Paper ST-12

Eyes on the Road: A Methodology for Analyzing Complex Eye Tracking Data

Mary Anne Bertola and Stacy A. Balk
Science Applications International Corporation (SAIC), McLean, VA

ABSTRACT

Distracted driving is a relevant social issue with potentially devastating consequences. In part due to recent calls from President Obama and United States Transportation Secretary LaHood to curb distracted driving, research on the topic is becoming more prevalent. The use of eye tracking devices in on-road vehicles is an invaluable resource to investigate driver situational awareness and attention capture. Such tools provide insight into where drivers are looking, both within and outside the vehicle, while traveling down a roadway. Data from eye trackers in a real world environment, however, present a unique set of analysis challenges. For example, there are multiple ways to quantify visual behavior (e.g., duration of fixations, percentage of time, etc.) and such quantifications are constrained to non-negative values since a driver cannot look at an object for a negative amount of time. Additionally, responses are correlated since it is general practice to use eye movement data from one person over a period of time, as opposed to one specific instance in time. The GENMOD procedure in SAS® lends itself to accommodating such analysis challenges of eye tracking data through the use of generalized estimating equations which allow for restrictions on the values of a response variable and account for correlated measurements. This paper demonstrates the application of generalized estimating equations through the GENMOD procedure to analyze driver visual behavior in the presence of different roadway environments. Eye tracking devices are implemented in a variety of settings (e.g., training flight simulators, software usability, etc.). As such, it is hoped that analytical methodologies presented in this paper are also useful in the analysis of a variety of other eye tracking applications.

INTRODUCTION

Eye tracking is a relatively simple methodology that has long been used as a way to quantify eye movements. Early eye tracking devices commonly involved participants wearing contacts with embedded wiring that covered a large portion of the sclera (the white of the eye). Not only were these devices cumbersome, they also had the potential to be quite dangerous. However, most modern equipment utilizes simple active infrared light (not visible to humans) to create reflections from the eye that are captured by small camera systems. This setup provides the ability to unobtrusively, and relatively inexpensively, track eye movements (see Duchowski, 2007 for overview).

Advancements in technology have not only provided a simpler way to investigate eye movements, but have also provided mobility that has allowed eye trackers to be implemented in a wide variety of environments. One such example is the integration of an eye tracking system which allows the eyes (and often head movements) to act as computer mouse and keyboard controls (e.g., Chin, Barreto, Cremades, & Adjouadi, 2008). This integration provides the opportunity for those people who do not have functional use of their hands to use the computer in a more naturalistic way than might otherwise be possible. Eye trackers can also be integrated to create gaze-contingent fisheye effects on computer screens. That is, an area of a screen can be visually enhanced or enlarged by simply shifting visual attention (e.g., Ashmore, Duchowski, & Shoemaker, 2005). This technique might be helpful to those with severe visual impairments or those who must closely visually inspect images (e.g., x-rays). Beyond these, eye tracking can be used in a variety of marketing, human factors, and ergonomics applications. For example, eye tracking can tell researchers where consumers look on a webpage, how people scan for specific objects in a cluttered array, or even how pilots visually search cockpits. This visual scanning information has the potential to be used to enhance designs to best capture observer attention, when and where it is appropriate.

While eye tracking has long been used in indoor or predictable viewing environments (Young & Sheena, 1975), implementation in more dynamic environments is relatively recent. One such environment is the on-road vehicle. The remainder of this paper focuses on eye tracking and eye tracking data analysis in the automobile. Further, it investigates a case study involving data from an instrumented vehicle. In-vehicle eye tracking systems have the potential to provide insight into *where* and *how long* drivers are looking at environmental objects, both inside and outside of the vehicle. For example, it has been shown that visual patterns differ between novice and experienced drivers (e.g., Mourant & Rockwell, 1972). Findings such as these can be applied to teaching novice drivers where visual attention should be directed while driving. Beyond this, in-vehicle eye trackers even have the ability to detect drowsy driving in real time (Barr, Howarth, Popkin, & Carroll, 2005). More pertinent to this paper, though, is the capability of eye trackers to perform real time calculation of drivers' time looking away from the road, which could potentially indicate in-vehicle distracted driving (Ahlström & Kircher, 2010).

When examining the possible items drivers could be looking at while driving, it is clear that most items within the cab of the vehicle are stable (i.e., speedometer location, review mirror, gas gauge, etc.). However, drivers also spend a great deal of time looking at somewhat unpredictable and ever-changing environments. As a result, analyses of on-road eye tracking data often necessitate the grouping of environmental stimuli into naturally formed categories (e.g., road ahead, left side of the roadway, leading vehicles, driveways, etc.). The following discusses an approach to analyzing such a partitioning of the surrounding scene. It is hoped that this approach will help others to more easily evaluate and interpret their on-road eye tracking data.

EYE TRACKING

EQUIPMENT

On-road eye tracking data is typically captured through the use of a vehicle instrumented with an in-vehicle eye tracker system. These systems consist of single or multiple cameras directed at the driver's face. As the number of face cameras increases, so does the ability of the system to capture larger and more dramatic head movements of the driver. Similarly, an increase in the number of face cameras positively affects the overall precision and accuracy of the eye tracker system since the probability of convergence of both eye movements increases. Simpler systems consisting of one or two cameras, however, are usually less expensive and easier to install than more complex systems. In addition to the interior face camera(s), there are most often forward-facing camera(s) that collect video of the outside environment and roadway. Modern eye tracker systems integrate the eye movement data from the face camera(s) with the forward scene video so analysts can visually inspect where a person is looking while driving down a roadway.

In the case study investigated here, an instrumented vehicle with a three-camera eye tracker system was used. Three forward scene cameras were mounted on the roof of the vehicle above the driver's head. Each scene camera captured approximately 26^0 x 40^0 for a combined panoramic field of view of approximately 78^0 X 40^0. Figure 1 depicts the setup of the eye tracker system used for the case study.

Figure 1. Images of three forward scene cameras (LEFT) and three internal face cameras (RIGHT) in instrumented vehicle

Eye movement data is captured at a variety of speeds, typically ranging anywhere from 20 Hz to 500 Hz (frames per second) depending on the setting of the eye tracker and the type of equipment used. Eye trackers in a laboratory setting, for instance, tend to capture data at a much higher rate than those used in a field setting. Data for the case study presented here was collected at 60 Hz.

DEPENDENT MEASURES

Eye tracking data can be quantified in many ways (see Figure 2). For instance, one can examine driver eye-glance behavior in a frame-by-frame fashion. Another way is in terms of looks, which are consecutive frames to the same area in the driving environment. These methods, however, ignore driver intent and do not allow for smoothing of noise (an unavoidable error in the eye tracking system due to accuracy, mechanism calibration, etc.). One frame of noise could prematurely end a look calculation and thus the duration of a look may be shorter than what the participant truly experienced. Some eye tracking data analysis software packages allow for the calculation of fixations, which occur when the image of the target falls on the fovea (part of the retina responsible for sharp central vision or resolving fine detail) for a given period of time. A fourth technique considers glances, or consecutive fixations to the same area in the driving environment. These later methods tend to filter out system noise and generate data which mirror what is historically known about eye-movements. When frames of noise are encountered, the frames are likely absorbed into the current fixation or glance and then re-coded appropriately. Duration variables (e.g., duration of looks, fixations, glances, etc.) are generally measured in seconds or milliseconds; seconds will be used in this text.

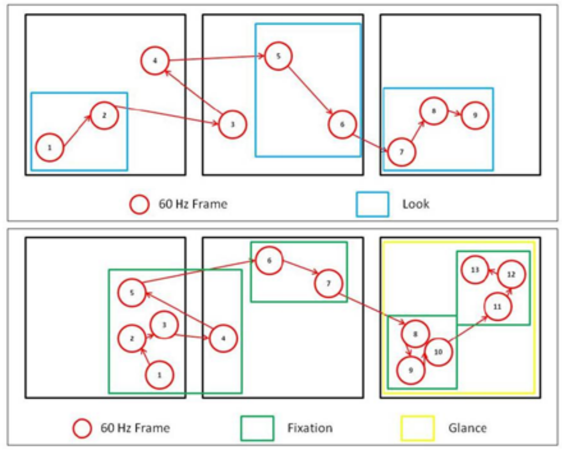

Figure 2. Visual interpretations of 60 Hz frame-by-frame measure (TOP and BOTTOM), look measure (TOP), fixation measure (BOTTOM), and glance measure (BOTTOM). Each of the three squares in the two images represents one area of the driving environment

Examples of typical summary information of eye tracking data include glance frequency, glance duration, and percentage of time to a target object. For ease of reading and calculation, mean percentage of time and mean fixation duration are discussed herein. However, the described methods can be applied to other dependent measures, as well.

CASE STUDY DESCRIPTION

Eye tracking data from 12 participants was collected while driving through 8 data collection zones (DCZs) which were approximately 960 ft in length apiece. Each DCZ was characterized as being in the absence or presence of off-premise vinyl billboards (ADVERTISING), which were large, vinyl signs advertising a business that was not present in the immediate area surrounding the sign. DCZs were also characterized by the level of non-essential visual stimuli presented in the scene, either low or high (CLUTTER). Examples of visual clutter included parking lots, schools, businesses, and roadside political advertisements. CLUTTER was a binary variable, with 0 representing low visual clutter and 1 representing high visual clutter. ADVERTISING was also a binary variable, with 0 representing the absence of an off-premise vinyl billboard and 1 representing the presence of an off-premise vinyl billboard. There were two DCZs representing each combination of the 2 (CLUTTER) x 2 (ADVERTISING) design. Figure 3 serves as a visual representation of two of the four combinations.

Figure 3. Example of low clutter DCZ with no off-premise advertising (LEFT) and high clutter DCZ with off-premise advertising (RIGHT)

For each 60 Hz frame, the location, or region, of where the participant was looking was recorded. Six regions of interest (ROIs) existed for this data set (see Table 1). ROIs can be defined as having static or dynamic movement, where fixations to dynamic ROIs tend to indicate that a driver was following an object with their eyes. Figure 4 depicts the ROIs visually.

Table 1. Descriptions of ROIs

ROI	Code	Description	Movement
Road Ahead	RA	Area of the forward roadway, approximately 26º x 40º in size	Static
Right Side of Road	Right	Area of the panoramic scene to the right of the forward roadway, approximately 26º x 40º in size (i.e., entering traffic, businesses, parking lots, trees, etc.)	Static
Left Side of Road	Left	Area of the panoramic scene to the left of the forward roadway, approximately 26º x 40º in size (i.e., on-coming traffic, businesses, parking lots, trees, etc.)	Static
Inside the Car	Car	Areas inside the cab of the vehicle (i.e., gas gauge, speedometer, etc.)	Static
Standard (Vinyl) Billboard	SBB	Predetermined off-premise vinyl billboard in the visual scene, as well as the area immediately surrounding the billboard	Dynamic
Unknown	Unknown	Areas where eye movements are tracking but the physical location is unknown (outside the panoramic scene)	Static

Figure 4. Visual display of ROIs used in the case study data

A sample of the frame-by-frame data set, Data_60, is shown in Table 2. There were approximately 7,200 frames per participant.

Table 2. Sample of 60 Hz data set, Data_60

Participant	Route	DCZ	ROI	Clutter	Advertising
1	1	1	Unknown	0	1
1	1	1	Unknown	0	1
1	1	1	SBB	0	1
1	1	1	RA	0	1
			…		
12	2	8	Left	1	1
12	2	8	Car	1	1
12	2	8	RA	1	1
12	2	8	RA	1	1

An eye tracking data analysis software program was used to determine fixation durations to ROIs. Although not determined by using the physiological definition of such eye movements, fixations were calculated using three criteria: (1) a minimum time (0.1 s), (2) a radius (0.03 radians), and (3) a maximum outlier time (0.08 s). In other words, if a series of frames met these three criteria then the software deemed the series a fixation. A sample of such a data set, Data_Fix, is shown in Table 3. There were approximately 225 fixations per participant.

Table 3. Sample of fixation duration data set, Data_Fix

Participant	Route	DCZ	ROI	Duration (s)	Clutter	Advertising
1	1	1	RA	0.2333	0	1
1	1	1	RA	0.3167	0	1
1	1	1	RA	0.3333	0	1
1	1	1	RA	0.4500	0	1
…						
12	2	8	RA	0.2667	1	1
12	2	8	RA	0.3500	1	1
12	2	8	RA	0.3333	1	1
12	2	8	RA	0.2167	1	1

Data_60 and Data_Fix are summaries of the same set of driver eye movements; however, each serves as a different quantification of those movements. Both will be used throughout this text to demonstrate the use of generalized estimating equations. Moreover, Data_60 will be used to analyze mean percentage of time and Data_Fix will be used to analyze mean fixation duration. The next sections describe this methodology in general, as well as its applicability to on-road eye tracking data.

GENERALIZED ESTIMATING EQUATIONS

By the nature of its data collection techniques, eye tracking data are repeated measurements from one individual and thus responses are correlated. In these cases, assuming independent observations tends to produce incorrect standard errors (Stokes, Davis, & Koch, 2000). To mitigate this inflation, correlations among responses must be taken into consideration. One statistical methodology which handles repeated measures data is Generalized Estimating Equations (GEEs), which uses a quasi-likelihood model instead of a maximum likelihood model. While a maximum likelihood model assumes a particular distribution for the response variable, a quasi-likelihood model only requires that the first two moments of the response variable be specified (Agresti, 2002). Given that the first moment is correctly specified, it can be shown that the resultant parameter estimates from the GEE technique are consistent. In other words, the GEE method is robust and parameter estimates are consistent even if the covariance matrix is incorrectly specified (Agresti, 2002). Such a property is useful since correlation structure is not always known ahead of time.

GEEs can handle missing observations, continuous explanatory variables, and time-dependent explanatory variables. In addition, GEEs reduce to Generalized Liner Model (GLM) estimating equations when an independent working correlation structure is used and are the maximum likelihood score equations for multivariate Gaussian data when an unstructured working correlation structure is used (Stokes et al., 2000). Furthermore, GEE relies on asymptotic theory; hence, the number of clusters used must support consistency of the parameter estimates. The determination of a “large enough” sample size also depends on the number of explanatory variables; if there are few predictors in the model, then a smaller number of clusters may be adequate (Stokes et al. , 2000).

One drawback of this method is that few model selection criteria exist since most options are maximum likelihood based. Pan (2001) proposed the Quasi-likelihood under the Independence model Criterion (QIC), which involves a slight alteration to the Akaike’s Information Criterion (AIC) that is typically used for model selection in linear regression involving independent observations. Beginning with SAS Version 9.2, the QIC statistic is included in the output from the GENMOD procedure.

BASIC PROGRAM: PROC GENMOD

In general, data analyzed using the GEE method can be modeled using a GLM, with the exception of correlated responses. Thus, one would assume a similar programming structure in SAS for implementing a GEE as one would use for implementing a GLM. In PROC GENMOD, the REPEATED statement is added in order to request a GEE analysis; the remainder of the procedure is typically the same as that used for a GLM. The basic program is as follows.

```
PROC GENMOD DATA=dataset;
   CLASS classification variable(s);
   MODEL response variable = predictor variable(s) / DIST=response distribution
      LINK=link function;
   REPEATED SUBJECT=clustering variable(s) / TYPE=working correlation
      structure;
RUN;
```

It is important to note that the clustering variable(s) listed in the SUBJECT= option must have a unique value for each cluster and also be listed in the CLASS statement. In the case of eye tracking data, the clustering variable will most likely be the unique identifier for the participants. However, for this particular case study the clustering variable will be DCZ within participant. By design, it is not simply the participant effect that is driving the response, but also the specific environment presented in the different DCZs. Thus, each DCZ for each participant will serve as one cluster of data, producing a total of 96 clusters (12 participants X 8 DCZs).

Table 4 demonstrates the allowable response distributions (DIST= option) and associated default link functions in SAS. Such a table can also be found in the SAS®/STAT User's Guide, Version 9.22 (2010), under PROC GENMOD. The default link function can be altered by using the LINK= option in order to account for the suspected nature of the response variable. If the default link function is desired, then including only the DIST= option is sufficient.

Table 4. Allowable response distributions and associated default link functions

Distribution	Default Link Function
Binomial	Logit
Gamma	Inverse
Geometric	Log
Inverse Gaussian	Inverse Squared
Multinomial	Cumulative Logit
Negative Binomial	Log
Normal	Identity
Poisson	Log
Zero-Inflated Poisson	Log/Logit
Zero-Inflated Negative Binomial	Log/Logit

Generally speaking, eye tracking data are restricted to non-negative values since a participant cannot look at an ROI of a negative amount of time. Hence, distributions which extend below zero are excluded. Similarly, percentages of time and fixation durations are continuous and as a result discrete distributions are also excluded. Thus, an appropriate response distribution would be either the Gamma distribution or the inverse Gaussian distribution since these are the only two allowable response distributions which meet both criteria. However, the support of both distributions exclude zero and so a small correction factor must be added (for example, 0.00001) in situations where the mean percentage of time or mean fixation duration equals zero (i.e., a DCZ when a participant does not look at a particular ROI). The Gamma distribution will be assumed here for simplicity.

Both the mean percentage of time and mean fixation duration response data are restricted to the $(0, \infty)$ range. Thus, choosing a (natural) log link transforms the lower bound to $-\infty$. As such, each parameter estimate corresponds to the additive change in the log of the mean response (either percentage of time or fixation duration). Similarly, e^{β} corresponds to the multiplicative factor by which the mean response changes. Figure 5 shows the corresponding GEE model for a single predictor (the log of the mean response; Equation 1.1) and the associated equation for the mean response (Equation 1.2).

$$log[\mu(x)] = \alpha + \beta x \qquad \text{(Eq. 1.1)}$$

$$\mu(x) = e^{\alpha+\beta x} = [e^{\alpha}][e^{\beta}]^{x} \qquad \text{(Eq. 1.2)}$$

Figure 5. Equations for the corresponding GEE model (log of the mean response; Equation 1.1) and the mean response (Equation 1.2) for a single predictor

Stokes et al. (2000) suggest several possibilities for the working correlation structure of the response (TYPE= option). Table 5 describes the various structures allowed in SAS, as well as a brief explanation of each. Given eye tracking data are in clusters, an appropriate working correlation structure would be the exchangeable structure.

Table 5. Allowable working correlation structures and descriptions

Independent	• Repeated observations for a participant are independent • GEE simplifies to GLM estimating equations • Default correlation structure in SAS
Fixed	• User-specified correlation matrix
Exchangeable	• Correlations between any two measurements within a subject is constant • Reasonable when repeated measurements are not obtained over time • Appropriate when cluster sampling is involved
Unstructured	• Provides most efficient estimator when there are relatively few observation times or conditions
***m*-dependent**	• Correlations depend on the distances between measures • Correlations diminish to zero for $s \geq m$
Auto-regressive (AR-1)	• Correlations depend on the distances between measures • Correlations diminish with increasing distance

Other options, such as the TYPE3 and WALD options, exist for this procedure. The TYPE3 option of the MODEL statement refers to Type 3 contrasts of the model effects. By default, PROC GENMOD produces likelihood ratio statistics for these contrasts. Given that GEE uses quasi-likelihood, a Wald statistic may be more preferable. This can be achieved by adding the WALD option to the MODEL statement, in addition to the TYPE3 option. For other available statements and options (e.g., CONTRAST, ESTIMATE, OUTPUT, etc.) refer to the PROC GENMOD section of the SAS/STAT User's Guide, Version 9.22 (2010).

ANALYZING ON-ROAD EYE TRACKING DATA: A CASE STUDY

The suitable response distribution, link function, and working covariance structure must be determined a priori in order to conduct the analysis. Although there are multiple ways of quantifying driver visual behavior, only mean percentage of time within a DCZ and mean fixation durations will be considered herein.

MEAN PERCENTAGE OF TIME DATA

When mean percentage of time data is analyzed as the dependent measure then PROC FREQ can be used to prepare the data for the GEE analysis. The program below assists in preparing Data_60 (see Table 2).

```
PROC FREQ DATA=Data_60 NOPRINT;
   BY Participant DCZ;
   TABLES ROI / NOFREQ NOCUM OUT=Freq1;
RUN;
```

It is important to note that there may be cases where a participant will not look at every ROI category in a DCZ. For instance, in those DCZs where no off-premise vinyl billboards are present (ADVERTISING = 0) there should not be frames coded as SBB and so the percentage of time to SBB for those DCZs should be zero. However, due to the nature of PROC FREQ, the absence of frames coded as SBB will not result in a zero percentage but rather the category will be ignored. Thus, it may be necessary to include "created data" as a way of ensuring that the resulting data set, Freq1, contains 576 observations (8 DCZ X 6 ROI = 48 observations per participant). In this case study, for example, "created data" consisting of 3 variables (PARTICIPANT, DCZ, and ROI) and 576 observations would be combined with Freq1 in a DATA step by using a MERGE statement and IN= option for the "created data." The IN= option controls which observations are included, hence guaranteeing that the resultant data set will have 576 observations. Next, one need only change the COUNT and PERCENT values from missing (.) to 0 for those observations which were not originally in Freq1. Freq1 will be used in the GEE analysis of mean percentage of time.

Through model selection techniques, the following program was used to analyze the mean percentage of time data. Since the glance behavior of a participant is being modeled, an intercept term is not practically interpretable and hence excluded by using the NOINT option on the MODEL statement.

```
PROC GENMOD DATA=Freq1;
   CLASS Participant DCZ ROI Clutter Advertising / ORDER=DATA;
   MODEL Percent = ROI|Clutter ROI|Advertising / DIST=GAMMA LINK=LOG NOINT
      WALD TYPE3;
   REPEATED SUBJECT=Participant*DCZ / TYPE=EXCH;
RUN;
```

Predictors for this model include ROI (R), CLUTTER (C), ADVERTISING (A), ROI*CLUTTER (R*C), and ROI*ADVERTISING (R*A). Equation 1.3, which is a modified version of Equation 1.1, can be used to estimate the log mean percentage of time based on these predictors (see Figure 6). Similarly, Equation 1.4 serves as a modified version of Equation 1.2 and can be used to estimate the mean percentage of time.

$$log[\mu(x)] = \beta_R x_R + \beta_C x_C + \beta_A x_A + \beta_{R*C} x_{R*C} + \beta_{R*A} x_{R*A} \qquad \text{(Eq. 1.3)}$$

$$\mu(x) = e^{\beta_R x_R + \beta_C x_C + \beta_A x_A + \beta_{R*C} x_{R*C} + \beta_{R*A} x_{R*A}} \qquad \text{(Eq. 1.4)}$$

Figure 6. Equations for the corresponding GEE model (log of the mean percentage of time; Equation 1.3) and the mean percentage of time (Equation 1.4) for ROI (R), CLUTTER (C), ADVERTISING (A), ROI*CLUTTER (R*C), and ROI*ADVERTISING (R*A)

Thirty five parameters were estimated in the model analyzing the mean percentage of time data. A sample of the parameter estimates table is shown by Table 6. Recall that rows of 0.0000 and missing values (.) indicate that a parameter was not estimated (see Intercept and Clutter – 0 rows in Table 6 for an example). Thus, the lack of a parameter estimate for the intercept confirms that the NOINT option is in effect.

Table 6. Sample parameter estimates for modeling mean percentage of time data (remainder of the table has been suppressed for space)

Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept		0.0000	0.0000	0.0000	0.0000	.	.
ROI	RA	4.4464	0.0162	4.4147	4.4781	274.83	<.0001
ROI	Right	0.8183	0.2382	0.3514	1.2852	3.43	0.0006
ROI	Left	0.9726	0.2222	0.5371	1.4080	4.38	<.0001
ROI	Car	1.8452	0.1759	1.5005	2.1899	10.49	<.0001
ROI	Unknown	1.1067	0.1586	0.7957	1.4176	6.98	<.0001
ROI	SBB	-11.3602	0.1241	-11.6034	-11.1171	-91.57	<.0001
Clutter	1	-0.2852	0.2447	-0.7648	0.1945	-1.17	0.2439
Clutter	0	0.0000	0.0000	0.0000	0.0000	.	.
			…				

The log of the mean percentage of time can be estimated when looking at a particular ROI category given the level of clutter and advertising in the surrounding scene by using Equation 1.3 (see Figure 6) and respective parameter estimates (see Table 6 for a partial listing). Similarly, Equation 1.4 can be used to estimate the mean percentage of time under the same conditions. Example calculations are shown in Figure 7 for the situation of looking at the road ahead (ROI = RA) in a high clutter environment (CLUTTER = 1) with no off-premise vinyl billboards (ADVERTISING = 0). Recall that the percentage corresponds to the percentage of time in one DCZ. Given these scene characteristics, a participant is expected to look at the road ahead approximately 82 percent of the time while traveling through one DCZ. The OUTPUT statement and PREDICTED specification can be used to automatically calculate the mean percentage of time for the various combinations of the explanatory variables.

$$log[\mu(x)] = 4.4464 - 0.2852 + 0.2494 + 0.0000 + 0.0000 = 4.4106$$

$$\mu(x) = e^{4.4106} = 82.3188$$

Figure 7. Example calculations for the estimated log mean percentage of time (TOP) and the mean percentage of time (BOTTOM) when looking at a particular ROI category given the level of clutter and advertising in the surrounding scene

Wald Type 3 analyses, shown in Table 7, indicate which predictors are significant in the GEE model for the mean percentage of time data at the 5% significance level. Both second-order interaction terms included in the model are influential, as well as the main effects for ROI category and level of ADVERTISING.

Table 7. Type 3 analysis for the mean percentage of time data (significant factors in italic bold)

Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**X^2**	**Pr > X^2**
ROI	5	8353.73	***<.0001***
Clutter	1	2.73	0.0984
Advertising	1	547.78	***<.0001***
ROI*Clutter	5	15.60	***0.0081***
ROI*Advertising	5	2622.88	***<.0001***

MEAN FIXATION DURATION DATA

If mean fixation duration is analyzed as the dependent measure then PROC MEANS can be used to prepare the data for the GEE analysis. The program below assists in preparing Data_Fix (see Table 3).

```
PROC MEANS DATA=Data_fix MEAN NOPRINT NWAY;
   CLASS Participant ROI;
   VAR Duration;
   OUTPUT OUT=Means1 MEAN=Mean_Fix_Dur;
RUN;
```

As was the case with the mean percentage of time data, there may be instances where a participant will not look at every ROI category in a DCZ. Hence, "created data" may be necessary for the mean fixation duration data, as well. The resulting data set, Means1, should contain 576 observations (8 DCZ X 6 ROI = 48 observations per participant). Means1 will be used in the GEE analysis of mean fixation duration.

Through model selection techniques, the following program was used to analyze the mean fixation duration data. Since the glance behavior of a participant is being modeled, an intercept term is not practically interpretable and hence excluded by using the NOINT option on the MODEL statement.

```
PROC GENMOD DATA=Means1;
   CLASS Participant DCZ ROI Clutter Advertising / ORDER=DATA;
   MODEL Mean_Fix_Dur = ROI|Clutter ROI|Advertising / DIST=GAMMA LINK=LOG NOINT
      WALD TYPE3;
   REPEATED SUBJECT=Participant*DCZ / TYPE=EXCH;
RUN;
```

Predictors for this model include ROI (R), CLUTTER (C), ADVERTISING (A), ROI*CLUTTER (RC), and ROI*ADVERTISING (RA). Equation 1.5, which is a modified version of Equation 1.1, can be used to estimate the log mean fixation duration based on these predictors (see Figure 8). Similarly, Equation 1.6 serves as a modified version of Equation 1.2 and can be used to estimate the mean fixation duration.

$$log[\mu(x)] = \beta_R x_R + \beta_C x_C + \beta_A x_A + \beta_{RC} x_{RC} + \beta_{RA} x_{RA} \qquad \text{(Eq. 1.5)}$$

$$\mu(x) = e^{\beta_R x_R + \beta_C x_C + \beta_A x_A + \beta_{R*C} x_{R*C} + \beta_{R*A} x_{R*A}} \qquad \text{(Eq. 1.6)}$$

Figure 8. Equations for the corresponding GEE model (log of the mean fixation duration; Equation 1.5) and the mean fixation duration (Equation 1.6) for ROI (R), CLUTTER (C), ADVERTISING (A), ROI*CLUTTER (R*C), and ROI*ADVERTISING (R*A)

Thirty five parameters were estimated in the model analyzing the mean fixation duration data. A sample of the parameter estimates table is shown by Table 8. The lack of a parameter estimate for the intercept confirms that the NOINT option is in effect.

Table 8. Sample parameter estimates for modeling mean fixation duration data (remainder of the table has been suppressed for space)

Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept		0.0000	0.0000	0.0000	0.0000	.	.
ROI	RA	-0.8035	0.0539	-0.9090	-0.6979	-14.92	<.0001
ROI	Right	-2.0968	0.2234	-2.5347	-1.6589	-9.38	<.0001
ROI	Left	-2.1058	0.2400	-2.5761	-1.6354	-8.77	<.0001
ROI	Car	-1.9207	0.2429	-2.3967	-1.4447	-7.91	<.0001
ROI	Unknown	-2.0870	0.3175	-2.7093	-1.4648	-6.57	<.0001
ROI	SBB	-11.5476	0.1826	-11.9054	-11.1898	-63.25	<.0001
Clutter	1	0.0706	0.3650	-0.6448	0.7859	0.19	0.8466
Clutter	0	0.0000	0.0000	0.0000	0.0000	.	.
			…				

The log of the mean fixation duration can be estimated when looking at a particular ROI category given the level of clutter and advertising in the surrounding scene by using Equation 1.5 (see Figure 8) and respective parameter estimates (see Table 8 for a partial listing). Similarly, Equation 1.6 can be used to estimate the mean fixation duration under the same conditions. Example calculations are shown in Figure 9 for the situation of looking at the road ahead (ROI = RA) in a high clutter environment (CLUTTER = 1) with no off-premise vinyl billboards (ADVERTISING = 0). Recall that fixation durations are in seconds. Given these scene characteristics, a participant would be expected to fixate on the road ahead in approximately 0.5 s intervals while driving through one DCZ. The OUTPUT statement and PREDICTED specification can be used to automatically calculate the mean fixation duration for the various combinations of the explanatory variables.

$$log[\mu(x)] = 0.8035 + 0.0706 - 0.0166 + 0.0000 + 0.0000 = -0.7495$$
$$\mu(x) = e^{-0.7495} = 0.4726$$

Figure 9. Example calculations for the estimated log mean fixation duration (TOP) and the mean fixation duration (BOTTOM) when looking at a particular ROI category given the level of clutter and advertising in the surrounding scene

Wald Type 3 analyses, shown in Table 9, indicate which predictors are significant in the GEE model for the mean fixation duration data at the 5% significance level. Both second-order interaction terms included in the model are influential, as well as the main effects for ROI category and level of ADVERTISING.

Table 9. Type 3 analysis for mean fixation duration data (significant factors in italic bold)

Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**X^2**	**Pr > X^2**
ROI	5	1534.34	***<.0001***
Clutter	1	2.17	0.1404
Advertising	1	111.72	***<.0001***
ROI*Clutter	5	15.73	***0.0077***
ROI*Advertising	5	731.73	***<.0001***

DISCUSSION

Results from the GEE methodology allow one to approximate how a driver will direct their visual attention. Data from the case study indicated that while the ROI category had a significant effect on how often the participant looked to that area, the interaction of the ROI category with the level of CLUTTER and ADVERTISING was also important. These results suggest that analysts need to consider the characteristics of the scene and not just the magnitude of the ROI in question. In other words, one must determine which predictors and interactions among the predictors have a significant effect on the dependent variable.

By design, the frame-by-frame data set, Data_60, depicts a multinomial sampling design since for each frame the response (ROI category) has six options. Thus, a GEE which assumed a multinomial distribution and logit link function (a multinomial logistic model) may be more appropriate. Kuss and McLerran (2007) note, though, that as of SAS Version 9.1, PROC GENMOD only allows this technique for multinomial ordinal data and the cumulative logit link function. In this particular case study, the data are nominal and lack a specific order so the use of an ordinal model is not practically useful. Kuss and McLerran (2007) propose using a multivariate binary model as an alternative; however, the authors were unable to successfully implement that approach with the case study data.

In some situations, creating a bridge between driver attention and eye movement behavior may be best understood by exploring the time spent looking away from the forward roadway. Given these circumstances, the ROIs used in the case study could be collapsed into two categories: Road Ahead (RA) and not Road Ahead (Right, Left, Car, SBB, and Unknown). Thus, a binomial response distribution and logit link function (a binomial logistic model) would be more appropriate so as to determine the probability of looking at the Road Ahead.

In summary, eye trackers are an invaluable resource for the research community. These devices shed light onto where participants are looking while driving, as in the case study, or while completing numerous other activities, such as exploring a website or using a flight simulator. Such knowledge helps researchers to understand how people visually navigate and become spatially aware of the scene in front of them. In the world of eye trackers, it is now possible to transport oneself into the eyes of a person and "see" what he/she sees. Consequences of this ability range from the possibility to detect drowsiness or inattention, in the case of driving, to assisting physically or visually impaired individuals in a variety of ways. Correct analysis of eye tracking data is imperative given these consequences and the first step is to understand the research question being asked. As was shown earlier, the interpretation of the ROIs plays a major role in which response distribution is chosen; in some instances only two categories are enough while in other cases a larger quantity is necessary in order to capture the true nature of the eye movements. In the end, GEE is a simple and effective technique that can be used to better understand and interpret eye tracking data. It is hoped that this method will help others to better understand their own complex eye tracking data and hence vastly expand the knowledge base about basic eye movements.

REFERENCES

Agresti, A. (2002). Analyzing repeated categorical response data. In Balding et al. (Eds.), *Categorical data analysis, 2nd edition* (pp. 455-490). Hoboken, NJ: John Wiley & Sons, Inc.

Ahlström, C., & Kircher, K. (2010). Review of real-time visual driver distraction detection algorithms. *Proceedings of Measuring Behavior.* Eindhoven, The Netherlands.

Ashmore, M., Duchowski, A. T., & Shoemaker, G. (2005). Efficient eye pointing with a fisheye lens. *Proceedings of the Canadian Human-Computer Communications Society (CHCCS)/ACM.* Victoria, BC, Canada.

Barr, L., Howarth, H., Popkin, S., & Carroll, R. (2005). A review and evaluation of emerging driver fatigue detection measures and technologies. *Proceedings of the 2005 International Conference on Fatigue Management in Transportation Operations.* Seattle, Washington.

Chin, C. A., Barreto, A., Cremades, J. G., & Adjouadi, M. (2008). Integrated electromyogram and eye gaze tracking cursor control system for computer users with motor disabilities. *Journal of Rehabilitation Research and Development, 45*, 161-174.

Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and Practice 2nd edition*. London: Springer-Verlag.

Kuss, O., & McLerran, D. (2007). A note on the estimation of the multinomial logistic model with correlated responses in SAS, *Computer Methods and Programs in Biomedicine*, 87(3), 262-269.

Mourant, R. R., & Rockwell, T. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors, 14*, 325-335

Pan, W. (2001). Akaike's Information Criterion in Generalized Estimating Equations, *Biometrics*, 57(1), 120-125.

SAS Institute, Inc. (Eds). (2010) *SAS/STAT User's Guide, Version 9.22*. Cary, NC: SAS Institute, Inc.

Stokes, M. E., Davis, C. S., & Koch, G. G. (2000). *Categorical data analysis using the SAS system* (2nd ed.). Cary, NC: SAS Institute, Inc.

Young, L., & Sheena, D. (1975). Methods and designs: Survey of eye movement recording methods. *Behavior Research Methods and Instruments, 7*, 397-429.

ACKNOWLEDGMENTS

This research was financially supported by FHWA contract #DTFH61-08-C-00006.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Mary Anne Bertola, M.S.
Science Applications International Corporation (SAIC)
6300 Georgetown Pike, Room F-215
McLean, VA 22101
Work: (202) 493-3379
Fax: (202) 493-3390
maryanne.bertola.ctr@dot.gov
www.saic.com

Stacy A. Balk, Ph.D.
Science Applications International Corporation (SAIC)
6300 Georgetown Pike, Room F-215
McLean, VA 22101
Work: (202) 493-3423
Fax: (202) 493-3390
stacy.balk.ctr@dot.gov
www.saic.com