SAE International

The Accuracy of Pedestrians in Estimating the Speed of a Moving Vehicle

2013-01-0785

Published 04/08/2013

Mark Strauss, James Carnahan
Ruhl Forensic, Inc. / Univ. of Illinois Urbana

Roland Ruhl
Ruhl Forensic, Inc.

doi:10.4271/2013-01-0785

ABSTRACT

This study was performed in order to evaluate the accuracy with which pedestrians estimate the speed of nearby automobiles. A total of 87 subjects were involved in this experiment, with 487 useful speed estimates being obtained from them. The vehicle speeds were measured using radar guns, and the moving cars were located upstream, downstream, and adjacent to the subjects. A survey form was used to obtain attributes about each subject. A multivariate regression was used to allocate portions of the variation in the percent error among the various factors that describe the subject and the moving vehicle. Statistical significance was discovered for the main effects and some interactions among: gender, driving experience, a self-assessment of ability to estimate such speeds, the use of the posted speed limit, the speed of the target vehicle, and the location of the vehicle on the roadway. The implications for speed estimation accuracy are presented, both for categories of subjects and for individual subjects. The standard deviation in the accuracy of individuals was remarkable for its large magnitude and its persistence over all the categories of subjects and also the different vehicle speeds and locations.

INTRODUCTION

People estimate the speeds of moving vehicles for a variety of reasons, such as the driver who is deciding when to make a left turn across moving traffic, the pedestrian who is deciding when it is safe to cross a street, or a person who sees a collision and provides information about what he saw. The purpose of this study was to provide a controlled environment in order to evaluate the ability of an ordinary pedestrian to estimate the speeds of vehicles that were approaching or moving away at different viewing angles, and at different vehicle speeds. To this end, an experiment was designed and conducted in which roadside pedestrians were asked to estimate the speeds of vehicles traveling in nearby lanes of actual traffic. Those vehicle speeds were measured using radar guns; the measured speeds varied from 37 to 81 kph (23 to 50 mph) during the experiment. Age, driving experience and other subject characteristics were obtained for each of the subjects and were used in the analysis.

BACKGROUND

A review of the literature was conducted and it was found that a great number of studies had to do with estimation of speed by drivers or passengers in a moving car, rather than by stationary pedestrians. In addition, some of the studies used a closed track for the experiment, rather than a public roadway.

Conchillo et al. (2006) studied the speed estimation capability of a passenger traveling in a car in different scenarios, including a closed track, and in real traffic conditions on a secondary road and highway. The passengers estimated the speed of the car in which they were traveling. The researchers found that their participants were more accurate when on the closed track without traffic than they were in traffic on a public road. That paper contains a review of published research for both simulated and real traffic conditions; it suggests that the two situations have not been shown to be equivalent. They note that the literature reports a general tendency to underestimate the speed of the vehicle in which the observer is traveling.

Recarte and Nunes (1996) performed an investigation in which passengers estimated the speed of a vehicle in which they were traveling; the drivers also attempted to produce a target speed which is also an estimation task. They also found that there was a tendency for drivers and passengers to underestimate speed, but that estimation error was reduced at higher speeds. They also found that error was reduced for subjects with more driving experience.

As mentioned, there are a number of papers that address the estimation of speed by a driver with a comparison to measured speeds. Although these are not very comparable to the study presented here, several are provided here in the event that the reader wishes to begin a literature search on this subject. A partial list of papers is: Evans (1970), Milosevic and Milic (1990), and Aberg et al., (1997). Some articles that specifically address the testimony of witnesses are: McAllister et al. (1988), Kebell and Giles (2000), and Kebell et al. (2002).

Scialfa et al. (1991) conducted an experiment with observers who estimated the speed of an isolated vehicle traveling between 20 and 84 kph on a closed track. The observers were categorized into 3 age groups. The stated goal for the study was to determine whether the ability to estimate vehicle velocity changes in a systematic way as people age. Among the many findings discussed, they reported a tendency for observers to underestimate lower speeds and overestimate higher speeds. This is the opposite of one of our findings and perhaps is due to the differences between a closed track and actual traffic.

An article by Hutchinson (1995) reviews 3 unpublished works that the authors of the present paper could not access in their original form. Nonetheless, Hutchinson's summaries of these works will be provided here because they address speed estimation by observers at the side of a road, as opposed to a driver or passenger in a moving vehicle, as was the case in many other studies. Hutchinson describes the work of Green (1954) as reporting high correlation between actual and estimated speed with a tendency to underestimate by about 8 kph. The work by Hurford (1968) also reported high correlation between true and estimated speed, again with consistent underestimation. Finally the work by Hinch (1967) was contrasting, since he indicated that his pedestrian subjects were unable to estimate speeds with any accuracy, with no bias toward underestimation.

Strauss and Carnahan (2010, 2010) have published multiple studies of the error that observers make when they estimate distances to static objects in a roadside setting. The results were very interesting and motivated the current work on speed estimation.

METHODS

Testing took place on three consecutive days, on a straight and flat 6-lane roadway that has a speed limit of 72 kph (45 mph). The pedestrians who were being tested stood at the point marked "PED" as shown in Figure 1. One of three very similar SUV's was driven in the right-most lane going either northbound or southbound. As this test vehicle was approaching or moving away, the pedestrians were asked not to look at the road. At the correct time, the pedestrians were asked to turn around and look at the test vehicle for 2-3 seconds when it was at one of 6 different locations, as noted in Figure 1 as "1" through "6." The observers then provided an estimate of speed for that vehicle, at that location. The test was repeated for each of the five other vehicle locations.

In addition to providing speed estimates, the subjects were interviewed and asked to rank themselves as to how well they believed they could estimate vehicle speed, and to provide other information such as their age, clarity of eyesight, work history, and hobbies. A total of 487 vehicle speed estimates were given by the 87 research subjects (41 female, 46 male) who were 18-76 years old.

The test subject sample set was comprised of volunteers who approached the study area and were then recruited to participate. They were not carefully selected from a larger pool to control variables such as age and sex. Although it was not possible to create a perfectly balanced experimental design for the characteristics of the volunteer test subjects, experimental design methods were used to conduct the rest of the experiment. The sequence of the locations of the vehicles presented to the subjects was randomized, along with the target speeds that the test vehicle drivers attempted to achieve. Those target speeds were described to the test vehicle drivers as: "48 kph" (30 mph), "64 kph" (40 mph), and "faster than 64 kph (40 mph) if possible, but otherwise with the flow of traffic." The radar gun measured speeds achieved by the test vehicles in response to those target speeds were, respectively: 43 - 64 kph (27 - 40 mph), 53 - 69 kph (33 - 43 mph), and 37 - 81 kph (23 - 50 mph). The *measured* vehicle speeds were used in all analyses to represent the actual vehicle speeds.

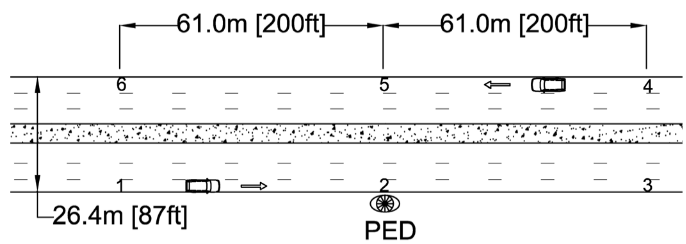

Figure 1. Testing configuration showing position of the subject pedestrian ("PED") and the locations (1 - 6) of the test vehicles when their speeds were estimated.

RESULTS

There were a total of 522 opportunities for vehicle speeds to be estimated (87 subjects and 6 vehicles each). The pedestrians did not provide speed estimates 29 times, leaving 493 occasions to compare an estimate with a measured speed in order to calculate the error. Most of the results in this paper are presented as percent error in speed estimation, which was calculated as:

$$\%\,Error = \frac{Estimated\ Speed - Actual\ Speed}{Actual\ Speed} * 100$$

Note that an overestimation of vehicle speed results in a positive percent error, and an underestimation results in a negative percent error.

Identification of Outliers

In order to investigate the possible presence of speed estimate outliers, a box plot procedure was utilized. It indicated the presence of 6 outlier responses with the following errors: −66%, −63%, 59%, 61%, 73%, 84%. The outlier criteria embedded in Minitab were used and they suggest outer bounds of −61.7% and 58.3%. The bounds follow standard practice and are based on the 25th and 75th percentiles ±150% of the interquartile range as discussed by Montgomery and Runger (1999), for example. These 6 speed estimates were removed from the data set; no two outliers came from the same subject. This left a total of 487 speed estimates that were used in the remainder of the analysis. These 487 speed estimates are referred to as the "overall" data set.

In order to explore the ramifications of excluding the 6 outlier responses in the data, nearly all of the calculations were repeated with the outliers included, and there were only minor differences. The adjusted R^2 decreased slightly from 25.3% to 23.8%, with the loss of statistical significance for the interaction term between the speed category and whether the subject used the speed limit in estimation. The p-value for that term increased from 0.048 (which, perhaps, was an indicator of borderline statistical significance to begin with) to 0.097, passing the commonly accepted threshold of 0.05. Otherwise the results were very similar, especially regarding the magnitudes of the numerous coefficient fits in the regression model and the consequent estimates of the least square means.

Analysis with Outliers Removed

As previously shown, there were six speed estimate responses that were found to be outliers. These six responses have been removed from the data set and will no longer be part of the analysis or results presented in this paper.

The data for percent error for all 487 speed estimates that make up the "overall" data set is shown as a histogram in Figure 2 along with the normal distribution curve associated with that mean and standard deviation. The normal distribution curve will be shown for all future graphs. Statistical tests showed a good fit to a normal model except at the extreme tails of the distribution. Figure 2 includes speed estimates for all the vehicles, at all speeds, at all 6 locations, except for the six outliers. The sample mean (−1.38%) is close to zero and slightly negative, which is a small underestimation of the actual vehicle speeds. It is important to realize that this result reflects the frequency of occurrence of the various characteristics of this sampled population, such as the varied driving experience, gender, and the subjects' self-assessment of their speed estimation ability.

The regression analysis shown later in this paper is used to approximate the mean percent error for the general population (not this particular sampled population) which results in a quite different estimate of −10.5%. This contrasting mean estimate is appropriate when the subject is equally likely to come from any of the factor categories: gender, extent of driving experience, a self-assessed rating of speed estimation ability, and whether subjects used their knowledge of the speed limit to estimate speed.

The standard deviation for the percent error in speed estimation for the "overall" data set was 21.7%. This is quite large and is a measure of the extreme variability amongst the 87 pedestrian observers in this subject population. This large standard deviation persisted throughout the various categorizations of subjects (which are shown in detail later), and is likely to be representative of the general population.

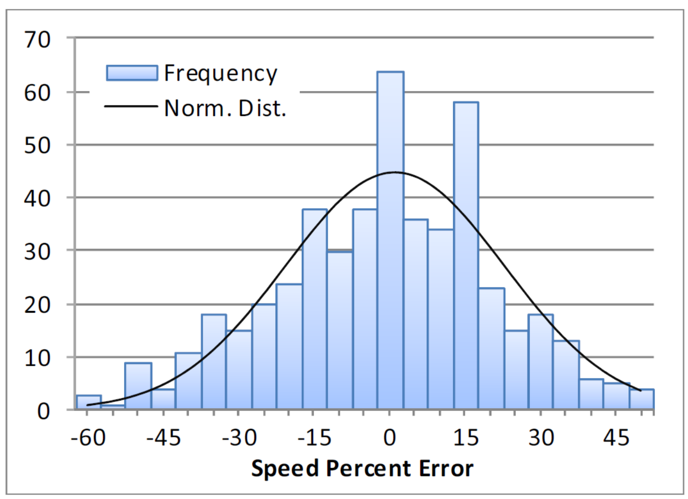

Figure 2. Histogram of the "overall" data sample; estimated speed percent error. (All vehicles, all speeds, all 6 locations, all subjects.) Mean = −1.38, StDev=21.68, N=487

The next step was to explore which factors affect speed estimation ability. Six factors were studied, four related to the subjects and two related to the vehicles: gender, years of driving experience, a self-assessment of speed estimation ability, whether test subjects replied that they had used their knowledge of the road speed limit during the estimation process, the vehicle speed, and the location of the vehicle when the speed estimate was given.

Separating the responses by gender, it can be seen that the histogram in Figure 3 for females is very similar to the histogram in Figure 4 for males. It is also observed that both sample means are close to zero and the sample standard deviations are nearly equal. So there were not important differences in estimation accuracy between males and females in the sample of subjects that were studied.

The amount of driving experience that the subjects had was categorized into 2 groups: 110 speed estimates from people who reported less than 2 years of driving experience (very little experience) and the other 377 estimates from people who reported 2 or more years of driving experience. As can be seen in Figure 5 and Figure 6, the sample of subjects with more driving experience have a mean error that is closer to zero (1.9%) compared with the mean of −12.5% for the subjects with very little driving experience. The sample standard deviation for the more experienced subjects is somewhat smaller, reflecting the tendency for their estimates to reside nearer to their mean. This is the first of several examples that show how the mean for the factor being examined can be different from the "overall" sample mean of −1.38%. Since the responses for people with more driving experience (and error closer to zero) were more numerous (377 compared with 110), this causes the "overall" sample mean to be pulled closer to zero than if there had been equal numbers in the two categories of driver experience.

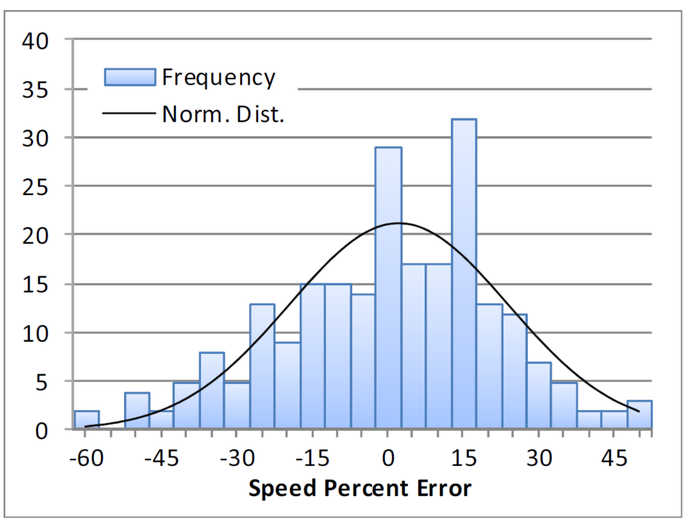

Figure 3. Females - estimated speed percent error. Mean= −0.23, StDev=21.70, N=231

Whether the subjects used their knowledge of the road speed limit to estimate the speeds of the moving vehicles was also studied and the histograms are shown in Figure 7 and 8. There were 343 responses from people who reported that they considered the road speed limit in their estimation process. Their mean error (1.77 %) was much closer to zero than the mean (−9.42%) for the 138 responses of those who claimed that they had not considered the road speed limit. The sample standard deviations for these two groups of responses were nearly equal.

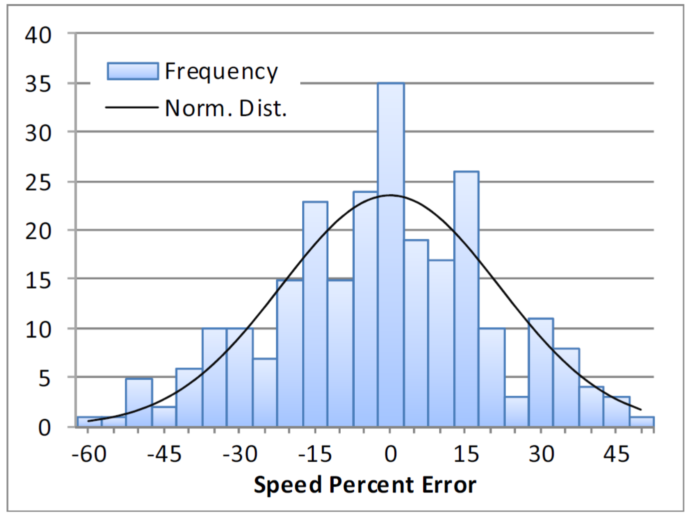

Figure 4. Males - estimated speed percent error. Mean= −2.41, StDev=21.65, N=256

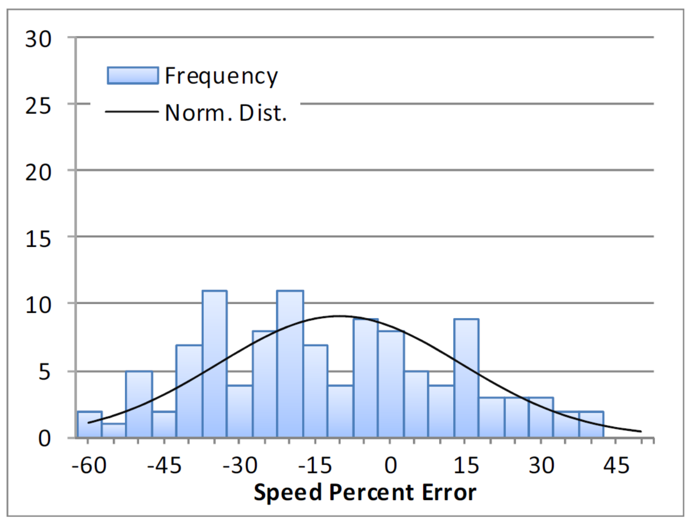

Figure 5. Driving experience < 2 years - estimated speed percent error. Mean= −12.47, StDev= 24.11, N=110

As previously seen with the factor "driving experience", it is now seen that the factor "speed limit" had two very unequally populated categories, and two very different means. Because more people considered the road speed limit when estimating vehicle speed, and their mean error was closer to zero, this had the effect of pulling the "overall" population mean error closer to zero.

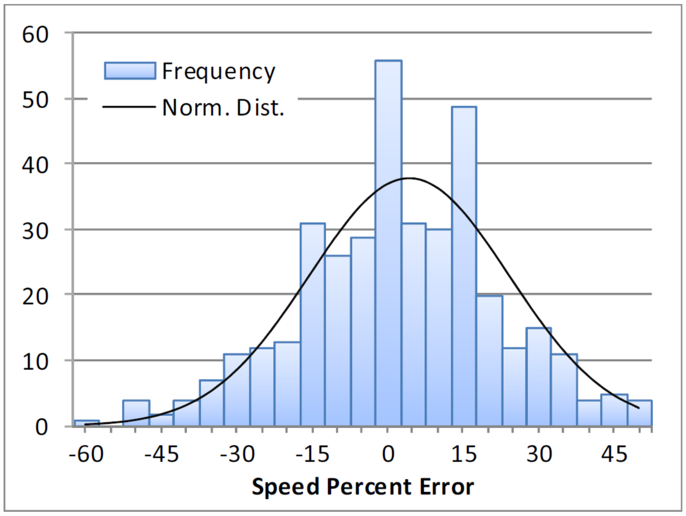

Figure 6. Driving experience ≥ 2 years - estimated speed percent error. Mean=1.86, StDev=19.82, N=377

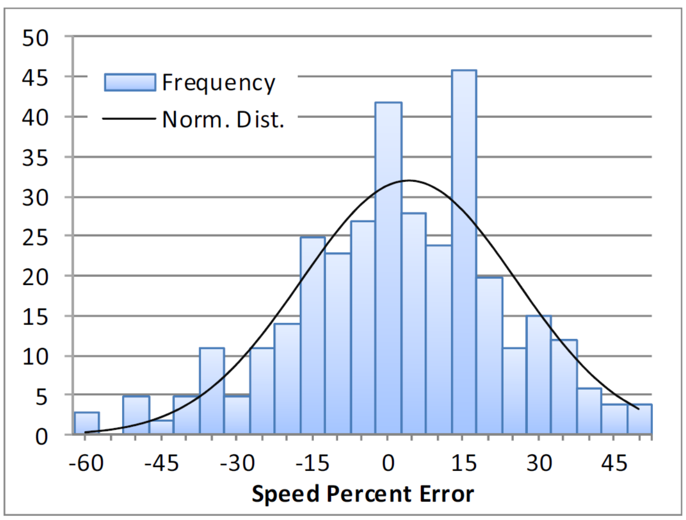

Figure 7. Subjects who DID consider the road speed limit during speed estimation - speed percent error. Mean=1.79, StDev=21.35, N=343

Subjects were asked to assess their own ability to estimate vehicle speeds using a Likert scale ranging from 1 (very poor) to 10 (excellent). A majority ranked themselves with a score greater than 4, resulting in 410 responses as presented in Figure 9. This group had a mean estimation error of 1.19%, whereas the 77 responses from the group who were less confident in their speed estimation ability (self-rating ≤ 4) had a mean estimation error of −15.04%, (see Figure 10). As is seen, there is quite a significant difference in the mean errors of speed estimation between the people who were more confident in their ability and those who were less confident. The sample standard deviations of the two groups were nearly equal. The effect on the "overall" sample mean is similar to that mentioned for the factor "driving experience".

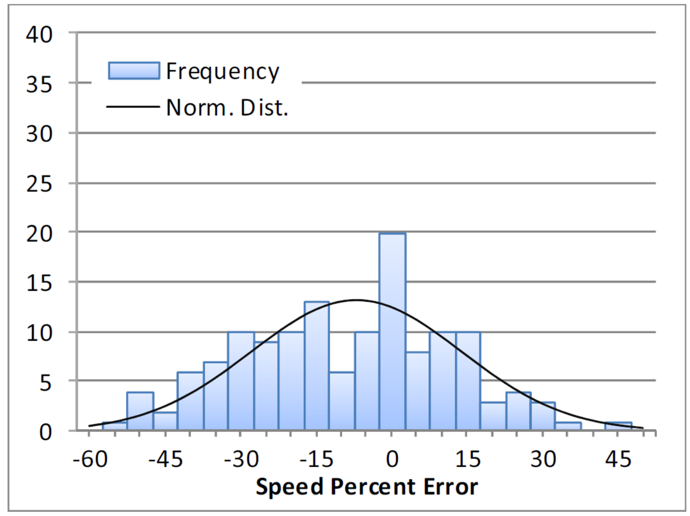

Figure 8. Subjects who DID NOT consider the road speed limit during speed estimation - speed percent error. Mean= −9.42, StDev=20.85, N=138

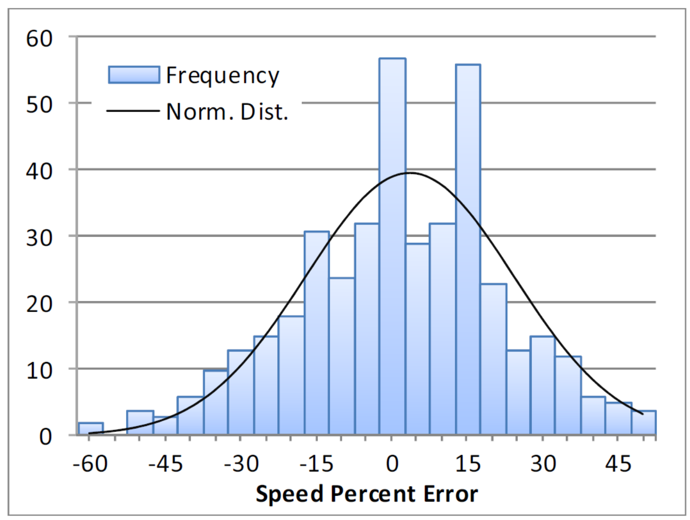

Figure 9. Subjects who had a self-perceived ability > 4 to estimate vehicle speeds - estimated speed percent error. (Based on Likert scale: 1= very poor; 10= excellent.) Mean=1.19, StDev=20.62, N=410

In addition to studying how the four subject specific factors affected their speed estimations, there was also an examination into whether the vehicle's speed and the vehicle's location at which the estimate was made affected the accuracy of the speed estimations. The vehicle speeds were categorized into 3 groups: low speeds less than 54.7 kph (34 mph), medium speeds between 54.7 kph and 71 kph (34 - 44 mph), or higher speeds greater than 71 kph (44 mph). The subjects tended to overestimate the low speeds and underestimate the high speeds, with the mean error for medium speeds being closer to zero as can be seen in Figures 11,12,13. Sample standard deviations tended to decrease as the speed went up, perhaps because some subjects took the speed limit into consideration when forming their speed estimates. The higher testing speeds were at about the speed limit.

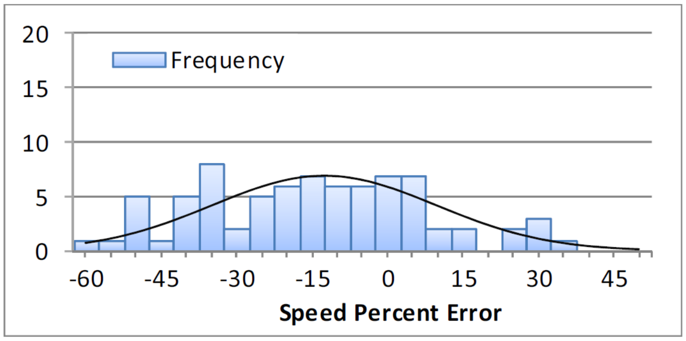

Figure 10. Subjects who had a self-perceived ability ≤ 4out of 10 - estimated speed percent error. Mean= −15.04, StDev=22.22, N=77

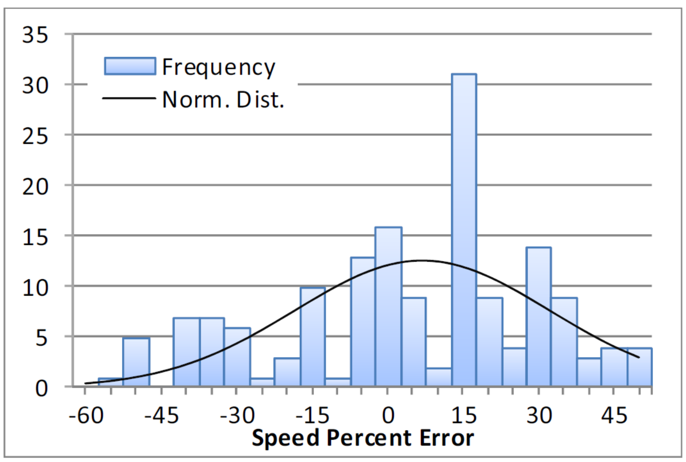

Figure 11. Estimated speed percent error when vehicles traveled at a LOW SPEED (speed < 54.7 kph). Mean=4.28, StDev=25.24, N=158

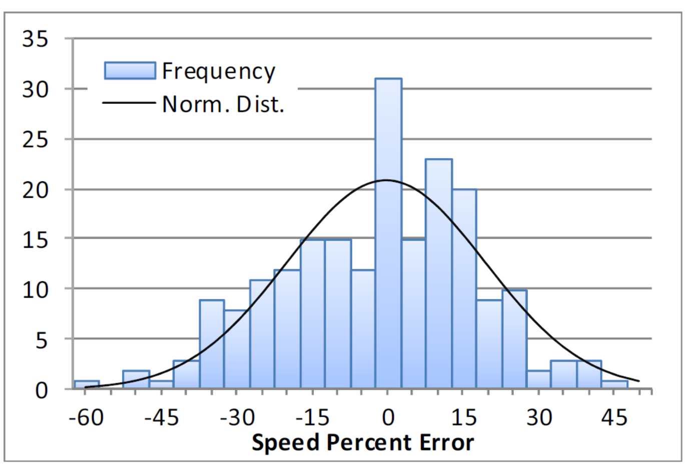

Figure 12. Estimated speed percent error when vehicles traveled at a MEDIUM SPEED (54.7 ≤ speed ≤ 71 kph). Mean= −2.57, StDev=19.79, N=207

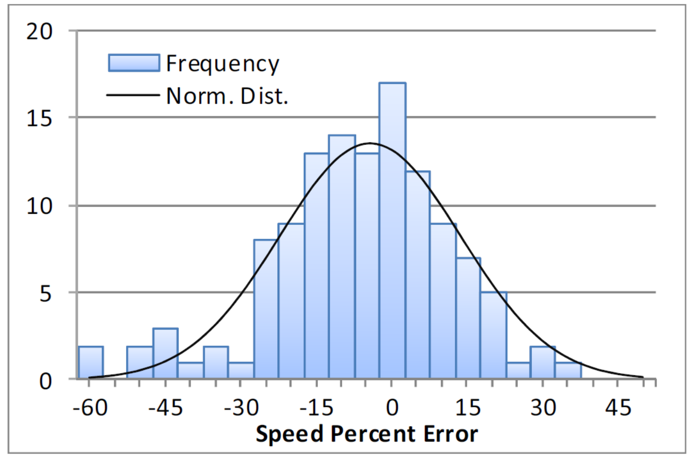

Figure 13. Estimated speed percent error when vehicles traveled at a HIGH SPEED (speed >71 kph). Mean= −6.68, StDev=17.95, N=122

As previously stated, each observer gave an estimate of the vehicle speed when a test vehicle was at one of six different locations. Three locations were on the same side of the street as the observer, and three were on the opposite side of the street. As shown in Figure 1, the test vehicle was passing directly in front of the test subject when the vehicle was in position 2 and 5; it was moving towards the subject in positions 1 and 4; and it was moving away from the subject in positions 3 and 6. These 6 vehicle locations were given names: "Near" or "Far" (for being in the near or far lane of traffic) and then followed by either "moving towards", "moving away" or "middle". As seen in Figures 14, 15, 16, 17, 18, 19, when the vehicle was in the near traffic lane there was a tendency to over-estimate the vehicle speed (between 2.13% and 3.98%), while in the far traffic lane the tendency was to under-estimate the vehicle speed (between −5.16% and −6.94%).

The information in Figures 14, 15, 16, 17, 18, 19, is presented in a different way in Figure 20 by making the histogram cells narrower; this makes the consistently large range of the estimation error more prominent. In addition, the reader can notice the shift in the data and the means for the Far lanes to the left, indicating more frequent speed underestimation than is observed for the Near lanes.

Table 1 provides a summary of the descriptive statistics that correspond to the 6 factors reported so far. The last column gives the result of a significance test for the difference in the means. The difference between males and females is not significant, while the differences in the means of the speed estimation errors for driving experience, self-perceived speed estimation ability, and the use of the speed limit to estimate vehicle speed were statistically significant, as indicated by their very small p-values. To further explore the significance of the errors associated between the means of the three vehicle speed categories and between the means of the six location categories requires multiple mean comparison procedures and are not included in this paper.

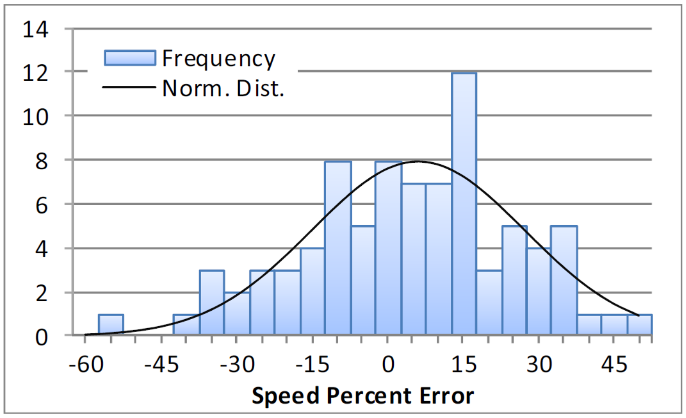

Figure 14. Vehicle NEAR and moving TOWARDS the subjects. Estimated speed percent error for location 1 as defined in Figure 1. Mean=3.98, StDev=20.53, N=79

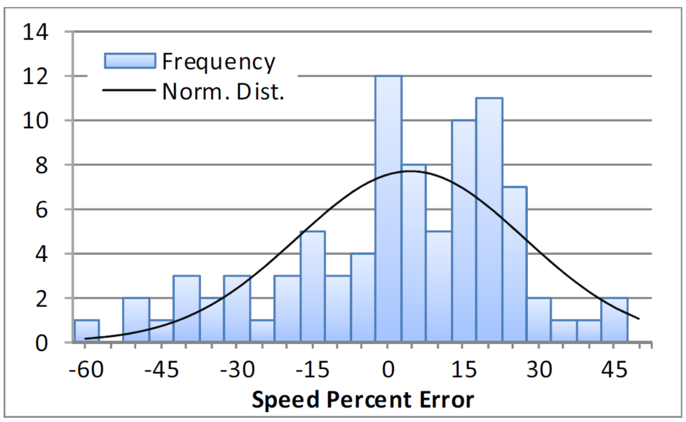

Figure 15. Vehicles at NEAR-MID location. Estimated speed percent error for vehicle location 2 as defined in Figure 1. Mean=2.13, StDev=22.58, N=87

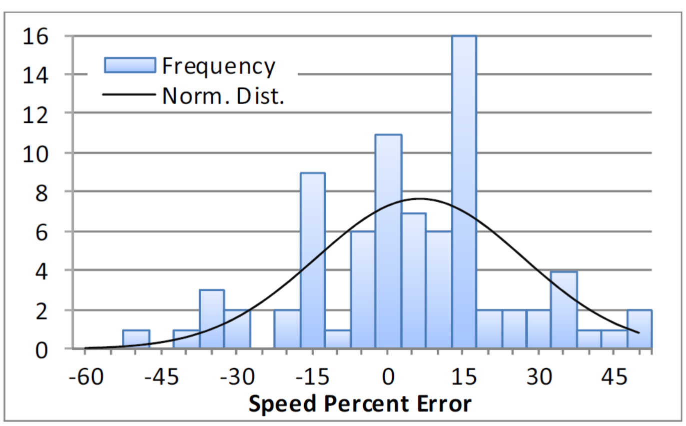

Figure 16. Vehicle NEAR and moving AWAY from the subjects. Estimated speed percent error for location 3, as defined in Figure 1. Mean=3.64, StDev=21.09, N=84

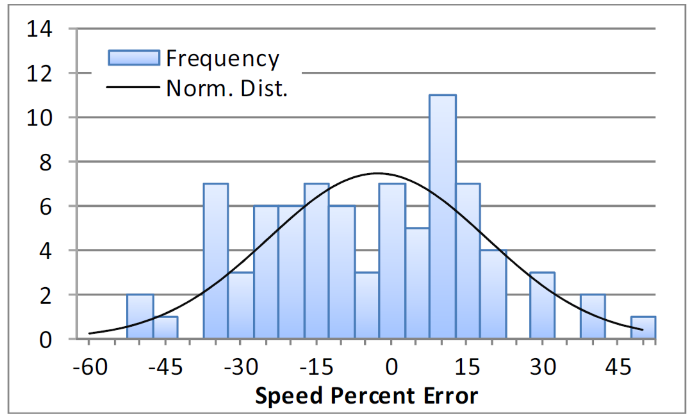

Figure 17. Vehicle FAR and moving TOWARDS the subjects. Estimated speed percent error for location 4 as defined in Figure 1. Mean= −5.16, StDev=21.67, N=81

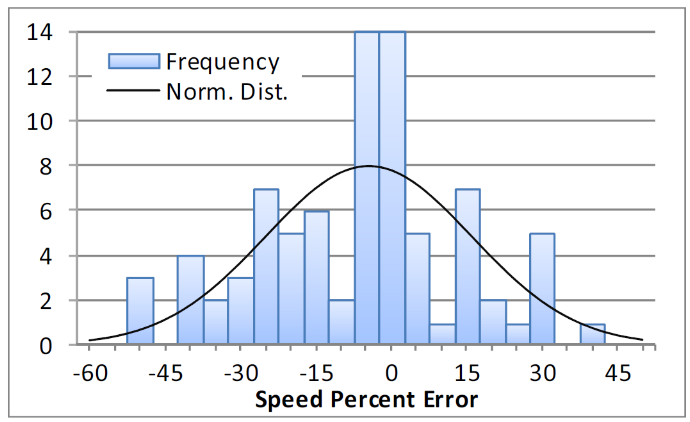

Figure 18. Vehicle at FAR-MID location. Estimated speed percent error for vehicle at location 5 as defined in Figure 1. Mean= −6.94, StDev=20.45, N=82

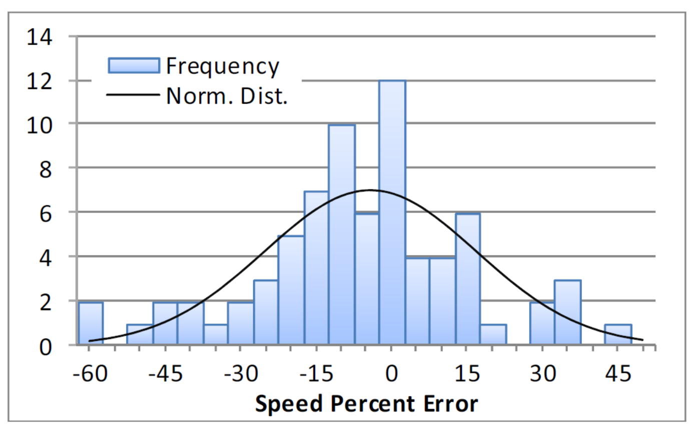

Figure 19. Vehicle FAR and moving AWAY from subjects. Estimated speed percent error for location 6 as defined in Figure 1. Mean= −6.61, StDev=21, N=74

Figure 20. Plot of estimated speed percent error categorized by vehicle location. (Location numbers refer to locations shown in Figure 1.)

A remarkable feature of the data is the large standard deviation in all categories of each of the six factors examined. An exception might be the high speed category, which could be influenced by the knowledge of the speed limit. Thus, even when the factor category means are quite different (as in low vs. high driving experience), the large standard deviations cause considerable overlap of the speed errors in both categories.

MULTIVARIATE REGRESSION MODEL

Up to this point this paper has reported on the means and standard deviations for the various categories for each of the six factors explored. However, these statistics for this sampled population of 87 people do not necessarily apply to the general population. This is due to what has sometimes been found to be a very unequal number of estimates for each category of a factor, and a large difference in the means. For example, the factor "driving experience" had two categories. One containing 377 estimates with a mean error of almost 2%, and the other category contained 110 estimates with a mean error of −12.5%. There are three times more data points in one category than the other which will create a bias when examining all the other factors. In order to separate out the effects of each of the factors in this study, independent of the other factors, a multivariate regression model was employed and the results are now discussed.

The Minitab (ver. 16) regression routine "General Linear Models" was used to fit a multivariate regression model using the six factors presented. In addition to the six main effects, there were three statistically significant interactions found between: gender and driving experience, self-perceived speed accuracy and driving experience, and the speed category and the use of the speed limit. Although the effect of gender (by itself) was statistically insignificant, it is retained in the model because of its interaction with driving experience.

Additional potential factors were studied beyond these six; however, their effect on the mean percent error was found to be statistically insignificant. In some cases, a factor was duplicious of a factor already in the model. An example is the subject's age, which provides the same kind of information as the reported years of driving experience. Similarly, whether the subject had a driver's license also turned out to be somewhat redundant with driving experience, and was therefore not further studied.

Table 1. Descriptive Statistics for Speed Percent Error for the "Overall" Data Set Categorized by Six Factors. (NS= Not statistically significant at p < 0.05)

Factor	Category	Mean (%)	Median (%)	Standard Deviation (%)	Subsample size	Significance Test
Overall	All data	-1.38	0.00	21.68	487	
Gender						
	Male	-2.41	0.00	21.65	256	NS
	Female	-0.23	0.00	21.70	231	(p = 0.268)
Driving Experience						
	< 2 years	-12.47	-15.65	24.11	110	significant
	≥ 2 years	1.86	2.13	19.82	377	(p < 0.0005)
Self-perceived Speed Estimation Ability						
	≤ 4	-15.04	-16.67	22.22	77	significant
	> 4	1.19	1.49	20.62	410	(p < 0.0005)
Used Speed Limit?						
	No	-9.42	-6.89	20.85	138	significant
	Yes	1.79	2.44	21.35	343	(p < 0.0005)
Speed Category						
	low	4.45	9.38	25.26	159	
	medium	-2.74	0.00	19.68	206	
	high	-6.68	-5.85	17.95	122	
Location of Vehicle						
	Near moving towards (1)	3.98	5.26	20.53	79	
	Near middle (2)	2.13	5.26	22.58	87	
	Near moving away (3)	3.64	3.95	21.09	84	
	Far moving towards (4)	-5.16	-3.23	21.67	81	
	Far middle (5)	-6.94	-4.26	20.45	82	
	Far moving away (6)	-6.61	-6.98	21.00	74	

A question that arose was whether there could be a form of learning during the testing process so that perhaps a subject was somehow more accurate when estimating the speed of the last car presented compared with earlier ones. To examine this effect, the vehicle location and speed were arranged in a factorial in the original experimental design so that their potential influence would be balanced. The order in which speed estimates were made was recorded and used as a factor in the regression. The results showed that the order of the estimates given was statistically insignificant in its effect on the percent error. In summary, the results in Table 2 list the six main factors and three interactions retained in the regression model, along with their p-values. It is worth noting that the statistical significance of the three interaction terms is not as conclusive as that of the main effects (excepting gender, which was not statistically significant).

Although many of the factors had extremely low p-values which indicate that they were statistically significant, the R^2 for the model is only 27.6%, with the corresponding "adjusted" R^2 of 25.3% that is calculated to correct R^2 in a way that account for the number of parameters estimated. The relatively low R^2 values reflect the very large standard deviations in the raw data that persisted even within the subcategories for the factors. The standard residual tests on the regression model were carried out and showed that the underlying assumptions for ANOVA-based models were satisfied: the residuals were very nearly normally distributed and had a pattern of constant variance.

In Table 3, the results of the regression model show estimates for the mean error for each of the subcategories for the six factors studied, free of the effects of the other five factors. It is seen that the confidence intervals for these subcategory means are moderately large, which is a mathematical result due to the low R^2 values from the regression. Later in this paper, the percent error in the speed estimate for an individual person's response will be shown to be accompanied by even larger confidence intervals.

Table 2. Results of Multivariate Regression Analysis.

Factor	Degrees of Freedom	p-value
Gender	1	0.827
Driving experience	1	<0.0005
Speed estimation assessment	1	<0.0005
Used speed limit?	1	<0.0005
Speed category	2	0.001
Location	5	<0.0005
Gender & Experience interaction	1	0.039
Speed estimation assessment and Experience interaction	1	0.023
Speed category & Used speed limit?	2	0.048

It needs to be emphasized that the regression model estimates for the categories of each factor in Table 3 are free of the effect of all the other factors. For example, for the two categories of speed estimation ability, this assumes that the subjects are equally likely to be female or male, were in either category of driving experience, used the speed limit (or not), be viewing a car in any of the three speed categories, at any of the six locations.

If all the factor categories in the sampled population had been balanced in their size, we can see from the regression model that the average error would be close to −10.5% (the average of the mean estimates for any factor). Comparing Table 1 with Table 3, we can see many examples of the multivariate regression model "correcting" the raw data to account for the imbalances in the subsample sizes within categories. This kind of situation is common in observational data and a regression model is often used to provide a clearer assessment of specific factor effects that are often clouded by other factors when only looking at the raw data.

To give an example of interpreting the information in Table 3, consider the category of subjects who assess their speed estimation ability as ≤ 4. The estimate of the mean error for this group is −18.1% and the 95% confidence interval bounds for the mean are from −22.5% to −13.7%. For the subjects who assess their speed estimation ability as > 4, the estimate of the mean error is −2.95% and the 95% confidence interval for the mean is −5.6% to −0.33%. It is clear that subjects who were not confident about their speed estimation ability were more likely to underestimate speeds (and have a greater error) than those who were more confident. This conclusion is based on both the difference in the modeled means and the confidence intervals for those means which are not overlapping.

In Table 3 the reader can see that there are three subject factors (driving experience, self-perceived speed estimation ability, used speed limit) with no overlap of the confidence intervals for the means of their categories. This provides evidence of distinct differences in certain sub-populations of subjects.

DISCUSSION

Up to this point, the discussion has focused on estimates of the mean error for *categories* of the characteristics of subjects and situations. Next, consideration is given to an *individual* subject, and the mean error and confidence interval for the percent error for that person's speed estimate. The regression model was used to predict the estimation error for a hypothetical person who has a chosen set of characteristics (gender, driving experience, estimation ability, use of the speed limit) and situations (vehicle speed, location). Table 4 shows a number of results when the regression model predicts the mean error and confidence interval for a hypothesized individual subject based on the values chosen for the six factors. It should be kept in mind that these are only 7 cases chosen for demonstration from the 288 different combinations of subject features and situations that are possible.

Table 3. Regression Model Estimates of Mean Speed Error by Factor Category

Factor	Category	Mean Estimate (%)	Standard Error of Mean Estimate (%)	95% confidence interval [lower, upper] (%)
Gender				
	Female	-10.75	1.72	[-14.19, -7.32]
	Male	-10.29	1.66	[-13.60, -6.98]
Driving Experience				
	< 2 years	-16.39	1.91	[-20.21, -12.57]
	≥ 2 years	-4.65	1.72	[-8.10, -1.20]
Self-perceived Speed Estimation Ability				
	≤ 4	-18.09	2.22	[-22.52, -13.66]
	> 4	-2.95	1.31	[-5.57, -0.33]
Used Speed Limit?				
	No	-14.97	1.88	[-18.73, -11.21]
	Yes	-6.06	1.40	[-8.86, -3.26]
Speed Category				
	low	-5.49	1.90	[-9.29, -1.69]
	medium	-11.84	1.70	[-15.23, -8.45]
	high	-14.23	2.14	[-18.50, -9.96]
Location of Vehicle				
	Near moving towards (1)	-6.20	2.37	[-10.93, -1.47]
	Near middle (2)	-7.72	2.30	[-12.32, -3.12]
	Near moving away (3)	-4.44	2.29	[-9.03, 0.15]
	Far moving towards (4)	-14.75	2.35	[-19.44, -10.06]
	Far middle (5)	-16.22	2.34	[-20.91, -11.53]
	Far moving away (6)	-13.78	2.41	[-18.61, -8.95]

The first two cases in Table 4 refer to female and male subjects with greater driving experience (≥ 2 years), higher self-assessment of speed estimation ability (> 4), who used the speed limit when they estimated the speed of vehicles with the speed categories and locations given. These two cases were chosen to show that the mean error for some individuals and situations can turn out to be close to zero (−0.1% and −0.2%). The width of the confidence interval is quite large (−38% to 37%), but it is centered near zero, so underestimation and overestimation for these individuals and situations is about equally likely. These results may not surprise, since the subject features of experience, self-assessment of estimation ability, and use of the speed limit could be expected to result in greater accuracy.

Next the regression model was used to try to identify the features of individuals with mean errors that were not close to zero. The female subject (3rd case in Table 4) chosen has lower driving experience, lower self-assessment of her speed estimation ability, and did not report the use of the speed limit when she estimated the speed of high speed car in the Far middle location. She would be expected to underestimate that vehicle's speed (since her mean error is −37.7%) and would be very unlikely to give a correct estimate of that speed, since the upper bound on the 95% confidence interval is less than zero. The corresponding male subject (4th case) also tended to underestimate speed, (his mean error is −34.9%), with some small chance of estimating the speed correctly since the upper bound on his confidence interval is 3.6%.

Table 4. Examples of Using the Regression Model to Estimate the Mean Speed Percent Error and Confidence Interval for 7 Individual Responses

	Gender	Driving Experience	Self-Perceived Speed Estimation Assessment	Used Speed Limit?	Speed Category	Location	Estimate of mean error (%)	95% Confidence Interval for mean error [lower, upper] (%)
1	Female	≥ 2 years	> 4	Yes	medium	Far middle	-0.1%	[-37.5%, 37.3%]
2	Male	≥ 2 years	> 4	Yes	high	Near moving towards	-0.2%	[-37.8%, 37.3%]
3	Female	< 2 years	≤ 4	No	high	Far middle	-37.7%	[-78.0%, -1.32%]
4	Male	< 2 years	≤ 4	No	high	Far middle	-34.9%	[-73.3%. 3.6%]
5	Female	≥ 2 years	> 4	Yes	low	Near moving away	20.5%	[-17.0%, 57.9%]
6	Male	< 2 years	≤ 4	Yes	high	Near moving towards	-22.8%	[-61.0%, 15.4%]
7	Male	≥ 2 years	≤ 4	No	low	Near moving away	-7.6%	[-45.7%, 30.6%]

The 5th case in Table 4 is a female subject with a mean overestimation error of 20.5% which was the largest overestimation encountered while exercising the regression model. The subject features are the same as the female who was the 1st case in the table, but in the 5th case overestimation turns out to be likely for a low speed vehicle in the near lane, moving away from the subject. This is one of many examples where the vehicle speed and location had a significant effect on accuracy of speed estimation.

Next consider the 6th case in Table 4 where a male has both limited driving experience and a low self-assessment of his speed estimation ability. He used the speed limit while estimating a high speed vehicle in the Near moving-toward location. The model estimates that this type of subject will underestimate the vehicle speed by an average of 22.8%. Thus if he saw a vehicle going 75 kph, the average (mean or expected) value of his speed estimate would be 58 kph. Using the confidence interval bounds, 95% of the time he would provide a speed estimate contained between 29 kph and 87 kph.

The 7th case in Table 4 is for a male with a high level of driving experience but low self-assessment of speed estimation ability, who did not report the use of the speed limit in estimating the speed of a lower speed car in the Near moving-away location. The regression model estimates the average error as −7.6%. So for a vehicle going 45 kph, this man's mean speed estimate would be about 42 kph with a 95% confidence interval that extends from 24 kph to 59 kph.

It is informative to compare the extensive overlap at the lower portions of the confidence intervals for estimated speeds corresponding to the 6th and 7th cases in Table 4. Both hypothetical subjects were males with low self-assessment of speed estimation ability. Their driving experience differed, along with the use of the speed limit and the speed of the car whose speed is being estimated (75 kph compared with 45 kph). Nonetheless, it is probable that many of their speed estimates would turn out to be similar in magnitude, despite the large difference in speed of the vehicles being estimated.

A general observation is that confidence intervals for individual responses are extremely wide, regardless of the mean value for percent error about which they are centered. This reflects great uncertainty in the error associated with the individual estimates. This level of uncertainty is primarily due to the large observed standard deviations in the raw data, presented earlier. Mathematically, it is due to the low R^2 values, which indicate a high level of the unexplained variation remaining even after the multivariate analysis has identified significant shifts in mean errors that can be attributed to the various factors and their interactions in the regression model.

SUMMARY AND CONCLUSIONS

In order to investigate the accuracy with which a stationary pedestrian can estimate the speed of a vehicle, an experiment was designed and conducted. The test subjects were of varied age, life experience, gender, education, and were part of the general population. Thus the subjects and their attributes varied significantly, just like the general population. To address the challenge of determining whether certain life attributes affect a person's ability to estimate the speed of a vehicle, a number of questions were asked in order to characterize each subject in various ways. To date, some of the demographic data has been analyzed along with the controlled vehicle and roadway configuration. Although there is more analysis to be performed on the background data, the results to date indicate that the percent error in vehicle speed estimation is related to:

- Driving experience
- Self-perceived speed estimation ability
- Whether subjects used their knowledge of the speed limit to estimate speed
- Vehicle speed
- Vehicle location
- Gender (as an interaction, not a main effect)

In addition, there was a significant statistical interaction between the subject's self-perceived speed estimation ability and their driving experience. This would make sense as it would be expected that the more experience a person has driving, the more opportunities for feedback they have to learn to judge the speed of a vehicle. A second interaction was found between the speed of the vehicle and whether the subject reported using the speed limit. It is not surprising that this would affect speed estimation error since an observer would be more inclined to use the speed limit as an upper bound for their estimation when cars are moving faster. And lastly, gender was not found to be a significant factor by itself, but it was influential in an interaction with driving experience. For instance, females with little driving experience were somewhat more likely to underestimate speeds than males with little driving experience.

The regression model was also used to predict the estimation error for hypothetical persons who had a chosen set of characteristics (gender, driving experience, estimation ability, use of the speed limit) and situations (vehicle speed, location). The differences in the mean speed estimation error for individuals from different categories of subjects were calculated and some interesting results were provided. It is seen that the confidence interval for all estimates is very wide. This was a characteristic of the "overall" raw data also.

Based on this analysis, categories of subjects could be selected that result in their mean percentage error being quite different for different situations. However, the range of the error for those individual speed estimates is likely to overlap significantly. So, given two different categories of individual subjects, we are likely to observe a great deal of commonality in the error of their speed estimates even though the means are different. Also, when presented with vehicles traveling at greatly different speeds, it is very possible for different subjects to provide speed estimates that are similar.

Lastly, consider the hypothetical situation where a subject is chosen at random from some large population about which very little demographic information is known. Then, while standing at the side of a roadway the subject is asked to provide a speed estimate for a moving vehicle, whose speed is not known, nor is anything precisely known about the position of the vehicle with respect to the subject. Based on this work, and under the test conditions specified, one would estimate the error associated with that estimate to have a mean of about −10.5% and probable bounds of −48% and +27%.

The experimental results and subsequent statistical analysis presented here quantify the unreliability of vehicle speed estimates given by pedestrians. This may be helpful for roadway designers, traffic engineers, and also accident reconstruction professionals who consider witness testimony in their analysis.

The authors would like to briefly compare these results to our previous work on distance estimation accuracy in a roadside setting. In that earlier work, it was found that subjects generally tended to provide underestimates of roadside distances, consistent with underestimating speeds. Although males were more accurate than females in estimating distances, we found no gender difference in speed estimation accuracy, except in an interaction with driving experience. Subjects who rated themselves highly with respect to distance estimation ability had smaller errors than the subjects who were less confident, which was a similar finding to the one in this speed estimation research. And similar to the work on distance estimation, it also found that there was a high standard deviation associated with the estimation error of individual subjects.

REFERENCES

1. Aberg, L. et al., "Observed Vehicle Speed and Drivers' Perceived Speed of Others", Applied Psychology: An International Review, 1997, Vol. 46, No. 3, pp. 287-302.

2. Conchillo, A. et al., "Comparing Speed Estimations from a Moving Vehicle in Different Traffic Scenarios: Absence versus Presence of Traffic Flow", The Spanish Journal of Psychology, 2006, Vol. 9, No. 1, pp. 32-37.

3. Evans, L. “Speed Estimation from a Moving Automobile”, Ergonomics, 1970, Vol. 13, No. 2, pp. 219-220.

4. Green, G. “The Pedestrian's Judgment of the Minimum Time Interval to Allow when Crossing the Road,” Transport and Road Research Laboratory unpublished note RN 2238, 1954.

5. Hinch, J., “A Study into the Estimation of Vehicle Speeds and Distances by Untrained Pedestrians.” MSc Thesis, University of Birmingham, 1967.

6. Hurford, R., “Speed Estimation”, MSc Thesis, University of Birmingham, 1968.

7. Hutchinson, T., “Witnesses' Estimates of the Speeds of Traffic Accidents”, Accident Analysis and Prevention, 1975, Vol. 7, pp. 27-35.

8. Kebbel, M. et al., “The Influence of Belief that a Car Crashed on Witnesses' Estimates of Civilian and Police Car Speed”, The Journal of Psychology, 2002, Vol. 136, No. 6, pp. 597-607.

9. Kebbel, M. and Giles, D., “Some Experimental Influences of Lawyers' Complicated Questions on Eyewitness Confidence and Accuracy”, The Journal of Psychology, 2000, Vol. 134, No. 2, pp. 129-139.

10. McAllister, H., et al., “Speed Estimates by Eyewitnesses and Earwitnesses; How Vulnerable to Postevent Information?” Journal of General Psychology, 1988, Vol. 115, No. 1, pp. 25-35.

11. Montgomery, D. and Runger, G., Applied Statistics and Probability for Engineers, 2nd Ed., John Wiley & Sons, New York, 1999.

12. Milosevic, S. and Milic, J, “Speed Perception in Road Curves”, Journal of Safety Research, 1990, Vol. 21, pp. 19-23.

13. Recarte M., and Nunes, L., “Perception of Speed in an Automobile: Estimation and Production”, Journal of Experimental Psychology: Applied, 1996, Vol. 2, No. 4, pp. 291-304.

14. Scialfa, C., et al., “Age Differences in Estimating Vehicle Velocity,” Psychology and Aging, 1991, Vol. 6, No. 1, pp. 60-66.

15. Strauss, M. and Carnahan, J., “How well can the average person and police estimate distance along a roadway?”, AREC Proceedings, Wildhaus, Switzerland, 2010.

16. Strauss, M. and Carnahan, J., “Observed Errors in Distance Estimation,” SAE Technical Paper 2010-01-0046, 2010, doi:10.4271/2010-01-0046.

CONTACT INFORMATION

For additional information, contact

Mark Strauss
mgstrauss@ruhl.com

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge Ruhl Forensic, Inc. for the financial and logistical support in order to conduct this research. Special appreciation goes to Jacqueline Orr and Mary Ruhl for the recruitment, coordination and oversight of the subjects and test drivers; and Mary K. Dooley-Owen for her editorial contributions.

The Engineering Meetings Board has approved this paper for publication. It has successfully completed SAE's peer review process under the supervision of the session organizer. This process requires a minimum of three (3) reviews by industry experts.

ISSN 0148-7191

Positions and opinions advanced in this paper are those of the author(s) and not necessarily those of SAE. The author is solely responsible for the content of the paper.

SAE Customer Service:
Tel: 877-606-7323 (inside USA and Canada)
Tel: 724-776-4970 (outside USA)
Fax: 724-776-0790
Email: CustomerService@sae.org
SAE Web Address: http://www.sae.org
Printed in USA

SAE International®