete, abstract), mental calculation (phone, experimenter), or auto-recall conversation (phone, experimenter). At the same time as these were going on, they had to detect brief visual targets (LEDS). Verbal learning involved actively listening to a recorded message and memorizing for use at a later time. Verbal production involved repeating from memory what they had heard in the learning phase. Concrete refers to physical descriptions, whereas abstract was "meaning" retention. For the mental calculation task, the subject was asked to convert the currency of Pesos into Euros after being told the amount over the phone or by the experimenter. The recall conversation task required the subject to give detailed accounts of events that had happened in the past.

What they found in this study was that the most cognitively involved tasks increased pupil dilations and decreased peripheral inspections (e.g. speedometer, mirror, etc.) similar to that observed in earlier Recarte & Nunes (2000) paper. None of the learning tasks seemed to influence eye movements, however the more cognitively intense tasks (as rated by subjects) such as “peso conversions” did note a significant change in scanning behavior which was also correlated to the detection rate. Of the six that produced an effect, the phone recall condition (f=24.13) was the greatest followed closely by the live peso conversion (f=22.55) then phone peso conversion (f=20.98). Again, the authors conclude that visual scanning is heavily dependent on task loading and that tasks relying on similar "pools" of attention will tend to produce the greatest amount of negative interference with each other.

1.8.1.2 Verbal Loading Tasks

Another common loading task is the verbal loading task. Two verbal loading tasks will be reviewed: (a) Mental Arithmetic and (b) N-Back task. Of the two, the mental arithmetic has been proven effective in driving situations; however neither task has been explored for its effect on eye-movements. These were chosen because both are the easily applied tasks that will work in dynamic settings.

1.8.1.2.1 Mental Arithmetic

Mental Arithmetic is a good example of a "verbal" attentional loading task. However, it should be noted that this has not been studied that extensively *while* eye-tracking. The only paper that explicitly measures this exact interaction is a study dealing with NASA pilots. Because piloting is logically similar in nature to driving, (still need to look at gauges, and external environment when landing etc), this data should be applicable to the driving field as well. Tole et al., (1983) conducted this study in the effort to determine how gauge navigation varied by experience in pilots. The loading task was a mental arithmetic where the pilots simply stated whether the previous number presented was either larger or smaller than the current. For example, when given the sequence (6,3) they would say "plus", but when given the sequence (3,6) they would say "minus" and so forth. Before they even analyzed loading effects, they immediately noted that experience seemed to be a major factor highly predefining where fixations were going to be directed. They noted many more fixations to informationally "less important" gauges (i.e. gauges not essential to maintain *immediate* flight) in the novices than the experienced pilots. However, even though the subjects tended to look at different gauges, they still exhibited similar fixation patterns when involved with a workload task. One important observation was that as mental loading increased, all pilots begin to only look at what they *felt* was the *most* relevant information. As loading increased further, they began to progressively dedicated attention solely to these “important” areas. Loading was characterized by an increase in fixation duration, decrease in saccadic amplitude, and very increased tendency to "stare".

One strength of the Tole et al., (1983) paper is that they defined a person's fixation pattern by its *entropy*. This is pretty much the same methodology used in this study except with slight variations in the math due to different state spaces being defined. It is important to note that when using the entropy metric; they *were able to observe significant entropy effects*. This gives validity to the attempt to apply this metric to a driving situation. More detail about the *entropy* metric will be discussed in future sections.

1.8.1.2.2 N-Back Task

Another form of verbal loading found in the literature is the N-Back task. The "N" part of the task refers to the number of digits or stimuli "back" in the previous sequence one must remember to make a comparison to the present stimuli. While this task has not been directly correlated with eye movement influence, there is no reason to believe it will not extrapolate into this field since it has shown loading effects in many other studies (Perlstein et al., 2003; Carlson et al., 1998; Callicott et al., 1999; Braver et al., 2001). The main strength of the "N-back" task is that it is a *proven technique for representing* *multiple levels of loading.* The N-Back concept is simple… "Does the current stimuli match a stimuli given N-instances back in time." Stimuli are presented in a continuous serial fashion and difficulty is modified by the length of the “memory string” the subject is required to keep in memory. In this paradigm, the subject simply responds "yes” or “no" if the current stimulus is the same as the stimulus presented N-instances back. The typical "yes" rate from the literature is balanced to 33% of the total responses (Perlstein et al., 2003; Carlson et al., 1998; Callicott et al., 1999; Braver et al., 2001) and usually has 90%+ accuracy.

For an example of this task, let’s use the sequence: **2,4,2,2**. A “1”-Back task would require the subject to compare the current number (last one presented) with the one previously given which in this case would be “(2 vs. 2) = same or yes”. A “2”-Back task asks the subject to compare the current number with the one presented 2 instances prior which would elicit the response “(2 vs. 4) = different or no”.

Most of the studies reviewed reported three common correlations when using N-Back tasks: (a) *Difficulty*, (b) *Reaction time*, and (c) *Error rate* all systematically increase as the N-back size increases. *Also, the largest size difference between of any of these three trends seems to most often occur between the 1-back and 2-back conditions* (Verhaeghen & Chandramallika, 2005; Carlson et al., 1998; Perlstein et al., 2003; Smith & Jonides, 1997). After considering this last summary finding, the current study was designed to use only the 1-back and 2-back conditions. The 1-back task represents a "low" cognitive loading condition and the 2-back task represents the "higher" cognitive loading condition.

1.9 Additional Sources of Information

For an alternative overview of what was just reviewed in this chapter (sections 1.0 through 1.8), refer to the following review papers: Crundall et al., (1998), Green (2002) and Serafin (1996). Of the three, Green is by far the most comprehensive and also the only one that is a true "review" paper. Although not classified as actual review papers, the Crundall and Serafin papers are also very good since they both do an *exceptional* job summing up previous studies before discussing the results of their own. Crundal reviews mainly experience-related phenomena, while Serafin should be referred to for summaries of straight and curved road driving.

1.10 Conclusions from the Review of the Eye Tracking Literature

There is no organizing theory or conceptual framework guiding current driving research. All analyses to date seem to rely on mostly descriptive measures such as frequencies or direct observations. Unfortunately these are not really comparable across studies, as they provide no real standard from which to compare. Every situation can be described in great detail, *but how can one compare across studies when each has different circumstances?* In order to bridge this gap, more concrete metrics are needed. Information theory will provide this bridge in the form of an entropy metric that can be used to assess *quantitative* estimates on driving behavior under workload. Quantitative metrics will allow studies to be compared by scaled metrics that can be exactly defined for each situation. Different situations will then have a common basis for comparison.

Even though information theory was ultimately determined as the metric of choice for exploring our narrow focused objectives, there is another method that could have been utilized: Spectral Analysis. In order to remain thorough, both of these concepts will be discussed in the next few sections.

2.0 Quantitative Metrics of Driver Eye Movement Behavior

2.1 Spectral Analysis

Spectral analysis could be used to break down scan patterns into component frequencies. This methodology would be similar to the decomposition of sound into its component frequencies, only substituting different scan patterns as the frequencies. Take out high frequencies (eye scan patterns least used) and one will be left with the low frequencies (patterns most often used) and vice versa. This would be good for looking for frequency shifts due to workload changes but only across relatively long periods as it requires vast amounts of data. Technically this approach could be used and what would essentially be observed is a few main sequence scanning patterns, but not much beyond that. One would then have to tease out just what exactly it all means which is much more complex than simply rating all movements on a single base scale. Nevertheless, spectral analysis could still prove useful if it is decided to look at relative component frequency shifts in response to a particular workload level or situation. For example, one case it could prove useful would be to identify how changes in induced workload will impact the relative frequencies of glance patterns towards peripheral targets verses center of the road targets.

In any event, what this study is looking to accomplish is merely the validation of global eye-scanning behavioral changes over time and not an evaluation of the individual visual components at any particular point in the paradigm. A more detailed spectral analysis will may be done at later date if enough data will be available, but not for the present study. What was ultimately chosen as the desired metric was another method taken from the information theory realm: entropy. The entropy metric was chosen because unlike spectral analysis, entropy produces a simple range of values that can be easily scaled. It also can do so with only a fraction of the amount of data that would be required for a spectral analysis. The entropy approach will be much simpler as the results will be computed into a single value for each situation verses the hundreds of component frequencies inherent in a spectral analysis. This information theory based scale will have an absolute minimum/maximum and the calculated value will always fall within this range.

2.2 Entropy: A Metric for Quantifying System Complexity

The *theory of information* (Shannon, 1948) provides a conceptual framework for quantifying the complexity of *any* control system – including the human saccadic eye movement system. Borrowing from the field of thermodynamics, this complexity metric has been termed *entropy*. The entropy of a system is directly proportional to the amount of information necessary to describe the behavior of that system. That is, the more information needed to specify the system (and the state-space it can occupy) – the more complex the system (Attneave, 1959; Edwards, 1969).

Since the entropy of a system is based upon units of information, any attempt to specify this construct must begin with a quantitative metric of information. According to Shannon's classic theory, the *amount of information* needed to describe a system's complexity can be quantified by calculating the number of yes/no questions needed – on average – to ascertain the state of that system at any given point in time. In a simple random system characterized by a finite state-space (whose alternatives are equiprobable), Shannon (1948) has shown that the number of yes/no questions needed to determine the current state can be represented by equation 1:

$$\mathbf{H = \#yes/no\ questions = log_2(n)}$$

(Equation 1)

where:
n = size of the equiprobable state-space
H = information (expressed in units known as *bits*)

Let's work through some concrete examples to see if we can obtain a more intuitive appreciation of this definition of the amount of information needed to define a system:

2.2.1 Example 1:

Consider the case of a very simple "seeing robot". This robot can position its "eyes" (a pair of video cameras) to one of only two states: either to the left or to the right (see Figure 15). How complex is this robot's eye movement system? According to information theory, its complexity is proportional to the number of yes/no questions (i.e., bits of information) needed to eliminate our uncertainty about the current state of the system. Since our robot's eyes can be in one of two states (n = 2), Eq. 1 specifies its informational content as being: $H = \log_2 (2) = 1$ bit. That is, by asking only a single yes/no question (e.g., "Is the robot looking to the left?"), we can completely circumscribe the system's status to one of all possible states. This analysis can be represented in the form of a decision tree as follows:

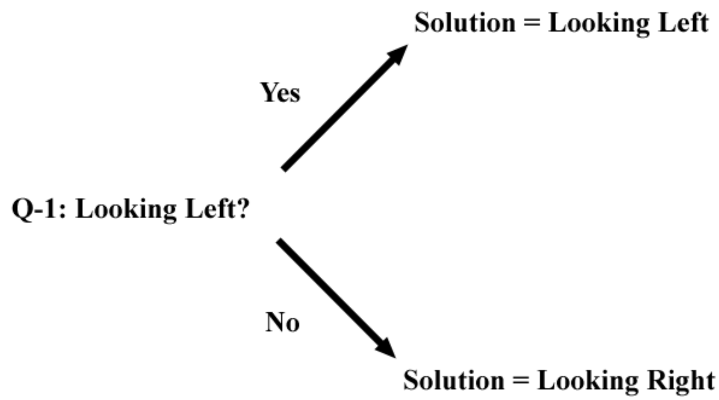


Figure 15. Decision-tree depicting the entropy of a 2-state system.

<u>*2.2.2 Example 2*</u>.

Now let's look at a slightly more sophisticated robot. This robot can look at *four* different places (see Figure 16) in the environment since it can move its eyes either up or down as well as to either the left or right, yielding the following equiprobable state-space:

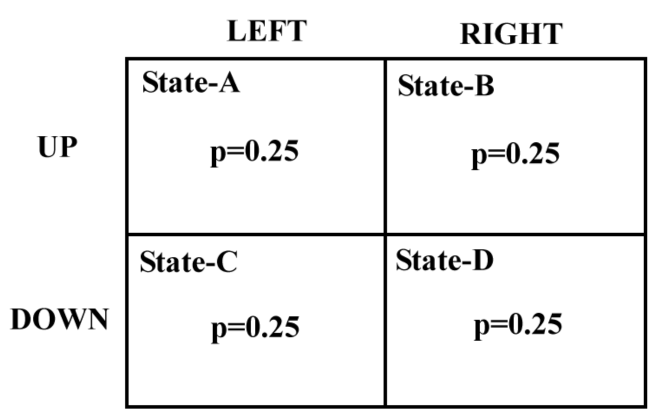


Figure 16. Equiprobable state-space of size = 4.

Assuming that each state is equiprobable, information theory holds that the current state of the system can be ascertained by asking $\log_2(n) = \log_2(4) = 2$ yes/no questions within the context of an optimal search strategy. Hence, the information needed to describe this system (and, hence, its complexity) can be quantified as 2 bits. Again, the logic of this approach can be depicted via the decision tree in Figure 17.

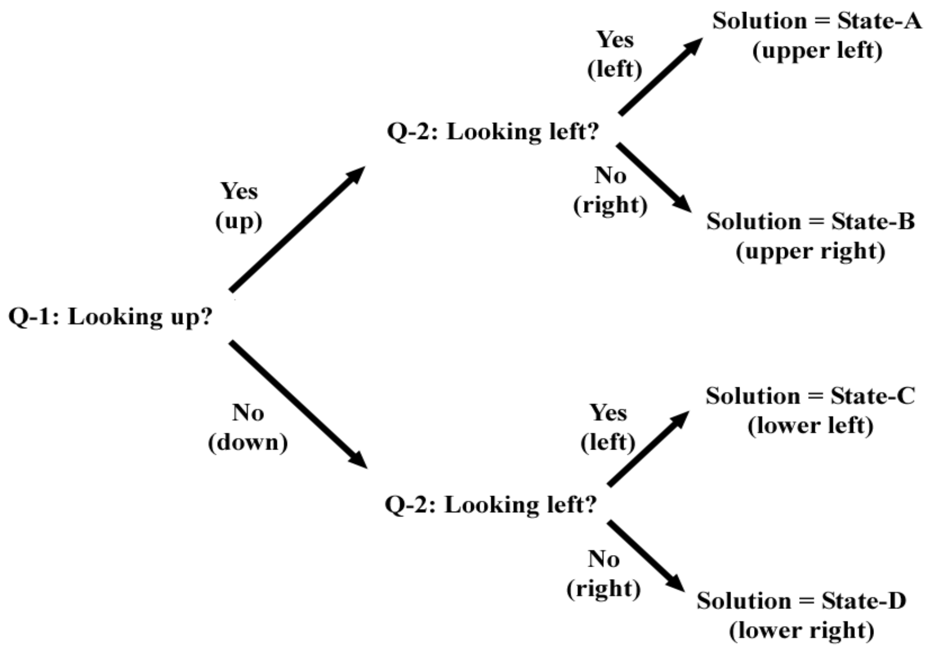


Figure 17. Decision-tree depicting the information theory approach for calculating the entropy of the equiprobable state-space from Example 2.

The examples of calculating the entropy of a system to this point have considered random systems in which each member of the state-space is equiprobable. Yet, many control systems of interest (such as the human eye movement system) are characterized by state-spaces in which the individual elements are not equiprobable.

The next example demonstrates how Shannon's (1948) concept of entropy can be generalized to such systems:

2.2.3 Example 3.

Suppose, as in Example 2 above, that the robot can move its eyes to one of four positions *but* that the probability of occupying each of these states varies according to the following probability matrix in Figure 18.

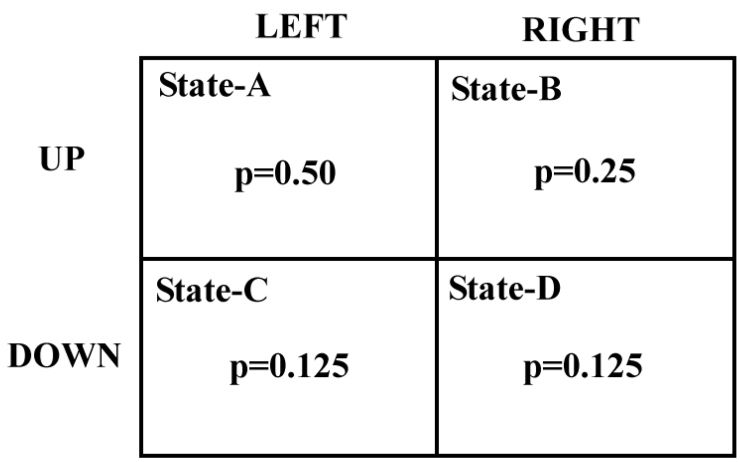


Figure 18. *Non-equiprobable* state-space of size = 4.

It can be shown that the entropy of a system with such a state-space can be ascertained according to the logic represented in the following decision tree seen in Figure 19.

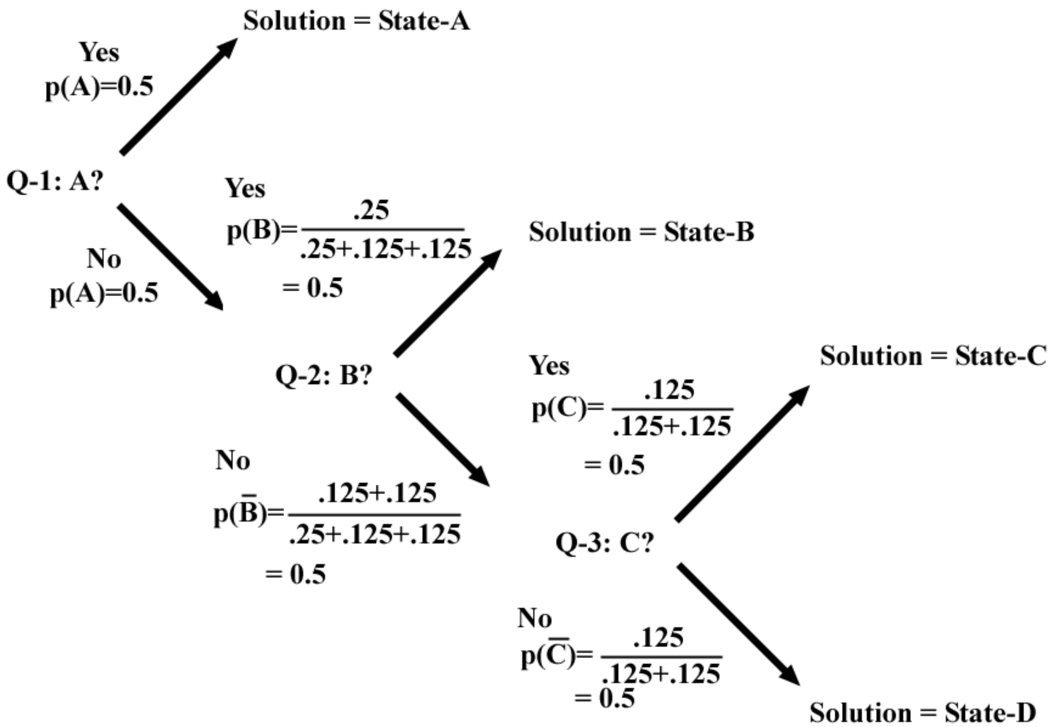


Figure 19. Decision-tree depicting the information theory approach for calculating the entropy of the non-equiprobable state-space from Example 3.

Inspection of this decision tree reveals that sometimes the optimal search process through this non-equiprobable state-space will "get lucky" and reach solution after making a single yes/no probe. However, in the worst case scenario, the optimal search algorithm requires three yes/no question probes in order to reduce uncertainty about the state of the system to zero. Shannon (1948) has shown that the entropy metric for such a system is specified by the state-space probability weighted average of all possible yes/no questions needed to specify the decision tree (see Figure 19). The computational algorithm for achieving this weighted average is summarized in Table 1.

State Solution	(nq_i) Number of yes/no questions	(p_i) state-space probability (weight)	(nq_i) (p_i)
A	1	0.5	0.5
B	2	0.25	0.5
C	3	0.125	0.375
D	3	0.125	0.375
		$H = \Sigma\,(nq_i)\,(p_i) =$	1.75 bits

Table 1. Computation of weighted average of all possible yes/no questions required to exhaustively search the non-equiprobable state-space from Example 3.

Shannon (1948) has shown that the entropy (**H**) of systems with non-equiprobable state-spaces can be computed via his *classic entropy equation*:

$$\underline{\textbf{Entropy}} = \mathbf{H} = \Sigma\, \mathbf{p_i} \log_2 (1/\mathbf{p_i})$$

(Equation 2)

Applying this equation to the computation of entropy for example 3 can be shown to yield an identical estimate of system complexity to the weighted-average approach described above (see Table 2).

State_i	(p_i)	$\log_2 (p_i)$	$(p_i)\log_2 (p_i)$
A	0.5	-1.0	-0.5
B	0.25	-2.0	-0.5
C	0.125	-3.0	-0.375
D	0.125	-3.0	-0.375
		$H = -\Sigma\,(p_i)\log_2 (p_i) =$	1.75 bits

Table 2. Computation of the entropy of the non-equiprobable state-space from Example 3 using Shannon's classic formula (Eq. 2).

Again, the entropy for this system computed using Eq. 2 is found to be equivalent to 1.75 bits of information. For the purposes of thoroughness, it can be shown that the simplified entropy equation (Eq. 1) can be derived directly from Shannon's general formula (Eq. 2) as follows:

$$
\begin{aligned}
H &= - \Sigma\ p_i \log_2 (p_i) && \text{[Eq. 2]} \\
&= - \Sigma\ 1/n \log_2 (1/n) \\
&= n\ -(1/n \log_2 (1/n)) \\
&= - \log_2 (1/n) \\
H &= \log_2 (N) && \text{[Eq. 1, Q.E.D.]}
\end{aligned}
$$

Edwards (1969) summarizes several important properties of the general entropy equation:

1. A value of H can be computed for any finite state-space.
2. The minimum value of H is zero. This can occur only in the case of a 1-state system (i.e., 0% uncertainty or 100% certainty).
3. The value of H is maximal (for a given state-space size) when all possible system states are equiprobable. This property is demonstrated in Figure 20 and will be elaborated in a subsequent section.

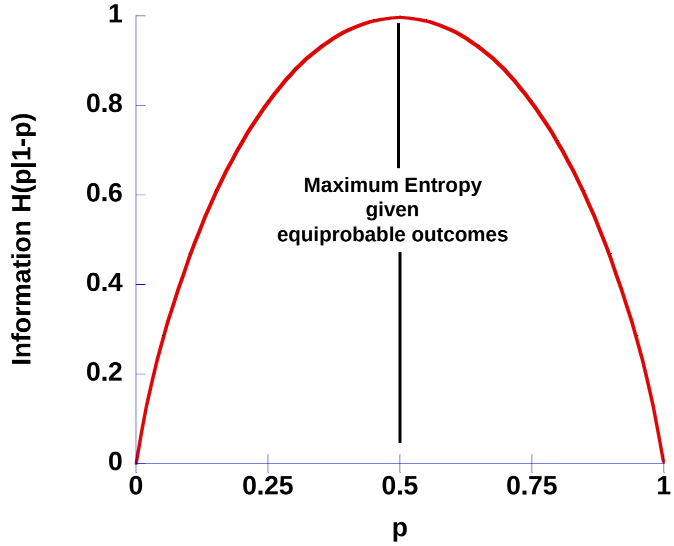


Figure 20. Entropy of a 2-state system for all possible combinations of probabilities (where p(q) = p and p(not q) = 1 – p(q)). Note that maximum entropy occurs when p(q) = p(not q) = 0.5 (i.e., equiprobability).

Finally, it should also be noted that conditions will always be different between systems; therefore it is imperative to establish a metric common to all. Entropy allows this via a simple normalization process. By dividing the empirically measured estimate of entropy ($H_{observed}$) by the theoretical maximum value of entropy (H_{max}) for the state-space being modeled (see Eq. 3). This *relative entropy* value allows one to compare results across groups and situations.

$$\text{Relative Entropy} = H_{observed} / H_{max}$$

(Equation 3)

2.3 Development of an Information theory Metric for Quantifying the Complexity of Driver Eye Movement Behavior

If the visual world is decomposed into a multi-element spatial grid (either arbitrarily or on the basis of theoretical considerations), then the global complexity of eye movement behavior can be quantified by using Shannon's (1948) entropy metric (Eq. 2). However, since the *a priori* probabilities of gaze to each region of this visual space are unknown, the probability matrix for this state-space must be estimated on the basis of empirical experiment. Tole et al., (1983) has demonstrated the feasibility of such an approach for quantitatively capturing global variations in pilot eye movement behavior as a function of task load. In the sections that follow, the logic for applying this approach to quantifying global complexity of driver eye movements – which will be called *visual entropy* – will be developed. First, the rationale for establishing the spatial boundaries of the driver's visual state-space will be developed. Next, the computation of the entropy metric for independent events will be generalized to the case of serially-related events (such as sequential eye movements) through the introduction of *Markov chain* models. Finally, we will propose a methodology for measuring visual entropy while driving; and, evaluating its potential sensitivity to variations in driver task load conditions.

2.3.1 Visual Entropy

"Visual Entropy" is the tendency of *fixation patterns* to behave randomly. A fixation pattern is defined by the *order of fixation transitions*. One can think of it as a case of "connect the dots" only without numbers. Any fixation links to another "dot" and each possible pathway is a potential transition. When entropy is high, the observers will look at *everything* an equal number of times and will *transition* between all possible combinations of stimuli with a near equal frequency. As a person begins to start focusing his or her attention over a narrower range of potential fixation points, the entropy will also decrease. This is because frequency of use for the other potential transitions has decreased. In a decreased Visual Entropy state, systematic fixation patterns will begin to emerge and previously random scanning patterns become more ordered. This is a *very* new approach to quantifying eye movement behavior with very little research supporting its use. This current study will add to the literature by applying this concept to a driving situation. This is important not only because it has never been done before, but also because it could provide a valid *quantitative* means for evaluation of a person's real-world driving behavior.

So why apply entropy as a reference measure for vision? First off, entropy doesn't require extensive theory in order to be applied. Entropy relies simply on *probabilities* of occurrence. This means that it can be applied to almost any situation as probability matrices can theoretically be taken from *any* situation. By definition, probabilities are quantitative in nature, meaning all results using this metric will yield highly interpretable data that can easily be scaled and modeled.

Entropy of eye movements was first explored in a paper by Tole et al., (1983) describing the looking behavior of eleven pilots. As explained before, this study entailed having the pilot simply scan the flight instrument panel while performing certain secondary loading tasks. It was observed that as workload went up, dwell times increased, fixations decreased, and that the dwell times were re-proportioned to the more "important" instruments. Important in this sense refers to the gauges absolutely needed to maintain *current* flight. This means that gauges like engine temp, water temp, etc. which usually may only get sampled minimally to begin with, began to be completely displaced by more looks to air speed, attitude, etc. As the workload increased, entropy decreased as they started to restrict the range of their eye movement behavior. In a high entropy state, one would expect less focused visual scanning with no predisposition to fixate at any one instrument more than another.

It was expected that the driving situation results of this study should be similar to that observed in the Tole et al., (1983). Subjects should start to decreasingly sample “less important” peripheral targets and begin focusing more on the informationally important regions of the road (roadway in front of driver) as workload increases. This should also be increasingly noticeable in the older populations since they are thought to have a less flexible working memory system resulting in more difficulty recruiting attention for a workload task. Ultimately any secondary task used should result in restricted sampling frequencies of the external visual environment which in turn will correlate to a decreased entropy value.

Visual Entropy in this study was measured in much the same fashion as outlined in the Tole et al., (1983) paper, but focused on aspects of the environment other than the instrument panel. In the current study, all eye movements were measured for a given test run. The eye movement data was then processed and separated into fixations that could be overlaid on a grid-work pattern representing the visual scene similar to that seen in Figure 21. In this figure one will notice a grid composed of 7 distinct regions of interest. One will also note that the boxes are skewed in size (due to projective geometry). The border between the “near” (regions 1, 2 and 3) and the “far” (regions 4, 5 and 6) visual world was selected in accordance with Donges’ (1978) influential theory of automobile steering behavior. In this theory, visual guidance necessary for optimal steering depends upon two sources: (a) optical flow from near range peripheral vision which operates in a closed-loop fashion to null out lane position errors; and (b) far range central vision that is sampled in open-loop fashion to detect hazards and unexpected variations in roadway geometry.

Placement of this near/far boundary will be at the position delineating a 2 sec advance preview time down the road (scaled to the driving speed used in this investigation). Recent research has shown that driver lane keeping performance improves as preview time increases and reaches asymptotic levels at between 1.8 and 2.0 sec (European Community, 1999). Thus, a preview time of 2 sec appears to represent the transition point between Donges' (1978) near versus far visual processes. The dividing lines delineating the central visual regions from the peripheral areas on the left and right are based upon the half-width of the useful field of view assessed while driving (see Schieber & Gilland, 2005). The final region of the driver's visual world, defined as number "7" in Figure 21, refers to the viewing area on interior vehicle surfaces like the instrument cluster, radio, dash board, panel, etc.

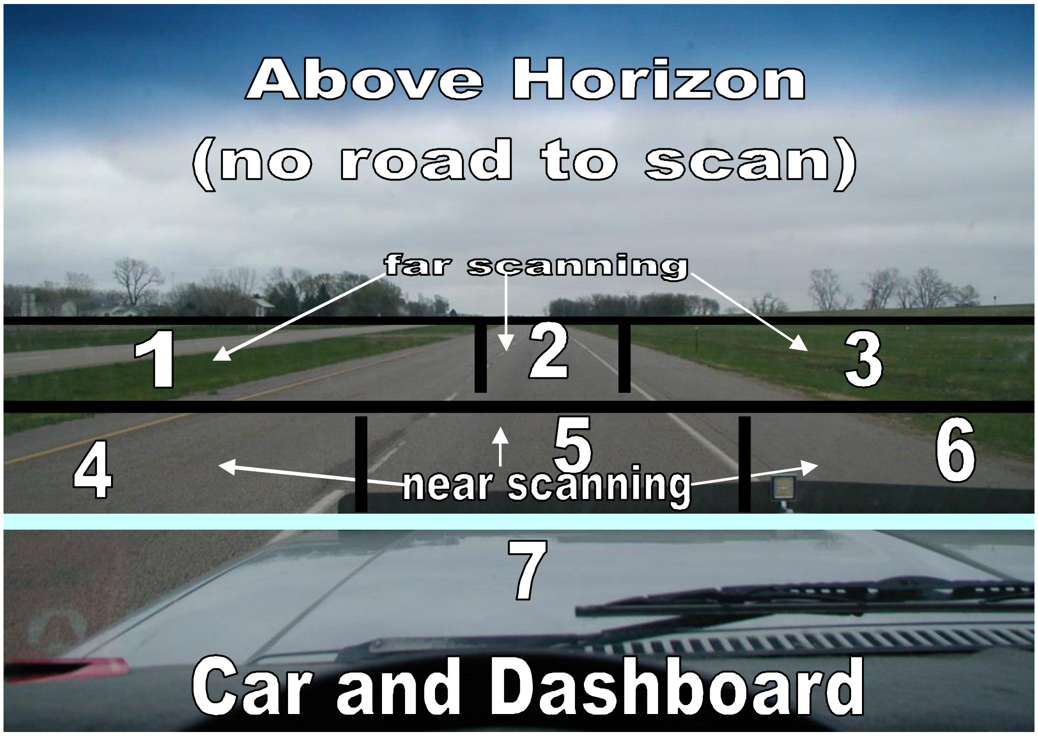


Figure 21. Example of partitions developed to represent possible transition state spaces. The first three spaces (1,2,3) represent FAR scanning transitions. The second row (#4,5,6) represent near transition spaces. Box #7 represents all glances to the inside of the car at instruments, etc. The cutoff point between near and far is determined by the distance needed to allow a 2 second preview time at a given speed since this is the distance needed for lane maintenance functions as discussed earlier.

From this "grid-wise" decomposition of the visual scene, we can classify the general area where the subject was looking during any given fixation. The grid-work provided a means of quantifying visual sequences to be evaluated by the entropy measure. Any fixation that moved from one grid box to the next would be reasoned to be due to scanning and recorded in series. The different combinations of “box” transitions are then added up and the relative proportion of each combination extracted for an Entropy calculation.

In the Tole et al., (1983) paper, it was found that fixation sequences of length = 2 yielded the most consistent results although any number could theoretically be used. The idea here is that data in such sequential "chains" could somehow be related, decreasing entropy because they are no longer statistically independent. We explored eye movement sequences with lengths of 2 and 3 using a technique called a *Markov chain model.* This will be explained in more detail in the next section of this paper. We expected results similar to that found in the Tole (1983) paper where they observed decreased entropy as workload increased and subjects restricted their range of gaze.

2.3.2 Markov Chain Modeling

Shannon's classic equation (Eq. 2) specifies the entropy for a state-space of statistically *independent* events (A more appropriate modification of the entropy equation will be presented below as Eq. 4). However, the eye movements used to scan the driving environment are most certainly not statistically independent from one another. This sequence of driver eye gazes is a stochastic process that can be modeled as a matrix of interrelated probabilities. Since the matrix describing the likelihood of transitioning from any one position (i.e., state) to every other possible position (in state-space) will be statistically dependent upon previous events in the series of eye movements, this stochastic series is called a *Markov process* (Karlin and Taylor, 1975; Tuckwell, 1988). It is presumed that the sequential relatedness of any given eye movement transition from one state to another extends over a finite range of serial events; and, thus, can be held to be *ergodic* in nature (i.e. all states are assumed to occur equally, be representative of the whole, and be presented at least once). The underlying mathematical structure of such an ergodic Markov process can be estimated given a representative sample of the system's sequential behavior (Edwards, 1969; p. 46).

In order to calculate the entropy of an ergodic eye movement system, one needs to estimate the correlation matrix that maps the probability of transitioning to each element of the state-space as a function of the prior state of the system (i.e., a network of sequential conditional probabilities). The previous state can refer to the immediate past gaze position (i.e., a 1st-order Markov process) or to all possible prior gaze sequences of length-n (i.e., nth-order Markov processes). Once this matrix of probabilities has been empirically estimated, Van der Lubbe (1997; p. 86; eq. 3.4) has shown that the entropy of the eye movement system's behavioral transitions can be calculated using the following formula:

$$\mathbf{H(Y|X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(Y_{ij}\,|X_{i})\log_2\, p(Y_{ij}\,|X_{i})\; p(X_{i})}$$

Equation 4. (X = the previous state space, Y = the current state space)

Markov chain models are essentially just an extension of the entropy Eq. #2 mentioned earlier. The difference is that now instead of just finding the probability of a single state to occur given a group of occurrences, you get a compound probability for multiple states to progressively occur over time, hence the "chain" part of the name. For every additional number you add to sequence length, you increase the order of magnitude needed to assess the probability of that chain. For example a 2-sequence chain predicts the probability of the second state given knowledge of only *one* previous state so it is termed a 1st-order chain. A three-sequence chain determines the probability based on *two* previous states, so it is termed a 2nd order chain, and so forth. One essentially gets a compound result as the first state's probability is used to determine the probability of a two-sequence set to occur and that in turn will be used to determine probability a three-sequence set, etc. The individual probabilities for each of these "state space sets" (1st order, 2nd order, etc) will then sum together to estimate an overall entropy value. In the Tole et al., (1983) paper they found that sequence lengths of two yielded the best results when studying eye movements. The current study expanded on the Tole paper and extended visual scanning out to three sequential state transitions (i.e., 2nd-order Markov models).

Using a chain length beyond three sequential transitions is possible, but not practical since the size of the state-space combinations rises exponentially with the size of the sequence. This sets a rational limit on research since resources have to be balanced with the reward. If other studies find only a negligible power increase in with sequence size, yet one now requires 8 times as much data to do the same analysis, there has to be a point where you just accept what you have. Since the Tole et al. (1983) paper found only marginal improvement beyond strings of 2, this study has already went above and beyond by including 2nd order entropy (strings of 3). An example of how a simple 7-state probability maps to a 3 set sequence can be seen in Table 7 in the next section. Notice how spread out this data has becomes. If even a single digit of sequence length is added, the data becomes dramatically spread out. As the number of potential possibilities increases, *much* more additional data will need to be sampled in order ensure accurate sample size and representation.

2.3.2.1 Demonstration of a Markov Model Calculation

What follows, is a demonstration of how to calculate a 1st order (2-sequence) Markov entropy calculation. The first step in any case is to pre-determine the “transition state” matrix. In these examples let’s assume 7 states, which therefore will create a matrix of 49 potential states as seen in Table 3. This matrix will eventually hold empirically derived values for actual field calculations, but in this example we will use equi-probability for all states in order to calculate H_{max} (Reminder: The entropy of a state-space is maximized (H_{max}) when all states are equally probable.)

In Table 3, the current state is represented by horizontal row and prior states by the vertical column. The way to read the chart is that when given prior state (X_i) has ALREADY occurred; the probability that it will be followed with the new state (Y_{ij}/X_i) is…

Prior State	**Present State $p(Y_{ij}\|X_i)$**						
(X_i)	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	------	0.166	0.166	0.166	0.166	0.166	0.166
2	0.166	------	0.166	0.166	0.166	0.166	0.166
3	0.166	0.166	------	0.166	0.166	0.166	0.166
4	0.166	0.166	0.166	------	0.166	0.166	0.166
5	0.166	0.166	0.166	0.166	------	0.166	0.166
6	0.166	0.166	0.166	0.166	0.166	------	0.166
7	0.166	0.166	0.166	0.166	0.166	0.166	------

Table 3. This figure is demonstrating the probability of occurrence of any given set (Y_{ij}/X_i). Notice that all states show equal probability because it is representing the calculation for H_{max}.

Once the state transition probability matrix has been established, we populate it with empirically estimated values. In this case we are calculating the $H_{max,}$ so equi-probability was used representing that no serial dependence is observable in any case. These probability values from the matrix are then placed into the “inner” portion of

Equation #4 represented by Eq. #4.1 (below) and the resulting summation placed in column b. of Table 4 for each prior state (X_i) as will be demonstrated in the paragraphs which follow.

$$-\sum_{j=1}^{m} p(Y_{ij} | X_i) \log_2 p(Y_{ij} | X_i) \qquad \textit{(Equation 4.1)}$$

Example calculation for sub-entropy with prior state X_1 using summation Eq. 4.1

$$\Sigma_1 = -(((.166)\log_2 (.166)) + ((.166) \log_2 (.166)) + ((.166) \log_2 (.166)) + ((.166) \log_2 (.166)) + ((.166) \log_2 (.166)) + ((.166) \log_2 (.166)))$$

$$\Sigma_1 = -((-.4300) + (-.4300) + (-.4300) + (-.4300) + (-.4300) + (-.4300))$$

$$\Sigma_1 = 2.58$$

a.	b.	c.	d.
X_i	$-\sum_{j=1}^{m} p(Y_{ij} \| X_i) \log_2 p(Y_{ij} \| X_i)$	$p(X_i)$	product
1	2.58	× 0.143 =	0.369
2	2.58	0.143	0.369
3	2.58	0.143	0.369
4	2.58	0.143	0.369
5	2.58	0.143	0.369
6	2.58	0.143	0.369
7	2.58	0.143	0.369

$$-\sum_{i=1}^{n}\sum_{j=1}^{m} p(Y_{ij} | X_i) \log_2 p(Y_{ij} | X_i)\, p(X_i) = 2.58 \text{ bits}$$

Table 4. Entropy derivation for H_{max}. Notice how Eq 4.1 is modified into Eq. 4 by simply adding a weighting value and another summation. The second summation is for the state spaces themselves instead of *within* each state space like the first summation.

Now repeat the same steps for each prior state row to get sub-entropy associated with each state. Each present state (Y_{ij}/X_i) is then weighted (multiplied) by the probability of the prior state (column c. from Table 4) to get a product (column d. from Table 4). One will remember that Eq. 4.1 was referred to as the "inner" summation (i.e., Σ_j). Now it is time to bring the whole equation into play and calculate the "outer" summation of Eq. 4 (i.e., Σ_i). This is simply the summation of the products of column b and c (found in column d of Table 4) for each prior state (X_i). This final calculation can be seen to yield 2.58 bits when all variables are equally probable. This is essentially saying that given a completely random system, we can – on average – determine what two-sequence event is currently being exhibited by asking no more than 2.58 questions (i.e., the essence of information theory). It should also be noted that this value is also the absolute highest entropy that this system can possess. *Any deviation* from pure randomness will result in a reduced “bit” quantity. Take the next example as a demonstration of the validity of this statement…

Example Calculation for entropy using non-random data

Prior State	**Present State $p(Y_{ij}\|X_i)$**						
(X_i)	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	------	0.01	0.23	0.23	0.166	0.166	0.198
2	0.23	------	0.198	0.166	0.230	0.166	0.010
3	0.23	0.23	------	0.166	0.166	0.01	0.198
4	0.01	0.23	0.23	------	0.166	0.198	0.166
5	0.01	0.23	0.23	0.166	------	0.198	0.166
6	0.89	0.01	0.01	0.040	0.001	------	0.004
7	0.01	0.89	0.01	0.040	0.001	0.004	------

Table 5. This figure is demonstrating the *simulated* probability of occurrence of a given set (Y_{ij}/X_i). Notice that all states DO NOT show equal probability.

The probability matrix depicted in Table 5 is a modified copy of the matrix for a completely random system shown previously in Table 3. The conditional probabilities ($Y_{ij}|X_i$) for prior states X_1 through X_5 represent a slight divergence from randomness, while those associated with X_6-X_7 represent a dramatic divergence. However, it should be remembered that ALL conditional probabilities are incorporated into the overall estimate of entropy in the final computation.

In a real data set (as opposed to a purely random data set generated to compute H_{max}), the proportion of the previous state occurrence weights (i.e., the X_i values) would be determined empirically by a frequency count of occurrences within the data set. In the computational table presented in Table 4, equiprobability was assumed resulting in a weighting of 16% for each state, but in this next example unequal probabilities will be used. The representation of the new weighting factors will again be seen again in column (c) in Table 6 below.

Prior State Probabilities of Occurrence

X_1= 10% X_2= 12% X_3= 13% X_4= 15%
X_5= 16% X_6= 17% X_7= 17%

Calculation of inner-summation (Eq. 4.1) applied to Prior State #1
(Mathematically the same as prior states 2, 3, 4 & 5)

$$\Sigma_1 = -(((.01) \log_2 (.01)) +((.23) \log_2 (.23))+ ((.23) \log_2 (.23))+ ((.166) \log_2 (.166))+ ((.166) \log_2 (.166))+ ((.198) \log_2 (.198)))$$

$$\Sigma_1 = - ((-.066)+ (-.487)+ (-.487)+ (-.4300)+ (-.4300)+ (-.463))$$

State #1,2,3,4,5 = Σ_1 = 2.363

Prior State #6 calculation
(Mathematically the same as prior state 7)

$$\Sigma_6 = -(((.89) \log_2 (.89)) +((.01) \log_2 (.01))+ ((.01) \log_2 (.01))+ ((.04) \log_2 (.04))+ ((.001) \log_2 (.001))+ ((.004) \log_2 (.004)))$$

$$\Sigma_6 = - ((-.149)+ (-.066)+ (-.066)+ (-.186)+ (-.0099)+ (-.0318))$$

State #6,7 = Σ_6 = .5087

a. (X_i)	b. $-\sum_{j=1}^{m} p(Y_{ij} \mid X_i) \log_2 p(Y_{ij} \mid X_i)$	c. $p(X_i)$	d. Product
1	2.363	.10	.2363
2	2.363	.12	.28356
3	2.363	.13	.30719
4	2.363	.15	.35475
5	2.363	.16	.3784
6	.5087	.17	.0864
7	.5087	.17	.0864

Entropy = 1.3765

Table 6. Entropy derivation of using non-random states.

In this second example it can be observed that even though the same state space was sampled, a decrease in entropy occurred since several of the sequences now occur more frequently than the others.

One can use the same calculations just discussed for any order (2nd, 3rd, 4th etc.) Markov Chain. The only difference is the complexity of the calculation since each "level" of entropy will also include those before it. For example, instead of just finding the probability of X occurring after Y, you now find the probability of T following an X that ALSO FOLLOWED Y which will be weighted by the relative occurrence of T's, X's and Y's. As one might guess, the probability of any given sequence consistently occurring will dramatically decline as set length increases. This is demonstrated in Table 7.

As you increase the length of the sequence you exponentially increase the number of possible solutions. In a 1st-order Markov model there are already 42 possible sequential outcomes given only 7 transition states. Bump that sequence length from 2 to 3 and you inflate the size of the state-space to 252 possible transition combinations. Another bump in the sequence length and you reach a state-space with 1512 possible combinations etc. As one can imagine, the number of possible combinations within the state-space begins to grow rather large and the size of the empirical data set that one needs to collect in order to estimate their probability structures limits our ability to use models beyond sequences of Markov length four.

Now that a general equal probable approach has been outlined lets recap and then expand on how to apply this to a real-world data set where conditions are not equi-probable. The first thing to start with is a data set. This is a string of numbers representing transitions of fixations between specified state-spaces, which in our case, are defined by the seven regions specified in Figure 23. Only transitions are counted which means that repeated fixations within one of the 7 predefined boxes would be collapsed upon itself. For example the *fixation* series (1,4,4,4,2,6,6,4,7,7,7) would be collapsed into the *transition* series 1,4,2,6,4,7. In order to keep things simple, a matrix population of only 6 transition data points will be used. Keep in mind, this would only represent about *4 seconds* of eye tracking data, but should still serve to emphasize the main points.

The data set 1,4,2,6,4,7 derived previously would be broken into the linked transition sets (1,4,2)(4,2,6)(2,6,4)(6,4,7) in a 2nd order Markov model and ~~(1,4)(4,2)(2,4)(4,2)(2,6) (6,2)(2,6)(6,4)(4,6)(6,4)(4,7)~~ (1,4)(4-2)(2-6)(6-4)(4-7) in a 1st order Markov model. These new data sets are then ordered into percentages and counts as can be seen below in Table 7.

1st Order Markov State Space	2nd Order Markov State Space					
Sequence/count	Potential Sequence/ observed count					
1,2 – 0	1,2,1 – 0	2,3,1 – 0	3,4,1 – 0	4,5,1 – 0	5,6,1 – 0	6,7,1 – 0
1,3 – 0	1,2,3 – 0	2,3,2 – 0	3,4,2 – 0	4,5,2 – 0	5,6,2 – 0	6,7,2 – 0
1,4 – 1	1,2,4 – 0	2,3,4 – 0	3,4,3 – 0	4,5,3 – 0	5,6,3 – 0	6,7,3 – 0
1,5 – 0	1,2,5 – 0	2,3,5 – 0	3,4,5 – 0	4,5,4 – 0	5,6,4 – 0	6,7,4 – 0
1,6 – 0	1,2,6 – 0	2,3,6 – 0	3,4,6 – 0	4,5,6 – 0	5,6,5 – 0	6,7,5 – 0
1,7 – 0	1,2,7 – 0	2,3,7 – 0	3,4,7 – 0	4,5,7 – 0	5,6,7 – 0	6,7,6 – 0
2,1 – 0	1,3,1 – 0	2,4,1 – 0	3,5,1 – 0	4,6,1 – 0	5,7,1 – 0	7,1,2 – 0
2,3 – 0	1,3,2 – 0	2,4,2 – 0	3,5,2 – 0	4,6,2 – 0	5,7,2 – 0	7,1,3 – 0
2,4 – 0	1,3,4 – 0	2,4,3 – 0	3,5,3 – 0	4,6,3 – 0	5,7,3 – 0	7,1,4 – 0
2,5 – 0	1,3,5 – 0	2,4,5 – 0	3,5,4 – 0	4,6,4 – 0	5,7,4 – 0	7,1,5 – 0
2,6 – 1	1,3,6 – 0	2,4,6 – 0	3,5,6 – 0	4,6,5 – 0	5,7,5 – 0	7,1,6 – 0
2,7 – 0	1,3,7 – 0	2,4,7 – 0	3,5,7 – 0	4,6,7 – 0	5,7,6 – 0	7,1,7 – 0
3,1 – 0	1,4,1 – 0	2,5,1 – 0	3,6,1 – 0	4,7,1 – 0	6,1,2 – 0	7,2,1 – 0
3,2 – 0	**1,4,2 – 1**	2,5,2 – 0	3,6,2 – 0	4,7,2 – 0	6,1,3 – 0	7,2,3 – 0
3,4 – 0	1,4,3 – 0	2,5,3 – 0	3,6,3 – 0	4,7,3 – 0	6,1,4 – 0	7,2,4 – 0
3,5 – 0	1,4,5 – 0	2,5,4 – 0	3,6,4 – 0	4,7,4 – 0	6,1,5 – 0	7,2,5 – 0
3,6 - 0	1,4,6 – 0	2,5,6 – 0	3,6,5 – 0	4,7,5 – 0	6,1,6 – 0	7,2,6 – 0
3,7 – 0	1,4,7 – 0	2,5,7 – 0	3,6,7 – 0	4,7,6 – 0	6,1,7 – 0	7,2,7 – 0
4,1 – 0	1,5,1 – 0	2,6,1 – 0	3,7,1 – 0	5,1,2 – 0	6,2,1 – 0	7,3,1 – 0
4,2 – 1	1,5,2 – 0	2,6,2 – 0	3,7,2 – 0	5,1,3 – 0	6,2,3 – 0	7,3,2 – 0
4,3 – 0	1,5,3 – 0	2,6,3 – 0	3,7,3 – 0	5,1,4 – 0	6,2,4 – 0	7,3,4 – 0
4,5 – 0	1,5,4 – 0	**2,6,4 – 1**	3,7,4 – 0	5,1,5 – 0	6,2,5 – 0	7,3,5 – 0
4,6 – 0	1,5,6 – 0	2,6,5 – 0	3,7,5 – 0	5,1,6 – 0	6,2,6 – 0	7,3,6 – 0
4,7 – 1	1,5,7 – 0	2,6,7 – 0	3,7,6 – 0	5,1,7 – 0	6,2,7 – 0	7,3,7 – 0
5,1 – 0	1,6,1 – 0	2,7,1 – 0	4,1,2 – 0	5,2,1 – 0	6,3,1 – 0	7,4,1 – 0
5,2 – 0	1,6,2 – 0	2,7,2 – 0	4,1,3 – 0	5,2,3 – 0	6,3,2 – 0	7,4,2 – 0
5,3 – 0	1,6,3 – 0	2,7,3 – 0	4,1,4 – 0	5,2,4 – 0	6,3,4 – 0	7,4,3 – 0
5,4 – 0	1,6,4 – 0	2,7,4 – 0	4,1,5 – 0	5,2,5 – 0	6,3,5 – 0	7,4,5 – 0
5,6 – 0	1,6,5 – 0	2,7,5 – 0	4,1,6 – 0	5,2,6 – 0	6,3,6 – 0	7,4,6 – 0
5,7 – 0	1,6,7 – 0	2,7,6 – 0	4,1,7 – 0	5,2,7 – 0	6,3,7 – 0	7,4,7 – 0
6,1 – 0	1,7,1 – 0	3,1,2 – 0	4,2,1 – 0	5,3,1 – 0	6,4,1 – 0	7,5,1 – 0
6,2 – 0	1,7,2 – 0	3,1,3 – 0	4,2,3 – 0	5,3,2 – 0	6,4,2 – 0	7,5,2 – 0
6,3 – 0	1,7,3 – 0	3,1,4 – 0	4,2,4 – 0	5,3,4 – 0	6,4,3 – 0	7,5,3 – 0
6,4 – 1	1,7,4 – 0	3,1,5 – 0	4,2,5 – 0	5,3,5 – 0	6,4,5 – 0	7,5,4 – 0
6,5 – 0	1,7,5 – 0	3,1,6 – 0	**4,2,6 – 1**	5,3,6 – 0	6,4,6 – 0	7,5,6 – 0
6,7 – 0	1,7,6 – 0	3,1,7 – 0	4,2,7 – 0	5,3,7 – 0	**6,4,7 – 1**	7,5,7 – 0
7,1 – 0	2,1,2 – 0	3,2,1 – 0	4,3,1 – 0	5,4,1 – 0	6,5,1 – 0	7,6,1 – 0
7,2 – 0	2,1,3 – 0	3,2,3 – 0	4,3,2 – 0	5,4,2 – 0	6,5,2 – 0	7,6,2 – 0
7,3 – 0	2,1,4 – 0	3,2,4 – 0	4,3,4 – 0	5,4,3 – 0	6,5,3 – 0	7,6,3 – 0
7,4 – 0	2,1,5 – 0	3,2,5 – 0	4,3,5 – 0	5,4,5 – 0	6,5,4 – 0	7,6,4 – 0
7,5 - 0	2,1,6 – 0	3,2,6 – 0	4,3,6 – 0	5,4,6 – 0	6,5,6 – 0	7,6,5 – 0
7,6 – 0	2,1,7 – 0	3,2,7 – 0	4,3,7 – 0	5,4,7 – 0	6,5,7 – 0	7,6,7 – 0

Table 7. This table shows how the state space size will change when evaluating the same series of numbers (1, 4, 2, 6, 4, 7) with two different Markov models. The left-most column depicts all possibilities associated with number sets of size 2. The six right-most columns represent the increased number of possibilities resulting by simply bumping up the sequence length to three (2nd order entropy.)

These counts of proportional values are what one would substitute into the example matrices (like in Table 5), instead of the generic equi-probable value examples seen in Table 3. This will produce different probability rates for each column (sections b. and d. from Table 4 or Table 6). One then simply sums the resulting product to get the entropy value associated with that level in the chain just like was done in the examples. One then adds up the different Markov chain levels (3rd order will also include results from 2nd and 1st, 2nd order will include the results from 1st, etc) to get an over all Entropy value associated with the experiment. Then take the observed entropy and divide it by the H_{max} to get the relative entropy for that subject. It is this value that we will use to compare all subjects against each other. High entropy will mean all combinations are close to equi-probable. Low entropy will mean redundancy and high probabilities of only a few combinations.

2.4 Experimental Summary

The main goal of this study was to evaluate a new statistical approach for quantifying how a driver's eye scan patterns change in *real world* situations. More specifically, it was sought to evaluate quantitative differences in eye scan patterns in a rural highway environment as a function of driver age and secondary task loading.

The current study estimated the statistical complexity of one's scan patterns using a metric borrowed from information theory (i.e. Entropy.) Entropy values were calculated using Markov chain models in order to account for possible dependencies between sequential eye movements. Observed entropy was divided by the highest possible entropy value to order to derive relative entropy, which was then compared across groups.

In this study two well-known secondary loading tasks were used: N-Back (verbal) and Clock face (visual spatial) task. The "N-back" task was chosen because it is a proven technique for representing *multiple levels* of loading or effort. The Visual Spatial task was chosen because it is reasoned to rely on a *similar attentional pool as vision*. Two levels of the N-Back task were used to help identify effort effects that may be present; however this was only done between young groups due to the difficulty in recruiting older subjects compounded with poor eye tracker performance once they did participate. The 1-back task represented a "low" mental loading condition while the 2-back represented a "high" mental loading condition.

VSM should form complementary results to the N-Back task since both are thought to rely on different attentional sources (Wickens, 1984.) This means they should both interact with vision differently. VSM is reasoned to rely directly on the same visual-spatial attention pool as driving because one has to “see” the image before making a criterion judgment, whereas the N-back task does not require the subject to mentally “visualize” anything.

From the literature review it can be inferred that both manipulations of mental loading should influence eye movements, however this can only be said *for certain* when referring *to laboratory settings*. There are very few studies of this nature that have been done while actually driving. A main goal of this current study was to observe how eye movements are affected in a real world setting.

2.4.1 Hypotheses

Hypothesis 1. General Workload Effect
Entropy values should decline in the presence of task loading.

Previous research has shown that as subsidiary workload demands increase, drivers demonstrate a constricted distribution of saccadic eye movement patterns. Any confinement of scanning patterns will directly impact entropy since entropy values are the highest when sampling is completely random (*no confinement*).

Hypothesis 2. Workload X Modality Effect
Entropy values should decrease more for the visual spatial loading task than for either N-back task.

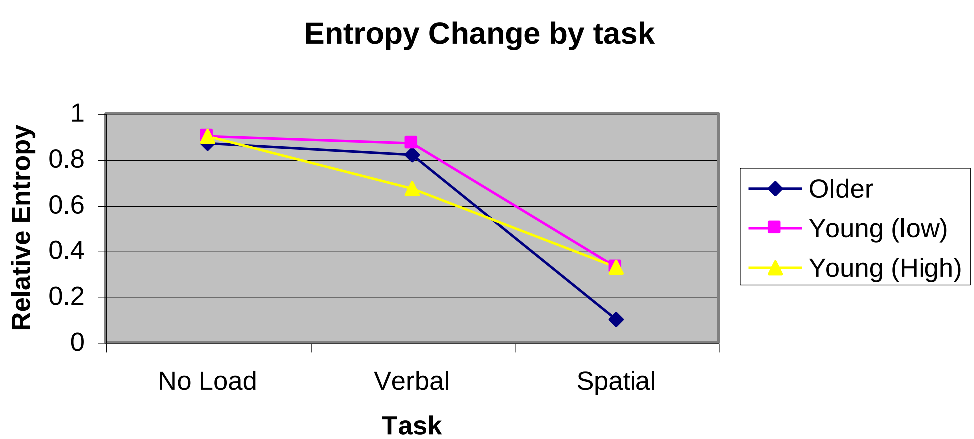
Driving is mainly a vision-oriented activity. Visual-spatial manipulation relies heavily on the same brain structures and/or attentional pools required for terrestrial navigation (i.e., driving). Tasks that rely on similar attentional resource pools should show large interference effects (i.e., it will be harder to concurrently share attention). The N-Back task will also show a decreased entropy value, but not to the degree as VSM due to dissimilar pools of attention resources being used. (Wickens, 1984)

Hypothesis 3. Age Effect
Entropy values should decrease with age.

Aged individuals generally show a decreased response rate in most situations. Under stress they have been shown to perform similarly but at a significantly decreased level from a younger group (see Figure 15). It has been observed in other studies that older subjects will miss many more targets in the periphery under high task load and begin to focus on narrower subsets of the visual scene. We expect this to translate to a general decrease in the breadth of visual scanning, which ultimately means decreased entropy.

Hypothesis 4. Age X Workload Effect
Effects of task load should be exacerbated in old groups.

It is expected that the difference between task types should be especially unfavorable to the older subjects doing the visual spatial task. They are already more prone to having less spare attention due to natural physiology, so tasks that borrow attention from the visual-spatial system should be especially taxing causing fewer scanning updates. Few scanning updates ultimately means fewer combinations that are sampled meaning a much more ordered (low entropy) system should be observed.



Other Hypotheses
1. Lowest average entropy will be observed in the Older subjects in all conditions
2. Verbal #1 (1-back) will have higher entropy than visual
3. Verbal #2 (2-back) will have higher entropy than visual
4. Verbal #1 will have a higher entropy than Verbal #2
5. The largest drop in entropy will be in the visual spatial condition in both age groups

3.0 Methods

3.1 Participants

There were a total of forty-two participants in this study. Each was recruited from around the Vermillion, South Dakota area. The younger subjects were recruited through use of the University of South Dakota's Sona-Systems Online (http://www.usd.sona-systems.com) subject recruitment system or by word of mouth. The older subjects were recruited from area community service organizations (e.g., University Emeritus Club, Lions Club, Rotary International, etc) or acquired from a list of participants who previously specified an interest while participating in other University studies.

Twenty eight of the participants were classified as "young" and fourteen were classified as "older". The young were separated into 2 separate groups in order to test a task difficulty question. Group 1 (Young) consisted of 5 females and 9 males with an average age of 27 (range = 19-35, SD = 4.3yrs). Group 2 (Young) was composed of 7 females and 7 males with an average age of 22 (range = 18-38, SD = 5.3yrs). The "older" participant group contained 6 females and 8 males with an average age of 75 (range = 67-86, SD = 5.6yrs). All subjects held a current driver's license and had previous driving experience. All data was acquired between July 1st 2006 and October 26th 2006 between the hours of 11am and 4pm.

3.2 Apparatus

The USD Toyota Avalon test vehicle (see Appendix A) was used in this experiment to house all data collection devices. Eye-tracking data were collected with a dash mounted ASL-ETS PC corneal reflection tracking system (see Appendix C).

Heading, steering, GPS, speed, and response data were sampled by several additional computer systems mounted in the trunk of the vehicle (see Appendix B). A (DC-AC) converter running off a deep cycle battery served to power all systems. Battery charge was maintained with a high capacity alternator that replaced the car's stock alternator.

Sound stimuli for both verbal and visual spatial tasks were verbally recorded onto a Windows 98 computer using Goldwave v5.14 (2006) software. Stimuli were then frequency filtered first using the "light hiss" noise reduction preset followed by a low-pass filter set at: (1) 4000khz initial cutoff, (2) static setting, with (3) *steepness* set at "20". This was done to minimize bias due to the natural hearing loss suffered by the common older subject. Since hearing frequencies are lost in the upper ranges with age, all stimuli was filtered to make the younger subjects with better hearing "hear" like the average frequency deficient older subject.

All collected data was synthesized into a master file via the pathways shown in Appendix E.

3.3 Procedure

Attaining human subject testing approval from the University of South Dakota was the first precursor to this project. This consisted of CITI certification (see Appendix F) and a full committee review of the test protocol. Once approval was received, we began processing subjects using the structure outlined in the next few paragraphs.

Upon arrival at the USD Psychology building, subjects were provided with a description of the study and told the requirements that would be placed upon them. They were informed that they may quit the study at anytime if they felt uncomfortable. All were then required to read and sign an informed consent form outlining all aspects of the experiment before they would be allowed to participate (see Appendix D for a sample of the informed consent document.) After consent was granted, we assigned them a number to assist with anonymity and began collecting background information such as age and accident history. Next, two vision screening tests were given to evaluate the eyes binocularly at both near and far distances. All eye measurements were done using a Bausch and Lomb Orthorater. For this study, a minimum far visual acuity of 20/40 was chosen because this value is the cutoff for driver licensing in many states (including South Dakota.) Failure to see at this minimal level would mean the study could not progress any further with that subject. Luckily, no subject was excluded from the study based on this criterion. The final step was to measure auditory sensitivity. This ensured that all subjects were physically able to hear our auditory delivered stimuli and also helped determine which ear was the most sensitive for stimuli delivery (left/right reversible headset used in car).

Next, the subject was taken to the University's Toyota Avalon test vehicle (see Appendix A for pictures) and told to adjust the seats, mirrors, etc. to a comfortable position. A brief learning session involving N-back tasks and clock tasks was then administered to familiarize the subject with these procedures. First, several sample trials of each task were demonstrated to the subject, then approximately 30 practice trials of each task were administered until the subject felt confident with each task (or error rate = 0 for 10 successive trials).

After the task learning phase, subjects were calibrated to the ASL-ETS-PC dash mounted eye-tracking system using standardized calibration poles mounted onto the front bumper of the test car by means of steel pegs. Mounting the calibration stimuli on the car guaranteed that all subjects received the exact same positional orientation of stimuli. It also helped to assure alignment the scene camera to exactly the same position for each subject.

The subject was then instructed to begin driving to the outskirts of the northern edge of Vermillion to begin the experiment. Each trial began and ended at Highway 50. All measurements were taken while driving north/south on North University Road. Each trial lasted approximately 15 minutes and covered 8 miles. All data was collected during three round-trips to the “bend” and back to the edge of town (see Appendix G for map of route). Each round-trip measured a different loading situation (no loading, Verbal task (1 or 2), or Visual Spatial task). The order of these task trials was randomly determined at the beginning of the experiment to balance fatigue and/or learning effects that may occur. For both loading task conditions, the subjects are presented with *auditory* recordings of loading stimuli while driving.

In the N-Back task condition, subjects were presented with a continuous stream of numbers (Inter Stimulus Interval (ISI) = 4 seconds) and asked to make a judgment concerning if the number was the same or different than the number presented 1 or 2 units previously. For example, given the presentation of the series (1, 3, 7, 2, 4, 2, 3, 5, 5), is the last digit presented (5) the same as the one presented 2-back (3)? In this case, the subject would report "NO". The error rate was calculated to insure that the subject was doing the task as intended and not simply guessing. This error rate also served as a reference for relative cognitive loading since we assume more cognitively involved processes should be more prone to error.

In the Visual Spatial Loading task (clock task), subjects were instructed to imagine the position of clock hands associated with a time that was again auditorily presented. Given the imagined clock hand positioning, is the smallest angle formed between those hands less than or greater than 90°? They were told to simply respond with a "yes" for greater than 90°, or "no" for less than 90°. For example, if one heard the time "1:15", they would imagine an analog clock with hands on the 1 and 3 positions. In this example the response would be "yes": the smallest angle formed is *less than 90°*. Again, as with the Verbal task, the error rate was computed as an indicator of workload and to assure the task was being attempted correctly. In this study the ISI for the Visual Spatial task was adjusted to 8 seconds as opposed to 4 seconds in the verbal task. This was because any ISI less than this frequently resulted in an overlap of stimuli presentation without the subject being able to make a decision between presentations.

The main concern in this study was the data derived from the eye scan patterns and how this changes in response to the loading task. Because of this, all data needed to be collected *without* cueing any unwanted searching strategies. The data was desired to be as naturalistic as possible so that *only* the loading task could be logically linked to performance changes. The subjects received all stimuli auditorily, whereupon they completed the task and reported the answer vocally which can be reasoned to be similar to the demands of an in-car conversation. Being spoken to and responding is a task that many already do while driving so subjects should be experienced. This should be minimally intrusive to the task of driving as noted before since most people should be experienced with conversing while driving.

The fixation data was initially stored as raw coordinates representing where the eye was looking at any given fixation during the course of the experiment. This data was later screened and transformed by ASL ETS-PC software into data we could analyze. Fixations were defined by durations >200msec and transitions between fixations as a movement >1°. Fixations were then arranged in a sequence list determined by transitions across the predefined boundaries (see Figure 21). Due to variations in calibrations between subjects it became necessary to define an anchor point individually to each subject from which to define the 7 proposed boundaries. A detailed account of how this was calculated can be viewed in Appendix H.

Upon completion of the three test runs, the car was driven back to the University where the subject was debriefed on specifics of the study. Data was then ported out of the car's computer via the use of a Memorex 2-gigabit memory stick.

There were 127 separate data files recorded in this study, each representing a single trial. All were coded into the format: subject number/trial type (N = no task, A = 1-back, B= 2-back, C=clock task.) All data was screened and manipulated using the following protocols:

1. Any trial without a full video file was immediately labeled void for the entire trial regardless if the data looked acceptable. Records 107N, 112N, 112A, 120B, 120C, 120N, 121C, 121N, 128C, 128N, 206C and 214C were excluded from *all* analyses due to this criterion.
2. The exact same sections of roadway were traveled in each trial. These were determined through the trial's scene video. Any data before or after these points was deleted. The Vermillion "cut" point was a speed limit sign about ½ mile

from Route 50. The north side turn-around “cut” point was a shelter belt on the west side of the road approximately ¼ mile from the turn.

3. Any data below 30mph was excluded if it occurred inside a designated testing range *and* was not within 10 mph of that subject’s overall average. Major speed changes usually occurred as a result of anticipation. This is to say they began slowing down for corners, cars, bikes, etc. Since the change was occurring in anticipation a future event, the scanning patterns may have also been compromised so it was better to error on the safe side and not include it. As a side note, there were two older subjects who had speeds regularly drop below 30mph. However, since this was a regular occurrence and not blatantly linked to an external event (as determined by experimenter riding the back seat), this data was still used for analysis.

4. All data occurring during a passing, following, or other potentially confounding situation was also cut from the main file. The goal was to have as naturalistic data as possible related *only* to straight road, non-guided driving at highway speeds.

5. The portions of the data files that remained after the first 4 edits were then screened again to make sure they had at least 175 usable state space transitions per trial. Any data files that could not meet this criterion were not used for visual entropy analyses as the opportunity to sample all available state spaces starts to become questionable for the higher order entropy analyses. Only fixation files with a minimum of 175 transitions were used in visual analyses, whereas any file with a valid video was eligible for the non-visual metrics (e.g. yaw rate, speed,

accuracy, etc.) Records 103A, 106C, 106N, 112A, 112N, 121A, 123B, 125B, 127B, 127C, 130B, 201A, 203A, 204A, 204C, 206A, 206N, 207A, 207C, 207N, 208A, 208C, 208N, 209A, 212A, 212C, 212N, 213A, and 213C were excluded from the visual entropy analyses based on this criterion.

6. After determining eligibility and screening using the above 5 criterion, an accuracy check was run on all responses given during the trials. Any trial with less than 70% accuracy was cut due to the chance that they may not be doing the trial as instructed or intended. Only 2 files (107C, 203C) were excluded based on criterion #6.

3.4 Results

This study employed a 2 (age) X 3 (task: no subsidiary task, verbal-memory task, visual-spatial task) experimental design. The between subjects variable was driver age while the within subjects variable was subsidiary task load condition. The main dependent variable was eye movement entropy; however, steering wheel movements, gyroscope output, and vehicular speed were also subjected to exploratory analyses.

Old	**Baseline**	**1-Back**	**Clock Task**
Young-1	**Baseline**	**1-Back**	**Clock Task**
Young- 2	**Baseline**	**2-Back**	**Clock Task**
	Control	**Verbal Coding Mode**	**Visual Coding Mode**

Table 8. Experimental Design.

The young participants were divided into two separate experimental groups. Half of the young participants (Group 1) were required to perform a 1-Back memory task during the verbal loading condition. The other half (Group 2) performed the more demanding 2-Back version of the memory task during the verbal secondary task loading condition. This manipulation was implemented in order evaluate the potential role of an age by task difficulty interaction effect across the verbal and visual load conditions should the visual-spatial task prove more demanding for the older participants than the verbal-memory task. Since this pattern of results was demonstrated, the data from the 2-Back sample (young Group 2) are analyzed and discussed in subsequent sections of this report.

SPSS's General Linear Model was used to test all main hypotheses, and uncorrected t-tests were used to implement *post hoc* comparisons between and across conditions. Since this experiment involved the factorial combination of a within-subjects factor with a between-subjects factor (i.e., a so-called mixed design) there was a potential for violations of the sphericity assumption in our data (i.e., homogeneity of covariance). This would mean that uncorrected ANOVAs would tend to exhibit type I error levels that exceeded the specified alpha levels for the F-statistic. To counter this possibility, a Huyhn-Feldt sphericity correction was used as suggested in Keppel (1973). Although there was some violation of the sphericity assumption in the data, the Huyhn-Feldt results were not inconsistent with the uncorrected ANOVA output and therefore not considered an issue in this study. Uncorrected values were therefore used.

In this compilation of the results, the following labeling conventions were used:

1. *Age Groups*

   (a) *Group 1 Young,* may also be referred to as "**Y1**"
   (b) *Group 2 Young,* may also be referred to as "**Y2**"
   (c) *Group 1 Old,* may also be referred to as "**O1**"

2. *Loading Tasks*

   (a) 1-Back task = verbal-memory task 1
   (b) 2-Back task = verbal-memory task 2
       (This task was ONLY done by Group 2 Young)
   (c) Clock Task = Visual Spatial task
   (d) No loading task (baseline) = None

As described in the Method section of this report, data were collected from fourteen participants in each of the three experimental groups. However, because of numerous difficulties encountered with the apparatus used to track eye movements, all of the analyses of eye movement behavior reported below were based upon smaller subsets of participants. The typical number of data files available for eye-movement analysis was ten in the younger groups, versus six in the older group.

The nature of the difficulties encountered with the eye tracking apparatus is documented below. Significant effort (and expense) will be required to minimize the eye movement data loss among older participants in any future studies employing the ASL ETS eye tracking system. Analyses of non-ocular driving performance metrics typically were based upon group sizes of 14.

3.4.1 Young (Y1) vs. Old: Visual Entropy Results

SPSS GLM (2) age group by (3) loading task ANOVA yielded significant results for both the loading task [$f\ (2,24)=16.402$, $p<.001$] and the loading task by age group interaction [f (2,24)=9.808,p<.002] when using the dependent variable of 1st order entropy. When compared by 2nd order entropy we still observed a slightly decreased but significant loading task effect [f (2,24)=13.293,p<.001] and loading task by age group interaction [f (2,24)=8.208,p<.003].

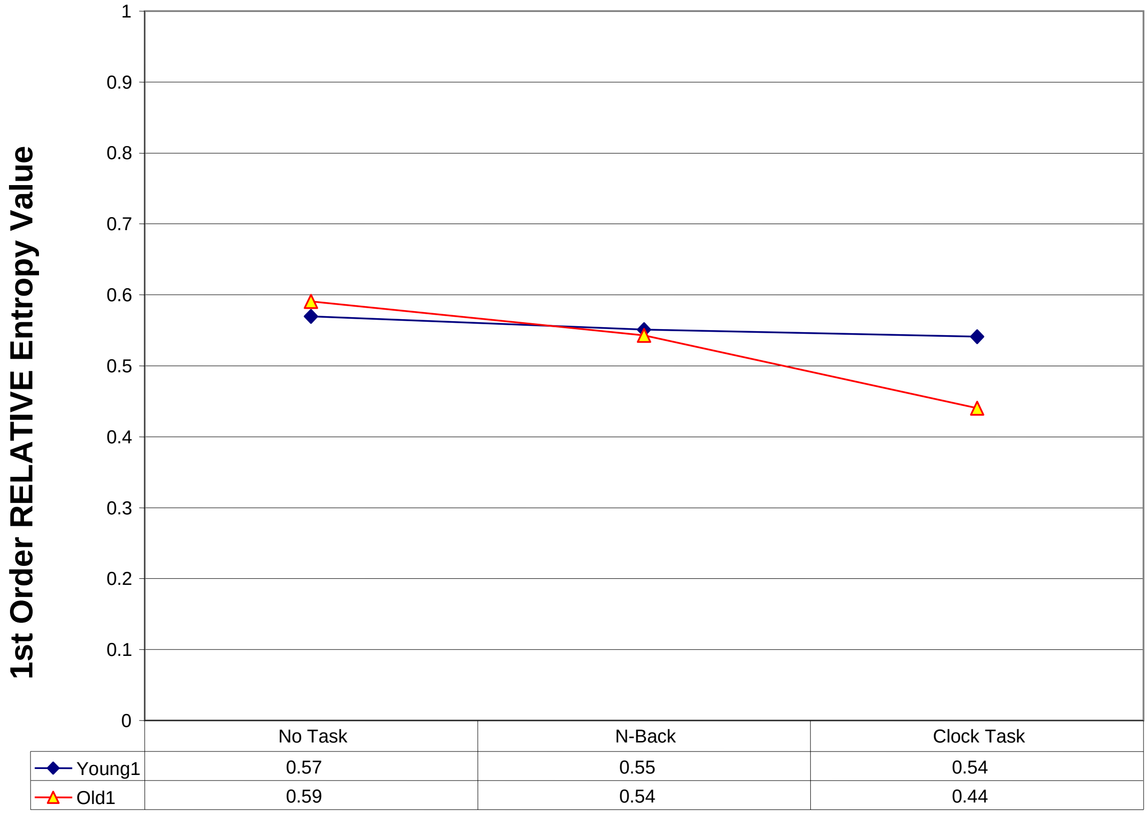


Figure 22. Mean 1st Order Relative Visual Entropy - Young vs. Older Groups.
(Actual divided by max to determine RELATIVE)

Paired sample t-tests of 1st order entropy data reveals significance between conditions:

(a) 1-Back vs. No Task in Y1 (t(9) = -4.271, p<.003)
(b) Clock vs. No Task in O1 (t(5) = -6.286, p<.002)
(c) *between* groups in Clock Task (t(5)= -3.011, p<.03).

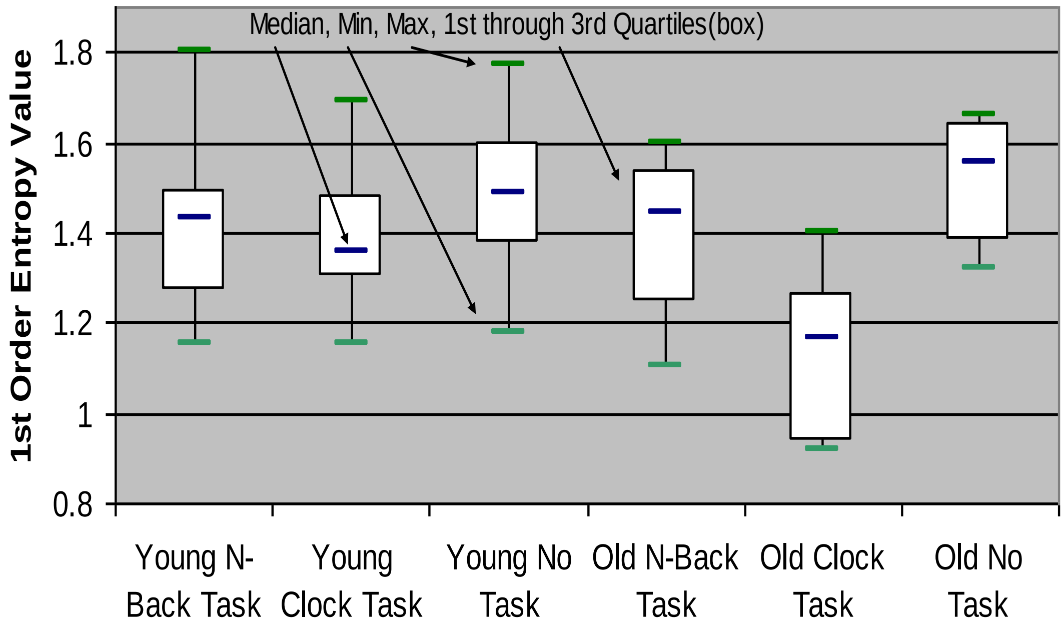


Figure 23. Standard box plots of 1st order Entropy by task load and age group.

(Actual entropy values, not converted to relative)

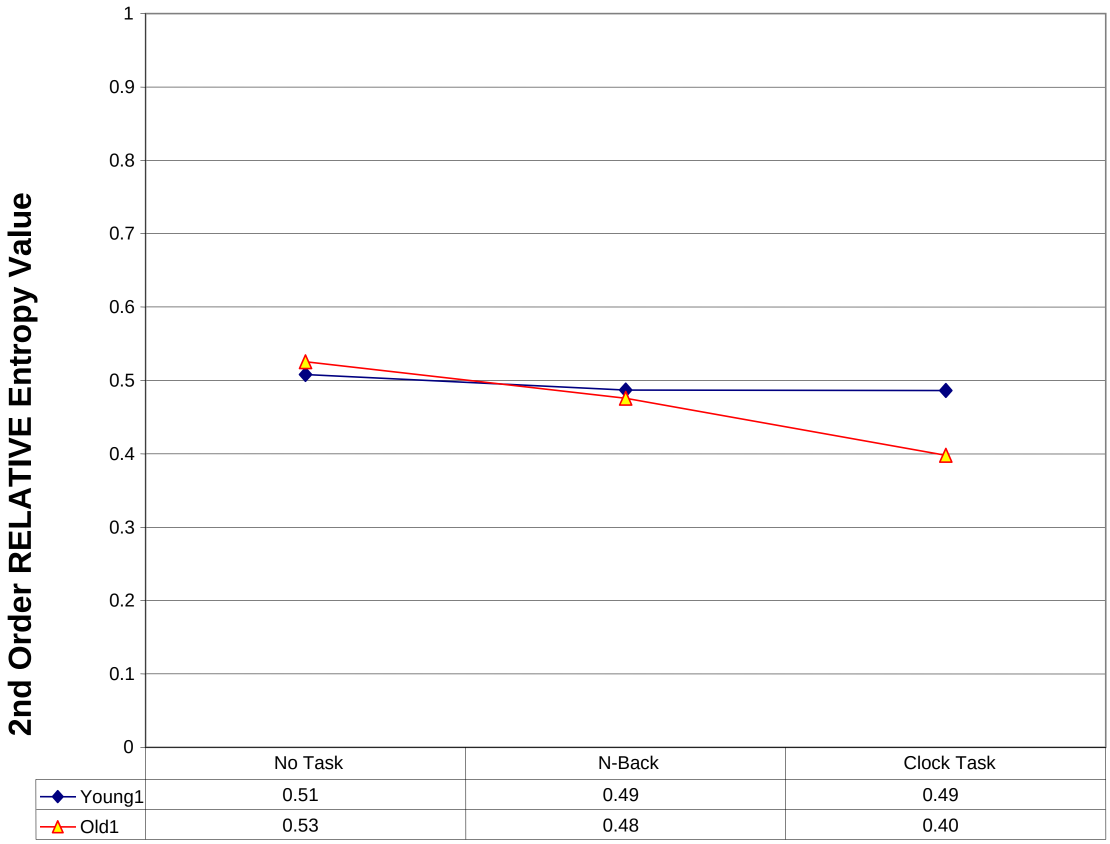


Figure 24. Mean 2nd Order Relative Visual Entropy – Y1 vs. O1 groups.

(Actual divided by max to determine RELATIVE)

Paired sample t-tests of 2nd order entropy data reveals significance between conditions:

(a) 1-Back vs. No Task in Y1 (t(9) = -3.025,p<.02)
(b) 1-Back vs. Clock in O1 (t(5) = -3.009,p<.003)
(c) Clock vs. No Task in O1 (t(5) = -5.878,p<.003)
(d) *Between* groups in Clock (t(5) = -3.064,p<.03)

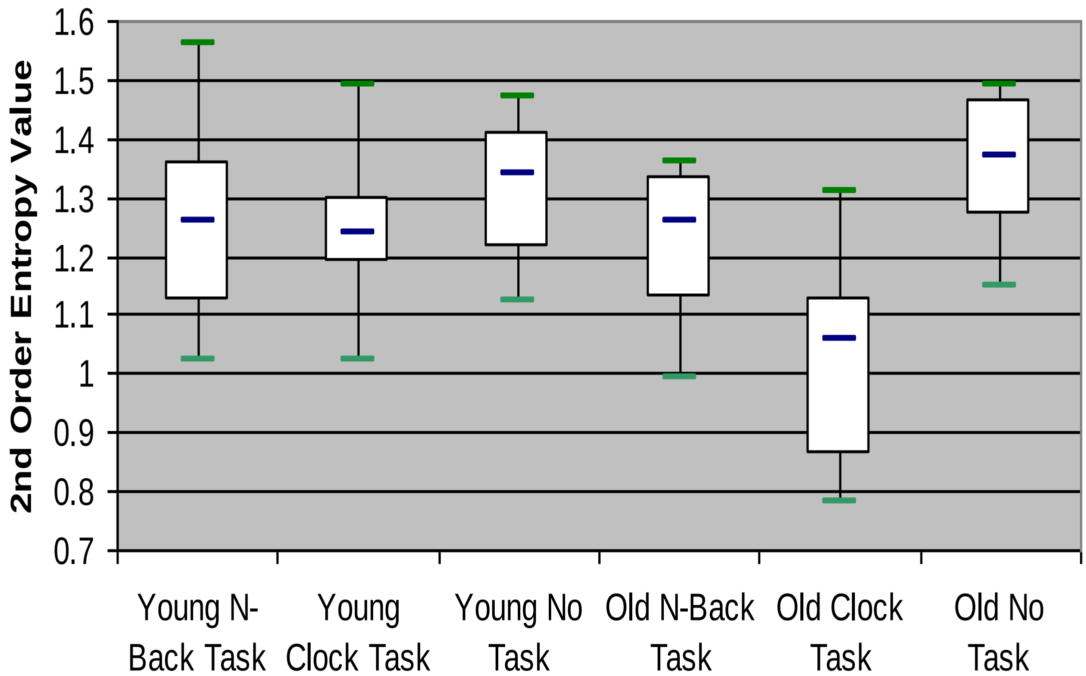
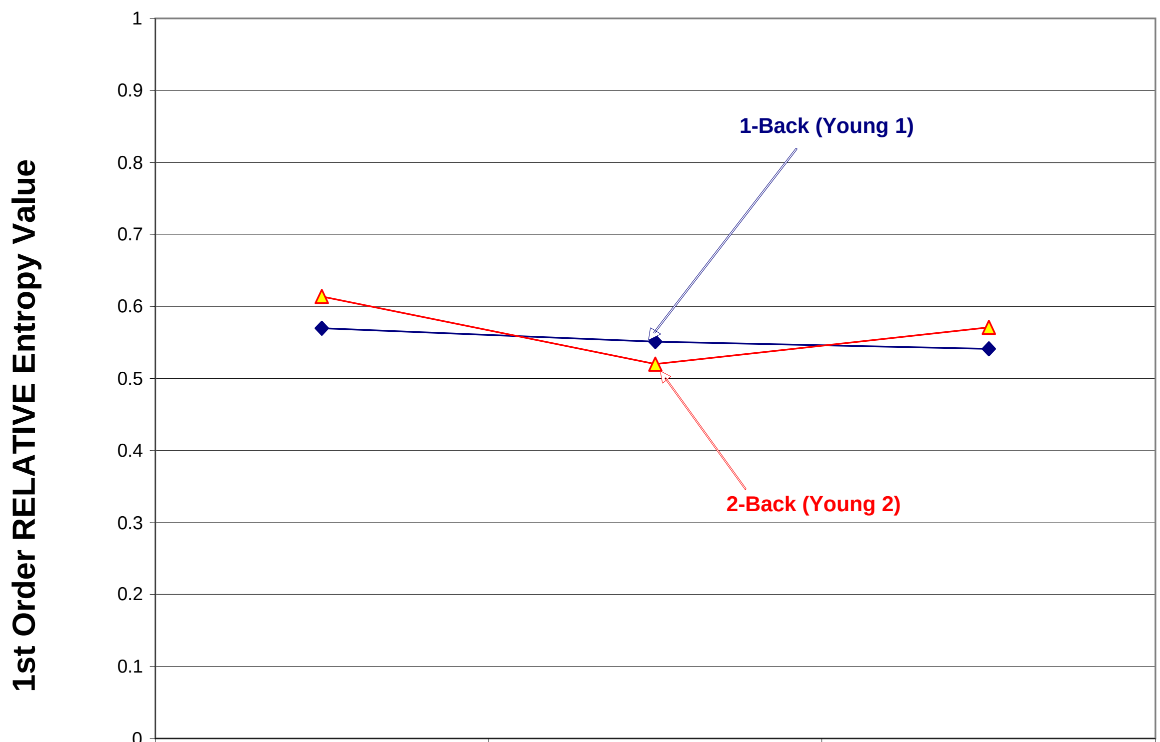


Figure 25. Standard box plots of 2nd order Visual Entropy by task load and age group.
(Actual entropy values, not converted to relative)

3.4.2 Y1 vs. Y2: Visual Entropy Analysis

SPSS GLM (2) group by (3) loading task ANOVA contrasting the two young groups revealed a significant loading task effect [f (2,30)=12.269,p<.001] and a significant loading task by group interaction [f (2,30)=3.768,p<.04] for 1st order Entropy. Analysis of the 2nd order visual entropy data yielded a similar pattern of results: both loading task [f (2,30)=14.443,p<.001] and the loading task by group interaction [f (2,30)=4.869, p<.02] were statistically significant.



	No Task	N-Back	Clock Task
Young1	0.57	0.55	0.54
Young2	0.61	0.52	0.57

Figure 26. Mean 1st order Relative Visual Entropy – Y1 vs. Y2 groups.

(Actual divided by max to determine RELATIVE)

Paired sample t-tests of 1st order entropy data reveal significance between conditions:

(a) 1-Back vs. No Task in Y1 (t(9) = -4.271,p<.003)
(b) 2-Back vs. No Task in Y2 (t(6) = 3.497,p<.02)

There was no significance differences noted between groups in any one category, just individual differences within each of the young participant groups.

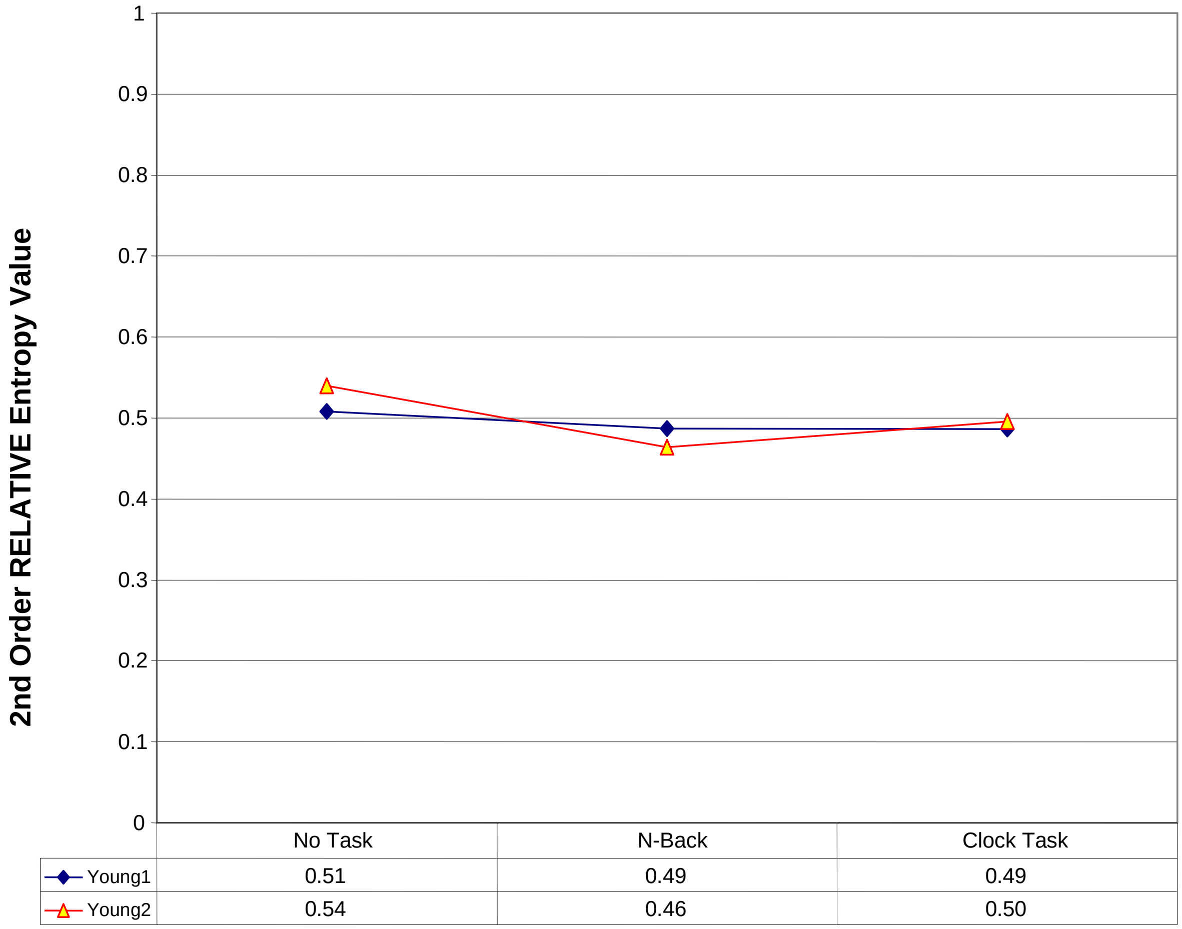


Figure 27. Mean 2^{nd} order Relative Visual Entropy – Y1 vs. Y2 groups.

(Actual divided by max to determine RELATIVE)

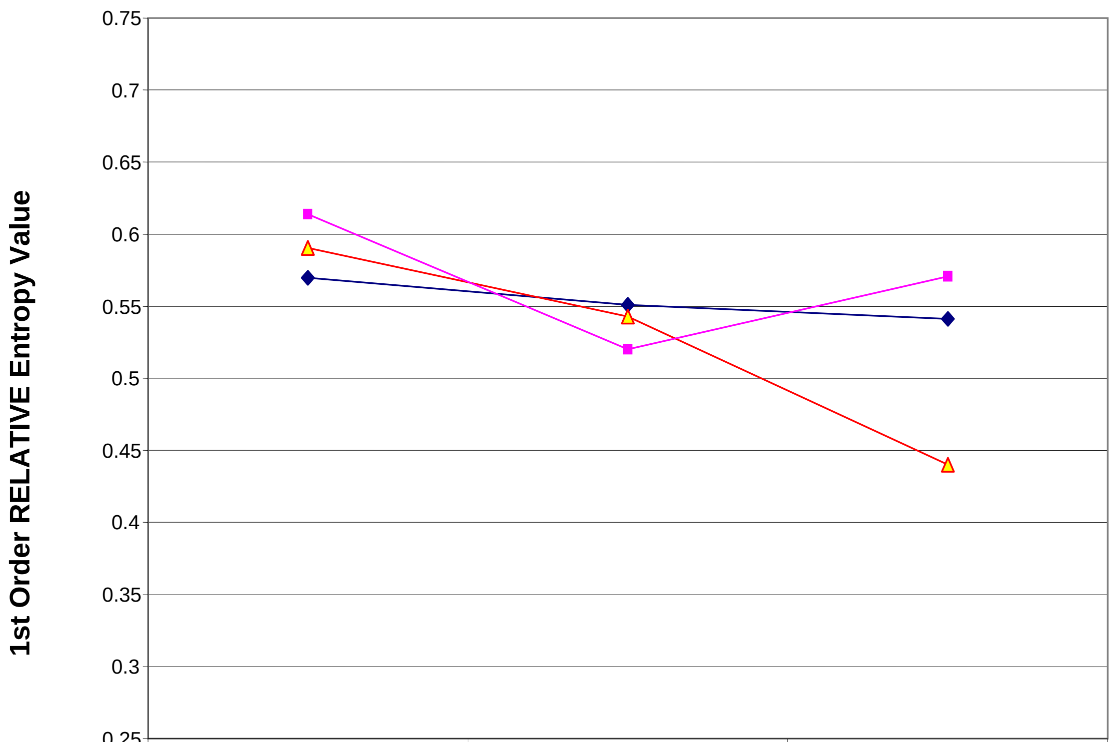
Paired sample t-tests of 2^{nd} order entropy data reveal significance between conditions:

(a) 1-Back vs. No Task in Y1 $(t(9) = -3.025, p<.02)$
(b) 2-Back vs. No Task in Y2 $(t(6) = -3.090, p<.03)$
(c) Clock vs. No Task in Y2 $(t(9) = -2.321, p<.05)$

Again, there were no statistically reliable differences noted between groups in any one category, just individual differences within each group of young participants.

3.4.3 Summary of Visual Entropy Effects

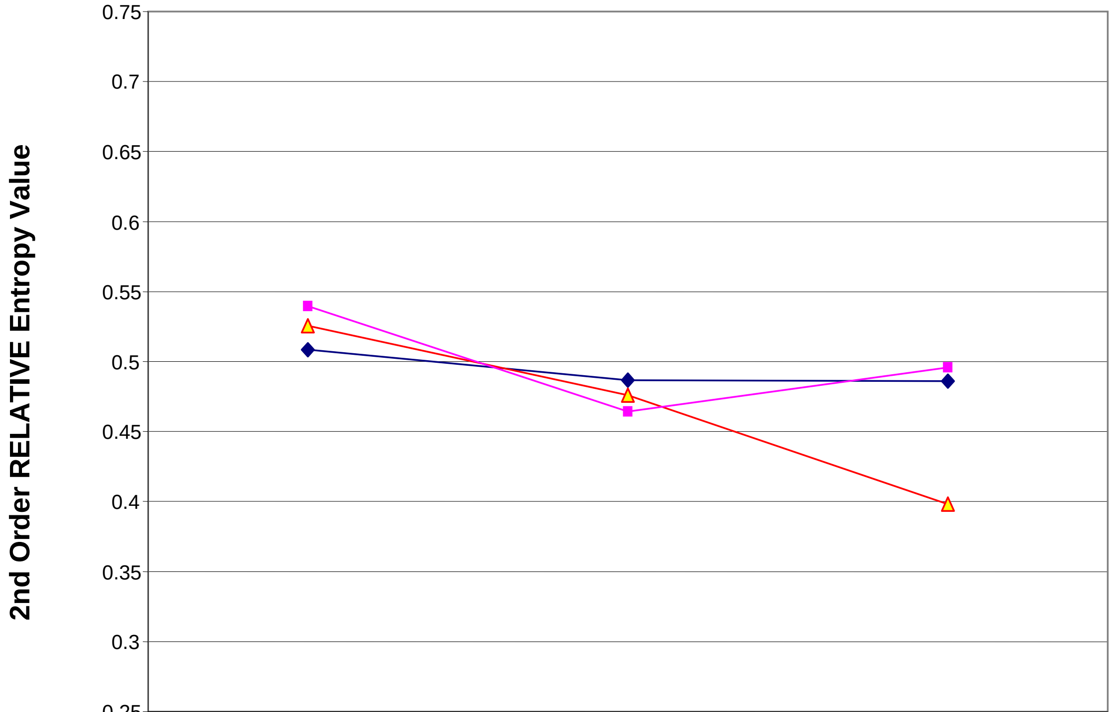
The next two figures (28 & 29) summarize the entropy associated with each task for all participant groups. Figure 28 represents 1st order entropy and Figure 29 represents 2nd order entropy. Note that 1st and 2nd order entropy both showed the similar pattern even though the 2nd order entropy was lower in overall value and exhibited less separation between each group.



	No Task	N-Back	Clock Task
Young1	0.57	0.55	0.54
Old1	0.59	0.54	0.44
Young2	0.61	0.52	0.57

Figure 28. All group 1st order Relative Entropy values by task.

(Actual divided by max to determine RELATIVE)



	No Task	N-Back	Clock Task
Young1	0.51	0.49	0.49
Old1	0.53	0.48	0.40
Young2	0.54	0.46	0.50

Figure 29. All group 2nd order Relative Entropy values by task.

(Actual divided by max to determine RELATIVE)

3.4.4 Eye Gaze Transition Matrices

What follows in the next three subsections is a descriptive analysis based upon the twenty “most frequently observed” eye gaze transition sequences for young 1(Y1), Old 1 (O1), and a composite of both. These findings are presented via Figures 30 through 35 to help the reader visualize each eye gaze transition state’s relative contribution to overall viewing behavior as well as to the computation of the visual entropy metric.

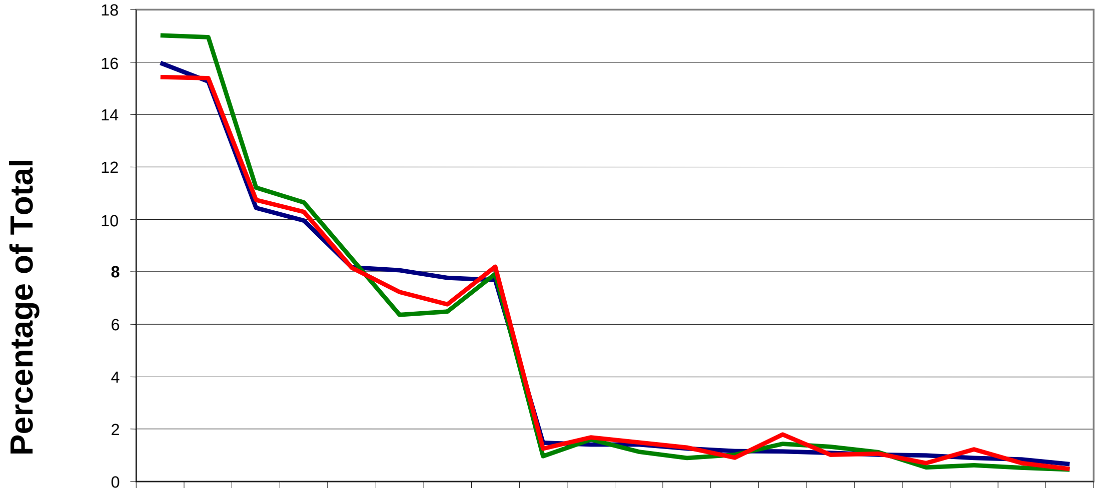
The most heavily viewed transition sequence was to move from gaze state 2 to 3 and then back to state 2 again. If one refers back to Figure 21, this is the sequence representing a gaze directly in front of the car and the “far right ditch”, respectively. This is a logical scanning pattern as one could imagine that these are the areas with the most information for this particular driving scenario. The road was narrow with no shoulder, posted at a high speed (55 MPH), and had very little traffic. The right ditch area was probably scanned frequently owing to the absence of a shoulder and the anticipated probability of occurrence random hazards such as animals, foreign objects, and pedestrians. In addition, drivers apparently did not need to scan the left side (gaze state #1) as much since there were not many hazards usually associated with that side of the road beyond other vehicles (see Figures 30 and 31).

The least frequently observed gaze state transitions were combinations such as looking from one “corner” to another within the defined visual frame. If a look to a corner did occur, it was almost always redirected to space #2 in the next transition. A brief look at the data suggests that state space #2 is the most often used. Of the top 20 combinations, this space follows or is followed 50% of the time. This should not be surprising as this is also the space that encompasses the defined “center” of our visual world. It is located immediately straight in front of the driver in the path that he or she is traveling. One will also note a high percentage of looks in the 2-5 or 5-2 areas (see figures 30 & 31). This is also somewhat predictable as it represents looking in front of the car and then scanning further out or vice versa. This is another area that is essential to the maintenance of heading of the vehicle.

There were also very few glances that started or ended in the lower two corners (i.e., states 4 or 6). When one did occur, state 4 was considerably more represented than gaze state 6. In straight road driving there really just isn't much going on in these areas so this may account for this decreased frequency. The dramatically decreased use of state space #6 seems to suggest that once a target leaves the realm of gaze state 3, it may no longer be considered a hazard and so is no longer monitored.

Glance transitions to gaze state 7 almost always were followed by a glance back to state 2. This is more than likely due to glances at the speedometer. One more interesting observation to point out is that the frequency of the "top 20" transition states seems to have remained relatively constant across conditions in group Y1, while a relatively large decrease in sampling space use occurred in the O1 group under the Clock load task condition.

To help with the interpretation of these graphs, simply look up the transition sequence located at the bottom of Figures 30 and 32 then refer to Figure 31 and 33, respectively, to see the corresponding vectors representing each gaze state transition within the visual space of the forward scene of the roadway.



	23	32	21	12	52	72	27	25	31	35	13	42	75	53	15	51	57	24	14	54
1-Back	16	15	10	10	8	8	8	8	1	1	1	1	1	1	1	1	1	1	1	1
Clock	17	17	11	10	9	6	6	8	1	2	1	1	1	1	1	1	1	1	1	<.5
No Task	15	15	11	10	8	7	7	8	1	2	1	1	1	2	1	1	1	1	1	<.5

Transition Order

Figure 30. "Top 20" Gaze State Transitions of the Young (Y1) group ordered by percentage of total transitions (rounded to nearest whole number).

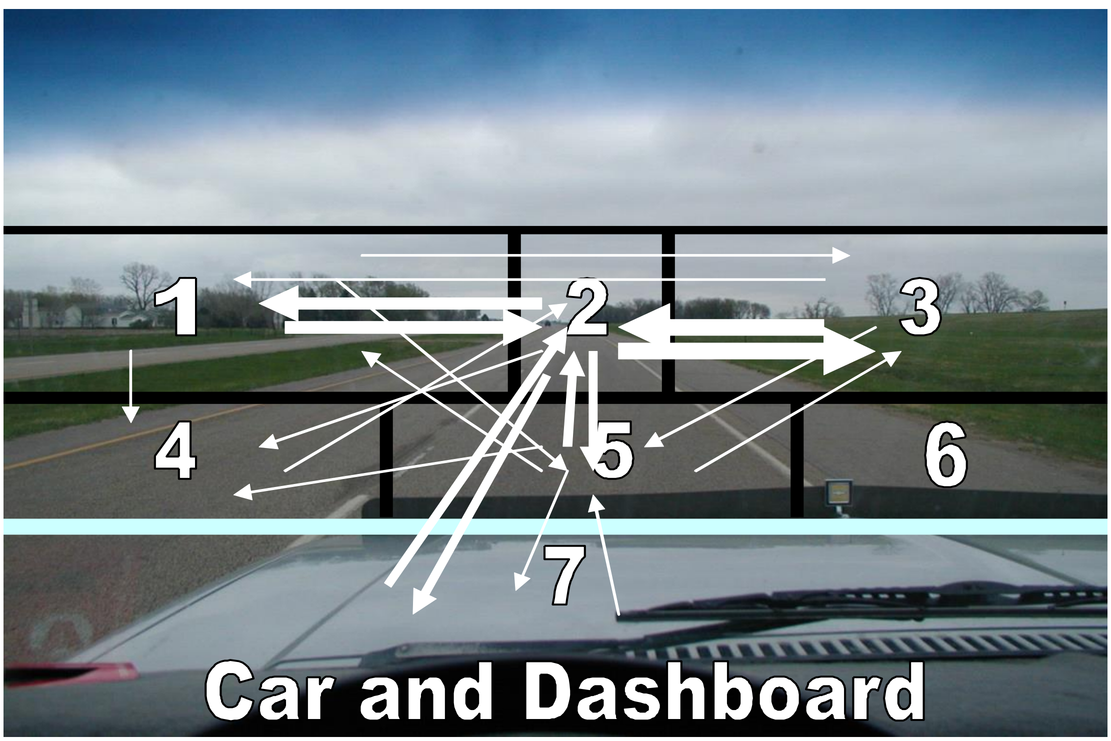
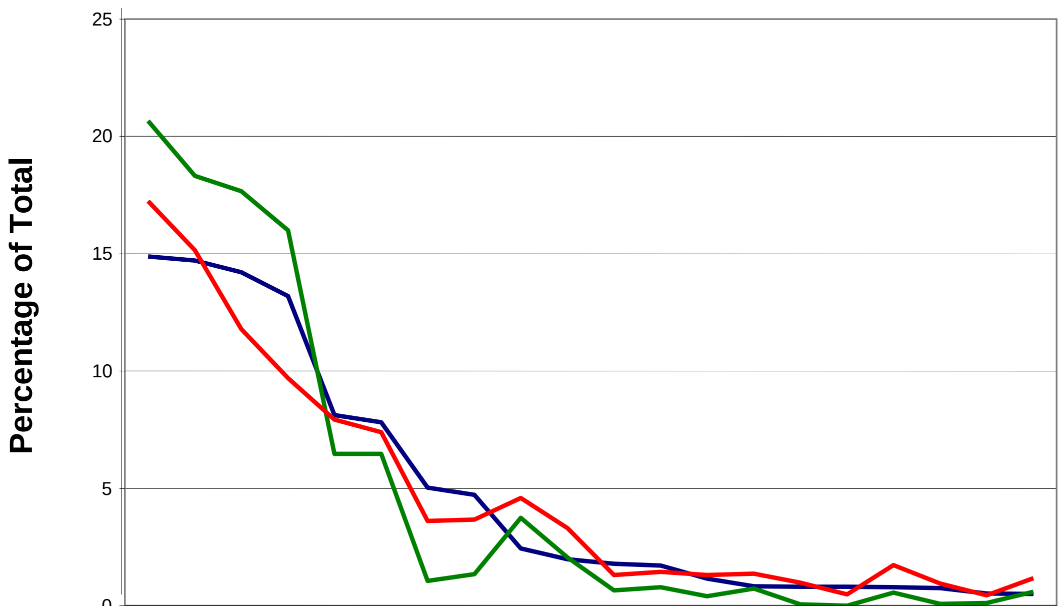


Figure 31. A picture of what the "top 20" transition sequences look like mapped out for the young (Y1) group. The higher the percentage of transitions accounted for, the more thick the indicator line.

The direction of the transition is indicated by an arrowhead in Figure 31. The highest percentage of transitions (approx 15-17%) took place between the 2-3 or 3-2 states as indicated by the heavy indicator line. The next highest frequency transition event was between the 2-1 and 1-2 states with 9-11% of the total. The next highest were the 2-7, 7-2, 5-2, and 2-5 states with 7-8% of the total. No other transition state sequence could individually account for more than 1.8% of the total. These transitions are represented by the narrowest vectors.



	23	32	52	25	21	12	27	72	35	53	24	42	13	15	75	41	31	14	37	51
1-Back	15	15	14	13	8	8	5	5	3	2	2	2	1	1	1	1	1	1	1	<.5
Clock	21	18	18	16	6	6	1	1	4	2	1	1	<.5	1	<.5	<.5	1	<.5	<.5	1
None	17	15	12	10	8	7	4	4	5	3	1	1	1	1	1	<.5	2	1	<.5	1

Figure 32. "Top 20" Gaze State Transitions of the Old (O1) group ordered by percentage of total transitions (rounded to nearest whole number).

.

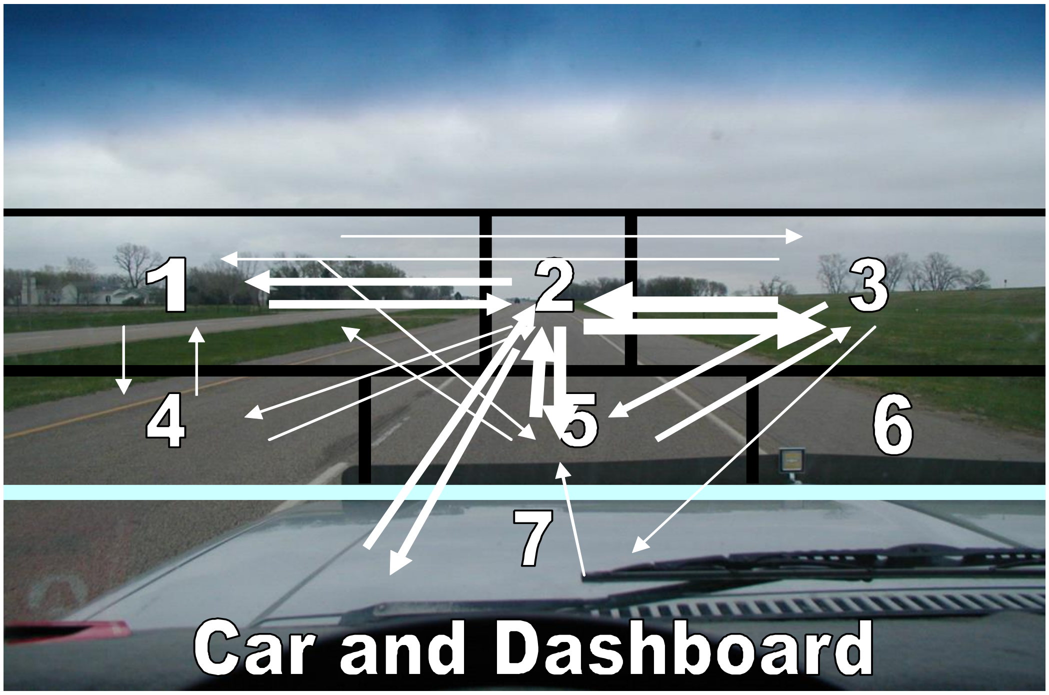
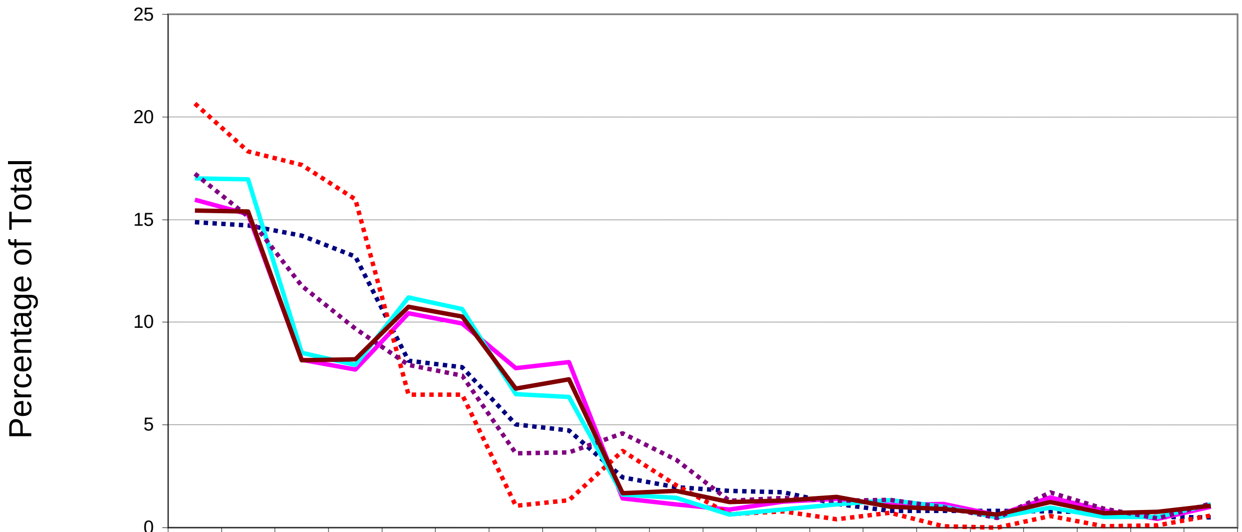


Figure 33. A picture of what the "top 20" gaze state transition sequences look like mapped out for the Old group.

Again in Figure 33, the direction of the transition is indicated by an arrowhead. The highest percentage of gaze state transitions (approx 14-20%) took place between the 2-3 or 3-2 states as indicated by the heaviest indicators. One trend that stands out immediately is a shift of attention more towards the 2-5 and 5-2 state transitions rather than the 1-2 and 2-1 transitions observed in the young (Y1) group data. These state transitions accounted for 9-18% of the total. The next highest were 2-7, 7-2, 5-2, and 2-5 state transitions with 7-8% of the total (these were 9-11% in the young). There is then a leveling out observed with transitions 2-1, 1-2, 2-7, 7-2, 3-5 and 5-3 accounting for approximately 2-8% of the total. No other transition sequence could individually account for more than 1.8% of the total.

3.4.4.3 Group Young 1 and Old 1



	23	32	52	25	21	12	27	72	35	53	24	42	13	15	75	41	31	14	37	51
O 1-back	15	15	14	13	8	8	5	5	3	2	2	2	1	1	1	1	1	1	1	<.5
Y 1-back	16	15	8	8	10	10	8	8	1	1	1	1	1	1	1	1	1	1	<.5	1
O Clock	21	18	18	16	6	6	1	1	4	2	1	1	<.5	1	1	<.5	1	<.5	<.5	1
Y Clock	17	17	9	8	11	11	6	6	2	1	1	1	1	1	1	1	1	1	1	1
O None	17	15	12	10	8	7	4	4	5	3	1	1	1	1	1	<.5	2	1	<.5	1
Y None	15	15	8	8	11	10	7	7	2	2	1	1	1	1	1	1	1	1	1	1

Transition Order

Figure 34. “Top 20” Gaze Transitions of the Old and Young (Y1) groups ordered by percentage of total gaze transitions (rounded to nearest whole number).

Figure 34 is simply the aggregate top 20 gaze transition sequences resulting from both groups to show the main transition sequence usage as a whole. Note that both groups demonstrate similar sequences in the same proportions with one exception. The function that really stands out is the Old group under the clock task. This function demonstrates decreased diversity on the scanning pattern with 73% of the total possible transition sequences occurring in the four most frequent transition states: (a) 2-3 (21%), (b) 3-2 (18%), (c) 5-2 (18%), (d) 2-5 (16%).

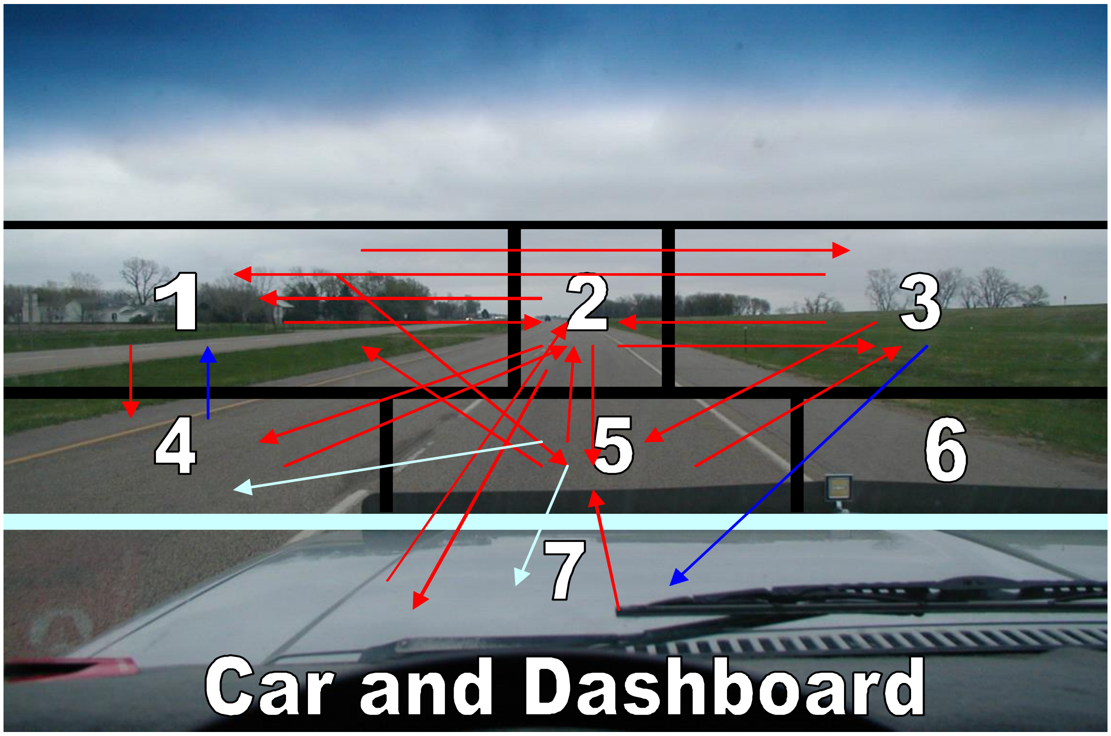


*Figure 35***.** A picture of what the "top 20" transition sequences looks like when mapped out for both the Young (Y1) and Old groups. The medium red lines represent transitions that both groups made. The dark blue represents transition sequences only relevant to the old. The light green represents transition sequences only relevant to the young.

One interesting observation that can be seen on Figure 35 is that the young had more glances originating from box 5 and the older groups had more transitions located at peripheral locations (3-7) and (4-1). There are 3 colors of arrows on this figure for the following reason: the top 20 sequences between both groups varied slightly. The young were the only ones to have a 5-7 and 5-4 sequence appear in the top 20, whereas the old where the only ones to have a 3-7 and a 4-1 appear. The remaining transitions were observed for both groups.

3.4.5 Saccade Amplitude Distributions

A SPSS GLM (2) age group by (3) loading task ANOVA of saccade amplitude yielded a significant loading task effect [f (2,24) =13.991, p<.001] and a significant loading task by age group interaction [f (2,24)=9.365,p<.002] for the young (Y1) versus Older participants data. Mean saccadic amplitudes are presented in Figure 36, while the overall distributions (in degrees) are presented in Figures 37 and 38.

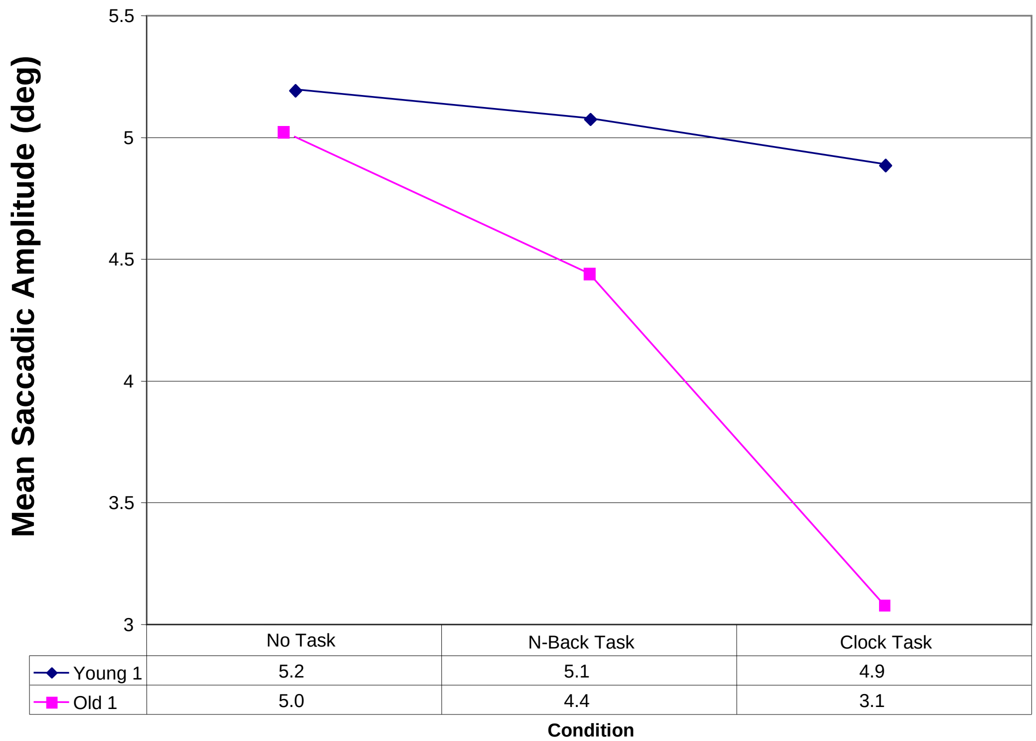


Figure 36. Mean saccade amplitude as a function of Age and Secondary Task Loading.

Paired sample t-tests of saccade amplitude by condition results showed significant differences in:

(a) 1-Back vs. Clock overall	(t(15) = -3.020,p<.01)
(b) Clock vs. No Task overall	(t(16) = -3.274,p<.006)
(c) Clock vs. No Task in O1	(t(5) = -6.201,p<.003)
(d) 1-Back vs. No Task in O1	(t(5) = -3.097,p<.03)
(e) *Between* groups in Clock	(t(5) = 2.571,*p*<.*05*)

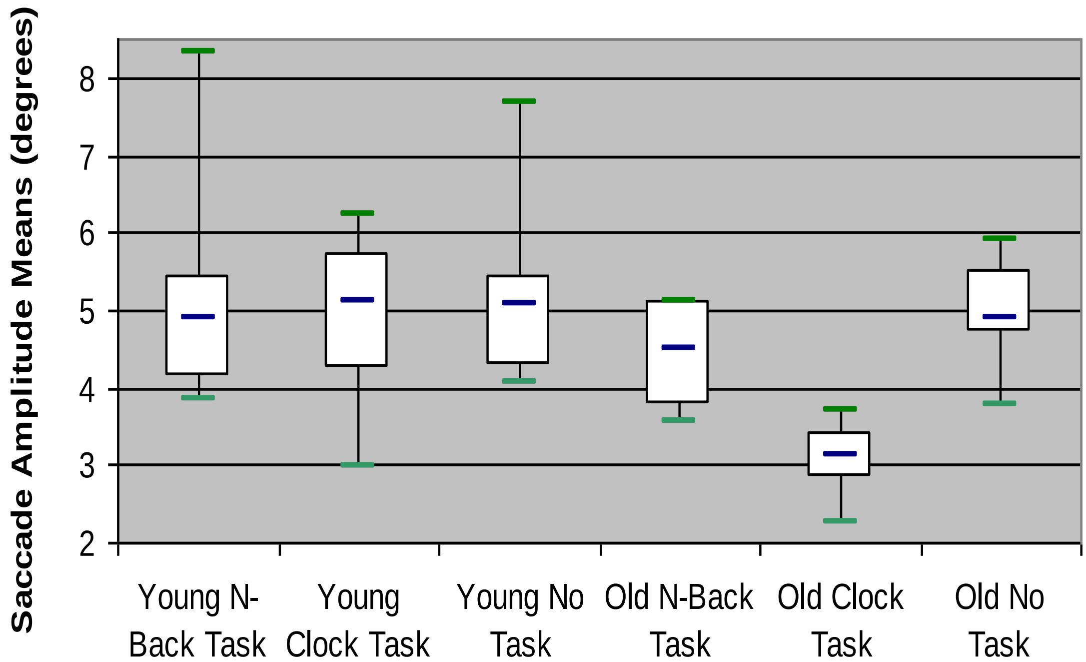


Figure 37. Standard box plots of saccade amplitude means. All are arranged by loading task and age.

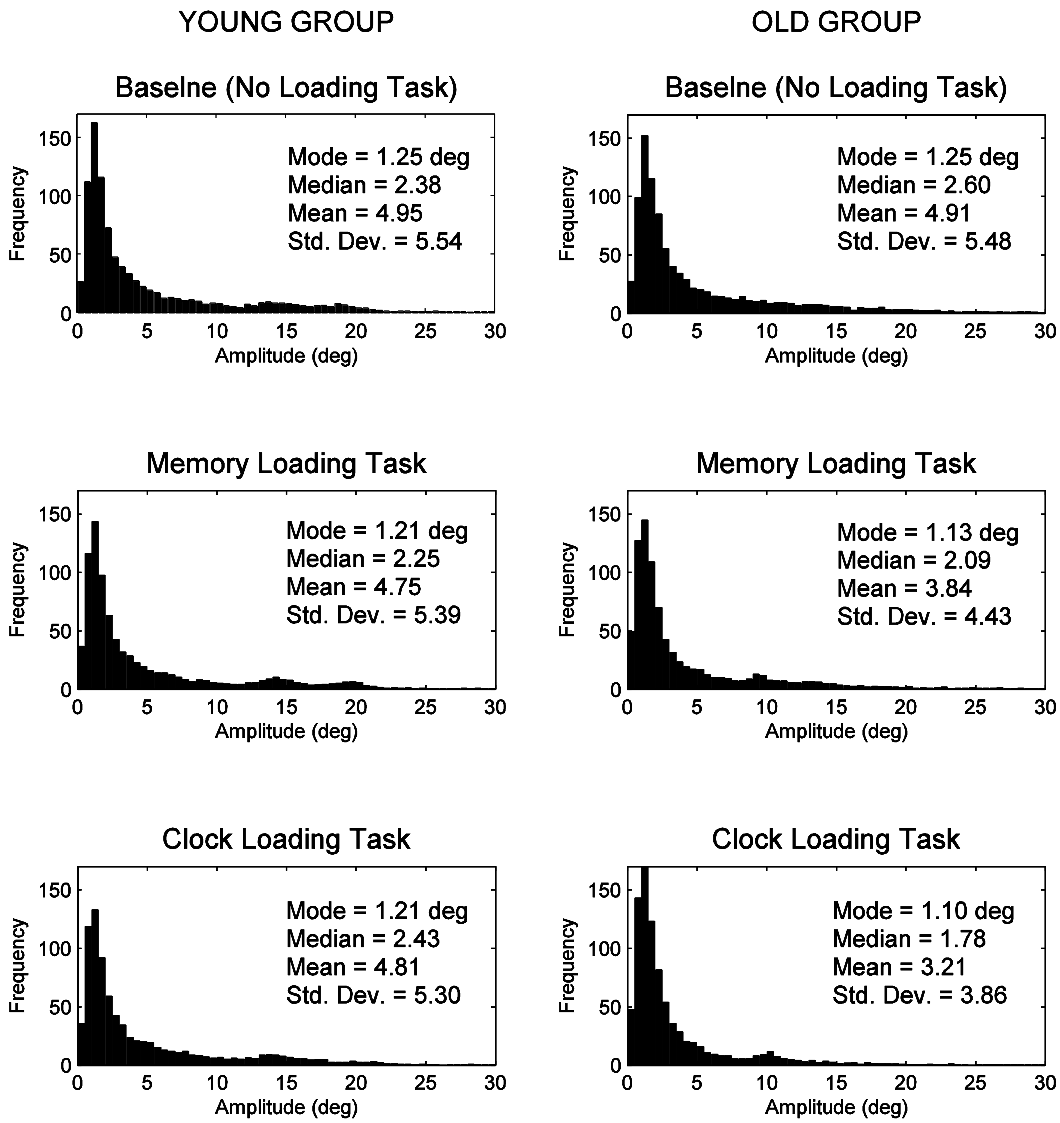


Figure 38. Saccade amplitude distributions as a function of Age and Secondary Task Loading condition.

3.4.6 Dwell Time Statistics

SPSS GLM (2) age group by (3) loading task ANOVA of fixation dwell time yielded a significant loading task effect [f (2,24) =8.367, p<.003] and a significant loading task by age group interaction [f (2,24)=4.783,p<.02] for the young (Y1) versus older participant data. Dwell time means are presented in Figure 39 and 40, while the actual dwell time distributions can be observed in Figure 41.

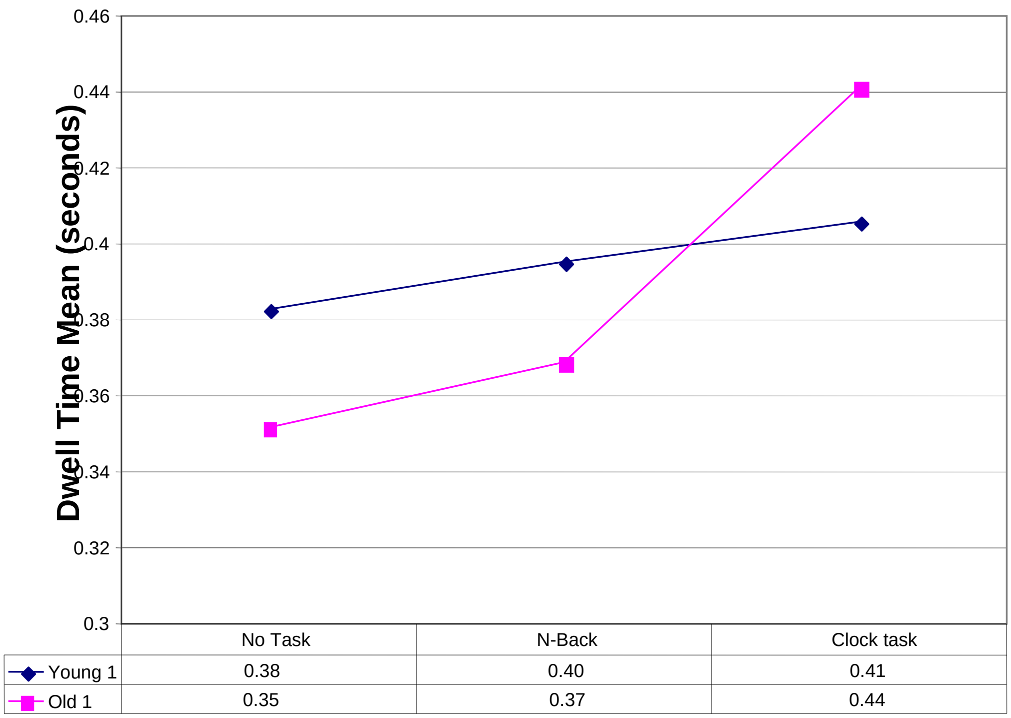


Figure 39. Mean fixation dwell time as a function of Age group and loading task conditions

T-tests of Dwell time means by condition results showed significant differences in:

(a) 1-Back vs. No Task as a whole (t(15) = 3.059,p<.009)
(b) Clock vs. No Task as a whole (t(16) = 3.354,p<.005)
(c) Clock vs. No Task in O1 (t(5) = 4.328,p<.009)

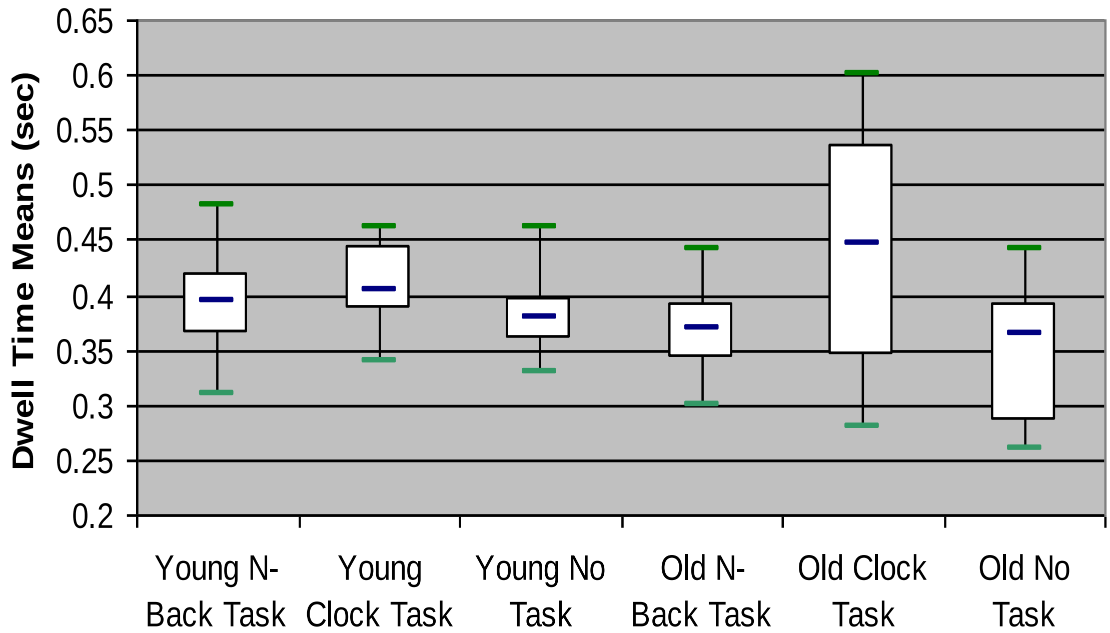


Figure 40. Box plot of dwell time means as a function of age group and task loading condition.

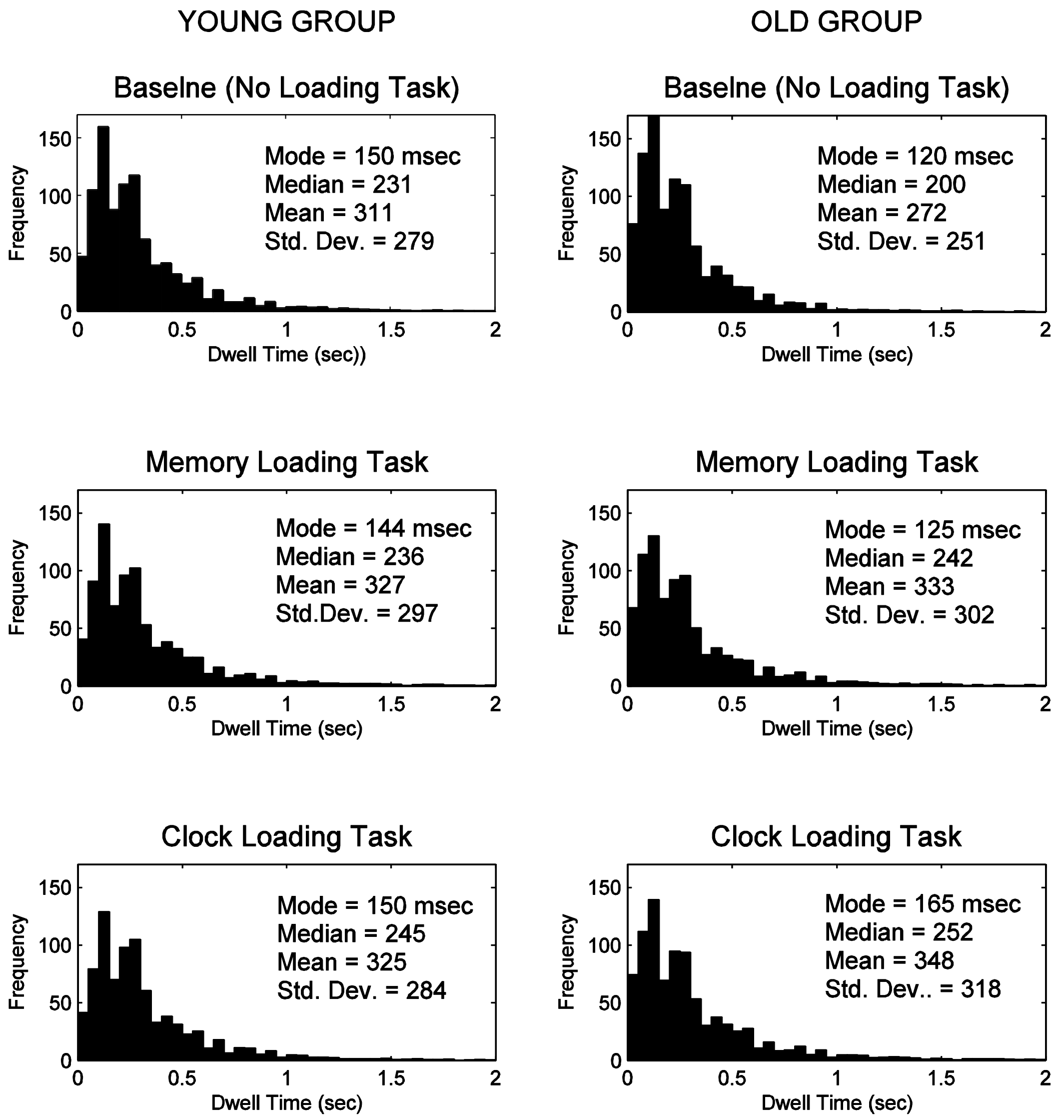


Figure 41. Fixation dwell Time distributions as a function of Age group and Secondary Task Loading condition.

3.4.7 Pupil Size

SPSS GLM (2) age group by (3) loading task ANOVAs performed on both mean and standard deviation of pupil size revealed no significant effects. Nonetheless, a general reduction in mean pupil size for the old group was observed (see Figure 42). This trend was expected due to the well known phenomenon of *senile miosis*. Unfortunately, there may have been a bias in this measure due to the way the ASL eye tracker computes pupil size. Since the eye scanning camera is stationary (as opposed to the more traditional head-mounted configuration), significant movements of the driver's head typically result in a geometric distortion in the effective shape of the pupil's observed shape. For example, the circular pupil transforms into an ellipse as the head is rotated from the left-to-right. As such, head rotations can yield a change in the measured pupil size when no such change in size actually occurs. Because of this potential bias, analysis of pupil size was restricted to the vertical diameter of the pupil. This approach should have minimized the effects of geometric distortions due to head movements since up-down head translations are much less frequent during driving than left-right rotations of the head.

However, there was no way to completely eliminate the co-variation of head movements and measurements of pupil size. Thus, the null findings regarding pupil size across the conditions of this experiment must be considered to be inconclusive.

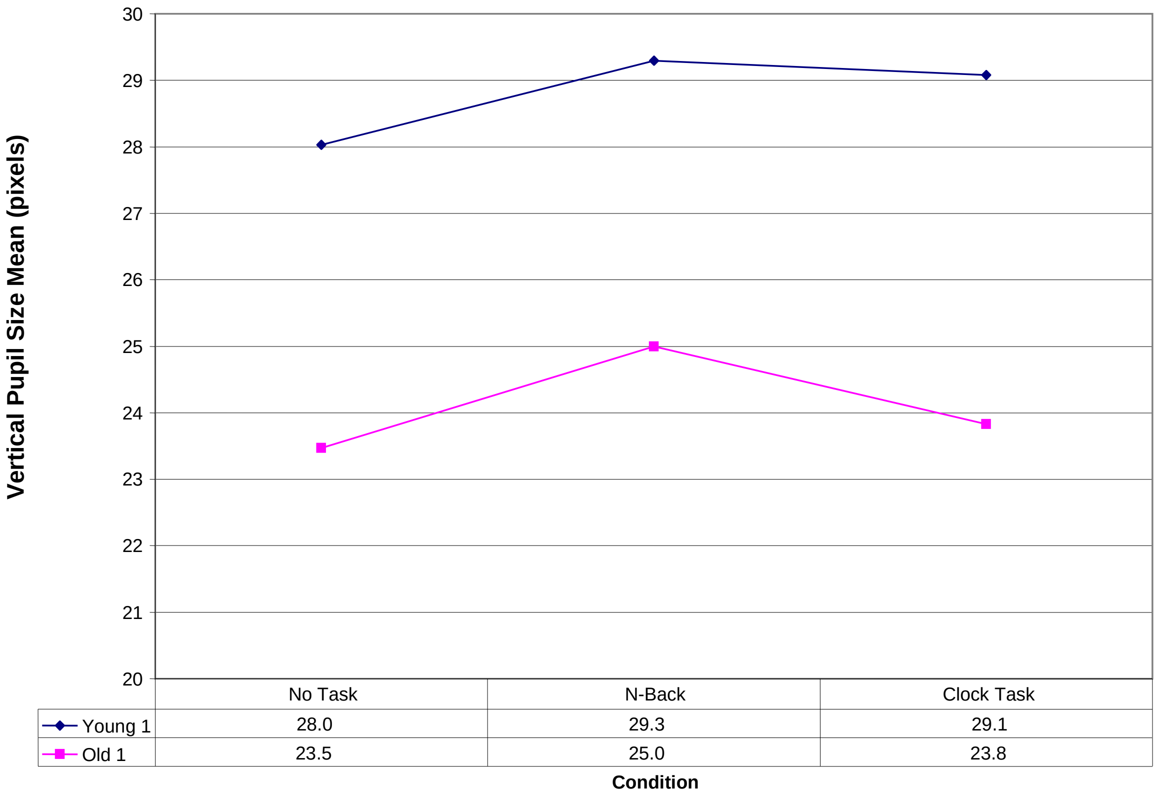


Figure 42. Mean pupil size as a function of age group and task loading condition.

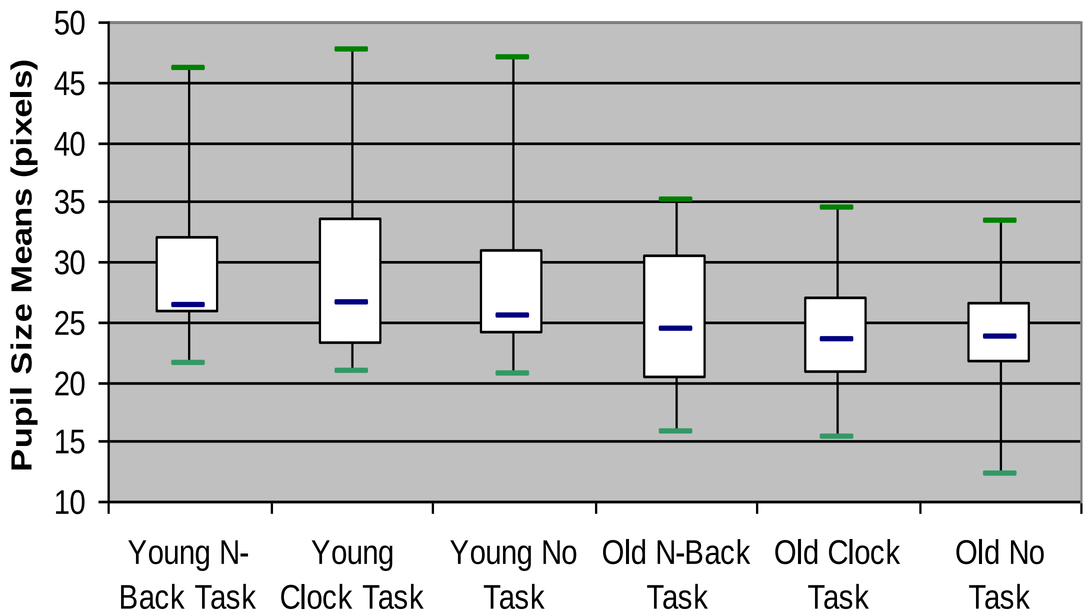


Figure 43. Box plot of pupil size as a function of age group and task loading condition.

3.4.8 Steering Entropy

Visual entropy is a relatively new concept and so we decided to contrast it to a more established metric that uses the same analytical technique: Steering Entropy. Since we already had the apparatus needed to measure steering wheel position, it was a logical step for validation of our paradigm. Using steering angle outputs from an absolute encoder mounted on the steering shaft, we did an analysis on steering entropy as defined by two well known leaders in the steering entropy discipline: Nakayama, et al. (1999) and Boer (2001, 2005). Both of these authors were previously able to elicit steering effects (*albeit in simulator studies*) while under conditions similar to this study. Since this was not covered in previous sections, a brief summary of the concept will be presented here before our results are examined.

Steering entropy is essentially a measure of the change in steering behavior as a function of time. Nakayama, et al. (1999) referred to this error correction as a change in steering "smoothness" or speed of error corrections. This has been demonstrated to reliably increase in value as cognitive load increases (Nakayama et al. 1999; Boer, 2001). This correlative link is believed to be attributed to the unequal sharing of attention (fewer cognitive "status" updates) which causes larger timing lapses between maintenance corrections. This is to say that if one monitors his heading *constantly* there is very little chance for a large heading correction being required because he is continually nullifying small errors as they occur. However, if one takes attention away, he or she will not be able to make these constant corrections resulting in an increased number of larger, quicker corrections being substituted.

Steering entropy measurement is similar to most other entropy measurements. In the steering paradigm, fluctuation in steering angle over time is the main variable. These fluctuations are directly linked to the maintenance of current heading which is the primary task when driving. Any task that takes attention away from this monitoring will lead to more time passing between update corrections. Since more time passes between corrections, they will tend to be larger and quicker than normal as each correction now needs to nullify a magnified amount of error in the same amount of time as it was previously done. As attention becomes increasingly taxed, the size of the steering corrections will undoubtedly increase, ultimately leading to a higher entropy value.

To determine the relative amount of error, the steering entropy value on the secondary task condition is compared to the size of the error values taken from a previous time series (e.g. Preceding X seconds or X number of data points). Given the assumption that all movements should be similar and fluidly connected, one can use a Taylor expansion to approximate what the next error value should be. For example, one would predict that since one has been turning 2° for the last 3 seconds that they will continue to turn at that rate in the next second. This would represent a low workload situation and be relatively predictable. Any errors that occur are small in magnitude. However, as driving complexity increases, steering adjustments tend to be more erratic which deviates further from the predicted hence yielding a larger error value. Instead of a steady 2° turn, one may now occasional observe 4°-5° spikes. It is this increased error deviation that is used to calculate differences between differing conditions. Every deviation from a "smooth" predicted will produce an error value that can be measured and compared.

More recent advancements in the steering entropy measurement paradigm have refined the ability to detect changes in entropy. Previous techniques sought mainly to identify high frequency deviations, however in a more recent paper written by Boer et al. (2005) it was shown that one could greatly increase power by sampling down to the 4Hz level which would detect low frequency deviations in addition to the high. The reasoning for this was that they observed several cases where different people had different expressions of entropy. Not all error corrections are instant; some are corrected over a longer time span than previously anticipated. While some individuals express a general entropy increase in all frequencies of movements, others may demonstrate increases in only one domain but not the other. In other words, a subject might show an increase in high **or** low frequency movements, but not always in both.

To sum up the steering entropy idea: increased secondary workload causes the need to borrow attention from heading monitoring (steering). Fewer heading updates means more extreme heading corrections. Using more recent models of steering entropy we capture steering corrections existing in two forms: high frequency quick corrections and low frequency slow corrections. By accounting for both types of corrections, a more complete composite of error correction for the steering task should be observed.

Unfortunately, neither of the two main steering entropy approaches yielded significant results across our experimental conditions. Our best explanation for these null outcomes is that several conditions in this experiment could potentially have combined to reduce the sensitivity of the steering entropy metric. The participants drove on a fairly narrow road without shoulders and at high speeds. Under these conditions one cannot afford the opportunity to “neglect” lane positioning. Steering entropy is essentially a measure of this neglect so its absence would be decisive. It is possible that with wider lanes and/or lower speeds, systematic changes in steering entropy could still be observed under these conditions. Nevertheless, the graphs of these results can be seen in the next few figures.

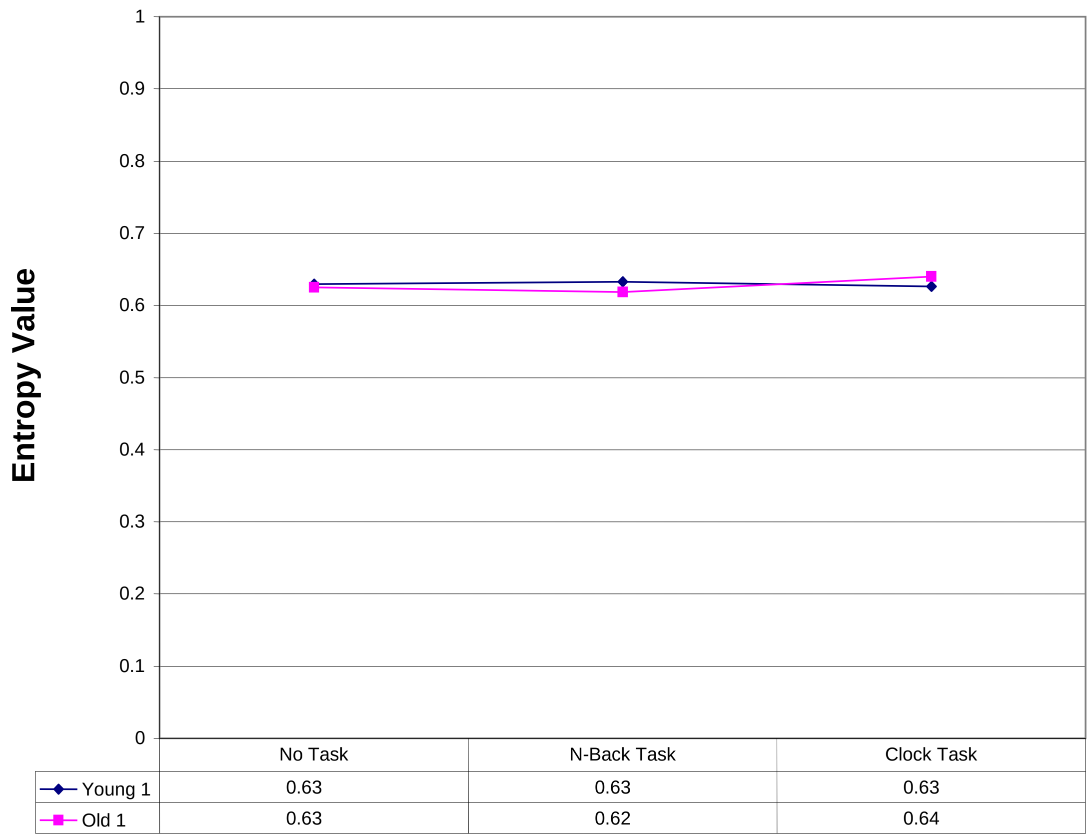


Figure 44. Nakayama steering entropy as a function of age group and task loading.

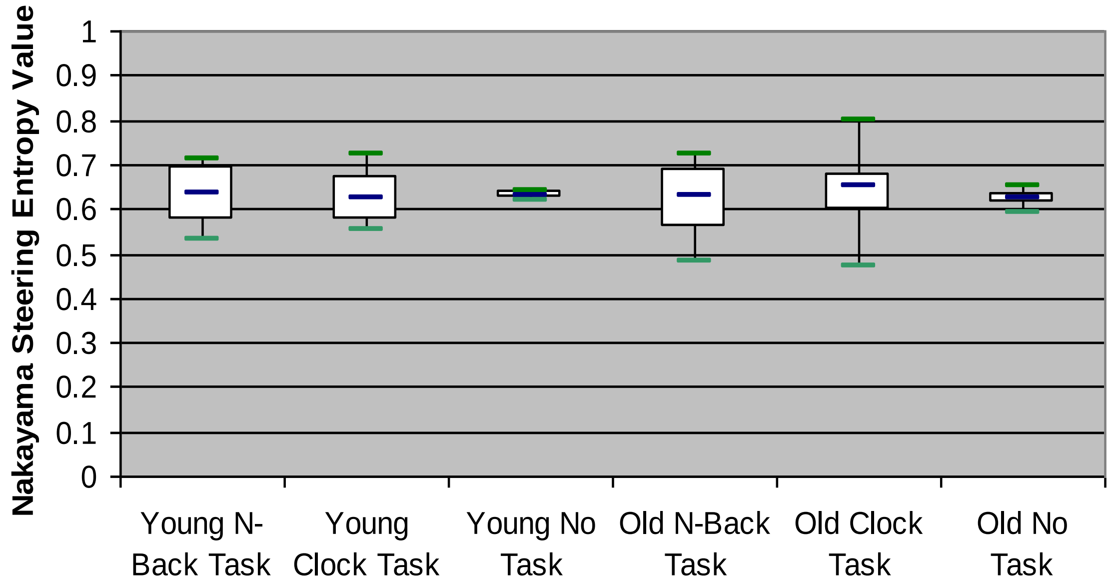
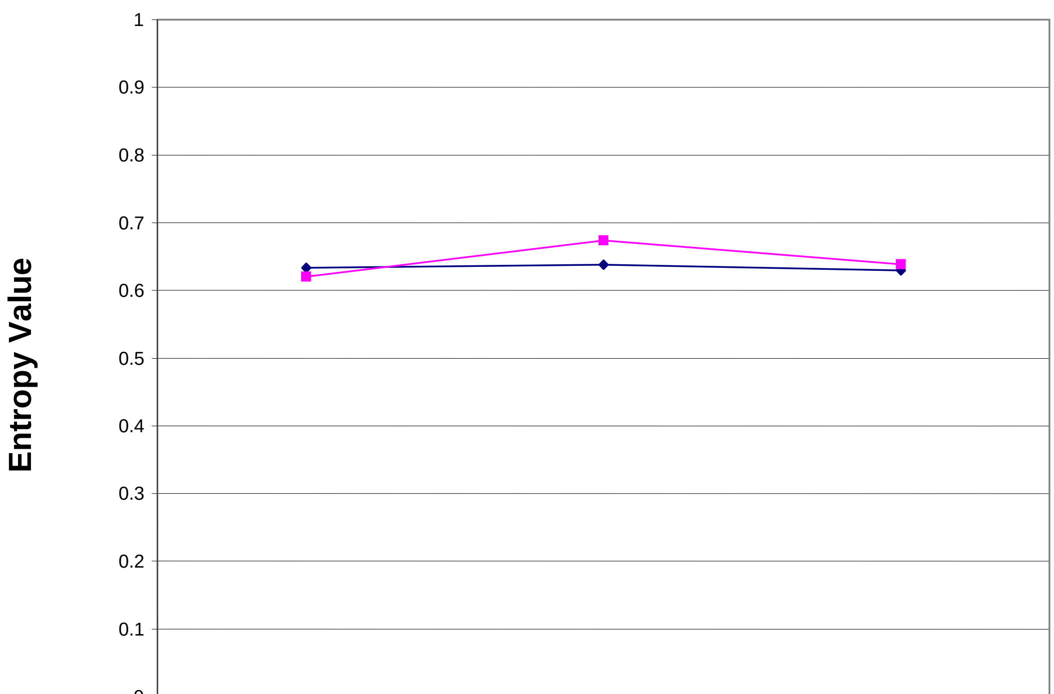


Figure 45. Standard box plots of Nakayama Steering Entropy metrics as a function of age group and task loading



	No Task	N-Back Task	Clock Task
Young 1	0.63	0.64	0.63
Old 1	0.62	0.67	0.64

Figure 46. Boer steering entropy as a function of age group and task loading.

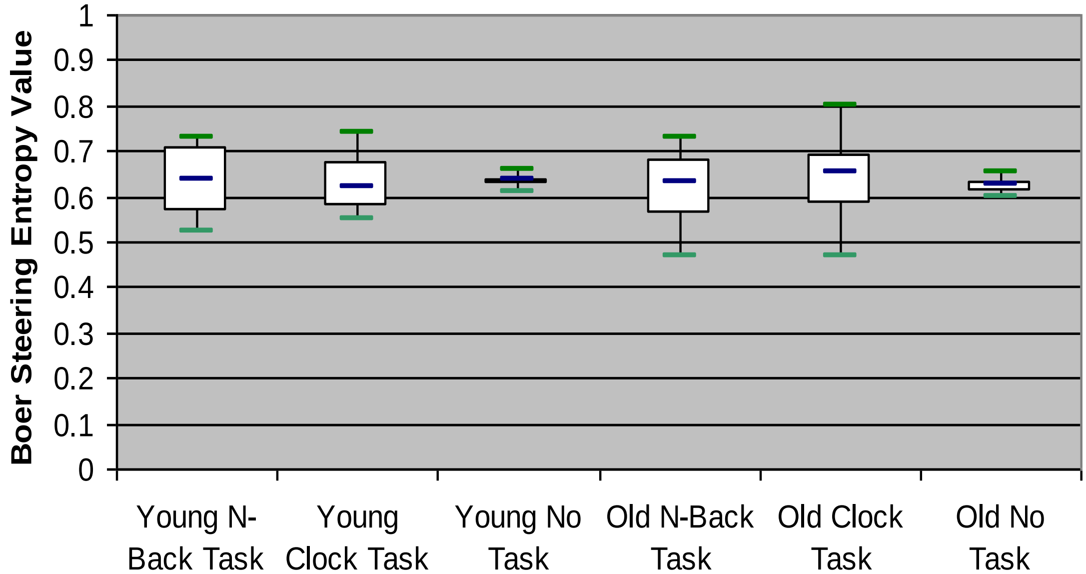


Figure 47. Standard box plots of Boer Steering Entropy values as a function of age group and task loading

3.4.9 Vehicle Yaw Rate Variation

SPSS GLM (2) age group by (3) loading task ANOVA of the standard deviation of vehicle yaw rate revealed a statistically significant loading task effect [f (2,44) = 3.812, p<.031] and a loading task by age group interaction [f (2,44)=3.751,p<.032] for data collected from the young (Y1) and older participants.

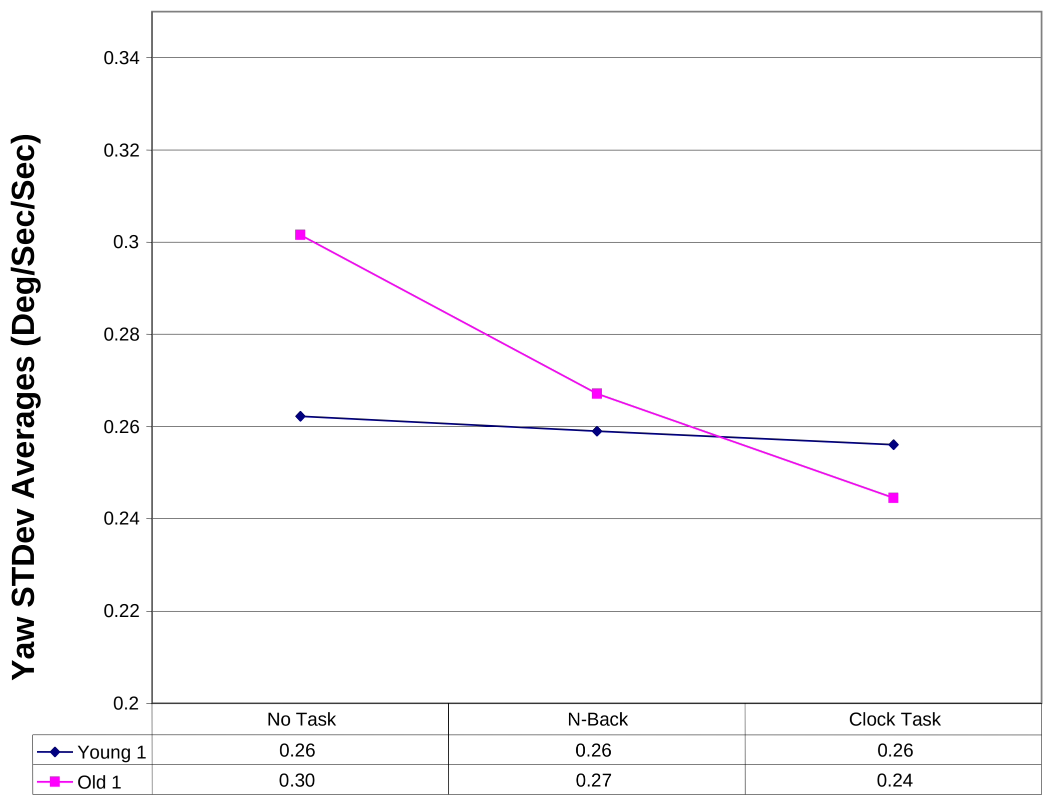


Figure 48. Standard deviation Averages of vehicle yaw rate as function of age group and loading task conditions. (Conversion factor of .003 applied to raw data to derive above results)

Post hoc t-tests revealed significance differences between:

(a) 1-Back vs. No Task in O1 (t(13) = -3.049,p<.01)
(b) Clock vs. No Task in O1 (t(11) = -3.594,p<.005)
(c) *Between* groups in No Task (t(11) = -2.204,p<.05)

No significant difference was noted in the young group across conditions.

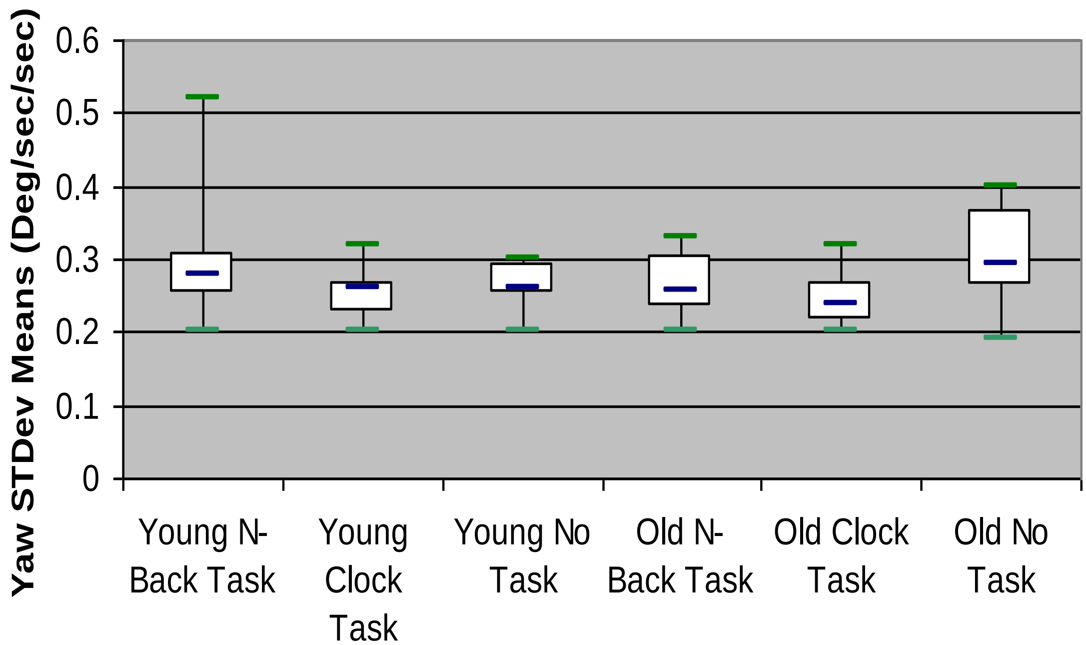


Figure 49. Box plot of standard deviation of vehicle yaw rate as a function of age group and loading task conditions.

3.4.10 Driving Speed

A SPSS GLM (2) age group by (3) task loading ANOVA of mean driving speed yielded a significant loading task main effect [f (2,44) = 8.648,p<.002] with a marginal loading task by age group interaction [f (2,44) = 3.035, p=.058] for the young (Y1) versus older participants. There was also a highly significant between group effect [f (1,22) = 20.405, p <.001] observed for speed means between these two groups. When looking at standard deviations, a significant task effect [f (2,44) = 6.334,p<.004] and between group effect [f (1,22) = 7.457, p=.012] are again noted, but the interaction is no longer significant.

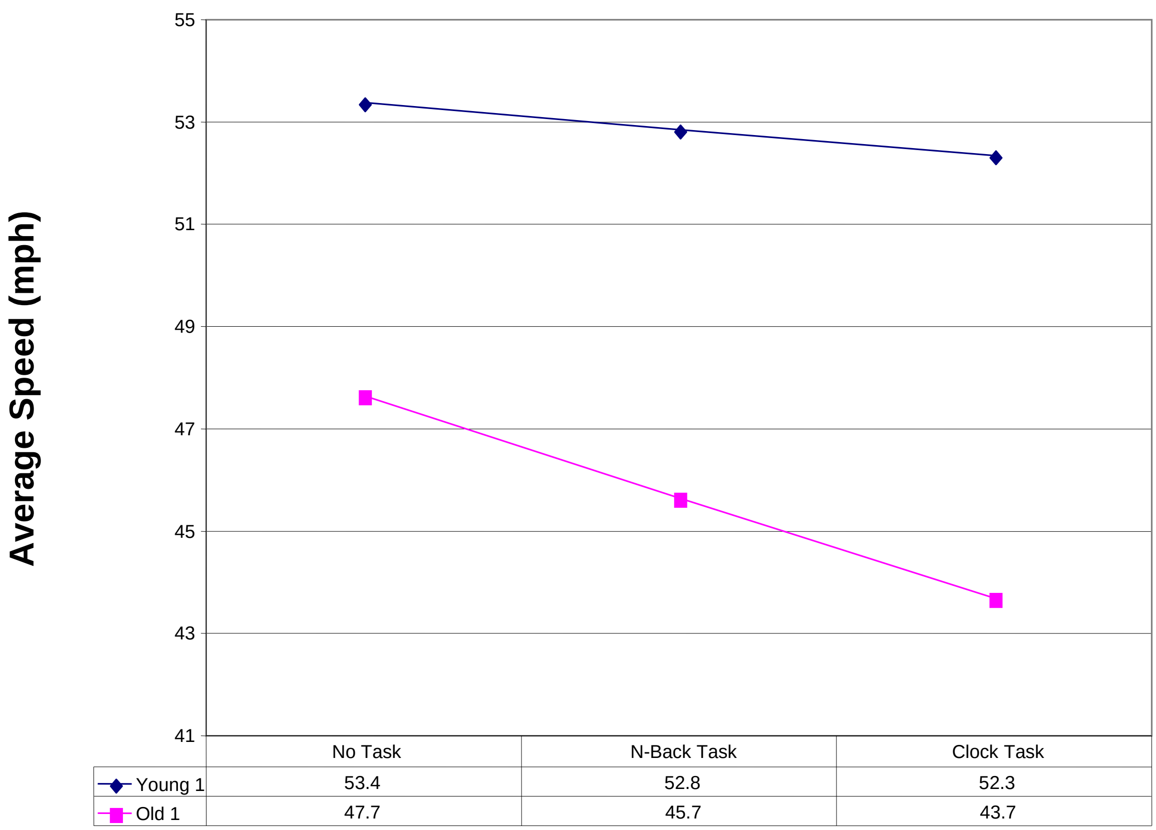


Figure 50. Average Speed by Condition.

Post hoc t-tests of average speed revealed the following significant differences:

(a) 1-Back vs. Clock Overall	($t(23) = 2.576, p<.017$)
(b) Clock vs. No Task Overall	($t(23) = -3.541, p<.003$)
(c) 1-Back vs. Clock in O1	($t(11) = 2.818, p<.018$)
(d) Clock vs. No Task in O1	($t(11) = -3.285, p<.008$)
(e) Between Groups in 1-Back	($t(12) = 3.994, p<.002$)
(f) Between Groups in Clock	($t(10) = 4.249, p<.002$)
(g) Between Groups in No Task	($t(11) = 4.028, p<.002$)

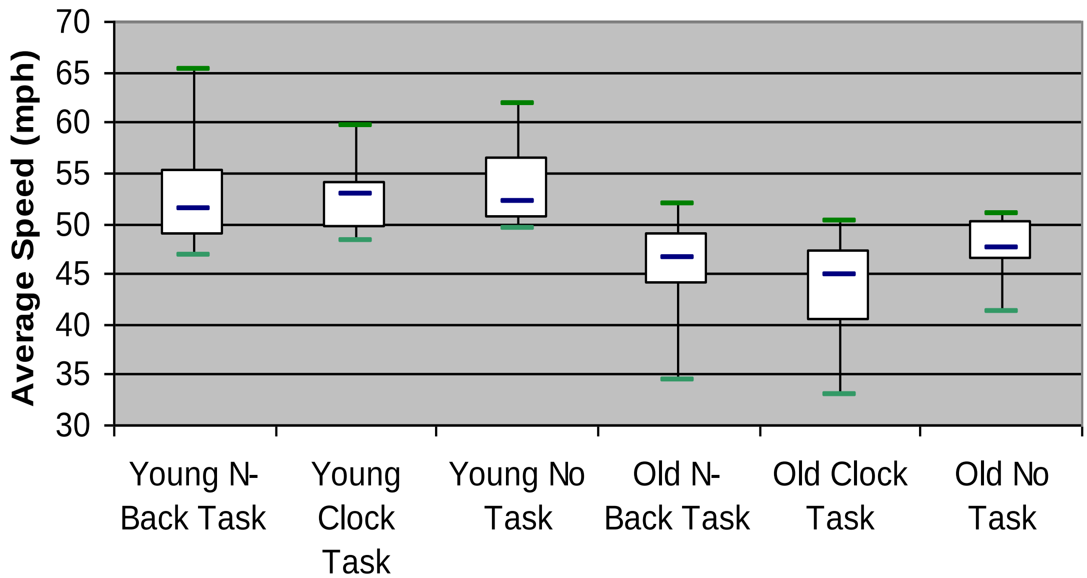


Figure 51. Box plots of average driving speed as a function of age and loading task.

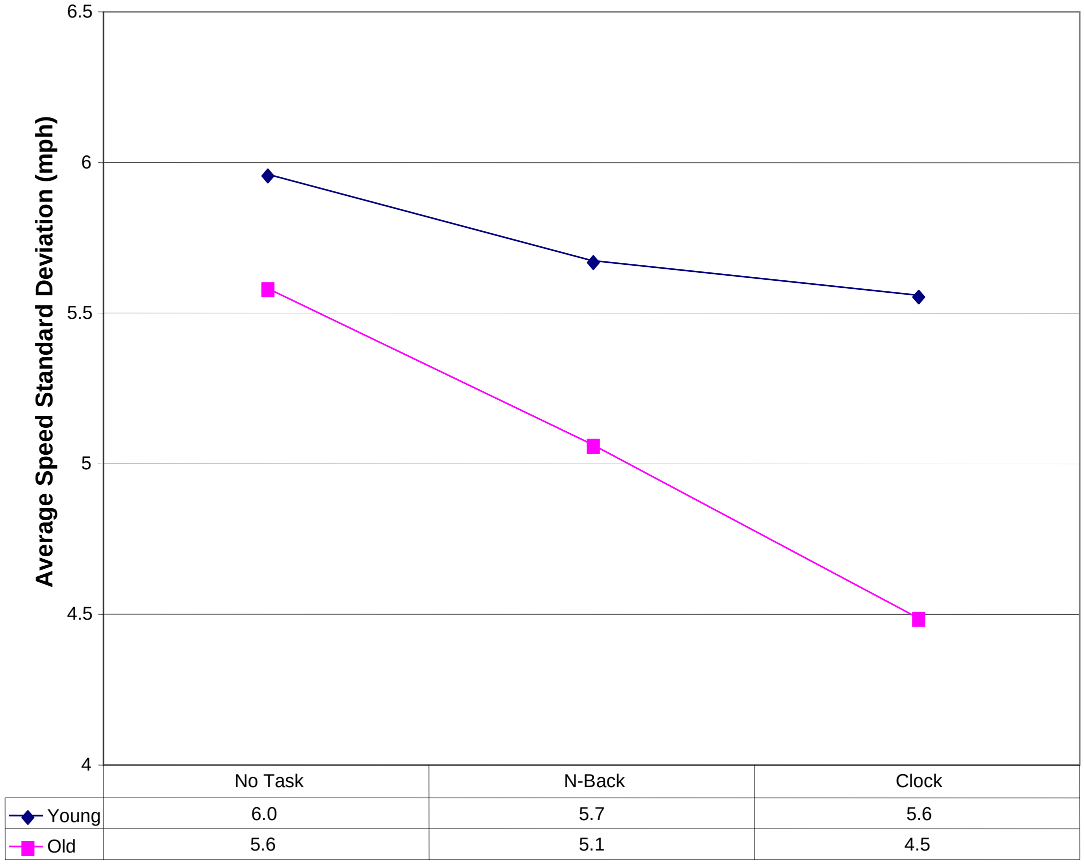


Figure 52. Average Speed Standard Deviation by Condition

Post hoc t-tests of speed standard deviations revealed the following significant differences:

(a) 1-Back vs. Clock task in O1 $(t(11) = 2.418, p<.034)$
(b) *Between* Groups in Clock task $(t(10) = 3.782, p<.004)$

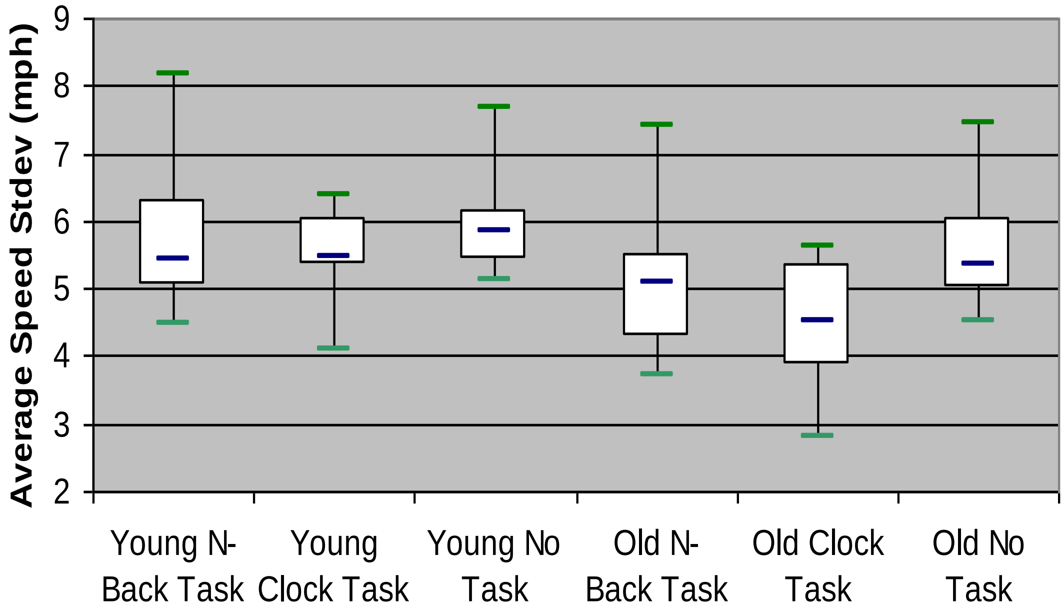


Figure 53. Box plots of Speed standard deviation averages as a function of age and loading task.

3.4.11 *Loading Task Performance Accuracy*

In order to assure that subjects were performing the secondary loading tasks as intended, task accuracy measures were taken. A SPSS GLM (2) age group by (2) loading task ANOVA of mean accuracy performance yielded a significant loading task effect [f (1,22) = 33.441, p<.001], loading task by age group interaction [f (1,22) = 9.392,p<.007], and a between group effect [f (1,22) =10.034,p=.004] for the young (Y1) and older participants. When comparing the two young groups (Y1 against Y2) a significant loading task effect [f (1,20) = 5.026,p<.04] was observed but no between group difference was found.

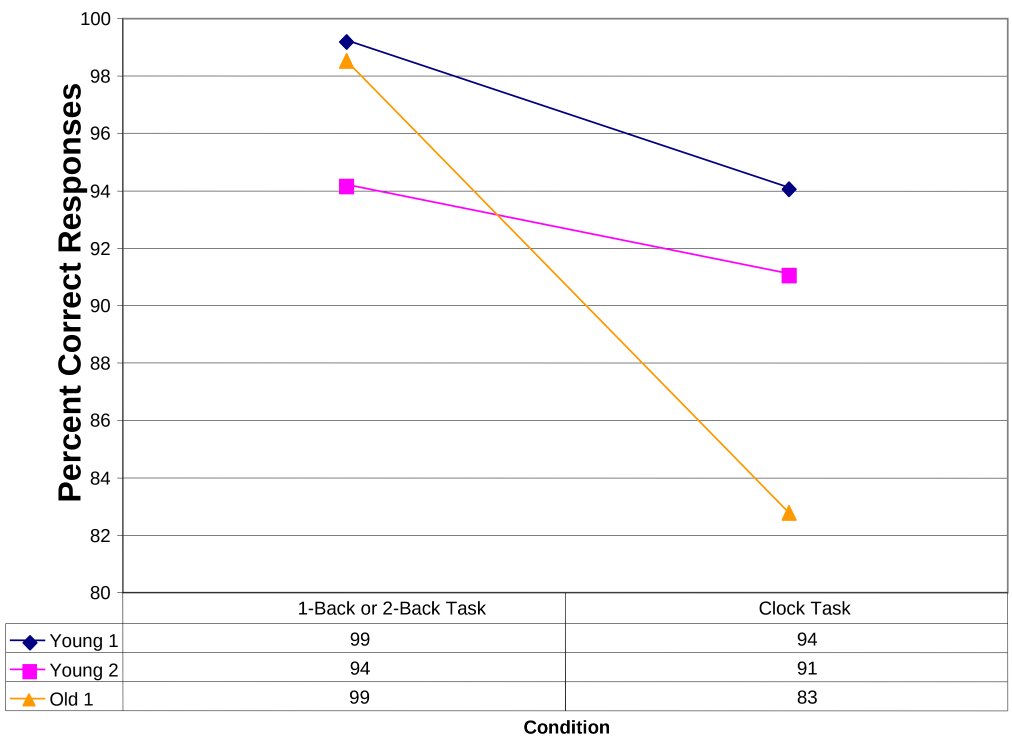


Figure 54. Secondary task accuracy as a function of group and loading task.

Post hoc t-tests revealed the following significant differences:

(a) 1-Back vs. Clock Task in Y1 (t(11) = 4.308,p<.002)
(b) 1-Back vs. Clock in O1 (t(11) = 4.662,p<.002)
(c) *Between* groups Y1/O1 in Clock Task (t(10) = 3.074,p<.013)

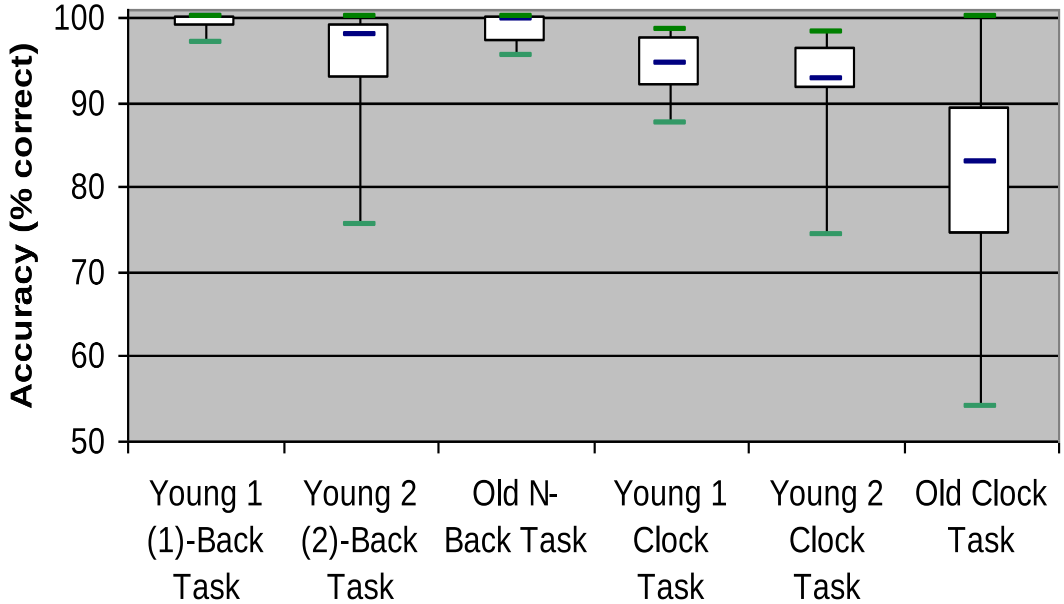


Figure 55. Box plots of secondary task accuracy as a function of group and secondary loading task.

For reference, a summary of all statistically reliable findings has been included in table 9.

		Young Group 1			Young Group 2			Old		
		No Task	1 - Back	Clock	No Task	2 - Back	Clock	No Task	1-Back	Clock
Young Group 1	No task	------	Vis Ent 1 Vis Ent 2					Yaw Rate Speed		
	1 - Back	Vis Ent 1 Vis Ent 2	------						Speed	
	Clock			------						Vis Ent 1 Vis Ent 2 Sac. Amp Speed
Young Group 2	No Task				------	Vis Ent 1 Vis Ent 2	Vis Ent 2			
	2 - Back				Vis Ent 1 Vis Ent 2	------				
	Clock				Vis Ent 2		------			
Old	No Task	Yaw Rate Speed						------	Sac. Amp Yaw Rate	Vis Ent 1 Vis Ent 2 Sac. Amp Dwell Yaw Rate Speed
	1- Back		Speed					Sac. Amp Yaw Rate	------	Vis Ent 2 Speed
	Clock			Vis Ent 1 Vis Ent 2 Sac. Amp Speed				Vis Ent 1 Vis Ent 2 Sac. Amp Dwell Yaw Rate Speed	Vis Ent 2 Speed	------

Table 9. Summary of findings: matrix of conditions observing statistical significance (p<.05) between situations and groups.

4.0 Discussion

4.1 Visual Entropy Metric

Entropy is a metric derived from information theory that provides a conceptual framework to mathematically describe the complexity of any uncertain system. This metric's utility for eye movements was first established in a study by Tole, et al. (1983) evaluating instrument panel gaze pattern changes in cognitively loaded *pilots*. In the current study, a similar entropy metric was used to quantify changes in visual scanning complexity in *automobile* drivers. Changes were categorized by variations in cognitive resource demand and driver age. The entropy metric proved a useful and powerful tool for differentiating age and task situations using eye-movement data.

Upon examination of the results, the first main observation was that the overall scanning complexity did not substantially differ between young and old drivers in either the baseline or verbal-memory (i.e. n-back task) conditions. Initially this was somewhat surprising as a global decrease in complexity across all conditions was predicted for the old group compared to the young. This was expected due to maturational decrements in mental efficiency inherently related to age (Korteling et al., 1990; 1991). Since an older adult has a much lower attentional threshold to begin with, performance decrements should be observed on all tasks. In actuality, the baseline and verbal task were not reliably separable between groups. Although entirely speculation, there is a possible explanation for this finding. It is plausible that the enormity of driving experience possessed by older driver allows him or her to compensate for any attentional interference since both these test situations involve highly over-learned activities. On average, any given subject from the older group had approximately 48 more years of driving experience than their younger counterpart. Over a timeframe this large, it can be assumed that an older driver has driven more miles and had many more conversations while driving than a younger subject. In section 1.3.3, several studies summarized just how influential the effect of experience can be.

The visual-spatial condition however, was a different story. A dramatic divergence was witnessed as a function of age when looking at the visual-spatial (i.e. clock task) condition. When engaged in the visual-spatial task, young drivers demonstrated only a marginal decrease in 1^{st}-order eye movement complexity (5%), while the older individuals engaged in the same task, demonstrated a considerably greater reduction in complexity (25%) compared to the baseline condition. This result suggests that during the baseline and verbal-memory conditions, all drivers had sufficient "spare visual capacity" to actively explore the visual environment. However, during the loading imposed by the visual-spatial subsidiary task, older drivers appeared to lack the spare capacity required to simultaneously maintain complex levels of visual scanning behavior.

Even though "experience" could help explain some of the between-group discrepancies, there are other portions of the results that are obviously independent of experience (such as those seen *within* each group). These within group results follow the logic from another common perspective: multiple resource theory (MRT). Considered from within the context of Wickens' (1984) multiple resource theory, the results of this study are consistent with an increased interference effect between similar pools of attention (visual spatial task interfered more with visual behavior than the verbal task did). In this respect, since the tasks rely on functionally different pools of attentional resources, they interact with visual scanning differently. As the attentional load on the visual modality increases due to the imposition of the visual-spatial subsidiary task, the ocular scanning patterns became impoverished so as to include only what was “necessary” to maintain vehicle heading and a minimal state of situation awareness. It seems as though the performance of the “visual spatial task” was interfering the most with the visual-spatial ability of scanning the roadway environment – especially among the older drivers. This magnitude of effect was not observed while engaged in the verbal memory task. Again, within the context of multiple resource theory, this interference occurred because the clock task “competes” for the same type of attentional resources required to support complex ocular scanning behavior. The visual-spatial subsidiary task essentially relied on a resource that was already heavily engaged due to the act of driving and therefore more prone to conflict.

The older subjects in this study were much more sensitive than their younger counterparts to a cognitive load induced by the visual-spatial condition. This age difference was striking in nearly all the different metrics evaluated (both ocular and non-ocular). In the visual-spatial condition 75% of all fixation sequences could be represented by only 4 distinct transitions states among the older drivers. These 4 transitions are as follow:

(a) 2-3 : Far down road transitioning to far left ditch (21%)

(b) 3-2 : Far left ditch transitioning to far down road (18%)

(c) 5-2 : Far down road transitioning to the front of the car (18%)

(d) 2-5: In front of the car transitioning to far down the road (18%)

However in the younger drivers, only up to 55% of the visual scanning behaviors could be defined by any 4 combinations of transition state sequences. In addition to fewer behaviors occurring in large proportions, the main 4 combination types are also reordered in the young group as follows:

(a) 2-3 : Far down road transitioning to far left ditch (17%)

(b) 3-2 : Far left ditch transitioning to far down road (17%)

(c) 2-1 : Far down road transitioning to the far right lane (11%)

(d) 1-2: Far right lane transitioning to far down the road (10%)

Instead of spending the majority of time looking at the left ditch and the road directly in front of the car (both near and far) like the older subjects, the younger group expanded their search to also include the far right side, etc. The visual gaze patterns in the older group suggest that a behavioral coping strategy is occurring during the visual spatial task. Again, from a MRT standpoint, this is occurring due to attentional interference between the task and driving. As attentional interference increases due to the visual spatial task's competition for visual attention, the older subjects respond by narrowing the fixation patterns to only the most relevant areas of the road absolutely needed for safe navigation such as:

1. *Potential hazard areas* like the left ditch. It is speculated that the right lane was not fixated as much as it was in the young because experience had taught the older drivers that while a car in the right lane is always a hazard, it is not as prone to coming into their lane as is oncoming traffic, etc. from the left side. If it is not as prevalent, then it should not need to be scanned as frequently.
2. *Future heading of the car* (i.e. look at point of optical expansion down the road). This is important because at 55mph, hazards and road changes in the

distance require immediate recognition in order to plan an avoidance course in time.

3. *Current road directly in front of the car*. This is important for lane maintenance of the car.

Although both groups' scope of *fixation* locations (see table 10) are somewhat similar and agree with that anticipated based on previous studies (see figure 4 or Serafin, 1994), the older group's *transitional* space was much more restricted than the young in the visual-spatial condition. When looking at the proportion of fixations occurring in each location on table 10 one will notice that in both the no-task and verbal conditions, both ages have a very similar scope of fixations dedicated to each "state space". However, when progressing to the visual spatial condition, the whole fixation cluster seems to shift up and left in the old, whereas the young still remains similar to the previous conditions. This shift is ultimately reflected as fewer state space transitions because the confined visual range limits the opportunity for a transition to occur.

11%	53%	17%	4%	62%	16%
1%	10%	1%	1%	15%	1%
7%	Young Clock Task		1%	Old Clock Task	
10%	53%	16%	8%	56%	16%
2%	9%	1%	2%	13%	1%
9%	Young N-Back		5%	Old N-Back	
10%	49%	18%	9%	46%	22%
2%	11%	1%	2%	14%	1%
9%	Young No Task		6%	Old No Task	

Table 10. Percentage of total *fixations* occurring in each state space (state spaces adapted from figure 21) separated by age and condition. Each number represents the approximate percentage of all fixations occurring in that state space during that particular trial type. This does not represent *transitions* between spaces only raw fixation location occurrences. Young participant trials are on the left side, older participant trials are on the right side.

Another interesting observation to come from this study involves the number of *overall* state space transition combinations represented (i.e. state 1 to 3, state 3 to 7, state 7 to 6, etc.). Older drivers used 14% fewer transition combinations in the visual-spatial condition than in the baseline condition, whereas younger drivers only used 2% fewer combinations across the same conditions (see figure 34 for graphical comparison). Again, this is consistent with more attentional resources being available in the young as only moderate interference of the visual-spatial task is observed (as indicated by very little decrease in overall visual sampling compared to baseline).

Overall, these findings are consistent with previous studies suggesting a reduction in available visual resources with age (see section 1.7), but how do the current findings regarding eye movement complexity map to the *a priori* experimental hypotheses derived from these earlier accounts in the scientific literature?

1. The *first hypothesis* stated that driver eye movement complexity should decline in the presence of any subsidiary loading task among young and older participants alike. This hypothesis assumes a single-resource model of attention. Little support for this simple hypothesis could be found given the analysis of the complexity data. Although the main effect of subsidiary task was statistically significant, subsequent analyses revealed that this effect could be accounted for by the complex age by subsidiary task interaction effect (see #4 below).
2. The *second hypothesis* stated that driver eye complexity should decrease more during the performance of the visual-spatial than the verbal-memory subsidiary task. This expectation was derived from multiple-resource theory given the assumption that the "clock task" and driver oculomotor scanning of the environment were expected to compete for the same pool of (visual-spatial) resources; while the verbal-memory task would draw resources from a separate and independent pool. Little support for the most general form of

this hypothesis could be found in the behavior of the young drivers. However, as will be seen in the paragraphs that follow, such an exaggerated interference effect of the clock task upon visual scanning behavior was found among the older drivers.

3. The *third hypothesis* stated that visual entropy should decrease as a general function of increased driver age independent of subsidiary task conditions. No evidence for such a general effect of age was observed. On the contrary, there was a small but non-significant elevation of eye movement complexity observed among the older participants in the baseline driving condition (see Figure 22).

4. The *fourth hypothesis* was the prediction of greatest theoretical as well as practical interest in this investigation. It was hypothesized that the complexity of eye movement behavior would decline disproportionately among the older drivers in the most visually demanding circumstances; namely, during performance of the visual-spatial subsidiary loading task. There was strong statistical support for confirming this hypothesis. The age by subsidiary task interaction was highly significant; with no significant age difference observed for the verbal-memory task but a sizable and significant age-related decline in visual complexity under the visual-spatial loading condition. However, despite this strong support for the fourth experimental hypothesis, the mechanism mediating the predicted pattern of results is less forthcoming.

Upon first glance, the results of this study appear consistent with the idea that fewer visual resources were available to the older drivers. However, there were additional findings that argue against this “strong” multiple-resource interpretation of the cognitive aging process. Specifically, these findings were those that support a “difficulty” explanation such as:

1. Young Group 2 - Verbal Task Control Results
2. Accuracy Results

The majority of the counterargument stems from a consideration of the data collected from the control group of young subjects (Young Group 2) who were given the 2-back version of the verbal-memory subsidiary task. Young Group 2 was given a more difficult version of the N-back task in an experimental attempt to disentangle the effects of *task difficulty* and *resource-specific competition*. The logic behind the inclusion of this group was as follows: It was expected that age-differences in eye movement complexity would be small for the verbal-memory task condition but much larger under the visual-spatial task condition. Yet, if such a predicted difference were observed, at least two competing explanations would be possible: (a) older drivers had less available visual resources to dedicate to both the clock task and scanning of the roadway environment; or, (b) the increased interference with eye scanning behavior in older drivers resulted for an increase in the *structural difficulty* of the clock task (compared to the verbal task) rather than its reliance upon visual resources, *per se*. The 2-back version of the verbal-memory task was used with Young Group 2 because it was more difficult than the 1-back version of the task. If the eye movement complexity observed for Young Group 2 under the 2-back verbal-memory task was less than that demonstrated while performing the clock task, then it would follow that *task difficulty* rather than visual resource depletion (as predicted by multiple-resource theory) might be mediating the expected age by task load condition interaction (Hypothesis 4). Indeed, Young Group 2 demonstrated a small but statistically significant decrease in visual entropy relative to their own performance under the visual-spatial subsidiary load task as well as compared to the performance of Young Group 1 in the 1-back version of the verbal memory loading task. This finding complicates the interpretation of the significant age by loading task interaction predicted by Hypothesis 4 (see above). However, the small size of the difficulty effect (approximately 5%) observed for the 2-back verbal-memory task across Young Groups 1 and 2 does not, in and of itself, nullify the aforementioned conclusion that older drivers exhibited a *measurable* decrease in visual attentional resource capacity. In fact, the very large size of the reduction in visual entropy (25%) observed among the older drivers during the concomitant performance of the clock task still provides a strong argument for this claim.

The second argument that slightly weakens the multiple-resource interpretation forms around the accuracy results. A statistically reliable reduction in accuracy in *both age groups* between conditions suggests again that the clock task may simply be more difficult. The young group drops 5% in accuracy between the verbal task and the visual spatial task whereas the old group drops 11%. *However*, when comparing results between groups, it is observed that accuracy levels are statistically equivalent during the verbal task. It is not until the visual-spatial condition that one can observe differences *between* groups. Again, if the verbal task is the same difficulty for both an older participant and a younger participant, there is no reason to believe that the visual spatial task should be anymore difficult for the older subject from a strictly difficulty perspective. This suggests although difficulty may be a contributor, it is not the entire case.

In summary, both cognitive tasks produced some degree of interference with the visual scanning system. This interference was the most perceptible during the visual-spatial task especially when older subjects were involved. The strength of the argument for modality-specific competition as a primary cause is slightly weakened by the results of the young 2 control group and accuracy results. However, there is still some theoretical strength for the main argument due to the sheer magnitude of shift in an older individual's visual-spatial task entropy values.

Due to the unanticipated results of young group 2, future studies should consider alternate means for distinctly defining visual-spatial *difficulty*. The argument for separating "resource interference" from "difficulty" was less powerful than anticipated and will remain an issue faced by all future studies. A possible solution is to create several "resource-based" tests using correlated levels of difficulty to balance between the tasks by adjusting difficulty level individually for each participant. Alternatively, one could include multiple levels of difficulty within each resource category of subsidiary load task. For example, if one has 10 well-defined difficulty levels for both a visual-spatial and verbal task, he or she should be able to compare the performance results of each task to define which levels of each task produce similar results. Once a more reliable control on difficulty level has been established, it should be a much easier to conclude that any change contrasting two "difficulty-matched" tasks should be due to something other than difficulty (such as an attentional competition or deficits).

4.2 Other Ocular Metrics

More conventional parameters used to describe distributions of eye movement behavior were also examined in this study. It was observed that mean saccadic amplitude and fixation dwell times varied systematically with manipulations of subsidiary task load and driver age. These relationships appear to complement the patterns observed using the global measure of visual complexity. However, there are subtle differences between these metrics. At first glance, the relationship between mean saccadic amplitude and the major experimental factors appeared to directly mirror the overall pattern of results observed for the visual entropy metric (compare Figure 36 to Figure 22). However, this qualitative similarity did not stand up to quantitative verification. Young drivers demonstrated only a small reduction in mean amplitude for the verbal (0.12 deg) and visual-spatial (0.3 deg) subsidiary tasks, however neither of these changes were statistically reliable. Older drivers on the other hand demonstrated statistically reliable declines of 0.56 and 1.9 deg for the verbal and visual-spatial conditions, respectively. Both of these values were significantly diminished relative to the mean saccadic amplitude observed under baseline conditions. No between age group differences were observed for any of the mean amplitude comparisons that were analyzed. Thus, compared to the visual entropy metric, mean saccadic amplitude was: (a) less sensitive at discriminating between subsidiary load tasks among young drivers, (b) more sensitive at discriminating between subsidiary tasks in older drivers, and (c) less sensitive at distinguishing age differences.

The only systematic effect revealed by the analysis of the mean fixation dwell time data involved a small but statistically significant increase in dwell time from 351 milliseconds in the baseline condition to 406 milliseconds during execution of the visual-spatial subsidiary task. However, this trend was significant only among the older drivers. *Post hoc* exploratory analyses revealed that this slowing was associated with a general reduction in the overall number of eye movements exhibited by the older drivers when required to simultaneously perform the clock task. Hence, this slight "slowing" in the fixation update rate complements the "tunnel vision" effects revealed by the analysis of the state transition matrices previously discussed.

Overall, the gaze strategy of the young drivers seemed relatively unaffected by the type of task. They sampled the roadway environment with similar mean saccade amplitudes and dwell times across all conditions. Even if trends appeared present when graphed, they were not statistically significant in either of the traditional ocular metrics (dwell time and saccadic size). The entropy metric on the other hand, detected slightly more difference within the young groups which allowed the observation of statistical reliability ($p < .05$) between the verbal-memory task and baseline driving conditions.

The older drivers followed the same tendencies as the young, only each "level" was more clearly defined. They made fewer fixations, have longer dwell times and position sequential fixations closer together during a task load. However, the largest differences between groups occurred when the old participants were engaged in the visual-spatial loading condition. The metric of dwell time was able to reliably differentiate the visual-spatial task from no task in the old group. The saccadic amplitude metric was slightly better in that it was able to discriminate *both* the visual-spatial and verbal task from the no task condition. However, neither metric (dwell time or saccadic size) could reliably differentiate *between* the loading tasks. The 2nd order entropy metric on the other hand was able to distinguish both between task loads and also between the visual-spatial and no-task condition; however, the entropy metric was not able to detect a consistent difference between the verbal and no task condition.

The results of the traditional eye measurement metrics (saccadic size and dwell times) do not conflict with the results from the entropy metric. In actuality, they form a complement since each confirms the same underlying structure of behavior, only from different perspectives. As dwell durations increase, one uses more time to scan less of the environment (per unit time). Saccadic amplitude decreases, by definition, mean that successive fixations are closer together (less overall size/area scanned). Part of the behavioral complexity captured by the entropy metric is a composite of both the *time* element captured by dwell time, and the *size* element captured by amplitude. State-space transitions are directly influenced by *both* saccade amplitude and dwell duration since each of these attributes can limit the maximum number of state-spaces that *can be or are* scanned during any unit of time.

Since the entropy metric is a composite metric, it has the *potential* to be more useful in the initial phases of data analysis before one switches to the more narrow focused traditional metrics. Dwell time and saccade size only measure a *single characteristic* which means that if a change does not occur within that dimension, it will not be noticed. In contrast, entropy encompasses both dimensions so one has twice the chance to observe an effect. However, a blessing is also a curse as a composite score also means that some effects might not appear as large as they actually are. Should an effect occur *entirely* in one dimension (i.e. saccadic size *but not* dwell time, or vice versa), any metric related solely to that data type will be more diagnostic than the entropy metric since that data is only *a piece* of the total picture in entropy. This is why the entropy metric is viewed as a complement to existing metrics rather than a replacement. This metric offers great promise for improving *exploratory* strength in future studies simply to show that an effect does in fact exist. Once the effect has been found, one can switch metrics in an attempt to narrow focus.

Pupil size was the last of the ocular metrics explored. On average, older subjects had smaller pupils than the young, but this in itself should not be surprising due to normal age-related reductions in pupil size (i.e., *senile meiosis*). On the other hand, the observation that pupil size was not significant in a tasked condition compared to a no-task condition *was* surprising. This is directly contrary to Recarte & Nunes (2000 & 2003) in which they used pupil size as a confirmation of cognitive loading in tasks and conditions similar to this study. Although they could not statistically differentiate between tasks (due to being "difficulty" matched), they could still reliably distinguish a no-task from a tasked condition using this metric.

The interpretation of this failure to find a significant relationship between pupil size and task load was complicated by several factors:

1. The first was the angular line of sight between the eye tracker's eye camera and the driver's face. This angular offset tended to distort the image of the pupil along both the horizontal and vertical axis. As a result the image of the pupil was an oval rather than a circle. This in itself is only a minor issue except when in tandem with the next factor…
2. The second contributing factor was uncontrolled driver head movements. In a non-dynamic environment, body movements can be easily controlled.

> However in a dynamic driving environment, one cannot anticipate the erratic movements of the operator due to bumps in the road, curves in the road, grade of the road, and/or other natural movement behaviors required to see that change based on the situation. From the eye camera's fixed vantage point, the pupil size appears to fluctuate from small to large based on *distance* from the camera sensor, not due to actual changes in pupil size. These fluctuations occur in differing degrees in both dimensions of measure (vertical and horizontal) depending on head orientation at the time of movement.

These distortions, taken together, introduced significant levels of noise into the on-line measurement of pupil size; thus, making the detection of small reliable changes nearly impossible.

4.3 "Non-Visual" Metrics

Steering Entropy was examined in a post hoc fashion because it is similar in theory to the visual entropy metric. Unfortunately, the results for these analyses bore no resemblance to what was expected based on the Nakayama, et al. (1999) and Boer (2001, 2005) studies. This null finding is believed to be due to the environment in which the tasks were framed. Given that the tasks were performed at high speed on a narrow road with no shoulders, the driver simply didn't have the opportunity to "neglect" lane maintenance. As a result, the drivers appeared to devote more attention to steering maintenance effectively correcting small heading errors as they occurred. Since attentional focus was not allowed to lapse between conditions, the steering entropy evaluation concluded no differences across task load conditions.

A second interpretation for the null steering entropy results is that this may simply be how one actually drives in the "*real" world*. To date, this may be the first attempt to validate the steering entropy paradigm outside of a simulator. It is nearly impossible to associate the same amount of risk perception inherent to the real world in a simulator. Therefore any effect noticed in a simulator (especially one affected by an individual's comfort level with risk) may not be directly applicable to the real world. Nevertheless, it is still worth exploring this concept in future studies due to the ease of data collection. It is hypothesized that by providing less restrictive driving conditions, one will be able to maximize the amount of neglect a subject can "afford" to suffer. As one neglects his or her actions with a greater frequency, steering perturbations should become much more defined which should lead to more distinguished results. As a logical next step, it is suggested that this metric be evaluated on a closed-course test track before application in the real world again. Under this more clearly defined environment; one should be able to more fully explore the actual extent of road condition influence.

Vehicular *yaw rate* was another metric explored since the apparatus was already in the car. This was felt to be a good "direct" measure of vehicle control that could be compared to the steering entropy metric. Surprisingly, this metric proved much more sensitive to variations in subsidiary task load than the steering entropy metric, even though it is highly related to the same variable: steering corrections. Unlike steering entropy, this data set allowed us to distinguish between groups for the No-task condition, and between the verbal/clock task and no-task within the old group. There was very little range variation in the young drivers (.0061 $^{deg/sec/sec}$) but much more (.057 $^{deg/sec/sec}$) in the older group (see figure 48). Another surprising finding was that both groups tended to have less gyro fluctuation when they are engaged in a cognitive task, which was the exact opposite of every single other metric that was assessed (visual and non-visual). The task with the least amount of variation was the clock task and most variation was the no-task condition. Why does the yaw rate show a systematic decrease in deviation for the old subjects depending on task but steering entropy showed that “neglect” was the same between conditions - Shouldn’t steering “neglect” correction and car “wobble” be directly correlated? Why is there a reversal of the graphical data compared to the other metrics? There were two speculations for this result:

1. It is possible that the pattern observed is a function of overall speed. Increased speed should impose a higher g-force load on the entire vehicle when it hits a bump or goes around curve, etc. On the surface, this logic seems sound. As speed goes up, so did the reported amount of gyroscopic wobble. However, the only vectors that can affect the gyro in this car must be lateral in nature as there is no sensor for vertical force. This rules out most situations other than curves of which there were none in this data. The second major issue with overall speed as a contributor is that the wobble phenomenon was highest in the older group whereas the speed was highest in the young group.
2. Another speculative answer is that it isn’t speed per se, but rather an increase in speed updates (i.e. pushing the gas pedal more erratically). This would have the same affect as raw speed, but be unrelated to the speed in general and act as a horizontal vector. In this hypothesis, it is the fluctuations in *velocity change* rather than a simple stable speed constituent. This would also be independent of steering. Subjects were not allowed to use cruise control and

so were forced to monitor speed. They were able to monitor speed more closely under the conditions that interfered the least with current visual load (i.e. baseline, then verbal, then visual spatial conditions). This is supported by the fact that in the baseline condition most participants drove the closest to the speed limit. Since they were instructed not to speed, they may have been making more corrections to keep within that speed range since they had no other tasks to side-track them. While engaged in a task however, most participants drove under the speed limit, presumably in order to compensate for the increased attentional load of a task. This behavioral shift may have impacted the amount of updates devoted to velocity because it was no longer a critical monitoring task in the grand scheme of the experiment. Without a loading task, it was really the only thing they had to truly be aware of independent of the road directly ahead. Unfortunately, this hypothesis also has a weak link. Given that velocity deviations are caused by lack of monitoring, one would expect the standard deviations of speed averages to be lowest in the baseline (since they are constantly being nullified). In reality, the standard deviations virtually mirrored the speed averages which would be opposite of expected given the hypothesis was true.

Unfortunately as of now, these questions and hypotheses (even if ruled out) are to be listed simply as discussion points for future studies since the true answers are still unknown. Additional modeling efforts and empirical testing will be necessary to better understand the mechanism underlying this pattern of curious results.

This leads into the final performance metric to be evaluated: *driving speed.* Surprisingly, this measure proved to be the most robust of any metric evaluated in this study. This metric was able to distinguish significance between tasks, age by task interactions, and even between groups. Again, the young performed at a somewhat elevated level to the older group. Young speed means ranged from 53mph in the no task condition to 52mph in the clock task, whereas the older subjects ranged between 48mph in the no-task condition to 44mph in the clock task condition (figure 50).

Driving speed has been used as an indicator of "effort" in the past. Generally speaking, as effort goes up, speed goes down in the attempt to slow down the total amount of global "information" (verbal, visual, etc) that needs to be processed per unit time. Under this assumption, our speed data supports the notion that the tasks types were of nearly equal "difficulty" because we note almost no variation between the young conditions. Of the two cognitive tasks, we can discern that the speed reductions were the greatest for the tasks most prone to visual interference (i.e. visual-spatial task) which was especially noteworthy in the older group. Again, this should also be expected given our hypotheses about visual-spatial interference and aging is true.

Yet again, we see another metric (driving speed) that restates a story similar to Visual Entropy:

1. Young = relatively unaffected between tasks
2. Old = exaggerated performance decrement during the clock task
3. No task condition = Highest Speed / Highest Entropy
4. Verbal Task = Medium Speed/ Medium Entropy
5. Clock task = Lowest speed, Lowest entropy

4.4 Known Shortcomings of the Study

There were several unforeseen aspects of this study that hindered its overall power and interpretability. The main issue was the eye-tracker's inflexibility both in terms of mounting configuration and software parameters. This proved especially troublesome when testing older participants as a huge drop-out rate was experienced. Only 6 out of the 14 older subjects recruited provided usable eye movement data. Two issues that contributed to this large data drop-out rate were associated with normal aging: namely, (a) the need for bifocal eye glasses and (b) ptosis.

The inability to adjust the angular line-of-sight between the eye tracker's eye camera and the participant's right eye represented a major problem for the study of older drivers. Due to constraints imposed by the configuration of the experimental vehicle's instrument panel, the eye camera was mounted in the center console well below the drivers' eye level. As a result, bifocal lenses (worn by the great majority of the older participants) tended to bisect the image of the pupil. This resulted in a geometric distortion in the shape of the pupil's image which caused in a high rate of data loss for this group. This distortion was instantly recognizable by its characteristic "mushroom" pupil shape of the in the eye-tracker's video screen.

In addition, many of the older participants had "drooping" eye lids (ptosis) that interfered with the ability to track eye position. When the eye lid drooped so much as to occlude the border of the pupil, the image processing algorithms implemented by the eye tracking software could no longer "recognize" the location of the pupil (making eye tracking impossible). In fact, the ASL software often "crashed" under these conditions necessitating a need to manually reboot the control computer at the expense of loosing all data stored in flash memory. Future attempts to use the ASL ETS eye tracking system to study older drivers will benefit from screening procedure that eliminate participants with ptosis and provide participants with custom eyeglasses *without* bifocal additions.

Finally, the last concern encountered in this study was the restrictive nature of roadway geometry. This was originally overlooked because it was the only stretch of road available that met all criteria for the visual entropy portion of the experiment (low traffic, straight, not looking into sun, similar illumination going both directions, unchanging highway speeds for at least 4 miles, etc). While classified as an "issue", it should be noted that it was mainly only an issue for the metrics examined post hoc (e.g., steering entropy, etc). Since the narrow lanes and absence of a shoulder did not allow as much opportunity to wander or let attention lapse (i.e., steering neglect), the data may be more truncated than it would have been if observed on a lower stress road. Future researchers will need to adjust protocols as necessary to maximize the value of the specific metrics they determine have the highest priority in their study. This will undoubtedly come at a cost to the other metrics and so will need to be balanced depending upon objectives.

4.5 Summary

The visual entropy metric has proven the ability to quantify and discriminate between age and task type in a highly dynamic environment such as driving. The validity of this metric was strengthened by the fact that almost all metrics both visual and non-visual seemed to converge with the same global story (see table 8.) The visual entropy metric proved more sensitive to attentional demands than all alternative visual metrics assessed (e.g. pupil size, dwell times, and saccadic amplitudes) and did so in a *real-world* setting.

The validation of the visual entropy metric supports the utility of multiple resource theory (MRT). This metric was able to detect differences between both tasks and age, all while following a predictable path of logic set by the multiple resource theory model. In this study, there was strong evidence that a visual-spatial cognitive task does in fact interfere with a visual motor task more than a verbal oriented task does; supporting the idea that these are functionally distinct pools of attention. In the verbal condition (different attentional pool than vision) very minimal interference with visual scanning was observed in either group. However, in the visual spatial condition (same attentional pool as vision), a very dramatic shift was observed in the older adults. There was also evidence that older adults are more prone to inter-modality interference than their younger counterparts as reflected in nearly all metrics evaluated (ocular and non-ocular). This finding supports the idea that an older adult has an increased disposition for attentional interference in concurrent tasks relying on similar pools of attention. This metric should definitely be considered for future eye-movement quantification studies given a more robust method for collecting data has been established.

The confirmation of this metric has potentially opened the door to many new types of applications, not necessarily exclusive to the driving domain. For instance:

1. In the realm of decision making, one might be able to correlate *actual* difficulty and *perceived* difficulty of a question simply by monitoring the change in scan pattern complexity after a question is asked. This is not entirely beyond the realm of imagination but would require that the entropy

metric be adapted to function with very small data series in a near real time environment.

2. One could also use the visual entropy metric as a quantitative rating metric in exploratory design studies. An example of this would be using the complexity of visual scanning to rate/rank the readability or navigability of text in an advertisement campaign or web-page. Current eye-tracking technology is often used in conditions such as this, but the results are sometimes vague and always subjective. With the entropy metric, an advertisement could be quantitatively rated by the complexity of scan patterns needed to complete some specified goal. An advertisement with a high entropy value would be characterized by poor navigational cues forcing the observer to search/read the entire page in order to arrive at some destination. An advertisement with low entropy would be well defined and require the observer to search very little in order to complete a goal.
3. Visual entropy as a rating metric could also be used to describe pictures and scenes rather than just text and web-page navigation. An example of this would be billboards and other exterior signage. In these situations, the goal would not be navigability or readability of text, but rather how effective a picture is at conveying an idea. Again, the more complex the visual search patterns, the higher the entropy and less effective that sign is for the transmission of a given concept.
4. Another potential application for the visual entropy metric would be to assess the “flash” or “cognitive engagement” of item. An example of this would be

to numerically rate the appeal of a series of products not by subjective assessment, but by the complexity of visual scan patterns observed when looking at them. If a product is designed to be engaging, then having a complex scanning pattern is very preferred. However, a product that produces little scan pattern changes would suggest that the item may be somewhat boring. Having a quantitative number confirm a subjective answer would allow one to have much more faith in a finding. Although much more futuristic of an idea, it is theoretically possible that this metric could also be used on animals as well as humans. By evaluating the scanning complexity of an animal observer, one may be able to rate the "flash" or "engagement" of an item from the perspective of the animal itself. This would be a huge leap for pet toy design. Using this technology, one could design more aggressive fishing lures or pet toys that encourage visual engagement.

5. There are also futuristic applications within the driving domain. It is possible that a real-time version of the visual entropy metric could someday be used in-vehicle. In this situation, the *car* would assess the current cognitive status of a driver based on his or her eye scan patterns in order to mediate data transfer between the car and driver. This would allow a car to positively enhance information transfer to the driver at times when he or she is deemed "cognitively loaded". An example of this might be the switching of a heads-up display road map to auditory directions when the car senses cognitive loaded in the form of eye-movement clustering.

6. Finally, one could potentially use the visual entropy metric as learning tool. Although not currently feasible, if a real-time version of the visual entropy metric was established, warnings could occur in the car and at the exact time the car senses a visual truncation is occurring. The car would warn the driver so that he or she becomes aware of this potentially dangerous behavior as it occurs. Since the context of the environment causing the behavior is known, the driver could theoretically train his or herself to better divide attention specifically in those types of situations. Retention should be much better when trained in this way since it is occurring in the environment where it will needs to be recalled later.

Now that the entropy metric has proven to work on scan patterns, other visual data sources might also be able to be tapped for future studies. Fixation durations, transitions and even saccadic size could be theoretically plugged into the entropy equation. As long as the data set can be defined as a distribution of probabilities, an entropy value can be assigned to describe that data’s complexity. This is the true power of this metric.

All in all, the visual entropy metric has now been validated by this study and the applications for its use are only limited by the imagination. Vision is simply the starting point.

5.0 References

Attneave, F. (1959). *Application of information theory to psychology: A summary of basic concepts, methods and results*. New York: Holt, Rinehart and Winston.

Ball, K., & Owsley, C. (1992). The useful field of view test: a new technique for evaluating age related declines in visual function. *Journal of American Optometry Association, 63,* 71-79.

Bhise, V., & Rockwell, T. (1973). Development of a methodology for evaluating highway signs. Project EES-315B Final Report. Engineering Experiment Station Bulletin 207. Engineering Publications, College of Engineering, Ohio State University, 2070 Neil Avenue, Columbus, OH 43210.

Boer, E. (2000). Behavioral entropy as an index of workload. Proceedings of the IEA 2000 / HFES 2000 congress.

Boer, E., Rakauskas, M., Ward, N., & Goodrich, M. (2005), Steering entropy revisited. In proceedings of the 3rd annual Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. June 27-30.

Braver, T., Barch, D., Kelley, W., Buckner, R., Cohen, N., Miezin, F., Snyder, A., Ollinger, J., Akbudak, E., Conturo, T., & Peterson, S. (2001). Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. *Neuroimage*, *14*, 48-59.

Brockmole, J., & Henderson, J. (2004). Prioritization of new objects in real world scenes: evidence from eye movements. *Michigan State University Eye Movement Laboratory Technical Report*, *1*, 1-13.

Buswell, G.T. (1935). *How people look at pictures*. Chicago, University of Chicago Press.

Callicott, J., Mattay, V., Bertolino, A., Finn, K., Coppola, R., Frank, J., Goldberg, T., & Weinberger, D. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex, 9*(1), 20-26.

Carlson, S., Martinkauppi, S., Rama, P., Salli, E., Korvenoja, A., & Aronen, H.(1998). Distribution of cortical activation during visuospatial n-back tasks as revealed by functional magnetic resonance imaging. *Cerebral Cortex*, *8*(8), 743-752.

Carpenter, R. (1988). *Movements of the Eyes*. London: Pion

Chapman, P., & Underwood, G. (1998) Visual Search of driving situations: Danger and experience. *Perception, 27,* 951-964.

Cohen, A., & Studach, H. (1977). Movements and driving cars around curves. *Perceptual Motor Skills. 44*(3), 683-689.

Crundall, D., Underwood, G., & Chapman, P. (1998) How much do novice drivers see? the effects of demand on visual search strategies in novice and experienced drivers. In G. Underwood (eds), *Eye Guidance in Reading and Scene Perception*, Elsevier Science Publishers, 395-417.

Crundall, D., Underwood, G., & Chapman, P. (2002) Attending to the peripheral world while driving. *Applied Cognitive Psychology, 16*, 459-475.

Cutting, J., Alliprandini, P., & Wang, R. (2000). Seeking one's heading through eye movements. *Psychonomic Bulletin & Review, 7*(3), 490-489.

Dishart, D., & Land, L. (1998). The development of eye movement strategies of learner drivers. In G. Underwood (eds.), Eye Guidance in Reading and Scene Perception, Elsevier Science Publishers, 419-429.

Donges E., (1978). A two level model of driver steering behavior. *Human Factors, 20*, 691-707.

Duren, L. (1993). Central Stimulus Processing during saccadic eye movements. In d'Ydewalle et al (Eds.) *Perception and Cognition: Advances in Eye Movement Research*. North Holland: Elsevier Science Publishers, 23-35.

Edwards, E. (1969) *Information Transmission: an introductory guide to the application of the theory of information to the human sciences.* London: Chapman and Hall Publishers.

European Communities. (1999). *COST 331: Requirements for horizontal road marking*. Luxembourg: Office for official publications of the European Communities.

Falkmer, T., & Gregersen, N. (2005) A comparison of eye movement behavior of inexperienced and experienced drivers in real traffic environments. *Optometry and Vision Science, 82*(8), 732-739.

Fildes, B., Charlton, J.L., Koppel, S., Andrea, D., & Hammer, M. (Winter 2005). Visual search strategies among older and younger drivers. *Aging Technical Group - Newsletter of the Human Factors and Ergonomics Society*, pg 6.

Galley, N., & Andres, G (1996). Saccadic eye movements and blinks during long-term driving on the autobahn with minimal alcohol ingestion. In A. G. Gale et al. (eds.), Eye Guidance in Reading and Scene Perception, Elsevier Science Publishers, 381-387.

Gilland, J. (2004) Age differences in the useful field of view during real world driving. Unpublished master's thesis. University of South Dakota

Green, P. (2002). Where do drivers look while driving (and for how long). *Human Factors in Traffic Safety*. 77-110.

Harbluk, J., & Noy, Y. (2002). The impact of cognitive distraction on driver visual behavior and vehicle control. TP# 13889 E Aussi disponible en francais.

Harms, L. (1986). Drivers' attentional responses to environmental variations: A dual-task real traffic study. In A.G. Gale, et al. (Eds.) *Vision in Vehicles*. Amsterdam: North-Holland. 131-138.

Harms, L. (1991). Variations in drivers' cognitive load: Effects of Driving through village and rural junctions. *Ergonomics, 34*, 151-160.

Harris, C.M., Hainline, L., Abramov, I., Lemerise, E., & Camenzuli, C. (1988). The distribution of fixation durations in infants and naïve adults. *Vision Research, 28*, 419-432.

Henderson, J. (1993). Visual attention and saccadic eye movments. In d'Ydewalle et al (Eds.) *Perception and Cognition: Advances in Eye Movement Research*. North Holland: Elsevier Science Publishers, 37-50.

Henderson, J., Williams, C.C., Castelhano, M., & Falk, R. (2003) Eye movmements and picture processing during recognition. *Perception & Psychophysics, 65*, 725-734.

Henderson, J., & Hollingworth., A. (1999). High-level scene perception. *Annual Review of Psychology*, 50, 243-71.

Henderson, J., & Hollingworth, A. (2003). Eye movements and visual memory: detecting changes to saccade targets in scenes. *Perception & Psychophysics, 65*, 58-71.

Henderson, J., & Hollingworth, A. (2003b). Global transsaccadic change blindness during scene perception. *Psychological Science*, *14*, 493-497.

Hollingworth, A., & Henderson, J.M. (2003) Testing a conceptual locus for the inconsistent object detection advantage in real-world scenes. *Memory and Cognition*, *31*, 930-940.

Horberry, T., Anderson, J., Regan, A., Triggs, T., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention*, *38(1),* 185 – 191.

Irwin, David, E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science, 5*(3), 94 - 101.

Karlin, S., & Taylor, H. (1975). *A first course in stochastic processes* (second edition). Academic Press. San Diego, California.

Keppel, G. (1973). *Design and analysis: a researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.

Korteling, J.E. (1990). Perception-response speed and driving capabilities of brain-damaged and older drivers. *Human Factors, 32*, 25-31.

Korteling, J.E. (1991). Effects of skill integration and perceptual competition on age related differences in dual-task performance. *Human Factors, 33*, 35-44.

Land, M. (1992). Predictable eye-head coordination during driving. *Nature, 359*, 318-320.

Land, M., & Horwood, J. (1996). The relationship between head and eye movements during driving. In Gale, A.G. et al. (eds.), Vision in Vehicles V. Elsevier Science Publishers, pg 153-160.

Land, M., & Lee. (1994). Where we look when we steer. *Nature, 369*, 742-744.

Land, M.F., & Tatler, B.W. (2001). Steering with the head: the visual strategy of a racing driver. *Current Biology, 11*, 1215-1220.

Louma, J. (1988). Drivers eye fixations and perceptions. In Gale, A.G. et al. (eds.), *Vision in Vehicles II*. Amsterdam: Elsevier Science Publishers, 231-237.

Mackworth, N. H, & Morandi, A.J. (1967). The gaze selects informative details within pictures. *Quarterly Journal of Experimental Psychology*. 35A, 187-198.

Maltz, M., Shinar, D., & Gurion, B. (1999). Eye Movements of younger and older drivers. *Human Factors*, *41*(1), 15-25.

McCarley, J., Vais, M., Pringle, H., Kramer, A., Irwin, D., Strayer, D. (2001). Conversation disrupts visual scanning of Traffic scenes. *Proceedings of the 9th International Conference on Vision in Vehicle*, Brisbane Australia.

Miura, T. (1990). Active function of eye movement and useful field of view in a realistic setting. In R. Groner, D. d'Ydewalle, & R. Parham (Eds.), *From eye to mind: information acquisition in perception, search, and reading*, 119-127.

Miura, T. (1992). Visual Search in Intersections: an underlying mechanism. *IATSS Research*, *16*(1).

Mourant, R., & Rockwell, T. (1970). Mapping eye movement patterns to the visual scene in driving: an exploratory study. *Human Factors, 12*(1), *81-87.*

Nakayama, O., Futami, T., Nakamura, T, & Boer, E. (1999). Development of a steering entropy method for evaluating driver workload. SAE Technical Paper reprinted from *Human Factors in Audio Interior Systems, Driving, and Vehicle Seating*, (SP-1426)

Olson, P., Battle, D., & Aoki, T. (1989). Driver eye fixations under different operating conditions. UMTRI-89-03.

Perlstein, W., Dixit, N., Cameron, C, Noll, D., & Cohen, J. (2003). Prefrontal cortex dysfunction mediates deficits in working memory and prepotent responding in schizophrenia. *Biological Psychiatry, 53*, 25-38.

Perryman, K., & Fitten, L. (1996). Effects of normal aging on performance of motor-vehicle operational skills. *Journal of Geriatric Psychiatry Neurology. 9*(3), 136-141.

Recarte, M., & Nunes, L. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology. 6*(1), 31-43.

Recarte, M., & Nunes, L. (2003). Mental Workload while driving: effects on visual search , discrimination, and decision making. *Journal of Experimental Psychology: Applied, 9*(2), 119-137.

Rockwell, T. (1988). Spare visual capacity in driving-revisited. In Gale, A.G. et al. (eds), *Vision in Vehicles II*. Amsterdam: Elsevier Science Publishers.

Schieber, F. (2005) Effects of driving speed and sign reflectance on nighttime highway sign Legibility Distance and Reading Time Requirements. HFES presentation.

Schieber, F., Burns, D., Myers, J., Willan, N., Gilland, J. (2004) Driver Eye fixation and reading patterns while using highway signs under dynamic nighttime driving conditions: effects of age, sign luminance, and environmental demand.

Schieber, F., & Gilland, J. (2005) Age differences in the useful field of view during real world driving. *Presented at the 2005 HFES conference in Orlando Florida.*

Serafin, C. (1994). Driver Eye Fixations on Rural Roads: Insight into Safe Driving Behavior. *UMTRI-94-21 technical report*.

Schnell, T., & Zwahlen, (1999). Driver preview distance at night based on driver eye scanning recordings as a function of pavement marking retroreflectivities. Transportation Research Record 1692, paper 99-1137, pg129-141.

Shannon, C. (1948). *The Bell system Technical Journal*. 27(July & October), 379-423 & 623-656.

Shinar, D., McDowell, E., & Rockwell, T. (1977) Eye movements in curve negotiation. *Human Factors, 19*, 63-71.

Shinar, D. (1978). Field dependence and driver visual search behavior. *Human Factors, 20*(5), 553-559.

Smith, E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*, 5-42.

Spijkers, W. (1992). Distribution of eye-fixations during driving. *IATSS Research 16*(1), 27-34.

Taoka, G. (1991) Distribution of driver spare glance durations. *Transportation Research Record #1318*, 142-146.

Tole, J., Stephens, A., Vivaudou, M., Ephrath, A., & Young, L. (1983) Visual scanning behaviour and pilot workload. *NASA CR-3717 technical report*. MIT: Division of Health Sciences and Technology.

Tuckwell, H. (1988). *Elementary Applications of Probability Theory*. Pub. Chapman and Hall Ltd. New York, New York. Reprint of original from 1943.

Tsimhoni, O., & Green, P. (1999). Visual demand of driving curves determined by visual occlusion, paper presented at the Vision in Vehicles 8 Conference, Boston, MA, August 22-25.

Underwood, G., Phelps, N., Wright, C., Van Loon, E., & Gaplin, A. (2005). Eye fixation scanpaths of younger and older drivers in a hazard perception task. *Opthalmic & Physiological Optics. 25*(4), 346-356.

Van der Lubbe, J.C.A. (1997). *Information theory*. New York: Cambridge University Press.

Verhaeghen, P., & Chandramallika, B. (2005). Aging and switching of the focus of attention in working memory: results from a modified N-Back task. *The Quarterly Journal of Experimental Psychology, 58A* (1), 134-154.

Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & R.Davies (Eds.), *Varieties of Attention.* New York: Academic.

Yarbus, A.L. (1967). *Eye Movements and Vision*. New York: Plenum.

Zwahlen, H (1980). Driver eye scanning behavior in rain and during an unexpected windshield wiper failure. *Zeitschrift fur Verkehrssicherheit*, 26(4), 148-155.

Zwahlen, H., & Schnell, T. (1998). Driver eye scanning behavior when reading symbolic warning signs. *Vision in Vehicles – VI,* Elsevier Science Ltd, pp. 3-11.

Zwahlen, H. & Schnell, T., (1999). Visual Target Detection Models for Civil Twilight and Night Driving Conditions, Paper No. 991545 presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC, January 10-14, 1999. Published in Transportation Research Record 1692, pp. 49-65.

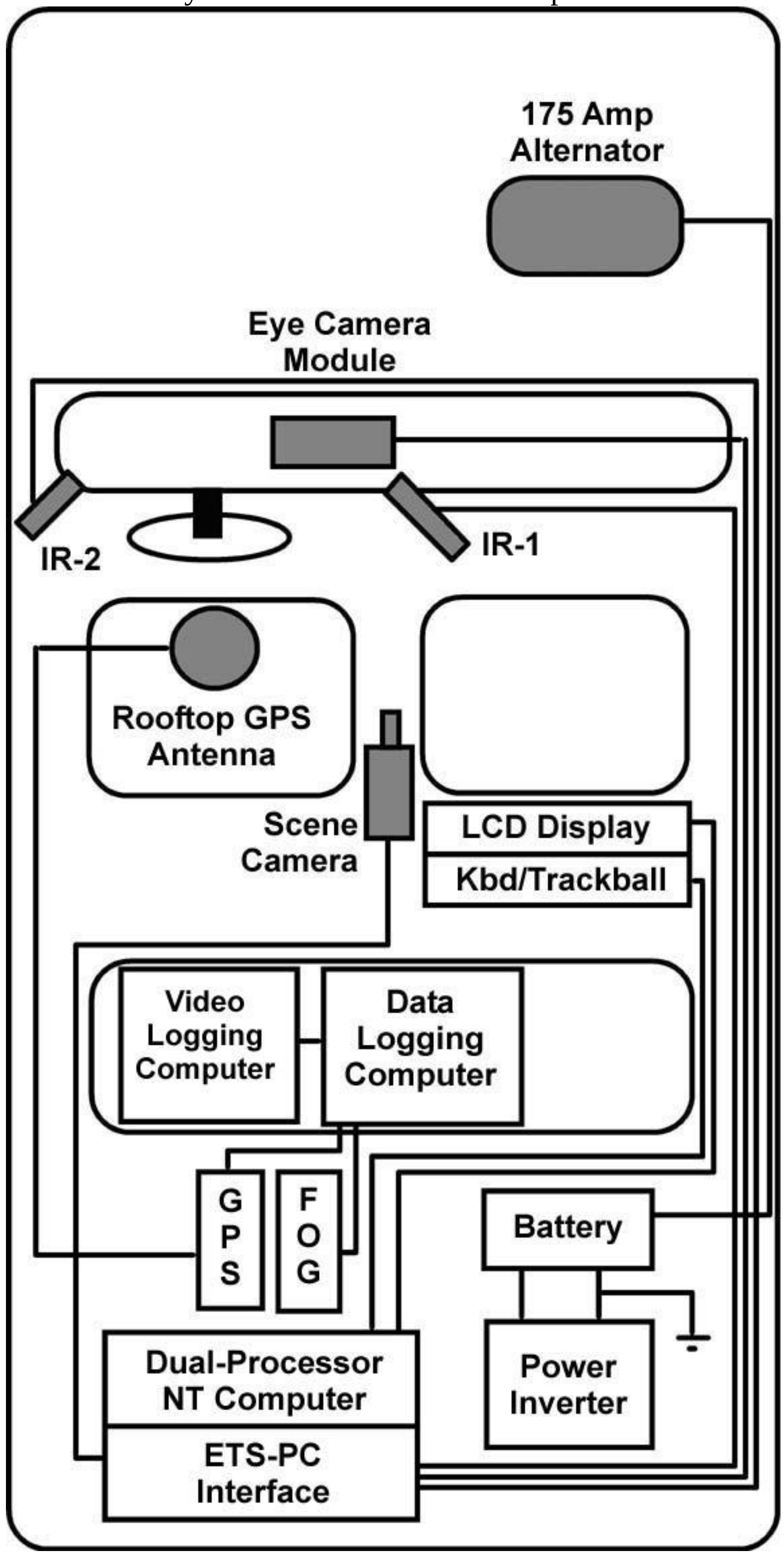
6.0 Appendix

6.0.1 Appendix A – Test Vehicle Pictures

Toyota Avalon (Subject driven test car)

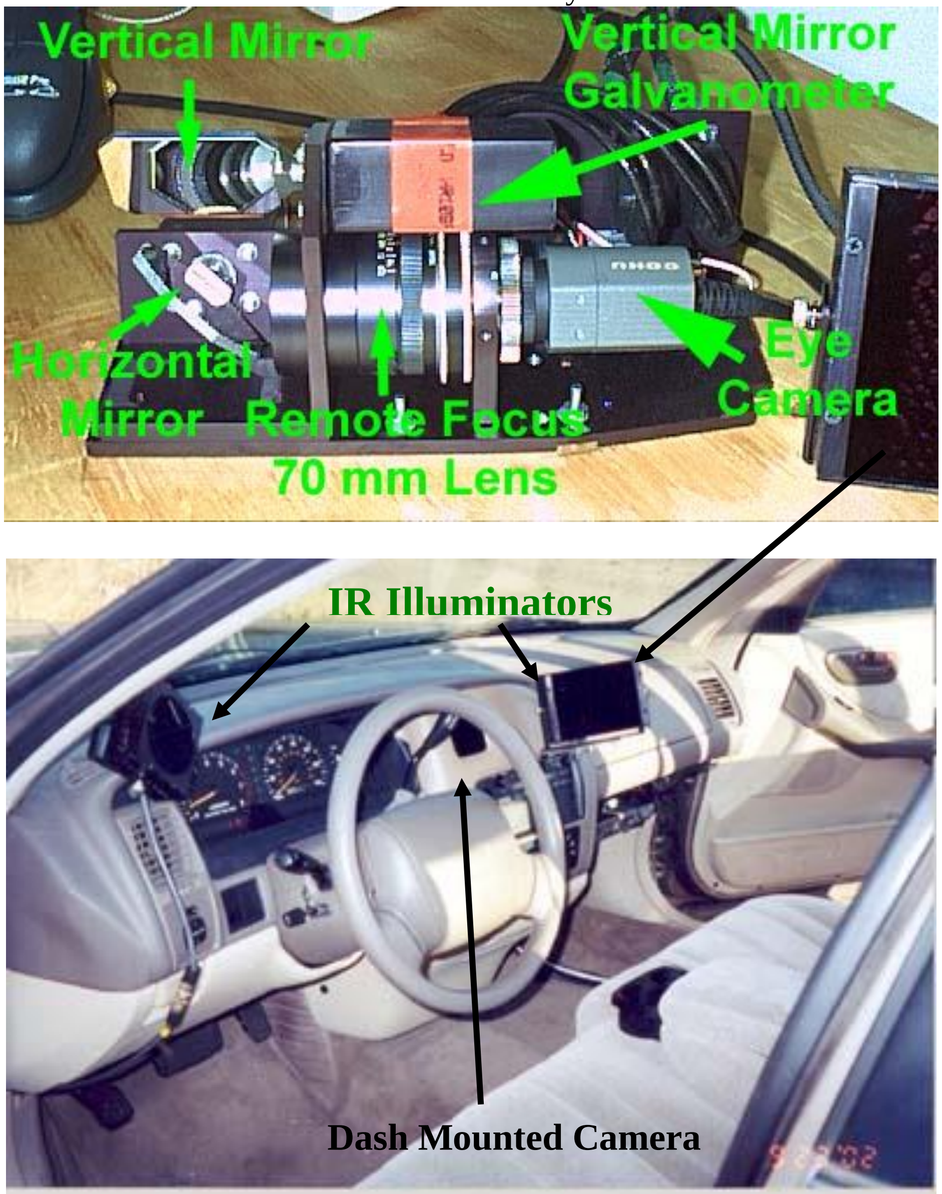
6.0.2 Appendix B – Diagram of Test Car Equipment

USD Toyota Avalon test vehicle setup schematic.



6.0.3 Appendix C – Eye-Tracking Camera and Location

ETS-PC dash mounted eye tracker



6.0.4 Appendix D – Informed Consent Document

Informed Consent
The University of South Dakota
Vermillion, SD 57069

TITLE: **Driver Eye Scan Patterns by Age and Workload.**
PROJECT DIRECTOR: **Jess Gilland**
Contact Info: **605-XXX-XXXX or jess.gilland@usd.edu**
Department: **Department of Psychology, Human Factors**

STATEMENT OF RESEARCH

It is a basic ethical principle that a person who is to participate in research must give his or her informed consent to such participation. This consent must be based on the understanding of the nature and risks of the research. This document provides information important for this understanding. Research projects include only participants who choose to take part. Please take your time to make this decision. If at any time you have questions please ask.

WHAT IS THE PURPOSE OF THIS STUDY?

You are invited to be in a research study whose purpose is to further understand how visual scanning patterns change under different driving conditions. The main focus will be the comparison of eye scanning behavior produced under three different workload conditions. The resulting data will also be compared by age and gender. You were selected as a possible participant because you satisfy the following requirements: you (1) currently hold a valid driver's license, (2) meet or exceed a minimum visual acuity of 20/40, (3) have not been involved in more than 1 vehicular accident in the past two years, (4) are covered by personal medical insurance or Medicare (*proof of insurance is required to participate*), and (5) have met certain age-related eligibility criteria.

HOW MANY PEOPLE WILL PARTICIPATE?

Approximately thirty-three (33) people will take part in this study. The ratio will be 22 young and 11 older participants.

HOW LONG WILL I BE IN THIS STUDY?

Your participation in the study will last approximately 2 hours. The experiment will begin within the Heimstra Human Factors Laboratory at the University of South Dakota, and then follow-up with driving of the University's research vehicle on the highway between Vermillion and I-29.

WHAT WILL HAPPEN DURING THIS STUDY?

You will first be given a brief description of the project and then be asked to complete the informed consent procedure if you agree to participate. Visual and auditory health will then be checked using standard visual acuity and auditory test procedures. Next, you will be taken to the parking lot and seated in the USD instrumented research vehicle (a 1998 Toyota Avalon). Here you will be allowed to adjust the seat and mirrors to ensure comfortable operation. Next we will align the computer's eye-tracking sensors to recognize your eye. After recognition is achieved, a brief calibration to a set of gaze points will follow. Upon completion of the calibration, you will be introduced to two different subsidiary tasks (visual spatial manipulation and verbal discrimination) that you will be asked to perform while driving later on. The visual spatial task will require a mental manipulation of auditory input to determine what the physically equivalent "answer" would be. For example, you will be asked to determine whether the angle between the hands of a standard clock face are more than or less than 90° when given times like 12:30, 1:45, 3:22, etc. The verbal discrimination task will require you to assess semantic, grammatical, or structural properties of auditory presented words. For example, in this task, you may be given an auditory presentation of words like "tree" and asked to make distinctions such as "does it start with the letter t?".

Once comfortable with these subsidiary tasks, you will begin driving to the outside of Vermillion via directions provided by the experimenter in the back seat. Once outside of town you will be instructed to drive on the highway towards I-29, turn around at the Coffee Cup gas station, and drive back to Vermillion. This will continue for three laps: one lap while performing a visual spatial secondary task, one lap while performing a verbal discrimination task, and one lap with no secondary task used. This will continue for three trips to the interstate and back. After successful completion of all trials you will return to the original parking lot and be debriefed regarding the experimental hypotheses of the study.

If at any time you wish to quit the study *for any reason* whatsoever, the experiment will be stopped and you will be allowed to leave. Please be aware that the experimental session will be recorded digitally to a computer file. These recordings will contain possibly identifiable information in the form of the vocal responses made for each task and/or video of the experiment as recorded from a scene camera in the car. It should be noted that this information will only be accessed by the Primary Investigator and his Academic Advisor. Once the necessary information is extracted from these recordings, their audio and video data (both hard data copies and digitized saved recordings) will be placed in a locked file cabinet so that participant confidentiality remains upheld.

WHAT ARE THE RISKS OF THIS STUDY?

As is the case whenever one drives an automobile in traffic, there is a risk that personal injury or death could result from an automobile crash while participating in this study. However, previous research and the past experience of the experimenters indicate that the chance you will be involved in a crash is no greater than that you would encounter during normal, everyday driving. In the event of an accident, the audio/video recording from your participation may be subpoenaed for investigation.

The University of South Dakota maintains general liability and automobile coverage, in the maximum amount of $1,000,000 per occurrence as currently provided by the Public Entity Pool for Liability (PEPL Fund) a self insurance liability pool established under SDCL ch. 3-22. However, this liability coverage should not be considered as a replacement for personal medical insurance to cover potential injuries to oneself. Your personal medical insurance should be considered as the primary source to all other available insurances. If you are not covered by personal medical insurance and/or Medicare you should not participate in this study.

If you are injured or become ill from taking part in this study, 24-hour emergency medical treatment is available at Sioux Valley Vermillion Medical Center. The University of South Dakota will provide compensation for research related injury if it is proven to be the direct result of negligence by a University of South Dakota employee. No other funds have been set aside to compensate you in the event of an injury.

WHAT ARE THE BENEFITS OF THIS STUDY?

You will not benefit personally from being in this study. However, we hope that in the future, other people might benefit from this study because it will increase our understanding regarding the limits of how visual attention is processed and distributed while driving. Such information may be useful for improving the design of highways, vehicles, driver screening tests in addition to also providing a validation to a new method for studying eye-tracking data.

ALTERNATIVES TO PARTICIPATING IN THIS STUDY

The participant's alternative is simply to not participate.

WILL IT COST ME ANYTHING TO BE IN THIS STUDY?

You will not incur any costs for being in this research study.

WILL I BE PAID FOR PARTICIPATING?

You will not be paid for being in this research study. There will however be class credit points awarded for those recruited from psychology classes at USD. Students expecting class credit in exchange for their participation should be sure to identify the specific class to which they want the credit to apply using Experimetrix Online.

WHO IS FUNDING THE STUDY?

The University of South Dakota and the research team are receiving no payments from other agencies, organizations, or companies to conduct this research study.

CONFIDENTIALITY

The records of this study will be kept private to the extent permitted by law. In any report about this study that might be published, you will not be identified. Your study record may be reviewed by the USD Compliance Office and the University of South Dakota Human Subjects Committee.

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. Confidentiality will be maintained by storing all results of visual acuity screening and all performance data electronically. This data will be coded by anonymous sequence number and gender only. Informed consent documents will be the only project documents containing personally identifying information and will be stored in a locked file cabinet in the office of the principle investigator.

If we write a report or article about this study, we will describe the study results in a summarized manner so that you cannot be identified. Any presentation or publication data will be in the form of anonymous group statistics.

IS THIS STUDY VOLUNTARY?

Your participation is voluntary. You may choose not to participate or you may discontinue your participation at any time without penalty or loss of benefits to which you are otherwise entitled. Your decision whether or not to participate will not affect your current or future relations with the University of South Dakota.

CONTACTS AND QUESTIONS?

The researchers conducting this study are Project Director: **Jess Gilland** and his Advisor: **Professor Frank Schieber.** You may ask any questions you have now. If you have questions later, you may contact **Jess Gilland** at **(605) XXX-XXXX** or **jess.gilland@usd.edu** during the day. You may also contact **Professor Frank Schieber** at **(605) XXX-XXXX** or **frank.schieber@usd.edu** during regular business hours.

If you have questions regarding your rights as a research participant, or research related injury you may contact the **University of South Dakota Institutional Review Board** at **(605) - 677-6184.**

General information about being a research participant can be found by clicking "Information for Research Participants" on the Research Compliance web site http://www.usd.edu/oorsch/compliance/.

Your signature indicates that this research study has been explained to you, that your questions have been answered, and that you agree to take part in this study. You will receive a copy of this form.

Participant's Name: ____

______________________________________ ______________________

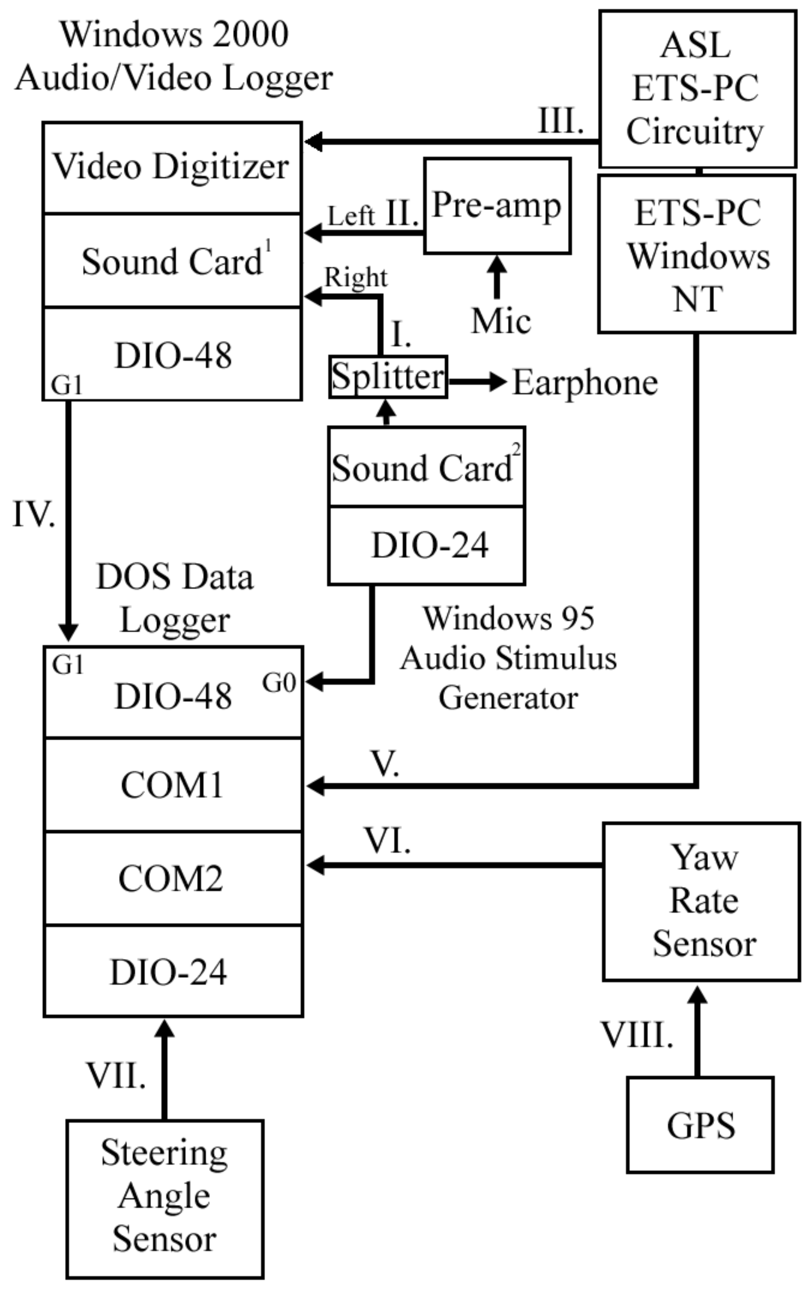
Signature of Participant Date

I have discussed the above points with the participant or, where appropriate, with the participant's legally authorized representative. It is my opinion that the participant adequately understands the risks, benefits, and procedures involved with participation in this study.

______________________________________ ______________________

Signature of Person Who Obtained Consent Date

6.0.5 Appendix E – Data Collection Pathways



USD Research Vehicle Instrumentation Diagram
Driver Visual Entropy Project (Summer 2006)

Functional Module	Part/Model Number
Signal Splitter	Radio Shack Baseband 1X3 amplifier (15-1103)
Pre-amp	Rolls Mini-Mic Preamp MP13[a]
Earphone/Microphone	Labtec AXIS-501 headset
Stereo Sound Card[1]	SoundBlaster Pro 5.1 (Model SB0220)
Stereo Sound Card[2]	Sound Blaster 16 (Model CT4170)
Video Digitizer	Hauppauge WinTV (PCI) Model 38101 Rev. B410
DIO-48 (2)	Acess PIO-DIO-48 48-bit digital I/O interface
COM1	Built-in RS232 Serial Port (IRQ-4)
COM2	Built-in RS232 Serial Port (IRQ-3)
DIO-24	Metrabyte DIO24 24-bit digital I/O interface
Steering Angle Sensor	OMRON Rotary Encoder E6C2-AG5C (9.5 bits)
Yaw Rate Sensor	KVH AutoGyro Fiber Optic Gyroscope
GPS	Starlink DNAV-212 D-GPS receiver

[a] Modified to supply power to Labtec AXIS-501 dynamic noise-canceling microphone.

Data Streams

I. Audio stimulus (secondary loading tasks)
II. Audio response channel (from headset microphone)
III. ETS-PC scene camera with gaze position overlay
IV. Digitized video unique frame number (20-bit) (30 Hz)
V. ETS-PC gaze position/pupil size data (60 Hz)
VI. Fiber-optic laser gyroscope data (with interleaved GPS) (10 Hz)
VII. Steering wheel angular position (9.5 bit graycode)
VIII. GPS – latitude, longitude, speed (1 Hz)

6.0.6 Appendix F – CITI Certification

COURSE COMPLETION CERTIFICATE

CITI PRESENTS THIS CERTIFICATE TO

Jess Gilland

jgilland@usd.edu

FOR THE SUCCESSFULL COMPLETION OF ALL
REQUIRED SOCIAL/BEHAVIORAL MODULES OF
THE CITI
WEB BASED COURSE ON THE PROTECTION OF
HUMAN
RESEARCH SUBJECTS ON 2/17/04

THIS DOCUMENT WAS PREPARED BY

Lisa Korcuska, BA, CCRP, CCRC
Research Compliance
Associate Compliance Officer

CONFIRMATION MAY BE OBTAINED AT
Research Compliance
Pardee Hall 301

2/25/04

Certification valid for 3 years from date of completion

6.0.7 Appendix G – Field Testing Route

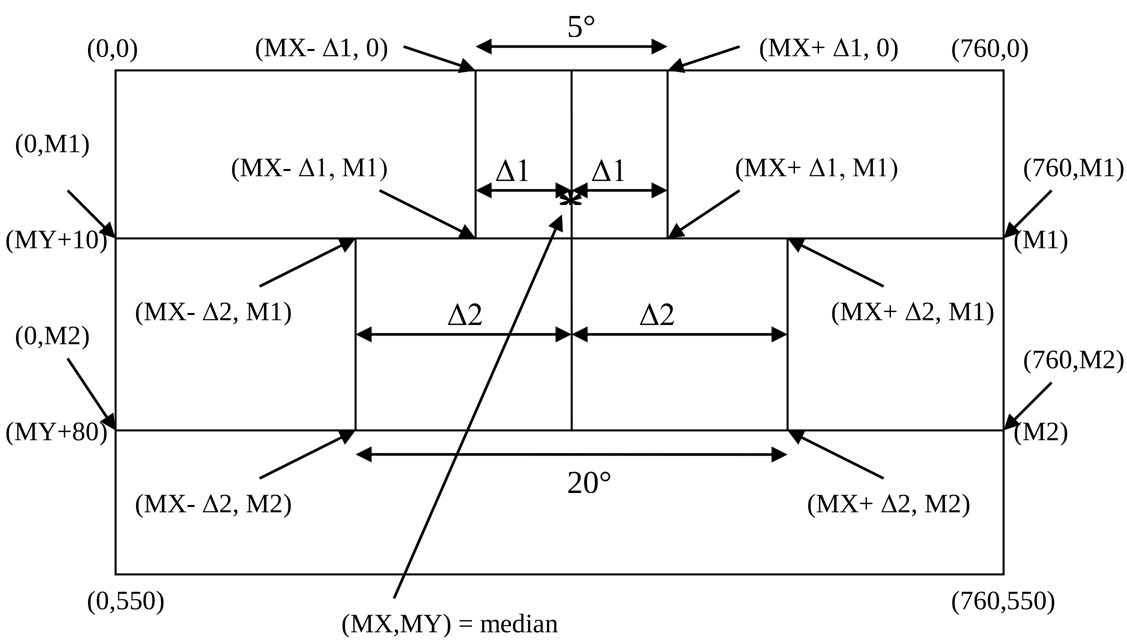
Field Testing Route

6.0.8 Appendix H – Matlab Fixation Calculation Code

Matlab Script Used to provide descriptives on secondary measures while also defining the Areas of Interest specific to each Subject. AOI's are then used for Entropy sequence calculations.
(Original Code written by Frank Schieber PH.D.)

```
function [edges]=modelETS(filename_prefix)
%
%read specified Gilland dissertation data file
%strip-out zeros and display median(x,y) values

%read contents of file
filename=[filename_prefix,'.dat'];
[frame,sample,x,y,hpupil,vpupil,gyro,speed,wheel,lat,long,snd,etime]=te
xtread(filename,...
    '%d,%d,%d,%d,%d,%d,%d,%f,%d,%f,%f,%d,%d');

%strip away zero values
xn=x(x>0);
yn=y(y>0);

%calculate percentage of zero data
percentx = ( length(xn)/length(x) ) * 100.0;
percenty = ( length(yn)/length(y) ) * 100.0;

%calculate eye tracker coordinate medians
MX=median(xn);
MY=median(yn);

%report results
display(['X-median = ',num2str(mx),'  (%zero data loss = ',num2str(100-
percentx),')'])
display(['Y-median = ',num2str(my),'  (%zero data loss = ',num2str(100-
percenty),')'])

%calculate descriptives of speed, gyro, and steering sensor

%mean/sd/min/max speed
speed=speed(speed>0); %filter-out gps drop-outs (if any)
smean=mean(speed);
ssd=std(speed);
smin=min(speed);
smax=max(speed);
disp(' ')
disp(['speed mean = ',num2str(smean)])
disp(['speed std  = ',num2str(ssd)])
disp(['speed min  = ',num2str(smin)])
disp(['speed max  = ',num2str(smax)])

%steering wheel
wmean=mean(wheel);
```

```
wmd=median(wheel);
wsd=std(wheel);
wmin=min(wheel);
wmax=max(wheel);
disp(' ')
disp(['wheel mean   = ',num2str(wmean)])
disp(['wheel median = ',num2str(wmd)])
disp(['wheel std    = ',num2str(wsd)])
disp(['wheel min    = ',num2str(wmin)])
disp(['wheel max    = ',num2str(wmax)])

%gyro
gmean=mean(gyro);
gmd=median(gyro);
gsd=std(gyro);
gmin=min(gyro);
gmax=max(gyro);
disp(' ')
disp(['gyro mean   = ',num2str(gmean)])
disp(['gyro median = ',num2str(gmd)])
disp(['gyro std    = ',num2str(gsd)])
disp(['gyro min    = ',num2str(gmin)])
disp(['gyro max    = ',num2str(gmax)])

%use previously calculated median(MX,MY) eye gaze position to
%construct AOI model of the driver's visual field
%
%                    AOI Calculation Summary Pictorial
%                        (Adapted from Figure 23)
%
```




```
%
%Coordinates of plane that are known: (0,0) top left to (760,550)
%bottom right.
%M1 = dividing line between near and far
```

```
%place it 1-deg x 10 pixels/deg below median-y
M1=my+10;
%M2 = dividing line between near and instrument-panel
%can't use fixed value of Y=386 because of DC offset error unique to
each individual
%define M2 as a specified angular distance below MY
%place it 8-deg x 10 pixels/deg below median-y
M2=my+80;
%delta1 = half-width of far/central zone
%plus/minus 2.5-deg x 10 pixels/deg
delta1=25;
%delta2 = half-width of near/central zone
%plus/minus 10-deg x 10 pixels/deg
delta2=100;
%
%
%far/left zone
fltop=0;
flbottom=M1;
flleft=0;
flright=mx-delta1;
%
%far/center zone
fctop=0;
fcbottom=M1;
fcleft=mx-delta1;
fcright=mx+delta1;
%
%far/right zone
frtop=0;
frbottom=M1;
frleft=mx+delta1;
frright=760;
%
%near/left zone
nltop=M1;
nlbottom=M2;
nlleft=0;
nlright=mx-delta2;
%
%near/center zone
nctop=M1;
ncbottom=M2;
ncleft=mx-delta2;
ncright=mx+delta2;
%
%near/right zone
nrtop=M1;
nrbottom=M2;
nrleft=mx+delta2;
nrright=760;
%
%instrument panel
iptop=M2;
ipbottom=550;
ipleft=0;
ipright=760;
```

```
model=[1,fltop,flbottom,flleft,flright; ...
       2,fctop,fcbottom,fcleft,fcright;
       3,frtop,frbottom,frleft,frright;
       4,nltop,nlbottom,nlleft,nlright;
       5,nctop,ncbottom,ncleft,ncright;
       6,nrtop,nrbottom,nrleft,nrright;
       7,iptop,ipbottom,ipleft,ipright];
model

%open new AOI file
aoi=[filename_prefix,'.aoi'];
disp(['Output filename: ',aoi]);
[fid,message]=fopen(aoi,'wt');
fprintf(fid,'\n\n[Area of Interest Header]\nfile_id=  V1.0\n');
fprintf(fid,'date_created= %s\n', date);
fprintf(fid,'title= %s\n', aoi);
fprintf(fid,'number_of_aois= 7\n');
fprintf(fid,'number_of_scene_planes= 1\n');
fprintf(fid,'eyehead_type_file= NO\n');
fprintf(fid,'aoi_file_scaled= NO\n\n');
fprintf(fid,'1=  1  0 "Far_Left   "  %7.3f  %7.3f  %7.3f  %7.3f\n',
fltop,flbottom,flleft,flright);
fprintf(fid,'2=  2  0 "Far_Center "  %7.3f  %7.3f  %7.3f  %7.3f\n',
fctop,fcbottom,fcleft,fcright);
fprintf(fid,'3=  3  0 "Far_Right  "  %7.3f  %7.3f  %7.3f  %7.3f\n',
frtop,frbottom,frleft,frright);
fprintf(fid,'4=  4  0 "Near_Left  "  %7.3f  %7.3f  %7.3f  %7.3f\n',
nltop,nlbottom,nlleft,nlright);
fprintf(fid,'5=  5  0 "Near_Center"  %7.3f  %7.3f  %7.3f  %7.3f\n',
nctop,ncbottom,ncleft,ncright);
fprintf(fid,'6=  6  0 "Near_Right "  %7.3f  %7.3f  %7.3f  %7.3f\n',
nrtop,nrbottom,nrleft,nrright);
fprintf(fid,'7=  7  0 "Inst_Panel "  %7.3f  %7.3f  %7.3f  %7.3f\n',
iptop,ipbottom,ipleft,ipright);
fprintf(fid,'[Area of Interest Gains and Offsets]\n');
fprintf(fid,'1= 0.000 0.000 0.000
fclose(fid);
```