Technical Report UMTRI-2006-6 May, 2007

Task Acceptability and Workload of Driving City Streets, Rural Roads, and Expressways: Ratings from Video Clips

SAfety VEhicles using adaptive Interface Technology (SAVE-IT Project)
Task 2C: Develop and Validate Equations

Jason Schweitzer and Paul A. Green

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UMTRI-2006-6		

4. Title and Subtitle	5. Report Date / 6. Performing Organization Code
Task Acceptability and Workload of Driving City Streets, Rural Roads, and Expressways: Ratings from Video Clips	5. Report Date: May, 2007 6. Performing Organization Code: accounts 049178, 049183
7. Author(s) Jason Schweitzer and Paul A. Green	8. Performing Organization Report No.
9. Performing Organization Name and Address The University of Michigan Transportation Research Institute (UMTRI) 2901 Baxter Rd, Ann Arbor, Michigan 48109-2150 USA	10. Work Unit no. (TRAIS) 11. Contract or Grant No.: Contract DRDA 04-4274
12. Sponsoring Agency Name and Address Delphi Delco Electronic Systems One Corporate Center, M/C E110 Box 9005, Kokomo, IN 46904-9005	13. Type of Report and Period Covered: 1/05-2/07 14. Sponsoring Agency Code

15. Supplementary Notes

SAVE-IT project, funded by U.S. Department of Transportation

16. Abstract

Subjects rated the workload of clips of forward road scenes (from the advanced collision avoidance system (ACAS) field operational test) in relation to 2 anchor clips of Level of Service (LOS) A and E (light and heavy traffic), and indicated if they would perform any of 3 tasks (dial a phone, manually tune a radio, enter a destination) in driving the scenes shown. After rating all of the clips, subjects rated a wider range of described situations (not shown in clips) and the relative contribution of road geometry, traffic, and other factors to workload.

Using logistic regression, predictive equations for the refusal to engage in the 3 tasks were developed as a function of workload, driver age, and sex.

Several equations were developed relating real-time driving statistics with workload, where workload was rated on a scale of 1 (minimum) to 10 (maximum). Some 87% of the rating variance was accounted for by the following expression: Mean Workload Rating=8.87-3.01(LogMeanRange)+ 0.48(MeanTrafficCount)+ 2.05(MeanLongitudinalAccleration), where range (to the lead vehicle) and traffic count were both determined by the adaptive cruise control radar. Other estimates were also generated from post-test ratings and adjustments, considering factors such as construction zones, lane drops, curves, and hills. From the results of this report alone, the workload estimates needed by a real-time workload manager could be developed using (1) the real time data, (2) look-up tables based on the clip ratings, (3) look-up tables based on the post-test data, or (4) some combination of those 3 sources.

17. Key Words	18. Distribution Statement
Distraction, Attention, Driving Performance, Crashes, ITS, Human Factors, Ergonomics, Safety, Usability, Telematics, Workload	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classify. (of this report)	20. Security Classify. (of this page)	21. No. of pages	22. Price
(None)	(None)	186	

Form DOT F 1700 7 (8-72) Reproduction of completed page authorized

	TASK ACCEPTABILITY AND WORKLOAD OF DRIVING CITY STREETS, RURAL ROADS, AND EXPRESSWAYS: RATINGS FROM VIDEO CLIPS **UMTRI Technical Report 2006-6, May, 2007**

Jason Schweitzer and Paul Green	**University of Michigan** **Transportation Research Institute** **Ann Arbor, Michigan, USA**

1 Primary Issues

1. How repeatable are the workload ratings within and between drivers?
2. How do workload ratings vary overall?
3. What is the relationship between workload ratings of driving situations and (1) road type (e.g., urban), (2) road geometry, (3) lane driven, (4) traffic volume (as measured by LOS), (5) driver age, and (6) driver sex?
4. How can workload ratings be estimated using the driving performance statistics developed from the ACAS FOT data set?
5. How do ratings of workload vary with the relative position of vehicles ahead on expressways?
6. What is the relative contribution of traffic, road geometry, visibility, and traction to ratings of workload?
7. How does the probability of a driver being willing to do a secondary task while driving (tune a radio, dial a phone, enter a destination) vary with (1) the overall ratings of workload and (2) road characteristics, traffic, & driver characteristics in question 3?

2 Methods

1. Practice & become familiar with entry tasks while driving UMTRI simulator (dial phone, tune radio, enter street address)
2. Rate workload of clips 1-to-10 scale with anchor clips for 2 and 6 (LOS A, E) 2 or 3 clips shown together (usually LOS A, C, E) Also say if would dial a phone, tune a radio, or enter street address for that clip <u>Roads Presented</u> 2-Lane rural (straight, curved) v. LOS (A, C, E) 4-Lane urban (straight, intersection) v. LOS (A, C, E) 6-Lane expressway (left, center, right) v. LOS (A, C, E) + merge (right only) v. LOS (C, E)

Anchors (LOS A, E)

3. Post-Test Ratings

(a) Workload of traffic on expressway (versus distance ahead)
(b) Workload for residential, urban, rural roads, and expressways
(c) Contribution of traffic, visibility, road geometry, and traction to total workload

3 Results and Conclusions

P(Not willing to do task) =1/(1+e^-(ax+b)), a=slope, b = intercept, x=clip workload rating (Q7: Estimating What Drivers Will Do Given Workload)							
Age	**Sex**	**Radio**		**Phone**		**Navigation**	
		Intercept	**Slope**	**Intercept**	**Slope**	**Intercept**	**Slope**
Young	Male	-83.19	11.33	-8.08	2.87	-5.18	5.40
Young	Female	-12.04	1.85	-10.59	2.37	-13.34	1.54
Middle	Male	-18.11	2.85	-6.57	2.12	-3.66	1.79
Middle	Female	-3.28	1.63	-12.47	4.41	-8.68	2.14
Older	Male	-6.45	1.53	-4.28	1.84	-3.78	3.29
Older	Female	-5.35	1.24	-10.23	4.48	-0.08	2.12
Mean		-21.40	3.40	-8.70	3.02	-5.79	2.71

Q3: Mean Workload Ratings (from Clips) (For Workload Manager Estimate)									
Age Group	**LOS**	**Rural**		**Urban**		**Expressway**		**LOS Mean**	**Age Mean**
Young	A	2.6	4.2	2.6	4.7	2.4	3.9	2.6	4.3
	C	4.1		5.2		3.7		4.3	
	E	5.9		6.2		5.7		5.9	
Middle	A	2.8	4.2	2.7	4.9	3.0	4.6	2.8	4.6
	C	4.0		5.4		4.5		4.7	
	E	5.8		6.6		6.4		6.3	
Older	A	2.6	3.7	3.1	4.7	3.2	4.4	3.0	4.2
	C	3.6		5.2		4.3		4.3	
	E	5.0		5.7		5.7		5.4	
Mean			4.0		4.8		4.3		4.4

Rural		**Urban**		**Expressway**	
Straight	4.0	Intersect	4.8	Left Lane (A, C, E)	4.8
Curved	4.1	Not	4.8	Middle (A, C, E)	4.3
				Right (A, C, E)	4.0
				Right Merge (C, E)	5.7

Q 2,3:Mean Workload Ratings from Clips (Workload Manager Table Look Up)								
Rural Roads		**Geometry, Traffic, and Subject Data**						**Road & Traffic Only**
LOS	**Geometry**	**Young**		**Middle**		**Old**		
		Female	**Male**	**Female**	**Male**	**Female**	**Male**	
A	Straight	1.9	2.3	3.0	2.0	2.4	2.7	2.7
	Curved	2.9	3.4	3.9	2.2	2.5	2.9	
C	Straight	4.4	3.9	4.0	4.0	3.2	4.0	3.9
	Curved	4.0	4.1	4.3	3.8	3.0	4.1	
E	Straight	6.7	5.4	6.0	5.8	4.5	5.6	5.5
	Curved	6.2	5.1	6.1	5.4	4.2	5.6	

Urban Roads		**Geometry, Traffic, and Subject Data**						**Road & Traffic Only**
LOS	**Intersection**	**Young**		**Middle**		**Old**		
		Female	**Male**	**Female**	**Male**	**Female**	**Male**	
A	No	3.0	2.8	3.3	2.8	3.1	2.5	2.9
	Yes	2.8	2.4	3.5	3.0	2.8	2.1	2.8
C	No	5.0	5.2	4.7	5.1	5.1	4.2	4.9
	Yes	6.1	5.4	5.0	6.1	6.8	4.7	5.7
E	No	6.7	7.0	5.7	6.5	6.7	6.4	6.5
	Yes	6.3	6.3	4.7	5.8	6.3	5.3	5.8

Expressways		**Geometry, Traffic, and Subject Data**						**Road & Traffic Only**
LOS	**Lane**	**Young**		**Middle**		**Old**		
		Female	**Male**	**Female**	**Male**	**Female**	**Male**	
A	Left	2.8	2.3	3.2	3.5	3.8	3.4	3.2
	Middle	2.8	2.4	2.7	3.1	3.3	3.3	2.9
	Right	2.3	2.0	2.9	2.8	3.1	2.4	2.6
C	Left	4.5	3.6	4.6	4.8	4.5	4.8	4.5
	Middle	4.1	3.7	4.8	5.0	4.5	4.8	4.5
	Right	3.1	3.3	4.0	3.8	3.9	3.3	3.6
	Right Merge	5.2	4.5	5.8	4.9	4.4	5.1	5.0
E	Left	6.6	6.1	7.3	6.9	5.7	7.1	6.6
	Middle	5.6	4.8	6.2	5.5	4.9	5.5	5.4
	Right	5.8	5.5	6.8	6.0	5.0	5.9	5.8
	Right Merge	6.6	5.8	6.9	6.7	5.9	6.6	6.4

Q3: Workload Estimates from Post-Test Ratings (Another Way to Estimate Workload, Potential Extension of Clip Rating Estimates)			
Road Situation	**Post-Test Workload Rating**	**Road Modifier**	**Rating Change**
Urban, downtown	72	Crash scene	23
Expressway, crash scene	71	Construction	22
Expressway, construction	70	Very curved or hilly	19
Rural, very curved or hilly	69	Downtown	16
Expressway, lane drop	64	Lane drop	15
Expressway, with 3-foot shoulder	62	Signaled intersection	9-15
Urban, corner commercial bldg	60	3-foot shoulder	14
Rural, stop sign for cross traffic	60	>25% parked cars	14
Residential, signaled intersection	59	Curved or hilly	5-10
Rural, 1-foot shoulder	58	Stop sign for cross traffic	10
Expressway, curved or hilly	58	Interchange	10
Expressway, interchange	58	1-Foot shoulder	9
Rural, signaled intersection	58	0-25% Parked cars	7
Residential, >25% parked cars	58	Corner commercial building	4
Residential, curved or hilly	55		
Rural, curved or hilly	54		
Residential, 0-25% parked cars	51		

Q3: Estimation of Workload from Post-Test Ratings Potential Correction Factors for Clip Rating Estimates								
Road Type & Mean	**Road Modifier**		**Lane Modifier**		**Traffic**		**Driveways**	
Rural Mean=58	-8	Base case	-1	2 Lanes	-5	None/Little		
	-3	Gentle curve/hill	1	3 Lanes (in left)	+5	Some		
	-3	1-ft shoulder	+2	4 Lanes (in left)				
	+1	At, approach light						
	+2	Stop sign for others						
	+11	Very hilly, curved						
Urban Mean=63	-7	Base case	-3	2 Lanes	-6	None/Little		
	-3	Corner business	-2	3 Lanes	-3	Some		
	+9	Downtown	+0	4 Lanes	+9	Heavy		
			+4	>=5 Lanes				
Xway Mean=61	-13	Base case	-1	Left	-12	None/Little		
	-3	Curved/hilly	0	Middle	0	Some		
	-3	Exit	+2	Right	+12	Heavy		
	0	Lane Drop						
	+1	Guardrail						
	+10	Construction						
	+10	Crash						
Residential Mean=54	-10	Base					-6	Few
	-2	Some parking					-1	Some
	+1	Curved/hilly					+5	Many
	+4	Many parked cars						
	+5	Intersection						

Sex	**Age**		
	Young	**Middle**	**Old**
Male	-14	+8	+3
Female	-9	+10	+4

Relationship between Clip and Post-Test Workload Ratings (For Comparison of Different Estimation Methods)			
Road Type	**Equation (Clip Rating =)**	**R2**	**# Data Points**
All Roads	-0.58 + 0.94*(post-test rating)	0.56	36
Expressway	0.0012 +0.090*(post-test rating)	0.73	22
Rural	-2.13 + 0.10*(post-test rating)	0.76	8
Urban	-8.68 +0.24*(post-test rating)	0.89	6

Q4: Estimation of Workload (of Clips) from Driving Performance (Real-Time Estimate of Workload)	
Condition	**Mean Workload Rating =**
All data, most strict entry requirement, 82% variance	8.86 -3.00(LogMeanRange125) + 0.47(MeanTrafficCount)
All data, looser entry, 87% variance	8.87 - 3.01(LogMeanRange125) + 0.48(MeanTrafficCount) + 2.05(MeanAxFiltered)
Some data, 85 % variance	8.07 – 2.72(LogMeanRange125) + 0.48(MeanTrafficCount) + 2.17(MeanAxFiltered) - 0.34(MinimumVpDot(0 removed))
Where:	
LogMeanRange125=	Logarithm mean distance (m) to the same-lane lead vehicles over 30 s interval. If no lead vehicle, mean distance = 125 m
MeanTraffficCount =	Mean # vehicles detected (15 deg FOV check degree field of view), over 30 s inteval
MeanAxFiltered =	Mean longitudinal acceleration (m/s2)
MinimumVpDot = (0 removed)	Min acceleration of lead vehicle (m/s2) over 30 s interval, exclude case of no lead vehicle.

Q6: Relative Contribution of Various Factors to Overall Workload (For Final Workload Calculation)					
Factor	**Road Type**				
	Xway	**Rural**	**Residential**	**Urban**	**Mean%**
Road geometry	1.3	2.1	1.7	1.6	17
Road surface condition	2.8	3.1	2.7	2.7	28
Visibility	2.5	3.4	2.8	2.5	28
Traffic	3.4	2.3	2.8	3.2	29

TABLE OF CONTENTS

INTRODUCTION

Over the last few years, the topic of driver distraction has received considerable attention in the scientific literature (Glaze and Ellis, 2003; Horrey and Wickens, 2003; Young, Regan, and Hammer, 2003; Uchiyama, Kojima, Hongo, Terashima, and Wakita, 2004; Victoria Road Safety Committee, 2006) and in the media (time.blogs.com/daily_rx/2006/06/talking_on_cell.html, www.nhtsa.dot.gov/nhtsa/announce/testimony/distractiontestimony.html, www.cartalk.com/content/features/Distraction/, www.cartalk.com/content/read-on/1999/10.21.html, and www.morganlee.org). The focus of public concerns has been on the dangers of driving and using a cell phone, though there are many other sources of distraction to consider. As use of cell phone features such as texting and web access increases, the distraction problems could increase as well.

The distraction problem is a specific legal concern, especially to vehicle manufacturers and suppliers in the U.S. because of product liability laws and judgments. An accepted standard is that products should be designed for reasonable and expected use and misuse. It is common knowledge that cell phones are used for various tasks while driving and that other tasks (e.g., using navigation systems) are performed while driving as well. Accordingly, given liability, those tasks should be designed so they can be performed safely while driving, or a context needs to be established so those tasks are not performed while driving.

This could be achieved in several ways. Drivers could be educated on the risks of performing distracting tasks while they drive. However, historically, driver education has only been effective in teaching drivers skills, not in teaching behavior. Another solution is to ban cell phone use entirely, which has proven to be politically challenging (www.ncsl.org/programs/transportation/cellphoneupdate05.htm). A third solution is to outfit vehicles with workload managers, systems that will determine the workload a driver is experiencing from the primary driving task, estimate the load of the second and potentially distracting task, and, combining that with other information, determine what is appropriate for drivers to do (Michon, 1993; Green, 2004).

For that to occur, one needs data on what drivers are willing to do as a function of workload, what is safe to do, and a means to determine driver workload as a function of traffic, road geometry, and other characteristics.

What Equations, Rules, and Other Evidence Have Been Developed to Predict Workload of Driving?

Initial U.S. Studies of Workload

The number of studies concerning driving workload is extremely lengthy. However, many of them concern topics such as the measurement of workload (e.g., Tijerina, Angell, Austria, Tan, and Kochhar, 2003; Young, Regan, and Hammer, 2003), test procedures such as from the Advanced Driver Attention Metrics (ADAM) project

(Breuer, Bengler, Heinrich, and Reichelt, undated), or the identification of statistical differences between conditions. Given resource limitations, only a few selected studies are described here, with selection biased towards studies that provide or could provide quantitative predictions, often in the form of regression analyses. A large number of studies use ANOVA to describe statistically significant differences, and making predictions based on those studies is often difficult.

Based on an analysis of the literature, Hulse, Dingus, Fischer, and Wierwille (1989) proposed a formulation for the demand of driving. Subsequently, 5 graduate students studying human factors engineering and well acquainted with the concept of workload participated in an experiment to validate the proposal. They were shown a map of the route and then drove it twice, once for familiarization and then to rate the driving demand on a scale from 1 to 9 (1 = able to look away from the road for long periods (4 s or more); 5 = able to look away for periods of 1 to 1.5 s; 9 = not able to look away at all). Ratings considered the extent to which drivers could look away from the road and the possibility of unanticipated traffic, intersections, and interactions with other vehicles. Correlations of the ratings and workload equation that follows were reasonably high.

Workload (from 0 to 100) = = 0.4A + 0.3B + 0.2C + 0.1D

where:

A = 20 log2(500/Sd) **(Sight Distance Factor)**

where Sd = sight distance (m)
if Sd > 500, then A = 0
if Sd < 15.6, then A=100

B = (100*Rmax) / R **(Curvature Factor)**

where R = radius of curvature
Rmax = maximum value of the radius of curvature
(set to 18.52 m (60.7 ft), the turn radius for a city street)

note: R = 360X / (2pa)
X = arc length along the curve (m)
a = change in direction (degrees)

C = -40So + 100 **(Lane Restriction Factor)**

where So = distance of closest obstruction to road (m)
(phone pole, fence, ditch, etc.)
if So > 2.5, then C=0

$$D = -36.5W + 267 \qquad \textbf{(Road Width Factor)}$$

where W = road width for 2 lanes (m)
if W > 7.3 (24 ft, 12 ft lanes), then D = 0
if W < 4.57 (15 ft, 7.5 ft lanes), then D = 100

However, workload is not just due to the road geometry as explored by Hulse et al. (1989). Nygren (1995) had 55 truck drivers make tradeoffs between pairs of 5 factors (traction, visibility, traffic, road, and lighting) that contribute to driving workload, assuming each factor could have 2 levels. For each pair of factors, there were therefore 4 possible combinations. For example, for traffic density and lighting, they are traffic density (low, high) paired with lighting (day, night). However, one does not need to ask subjects to know that high traffic density paired with night lighting is the highest workload and low traffic density with daylight is the lowest workload, which simplified the experiment. Only the middle pairs needed comparison. (Which leads to greater workload, low traffic density at night or high traffic density during the day?) These pairwise judgments were analyzed using conjoint analysis, a multidimensional scaling technique. Table 1 shows the results. Notice that traction accounts for more than half of the total importance (at least for truck drivers).

Table 1. Relative Importance of Workload Factors.

Importance Rank	Relative Importance	Factor	Levels
Most	52%	Traction	Good, poor
	26%	Visibility	Good, poor
	13%	Traffic density	Low, high
	6%	Road	Divided, not divided
Least	3%	Lighting	Day, night

How could those designing workload managers use the results from these 2 experiments? From Nygren's results, one could compute a total workload score, weighting the 5 factors based on their relative importance (Table 1).

From Hulse's results, one could estimate the workload related to visibility using the A factor from Hulse's workload equation, where the value for visibility is proportional to the log of sight distance. In addition, data phase 1 of this project (Cullinane and Green, 2006) described later, could also be used.

The "road" factor could be the sum of the other factors in the equation (B+C+D). Interestingly, this suggests very different weights than those suggested by Hulse et al., where A, B, C, and D had equal weights. Currently, data for B, D, and D either can or will be obtained from a GPS navigation system.

Quantitative estimates for other factors could come from the literature or be developed by asking technical experts to generate values associated with good/poor for each situation and assuming the effects of each factor on the rating is linear. For example, for traffic, good might be considered LOS A and poor LOS E (though the scale goes to LOS F-failing). Data on traffic (vehicles / lane / hour) could be obtained in real time from traffic message broadcasts, estimated from previous traffic counts on an hour-by-hour basis, or estimated from ACC radar system returns.

For traction, coefficient of friction (mu) values of greater than or equal to 0.7 might be considered good and those less than 0.3 poor, but the relationship between workload and friction is unlikely to be linear (Fancher, 2007, personal communication; Karamihas, 2007, personal communication). For example, changing from a surface of 0.7 to 0.6 will have only a very modest effect, but changing from 0.3 to 0.2 (slippery snow) will have a major effect, and going to 0.1 (wet ice), even more so. So a function such as workload = constant x (mu.max – mu.now) will overpredict workload at high mus and underpredict at low values. An expression such as workload = -1 + e^^kx, where K>0 and a function of mu.max and mu.now might give a better fit to the effect of traction on workload. Furthermore, keep in mind that traction is vehicle specific and depends on vehicle handling characteristics, the tires and their wear, and the road surface. Fortunately, once that relationship is known, GPS-linked weather data from the U.S. DOT-proposed CLARUS system (www.its.dot.gov/clarus/index.htm), along with wheel spin data from traction control and dynamic stability control systems, could be used to make predictions about traction-related workload.

For lighting, the situation is also complicated. At night (Nygren's poor condition), driving is often data limited (Norman and Bobrow, 1975: Flannagan, 2007, personal communication). People do not know what they are missing. Furthermore, what people can see in using focal vision (to guide the vehicle) and ambient vision (to detect moving threats) changes in nonlinear ways with respect to ambient illumination. (See Liebowitz and Owens, 1977 for a discussion of these 2 visual systems.) Thus, using linear functions for these characteristics to estimate workload can be both misleading and difficult. Nonetheless, as a first approximation one could use the state of the headlight switch or ambient illumination sensors (where provided) to determine if it is day or night, and treat this variable as binary.

EU Research on Workload and Workload Managers

Starting in the 1990s, a large number of studies were conducted in Europe to develop workload managers to reduce telematics-induced distraction, which are comprehensively reviewed in Hoedemaeker, de Ridder, and Janssen (2002). Major topics include (1) the measurement of driver behavior and performance, (2) how to manage workload, (3) how to create a workload manager, and (4) how to achieve driver acceptance of workload managers. Projects discussed in detail include GIDS, ARIADNE, GEM, IN-ARTE, and COMUNICAR. (See Table 2.)

Table 2. Major EU Projects relating to Workload

Project	Partners	Objectives/ summary
GIDS (1990-1992)	U of Groningen, Delft U of Technology, INRETS-LEN, Philips, Saab, Yard Ltd, Renault, VTI, U of the Bundeswehr, U College Dublin, TNO Human Factors	Determine requirements & design standards for co-driver, included navigation system & cell phone, 2 demonstrators (1 car, 1 simulator)
ADRIADNE (1992-1994)	Rover, British Aerospace, Philips Research Labs, CARA Data Processing, U of Groningen, MRC Applied Psychology Unit, TNO Human Factors, VTI	
GEM (1994-1995)	Rover, British Aerospace, Philips, TNO Human Factors, Acit, TRC Groningen, U of Leeds, VTI	
IN-ARTE (1998-1999)		
COMUNICAR	CRF-Fiat, Volvo, Daimler Chrysler, Mertavib, Frauenhofer IAO, Bord, BAST, U of Genoa, U of Siena, Technical U of Athens, TNO Human Factors	Formerly www.comunicar-eu.org/ interface is central display, panel cluster, haptic knob
CO-DRIVE	TNO	

Overall, Hoedemaeker, de Ridder, and Janssen (2002, page 5) conclude that with regard to measurement, “Efforts to monitor momentary driver workload by more or less intrusive means will not succeed, or will never be suitable for practical applications, even though such methods might be theoretically best.” They report that workload has been estimated both by looking at driver actions and monitoring the effect on performance (e.g., headway), and by monitoring the driving situation and estimating workload using a lookup table. Key aspects of driver-vehicle interaction include the initiation and control of interaction sequences (driver or the vehicle), the total glance time to the display, the mental workload of the interaction, and the number and precision of movements required. Indicators of workload have also been obtained from driver actions (use of brakes, steering wheel, turn signal, etc.) and the environment (wiper, fog light status) that are easy to sense. Unfortunately, that report and many of the reports cited (or at least those that are publicly available) do not provide quantitative information on the relationship between the measures of interest and workload, information needed to build a workload manager.

Review of some of the web sites (or at least, those that are still active) and reports for these projects do provide some information about how workload is estimated, but the information desired (the particular parameters, measures, and equation used to determine workload) are rarely provided. Even the GIDS book, the first significant effort to develop a workload manager, states the following:

> "The following (continental) situations may require the system to intervene:
>
> Car following
> (1a) The car is too close to a vehicle in front that is in the same lane
>
> Rear vehicle
> (2a) The rear vehicle is close to the car which is decelerating too hard."
> (Michon, 1993, p. 101).

Unfortunately, terms such as "too fast," "too close," and "too hard" are never defined.

One noteworthy exception is a workload calculation described in Piechulla, Mayser, Gehrke, and König (2002) from the SANTOS project. Their calculation is based on data from subjects driving a test route that had been coded using Fastenmeier's (1995) taxonomy of traffic situations. Situations were coded on 6 dimensions: (1) road type (5 highway classes, 2 rural road classes, 7 city classes) (2) horizontal layout (curve versus no curve) (3) vertical layout (slope versus plane route) (4) intersections (4 classes) (5) route constrictions (yes/no) and (6) driving direction (straight ahead, turn left, turn right). On the test route, there were 186 scenarios, which were grouped into 22 unique situations using the Fastenmeier scheme. While driving, subjects looked for text on a slowly scrolling visual display. The dependent measure was the number of glances per second averaged over subjects for each of the 22 situation classes, which varied from 0.803 to 0.476. As fewer glances per second were associated with greater workload, workload was defined as the 1-mean glance frequency. Unfortunately, the authors of this report do not list those 22 situations, the glance data, or the workload estimates for them.

Data for those 22 situations are the core of a very thoughtful workload manager described in Piechulla, Mayser, Gehrke, and König (2003). One can get a sense of how his workload manager functions from an on-line demo (www.walterpiechulla.de/workloadpages/index.html). As shown in Figure 1, the workload manager begins by doing a table look-up of the workload due to the road segment being driven using the 6 dimensions of the Fastenmeier coding scheme. However, the workload incurred is both due to the road segment at the moment and planning for the road ahead. Piechulla et al. postulate that looking about 5 s ahead is reasonable, and that workload experienced decays exponentially with time y=2.71866e^^(-x/4.72657), where x and y are not defined. Figure 1 shows the calculation procedure proposed, presumably only for a vehicle fitted with an ACC (adaptive cruise control) system similar to that in the BMW test vehicle (pre-2003). In brief, the calculation involves determining if a vehicle is in range (120 m). If yes, then the workload is increased by 10 percent. If an intersection is in view (presumably 5 seconds), then the workload is also increased by 10 percent. Hard braking (in excess of 1 m/s2 or 0.1 g) also increases workload, and ACC operation (or at least the ACC system in Piechulla's pre-2003 BMW) reduces it (by 8 percent). As shown in the figure, passing (overtaking) and rapid approach all alter workload.

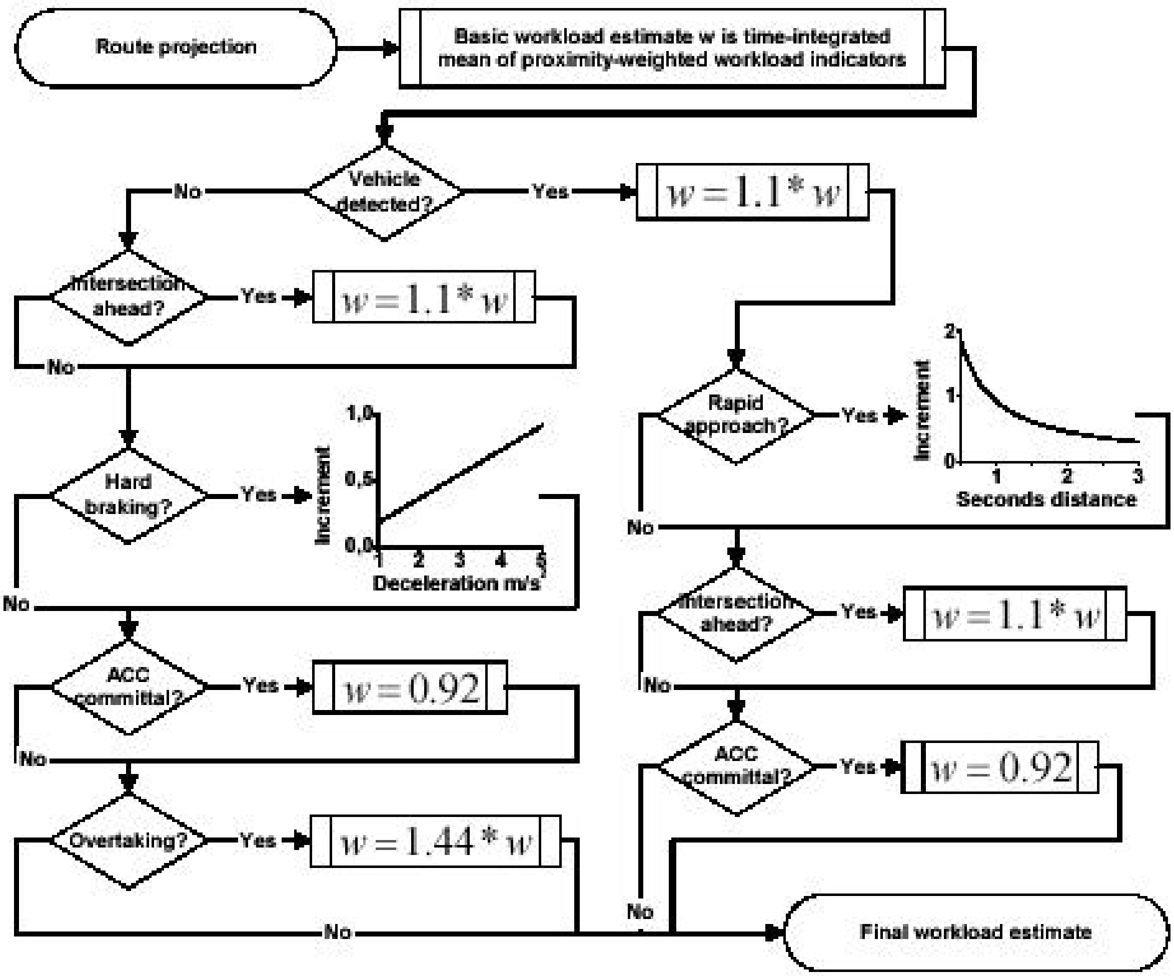


Figure 1. Adjustment of Workload Estimates in Piechulla Model

The model proposed by Piechulla et al. is quite interesting and represents a significant step beyond Hulse et al. and Nygren in that it presents quantitative workload estimates for real roads and for a wide range of driving situations. It also introduces the idea that workload is due in part to the road segment being approached. In terms of SAVE-IT, the model includes heading control and ACC, whose impact has not been given much consideration. Interestingly, the model only considers a single lead vehicle, not multiple vehicles as traffic, and includes overtaking maneuvers. Overtaking is assumed to mean going past another vehicle in another lane, not a flying pass that involves a lane change. This is an important assumption because overtaking leads to one of the largest increments in workload.

A more detailed model from an earlier paper, translated here (Milla, 2007, personal communication) from the German original (Piechulla, Mayser, Gehrke, and König, 2002), appears in Figure 2. In contrast to the work of Nygren, Piechulla et al. (2002)

suggest only very modest increases in workload due to darkness (2.6%), rain (5%), a wet surface (2.5%), and ice (10%).

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Figure 2. Model Presented in Piechulla, Mayser, Gehrke, & König, 2002 (translated)

Motorola Driver Advocate Project

The goal of the Motorola project was to determine if the driver was distracted, not to measure workload per se. In contrast to the approach used by Piechulla et al., that classified driving situations, the Motorola work by Torkkola et al. examined correlations between driving performance statistics and driver state (distracted vs. attentive) based on where drivers looked (toward or away from the road).

More specifically, Torkkola, Massey, and Wood (2004) describe an experiment in which subjects drove in the middle lane of a simulated 3-lane expressway (at 55 mi/hr in "heavy" traffic). The road surface was dry and driven in the daylight. At various times subjects were cued to look at images in their blind spot (left or right) for up to 5 seconds. They were paid a bonus when they correctly identified characteristics of the image in the blind spot (its color, kind of vehicle, etc.) in response to post-glance experimenter questions. Driving performance was recorded using sensors that would be present in an otherwise ordinary vehicle with a collision avoidance system, sampling at 60 Hz. Table 3 shows 7 basic measures recorded and Table 4 shows 5 statistics computed for each of them. Statistics were selected to provide estimates of typical values, trends, variability, and rate of change for the 7 basic measures.

Table 3. Measures Used by Torkkola, Massey, and Wood (2004)

Abbrev.	**Statistics (all sampled at 10 Hz)**	**Comment**
SWa	Steering wheel angle	Units known
Ap	Accelerator position	Measure (angle?), units unknown
LLEd	Left lane edge distance (=left front wheel from left lane edge)	From where on the tire to where on the line
CLa	Cross lane (lateral) velocity (=rate of change of distance to left lane edge)	Units unknown
CLv	Cross lane (lateral) acceleration (=rate of change of cross lane velocity)	Units unknown
Se	Steering error (=difference between current wheel angle and angle for travel parallel to lane edges)	Units unknown
Lb	Lane Bearing (Vehicle heading=angle of vehicle to angle of road 60 m ahead)	Units unknown

Table 4. Statistics Computed by Torkkola, Massey, and Wood (2004)

Statistic	Definition	Comment
Ra9	Moving mean of sign over 9 previous samples	Typical value - smoothed version of signal
Rd5	Moving difference 5 samples apart	Trend
Rv9	Moving standard deviation of 9 previous samples	Variability
Ent15	Entropy of error for linear predictor of signal	Randomness/ Unpredictability/variability
Stat3	Multivariate stationarity of a number of variables 3 samples apart	Overall rate of change of a group of signals, 1 for none change, 0 for drastic change

The 7 basic variables, plus 13 statistics based on them (20 total, Table 5), were used to predict if the driver was attentive and if so if the driver was looking left or looking right. This atheoretic approach did quite well, detecting 78% of the inattentive time segments (to the nearest 0.1 s) and 98.4% of the attentive time segments (Table 6). Notice there is some change in the order between the 2- and 3-state detectors (Table 5). The authors do not suggest how which factors to include or their importance would change with road type, weather, road surface conditions, visibility or other factors that affect workload and attention to driving.

Table 5. Importance of Signals for Inattention Detector

Variable	Importance	
	2-State (attentive or not)	3-State (attentive left, right, not)
distToLeftLaneEdge_rd5_ra9	100.00	69.87
steeringWheel_rv9	99.94	57.17
Accelerator	98.72	100.00
Stat3_of_steeri ngWheel_accel	95.06	61.09
crossLaneVelocity	94.79	65.64
steeringWheel_ent15_ra9	90.37	57.32
distToLeftLaneEdge	80.62	55.85
aheadLaneBearing_rd5_ra9	79.90	71.22
distToLeftLaneEdge_rv9	77.80	60.35
aheadLaneBearing	75.24	71.22
steeringWheel	70.90	64.80
steeringError	68.26	58.77
crossLaneVelocity_rv9	68.13	68.68
Stat3_of_steeringErrorcrossLaneVelocity distToLeftLaneEdgeaheadLaneBearing	60.84	49.52
steeringWheel_rd5_ra9	56.12	51.74
steeringError_rd5_ra9	47.91	54.38
Accelerator_ent15_ra9	40.96	41.79
Accelerator_rv9	38.35	43.55
crossLaneAcceleration	34.54	36.95
Accelerator_rd5_ra9	31.33	38.24

Table 6. Detection Matrix for Attention/Inattention Detectors

2-State Detector		
Actual	Predicted	
	Attentive	Inattentive
Attentive	19988=98.4%	319=1.57%
Inattentive	355=21.58%	1290=78.42%

3-State Detector			
Actual	Predicted		
	Attentive	Inattentive Left	Inattentive Right
Attentive	9230=99.79%	4=0.04%	15=0.16%
Inattentive Left	30=14.78%	173=85.22%	0
Inattentive Right	54=18.82%	0	233=81.18%

As a follow-on to this work, Torkkola, Venkatesan, and Liu (2004) attempted to identify individual maneuvers using machine learning. The first step was to identify which sensors should be used. Four subjects drove for 15 minutes each in a world that consisted of 2- and 3- lane expressways, and 2- and 4-lane urban, suburban, industrial, and rural roads. Traffic was present and vehicle speeds varied. Drivers performed 12 types of maneuvers (ChangeLeft, ChangeRight, CrossShoulder, NotOnRoad, Pass, Reverse, MoveSlow, Start, Stop, Tailgate, TurnRight, and UTurn). Some maneuvers overlapped (e.g., Pass=ChangeLeft followed by ChangeRight).

In their analysis Torkkola et al. examined (1) a base set of 15 variables (Table 7), (2) all quadratic terms (cross products and squares of those 15), (3), all derivatives of the 13 continuous variables, (4) short time entropies for steering, brake, and accelerator, (5) multivariate stationarity with delta=2 and 3, and (6) the output of a quadratic classifier trained using a least squares method for the 13 continuous variables. (Turn signal and VehicleAhead were the only discrete variables.)

Table 7. Variables Used by Torkkola, Venkatesan, and Liu (2004)

Variable	**Description**
Accelerator	Normalized accelerator input value
Brake	Normalized brake input value
Speed	Speed of the subject (m/s)
Steer	Normalized steering angle (deg)
Turn Signal	Status of indicator lights
AheadLaneBearing	Bearing of the current lane 100 meters ahead
CrossLaneAcceleration	Acceleration perpendicular to the lane (m/s2)
CrossLaneVelocity	Velocity perpendicular to the lane (m/s2)
RightLaneEdgeDistance	Distance to the right edge (m)
LeftLaneEdgeDistance	Distance to the left edge (m)
LaneOffset	Offset relative to the center of the lane (m)
LateralAcceleration	Acceleration perpendicular to the vehicle (m/s2)
HeadwayDistance	Distance from the subject's front bumper to the rear bumper of any vehicle ahead (m)
HeadwayTime	Time to the vehicle ahead (s)
VehicleAhead	Name of the closest vehicle ahead of the subject in the same lane

For all maneuvers, turn signal and speed were important, and for some stationarity of the sensors and entropy of steering and braking were high. Table 8 shows the sensor-derived measures associated with some of the maneuvers. The image in that table, pasted from the original source, is the best available. Those interested in further details should see the original source (Torkkola, Venkatesan, and Liu, 2004).

Table 8. Maneuvers and Associated Measures

TABLE II
FEATURES SELECTED

Random Forests				CfsSubsetEval-BestFirst	
Maneuver	**All Maneuvers (cont)**	**ChangingLeft (cont)**	**ChangingRight (cont)**	**Maneuver**	**All Maneuvers (cont)**
quadClassValue	stat2	distToRightLane_ra3_rd3_ra3_rd3	aheadLaneBearing*aheadLaneBearing	crossLaneVelocity*crossLaneVelocity	steer_ent15
speed	turnSignal	accelerator*speed	stat3	crossLaneVelocity*HeadwayTime	stat3
speed*speed	aheadLaneBearing*laneOffset	speed*aheadLaneBearing	accelerator*aheadLaneBearing	distToLeftLane*distToRightLane	**ChangingLeft**
stat3	speed*distToLeftLane	aheadLaneBearing	speed*distToLeftLane	distToRightLane*laneOffset	turnSignal
stat2	brake_ent15	crossLaneVelocity_ra3_rd3_ra3_rd3	speed	quadClassValue	distToRightLane_ra3_rd3
steer*aheadLaneBearing	accelerator*speed	crossLaneVelocity*distToLeftLane	distToRightLane_ra3_rd3_ra3_rd3	brake_ent15	distToLeftLane_ra3_rd3_ra3_rd3
turnSignal	crossLaneVelocity*crossLaneVelocity	aheadLaneBearing*aheadLaneBearing	aheadLaneBearing_ra3_rd3_ra3_rd3	steer_ent15	HeadwayDist_ra3_rd3_ra3_rd3
steer_ent15	steer*steer	speed*speed	speed*speed	stat3	accelerator*crossLaneVelocity
speed*distToLeftLane	aheadLaneBearing*aheadLaneBearing	stat3	turnSignal	**All Maneuvers**	accelerator*HeadwayDist
brake_ent15	aheadLaneBearing*lateralAcceleration	laneOffset_ra3_rd3_ra3_rd3	distToLeftLane	speed	speed*aheadLaneBearing
aheadLaneBearing*lateralAcceleration	aheadLaneBearing	crossLaneVelocity*HeadwayTime	lateralAcceleration_ra3_rd3	turnSignal	speed*crossLaneVelocity
aheadLaneBearing*laneOffset	steer*aheadLaneBearing	aheadLaneBearing*HeadwayDist	aheadLaneBearing*distToRightLane	aheadLaneBearing	speed*distToRightLane
steer*steer	speed_ra3_rd3	accelerator	crossLaneVelocity*crossLaneVelocity	steer_ra3_rd3	crossLaneVelocity*HeadwayDist
laneOffset*laneOffset	steer	accelerator*aheadLaneBearing	steer_ra3_rd3	distToLeftLane_ra3_rd3	crossLaneVelocity*HeadwayTime
crossLaneVelocity*crossLaneVelocity	aheadLaneBearing*distToRightLane	accelerator*HeadwayDist	aheadLaneBearing*crossLaneVelocity	HeadwayDist_ra3_rd3	stat3
distToRightLane*distToRightLane	laneOffset*laneOffset	aheadLaneBearing*HeadwayTime	stat2	brake_ra3_rd3_ra3_rd3	**ChangingRight**
distToRightLane	distToRightLane	speed*distToRightLane	accelerator*laneOffset	speed_ra3_rd3_ra3_rd3	turnSignal
accelerator*speed	speed*distToRightLane	steer*aheadLaneBearing	**CfsSubsetEval-BestFirst**	steer_ra3_rd3_ra3_rd3	aheadLaneBearing
distToRightLane_ra3_rd3_ra3_rd3	crossLaneVelocity	accelerator*HeadwayTime	**Maneuver**	aheadLaneBearing_ra3_rd3_ra3_rd3	crossLaneVelocity
speed*distToRightLane	distToRightLane*laneOffset	distToLeftLane_ra3_rd3_ra3_rd3	speed	distToRightLane_ra3_rd3_ra3_rd3	distToRightLane_ra3_rd3
speed_ra3_rd3	speed*crossLaneVelocity	accelerator*accelerator	steer	HeadwayDist_ra3_rd3_ra3_rd3	distToRightLane_ra3_rd3_ra3_rd3
distToLeftLane*distToRightLane	distToRightLane*distToRightLane	aheadLaneBearing*lateralAcceleration	turnSignal	accelerator*speed	laneOffset_ra3_rd3_ra3_rd3
laneOffset_ra3_rd3_ra3_rd3	distToRightLane_ra3_rd3	**ChangingRight**	steer_ra3_rd3	accelerator*steer	accelerator*crossLaneAcceleration
aheadLaneBearing*aheadLaneBearing	distToLeftLane	speed*crossLaneVelocity	laneOffset_ra3_rd3	accelerator*crossLaneVelocity	brake*crossLaneVelocity
distToRightLane_ra3_rd3	distToLeftLane*distToRightLane	crossLaneVelocity	brake_ra3_rd3_ra3_rd3	brake*lateralAcceleration	speed*aheadLaneBearing
distToLeftLane_ra3_rd3	distToLeftLane_ra3_rd3_ra3_rd3	crossLaneVelocity*distToRightLane	speed_ra3_rd3_ra3_rd3	speed*speed	Speed*crossLaneAcceleration
distToLeftLane_ra3_rd3_ra3_rd3	speed*aheadLaneBearing	distToRightLane_ra3_rd3	steer_ra3_rd3_ra3_rd3	speed*crossLaneVelocity	speed*crossLaneVelocity
brake*speed	**ChangingLeft**	crossLaneVelocity*distToLeftLane	aheadLaneBearing_ra3_rd3_ra3_rd3	speed*distToLeftLane	speed*distToLeftLane
aheadLaneBearing*crossLaneVelocity	speed*crossLaneVelocity	accelerator*crossLaneVelocity	distToRightLane_ra3_rd3_ra3_rd3	speed*distToRightLane	speed*distToRightLane
aheadLaneBearing*distToRightLane	crossLaneVelocity	speed*aheadLaneBearing	HeadwayDist_ra3_rd3_ra3_rd3	steer*aheadLaneBearing	speed*laneOffset
distToLeftLane	turnSignal	distToLeftLane_ra3_rd3	HeadwayTime_ra3_rd3_ra3_rd3	steer*distToLeftLane	crossLaneAcceleration*laneOffset
distToRightLane*laneOffset	accelerator*crossLaneVelocity	laneOffset_ra3_rd3	brake*lateralAcceleration	steer*laneOffset	CrossLaneVelocity*HeadwayDist
All Maneuvers	crossLaneVelocity*distToRightLane	speed*laneOffset	speed*speed	aheadLaneBearing*laneOffset	distToLeftLane*distToRightLane
speed	distToLeftLane_ra3_rd3	aheadLaneBearing	speed*distToLeftLane	crossLaneAcceleration*crossLaneAcceleration	stat3
speed*speed	distToRightLane_ra3_rd3	distToLeftLane_ra3_rd3_ra3_rd3	steer*aheadLaneBearing	crossLaneVelocity*HeadwayDist	stat2
quadClassValue	speed	laneOffset_ra3_rd3_ra3_rd3	steer*laneOffset	distToLeftLane*distToRightLane	
stat3	crossLaneVelocity*HeadwayDist	crossLaneVelocity_ra3_rd3_ra3_rd3	aheadLaneBearing*lateralAcceleration	quadClassValue	
steer_ent15	laneOffset_ra3_rd3	speed*distToRightLane	crossLaneAcceleration*crossLaneVelocity	brake_ent15	

ra3 / rd3 – Running Average / Difference with window size of 3 samples; a*b – Feature *a* multiplied with Feature *b* (Cross product); a_ent15 – Running entropy of the prediction error of feature *a*; statN – stationarity with delta = N; quadClassValue – The output of the quadratic classifier

Torkkola, Venkatesan, and Liu (2005) used the same data, variables, and statistics as the previous experiment, but focused on only 6 maneuvers (ChangeLeft, ChangeRight, Pass, Start, Stop, Tailgate). Instead of using random-forest based feature selection, they used hidden Markov models. An important part of the process was to identify the common subunits of maneuvers (drivemes). Based on the figures presented, the results from this approach make sense, but the authors do not provide enough information to build a maneuver identifier for a workload manager.

That work has continued at Motorola; the most recent summary is Torkkola, Gardner, Schreiner, Zhang, Leivian, and Summers (2006). In this paper, the focus is on classifying 29 different maneuvers as shown in Table 9. Figure 3 shows their classification algorithm in operation, where the time scale is 100 ms increments. Based on this example, the performance of their algorithm looks quite good.

Table 9. Maneuvers Classified by Torkkola et al. (2006)

ChangingLaneLeft	ChangingLaneRight	ComingToLeftTurnStop
ComingToRightTurnStop	Crash	CurvingLeft
CurvingRight	EnterFreeway	ExitFreeway
LanChangePassLeft	LaneChangePassRight	LaneDepartureLeft
LaneDepartureRight	Merge	PanicStop
PanicSwerve	Parking	PassingLeft
PassingRight	ReversingFromPark	RoadDeparture
SlowMoving	Starting	StopAndGo
Stopping	TurningLeft	TurningRight
WaitingForGapInTurn	Other (Cruising)	

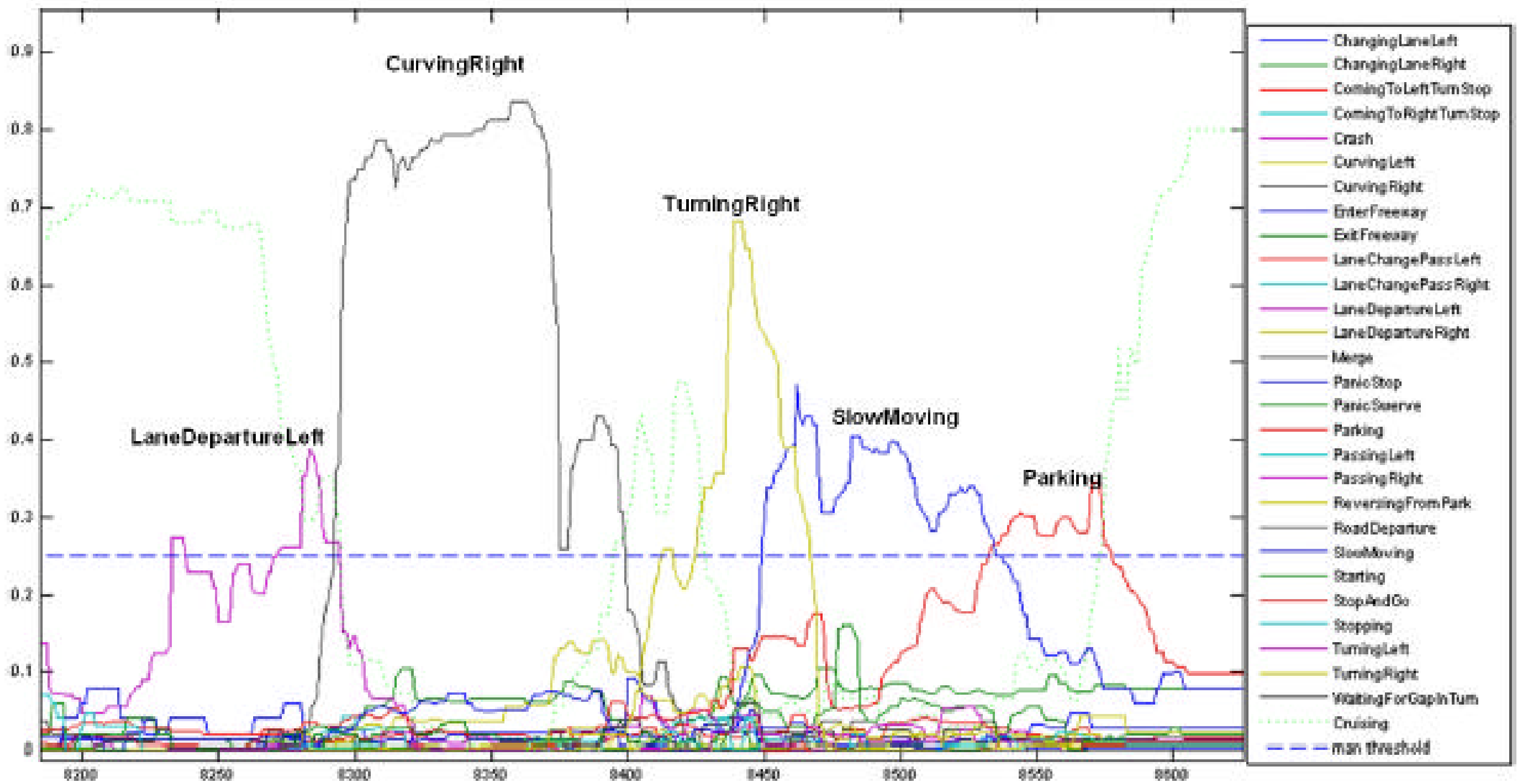


Figure 3. Maneuver Probability Example from Torkkola et al. (2006)

What Factors Affect the Workload of Secondary Tasks?

The focus of the experiment in this report is on quantifying the demands of the driving task. However, as part of that experiment, subjects were asked if they would be willing to do certain secondary tasks in particular situations. Therefore, some mention of the factors affecting secondary task demand is needed. In brief, the extent to which tasks add to driver workload depends on (1) driver exposure, (2) task intensity and its demand on the resources shared with driving, (3) driver experience with the tasks, (4) the engagement of those tasks, and, some have argued, (5) task interruptability. Some discussion of each of those points follows.

Driver exposure is a function of secondary task duration (longer exposure leads to greater load over time) and frequency (more often leads to greater load). It has been argued that when performed statically (with a vehicle parked), visual-manual tasks requiring more than 15 seconds to complete should not be performed while driving. That requirement is part of SAE Recommended Practice J2364 (Society of Automotive Engineers, 2004a). There is evidence, however, supporting even shorter task durations (Society of Automotive Engineers, 2004b).

Task intensity and resources for a number of common in-vehicle tasks examined in Yee, Nguyen, Green, Oberholtzer, and Miller (2007), an analysis conducted in phase 2 of this project. In brief, in accomplishing a task, people may utilize visual, auditory, cognitive, and psychomotor (VACP) resources. According to multiple-resources theory, overload may occur when any one of those resources is overloaded (Wickens, Gordon, and Liu, 1998), such as when 2 tasks make high demands for the same resource. The multiple-resources theory underlies tools such as IMPRINT (Mitchell, 2000). Though data on the time varying demands of the primary task are not available, data on the demands and the frequency of occurrence of many secondary subtasks that occur while driving (e.g., picking up a cell phone) are provided in that report.

Also important is the extent to which a task engages a driver. In some sense, this is the core of a distraction, something that attracts driver attention. Tasks such as dealing with a bee in a car or a crying baby are good examples of tasks that are engaging, that draw the drivers' attention. Quite frankly, this characteristic has not been given much consideration in the driving literature, and it certainly has not been quantified. Key aspects include risk to the safety of the driver and passengers (such as the bee in the car or a crash warning message), potential vehicle damage (such as from an unattended spill), if the task has financial or business consequences, the relevance of the task to the trip (such as route guidance), the time for which information is available or how soon it is needed (such as seeing an exit ahead and needing to make a decision before it is reached), if the task is initiated by the driver or externally, if the task involves verbal communications, and so forth. (See Lerner, 2005.)

Task experience matters. With practice, people do tasks more rapidly and accurately, and often the demands for visual and cognitive resources are reduced. However, for many of the tasks of interest, except probably those related to dialing, texting, and some entertainment system tasks, experience with the task can be limited.

Finally, driver interfaces that are not interruptible (for example those with limited timeouts that force a driver to continue a task, such as a navigation data entry screen that would blank after 2 seconds of no input) are a bad idea. Fortunately, such interfaces are rare. However, the assertion is that drivers perform secondary tasks in almost a casual manner—they enter a state, and that after the driving conditions are ideal, they enter the city, and they wait a while and then… In fact, observations of drivers indicate people do not behave that way, though published research documenting this, one way or the other, is absent in the open literature. Once starting an in-vehicle task, drivers are fairly persistent in completing it. Quite frankly, it could be differences of opinion on this may reflect different personal experiences, namely observations of German drivers versus American drivers. Data to resolve the extent to which secondary tasks are interrupted in naturalistic driving by drivers in different countries are needed.

A more extensive review of the factors that affect the demands of secondary tasks appears in a report in phase I of the SAVE-IT project (Zhang and Smith, 2004), focusing on mean task times and task time variance. In terms of secondary tasks that drivers should not do or do not want to do while driving, they identify (1) Rockwell's 2-second rule (drivers are reluctant to look away from the road for more than 2 seconds at a time) and (2) the SAE J2364 15-second rule.

As a Function of Driving Workload, Which Tasks Do Drivers Find Acceptable to Do and When?

Since the Phase 1 report was completed, one particularly noteworthy study of direct relevance to this report has been completed. Lerner conducted 6 focus groups and an on-road experiment to address what drivers find acceptable (Lerner, 2005; Lerner and Boyd, 2006). Those 6 groups consisted of teenagers, young drivers, 2 middle-aged groups, older drivers, and navigation system users (a total of 45 drivers). The focus groups considered what drivers take into account when engaging in a secondary task, close calls drivers might have experienced, whether drivers are aware of when they are distracted, and other topics related to driving risk. A key finding was that "task motivations" seemed to be the predominant factors in deciding to a engage in a task followed by task attributes. Driving-related issues were the least predominant factor. Participants showed little concern for impending road conditions.

In the on-the-road experiment, 88 drivers equally drawn from 4 age groups (teen, young, middle, old) familiar to some degree with technology drove their own vehicle on a variety of roads. They identified their willingness to engage in various tasks while driving at each particular moment on a 1-to-10 scale ("1 = I would absolutely not do this task now, 10 = I would be very willing to do this task now with no concerns at all"). The precision with which subjects responded (nearest integer, tenth, hundredth) is not described, though it appears integers were suggested. However, the mean ratings are reported to the nearest hundredth of a point.

In addition, ratings of risk were also obtained. The devices were not used when the question was asked. A total of 81 of the 154 combinations of the 14 in-vehicle tasks (Table 10) with the 11 driving situations were explored. (Greater detail is provided in Lerner and Boyd, 2005.) At home, subjects subsequently completed a booklet that (1) examined why they rated the 11 driving situations as they did, (2) requested ratings of risk and if they were willing to engage in various tasks for various situations (5 duplications of on-road situations, 15 modifications of situations, and 20 new situations involving weather, passengers, etc. not tested on the road), (3) collected ratings for 32 tasks and 10 driving situations (and reasons why), (4) determined familiarity with their knowledge of the technology and associated tasks, and (5) collected ratings for personal characteristics such as aggressiveness, impulsiveness, and ability to perform multiple tasks concurrently.

Table 10. In-Vehicle Tasks and Driving Situations from Lerner and Boyd (2005)

In-Vehicle Tasks	Driving Situations
Cell phone: answer call	Freeway: proceed on mainline
Cell phone: key in call	Freeway: entrance/merge
Cell phone: personal conversation	Freeway: exit
Cell phone: key text message	Arterial: proceed on mainline
PDA: look up stored number	Arterial: unprotected left turn
PDA: pick up & read email	Arterial: protected U-turn
PDA: key in & send email	Arterial: stopped at red signal
Navigation system: key new destination	Parking lot: exit onto arterial
Navigation system: call up stored destination	Parking lot: search for space
Navigation system: search for Starbucks	2-Lane hwy: proceed, curvy
Select/insert CD	Residential street: proceed
Converse with passenger	
Drink hot beverage	
Unwrap/eat taco	

The discussion of the key results will emphasize the willingness-to-engage ratings as they were highly correlated (r=-0.98) with risk ratings. As shown in Figure 4, their mean willingness-to-engage ratings varied from about 9.5 (middle-aged driver, conversing with passenger) to about 2.2 (older drivers, using PDA to key and send email). For example, ratings for text messaging were just below 4, whereas conversation on a phone was in excess of 8. As shown in Figure 4, those ratings varied substantially with driver age, with the willingness to engage in tasks decreasing with age, but were relatively invariant with the type of road being driven (Figure 5). As a footnote, all subjects were familiar with cellular phones, two-thirds were familiar with PDAs, but just over half were familiar with navigation systems. (Even though participants viewed video clips demonstrating each task, the lack of actual task experience is a concern).

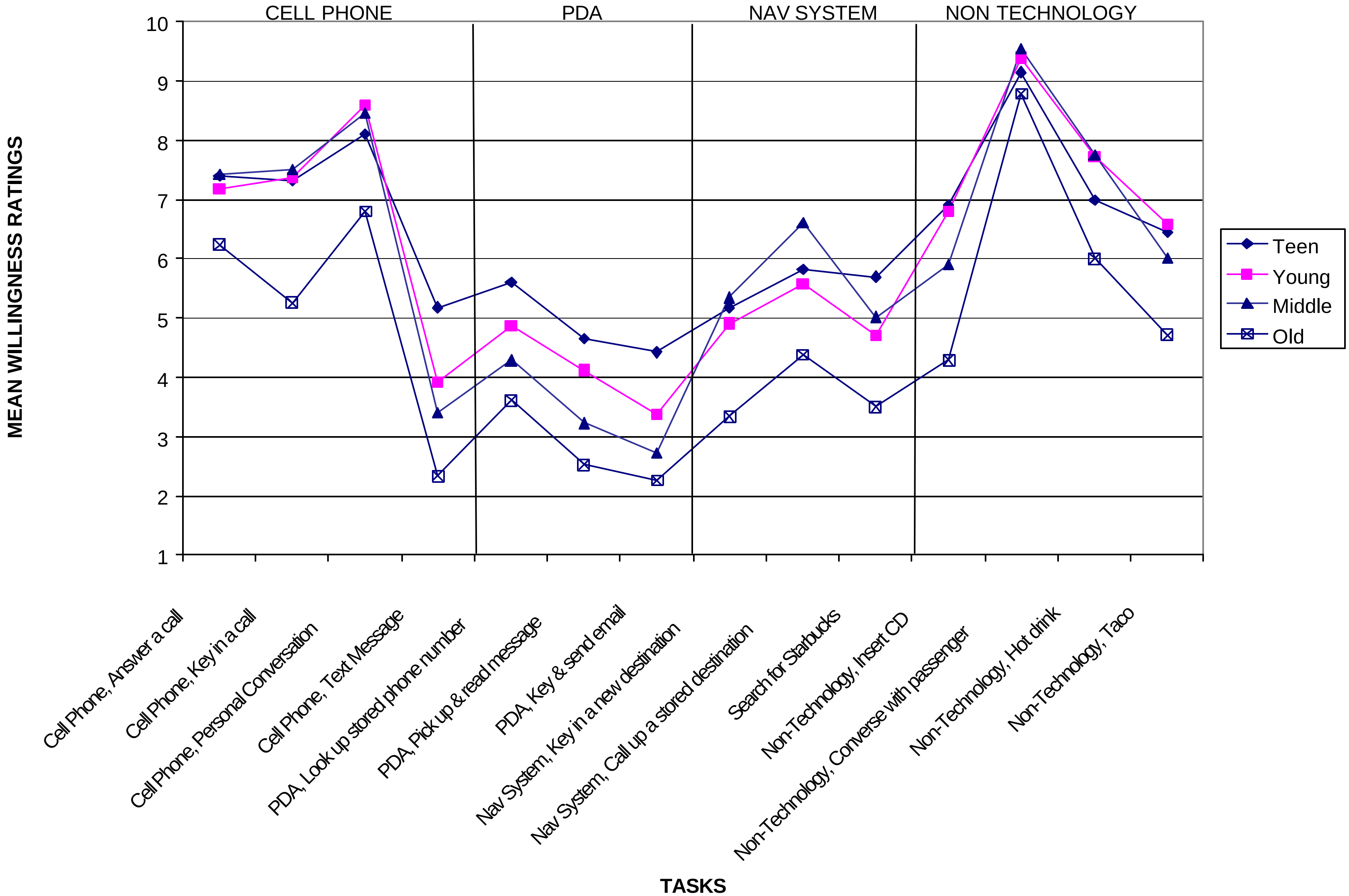


Figure 4. Willingness to Engage in Tasks as a Function of Driver Age

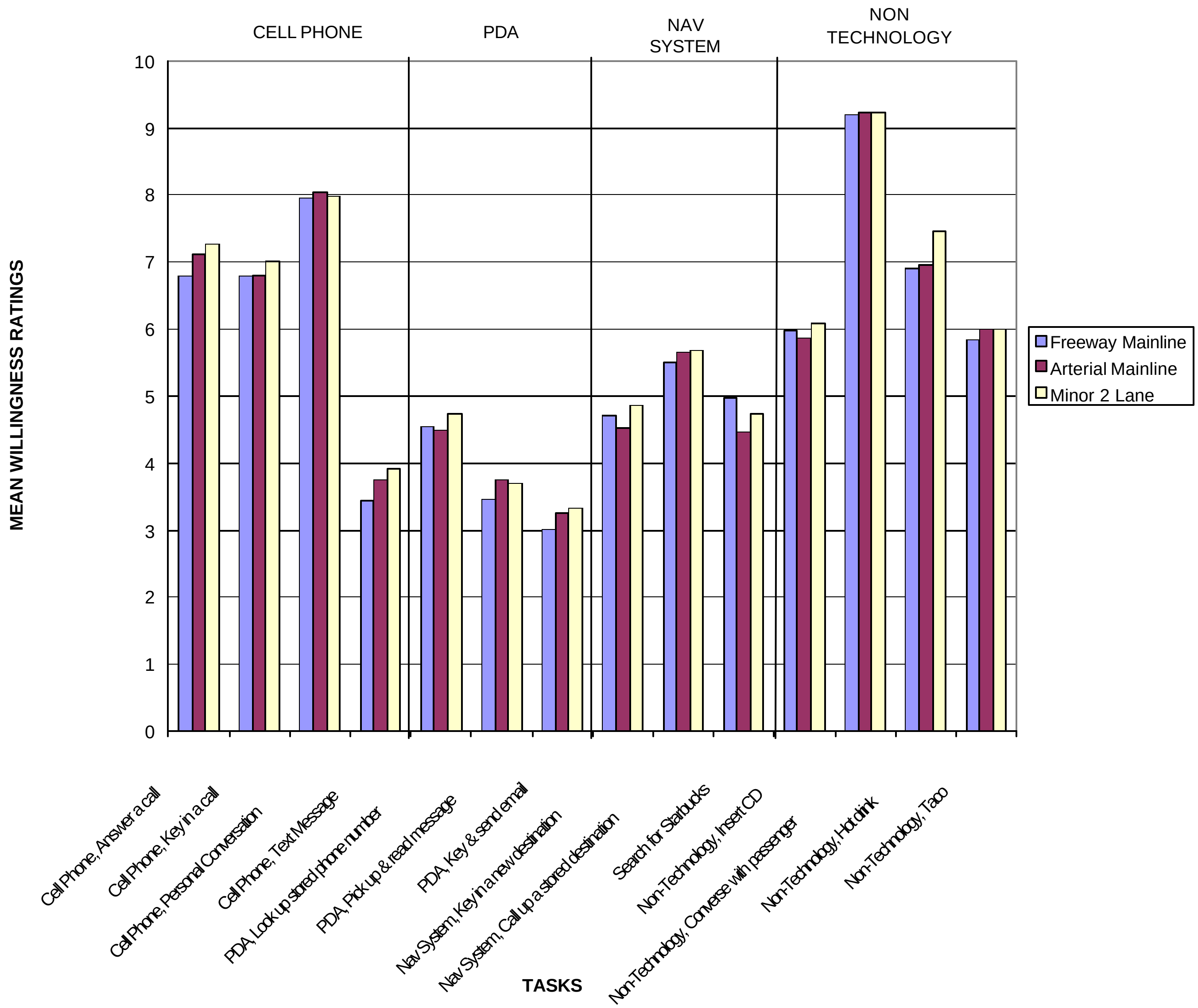


Figure 5. Willingness to Engage in Tasks as a Function of Road Type

Also of particular interest to the SAVE-IT project are the mean risk ratings for 32 in-vehicle tasks (Table 11). Notice that the riskiest tasks are associated with using a PDA and the next riskiest are tasks associated with navigation systems. Even the highest nontechnology tasks (eating a taco, dealing with children) were in the middle of the range of risk ratings.

Table 11. Mean Risk Ratings for All Drivers for Various In-Vehicle Tasks
Source: Lerner and Boyd (2005b)

In-Vehicle Task	Mean Risk Rating
Search the Internet using a PDA	8.93
Key in and send an email on PDA	8.33
Schedule a meeting using PDA	8.24
Open and read email on PDA	7.94
Take notes during a phone conversation	7.67
Check your schedule on PDA	7.51
Look up an entry in address book on PDA	7.29
Key a new destination into Nav System	6.93
Read a paper map	6.92
Alter your route preferences on Nav System	6.42
Find an alternate route on Nav System	6.31
Search for the nearest Starbucks on Nav Sys.	6.29
Retrieve a stored destination on Nav System	5.55
View an electronic map on Nav System	5.51
Eat something sloppy (like a taco)	5.51
Deal with children	4.53
Look up a stored phone number in a cell phone	4.50
Open and listen to voice mail on cell phone	4.41
Key in a cell phone call	4.17
Drink something hot	3.59
Have an extended phone conversation	3.50
Insert a CD, tape, or video	3.14
Find radio station that is not pre-programmed	2.97
Have a brief phone "exchange of information"	2.74
Place a cell phone call using speed dial	2.72
Answer a cell phone call	2.64
Eat something neat (like a cookie)	2.47
Drink something cold	2.39
Turn up the temperature	1.77
Talk with a passenger	1.69
Adjust the loudness of a sound system	1.69
Check the speedometer	1.37

Lerner and Boyd (2005) focus on the resource demands as suggested by subjects in their explanations of their risk ratings (Table 12). Added to the table is a column for demand, using terms common to VACP analysis. Unfortunately, several of the tasks examined by Lerner et al. were not examined by Yee et al (2007). Of the task characteristics leading to high demand in Lerner and Boyd (2005), cognitive demands were cited most often and auditory demands were not cited at all.

Table 12. Reasons for Ratings in the On-Road Evaluation

Reason	% Subjects Citing at Least Once	Demand
Attention taken from driving task	52	Cognitive
Interferes with visual monitoring	36	Visual
Physical requirements	23	Psychomotor
Length of task	21	
Task characteristics (complexity, error, type of task)	11	Maybe cognitive
Other	8	
Demands of reading	3	Visual/cognitive

In addition to the focus on secondary task demands, Lerner et al. also explored primary tasks demands. Table 13 shows the mean risk ratings, by driving situation, from the on-road evaluation. Merging has the highest rating.

Table 13. Mean Driving Risk Ratings for All Subjects for Various Situations
Source: Lerner and Boyd (2005)

Driving Situation	Mean Risk Rating
Merging from one freeway to another	6.62
Getting onto a freeway from an arterial road	6.22
Turning left across oncoming traffic from an arterial road	5.93
Driving on a two-lane curvy road	5.66
Exiting a freeway onto an arterial road	5.41
Driving on a major freeway	5.02
Exiting a parking lot & turning right onto arterial road	4.75
Driving on an arterial road	4.13
Driving on a local/residential road	3.51
Stopped at a red light on an arterial road	2.60

Table 14 lists how often subjects said the risk was great and associated reasons. Notice that reasons related to traffic were most common, followed by road geometry and visibility. Illumination and road surface condition were not mentioned. This may be because dry conditions and daylight were assumed.

Table 14. Reasons Given by Subjects for High Risk Ratings

Reason	% Subjects Citing at Least Once	Demand
Merging/interacting with other traffic	32	Traffic
High speed of traffic	26	Traffic
Behavior of other drivers (improper, risky, hard)	24	Traffic
Difficulty of visual and temporal judgments	20	
Maneuver requires concentration, awareness	20	
Opposing traffic	19	Traffic
Limited sight distance	13	Visibility
Demands of vehicle control, staying on path	13	Road geometry
Volume of traffic	11	Traffic
Unfamiliarity	10	
Limited maneuver time	5	
Presence of children, pedestrians	4	Traffic
Slow or stopped vehicles	2	Traffic
Unfamiliarity	2	
Presence of roadside hazards (e.g., trees)	2	

In the take-home rating booklet, the willingness-to-engage ratings for driving situations were slightly greater (by less than half of a point on the 10-point scale) than those collected on-road, and there were some interactions of evaluation method with the situation. Rain decreased the willingness to do tasks by about 0.6 on average, but this trend was less pronounced for tasks drivers were initially unlikely to do (ratings below 4), probably because of floor effects. Construction led to a slightly larger drop, about 0.7. Interestingly, peers in the vehicle, children in the vehicle, night conditions, congestion, and urgency had almost no effect on ratings. For the purposes of the SAVE-IT project, an equation to estimate driving situation risk would have been particularly useful. The authors have some concerns about these differences given the differences in ratings in the on-road versus booklet situations, the absence of ratings for the more difficult on-road conditions, and the subjects' prior experience with many of the tasks evaluated. (Of course, providing that experience would have increased the cost and duration of the study considerably.)

Issues Examined

Ideally, to predict workload and risk, one would have information on the demands of the primary driving task (traction, visibility and lighting, traffic density, road geometry), information on the secondary task (task duration and driver exposure, task intensity and resource demands, driver experience with tasks, task engagement, and possibly interruptability) and information about the driver. When this phase of the project was initiated, only the Nygren and Hulse studies were completed, so there were differing views of the relative importance of various factors in determining workload. If anything, the more recent work of Lerner adds to the disagreement. More significantly, none of the prior work provided comprehensive public data on the relative workload for a wide range of driving situations, which is necessary to develop a workload manager. That gap served as the primary motivation for this experiment.

Thus, given (1) the lack of published data regarding workload estimates for a wide range of driving conditions and (2) the availability of data from only 1 study (actually conducted in parallel with the project) on the willingness to engage in tasks, this experiment was conducted. To accomplish the project goal of building a workload manager, data was needed to determine the relationship among road types, traffic, other descriptors of the driving situation, and driving workload. The basic idea was that ratings of workload would be informative, and they could be readily obtained for the most common driving situations.

More specifically, the following questions were addressed:

1. How repeatable are the workload ratings within and between drivers?

2. How do workload ratings vary overall?

3. What is the relationship between workload ratings of driving situations and (1) road type (e.g., urban), (2) road geometry, (3) lane driven, (4) traffic volume (as measured by LOS), (5) driver age, and (6) driver sex?

4. How can workload ratings be estimated using the driving performance statistics developed from the ACAS FOT data set?

5. How do ratings of workload vary with the relative position of vehicles ahead (traffic) on expressways?

6. What is the relative contribution of traffic, road geometry, visibility and lighting, and traction to ratings of workload?

7. How does the probability of a driver being willing to do a secondary task while driving (tune a radio, dial a phone, enter a destination) vary with the overall ratings of workload and (b) road characteristics, traffic, and driver characteristics as in question 3?

TEST ACTIVITIES AND THEIR SEQUENCE

Overview

This study focuses on workload ratings given by drivers, and their perceived level of safety for 3 in-vehicle tasks. Subjects sat in a driving simulator and watched video clips of several different driving scenes. They provided a workload rating for each scene and noted if they would perform each of the 3 in-vehicle tasks while driving the scenes shown. After rating all of the clips, subjects provided ratings for a wider range of situations than was shown in the clips and overall ratings of the relative contribution of road geometry, traffic, and other factors to workload.

Clips from the existing ACAS dataset (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, Winkler, 2005) were used. Associated with the clips of the road scene were 400 engineering variables (speed, number of vehicles ahead, etc.), samples of face clips (showing where the driver was looking), and other information that might be useful in linking the driving situation to ratings of workload.

The disadvantage of these clips was that they were recorded at 1 Hz, making it difficult to readily determine the progress of events (such as a lane change or lead vehicle decelerating). In addition, the clips were recorded in black and white. In night scenes, oncoming headlights could not be distinguished from taillights of vehicles ahead. Since night scenes could not be reliably judged, they were not considered.

In planning this study, there was discussion of collecting an entirely new set of forward scene clips using an instrumented car sampled at a higher rate, in color, and with a wider field of view. Another option was to program the desired scenarios in the driving simulator. However, the effort to collect new data using either method was well beyond the cost and schedule of this project. Furthermore, there were so many unanswered questions about how to collect new data that focusing on the available data made sense.

Sequence of Test Activities

A summary of the sequence of activities appears in Table 15 and the complete instructions appear in Appendix A. The experiment consisted of a sequence of activities that took approximately 2-1/2 hours per subject. Upon arrival, participants were given consent and biographical forms to complete (Appendix B). The biographical form concerned their experience with driving as well as with the 3 in-vehicle tasks. Subjects were also given a vision test to verify that they had at least 20/40 eyesight, the common minimum requirement to drive in the U.S.

Participants then sat in the driver's seat of the UMTRI driving simulator and were instructed in the performance of the 3 in-vehicle tasks. They performed the tasks for about 2-3 trials until they no longer needed help. After driving a loop to become accustomed to the simulator, subjects completed 2 practice trials of each in-vehicle task while driving the simulator.

Table 15. Experiment Sequence Summary

Major Activity	Action	Estimated Duration (minutes)
Introduction	Greet Subject	2
	Fill out Consent Form	5
	Fill out Biographical Form	8
	Vision Test	2
	Seat Subject	2
	Give Subject Instructions	5
Practice	Practice Tasks	10
	Practice Driving	5
	Practice Tasks while Driving	8
Test Block 1	Rate Half of Clips	30
Break	Break	5
Test Block 2	Rate Second Half of Clips	30
Post-test	Fill out Post-Test Ratings	20
	Questions/Comments	2
	Pay Subject $70	2
	Total	**136**

Subsequently, 2 anchor video clips were looped and shown on the left side of the front screen while 3 clips whose workload was to be rated (for practice) were shown in the center of the screen. Using those anchors, subjects rated the workload of a large number of triples of test clips, grouped into 2 blocks.

Finally, subjects completed a post-test form, rating the workload of a large number of situations, and, upon completion, were paid.

Test Participants

The 24 subjects, 8 each from 3 age groups (18-30, 35-55, and 65+), were equally balanced for sex. The subjects either responded to a classified advertisement placed in The Ann Arbor News regarding a driving study, or were from a list of past participants.

The subjects, all native English speakers, were representative of the U.S. driving population in several ways. Although the study was conducted at a university, there was a deliberate effort not to recruit college students, and, in fact, only 3 took part in the study. The mean mileage reported by U.S. drivers is about 13,000 miles per year (www.fhwa.dot.gov/ohim/hs97/nptsdata.htm), and participants reported driving 2,000 to 40,000 miles per year (mean of 13,000). Seven subjects reported having more than 1 moving violation in the past 5 years, and 11 subjects had been in 1 crash within the past 5 years. Subjects were very slightly more aggressive/risk taking than normal, with 9 subjects preferring the left lane, 10 subjects the middle lane, and 5 subjects the right lane on an expressway with 3 lanes in each direction.

All but 1 subject reported being familiar with touch screens, and all of the subjects stated they were familiar with tuning the radio and setting preset stations on their car radios. Of the 24 subjects, 20 owned cell phones. None of the subjects owned a vehicle with a navigation system, hence the need for practice with the destination entry task.

More than 80 percent of the subjects wore contacts or glasses for reading or driving. Each subject's near and far visual acuity was tested with the following results: far visual acuity averaged 20/25, with a range of 20/13 to 20/50 (20/70 is minimum acuity required by State of Michigan for daytime driving). Near visual acuity averaged 20/27, with a range of 20/13 to 20/70.

Test Equipment

The experiment took place in the third-generation UMTRI driving simulator (www.umich.edu/~driving/simulator.html). The simulator consisted of a full-size cab, computers, video projectors, cameras, audio equipment, and other items (Figure 6). The simulator has a forward field of view of 120 degrees (3 40-degree channels) and a rear field of view of 40 degrees (1 channel). The forward screen was approximately 16-17 feet (4.9-5.2 m) from the driver's eyes (depending on seat adjustments), close to the 20-foot (6 m) distance often approximating optical infinity in accommodation studies. For the driving practice portion of the experiment, all 4 screens were used. For the workload rating segment, only the front and left screens were used.

Figure 6. Simulator Screen, Cab, and Control Room

The vehicle mockup consisted of the A-to-B pillar section of a 1985 Chrysler Laser with a custom-made hood and back end. Mounted in the mockup were a torque motor connected to the steering wheel (to provide steering feedback), an LCD projector under the hood (to show the speedometer/tachometer cluster), a touch-screen monitor in the center console (for in-vehicle tasks), a 10-speaker sound system (for auditory warnings), a sub-bass sound system (to provide vertical vibration), and a 5-speaker surround system (to provide simulated background road noise). The 10-speaker sound system (for in-vehicle tasks) was from a 2002 Nissan Altima and was installed in the A-pillars and lower door panel, and behind each of the two front seats. The stock amplifier (from the 2002 Nissan Altima) drove the speakers. The main simulator hardware and software was a DriveSafety simulator running version 1.6.2 software. The GeForce3 display cards did not support anti-aliasing.

The simulator was controlled from an enclosure on the driver's side of the vehicle and behind it. The enclosure contained a large table with multiple quad-split video monitors to show the output of every camera and computer, a keyboard and LCD for the driving simulator computers, and a second keyboard and LCD to control the instrument panel and touch-screen software. Also in the enclosure was a 19-inch rack containing all of the audio and video equipment (audio mixers, video patch panel and switchers, distribution amplifiers, VCR, quad splitter, etc.) and 2 separate racks for the instrument panel and touch-screen computers, the simulator host computers, and the 4 simulator image generators. The instrument panel and center console computers ran under the Mac OS. The user interface to the simulator ran under Windows and the simulators ran under Linux. Additional information on the simulator (e.g., a plan view of the facility with dimensions and the manufacturer and model numbers of key components) appears in Appendix C.

Video Clips Examined

Clips were presented for 3 classes of roads: expressways, rural roads, and urban roads. These classes roughly correspond to interstates and freeways, rural major and minor arterials, and urban major and minor arterial classes used in other studies in this project. Because of low traffic volumes, collectors and local roads, in general, were not considered.

For each road category, the goal was to explore three (A, C, E) levels of service (LOS), a term used by civil engineers to classify the traffic volume on a road. Shown in Table 16 are some example definitions for all LOS values (www.wsdot.wa.gov/ppsc/hsp/Survey/RegionRDP/ NCR-RDP/SR28-281-RDP/SR28-281-RDP-ExecSum.PDF). These terms are more precise than describing traffic as light, medium, or heavy, which depends on local experience. For example, heavy traffic in the upper peninsula of Michigan (sparsely populated) might be considered as moderate/medium in lower parts of the state (more densely populated) and as light traffic in Japan (densely populated). In fact, the definition of LOS is specific to the type of road being driven and is determined by the number of vehicles/lane/hour. For the data from the Highway Capacity Manual (Transportation Research Board, 2000) used to determine the LOS for each road class examined, see Appendix D.

Table 16. Level of Service Sample Definitions

Level of Service	Description
A	A condition of free flow in which there is little or no restriction on speed or maneuverability caused by the presence of other vehicles.
B	A condition of stable flow in which operating speed is beginning to be restricted by other traffic.
C	A condition of stable flow in which the volume and density levels are beginning to restrict drivers in their freedom to select speed, change lanes, or pass.
D	A condition approaching unstable flow in which tolerable average operating speeds are maintained but are subject to sudden variations.
E	A condition of unstable flow in which operating speeds are lower with some momentary stoppages. The upper limit of this LOS is the capacity of the facility.
F	A condition of forced flow in which speed and rate of flow are low with frequent stoppages occurring for short or long periods of time; with density continuing to increase causing the highway to act as a storage area.

Table 17 shows the urban situations examined, combinations of the most common factors: (1) traffic volume as assessed by LOS and (2) the presence/absence of traffic signals. Urban roads were defined as roads with 4 lanes, commercial entrance and exit points, and occasional intersections with traffic signals. The number in the cell (2) indicates 2 instances (different roads) seen by each subject. Each of those 2 instances was seen twice by each subject to determine the consistency of workload ratings.

Table 17. Urban Situations Examined

Situation	4 Lanes		
	A	C	E
Straight	2	2	2
Intersection 4 lanes, traffic signal (green for subject)	2	2	2

Figure 7 shows a typical frame from an urban clip. Notice that the field of view is sufficiently wide to capture the key information the driver would use in making decisions about workload.

Figure 7. Sample Frame from an Urban Road Video Clip

Originally, examining various turn-lane combinations was also considered, but there were few of them in the dataset, and over the 30 s window sampled, the associated workload was not stable. Also considered were clips where all intersections were consistently of 1 type (e.g., all 2 lanes or all 4 lanes). Such clips were difficult to find in the set, and, of course, more lanes at intersections usually meant more traffic on the main road, which was a confounding situation. Accordingly, intersection variations were not examined.

Urban areas tend to develop on flat land because that is often the least costly land to develop. Curves often occur as a means to avoid natural features such as mountains and valleys, which are less common in urban areas. Given the relatively low frequency of curves in urban roads in southeast Michigan, curves on urban roads were not examined.

Table 18 shows the situations explored for rural roads. Rural roads were defined as roads with 2 lanes and very few (less than 1) access points. Only 2-lane roads were considered because once they become 4 lanes (and are undivided), at least in southeastern Michigan, the road often becomes urban. For rural roads, there are few traffic signals, but curves are more common and were therefore considered. Figure 8 shows a sample frame from a rural road video clip.

Table 18. Rural/Open Road Situations Examined

Situation	**2-Lane Road Driven**		
	A	C	E
Straight	2	2	2
Curved	2	2	2

Figure 8. Sample Frame from a Rural Road Video Clip

Table 19 shows the situations examined for expressways. In contrast to rural roads, the curves on expressways are gentle and should have a small effect on workload, so curves were not considered. Expressways were 6 lanes (3 in each direction), with no access points (except for during a merging situation clip). The effect of lane driven was unknown and was explored.

Table 19. Expressways Situations Examined

Situation	6 Lane Road Driven								
	Left **Lane**			**Center Lane**			**Right Lane**		
	A	C	E	A	C	E	A	C	E
Straight	2	2	2	2	2	2	2	2	2
Merging								2	2

Also, to limit the number of clips to be rated, only 6-lane expressways (3 lanes per direction) were considered. In many ways, driving the left lane of a 4-lane expressway resembles driving the left lane in a 6-lane expressway. The same is true for the right lanes in both cases, though the 6-lane expressway has the added demand of traffic 2 lanes away for the outer lanes. Figure 9 shows a sample frame from an expressway video clip.

Figure 9. Sample Frame from an Expressway Video Clip

For expressways, the major demand is often from merging traffic, and then primarily in the right lane only. However, merging traffic for LOS A was not possible as by definition any merging traffic that would affect the right lane is at least LOS B. Thus, only 11 (not 12) combinations needed to be considered.

The probability of a crash increases significantly in work zones (Sullivan, Winkler, and Hagan, 2005) and so should the associated workload as the driver deals with lane shifts, lane drops, and construction equipment. However, there were too few instances of work zones in the ACAS dataset, so their full consideration was left to future research (though they were examined in the post-test ratings described later).

Thus, although there are a large number of traffic combinations that could be explored by road type, road geometry, number of lanes, and traffic combinations, the 23 examined here capture many of the common situations in which workload is an issue.

Clips were presented in an order counterbalanced for age and age effects. See Appendix E for the complete clip sequence.

Test Trial Ratings of Workload

When workload was to be rated, usually 5 clips appeared in front of subjects (Figure 10). All clips were 15 s long, 30 s of real video recorded at 1 Hz but played back at 2 Hz to provide a sense of continuity. (Clips played at the next higher speed, 4 Hz, were cartoonish, which was thought to degrade the credibility of the study.) See Appendix F for the additional information on playback speed issues.

Figure 10. Perspective View of Left and Center Screens during Workload Rating

On the left screen were 2 anchor clips of relatively low and high workload (Figure 11). These clips were looped to play continuously. The lowest workload (LOS A) was assigned a value of 2 and shown on the top portion of the left screen. That clip was of a fairly empty expressway (3 lanes, straight, 1 vehicle about 200 m ahead) with the subject in the right lane. The highest workload anchor clip (LOS E) was assigned a value of 6 and shown on the bottom portion of the screen. That clip showed a 4-lane expressway with the driver in the left lane, passing traffic to the right, and 6 cars visible ahead (in all lanes) within approximately 200 meters. These anchor clips were selected because (1) they represented reasonable but not extreme ends of the range of workload, (2) the workload was reasonably stable in the clips, (3) they were free from artifacts (e.g., a person driving erratically), and (4) the anchor roads resembled the roads in the test clips.

Figure 11. Screen Showing Anchor Clips

The center channel provided usually 3 but sometimes 2 clips with LOS values of A and C, with a clip in the E range when 3 clips were provided (Figure 12). So ratings of workload would be consistent, each triple (or sometimes pair) of clips showed the same driving situation (e.g., left lane of a 4-lane urban road). However, in all cases, the pair or triple of clips were always ordered with the lowest workload clips at the top and the highest workload clips at the bottom (e.g., LOS A on top, LOS C in the middle, and LOS E on the bottom), an order consistent with the anchor clips. To avoid confounding, an effort was made for each triple/pair to represent roads that were geometrically similar (same lane width, same shoulder width, same curvature, etc.), and sometimes it was the same road. That was not possible in all cases, given the content of the database and the schedule.

Figure 12. Center Screen Showing Test Clips

Subjects were told:

"You will be rating the demand of driving on expressways, rural roads, and urban streets as shown on video clips. Please rate the demand of actually driving the situation shown in the clip, not the demand of just watching the video. Also, state how safe you feel it is to (1) manually tune the radio, (2) manually dial a phone number, and (3) enter a navigation destination while in the situation shown.

The rating scale is from 1 to 10. To help rate the driving workload, reference clips will be continually shown on the left screen and you can look at them whenever you want. These clips have workloads of 2 (on the top) and 6 (on the bottom), where larger values mean more workload."

Subjects were also asked whether they would feel safe using each of the in-vehicle tasks in the current driving situation. Thus, drivers gave 4 responses for each clip: one workload rating and 3 yes or no answers corresponding to tuning a radio, dialing a phone, and entering a street address. These 3 tasks had been examined on the road in a prior SAVE-IT study (Zylstra, Tsimhoni, Green, and Mayer, 2004) and spanned a reasonable range of task times. Prior to data collection, subjects rated the workload for 1 triple of practice clips to verify they understood the rating task.

In-Vehicle Tasks

Secondary Task Menu

All 3 tasks began by selecting a task category from a hierarchical menu. To begin a task, the subject pressed the start button, which brought up 3 menu headings: radio, phone, and navigation (Figure 13). Pressing each of the main menu entries brought up a context-specific menu of 4 to 6 options. The submenu item "tuner" displayed the touch screen radio interface, "dial" displayed a phone keypad, and "address entry" displayed the navigation interface. An error tone was played for selecting an incorrect menu item. All tasks were presented during the practice sessions to make sure all subjects had a common appreciation for the demands of the 3 tasks (dialing a phone, manually tuning a radio, entering a destination) to overcome a lack of knowledge (because they had not done the task before) or biases due to particular user interfaces with which they were familiar.

Figure 13. Touch Screen with All Menu Options Displayed (6.2 x 3.6 inches)

Radio Tuning Task (Short Duration Task)

To begin, an index card displaying the decimal FM frequency (99.5) was presented to the subject atop the center stack and the subject was instructed to set preset number 1 to that station by using the up and down arrows on the right side of the radio (Figure 14), to increase or decrease the frequency by 0.2 per key press. Each station was either 2.8 Hz (14 button presses) or 4.2 Hz (21 button presses) above or below from the initially-displayed station frequency. Once the subject selected the appropriate station, they pressed the button for preset number 1 and feedback was given to indicate correct (celebratory sound) or incorrect entry (buzzer).

Figure 14. Radio from 1991 Honda Accord Station Wagon (5.8 x 1.9 inches), Presented on the Touch Screen as a jpeg Image

Phone Dialing Task (Medium Duration Task)

To begin, an index card displaying a 10-digit phone number was presented to the subject atop the center stack for the subject to enter using the keypad (Figure 15) on the touch screen. The sequence entered was shown in the blank area above the 3 function keys. Errors made by the subject could be corrected by using the Del key to go back and remove errors. Once the entire number was entered, the subject pressed the Talk key and feedback was given to indicate if the number was entered correctly (a ringing phone) or incorrectly (error tones).

Figure 15. Touch-Screen Telephone Interface Used for Dialing Task (2.6 x 3.2 inches)

Destination Entry Task (Long Duration Task)

To begin, the subject was presented with an index card with address information (city, street, number) in that order, the order in which information was to be entered. The index card was placed atop the center stack in the same location as for previous tasks. Subjects then entered the entire address using a QWERTY keyboard on a touch screen. (Figure 16). All of the addresses contained 20 total characters for road name, city name, and number, but were balanced with varying street and city name lengths. The line being entered had a white background whereas the other two lines had a gray background. After each line was complete, subjects pressed "return" to advance to the next line and the previous line became gray. Errors could be corrected using the back arrow. Pressing "return" on the third line ended the task and provided feedback as to whether the address was entered correctly (celebratory sounds) or incorrectly (buzzer sounds). (See Appendix A for additional details on all tasks.)

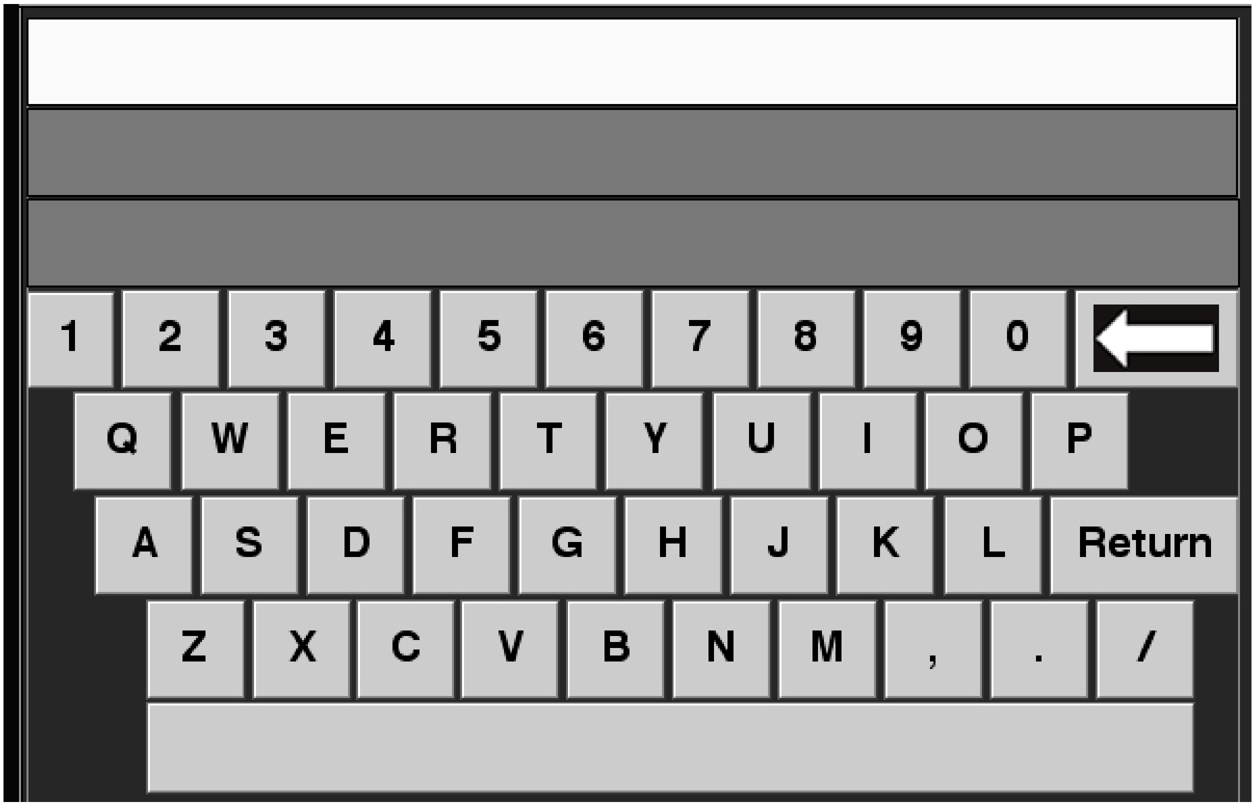


Figure 16. Touch-Screen Interface Used for the Destination Entry Task (6.1 x 3.1 inches)

Post Test Ratings

After watching all of the clips, subjects filled out a post-test form concerning the estimated workload for many situations that might be encountered while driving on urban, rural, and residential roads as well as expressways on a scale of 0 ("no demand") to 100 ("completely requires all of your capacity to just drive"). This form examined many situations that were not captured on tape, to allow rating subtly different situations (e.g., residential streets with no parked cars versus those with parked cars on 25% of the curb space). Where traffic levels might vary, multiple workload levels were examined. In addition, subjects rated how the distance to vehicles in various lanes on an expressway influenced workload, and how traffic, road geometry, visibility and illumination, and traction contributed to overall ratings of workload. See Appendix B for the post-test forms.

RESULTS

Note: In the instructions, the word “demand” is used to describe the rating requested of the subject. Here, the term workload is used. There could be differences in what the 2 terms mean, but for convenience and consistency with the literature, the term workload is used in the results and conclusions.

How Did the Test Trial Workload Ratings (of Clips) Vary Overall?

As a reminder, subjects rated the workload of clips (usually triples) given anchor clips of 2 (low) and 6 (high). Figure 17 shows the overall distribution of ratings. Notice that the clips are widely distributed in the ratings, which was an experimental goal.

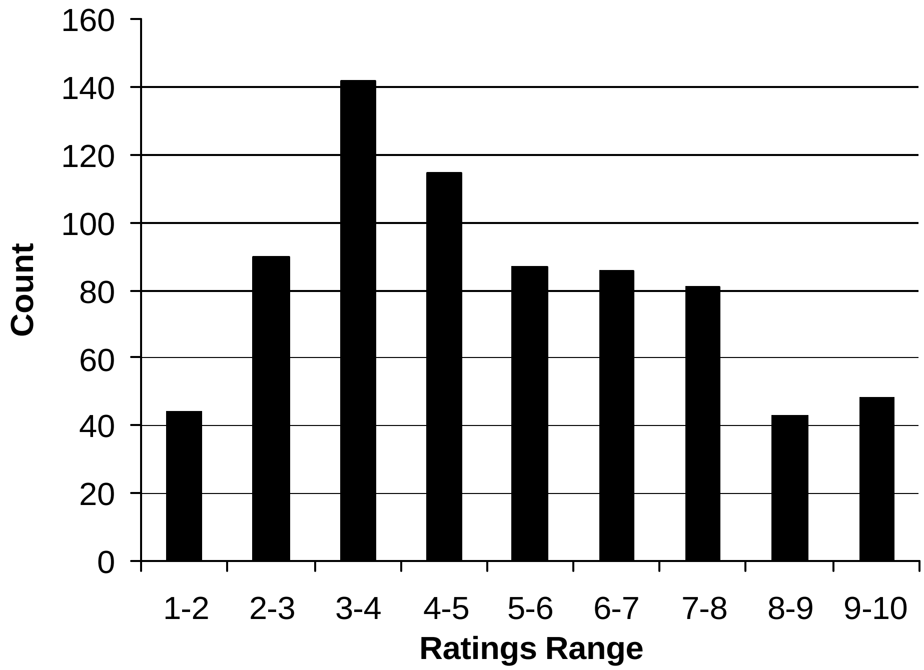


Figure 17. Distribution of Workload Ratings of Clips

Figure 18 shows the workload ratings split by LOS. Keep in mind that LOS is not an exact value but a range, and that is reflected in the spread of the ratings data. However, the ratings were consistent in that values for LOS A were usually less than those for LOS C, which in turn were usually less than those for LOS A (means of 2.8, 4.5, and 6.0, respectively, as shown later in Table 23). Also, LOS’s were spread across different road types, so a range of values makes sense for each LOS.

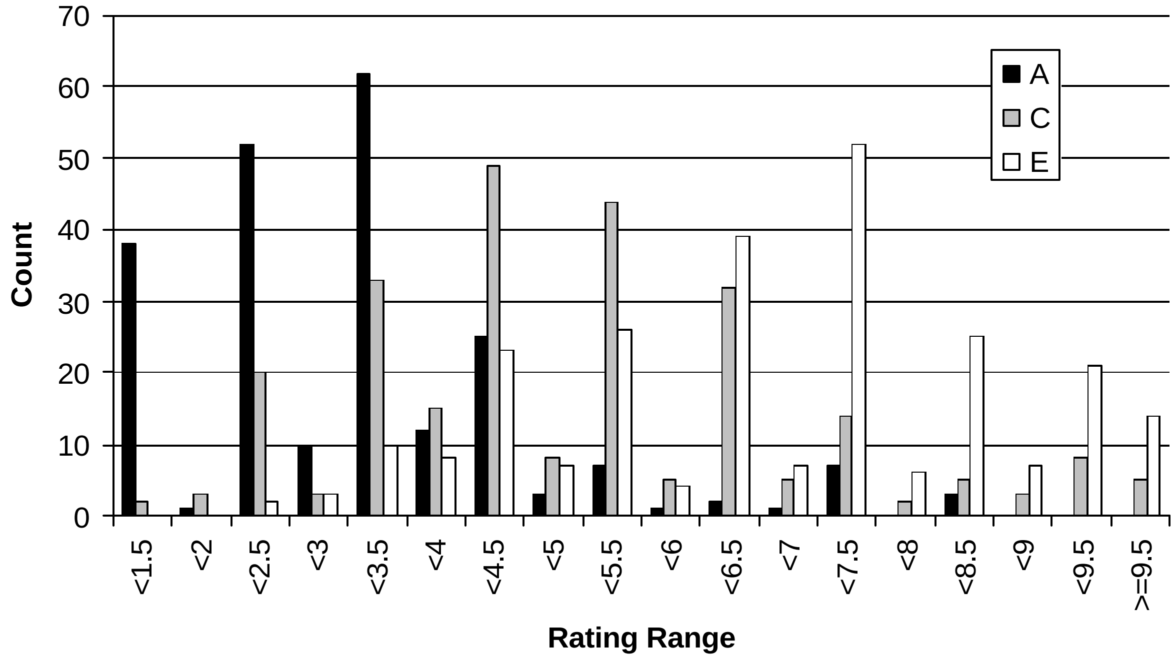


Figure 18. Workload Ratings Distribution by LOS

How Repeatable Were the Workload Ratings (of Clips) within and between Drivers?

Subjects were quite consistent between repetitions of clips. Those differences were determined 2 ways. Each subject saw each clip twice. The means of the absolute value of the differences for rural roads are shown in Table 20. The difference was typically 0.2 to 0.3, quite small considering ratings were usually given to the nearest 0.5. Subjects were not instructed to round to the nearest half point, but subjects tended to overwhelmingly round their answers to the nearest half point. In addition, each subject saw 2 clips representing each situation (e.g., a 2-lane straight rural road with LOS A). As one would expect, the mean differences were larger, 0.3 to 0.9, though relative to the measurement accuracy, still small. There was no apparent pattern to those differences.

Table 20. Rural Rating Consistency

Road Type	Level of Service	Mean Rating	Mean Difference between Repetitions	Mean Difference within Clip Type
Rural Straight	A,C,E	4.0	0.2	0.5
	A	2.4	0.2	0.4
	C	3.9	0.2	0.8
	E	5.7	0.3	0.3
Rural Curved	A,C,E	4.1	0.3	0.5
	A	3.0	0.3	0.9
	C	3.9	0.3	0.3
	E	5.4	0.3	0.3

For urban roads (Table 21) the mean differences between repetitions were larger (0.2 to 0.5) and the mean differences with clip type were considerably larger (0.8 to 2.1). In part, this is due to the larger mean ratings (which are generally accompanied by greater variability), but that still does not account for all of the increase.

Table 21. Urban Streets Rating Consistency

Road Type	Level of Service	Mean Rating	Mean Difference between Repetitions	Mean Difference within Clip Type
Urban no intersection	A,C,E	4.7	0.4	0.9
	A	2.9	0.2	0.8
	C	4.9	0.4	1.2
	E	6.5	0.5	0.8
Urban with intersection (with light)	A,C,E	4.7	0.4	1.5
	A	2.8	0.4	1.2
	C	5.7	0.4	1.3
	E	5.8	0.5	2.1

For expressways (Table 22), the differences between repetitions were less on average than those for urban streets (0.2 to 0.6) and the mean difference within clip type was clearly less (0.3 to 1.8). Interestingly, the mean workloads for expressways were close to those for urban streets.

Table 22. Expressway Roads Rating Consistency

Lane	Level of Service	Mean Rating	Mean Difference between Repetitions	Mean Difference within Clip Type
Left	A,C,E	4.7	0.5	0.8
	A	3.2	0.6	0.8
	C	4.5	0.6	1.3
	E	6.6	0.5	0.3
Middle	A,C,E	4.3	0.4	0.7
	A	2.9	0.4	1.0
	C	4.5	0.3	0.7
	E	5.4	0.3	0.5
Right	A,C,E	4.0	0.4	0.2
	A	2.6	0.2	0.6
	C	3.6	0.4	0.0
	E	5.8	0.5	0.0
Right lane w/ merging traffic	C,E	5.7	0.3	1.2
	C	5.0	0.3	0.6
	E	6.4	0.4	1.8

How Did the Rated Workload (of Clips) Vary with the Road Type, Geometry, Lane Driven, and Traffic?

Though there may be a more elegant manner to examine the factors that affect clip ratings, each of 3 road types was examined in a separate ANOVA for ease of computation. All 3 analyses shared the same subjects effects—age group (young, middle, old), sex (men, women), age * sex, and subjects nested within age, as well as traffic (LOS, usually A, C, E) but not always 3 levels), and age interacting with other factors. However, other differences were specific to each road type (road geometry of rural roads, intersection presence for urban streets, and lane and merging traffic for expressways). In all 3 ANOVAs, the main effects were examined as well as all interactions with subject age and LOS, which are variables with large effects.

One of the consequences of those separate analyses is that there were no overall statistics examining workload. As shown in Table 23, the workload of rural roads was slightly less than other roads for LOS C and E, and the workload ratings for expressways were in between. Keep in mind that clips were selected for each road type to meet particular conditions and were not a random selection of that LOS for that type of road. This could be the source of the differences. Furthermore, the relative real-world exposure of drivers to each LOS for each road type is not available.

Table 23. Mean Workload Rating by Road Type and LOS

Level of Service	Mean Workload Rating				
	Rural	Urban	Expressway		Mean
			Not Merging	Merging	
A	2.7	2.8	2.9	-	2.8
C	3.9	5.3	4.2	5.0	4.5
E	5.5	6.1	5.9	6.4	6.0
Mean	4.0	4.7	4.3	5.7	

Note: The mean workload for each LOS was computed based on how often each LOS occurred in the raw data, not the mean LOS for each road type. Had that not been done, then the expressway merging results would have dominated the data disproportionately.

Table 24 shows the results from the 3 ANOVAs, with the independent variables common to multiple analyses shown in the same row. As a reminder, urban roads were defined as roads with 4 lanes, commercial entrances and exits, and occasional intersections with traffic signals. Rural roads were defined as roads with 2 lanes and very few (less than 1) access points. Expressways were 6 lanes (3 in each direction), with no access points other than merging ramps (and exits).

Table 24. Summary of ANOVAs for Workload Ratings of Clips

Rural		**Urban**		**Expressway**	
Factor	**P**	**Factor**	**P**	**Factor**	**P**
LOS	<.0001	LOS	<.0001	LOS	<.0001
Road Geometry (straight vs. curved)	0.2564				
		Intersection	0.9343		
				Lane	<.0001
Age Group	<.0001	Age Group	0.0713	Age Group	<.0001
ßSex	0.5315	Sex	0.0394	Sex	0.0947
Subject [Age,Sex]	<.0001	Subject [Age, Sex]	<.0001	Subject [Age, Sex]	<.0001
Age * Sex	<.0001	Age * Sex	<.0001	Age * Sex	0.0021
LOS * Age	0.0256	LOS *Age	0.0019	LOS * Age	0.0001
		LOS * Intersection	<.0001		
		Intersection * Age	0.9256		
Road Geometry * LOS	0.0006				
Road Geometry * Age	0.5885				
				Lane * LOS	<.0001
				Lane * Age	0.6832

Note that in all 3 cases, the LOS was highly significant. As was noted previously, the mean workload was 2.8 for LOS A, 4.5 for LOS C, and 6.0 for LOS E. Interestingly, road characteristics were often not a significant factor in workload. For rural roads, road geometry (straight vs. curves) was not significantly different, though curved sections had a very slightly higher workload (4.1 vs. 4.0). Most of the curves were very gentle. However, as shown in the ANOVA, there was a significant interaction between LOS and curvature, with largest difference (curved greater than straight) noted for LOS A (Figure 19). It could be that for LOS A, the road ahead is visible and curves are easily seen. For greater LOS levels, the demand of traffic is such that the driver focuses on traffic and not on road geometry.

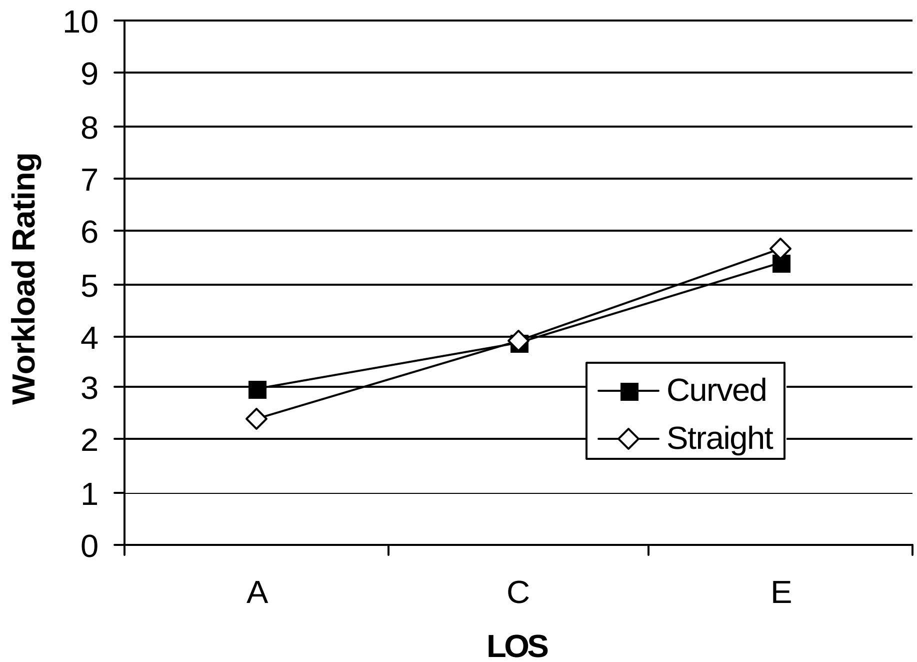


Figure 19. Mean Workload Rating vs. LOS and Road Curvature on Rural Roads

For urban streets, those with intersections and traffic lights had no greater workload than those without intersections (both 4.7). This could be because the clips were black and white (and not high resolution), so green, yellow, and red traffic signals were difficult to see. (See Figure 20.) Furthermore, the limited camera field (approximately 8 degrees vertically, and 40 degrees horizontally) meant that traffic signs were not in view close to an intersection as was crossing traffic.

Figure 20. Sample Frame from an Intersection Video Clip

Interestingly, there was a significant LOS * intersection interaction (Figure 21), with streets with intersections having a greater workload for LOS C but less for LOS E. In the absence of any other explanation, this could be just random variation. Keep in mind that within LOS categories, there is some variation of traffic levels.

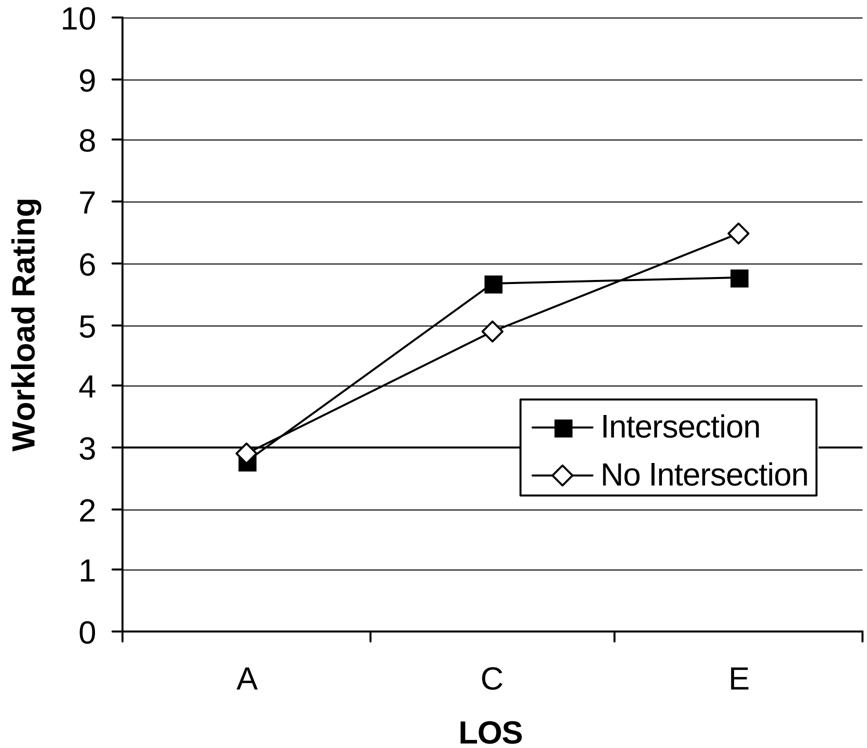


Figure 21. Mean Workload vs. LOS and Intersection on Urban Roads

For expressways, the lane had a significant effect on workload, at about 4.0 for the left, 4.2 for the middle, and 4.8 for the right, excluding merging scenarios. There also was a significant lane * LOS interaction (Figure 22), primarily because of somewhat low ratings for the workload of middle lanes for LOS E. The best explanation is that it is a statistical artifact. The presence of merging traffic also increased workload significantly, on average by about 1.0 above the workload for the right lane alone.

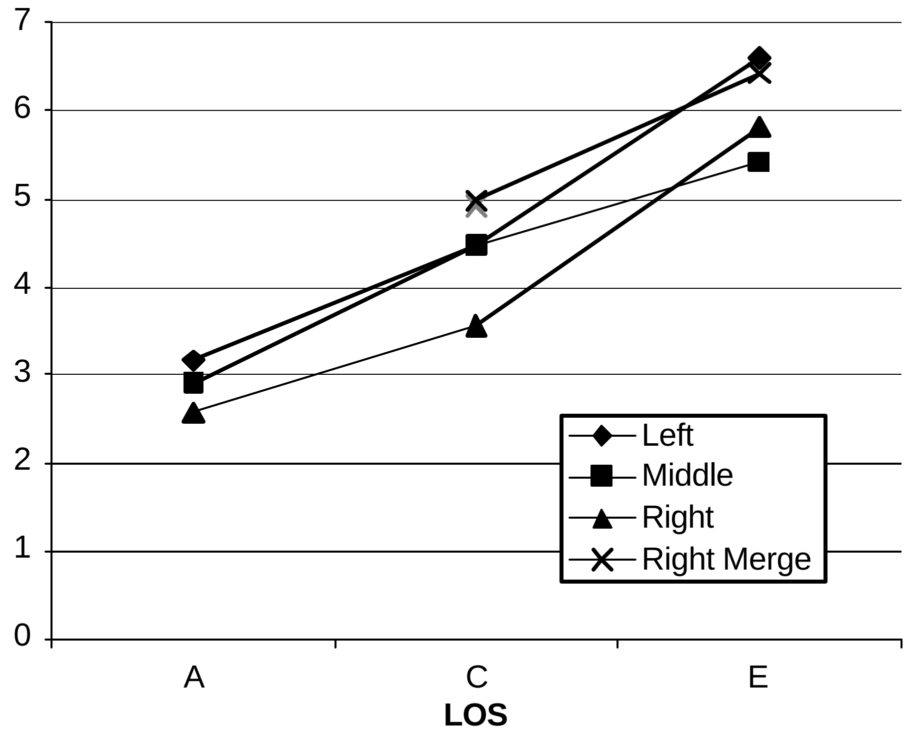


Figure 22. Mean Workload vs. LOS and Lane on Expressways

How Did the Rated Workload (of Clips) Vary with Driver Age and Sex?

As shown in Table 25, nontraffic factors contributed to workload in a quite complex manner. For every road type, there were significant age differences, with the overall middle age ratings being greater than those for younger and older subjects. However, the spread of the means was somewhat different for expressways, reflecting an age * road type interaction. Again, readers should keep in mind that (1) examples of roads in each of the 3 categories were selected to include particular feature combinations of that road type (e.g., left lane of a 3-lane expressway with LOS A), not as a random sample of all instances of that type of road, and (2) within LOS categories, traffic varied.

Table 25. Mean Workload by LOS, Age Group, and Traffic

<table>
<tr><th>Age Group</th><th>LOS</th><th colspan="2">Rural</th><th colspan="2">Urban</th><th colspan="2">Expressway</th><th>LOS Mean</th><th>Age Mean</th></tr>
<tr><td rowspan="3">Young</td><td>A</td><td>2.6</td><td rowspan="3">4.2</td><td>2.6</td><td rowspan="3">4.7</td><td>2.4</td><td rowspan="3">3.9</td><td>2.6</td><td rowspan="3">4.3</td></tr>
<tr><td>C</td><td>4.1</td><td>5.2</td><td>3.7</td><td>4.3</td></tr>
<tr><td>E</td><td>5.9</td><td>6.2</td><td>5.7</td><td>5.9</td></tr>
<tr><td rowspan="3">Middle</td><td>A</td><td>2.8</td><td rowspan="3">4.2</td><td>2.7</td><td rowspan="3">4.9</td><td>3.0</td><td rowspan="3">4.6</td><td>2.8</td><td rowspan="3">4.6</td></tr>
<tr><td>C</td><td>4.0</td><td>5.4</td><td>4.5</td><td>4.7</td></tr>
<tr><td>E</td><td>5.8</td><td>6.6</td><td>6.4</td><td>6.3</td></tr>
<tr><td rowspan="3">Older</td><td>A</td><td>2.6</td><td rowspan="3">3.7</td><td>3.1</td><td rowspan="3">4.7</td><td>3.2</td><td rowspan="3">4.4</td><td>3.0</td><td rowspan="3">4.2</td></tr>
<tr><td>C</td><td>3.6</td><td>5.2</td><td>4.3</td><td>4.3</td></tr>
<tr><td>E</td><td>5.0</td><td>5.7</td><td>5.7</td><td>5.4</td></tr>
<tr><td colspan="2">Mean</td><td></td><td>4.0</td><td></td><td>4.8</td><td></td><td>4.3</td><td></td><td>4.4</td></tr>
</table>

Rural mean is 4.0, not 2.6

Finally, there were many cases where age and LOS interacted as shown in Figures 23, 24, and 25. There was no consistent pattern to those interactions and the authors have no explanation for them.

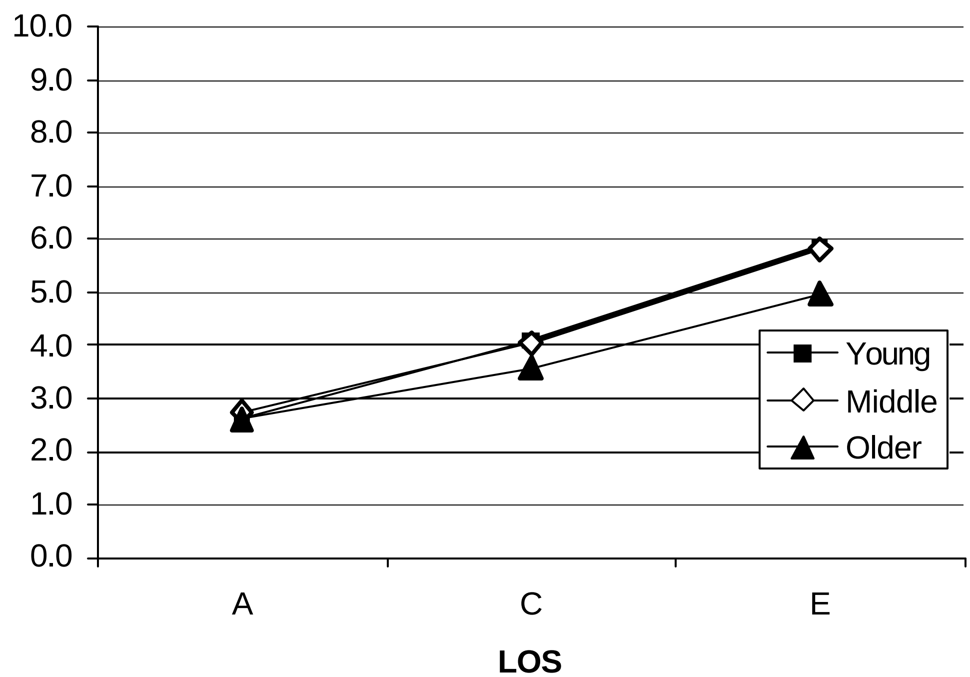


Figure 23. Mean Workload Rating vs. LOS and Age on Rural Roads

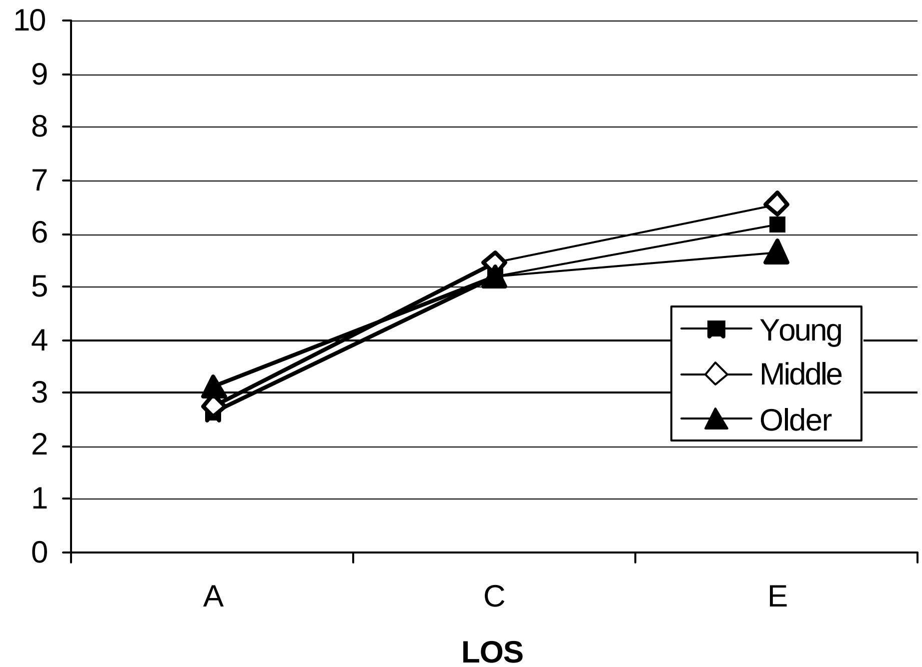


Figure 24. Mean Workload Rating vs. Age and LOS on Urban Roads

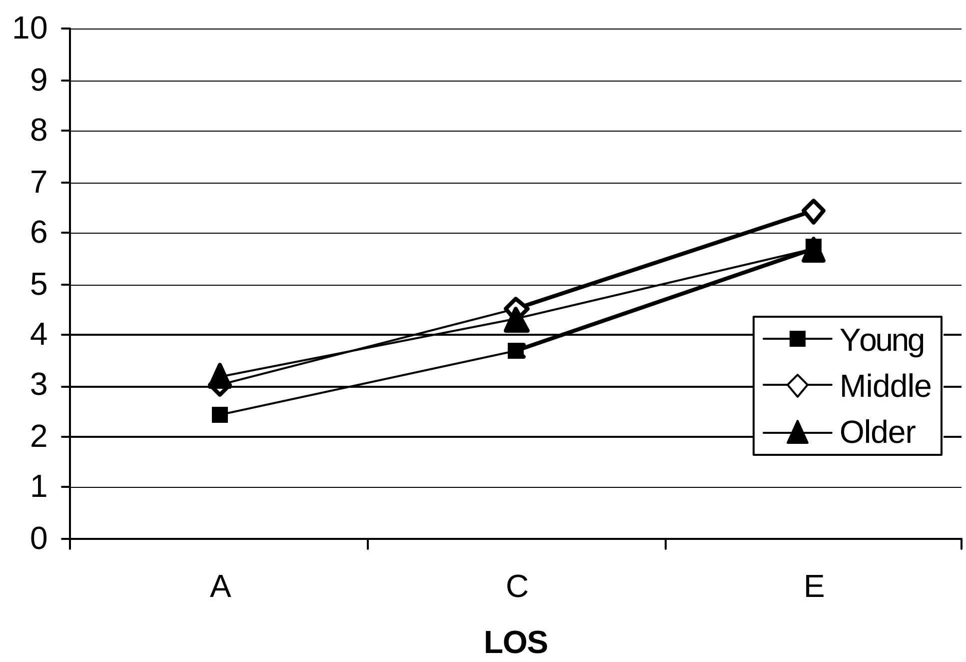


Figure 25. Mean Workload Rating vs. Age and LOS on Expressways

For all 3 road types, there were age * sex interactions, a common finding in many human factors studies (Figures 26, 27, and 28). The general trend was for women's workload ratings to be greater than men's for younger and middle-age drivers, but for men's to be greater for older subjects. Some differences were as large as 1, but many were less than 0.5. It may be a reflection of a "testosterone effect" where the younger and middle-age males want to demonstrate their driving prowess and give roads of a particular workload a lower rating, whereas the older men, often in poorer health than women of their age, give higher ratings because they have more difficulty driving.

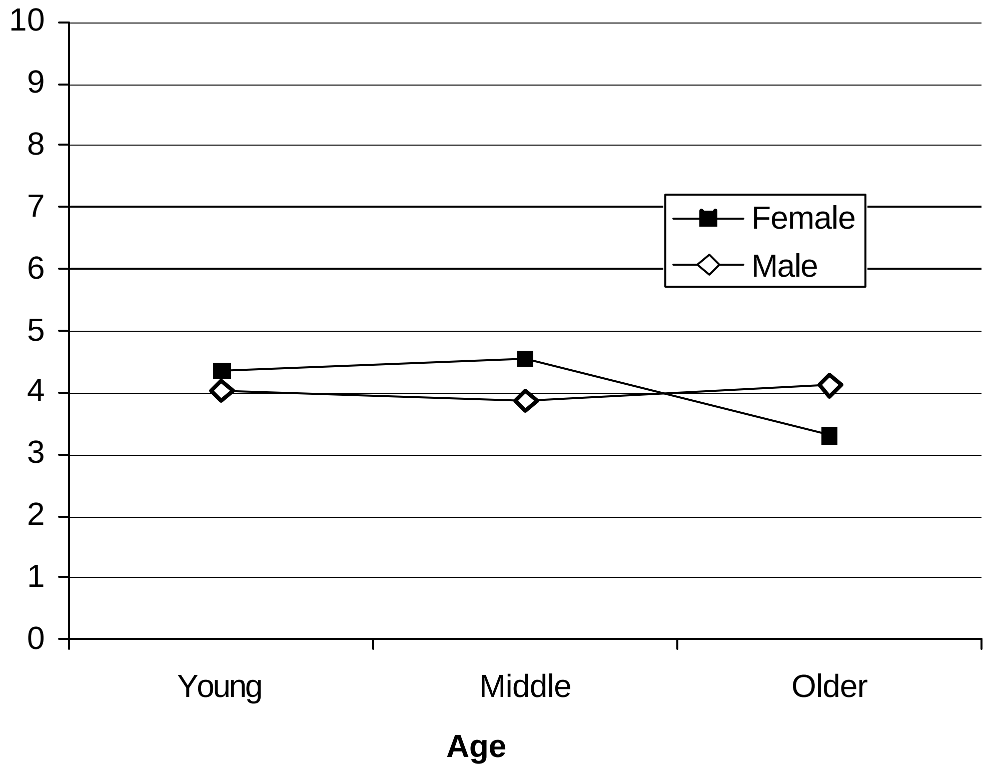


Figure 26. Mean Workload Rating vs. Age and Sex for Rural Roads

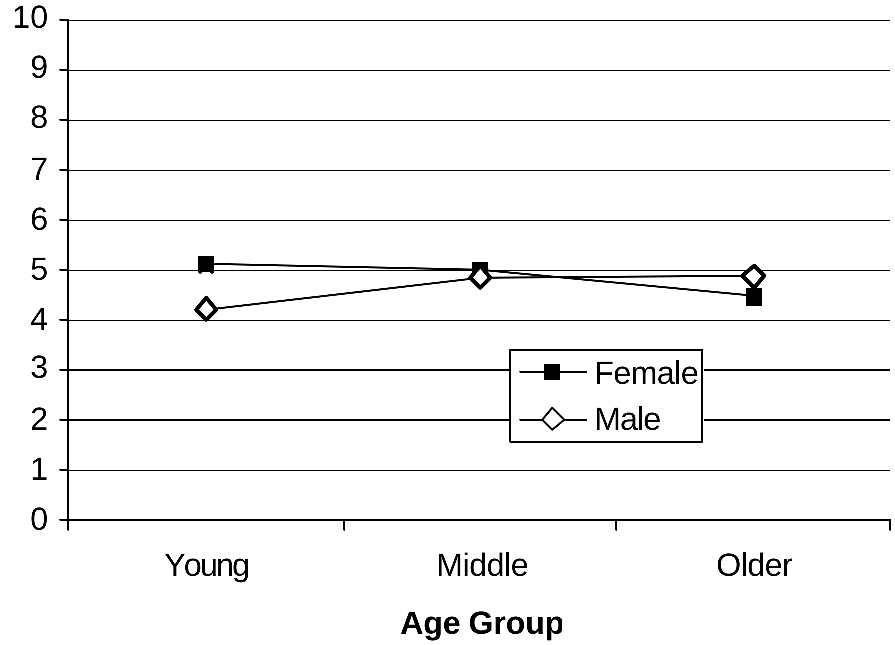


Figure 27. Mean Workload Rating vs. Age and Sex for Urban Roads

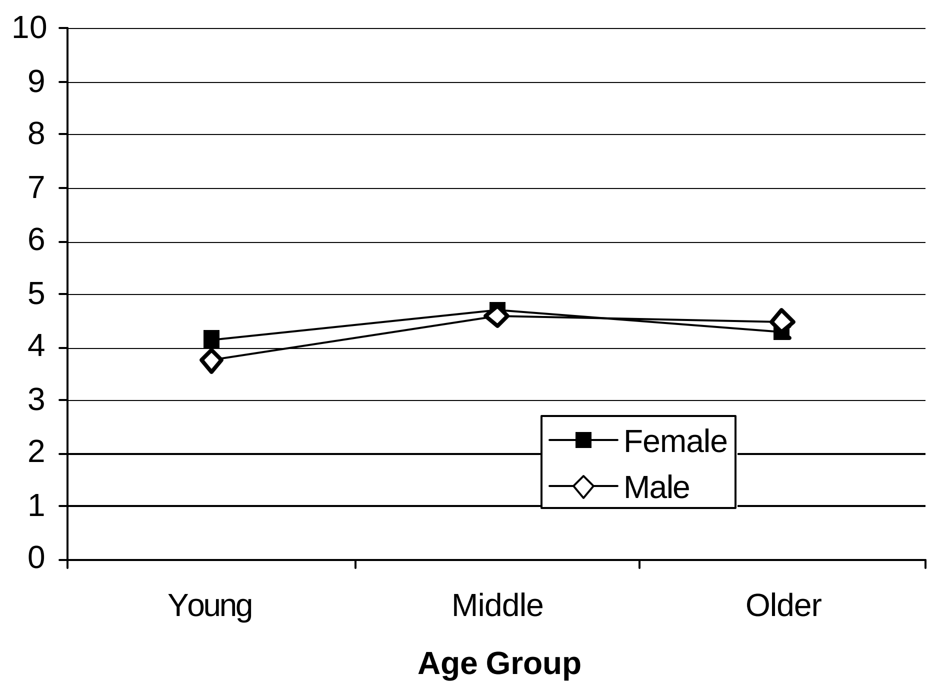


Figure 28. Mean Workload Rating vs. Age and Sex for Expressways

Thus, the general interpretation of these results is that (1) workload rating is most markedly affected by LOS, (2) the only geometry feature to affect rated workload when LOS was specified was merging, (3) in general, the ratings from lowest to highest were younger, older, and middle aged, and (4) young- and middle-age women rated workload higher than men their age, but the reverse was true for the older age group.

Using Lookup Tables, What is the Estimated Workload for Various Driving Situations as a Function of Road Geometry, Traffic, and Driver Characteristics Derived from the Clip Ratings?

How the commonly used approach for prediction, stepwise regression, should be applied in this case is not straightforward. The major difficulty is that the workload ratings do change with age, but the largest ratings are from the middle-age group. Furthermore, some of the underlying factors were unique to each road category.

Furthermore, it makes sense to think of how an expression developed here would be used in practice. Basically, there are 2 cases, (1) road and traffic data are available (curvature, presence of intersections, LOS, etc.) and (2) that data is available as well as information on the driver (age and sex). For the first case, a vehicle would need to be outfitted with a navigation system (to provide information about curvature, intersections, etc.) and an ACC for traffic estimates. A vehicle outfitted with a workload manager is likely to have both.

In terms of implementation, the simplest approach would be a lookup table as opposed to an equation. Given how the tables would be used, 2 sets of tables were developed, one set that included all statistically significant geometric factors and a second set that also included driver characteristics. So for example, for rural roads LOS, Age Group, Sex, LOS * Age, and LOS * Geometry (straight vs. curved) were significant. Since LOS * Age and LOS * Geometry were significant, then all combinations of LOS, Age, and Geometry needed to be in the table. Furthermore, since age and sex interacted, then the table needed all combinations of LOS, Age Group, Geometry, and Sex. The tables that follow are organized so the rows pertain to geometry and traffic, and the columns are for age and sex.

Tables 26 and 27 are for rural roads. Notice that for geometry alone, there was only 1 significant factor, LOS.

Table 26. Mean Workload Ratings for Rural Roads, Road and Traffic Data Only

LOS	Mean
A	2.7
C	3.9
E	5.5

Table 27. Mean Workload Ratings for Rural Roads, All Data Available

LOS	Geometry	Young		Middle		Old	
		Female	Male	Female	Male	Female	Male
A	Straight	1.9	2.3	3.0	2.0	2.4	2.7
	Curved	2.9	3.4	3.9	2.2	2.5	2.9
C	Straight	4.4	3.9	4.0	4.0	3.2	4.0
	Curved	4.0	4.1	4.3	3.8	3.0	4.1
E	Straight	6.7	5.4	6.0	5.8	4.5	5.6
	Curved	6.2	5.1	6.1	5.4	4.2	5.6

Tables 28 and 29 are for urban roads. As with rural roads, when road geometry and subject variations are considered, 36 cells are needed to capture the underlying variation.

Table 28. Mean Workload Ratings for Urban Roads, Road and Traffic Data Only

LOS	Intersection	Mean
A	No	2.9
	Yes	2.8
C	No	4.9
	Yes	5.7
E	No	6.5
	Yes	5.8

Table 29. Mean Workload Ratings for Urban Roads, All Data Available

LOS	Geometry	Young		Middle		Old	
		Female	Male	Female	Male	Female	Male
A	No	3.0	2.8	3.3	2.8	3.1	2.5
	Yes	2.8	2.4	3.5	3.0	2.8	2.1
C	No	5.0	5.2	4.7	5.1	5.1	4.2
	Yes	6.1	5.4	5.0	6.1	6.8	4.7
E	No	6.7	7.0	5.7	6.5	6.7	6.4
	Yes	6.3	6.3	4.7	5.8	6.3	5.3

Tables 30 and 31 are for expressways. Since there were no road * subject related interactions, Table 29 is actually the row means of Table 28. Note that there is no right merge case of LOS A since a merging vehicle increases the workload to LOS C.

Table 30. Mean Workload Ratings for Urban Roads, Road and Traffic Data Only

LOS	Lane Combination	Mean Rating
A	Left	3.2

	Middle	2.9
	Right	2.6
C	Left	4.5
	Middle	4.5
	Right	3.6
	Right Merge	5.0
E	Left	6.6
	Middle	5.4
	Right	5.8
	Right Merge	6.4

Table 31. Mean Workload Ratings for Urban Roads, All Data Available

LOS	Lane	Young		Middle		Old	
		Female	Male	Female	Male	Female	Male
A	Left	2.8	2.3	3.2	3.5	3.8	3.4
	Middle	2.8	2.4	2.7	3.1	3.3	3.3
	Right	2.3	2.0	2.9	2.8	3.1	2.4
C	Left	4.5	3.6	4.6	4.8	4.5	4.8
	Middle	4.1	3.7	4.8	5.0	4.5	4.8
	Right	3.1	3.3	4.0	3.8	3.9	3.3
	Right Merge	5.2	4.5	5.8	4.9	4.4	5.1
E	Left	6.6	6.1	7.3	6.9	5.7	7.1
	Middle	5.6	4.8	6.2	5.5	4.9	5.5
	Right	5.8	5.5	6.8	6.0	5.0	5.9
	Right Merge	6.6	5.8	6.9	6.7	5.9	6.6

What is the Relationship between Rated Workload (of Clips) and Statistics Summarizing Driving Performance Developed from the ACAS FOT Dataset?

Which measures should be considered for inclusion in the workload prediction?

The equations to predict workload ratings were developed in a 3-step process. The first step was to identify potentially predictive measures in the ACAS dataset and develop a rationale for why particular measures should be included, consistent with the project requirements. That rationale is summarized in Table 32. Appendix I provides a precise description of how each measurement was defined and recorded.

Table 32. Rationale for Measures Examined

Category	Measurement or Derivative	Comment
Subject Vehicle	Speed	Driving faster should lead to greater workload overall. However, when workload is high, drivers slow down.

Category	Measurement or Derivative	Comment
		Risk homeostasis theory suggests speed might not have any relationship.
	Longitudinal acceleration	When the vehicle is braking or accelerating, but especially braking, the workload is greater.
	Lateral acceleration	When lateral acceleration is high, the driver is maneuvering (changing lanes, turning, merging) and workload is greater.
	Lane position (Distance to lane edge)	The closer to the lane edge, the more likely the driver is not to be attending to the road because they are overloaded.
	TLC	When the driver does not attend to driving (is distracted), TLC should decrease.
	Steering wheel angle	Large angles are associated with turns, lane changes, etc., higher workload situations.
	Throttle angle	Greater throttle angle results in greater speed, so the effects of throttle angle should be the same as speed.
	Steering reversals	Large corrections are associated with greater distraction (workload) as they can indicate inattention to the road.
	Steering entropy	Several studies have shown that greater entropy indicates greater workload.
Subject Driver	Age	Older drivers are less capable of dealing with workload, and rate situations as more difficult relative to young and middle-age drivers.
	Sex	Sometimes there is an age-sex interaction, with young men saying the driving is easy (low workload), but because of relatively poorer health, older men being more challenged (and giving higher workload ratings) than older women.
Other Vehicles	Number (detected by radar)	The greater the number of vehicles ahead, the greater the workload.
	Density	Greater traffic density leads to greater workload.
	Lead vehicle speed	The greater the speed of a lead vehicle, the greater the workload.
	Lead vehicle longitudinal acceleration	The greater the change in speed of a lead vehicle, the greater the workload.
Subject Vehicle–Other Vehicle Relationship	Gap (distance) to lead vehicle	The greater the distance to the lead vehicle, the less the workload.
	Range rate (gap rate)	The change in speed of a lead vehicle, especially deceleration, increases workload.
	TTC	Decreasing TTC increases workload.

Category	Measurement or Derivative	Comment
Road	Number of lanes	Increasing the number of lanes increases the number of vehicles the driver considers and, therefore, workload. However, the highest capacity roads (expressways) have been designed to minimize demand (wide lanes, gentle curves, few crossings of traffic streams), so the opposite relationship could occur.
	Class	Same as above.
	Posted speed	Driving faster than the posted speed may indicate lower workload (less traffic, less demanding geometry, etc.).

Subsequently, the desired statistics for each measurement were correlated with mean workload ratings for the 46 test clips (each of the 23 situations occurring twice, averaged across 24 subjects), or for whatever number of clips for which there was data. So, for example, for speed of the subject vehicle, those statistics included the minimum, mean, and maximum. Based on their correlations with workload and other reasons described in detail later, measures were then selected for inclusion in the regression equation, which was computed in step 2. Only correlations with mean workload ratings of 0.40 or greater were considered for inclusion in the workload prediction. That level was one at which predictors could offer some small amount of useful prediction while providing a manageable number of predictors to consider.

This process could have been done in other ways, but the intent was to proceed in a manner that considered reasons why, based on theory and prior research, particular statistics should be correlated with workload, and not to examine the correlations of everything with everything else. Without those limitations, there were too many opportunities for high correlations to occur by chance.

Which statistics for each measure should be considered for inclusion in the workload prediction?

Subject Vehicle-Subject Speed –There are arguments that workload could either increase or decrease with speed. Keep in mind there are 2 aspects to speed, the speed range, determined by the posted speed for a road, and how fast a particular driver goes on a particular road. The strongest argument, from risk homeostasis theory, is that people drive as fast as they can until they reach some desired level of workload. Since few drivers drive a maximum acceptable workload, speed is primarily determined by the posted limit. Accordingly, across conditions, speed should be unrelated to mean workload rating. However, variations in speed (e.g., standard deviation over the 30 s period) should reflect changes that needed to be made to adjust for workload in the scene. Correlations for the minimum, mean, maximum, standard deviation of speed, and standard deviation divided by the mean speed with mean workload ratings were -0.34, -0.31, and -0.23, 0.44, and 0.41. The standard deviation was divided by the mean to stabilize the statistic, as normally the standard deviation of speed increases with the mean because the potential range of speeds increases as speed increases.

Correlations related to the standard deviation of speed with mean rated workload were much greater then those related to first order estimates. This makes sense. The greater the workload, the more the driver needs to adjust speed in response to the road situation. Accordingly, the standard deviation of speed (Figure 29) and that value divided by the mean (Figure 30) were included in further analyses. Uncharacteristically, in this case, dividing by the mean decreased the correlation. Notice that in Figure 37 there may be 2 outcomes mixed together, one in which the SD/mean is unrelated to workload (the stack of points around 0) and others where there is a correlation. Nonetheless, both of these statistics were included in the next step.

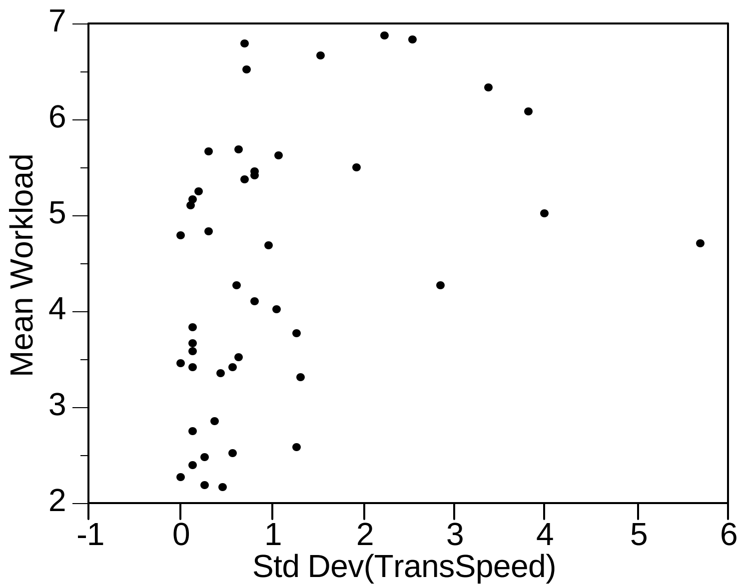


Figure 29. Standard Deviation of Speed vs. Mean Rated Workload

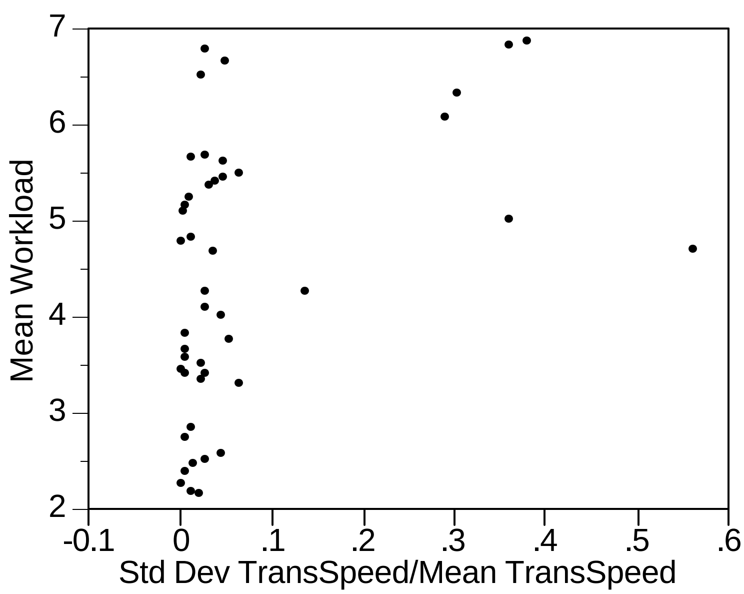


Figure 30. Standard Deviation/Mean Speed vs. Mean Rated Workload

Subject Vehicle-Longitudinal Acceleration-In response to traffic, the subject vehicle needs to accelerate and decelerate. Correlations were computed for filtered versions of minimum (-0.50), mean (0.02), absolute value (0.18), maximum (0.41), and standard deviation (0.47) of lateral acceleration. Distributions for the minimum, maximum, and standard deviation are shown in Figures 31, 32, and 33. The minimum reflects braking and the maximum reflects acceleration over the sampled interval.

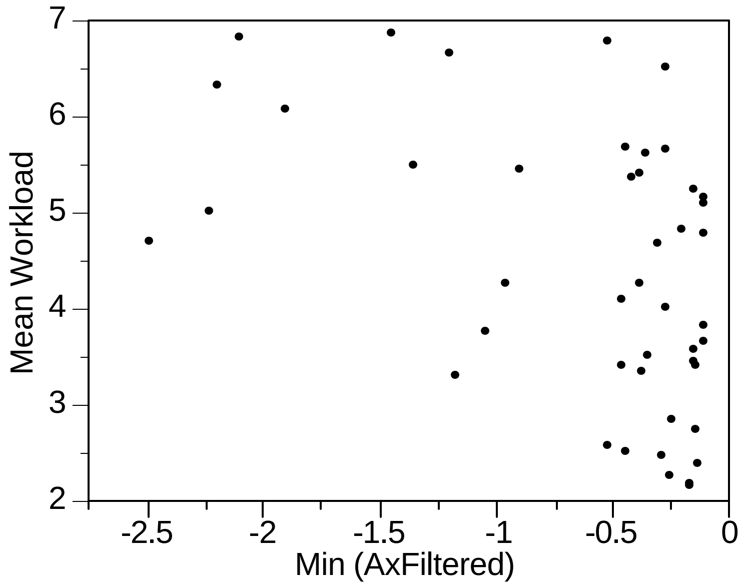


Figure 31. Minimum Longitudinal Acceleration (Filtered) vs. Mean Rated Workload

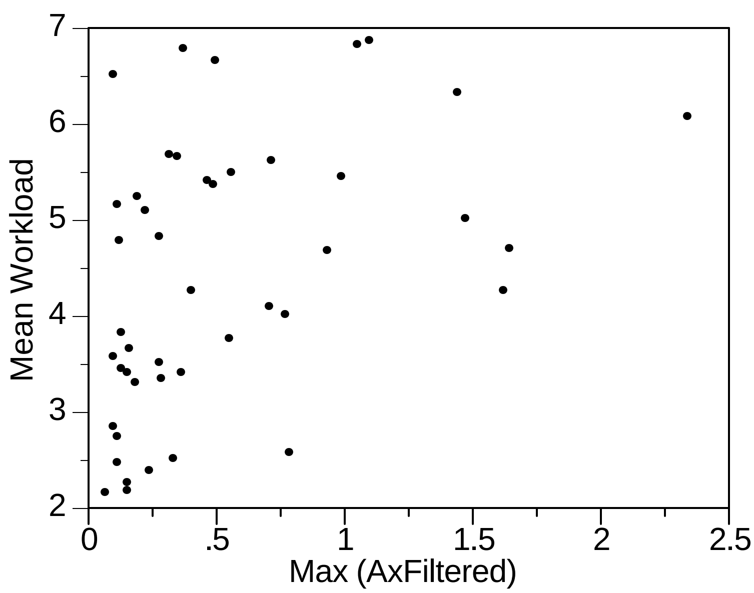


Figure 32. Maximum Longitudinal Acceleration (Filtered) vs. Mean Rated Workload

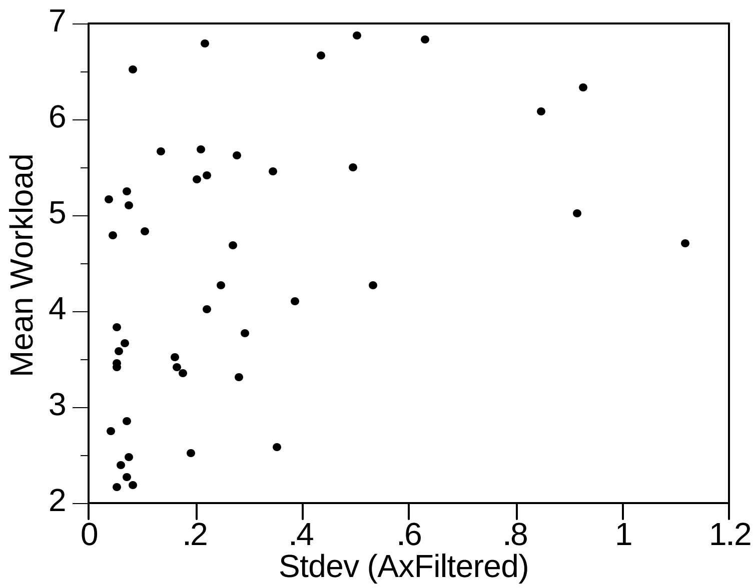


Figure 33. Standard Deviation of Longitudinal Acceleration (Filtered) vs. Mean Rated Workload

Subject Vehicle-Lateral acceleration differences were thought to be small because the mean was often 0 and the minimum and maximum were quantized at a few levels. That proved to be the case with minimum, mean, maximum, and the standard deviation of lateral acceleration correlating at -0.22, 0.01, 0.20, and 0.19 with mean workload rating respectively. Accordingly, they were not considered further.

One could argue if lane position is variable or if the mean value is very large or small (the driver was out of position, e.g., off to a side of a lane) then workload may be greater than normal. However, in this experiment, subjects were observers not engaged in driving, so lane position measures were thought to be less likely to be predictive of mean workload ratings. In fact, the correlations for all of the measures examined—mean lane offset (0.00), standard deviation of lane offset (or standard deviation of lane position) (0.06), mean distance to lane edge (0.19), minimum distance to lane edge (–0.09), maximum distance to lane edge (0.23), and standard deviation of distance to lane edge (0.11)—were extremely small. Hence, these measures were not considered further.

Subject Vehicle-Steering reversals have classically been used as a measure of workload, with larger reversals indicating a significant activity of the driver to correct their path (McLean and Hoffman, 1975.) Given the subject was an observer, low correlations were expected. In fact, that proved to be the case (r=0.19), so steering reversals were not considered further.

As the driver becomes inattentive to the primary task due to increased workload, time to line crossing (TLC) (Godthelp, Milgram, and Blaauw, 1984) decreases. All correlations of TLC measures with mean workload rating—minimum (0.10), mean (0.09), maximum (0.00) and standard deviation (-0.12)—were quite low, so TLC was not considered further.

Of the throttle measures, one might suggest that standard deviation of throttle might indicate variability in the road situation and greater workload, though in this case, given drivers were observers, correlations were expected to be low. In fact, that was somewhat the case with correlations of minimum, mean, maximum, and standard deviation of throttle being -0.32, -0.19, 0.08, and 0.23 respectively. The authors do not have a good explanation for a negative correlation of workload with minimum throttle. Given the low correlations, throttle measures were not considered further.

Subject Vehicle-Steering entropy is a measure of the randomness of the steering signal. The idea is the greater the extent to which future steering angle cannot be predicted from the past, the greater the random input from the driver. Since subjects were not driving, steering entropy was not expected to be correlated with mean workload rating. In fact, that was the case, with the 6 correlations all being less than 0.06 (SE 10 bins a=.4, SE 10 bins a=.2, SE 10 bins a=.05, SE 14 bins a=.4, SE 14 bins a=.2, SE 14 bins a=.05), so steering entropy was not considered further.

Subject Driver

Given that the focus of this analysis was on what the vehicle could measure in real time, and means were collapsed across driver age and sex, those factors were not considered at this step.

Other Vehicles (Traffic), Especially Lead Vehicle

Traffic has been well established in the literature as a source of workload as is described in the introduction. All measures of traffic including mean density, and minimum, mean, and maximum counts were all well correlated with mean workload rating (0.56, 0.61, 0.72, 0.68). (Minimum and maximum density could also have been considered, but the data was not available.) Interestingly, mean traffic count alone in the 30 s period accounts for half of the variance in predicting workload rating. Since subjects were observers not engaged in the driving task, the findings from this study may not fully reflect actual driving. Figures 34, 35, 36, and 37 show these findings. Notice the quantization of the minimum and maximum because they are counts of the distributions.

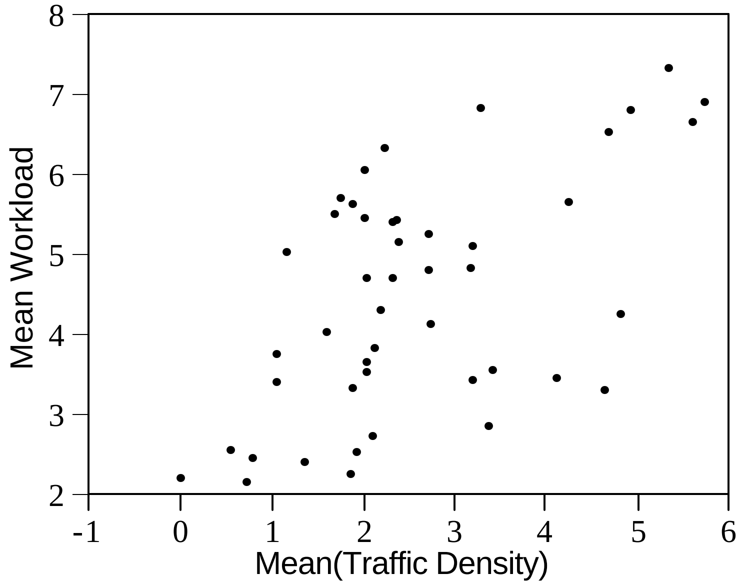


Figure 34. Mean Traffic Density vs. Mean Rated Workload

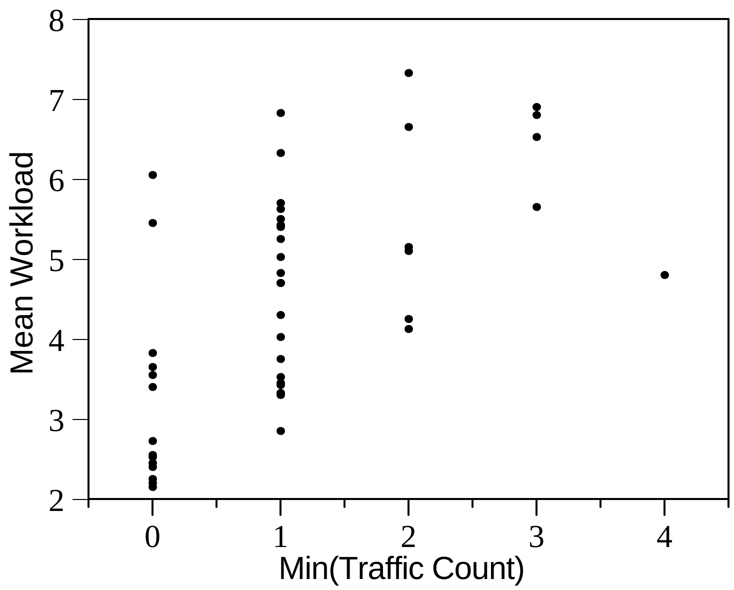


Figure 35. Minimum Traffic Count vs. Mean Rated Workload

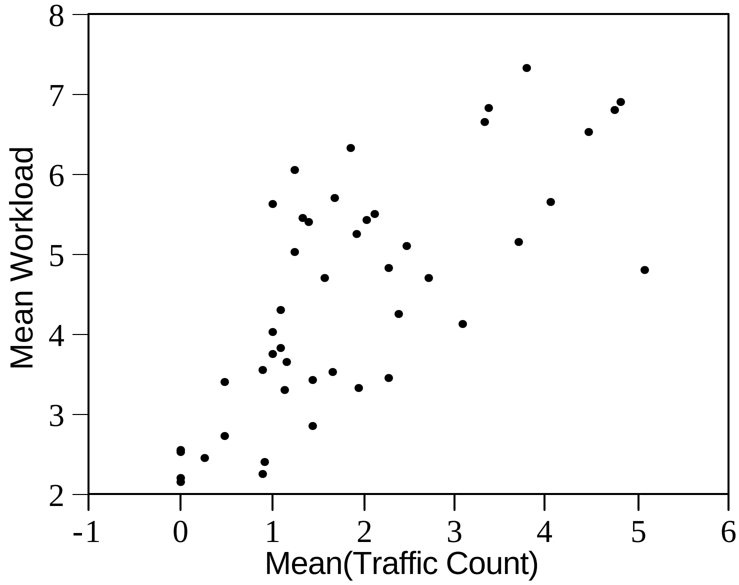


Figure 36. Mean Traffic Count vs. Mean Rated Workload

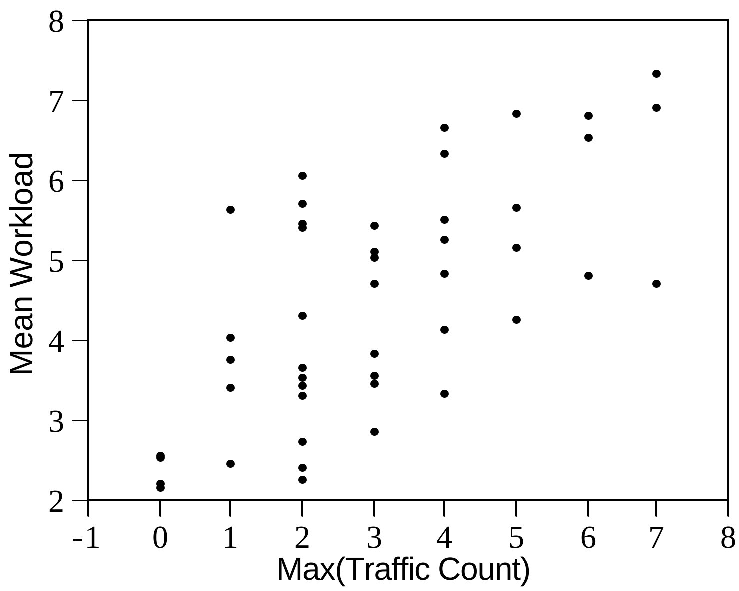


Figure 37. Maximum Traffic Count vs. Mean Rated Workload

How much workload a driver experiences should depend on how a lead vehicle behaves. For 9 of the 46 clips, there was no lead vehicle, so those cases were ignored. As was noted earlier, both higher and lower speeds could be associated with increase workload. In fact, correlations of the minimum, mean, and maximum <u>lead vehicle speed</u>

(Principal velocity, Vp) in the 30-s sample interval with mean workload rating were -0.10, -0. 13, and -0.11, so lead vehicle speed was not considered further.

For lead vehicle acceleration (Vpdot), negative values represent deceleration of the lead vehicle, leading the driver to brake, a contributor to workload. Positive values represent the lead vehicle moving away, less of a concern. In fact, the correlations of the mean workload rating with the minimum, mean, and maximum of lead vehicle acceleration were -0.55, -0.29, and 0.29 respectively. Obviously, the lead vehicle acceleration could not be computed for the 9 cases where there was no lead vehicle. Based upon these results, only the minimum lead vehicle acceleration (Figure 38) was considered further.

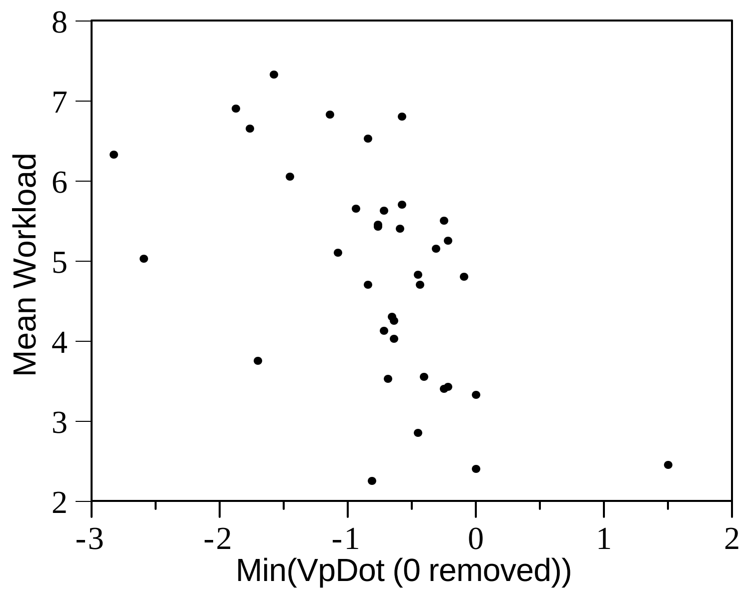


Figure 38. Minimum Lead Vehicle Acceleration vs. Mean Rated Workload

Subject Vehicle-Other Vehicle Relationship

The gap, the distance between the front bumper of the subject's vehicle and the vehicle ahead, is sometimes referred to headway distance in the literature. Internal to ACAS, the variable name was CIPV (Closest In-Path Vehicle) range. Clearly, the closer a lead vehicle is to the subject, the greater the workload. However, keep in mind that drivers have a preferred time headway and find extremely short headways heavily loading. Furthermore, when a vehicle is quite far away, having it even further away does not reduce workload very much.

Shown in Figures 39, 40, and 41 are the relationships between Mean Rated Workload and the Minimum, Mean, and Maximum Range for all 46 clips, with correlations of -0.77, –0.76 and -0.60 with rated workload, thus indicating that both minimum and mean distance to the lead vehicle are important, but maximum is the least important. Note the numerous reported ratings at 125 meters, maximum range of the radar. This was because when there was no target detected by the radar, 125 m was substituted, assuming the target was out of range. This substitution occurred for 9 of the 46 minimum values, 9 of the means, but 23 of the maximum values. This rather large number of substitutions may partially explain the lower correlation with mean rated workload.

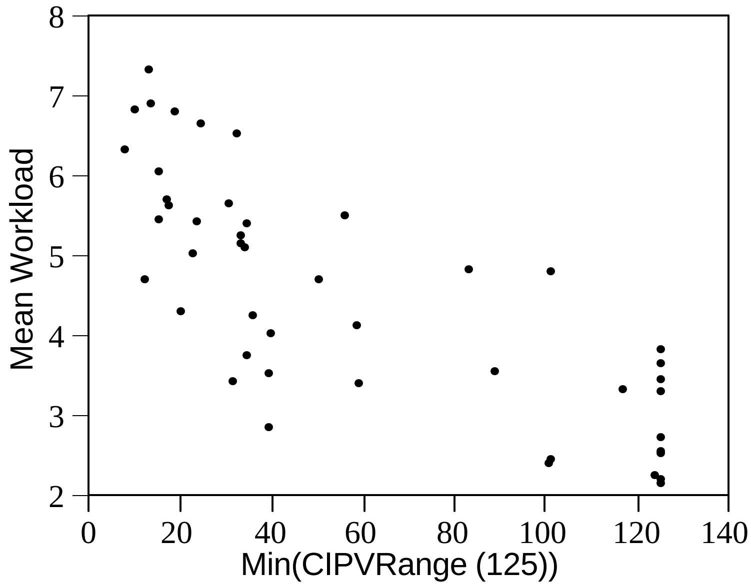


Figure 39. Minimum Gap vs. Mean Workload Rating

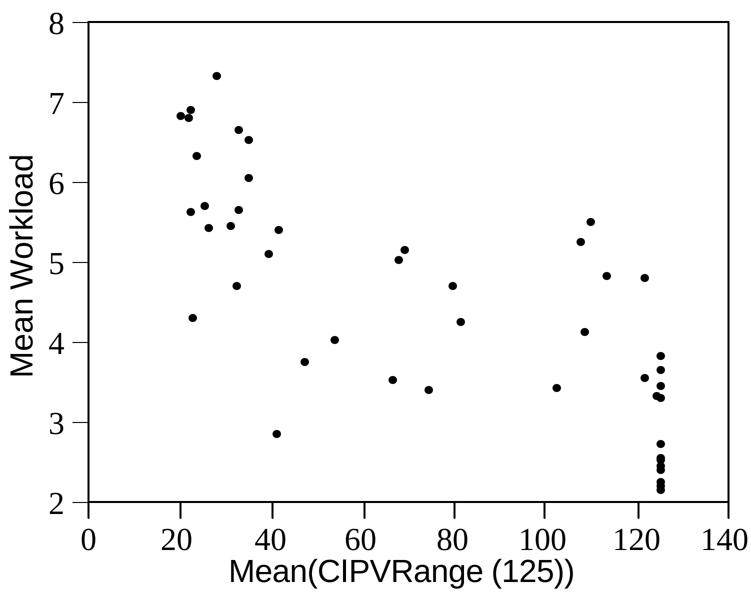


Figure 40. Mean Gap vs. Mean Workload Rating

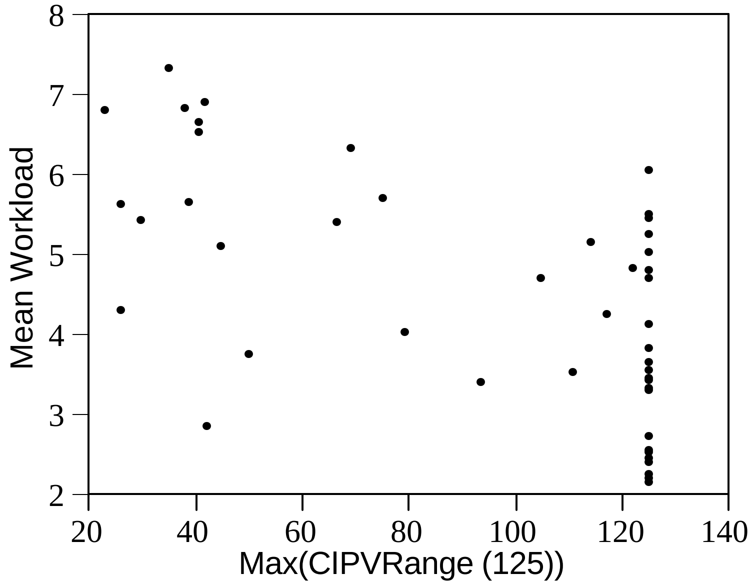


Figure 41. Maximum Gap vs. Mean Workload Rating

Of those 2 statistics, minimum range is probably the best choice, but the mean was also considered further. There were several times where no vehicle was ahead of the subject in their lane within range of the radar, and then a vehicle cut in from an adjacent lane. In that situation, there were thus 2 different workload periods.

Careful examination suggests the relationships between the mean gap and workload and minimum gap and workload are not linear and should not be given the reasoning presented earlier. In fact, research by Wierwille, and UMTRI's research in phase 1 suggest a log relationship. Figure 42 shows the relationship between log minimum gap and workload (r=-0.81), which appears quite linear. Why some nonlinearity remains in relationship between log mean range and mean workload rating (r=-0.77) is unknown (Figure 43). Figure 44, showing the maximum, indicates greater scatter (r=-0.59). As a footnote, the correlations change slightly, depending on how they are calculated. For example, when the 9 cases where the target was out of range are treated as missing (instead of using the max range), the minimum, mean, and maximum correlations decrease to -0.74, -0.73 and -0.51.

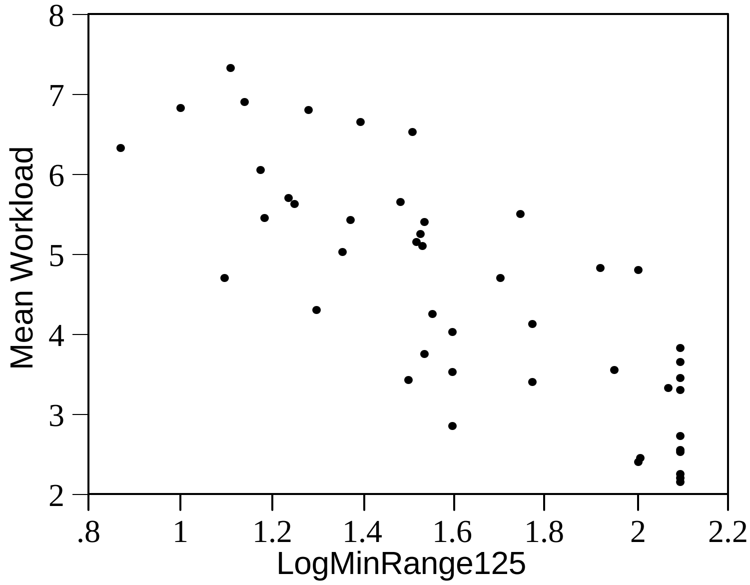


Figure 42. Log Minimum Gap vs. Mean Rated Workload

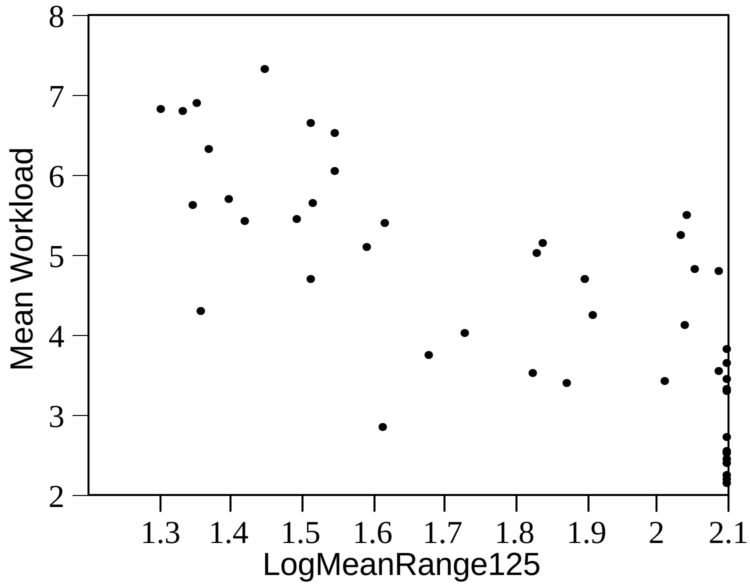


Figure 43. Log Mean Gap vs. Mean Rated Workload

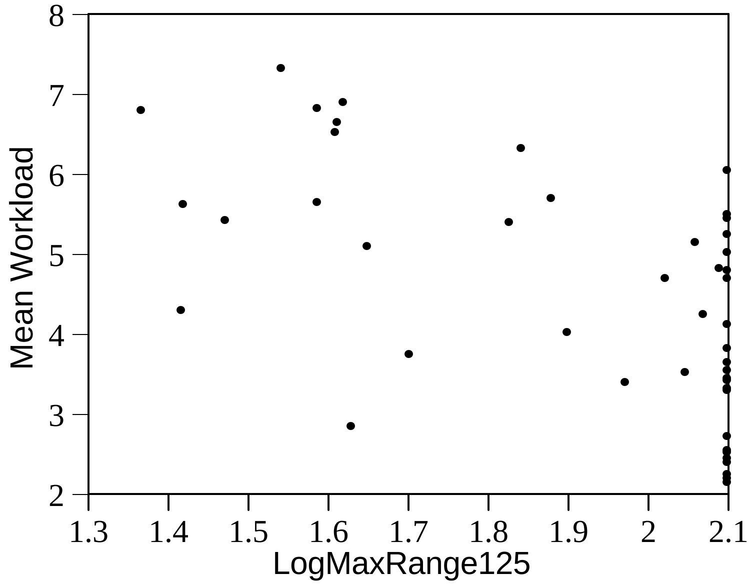


Figure 44. Log Maximum Gap vs. Mean Rated Workload

So how do drivers think about the impact of the distance to a lead vehicle on workload? One interpretation of Figure 43 is that the relationship is piecewise linear, with drivers being very concerned from 20 m (the closest following distance in the clips) out to about 50 meters, and increases in distance leading to marked reductions in workload. For expressways, 50 m corresponds to a 1.8 to 2.0 s headway, a typical following distance. Beyond that distance, increases in distance led to lesser reductions in workload.

However, splitting the data this way (at 50 m) and computing regression did not lead to a statistically significant relationship for the <50 m section, primarily because of variability in the workload ratings. Recall that when individuals re-rated clips, those ratings often differed by a value of 1, and when different clips for the same situation were rated, the range of values sometimes approached 2. Thus, trying to predict a workload rating to less than 1 may be beyond the limits of these data, especially since ratings were generally estimated to the nearest 0.5 on the workload scale.

When workload is to be estimated, what values should be assigned when the range exceeds 125 m (and is missing)? One strategy is to compute a linear relationship between log minimum range without those points (a linear function), and then using mean workload of the missing gap data, estimate the gap using the workload equation.

For the 9 cases where the range exceeded 125 m, the mean rated workload was 2.9. Based on regression analysis without those 9 cases, workload = 9.6 – 3.2(LogMinimumGap) or rearranging, MinimumGap = 10^((workload-9.6)/-3.2). The rated mean workload for the 9 points where the minimum was missing (vehicles were always out of range) was 2.9. Substituting, that leads to an estimate of 124 m. (Note:

Substituting nonrounded values leads to an estimate of 134 m. Thus, if the gap exceeds 125 m, assuming the gap is 125 m seems reasonable.

Also of interest is the derivative of gap, commonly referred to as range rate. Consistent with the analysis of gap, looking at the maximum makes sense (and this was verified by correlations, 0.02 for the minimum, 0.29 for the mean, 0.43 for the maximum with rated workload (Figure 45), computed where the missing cases were ignored). However readers should keep in mind that since the lead vehicle was out of range for 9 of the test clips, there is no range rate data for those clips.

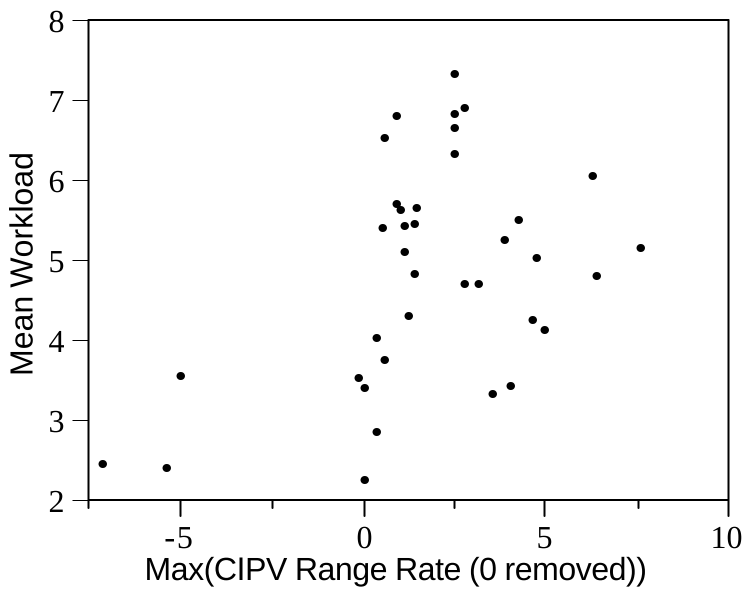


Figure 45. Maximum Range Rate vs. Mean Rated Workload

Gap time is commonly called headway time (and could be called range time). It was defined as CIPV Range / TransSpeed, where CIPV Range was in meters, and TransSpeed was in m/s. In contrast to gap distance, gap time considers the relative velocities of the 2 vehicles. Gap time is undefined when no lead vehicle is present, as was the case for 9 of the clips. One can therefore proceed 2 ways, (1) treating those 9 cases as missing, or (2) assuming the range is 125 m in those cases and computing a value. Consistent with earlier calculations, substitution was preferred. Using that approach, correlations of the minimum, mean, and maximum gap time with the mean workload were -0.72, -0.60, and -0.11 respectively, so the minimum and mean gap time (Figures 46 and 47) were considered for further examination (though including them reduces the sample size for analysis. Interestingly, eliminating the 9 substituted cases reduced the correlations to -0.58, -0.44 and -0.01).

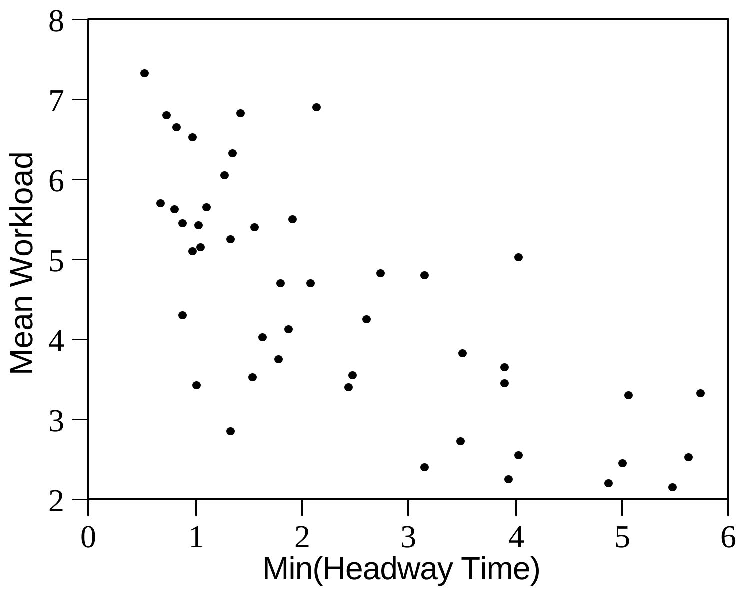


Figure 46. Minimum Headway Time vs. Mean Workload Rating

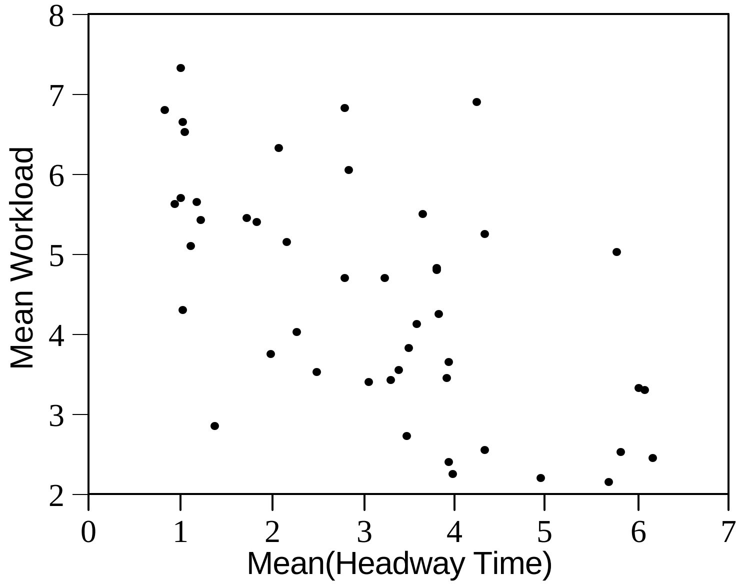


Figure 47. Mean Headway Time vs. Mean Workload Rating

Just as for gap, a log transform was considered, for the same reason that transform could make sense here, in particular as suggested by the minimum headway time. Taking logs improves the correlations of the minimum and the mean slightly, but not the maximum with values of -0.73, -0.61, and -0.31. Figures 48 and 49 show the Log transforms for the minimum and mean.

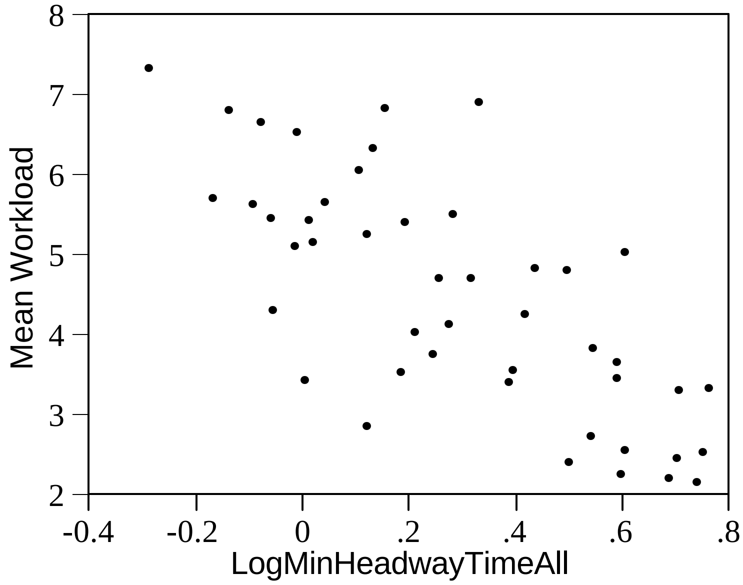


Figure 48. Log Minimum Headway Time vs. Mean Rated Workload

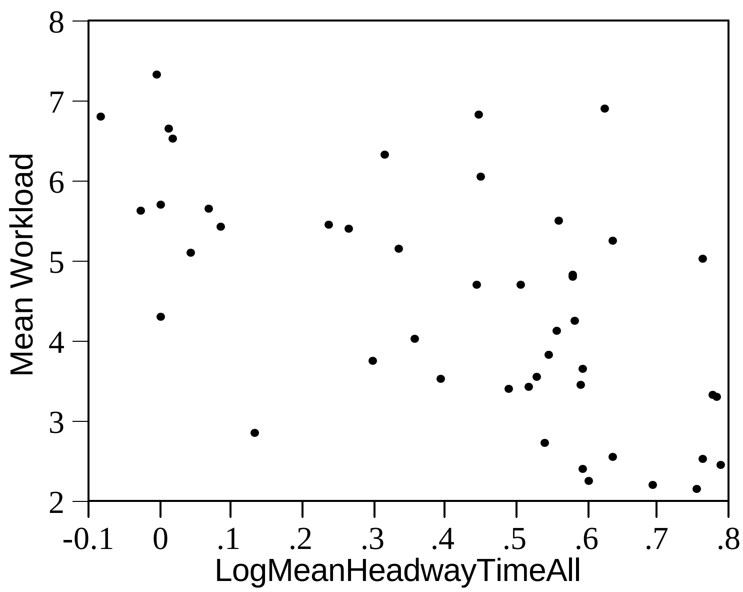


Figure 49. Log Mean Headway Time vs. Mean Rated Workload

TTC (Time to Collision)

The time to collision is defined as minus the range (gap) to the lead vehicle divided by range rate, and is valid for only when the range rate is negative, when the subject is closing on a lead vehicle. (The minus in front makes it positive.) It differs from gap time, also called time gap and headway time, defined as range divided by speed (forward velocity). Readers should be aware that none of these measures are perfect. For example, here range (the distance from the subject's front bumper to the lead vehicle rear bumper) is determined by a scanning laser or radar. What the sensor picks up as the lead vehicle depends on the vehicle, so the range value can be slightly in error. In some cases, the rear bumper might be detected, but plastic bumpers are not good targets. The senor might detect the taillights, or the back of the trunk or the tailgate of a pickup truck, or maybe even the rear axle.

Interestingly, time-to-collision and gap time were only moderately correlated (r=-0.46). One of the difficulties with TTC is that depending on how it is calculated, there is a reasonable chance TTC is zero in a 30 s interval, and for the 46 clips, the minimum was 0 for 17, constraining calculations. Removing the cases where TTC=0, the minimum, mean, and maximum were -0.45, -0.20, and 0.02. (See Figure 50.) Since including TTC, the minimum being the best choice, would have reduced the number of samples in the regression analysis considerably, TTC was not included in the regression calculations here, though it could be in future analyses.

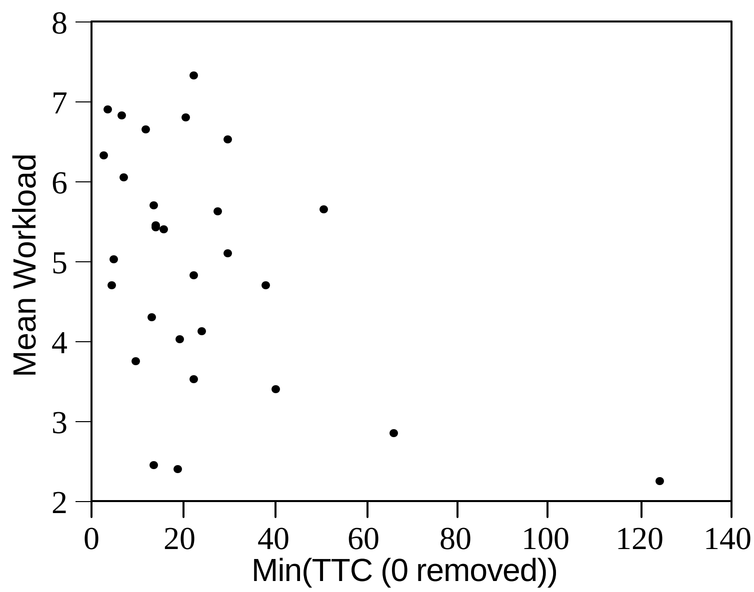


Figure 50. Minimum TTC vs. Mean Workload Rating

Road

<u>Lane width</u> was thought to be a predictor, as wider lanes are easier to drive, providing more room for steering error, but less so than other factors such as traffic. In fact, the correlations for minimum, mean, and maximum were 0.15, 0.13, and 0.14 respectively. Given the low correlations, lane width statistics were not considered further.

It was hypothesized that with more lanes would handle traffic, and thus be more demanding to drive. In fact, the minimum, mean, and maximum number of lanes in every segment was the same, and the correlations of the number of lanes with mean workload rating was 0.00, so those statistics were not included in further analysis.

What Are the Equations That Predict Workload of Driving (of the Clips Observed) from the Driving Statistics?

Summarizing the previous section, Table 33 shows the statistics whose correlations absolute values with mean workload rating were greater than 0.40 and were therefore candidates for inclusion in an equation to predict mean workload rating. Interestingly, minima and maxima are more common than means or standard deviations.

Table 33. Statistics with Correlated (r>0.4) with Mean Workload Rating

Category	Measure	Statistic	Correlation with Mean Workload	# Cases	Included in Regression Analysis?
Subject vehicle	Subject speed (TransSpeed)	Standard deviation (sd)	0.44	46	yes
		sd/mean	0.41	46	
	Long acceleration (Ax filtered)	Minimum	-0.50	46	yes
		Maximum	0.41	46	
		Standard deviation	0.47	46	
Other vehicles	Density	Mean	0.56	46	yes
	Count	Minimum	0.61	46	yes
		Mean	0.72	46	
		Maximum	0.68	46	
	Lead vehicle acceleration (VpDot)	Minimum	-0.55	37	sometimes
Subject-Other Vehicle Relationship	Gap (CIPVRange 125)	Minimum	-0.77	46*	no; note: if gap was missing, 125 was substituted
		Mean	0.76	46*	
		Maximum	-0.60	46*	
	Log gap	Minimum	-0.81	46*	yes, log better predictor than untransformed value
		Mean	-0.77	46*	
		Maximum	-0.59	46*	
	Gap rate (CIPVRange RateBlank)	Maximum	0.43	37	sometimes (note: cannot compute range rate when no lead vehicle)
	Gap time (Headway	Minimum	-0.70	46*	yes
		Mean	-0.60	46*	

	Time)				
	Log gap time	Minimum	-0.73	46*	log better predictor than untransformed value
		Mean	-0.61	46*	
	TTC (0 removed)	Minimum	-0.45	25	no, too many missing cases

In the previous analyses, there were a number of decisions that were made that influence subsequent regression computations. First, what should be done when there is no lead vehicle present? Looking at the figures, when there was no lead vehicle present, substituting the maximum range of the sensor, 125 m, looked consistent with the pattern of the data and slightly improved the correlations. Therefore, that adjustment was made, increasing the number of cases for those statistics to 46.

Second, what should be done in other instances when they are less than 46 cases? The greater the number of cases, the better the prediction, though requiring that all 46 cases be examined means deleting measures for which all 46 cases are not present, here lead vehicle acceleration and TTC. Examining ½ of the data so TTC could be included did not make sense. Therefore, there were 2 regression analyses, 1 with all the data (potentially 46 cases) and 1 where lead vehicle acceleration and gap rate were included (potentially 37 cases).

Third, should the original or log transformed value be used for gap related measures? Where the data were available, only the log-transformed statistics were used. The log transforms more closely approximate how drivers use gap related information and that point is supported by the correlations. There was no reason to use log transforms for derivatives (e.g., range rate).

Finally, in prior analyses, it was observed that the right merge cases were slightly different, adding about 1.0 to the mean workload estimate. Therefore, those 4 cases were omitted from the regression computation but manually added later. Stepwise regression was therefore used to predict mean rated workload using all the variables in Table 33 for which there were initially 46 cases. After the first 2 steps (with entry $p<.0001$), the prediction equation was:

Mean Workload Rating = 8.86 -3.00(LogMeanRange125) + 0.47(MeanTrafficCount)

Where:
LogMeanRange125 = Logarithm of the mean of the distances in meters to the lead vehicles in the same lane as the subject averaged over 30 sec. If there was no vehicle within 125 m, the range of the radar, the distance was set to 125 m.
MeanTraffficCount = Mean number of vehicle detected by the subject vehicle radar (15 degree field of view) averaged over 30 s.

This equation accounts for over 82% of the variance in the mean workload ratings, exclusive of the right merge situations. For them, add 1 to the computed workload rating.

Figure 51 shows the difference between the predicted and actual values for all 46 points. There were only 5 cases where the 2 values differed by more than 1--straight section of a rural road (LOS C), straight sections of an urban road (LOS C,E), and straight sections of an expressway in the right lane (LOS A,E). In all cases, only 1 of the 2 clips for each situation had errors of this size, suggesting the errors were clip specific and there were no relationships between prediction errors and road type, LOS, geometry, or other factors.

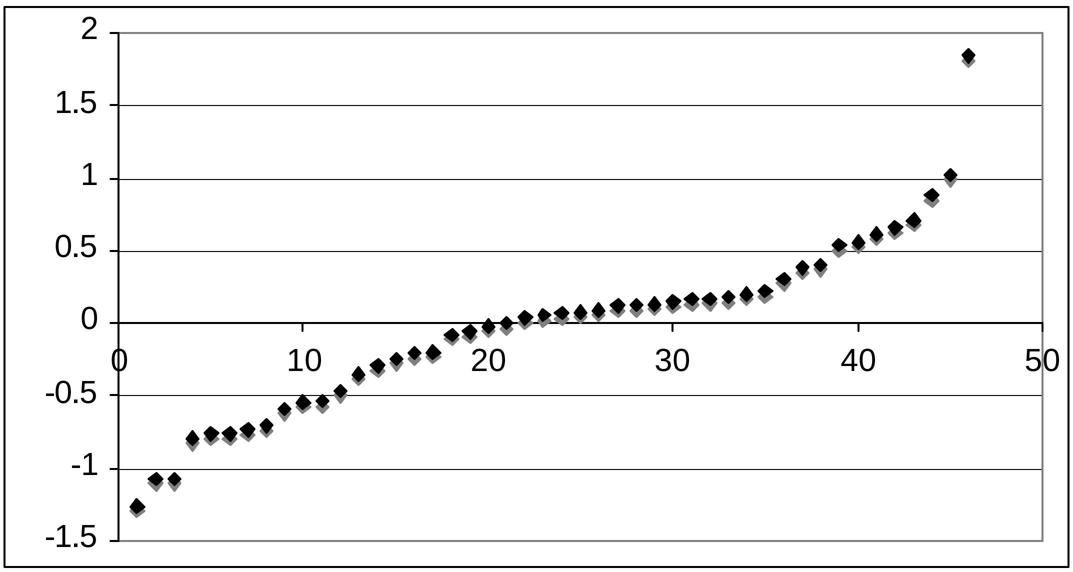


Figure 51. Residuals for the 2-Term Equation

In the next step in the analysis, (with p=.013 for entry), the resulting equation was:

Mean Workload Rating = 8.87 - 3.01(LogMeanRange125) + 0.48(MeanTrafficCount) + 2.05(MeanAxFiltered)

where:

LogMeanRange125 = Logarithm of the mean of the distances to the lead vehicles in the same lane as the subject averaged over 30 sec. If there was no vehicle within 125 m, the range of the radar, the distance was set to 125 m.

MeanTraffficCount = Mean number of vehicle detected by the subject vehicle radar (15 degree field of view) averaged over 30 s.

MeanAxFiltered = Mean longitudinal acceleration (m/s2), filtered

This equation accounted for 87% of the variance of the mean workload rating, an extremely large value. As an aside, the authors are still attempting to determine how the longitudinal acceleration was filtered by an internal GM algorithm, though it is known that the reported mean is based on 10 data points.

In subsequent steps, entry probability exceeded 0.05, the cutoff.

As shown in Figure 52 for the second regression equation, there were only 3 data points where the predicted and actual value difference by more than 1.0, remarkably close. They were for a straight section of an urban road (LOS C), and 2 situations for the right hand lane straight sections of an expressway (LOS A, E). Again, there were no instances were both clips of a situation had large prediction errors and there was no apparent pattern to the residuals.

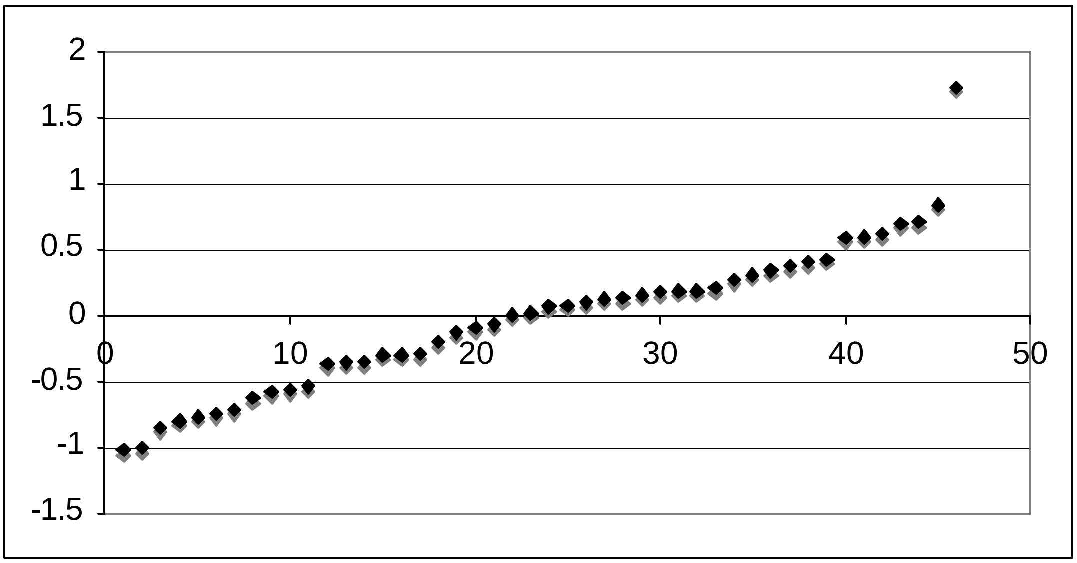


Figure 52. Residuals for the 3-Term Equation

A third regression equation was computed, using the stepwise entry criteria of the second model, but including the minimum lead vehicle acceleration and the maximum gap rate in the data et. The cost of adding these 2 variables was to reduce the number of cases from 42 to 31. In that analysis, LogMeanRange125 entered the model first, followed by MeanTrafficCount, and then MeanAxFiltered as before, with slightly different coefficients resulting. In the 4th step, MinimumVpDot(0 removed) entered the model (p=0.034), the last term before the 005 cutoff. That equation, which accounted for 85% of the mean workload rating variance was:

Mean Workload Rating = 8.07 – 2.72(LogMeanRange125) + 0.48(MeanTrafficCount) + 2.17(MeanAxFiltered) - 0.34(MinimumVpDot(0 removed))

where:

LogMeanRange125 = Logarithm of the mean of the distances to the lead vehicles in the same lane as the subject averaged over 30 sec. If there was no vehicle within 125 m, the range of the radar, the distance was set to 125 m.

MeanTraffficCount = Mean number of vehicle detected by the subject vehicle radar (15 degree field of view) averaged over 30 s.
MeanAxFiltered = Mean longitudinal acceleration (m/s2)
MinimumVpDot(0 removed) = Minimum acceleration of a lead vehicle in m/s2 averaged over a 30 s interval, with deceleration of the lead vehicle being negative values. Cases where there was no lead vehicle were not included in the computation.

Thus, system developers have 3 equations to choose from, one with strict entry criteria that includes only log mean range and traffic count, a second that also includes longitudinal acceleration of the subject vehicle, and a third that also includes minimum acceleration of the lead vehicle. All of these equations predict in excess of 80% of the variance of the workload-rating estimate. In most human factors studies, accounting for 50% of the variance is considered good, so these results are remarkable.

According to the Post-Test Ratings, How Does the Workload of Driving Vary as a Function of Road Geometry and Traffic?

Ideally, one would like to be able to estimate workload from data collected by the vehicle. One of the limitations of the clip rating data is that a large number of situations encountered in real driving were not examined, a limitation due to the time available to test subjects and the clip data base. To obtain the needed information, workload ratings for a wider range of situations than was shown in the clips was collected after the clip rating task. Specifically, subjects filled out a post-test survey specifying the workload on a 0 (“no demand”) to 100 (“completely requires all of your capacity to just drive”) scale for urban, rural, and residential roads, as well as expressways. Each subject rated 200 situations for a total of 4800 data points. In contrast to the clip rating data, there were no instances of subjects rating the same situation twice, so there are no reliability data.

Mean ratings (averaged by subject) ranged from 38 to 92 (overall mean 56, indicating most of the scale range was used, a desired result. A different scale was used than in the clip rating experiment to encourage subjects to think independently and more broadly about driving workload, not just about the clips they had seen. Further, because there were few residential roads in the ACAS database, clips of them were not rated but residential roads were rated in the post-test.

In an ANOVA of the post-test ratings, Age, Sex, Age*Sex, and Subjects within Age*Sex were examined as well as Road differences, and interactions of Road with Age, Sex, and the Age * Sex interaction. All factors were significant at $p<.0001$, except Age* Sex, which was significant at $p=.03$. In contrast to the clip-rating task, the order of ratings (Figure 53) from high to low was middle-aged subjects (mean=63) followed by older subjects (63) followed by younger subjects (49). The largest sex disparity was for younger subjects (men rating the workload to be less), with the size of the difference decreasing with age.

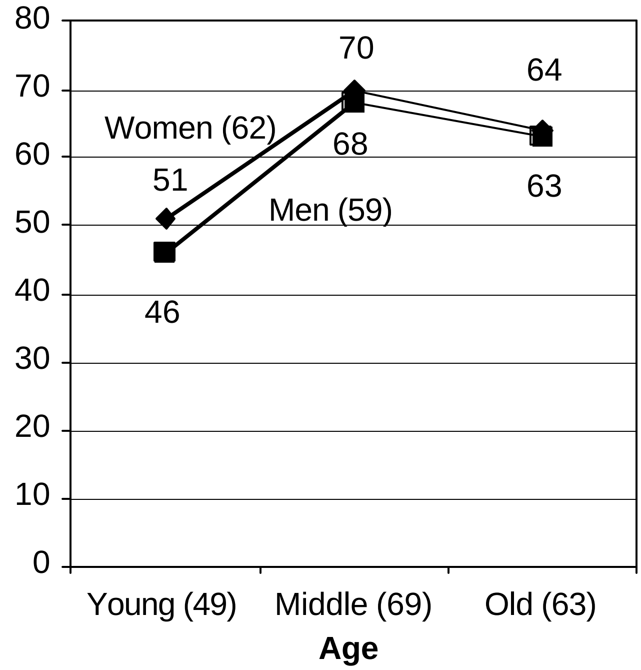


Figure 53. Mean Post-Test Workload Ratings by Age and Sex
(Averaged Across All Roads, Means in Parentheses)

One can argue if treating road type as a factor makes sense or not. In contrast to the clip-rating task, there was an effort to broadly sample the conditions found on real roads and for the traffic levels likely to be experienced. However, no effort was make to make those samples statistically representative. Accepting the premise of reasonableness, the means were 58 for rural roads, 63 for urban roads, 61 for expressways, and 54 for residential roads. Since the clips were monochromatic, traffic lights were relatively less evident in clips than in real scenes, potentially depressing the urban road ratings.

The correlations of the post-test ratings with the clip ratings suggest the port-test ratings for the different road types may have some independence. Nevertheless, it is still interesting to view all road types together (Table 34). Notice that downtown driving has the highest rated workload, but there are expressway and rural scenarios that also have high ratings. The situation with the lowest rating is residential.

Table 34. Rank Order of Road Situations

Road Modifier	Rating
Urban, Downtown	72
Expressway, with Crash Scene	71
Expressway, with Construction	70
Rural, Very Curved or Hilly	69
Expressway, with Lane Drop	64
Expressway, with 3-foot shoulder	62
Urban, w/ Commercial Building on Corner	60
Rural, w/ Stop Sign for Cross traffic	60
Residential, w/ Signaled Intersection	59
Rural, w/ 1-foot Shoulder	58
Expressway, Curved or Hilly	58
Expressway, with Interchange	58
Rural, w/ Signaled Intersection	58
Residential, w/ 25+% Parked Cars	58
Residential, Curved or Hilly	55
Rural, Curved or Hilly	54
Residential, w/ 0-25% Parked Cars	51

Table 35 shows the various road modifiers, with crashes and construction adding most to the workload ratings, being about 50% of the total workload for the lowest workload situations. The value of the modifiers differ slightly between road types, for example hilly or curvy adds 10 in 1 case, 11 in another.

Table 35. Overall Rank Order of Road Modifiers

Road Modifier	Rating
Crash Scene	23
Construction	22
Very Curved or Hilly	19
Downtown	16
Lane Drop	15
Signaled Intersection	15
3-foot shoulder	14
25+% Parked Cars	14
Curved or Hilly	10
Stop Sign for Cross traffic	10
Interchange	10
1-foot Shoulder	9
Signaled Intersection	9
0-25% Parked Cars	7
Curved or Hilly	5
Commercial Building on Corner	4

Table 36 shows the mean workloads averaged across subjects for each of the situations examined for rural roads. The data have been resorted from the original survey in ascending order by situation. Because there were 3 groups of subjects with only 8 subjects each, these data have not been partitioned by age. If so desired, these values could be corrected for age using the data from the previous figure (subtract 11 for young subjects, add 9 for middle-aged subjects, add 3 for older subjects), though there were road * age interactions. As a reminder, the largest difference for the post-test ratings was between the younger and other subjects.

Table 36. Mean Post-Test Workload Ratings for Rural Roads

Two values are for: (a) no or little traffic / (b) some traffic.

a / b

Only 2 values were considered since heavy traffic is rare on rural roads.

Situation	Total # Lanes			
	2	3 (Center Pass/Turn Lane)	4 (in Left Lane)	Mean
Base case=straight road 8 foot paved shoulder + 8 foot grass beyond that	40 / 54	44 / 56	45 / 57	43 / 56
Base case except gentle curves or hill	47 / 59	49 / 60	50 / 61	49 / 60
Base case with 1-foot shoulder, mailboxes, rocks, vegetation beyond	53 / 62	53 / 64	54 / 64	53 / 63
Base case + at or approaching intersection with traffic light	51 / 62	52 / 63	55 / 64	53 / 63
Base case + at or approaching intersection with a stop sign for the crossing road only	53 / 62	54 / 65	55 / 67	54 / 65
Base case except very curved or hilly road (mountain road)	64 / 74	65 / 74	63 / 74	64 / 74
Mean	51 / 62	53 / 64	54 / 65	53 / 63

For rural roads, mean ratings spanned a range of 34 (40 to 74 on a 0 to 100 scale). Narrowing the shoulder to 1 foot (from 8) increased the workload to a similar level of approaching a stop sign or traffic light (all changes of roughly 10 points). Changing to a mountain road (from the base case) led to an increment of roughly double the previous situations. Interestingly, the effect of traffic (from none or little to some) was also about a 10-point change. Adding lanes in themselves had smaller effects, between 1 and 2 points for each additional lane.

Table 37 provides the data for urban roads, ranging from 45 to 84. Going from the base case to some commercialization increased the ratings by about 4 on average, whereas the next increment (to “downtown”), increased ratings by about 13. As before, each addition lane adds about 1 or 2 to the rating. Traffic effects were substantial, with the increase from no/little to some increasing ratings by 8 points and the next increment to heavy being another 8 points. This increase was fairly consistent across conditions.

Table 37. Mean Post-Test Ratings for Urban Streets

Three values are for: (a) no or little traffic / (b) some traffic / (c) heavy traffic

a / b / c

Situation	# Lanes				
	2	**3 (Center Turn)**	**4 (Includes Turn Lane)**	**5 or More**	**Mean**
Base case=straight rd, cars parked on side, 10 intersect/mi, most with lights, no or few pedestrians, no stores	45/53/63	47/54/63	49/56/64	52/61/70	48 / 56 / 65
Base case but stores or gas station on corner	49/57/67	51/58/67	52/59/68	56/63/73	52 / 59 / 69
Base case but numerous stores & pedestrians ("downtown"), midblock driveways, no double parking	62/69/76	64/71/78	65/73/81	70/76/84	65 / 72 / 80
Mean	52/60/69	54/61/69	55/63/71	59/67/76	55 / 63 / 71

Table 38 shows the rating for expressways, ranging from 30 to 82. The expressway case included the most difficult situation, driving through construction in heavy traffic. Interestingly, this was rated as more demanding then a mountain road. Probably the mountain road would have been rated higher if there were constraints on shoulders and sheer drop offs.

Table 38. Mean Post-Test Ratings for Expressways

Three values are for: (a) no or little traffic / (b) some traffic / (c) heavy traffic

a / b / c

Situation	Total # Lanes (So 6=3 per Direction)			
	6 (in Left Lane)	**6 (in Middle Lane)**	**6 (in Right Lane)**	**Mean**
Base case = straight road, 1-lane paved shoulder on each side, wide grassy median, no guardrails needed	30 / 43 / 63	32 / 49 / 64	35 / 49 / 68	32 / 47 / 65
Base case+ Curved or hilly	45 / 58 / 72	45 / 59 / 70	46 / 59 / 71	45 / 59 / 71
Base case + Interchange (entrance/exit) in view or at it	40 / 54 / 72	44 / 56 / 73	48 / 61 / 75	44 / 57 / 73
Base case + Lane drop (e.g., 3 to 2 lanes) in your or adjacent lane	50 / 58 / 74	46 / 60 / 73	51 / 62 / 75	49 / 60 / 74
Base case but 3-foot shoulder & guardrail instead	49 / 61 / 74	47 / 61 / 73	51 / 63 / 79	49 / 62 / 75
Base case + Construction: Approaching or driving in lane shift or narrow lanes with concrete barriers, no shoulder	59 / 69 / 80	60 / 71 / 80	61 / 72 / 82	60 / 70 / 81
Base case + Approach or driving through crash scene	62 / 69 / 80	61 / 71 / 81	63 / 70 / 81	62 / 70 / 81
Mean	47 / 58 / 73	47 / 61 / 74	51 / 62 / 76	49 / 61 / 74

As before, situations have been listed in increasing order, not the order they were listed on the data collection sheets.

Table 39 shows the residential data, again sorted in order of increasing workload, with ratings ranging from 38 to 64. There was no particular pattern to the situations examined, though there was a marked increase from the base case (no parked cars) to any other situation. Each increment in the number of driveways increases the workload by about 6.

Table 39. Mean Post-Test Ratings for Residential/Suburban Streets
Since suburban streets rarely have traffic, only no or little traffic was considered.

Situation	Driveways (per Side of the Road)			
	0-<2 / Block (0.1 miles)	**2-5 / Block**	**> 5 / Block**	**Mean**
Base case, straight road, no parked cars, no intersection nearby	38	44	50	44
Base case, but >0 - 25% of curb has parked cars	46	51	58	52
Base case, but curved or hilly	50	54	60	55
Base case, but >25% of curb has parked cars	52	58	64	58
Base case, but at or approaching signed intersection, where you need to stop	55	59	64	59
Mean	48	53	59	54

Examination of the tables indicated that subjects rated workload using an additive model. Accordingly, Table 40 can be used to estimate workload ratings consistent with the post-test data. Some of the error in the table was due to rounding up. As an example, the prediction of workload for a rural road minimum case is 58 (mean) + road modifier (base case, -8) + lane factor (2 lanes, -1) + traffic (little/none, -5) for a total of 44, versus 40 provided by subjects. At the other end, for a 4 lane mountain road with some traffic, the table based total is 58+11+2+5=76 (versus 74 in the table. What matters most in these data is the relative size of various factors. For example, the data indicate that construction zones and crash scenes increase workload significantly, so attention is needed to these situations.

Table 40. Post-Test Workload Estimate

Road Type & Mean	Road Modifier		Lane Modifier		Traffic		Driveways	
Rural Mean=58	-8	Base case	-1	2 Lanes	-5	None/Little		
	-3	Gentle curve/hill	1	3 Lanes (in left)	+5	Some		
	-3	1-ft shoulder	+2	4 Lanes (in left)				
	+1	At, approach light						
	+2	Stop sign for others						
	+11	Very hilly, curved						
Urban Mean=63	-7	Base case	-3	2 Lanes	-6	None/Little		
	-3	Corner business	-2	3 Lanes	-3	Some		
	+9	Downtown	+0	4 Lanes	+9	Heavy		
			+4	>=5 Lanes				
Xway Mean=61	-13	Base case	-1	Left	-12	None/Little		
	-3	Curved/hilly	0	Middle	0	Some		
	-3	Exit	+2	Right	+12	Heavy		
	0	Lane Drop						
	+1	Guardrail						
	+10	Construction						
	+10	Crash						
Residential Mean=54	-10	Base					-6	Few
	-2	Some parking					-1	Some
	+1	Curved/hilly					+5	Many
	+4	Many parked cars						
	+5	Intersection						

How Well Do the Workload Ratings (of Clips) Agree with the Post-Test Ratings of Similar Situations?

Figure 54 shows that the ratings collected after the experiment was completed (post-test) and the workload ratings collected immediately after each clip were reasonably well related (r=0.75). This compares very favorably to the correlation obtained by comparing the ratings for each of the 23 situations in the clip-rating task with each other (r=0.76, Figure 55). Furthermore, keep in mind that post-test ratings were for descriptions of clips that were reasonably well matched, but not perfectly matched to 36 of the 46 clips that were shown in the clip-rating task.

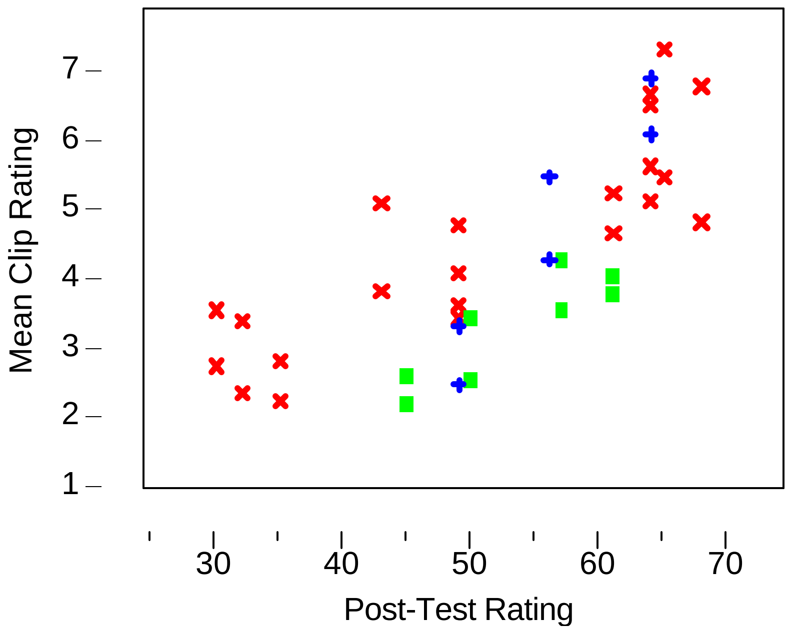


Figure 54. Correlation of Post-Test Rating with Mean Workload (Clip) Rating
Note: X=Expressway; box=Rural; +=Urban

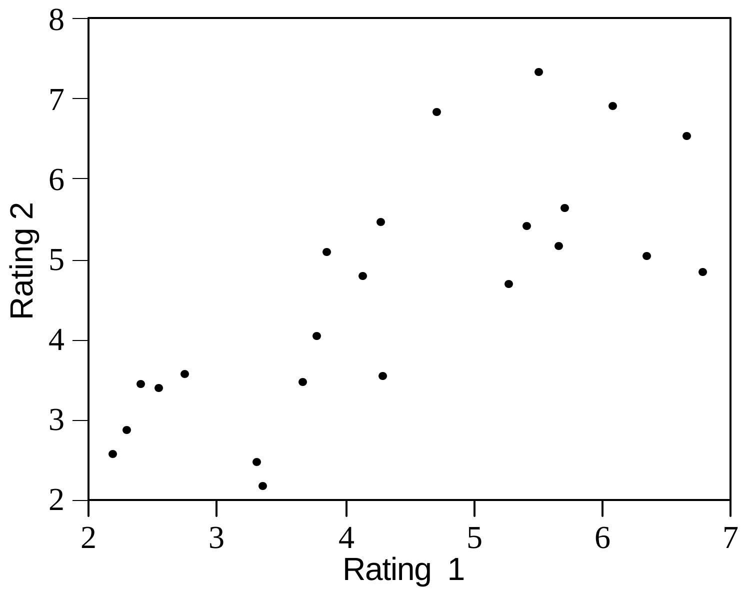


Figure 55. Correlation of First Clip Rating with Second Clip

What then, is the relationship between the post-test ratings and the clip ratings? According to linear regression, the mean clip rating = -0.58 + 0.94*(post-test rating), with r2=0.56 when all 36 data points are included. However, review of Figure 54 shows there are 3 different relationships, 1 for expressways, 1 for rural roads, and 1 for urban roads. Accordingly, the regression equations were computed for each road type. (See Table 41.) Notice that the intercepts and slopes differ considerably.

Table 41. Post-Test to Clip Rating Regression Equations

Road Type	Equation (Clip Rating =)	R2	# Data Points
Expressway	0.0012 +0.090*(post-test rating)	0.73	22
Rural	-2.13 + 0.10*(post-test rating)	0.76	8
Urban	-8.68 +0.24*(post-test rating)	0.89	6

How Does Rated Workload Vary with the Relative Position of Vehicles Ahead (Traffic) on an Expressway?

Pilot data, prior analyses, and the literature indicated that traffic was going to have a significant effect on workload, possibly more than any other factor. For that reason, traffic was examined in detail, though the effort was exploratory. The subject vehicle location was in the middle lane of a three-lane expressway. As a reminder, there was a vehicle in the left lane even with the subject vehicle, and then the position of a lead vehicle and a vehicle in the right lane at various distances from the subject, and the workload for each situation was rated on a 0 to 100 scale.

Figure 56 shows the distribution of ratings, whose mean was 61. Notice that there were a few cases where the maximum value was used, but overall, the ratings were widely distributed, and there were minimal problems with range limitations. In some discussions, with subjects, the question of how to deal with ceiling effects arose. "If some combinations of conditions leads to maximum workload, and conditions become worse, it is still the maximum."

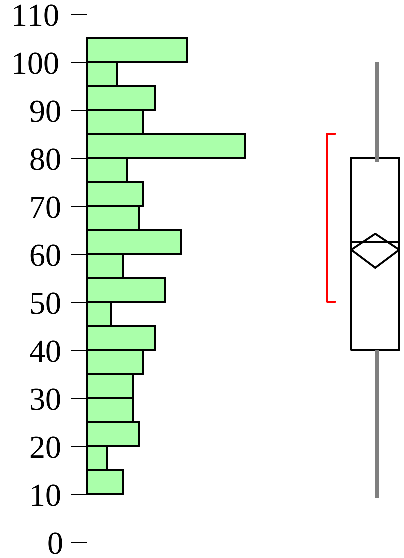


Figure 56. Distribution of Ratings of Demand for Expressway Traffic

An ANOVA was computed where Age group, Sex, Age * Sex interaction, Middle lane distance, Right lane distance, and the Middle * Right distance interaction were included. All factors were significant except of the Middle * Right interaction. Of those remaining all were at the $p<.0001$ level, except for Age * Sex ($p<.05$).

Figures 57 and 58 show the original data and a version where the log transforms of the middle lane distance are used, respectively. The rating clearly declines with distance in a log like manner as was found earlier, at least as well as can be determined by sets of 3 points.

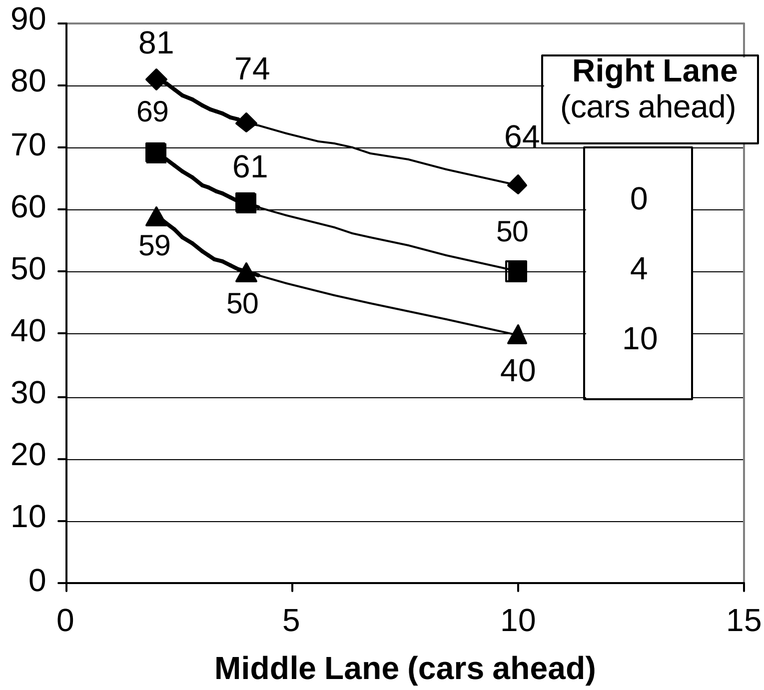


Figure 57. Rating of Traffic Demand Due to Location in the Lane
0, 4, and 10 refer to the number of car lengths ahead of the vehicle in the right lane.

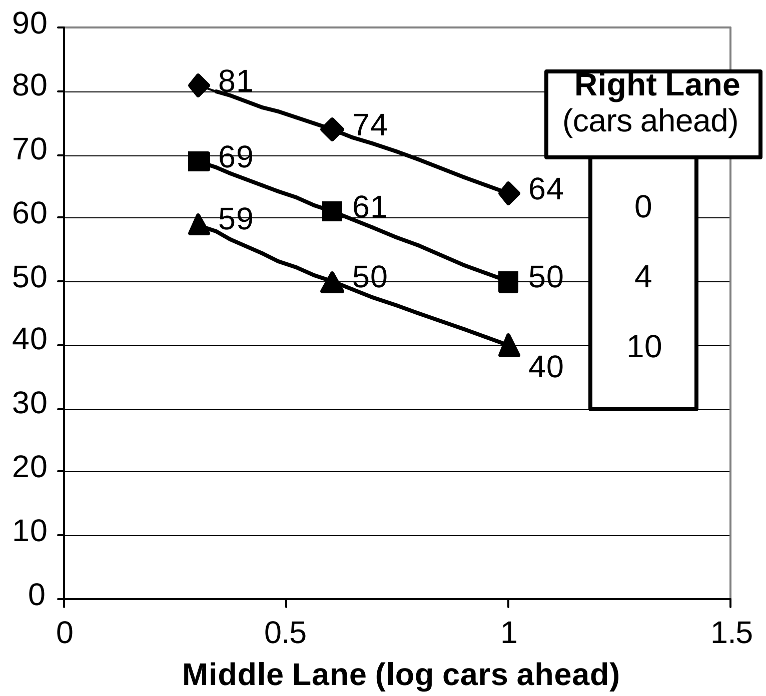


Figure 58. Rating of Traffic Demand Due to Location in the Lane (Log)
0, 4, and 10 refer to the number of car lengths ahead of the vehicle in the right lane.

In a simple linear regression, ignoring age and sex effects, led to the following equation:

Rating of traffic demand – 89.0 -25.9 (Log middle ahead) -2.4 (Right ahead)

Where:
Log middle ahead = log of the distance to the vehicle in the same lane in car lengths
Right ahead = distance to vehicle in right lane in car lengths

The r^2 was quite low, only 0.21, though their inclusion in the model was very highly significant. Much of what was not accounted for are individual differences, some of which could be adjusted for using age and sex corrections suggested by the means for each group.

Thus, these data show that distance effects are logarithmic, and at least for 2 vehicles, their demands are additive. However, looking at the range remaining, the addition of a 3rd vehicle in close should be very close to a workload of 100, the scale maximum, which makes sense. It is unknown how precisely these post-test ratings match experienced demands as there was not the opportunity to check such in this project.

It seems reasonable that subjects may feel very differently about workload when a vehicle was very close to them and in their lane, a condition not explored. It may be that asking subjects to imagine the situation without a reference image is too remote from the actual driving situation.

What is the Relative Contribution of Road Geometry, Road Surface Condition, Visibility and Lighting, and Traffic to Ratings of Total Workload?

Nygren (1995) suggests that workload is determined by 5 factors: traffic, ambient lighting, road geometry, visibility, and traction. In his conjoint analysis of the workload truck drivers perceive, those factors accounted from 52%, 26%, 13%, 6% and 3% respectively. To some extent, ambient lighting and visibility are inseparable as they both determine what the driver can see, and for that reason they were combined in the analysis that follows.

How could one consider all of those factors in a workload estimate? Because the ACAS images were recorded in black and white and of limited resolution, important cues at night (such as distinguishing white headlights from red taillights) were not visible. It was therefore not feasible to reliably assess the effects of visibility limitations from the ACAS tapes. Furthermore, since watching the tapes was a passive activity, the cues needed to assess traction were not provided, assessing them directly also did not make sense. However, some sense of their contribution was needed to develop a workload manager. To get a sense of those effects and make the predictions developed more generalizeable, subjects were asked to estimate the relative importance of those factors, in many ways an extension of Nygren's work.

Specifically, subjects assigned a weight (a number between 1 and 10, with the sum of all factors being 10) factors for the expressway, residential, rural, and urban road types.

To get an initial impression of subject differences, an ANOVA was run for each of the 4 factors (separately), with Age Group, Sex, Age Group*Sex, Subject nested in Age Group*Sex, and Road Type and the factors. Road type was significant for Road Geometry, Traffic, but not Visibility or Road Surface Condition, reflecting that some road types, such as rural, are highly variable in their geometry, whereas expressway have good geometry that varies little. Other than subjects in 2 cases, there were no other important statistically significant differences. In part, that may be because a few subjects did not think of workload as being additive. To paraphrase, “if it is really slippery, I cannot drive, so I am going to give almost all of the points to road surface condition.”

Although there are more elegant methods that consider the nonindependence of these ANOVAs, the key point is the means change very little due to individual differences, that is, they represent consistent and generalizeable factors that contribute to workload.

Therefore, it is appropriate to focus on the means in Tables 42 and 43. Notice these data do not match those of Nygren, where he found traction to be much more important and traffic to be much less important. In part, that may be because of the data collection method used here, where the images varied primarily in terms of road geometry and traffic, causing subjects to focus on those factors.

Table 42. Factor Weight Mean by Age Group

Factor	Age Group			Mean
	Young	Middle	Old	
Road Geometry-includes lane width, curvature, hills, intersections, merging & turn lanes	1.9	1.7	1.5	1.7
Road Surface-from dry to wet or icy, also includes road roughness, tire condition and vehicle factors that affect braking and handling	3.0	2.8	2.6	2.8
Visibility-how well you can see-determined by rain, snow, or fog, windshield condition, mirror design	2.3	2.4	3.6	2.8
Traffic-number of vehicles in your lane, adjacent lanes, oncoming, merging and intersecting, also includes pedestrians and bicyclists	2.8	3.1	2.9	2.9

Table 43. Relative Factor Importance

Factor	Road Type					
	Xway	Rural Road	Residential Street	Urban Street	Mean	%
Road geometry	1.3	2.1	1.7	1.6	1.7	17
Road surface condition	2.8	3.1	2.7	2.7	2.8	28
Visibility	2.5	3.4	2.8	2.5	2.8	28
Traffic	3.4	2.3	2.8	3.2	2.9	29

This provides some perspective to the overall experimental work. The video clip ratings focused on most important factor, namely traffic, as well as road geometry. Unfortunately, because of limited image quality, visibility could not be accurately assessed. However, the occlusion experiment from the first phase provides some relative sense of how sight distance contributes to workload. Road surface condition is much more difficult. This assessment requires either an extremely high quality moving base driving simulator or a test track with controllable road surface conditions or a special vehicle that can simulate varying surface conditions, all expensive requirements.

How Does the Probability a Driver Is Willing to Do a Task while Driving (Tune a Radio, Dial a Phone, Enter a Destination) Vary with Rated Workload, Road Geometry, and Traffic, and with Driver Age and Sex?

In addition to rating workload for each clip, subjects stated if they would (1) manually tune a car radio, (2) dial a 10-digit phone number, and (3) enter an address into a navigation system. Except for 8 responses, ratings were to the nearest 0.5 on the workload scale. To simplify the analysis, those 8 values were rounded off to the nearest 0.5.

Figure 59 shows the overall probability drivers would not do each of the 3 in-vehicle tasks (p(no)) averaged for each road type. Error bars are not shown to avoid clutter, but readers should keep in mind that that each of the 24 subjects saw 4 instances of each of the 23 scenarios, for a total of 2,208 responses. Though not distributed uniformly in Figure 59, each data point on average represents 85 yes/no decisions, so the probability estimates are reasonably stable. Differences among road types were of no practical consequence, and the function has a slight curvature. Notice that even at extremely low workload levels, 10-20% of all responses indicated some of these in-vehicle tasks should not be performed, and, similarly, at extremely high workload, some of these tasks were deemed safe to do, but only by very few drivers.

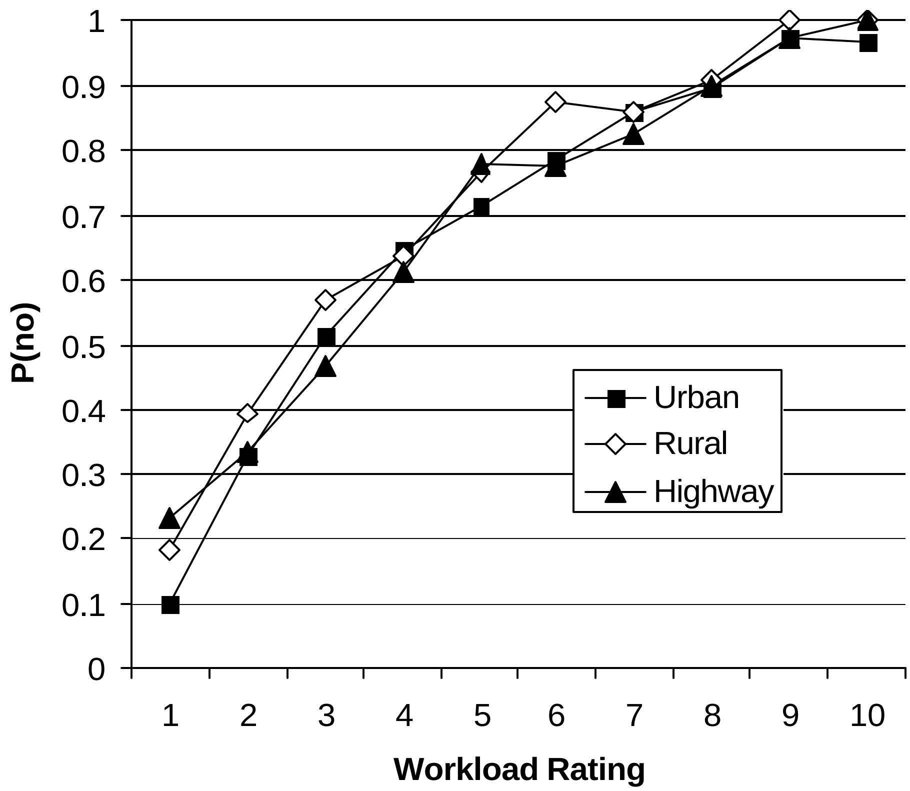


Figure 59. P(no) vs. Rating by Road Type for all Devices

Figure 60 shows the p(no) data partitioned by task. Manually tuning a radio was more acceptable (in terms of safety) than dialing a phone, which was more acceptable than entering a destination. Tuning a radio becomes unacceptable to 90% of the participants at workload of about 8.5, dialing a long distance phone number at about 7.5, and entering a street address at 3. To provide perspective, 2 was the workload for the light traffic anchor clip (an expressway with LOS A) and 6 was workload for the anchor clip for heavy traffic (an expressway with LOS E). Note that even at the lowest level of workload, approximately 45% of the sample (of 24 people) said destination entry should not be performed. This provides strong support for locking out destination entry under all circumstances when the vehicle is moving.

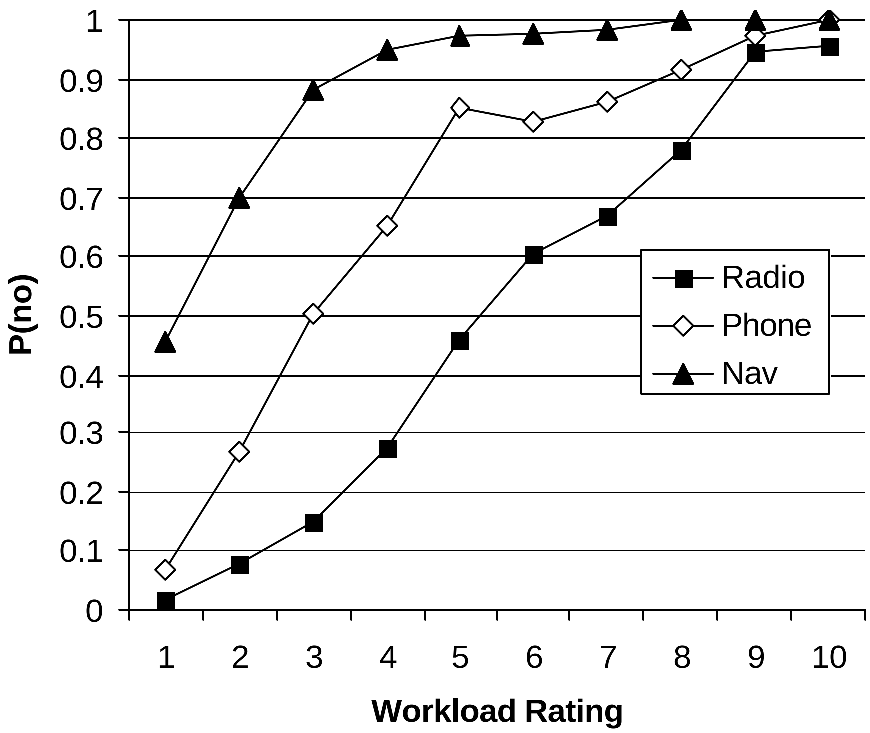


Figure 60. P(no) vs. Rating by Device for all Road Types

Figure 61 shows there were some differences due to age, with older drivers being more likely to say no than younger drivers (by 15-20%), and, surprisingly, younger drivers being more likely to say no than middle-age drivers by about 5-10%. Given they tend to take more risks, it was expected that p(no) would be lowest for younger drivers. Keep in mind, however, that there were only 8 subjects in each age group. For additional details, showing differences due to task and age group for each type of road, see Appendix G.

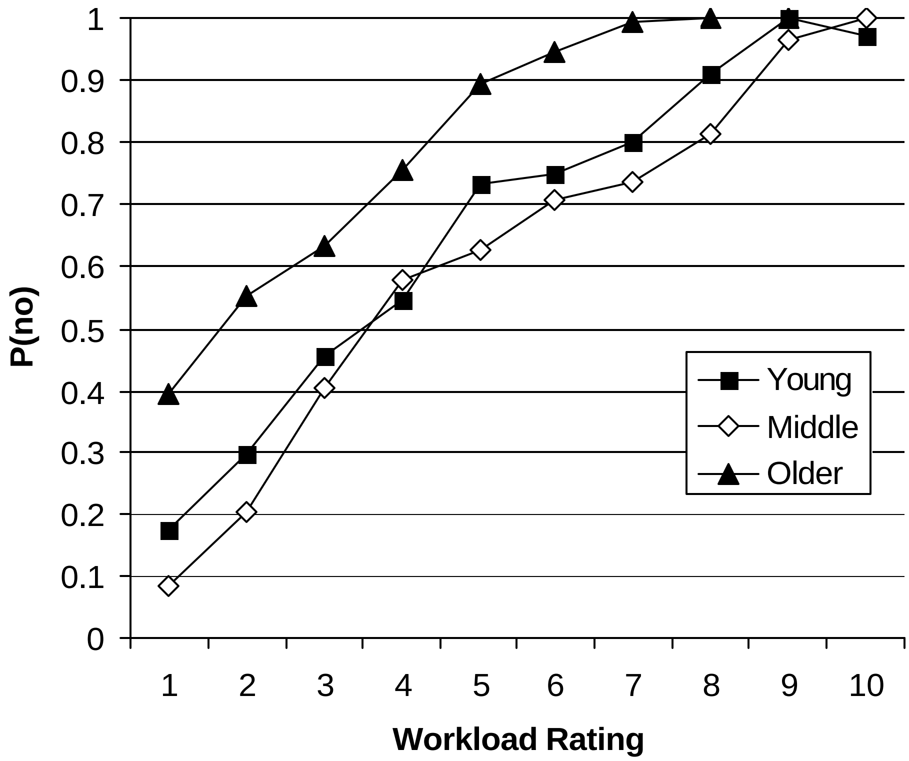


Figure 61. P(no) vs. Rating by Age Group for all Road Types and all Devices

Figure 62 shows there were small differences due to LOS independent of the workload rating. In theory, greater traffic volume should lead to a larger workload value, so there should be no independent effect of LOS. Furthermore, keep in mind that each LOS category is a range of traffic volumes, not a point. However, there are limits to that range, with LOS A associated with low workload ratings and LOS E with high workload ratings. By definition, LOS values should not be associated with the full range of workloads, though surprisingly, some of the LOS A situations had a workload of 8 and some of the LOS E situations were rated as 2. Had the effect of LOS been fully included in workload ratings, the curves for the 3 LOS levels would have been completely superimposed on one another. As it turned out, for any given workload, p(no) was slightly less for LOS E than for LOS C which in turn was slightly less than for LOS A (by about 0.1 in many cases). At this point, it is unclear why this occurred, though there could be confounding of LOS levels with other factors. Thus, a workload manager, to accurately predict driver preferences, should include both workload estimates and LOS when determining which tasks drivers should not do. As is shown later, using actual values for traffic volume leads to very accurate workload estimates.

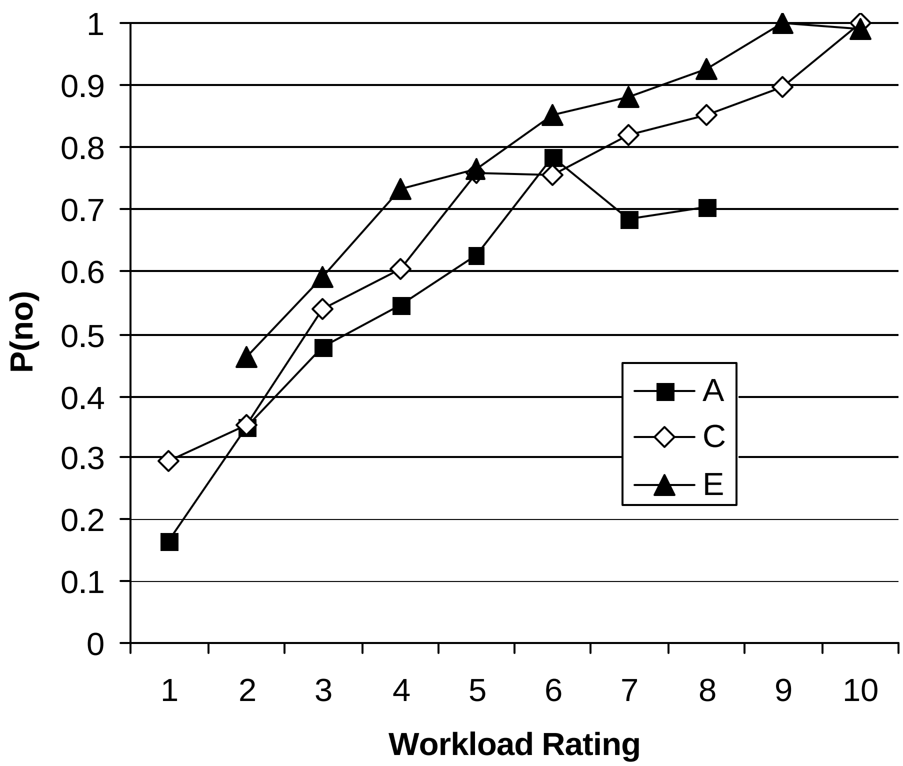


Figure 62. P(no) vs. Rating by LOS for all Road Types and all Devices

For each of the road types, at least one factor in addition to LOS and those related to subjects was considered. For rural roads, that factor was curvature and it had no effect on p(no) independent of workload (Figure 63). As was noted earlier, the effect of curvature on workload was only evident for LOS A.

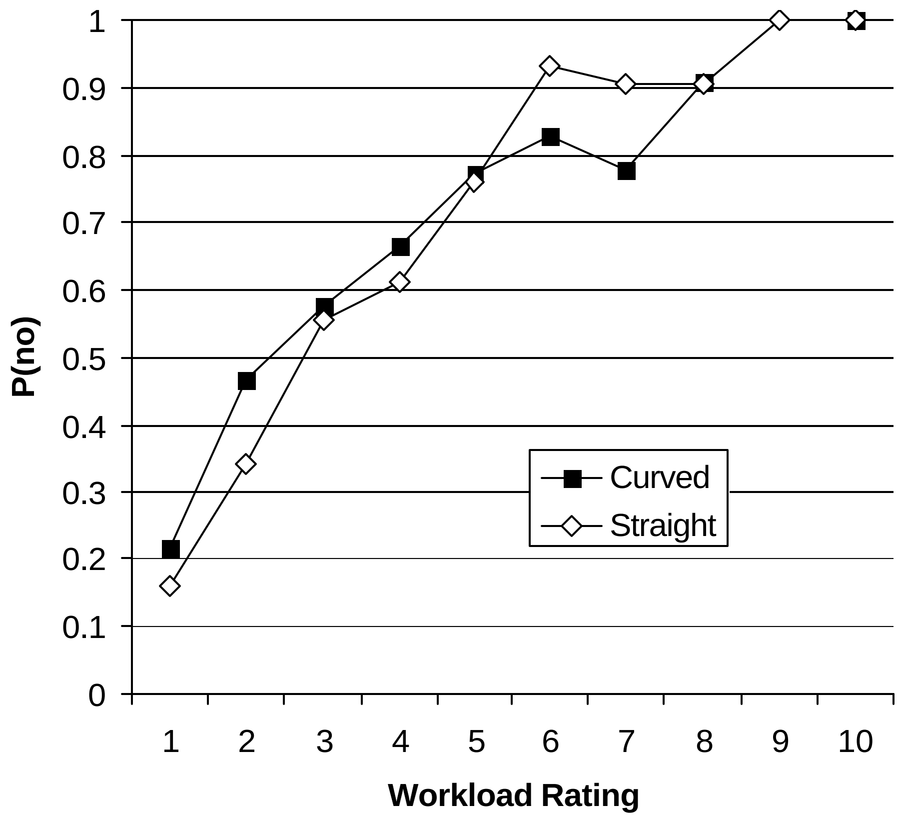


Figure 63. P(no) vs. Rating for Rural Roads – Straight vs. Curve Roads

For urban streets, that factor was the presence or absence of intersections. As shown in Figure 64, the effect of the intersection on the decision of which tasks to perform was included in the workload rating. Keep in mind that this experiment was a weak test of the effects of intersections as the signal color was not visible and the field of view was narrow, so the effect of intersections may have been underrated.

Figure 64. P(no) vs. Rating for Urban Streets – Intersections vs. Non-Intersections

Also examined for urban roads was the number of lanes (4 vs. 5, Figure 65), which seemed to have no effect independent of workload. This factor was not balanced and apparently not significant, so it is not included in other analyses or figures.

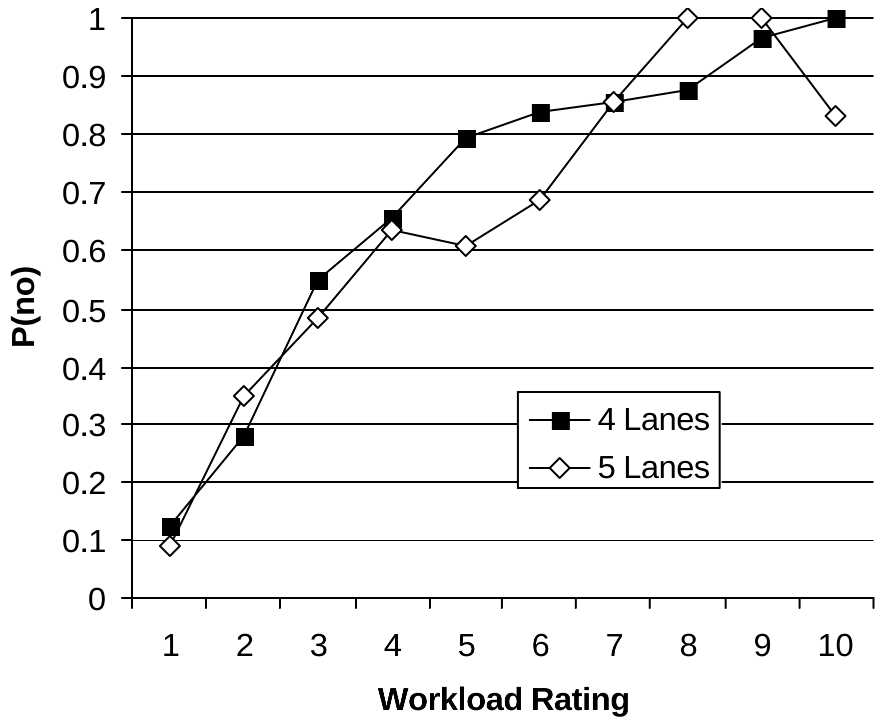


Figure 65. P(no) vs. Rating – 4 Lanes vs. 5 Lanes

Finally, for expressways, Figure 66 shows the effect of the lane driven (and merging) on p(no). This factor too seems to show few differences once workload is considered, with 1 exception, right merge, where p(no) was about 10 percent greater than for other lanes (i.e., subjects were 10 percent more likely to say no).

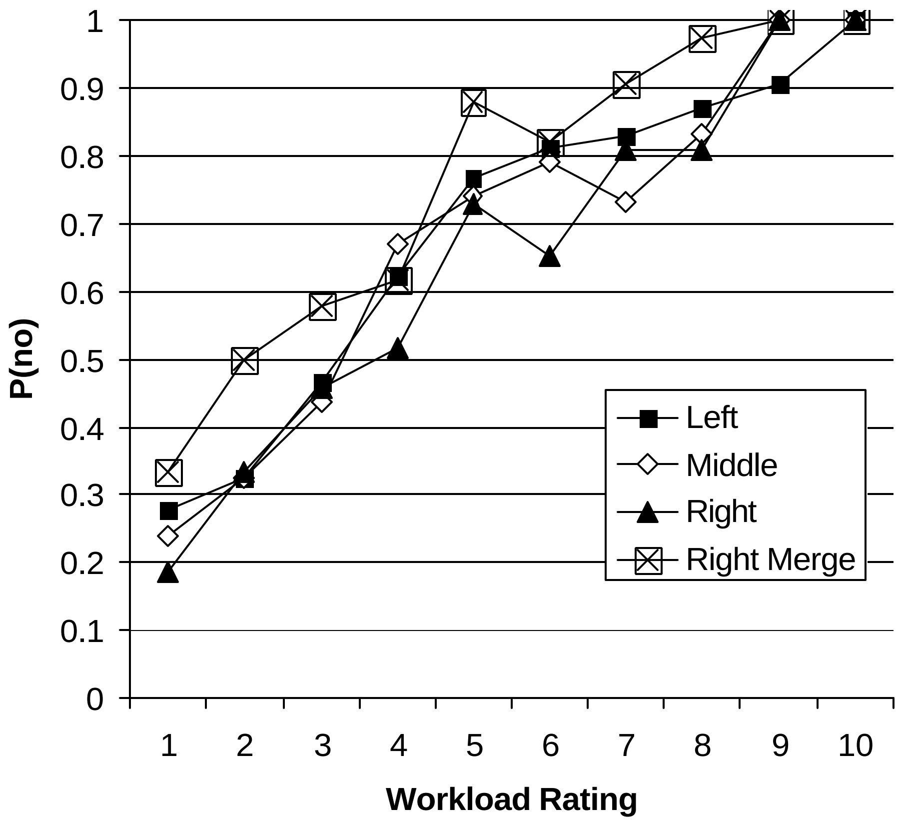


Figure 66. P(no) vs. Rating for Expressways by Lane

Thus, as a whole, these data show that in most cases, p(no), the probability that subjects say a task should not be undertaken, depends primarily on the workload, the driver age, and the task. Small adjustments are needed for the case of a merging vehicle on the right and for LOS A (both about 10% increases).

As a next step in the analysis, logistic regression was used to determine the relationship between p(no) and rated workload. As a reminder, logistic regression is the most appropriate form of regression analysis when the dependent variable (would drivers engage in a task, yes or no) is binary. In logistic regression p(no) =1/(1+e^-(ax+b)), where x=workload rating, and a and b are the slope and intercept of the regression function respectively.

Using logistic regression, the slope and intercept for each subject for each task was determined (Appendix H). If the subject always said no, then in theory the slope would be infinite and the intercept zero, and in those cases a logistic regression could not be computed for that individual. Therefore, a workaround was used to compute age * sex group means needed for a workload manager. For mean slopes, where no slope was provided by the logistic regression for a particular subject, the slope for that subject was set to zero and the mean was computed. For the mean slope, the inverses of the slopes for each subject were used. (The inverse of infinite slope is zero.) Results are shown in Table 44.

Table 44. Intercepts and Slopes for Each Age * Sex * Device Combination

(Cells where inverses were used are shown in bold)

Age	Sex	Radio		Phone		Navigation	
		Intercept	**Slope**	**Intercept**	**Slope**	**Intercept**	**Slope**
Young	Male	-83.19	11.33	-8.08	2.87	-5.18	**5.40**
Young	Female	-12.04	1.85	-10.59	2.37	-13.34	**1.54**
Middle	Male	-18.11	2.85	-6.57	**2.12**	-3.66	**1.79**
Middle	Female	-3.28	1.63	-12.47	4.41	-8.68	**2.14**
Older	Male	-6.45	1.53	-4.28	1.84	-3.78	**3.29**
Older	Female	-5.35	1.24	-10.23	4.48	-0.08	**2.12**
Mean		-21.40	3.40	-8.70	3.02	-5.79	**2.71**

As an aside, there are many other ways the data could have been adjusted to compute slopes and intercepts. For example, in situations where subjects would never do any task, the response for 1 case for the highest workload rated by that subject could be changed from no to yes. However, to balance that change, one could also argue for changing a similar response for another subject in the same Age Group * Sex group for the same device from yes to no. Given the variety of options and the limited sample size, the inverse approach is as reasonable as any. If greater precision is desired, a larger sample size is needed. For this initial investigation, the sample size was sufficient.

What is apparent from this table, but not from preliminary figures, are hints of an Age Group * Sex interaction for younger drivers. Keep in mind that each Age Group * Sex group is only 4 drivers, and that a single outlier, as may be the case for the young men, could create the impression of an interaction.

In practice, p(no) could be determined from the figures presented earlier, using that figure to create a lookup table for each device and age, or it could be computed here using Table 44 and the logistic equation.

CONCLUSIONS

How repeatable are the workload ratings within and between drivers?

The workload ratings were very consistent. Since each subject saw 2 clips for each scenario and each was seen twice, reliability can be computed several ways. Across clips with the correlation of the first rating of each driving scenario (averaged across subjects and the 2 instances) with the second was 0.76.

The mean absolute difference in ratings between the first and second time a subject saw a clip varied with the road type. For rural roads, the difference was typically 0.2 to 0.3, quite small considering ratings were usually given to the nearest 0.5. (As an aside, there were no comments that subjects remembered seeing a clip before, and therefore used the same rating.) When the comparison was for 2 clips representing each situation (e.g., a 2-lane straight rural road with LOS A, again within subject) the mean differences were larger, 0.3 to 0.9, though relative to the measurement accuracy, still small.

For urban roads, the mean differences between repetitions were larger (0.2 to 0.5) and the mean differences with clip type were considerably larger (0.8 to 2.1). For expressways, the differences between repetitions were less on average than those for urban streets (0.2 to 0.6) and the mean difference within clip type was clearly less (0.3 to 1.8). Interestingly, the mean workloads for expressways were close to those for urban streets.

How do workload clip ratings vary overall?

The mean workload ratings varied from 1 to 10 with a mean of 4.3 on a 1 to 10 scale. As shown in Figure 67, ratings were well distributed throughout the range.

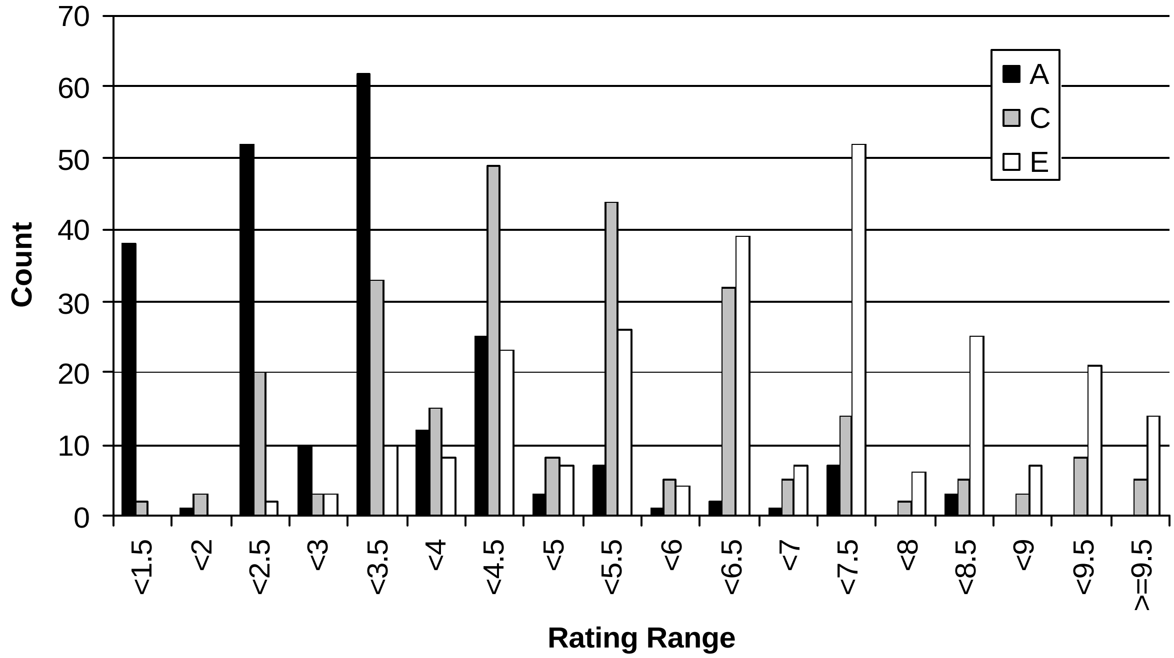


Figure 67. Workload Ratings (of Clips) Distribution by LOS

What is the relationship between workload ratings (of clips) of driving situations and (1) road type (e.g., urban), (2) road geometry, (3) lane driven, (4) traffic volume (as measured by LOS), (5) driver age, and (6) driver sex?

The factors that affected workload ratings varied somewhat with the type of road. However, fairly consistently, LOS, Age Group, Age Group * Sex, and LOS * Age Group were statistically significant. For expressways, the only road for which it was assessed, there were significant differences due to the lane driven, but interestingly there were not significant differences due to intersections for urban roads (because they were not salient in the monochromatic images) and curves on rural roads.

Table 45 shows that workload ratings increased with LOS (about 2.8 for A, 4.5 for LOS C, and 6.0 for LOS E), the largest single effect. Interestingly, mean ratings for young (4.3) and old (4.2) subjects were almost the same, though ratings for middle aged subjects were slightly greater on average (4.6). The authors have no explanation why this occurred. In terms of road types, the mean values were 4.0 for rural roads, 4.3 for expressways, and 4.8 for urban roads. There were no overall sex differences or differences due to intersections worthy of note. For intersections, this was probably because the traffic lights were not very evident in the clips. For expressways, the lane means were 4.7 for the left, 4.3 for the middle, and 4.0 for the right (ignoring merge situations). Thus, in terms for the primary factors, the range was 3.2 for traffic, 0.8 for road type, 0.7 for lane (on expressways), and 0.5 for age, so traffic is by far the most important factor of those examined.

Table 45. Mean Workload Ratings of Clips for Various Roads

Age Group	LOS	Rural		Urban		Expressway		LOS Mean	Age Mean
Young	A	2.6		2.6		2.4		2.6	
	C	4.1	4.2	5.2	4.7	3.7	3.9	4.3	4.3
	E	5.9		6.2		5.7		5.9	
Middle	A	2.8		2.7		3.0		2.8	
	C	4.0	4.2	5.4	4.9	4.5	4.6	4.7	4.6
	E	5.8		6.6		6.4		6.3	
Older	A	2.6		3.1		3.2		3.0	
	C	3.6	3.7	5.2	4.7	4.3	4.4	4.3	4.2
	E	5.0		5.7		5.7		5.4	
Mean			4.0		4.8		4.3		4.4

What is the relationship between workload ratings (based on the post-test data) road characteristics, traffic, and driver characteristics?

As shown in Table 46, the post-test workload ratings indicated that urban situations were associated with the highest workload, expressways second, rural roads third and residential roads fourth, though there were many exceptions.

Table 46. Post-Test Workload Ratings Ranked by Road Situation

Road Situation	Workload Rating
Urban, Downtown	72
Expressway, with Crash Scene	71
Expressway, with Construction	70
Rural, Very Curved or Hilly	69
Expressway, with Lane Drop	64
Expressway, with 3-foot shoulder	62
Urban, w/ Commercial Building on Corner	60
Rural, w/ Stop Sign for Cross traffic	60
Residential, w/ Signaled Intersection	59
Rural, w/ 1-foot Shoulder	58
Expressway, Curved or Hilly	58
Expressway, with Interchange	58
Rural, w/ Signaled Intersection	58
Residential, w/ 25+% Parked Cars	58
Residential, Curved or Hilly	55
Rural, Curved or Hilly	54
Residential, w/ 0-25% Parked Cars	51

Of the various road modifiers (Table 47), crashes and construction added most to the workload ratings, being about 50% of the total workload for the lowest workload situations. The value of the modifiers varies slightly between road types.

Table 47. Overall Rank Order of Road Modifiers

Road Modifier	Rating
Crash Scene	23
Construction	22
Very Curved or Hilly	19
Downtown	16
Lane Drop	15
Signaled Intersection	15
3-foot shoulder	14
25+% Parked Cars	14
Curved or Hilly	10
Stop Sign for Cross traffic	10
Interchange	10
1-foot Shoulder	9
Signaled Intersection	9
0-25% Parked Cars	7
Curved or Hilly	5
Commercial Building on Corner	4

How Can Workload Ratings Be Estimated Using Mean Ratings for Clips?

One of the ways to estimate workload is from a description of the driving situation. If a vehicle is fitted with ACC and navigation systems, the vehicle should know the type of road being driven, if an intersection was being approached, the lane driven (all from the navigation system), and the distance to vehicles ahead (from the ACC radar). If the vehicle was also fitted with a driver personality module, it would know the driver's age and sex. Given the expected implementation of workload managers, it is unlikely, at least initially, that workload managers would be implemented in vehicles without both GPS navigation and ACC systems. However, currently, personality modules with driver specific data of this type are uncommon, so it makes sense to consider 2 different workload manager implementations.

Consistent with the logic, tables follow that provide mean workload rating for these 2 implementations (1) road and traffic data only, and (2) that data plus driver characteristics. Because there were different features for each road type, 1 table was created for each class (Tables 48, 49, 50), with each table having 2 parts, 1 part for road and traffic data only, 1 that includes subject data. Cells entries are provided only for factor combinations that were significant in the ANOVA. So, for example, for rural roads, of the road and traffic factors, only LOS was statistically significant, so only those means are provided.

Table 48. Mean Workload Ratings for Rural Roads

LOS	Geometry	Geometry, Traffic, and Subject Data						Road & Traffic Only
		Young		Middle		Old		
		Female	Male	Female	Male	Female	Male	
A	Straight	1.9	2.3	3.0	2.0	2.4	2.7	2.7
	Curved	2.9	3.4	3.9	2.2	2.5	2.9	
C	Straight	4.4	3.9	4.0	4.0	3.2	4.0	3.9
	Curved	4.0	4.1	4.3	3.8	3.0	4.1	
E	Straight	6.7	5.4	6.0	5.8	4.5	5.6	5.5
	Curved	6.2	5.1	6.1	5.4	4.2	5.6	

Table 49. Mean Workload Ratings for Urban Roads

LOS	Intersection	Geometry, Traffic, and Subject Data						Road & Traffic Only
		Young		Middle		Old		
		Female	Male	Female	Male	Female	Male	
A	No	3.0	2.8	3.3	2.8	3.1	2.5	2.9
	Yes	2.8	2.4	3.5	3.0	2.8	2.1	2.8
C	No	5.0	5.2	4.7	5.1	5.1	4.2	4.9
	Yes	6.1	5.4	5.0	6.1	6.8	4.7	5.7
E	No	6.7	7.0	5.7	6.5	6.7	6.4	6.5
	Yes	6.3	6.3	4.7	5.8	6.3	5.3	5.8

Table 50. Mean Workload Ratings for Expressways

LOS	Lane	Geometry, Traffic, and Subject Data						Road & Traffic Only
		Young		Middle		Old		
		Female	Male	Female	Male	Female	Male	
A	Left	2.8	2.3	3.2	3.5	3.8	3.4	3.2
	Middle	2.8	2.4	2.7	3.1	3.3	3.3	2.9
	Right	2.3	2.0	2.9	2.8	3.1	2.4	2.6
C	Left	4.5	3.6	4.6	4.8	4.5	4.8	4.5
	Middle	4.1	3.7	4.8	5.0	4.5	4.8	4.5
	Right	3.1	3.3	4.0	3.8	3.9	3.3	3.6
	Right Merge	5.2	4.5	5.8	4.9	4.4	5.1	5.0
E	Left	6.6	6.1	7.3	6.9	5.7	7.1	6.6
	Middle	5.6	4.8	6.2	5.5	4.9	5.5	5.4
	Right	5.8	5.5	6.8	6.0	5.0	5.9	5.8
	Right Merge	6.6	5.8	6.9	6.7	5.9	6.6	6.4

The tables could be stored by a workload manager and used to look up the workload for any situation. For example, if a older male was driving on an expressway in the right lane in LOS E, and all the subject and road and traffic data were available, then the workload would be estimated to be 5.9. If the subject data were not available, then the workload would be estimated to be 5.8.

How Can Workload Be Estimated Using the Post-Test Ratings?

The post-test data provides more detail regarding the driving situation than the clip ratings, but because it is less well anchored, is probably less accurate. One could store the table that follows (Table 51) which contains means and adjustment factors for each situation. So, for the example of driving in the right lane of lane of an expressway in heavy traffic, the estimate would be 61 (expressway mean) -13 (base case) +2 (right lane) + 12 (heavy traffic) = 62.

Table 51. Post-Test Workload Estimate

Road Type & Mean	Road Modifier		Lane		Traffic		Driveways	
Rural Mean=58	-8	Base case	-1	2 Lanes	-5	None/ Little		
	-3	Gentle curve/hill	1	3 Lanes (in left)	+5	Some		
	-3	1-ft shoulder	+2	4 Lanes (in left)				
	+1	At, approach light						
	+2	Stop sign for others						
	+11	Very hilly, curved						
Urban Mean=63	-7	Base case	-3	2 Lanes	-6	None/ Little		
	-3	Corner business	-2	3 Lanes	-3	Some		
	+9	Downtown	+0	4 Lanes	+9	Heavy		
			+4	>=5 Lanes				
Xway Mean=61	-13	Base case	-1	Left	-12	None/ Little		
	-3	Curved/hilly	0	Middle	0	Some		
	-3	Exit	+2	Right	+12	Heavy		
	0	Lane Drop						
	+1	Guardrail						
	+10	Construction						
	+10	Crash						
Residential Mean=54	-10	Base					-6	Few
	-2	Some parking					-1	Some
	+1	Curved/hilly					+5	Many
	+4	Many parked cars						
	+5	Intersection						

If the driver's age and sex were known (the personality module case), then adjustments for age and sex (based on the age-sex means given in the results) shown in Table 52 could be used. For an older male, the adjustment would be +4, so the estimated workload rating would be 66.

Table 52. Post-Test Age-Sex Adjustments

Sex	Age		
	Young	Middle	Old
Male	-14	+8	+3
Female	-9	+10	+4

An alternative would be have the workload manager store the entire table of the mean post-test ratings for each situation and for each type of road (Tables 53, 54, 55, 56). Keep in mind that each data point represents 1 rating from each of the 24 subjects in the experiment, a reasonably stable value. However, if a personality module was available, then data by age and sex would be desired, which in this case would be a mean from groups of 4 subjects, a very small sample. Again, should such adjustments be desired, they should be based on the age and sex means in the previous table.

Table 53. Mean Post-Test Workload Ratings for Rural Roads

Two values are for: (a) no or little traffic / (b) some traffic.

a / b

Only 2 values were considered since heavy traffic is rare on rural roads.

Situation	Total # Lanes			Mean
	2	3 (Center Pass/Turn Lane)	4 (in Left Lane)	
Base case=straight road 8 foot paved shoulder + 8 foot grass beyond that	40 / 54	44 / 56	45 / 57	43 / 56
Base case except gentle curves or hill	47 / 59	49 / 60	50 / 61	49 / 60
Base case with 1-foot shoulder, mailboxes, rocks, vegetation beyond	53 / 62	53 / 64	54 / 64	53 / 63
Base case + at or approaching intersection with traffic light	51 / 62	52 / 63	55 / 64	53 / 63
Base case + at or approaching intersection with a stop sign for the crossing road only	53 / 62	54 / 65	55 / 67	54 / 65
Base case except very curved or hilly road (mountain road)	64 / 74	65 / 74	63 / 74	64 / 74
Mean	51 / 62	53 / 64	54 / 65	53 / 63

Table 54. Mean Post-Test Ratings for Urban Streets

Three values are for: (a) no or little traffic / (b) some traffic / (c) heavy traffic

a / b / c

Situation	# Lanes				Mean
	2	**3 (Center Turn)**	**4 (Includes Turn Lane)**	**5 or More**	
Base case=straight rd, cars parked on side, 10 intersect/mi, most with lights, no or few pedestrians, no stores	45/53/63	47/54/63	49/56/64	52/61/70	48 / 56 / 65
Base case but stores or gas station on corner	49/57/67	51/58/67	52/59/68	56/63/73	52 / 59 / 69
Base case but numerous stores & pedestrians ("downtown"), midblock driveways, no double parking	62/69/76	64/71/78	65/73/81	70/76/84	65 / 72 / 80
Mean	52/60/69	54/61/69	55/63/71	59/67/76	55 / 63 / 71

Table 55. Mean Post-Test Ratings for Expressways

Three values are for: (a) no or little traffic / (b) some traffic / (c) heavy traffic

a / b / c

Situation	**Total # Lanes (So 6=3 per Direction)**			**Mean**
	6 (in Left Lane)	**6 (in Middle Lane)**	**6 (in Right Lane)**	
Base case = straight road, 1-lane paved shoulder on each side, wide grassy median, no guardrails needed	30 / 43 / 63	32 / 49 / 64	35 / 49 / 68	32 / 47 / 65
Base case+ Curved or hilly	45 / 58 / 72	45 / 59 / 70	46 / 59 / 71	45 / 59 / 71
Base case + Interchange (entrance/exit) in view or at it	40 / 54 / 72	44 / 56 / 73	48 / 61 / 75	44 / 57 / 73
Base case + Lane drop (e.g., 3 to 2 lanes) in your or adjacent lane	50 / 58 / 74	46 / 60 / 73	51 / 62 / 75	49 / 60 / 74
Base case but 3-foot shoulder & guardrail instead	49 / 61 / 74	47 / 61 / 73	51 / 63 / 79	49 / 62 / 75
Base case + Construction: Approaching or driving in lane shift or narrow lanes with concrete barriers, no shoulder	59 / 69 / 80	60 / 71 / 80	61 / 72 / 82	60 / 70 / 81
Base case + Approach or driving through crash scene	62 / 69 / 80	61 / 71 / 81	63 / 70 / 81	62 / 70 / 81
Mean	47 / 58 / 73	47 / 61 / 74	51 / 62 / 76	49 / 61 / 74

Table 56. Mean Post-Test Ratings for Residential/Suburban Streets
Since suburban streets rarely have traffic, only no or little traffic was considered.

Situation	**Driveways (per Side of the Road)**			**Mean**
	0-<2 / Block (0.1 miles)	**2-5 / Block**	**> 5 / Block**	
Base case, straight road, no parked cars, no intersection nearby	38	44	50	44
Base case, but >0 - 25% of curb has parked cars	46	51	58	52
Base case, but curved or hilly	50	54	60	55
Base case, but >25% of curb has parked cars	52	58	64	58
Base case, but at or approaching signed intersection, where you need to stop	55	59	64	59
Mean	48	53	59	54

By way of comparison, the mean workload for the expressway base case for driving in the right lane in heavy traffic is 68, versus 62 estimated from the means, just under a 10 percent difference.

What is the Relationship between Ratings of Workload of Clips of Driving and Post-Test Ratings of Workload?

Since there are 2 sets of ratings, they both could be used by a workload manager. Visual inspection of correlation plots of the 2 sets of measurements suggests the relationship between the 2 ratings is linear. However, careful inspection suggests that a different linear relationship might exist for each type of road, thought that judgment is based on a very small number of data points. Accordingly, linear regression was used to compute equations to relate the 2 sets of data (Table 57), both for each road type and overall. As a reminder, there is no equation for residential roads because there were no clips for them that were rated. (There were too few of them in the original ACAS data set.) Notice that the intercepts and slopes differ considerably due to road type.

Table 57. Post-Test to Clip Rating Regression Equations

Road Type	**Equation (Clip Rating =)**	**R2**	**# Data Points**
All Roads	-0.58 + 0.94*(post-test rating)	0.56	36
Expressway	0.0012 +0.090*(post-test rating)	0.73	22
Rural	-2.13 + 0.10*(post-test rating)	0.76	8
Urban	-8.68 +0.24*(post-test rating)	0.89	6

Using example of driving on an expressway in the right lane in heavy traffic (assumed to be LOS E) for an older driver, the post-test mean workload rating was 68 (using Table

51 and the age-sex adjustment in Table 52) and estimated to be 62 just using adjustments (Table 51). Using the regression equations in Table 57, the clip rating is estimated to be 6.1 and 5.6 respectively. The clip rating was actually 5.9, midway between the 2 estimates and quite close to each, especially since clip rating estimates were to the nearest 0.5, and repeated ratings of the sample clip varied by that much, well within the limits of measurement error. Furthermore, keep in mind that neither the clip ratings nor the post-test ratings are the true value of workload. They are just estimates.

How can workload ratings be estimated using the driving performance statistics developed from the ACAS FOT data set?

An alternative method to estimate workload is to continually measure driving performance and use those data to predict workload ratings. Three predictive equations are suggested. For the first, the entire data set (46 cases = 2 ratings for 23 situations, sometimes reduced to 42 cases because right merge was initially excluded) was used in the equation, with the entry criterion being extremely strict, $p<.0001$. When no lead vehicle was present, at least detected by the radar, a vehicle was assumed present at the maximum range of the radar, 125. Using that assumption and the .0001 criterion, the prediction equation was:

Mean Workload Rating = 8.86 -3.00(LogMeanRange125) + 0.47(MeanTrafficCount)

Where:
LogMeanRange125 = Logarithm of the mean of the distances in meters to the lead vehicles in the same lane as the subject averaged over 30 sec. If there was no vehicle within 125 m, the range of the radar, the distance was set to 125 m.
MeanTraffficCount = Mean number of vehicle detected by the subject vehicle radar (15 degree field of view) averaged over 30 s.

This equation accounts for over 82% of the variance in the mean workload ratings, exclusive of the right merge situations. For them, add 1 to the computed workload rating.

For a looser entry criteria (with $p=.013$ for entry), the resulting equation was

Mean Workload Rating = 8.87 - 3.01(LogMeanRange125) + 0.48(MeanTrafficCount) + 2.05(MeanAxFiltered)

where:
LogMeanRange125 = Logarithm of the mean of the distances to the lead vehicles in the same lane as the subject averaged over 30 sec. If there was no vehicle within 125 m, the range of the radar, the distance was set to 125 m.
MeanTraffficCount = Mean number of vehicle detected by the subject vehicle radar (15 degree field of view) averaged over 30 s.

MeanAxFiltered = Mean longitudinal acceleration (m/s2)

This equation accounted for 87% of the variance of the mean workload rating, an extremely large value. As was noted earlier, the filtering algorithm (GM internal) was intended to remove artifacts in the estimated acceleration (from differentiating the speed signal) when the sampling rate was too low.

For third regression equation, the minimum lead vehicle acceleration and the maximum gap rate in the data set, which reduced the number of cases to 31 (because those values could not be determined when a lead vehicle was absent). The cost of adding these 2 variables was to reduce the number of cases from 42 to 31. The resulting equation, which accounted for 85% of the mean workload rating variance was:

Mean Workload Rating = 8.07 – 2.72(LogMeanRange125) + 0.48(MeanTrafficCount) + 2.17(MeanAxFiltered) - 0.34(MinimumVpDot(0 removed))

where:
LogMeanRange125 = Logarithm of the mean of the distances to the lead vehicles in the same lane as the subject averaged over 30 sec. If there was no vehicle within 125 m, the range of the radar, the distance was set to 125 m.
MeanTraffficCount = Mean number of vehicle detected by the subject vehicle radar (15 degree field of view) averaged over 30 s.
MeanAxFiltered = Mean longitudinal acceleration (m/s2)
MinimumVpDot(0 removed) = Minimum acceleration of a lead vehicle in m/s2 averaged over a 30 s interval, with deceleration of the lead vehicle being negative values. Cases where there was no lead vehicle were not included in the computation.

Readers are reminded that these equations estimate workload for the driving situations recorded in the ACAS FOT (not much residential driving) during daylight hours in good weather, with data on the lead vehicle and traffic being reported by a particular radar (15 deg FOV, 125 m maximum range) and a particular algorithm to compute longitudinal acceleration. Furthermore, keep in mind that all estimates are means for a 30-second time period.

How do ratings of workload vary with the relative position of vehicles ahead (traffic) on expressways?

Because traffic was an important factor, additional ratings for traffic were collected in the post test for expressways only to begin to understand its effects on workload rating. As shown in Figure 68, the relationship between the rating of demand (workload) and the distance to lead vehicles (in car lengths) was logarithmic.

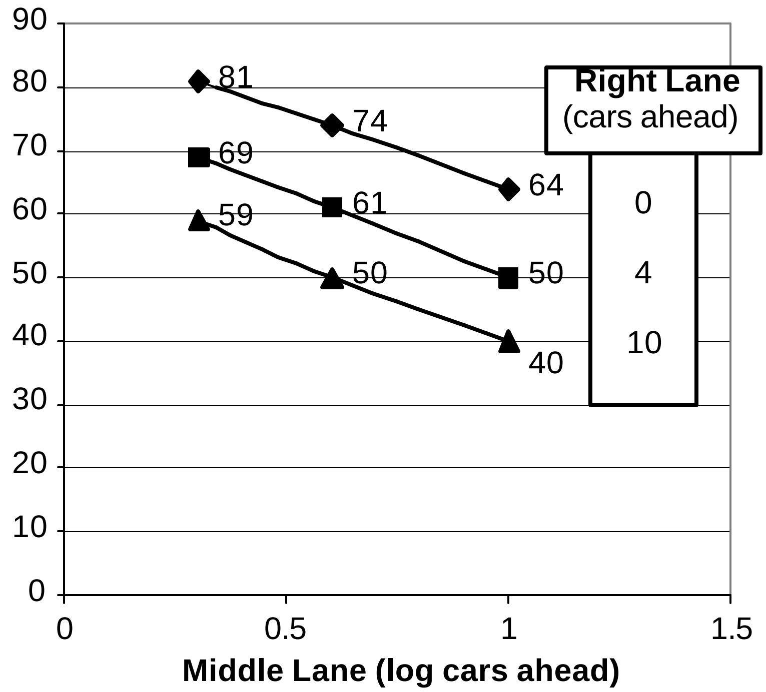


Figure 68. Rating of Traffic Demand Due to Location in the Lane (Log)
0, 4, and 10 refer to the distance ahead of the vehicle in the right lane.

Based on that data, the following equation was developed relating demand (workload) to the position of vehicles ahead. Notice that the contribution of the vehicle ahead in the same lane is about 10 times greater than that in an adjacent lane.

Rating of traffic demand – 89.0 -25.9 (Log middle ahead) -2.4 (Right ahead)

Where:

Log middle ahead =	log of the distance to the vehicle in the same lane in car lengths
Right ahead =	distance to vehicle in right lane in car lengths

What is the relative contribution of traffic, road geometry, visibility and lighting, and traction to ratings of workload?

As shown in Table 58, the factors contributing to workload, based on the post-test ratings were traffic, visibility, road surface condition, and road geometry in that order, though road surface condition and visibility were equal. In this set of ratings, sight distance and illumination were combined to form visibility. This result is quite different from that of Nygren, where traction was half of the rating, visibility and lighting just under 30%, traffic half of that, and road geometry only 6%.

Table 58. Relative Factor Importance

Factors	Road Type					
	Xway	Rural Road	Residential Street	Urban Street	Mean	%
Road geometry-includes lane width, curvature, hills, intersections, merging & turn lanes	1.3	2.1	1.7	1.6	1.7	17
Road surface condition-from dry to wet or icy, also includes road roughness, tire condition and vehicle factors that affect braking and handling	2.8	3.1	2.7	2.7	2.8	28
Visibility-how well you can see-determined by rain, snow, or fog, windshield condition, mirror design	2.5	3.4	2.8	2.5	2.8	28
Traffic-number of vehicles in your lane, adjacent lanes, oncoming, merging and intersecting, also includes pedestrians and bicyclists	3.4	2.3	2.8	3.2	2.9	29

How does the probability of a driver being unwilling to do a secondary task while driving (tune a radio, dial a phone, enter a destination) vary with the overall ratings of workload and (b) road characteristics, traffic, and driver characteristics as in question 3?

A workload manager may suggest or decide which task a driver may or can do as a function of the driving situation. To provide information for such a decision based on workload, logistic regression was used to link the probability a driver was not willing to do a task with workload and other factors. In fact, once workload was known, the type of road driven, its geometry, etc. have very little impact on drivers' willingness to engage. However, driver characteristics were important so they were included in the prediction.

More specifically, the probability a driver is not willing to do a task is estimated using the logistic regression p(no) =1/(1+e^-(ax+b)), where x=workload rating (from clips), and a and b are the slope and intercept of the regression function respectively in Table 59.

Table 59. Logistic Regression Intercepts and Slopes
for Each Age-Sex-Task Combination

(Cells where inverses were used are shown in bold)

Age	Sex	Radio		Phone		Navigation	
		Intercept	Slope	Intercept	Slope	Intercept	Slope
Young	Male	-83.19	11.33	-8.08	2.87	-5.18	5.40
Young	Female	-12.04	1.85	-10.59	2.37	-13.34	1.54
Middle	Male	-18.11	2.85	-6.57	2.12	-3.66	1.79
Middle	Female	-3.28	1.63	-12.47	4.41	-8.68	2.14
Older	Male	-6.45	1.53	-4.28	1.84	-3.78	3.29
Older	Female	-5.35	1.24	-10.23	4.48	-0.08	2.12
Mean		-21.40	3.40	-8.70	3.02	-5.79	2.71

How Could Workload Manager Function Given the Information in This Report?

To function, a workload manager needs to know (1) which factors contribute to total workload and how they should be combined, (2) how the workload of each factor is computed or estimated, (3) how the workload of each task is determined, and (4) the rules by which as task is deemed excessive. Finally, once an excessive task is identified, something needs to be done about that problem. What that will be (locking out tasks, informing the driver the task is excessive, etc.) is being determined in related studies being conducted at the University of Iowa (e.g., Donmez, Boyle, Lee, and McGehee, 2004a,b).

In terms of how various factors contribute to workload, the post-test ratings indicate the 4 factors examined, (1) traffic, (2) visibility, (3) road surface condition, and (4) road geometry contribute 39, 29, 28, and 17% to the total workload, respectively, assuming total workload is strictly additive. Other work such as that of Nygren and Hulse suggest other weightings. So, if one had the scores of each of these 4 dimensions, one could determine aggregate workload using these percentages as weights. However, it is important to realize from the perspective of multiple resource theory, that overload can occur when either the visual, auditory, cognitive, or psychomotor resources are overloaded. For driving, the critical resources are usually visual and cognitive, and as shown in Yee, Nguyen, Green, Oberholtzer, and Miller (2007), demand for those resources in driving are often coupled.

The factors of traffic and road geometry were given considerable emphasis in the project and their demands have been lumped together. Their combined demands can be determined in 3 ways. From the clip rating data, equations were developed that utilized real-time driving data (statistics based on distance to the lead vehicle, the number of vehicles ahead, and longitudinal acceleration for example) to determine estimated workload.

A second approach is to use the data from the post-test ratings, either the simple look up table or the means for each situation. If data on driver age and sex were available, the ratings could be adjusted to account for those 2 factors.

A third approach would be to rely on the clip-based ratings for situations where they fit and adjust and extend them using the post-test ratings. The data available should provide a reason first cut estimate of traffic and road geometry associated workload for a wide range of daylight conditions.

In terms of visibility, one needs to consider both ambient light levels and atmospheric conditions such as fog, rain, and snow, which in combination determine sight distance. From Hulse's work and from work from phase 1 (Cullinane and Green, 2006), it is apparent that the workload associated with visibility is proportional to the log of sight distance, and presumably a relationship could be developed from that prior work, at least for daytime. When nighttime is considered, the situation is complicated because is some sense workload is reduced by what the driver cannot see, so nighttime workload may be less in some situations, a paradoxical outcome.

Road surface condition is a bit more challenging because it has not been explored. As was noted in the introduction, the relationship between workload and friction is unlikely to be linear (Fancher, 2007, personal communication; Karamihas, 2007, personal communication). Possibly an expression such as workload = -1 + e^^kx, where K>0 and a function of mu.max and mu.now might give a better fit to the effect of traction on workload. Also, as was noted in the introduction, traction is vehicle specific and depends on vehicle handling characteristics, the tires and their wear, and the road surface. That data should be available from the published literature. Given some further thought, one should be able to develop a real time prediction of traction related workload that utilized information such as wheel spin data from traction control and dynamic stability control systems, to make real-time predictions about traction-related workload.

So, the presentation of a single unified equation or look up table to predict the workload of the primary task of diving is beyond the scope of this report. However, much of the information needed to develop such exists.

As was noted earlier, to determine workload, a workload manager needs to know the workload of each task. Research pertaining to that topic appears in Green and Shah, 2004; Yee, Nguyen, Green, Oberholtzer, and Miller (2007) and other reports produced for phase 2 of this project.

Finally, some means is needed to determine when a task should not be performed. That can be done using a VACP analysis of the primary and secondary tasks, or subjectively, either using the willingness to engage data from Lerner (2005) and Lerner and Boyd (2005), or using the workload estimates from the clip ratings (or post-test ratings that were converted into clip ratings by linear regression), and then using the logistic regressions for each task, age and sex group to compute p(no). Further work is needed to connect Lerner's research with the research conducted here so p(no) can be computed for a wider range of tasks. Also remaining is a decision as to at what level of p(no) the rules of the workload manager should change (for example allowing or not allowing a task).

What Is the Current Status of Workload Prediction and What Should Be Done Next?

This report provides a solid first step for developing a real-time workload manager. Ratings of workload were determined in 2 different ways and led to predictions that were consistent with each other and internally consistent. The precision of the rating predictions was the same as the precision of the rating in prediction themselves. In terms of real-time predictions of workload, those predictions required anywhere from 2 to 5 statistics that could be easily obtained from a vehicle with ACC and navigation systems. As an alternative, estimates could be obtained using a look-up table.

Those data, using weights determined in this report, could be combined with data on visibility (from prior studies of this project) and with data on traction to determine overall workload.

Finally, the probability a driver was not willing to do selected tasks was determined as a function of workload.

Thus, at this point, the development of a quantitative workload model is feasible. Thus, this report was successful in addressing all of the issues posed and making progress towards the higher-level goal.

To make additional progress, the following questions need to be addressed?

What is the relative contribution of traction, visibility, traffic, and road geometry to aggregate workload, both determine subjectively and from VACP data?

How can the workload associated with visibility and traction be computed?

What is the workload for nighttime situations?

How would the workload of urban driving scenes change if color scenes were used?

How can the workload of driving be estimated using a theoretical model of driving?

How do the position, movement, and types of vehicles in the traffic stream contribute to the workload associated with traffic?

How do the contributions of various factors change if the time period for the workload calculation is other than 30 seconds?

How does the estimate of traffic demand change as a function of the range and field of view of radar sensors that provide the data?

A great deal of progress has been made so far, but there is much yet to do. Saving some of the 42,000 people who die in the U.S. each year in motor vehicle crashes and estimated 1,000,000 worldwide depends on it.

REFERENCES

Alliance of Automobile Manufacturers (2003). Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems (version 3, June 17), Washington, D.C.: Alliance of Automobile Manufacturers.

Breuer, J., Bengler, K., Heinrich, C., & Reichelt, W. (in press). Development of Advanced Driver Attention Metrics (ADAM). DaimlerChrysler and BMW, Germany.

Cullinane, B. and Green, P. (2006). Visual Demand of Curves and Fog-Limited Sight Distance and Its Relationship to Brake Response Time (Technical Report), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation (http://www.volpe.dot.gov/hf/roadway/saveit/docs.html).

Donmez, B., Boyle, L., Lee, J.D., McGehee, D.V. (2004a). A Literature Review of Distraction Mitigation Strategies (Technical Report), Washington, D.C.: U.S. Department of Transportation.

Donmez, B., Boyle, L., Lee, J.D., McGehee, D.V. (2004b). Experiments for Distraction Mitigation Strategies, (Technical Report), Washington, D.C.: U.S. Department of Transportation.

Eoh, H., Green, P.A., Schweitzer, J., and Hegedus, E. (2006). Driving Performance Analysis of the ACAS FOT Data and Recommendations for a Driving Workload Manager (Technical Report UMTRI-2006-18), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Glassbrenner, D. (2005). Driver Cell Phone Use in 2005-Overall Results (Research Note DOT HS 809 967), Washington, D.C.: U. S. Department of Transportation (www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/RNotes/2005/809967.pdf).

Glaze, A. L., & Ellis, J. M. (2003). Pilot Study of Distracted Drivers (Technical Report prepared for Virginia Commonwealth University, Transportation and Safety Training Centre, VA, USA.

Godthelp, J., Milgram, P. & Blaauw, G.J. (1984). The Development of a Time-related Measure to Describe Driving Strategy, Human Factors, 26, 257-268.

Green, P. (1998). Visual and Task Demands of Driver Information Systems (Technical Report UMTRI-98-16). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Green, P. (1999a). Estimating Compliance with the 15-Second Rule for Driver-Interface Usability and Safety, Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society, CD-ROM.

Green, P. (1999b). The 15-Second Rule for Driver Information Systems. ITS America Ninth Annual Meeting Conference Proceedings, Washington, D.C.: ITS America, CD-ROM.

Green, P. (2004). Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions. (SAE paper 2004-21-0022), Proceedings of the 2004 International Congress on Transportation Electronics (Convergence 2004, SAE publication P-387), 165-180. Warrendale, PA: Society of Automotive Engineers

Green, P., Cullinane, B., Zylstra, B., and Smith, D. (2004). Typical Values for Driving Performance with Emphasis on the Standard Deviation of Lane Position: A Summary of the Literature (Technical Report), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation (http://www.volpe.dot.gov/hf/roadway/saveit/docs.html).

Green, P. and Shah, R. (2004). Task Time and Glance Measures for the Use of Telematics: A Tabular Summary of the Literature (Technical Report), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation (http://www.volpe.dot.gov/hf/roadway/saveit/docs.html).

Green, P. E., Wada, T., Oberholtzer, J., Green, P.A., Schweitzer, J., and Eoh, H. (2006). How Do Distracted and Normal Driving Differ: An Analysis of the ACAS FOT Data (Technical Report UMTRI-2006-35), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Hoedemaeker, M., de Ridder, S.N., and Janssen, W.H. (2002). Review of European Human Factors Research on Adaptive Interface Technologies for Automobiles (TNO Report TM-02-C031), Soesterberg, The Netherlands: TNO Institute for Perception.

Horrey, W. J. and Wickens, C. D. (2003). Multiple Resource Modeling of Task Interference in Vehicle Control, Hazard Awareness and In-Vehicle Task Performance. Driving Assessment 2003: Proceedings of the 2nd International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Iowa City, Iowa: University of Iowa, 7-12.

Lerner, N (2005). Deciding to Be Distracted, Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Iowa City, Iowa, University of Iowa, 499-506.

Lerner, N. and Boyd, S., (2005). On-Road Study of Willingness to Engage in Distracting Tasks (Technical Report DOT HS 809 863), Washington, D.C.: National Highway Traffic Safety Administration, U. S. Department of Transportation.

Leibowitz , H. W. and Owens, D.A. (1977). Nighttime Driving Accidents and Selective Visual Degradation, Science, New Series, July 29 197(4302), 422-423.

McLean, J.R. & Hoffmann, E.R. (1975). Steering Reversals as a Measure of Driver Performance and Steering Task Difficulty, Human Factors, 17, 248-256.

Michon, J.A. (ed.) (1993). Generic Intelligent Driver Support, London, U.K.: Taylor and Francis.

Mitchell, D.K. (2000). Mental Workload and ARL Workload Modeling Tools (ARL-TN-161), Aberdeen Proving Ground, Maryland, Human Research and Engineering Directorate, US. Army Research Laboratory.

Norman, D.A. and Bobrow, D.G. (1975). On Data-Limited and Resource Limited Processes. Cognitive Psychology, 7, 44-64.

Nygren, T. E. (1995). A Conjoint Analysis of Five Factors Influencing Heavy Vehicle Drivers' Perceptions of Workload, Proceedings of the Human Factors and Ergonomics Society, 39th Annual Meeting, Santa Monica, CA: Human Factors and Ergonomics Society, 1102-1106.

Oberholtzer, J., Yee, S., Green, P.A., Nguyen, L., and Schweitzer, J., (2006). Frequency of Distracting Tasks People Do While Driving: An Analysis of the ACAS FOT Data (Technical Report UMTRI-2006-17), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Piechulla, W., Mayser, C., Gehrke, H. & König, W. (2002). Online-Fahrerbeanspruchungsschätzung. 38. BDP-Kongress für Verkehrspsychologie. Regensburg, Germany, 12.-14.09.2002.

Piechulla, W., Mayser, C., Gehrke, H., and Koenig, W. (2003). Reducing Drivers Mental Workload by Means of an Adaptive Man–Machine Interface, Transportation Research Part F, 6, 233–248.

Schweitzer, J. and Green, P.A. (2006). Task Acceptability and Workload of Driving Urban Roads, Highways, and Expressway: Ratings from Video Clips (Technical Report UMTRI-2006-19), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Society of Automotive Engineers (2004a). Navigation and Route Guidance Function Accessibility While Driving (SAE Recommended Practice J2364), Warrendale, PA, Society of Automotive Engineers.

Society of Automotive Engineers (2004b). Navigation and Route Guidance Function Accessibility While Driving Rationale (SAE Information Report J2678), Warrendale, PA, Society of Automotive Engineers.

Sullivan, J.M., Winkler, C.B., and Hagan, M.R. (2005). Work Zone Safety ITS Smart-Barrel for an Adaptive Queue-Warning System (Technical Report 2005-3), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Tijerina, L., Angell, L., Austria, A., Tan, A., and Kochhar, D. (2003). Driver Workload Metrics Literature Review, Washington, D.C: National Highway Traffic Safety Administration, U.S. Department of Transportation.

Torkkola, K., Massey, N., and Wood, C. (2004). Driver Inattention Detection through Intelligent Analysis of Readily Available Sensors, Proceedings of the 7th Annual IEEE Conference on Intelligent Transportation Systems (ITSC 2004), Washington, D.C.: Institute of Electrical and Electronics Engineers.

Torkkola, K., Venkatesan, S., and Liu, H. (2005). Sensor Sequence Modeling for Driving, Proceedings of the 18th International FLAIRS Conference, Clearwater Beach, FL, AAAI Press.

Torkkola, K., Venkatesan, S., and Liu, H. (2004). Sensor Selection for Maneuver Classification (paper TuC3.5), 2004 IEEE Intelligent Transportation Systems Conference, Washington, D.C.: Institute of Electrical and Electronics Engineers, 636-641.

Uchiyama, Y., Kojima, S., Hongo, T., Terashima, R., and Wakita, T (2004). Voice Information System that Adapts to Driver's Mental Workload, R&D Review of Toyota CRDL, 39(1),16-22,(www.tytlabs.co.jp/office/library/review/ rev391pdf/391_016uchiyama.pdf).

Transportation Research Board (2000). Highway Capacity Manual 2000, Washington, D.C.: Transportation Research Board, National Academy of Sciences.

Victoria Road Safety Committee, (2006). Inquiry into Driver Distraction (Parliamentary Paper 209, session 2003-2006), Melbourne, Victoria, Australia: Parliament of Victoria.

Wickens, C.D., Gordon, S.E., and Liu, Y. (1998). An Introduction to Human Factors Engineering, New York, NY: Addison-Wesley.

Wierwille, W.W. (1995). Development of an Initial Model Relating Driver In-Vehicle Visual Demands to Accident Rate, Proceedings of the 3rd Annual Mid-Atlantic Human Factors Conference, Blacksburg, VA: Virginia Polytechnic Institute and State University, 1-7.

Yee, S., Nguyen, L., Green, P.A., Oberholtzer, J. and Miller, B. (2007). The Visual, Auditory, Cognitive, and Psychomotor Demands of Real In-Vehicle Tasks (Technical Report UMTRI-2006-20), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Yee, S., Nguyen, L., Green, P.A., Schweitzer, J., and Oberholtzer, J. (2006). Second Generation UMTRI Scheme for Classifying Driver Activities in Distraction Studies and Coding ACAS Video Clips (Technical Report UMTRI-2006-16), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Young, K., Regan, M., and Hammer, M. (2003). Driver Distraction: A Review of the Literature (Technical Report 206), Victoria, Australia: Accident Research Center, Monash University.

Zhang, H. and Smith, M. (2004). Literature Review of Visual Distraction Research (Technical Report), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation (http://www.volpe.dot.gov/hf/roadway/saveit/docs.html).

Zylstra, B., Tsimhoni, O., Green, P., and Mayer, K. (2004). Driving Performance for Dialing, Radio Tuning, and Destination Entry while Driving Straight Roads (Technical Report), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation (http://www.volpe.dot.gov/hf/roadway/saveit/docs.html).

APPENDIX A –
INSTRUCTIONS: TASK 2C SIMULATOR EXPERIMENT

Experiment Setup

Pre-Subject Setup

- Make sure all the forms are present for the subject
- Consent Form
- Biographical data form
- Instructions
- Post test form
- Payment form
- In-car Ratings Sheet

Simulator Setup

- Start up Hyperdrive Computer
- Start Hyperdrive 1.6.2 and open Practice Road
- Start Up Host and Video Channels
- Turn on Video and Audio power strips on the AV rack.
- Turn on VCR
- Turn on Screen above Hyperdrive Monitor
- Switch on both light switches on the bottom of the BUC
- Start up Secondary Task computer
- Start Video Displayer program on Secondary Task Computer
- Start up HP Laptop
- Run "MenuWorksQ" program
- Set up subject file for HP Laptop
- Make sure appropriate cards are in the glove box (Radio, Address and Phone)
- Turn on Projectors (IP, and 4 channels)
- Make sure IP computer is on and the IP program is running.
- Put headphones on desk in control room for experimenter

Subject Setup

Subject Greeting

- Meet the subject in the lobby
- Introduce yourself and verify the subject:
  "**Hello, my name is** - **State your name, You must be -** **State Subject Name"**
- Ask if the subject wants to go to the restroom or get a drink
- Go to the Sim room
- Flip Sign
- Verify the subject's and experimenter's cellular phone / pagers are OFF

Subject Forms

- **"Since this experiment involves driving, we need to verify you are a licensed driver. May I please see your driver's license?"** Check driver's license for vision restrictions and correct date of birth.

- Fill out Consent Form
- Fill out bio form
- Return driver's license

Vision test

  - Clean with alcohol swabs
- **"Since how well you drive depends on how well you see, we need to check your vision. For the entire test, please keep looking straight ahead"**
- Test visual acuity (FAR #2)

- **"Can you see in the first diamond that one of the circles is complete but the other three are incomplete? For each diamond, tell me its number and the location of the complete circle - Top, Bottom, Left, or Right.**

- Test near vision (80 cm) (FAR #2) with Lenses
- **"Can you see in the first diamond that one of the circles is complete but the other three are incomplete? For each diamond, tell me its number and the location of the complete circle - Top, Bottom, Left, or Right.**
- Color-abnormality (FAR #6)
- **"In each circle, there is a number. Starting with Circle A, could you tell me the number?"** (Circle F does not really have a number).

In-Simulator: Parked

Preparation

- Move Seat Back
- Seat the subject in the car
- Adjust seat
- Buckle up
- Adjust rear and side view mirrors
- Adjust all cameras
- Start Recording

Quick Overview:

"In the session today, you will evaluate the driving workload shown on video clips of a wide variety of road types and traffic situations. In addition, you will state how safe you feel it is to manually tune the radio, enter a navigation destination, and manually dial a phone number in those situations.

To give you a sense of the difficulty of those tasks, you will try to do them while driving the simulator on a fairly easy road."

<u>Adjust Sound</u>

Verify volume by playing a beep. "**Can you hear this easily?"**

This is the sound you will hear if you correctly enter the radio frequency and navigation destination. * a beep plays**. If the information is incorrect, you will hear the game show sound. ***Play buzzer**. For dialing a 10-digit number, incorrect sounds like** *incorrect number dialed tones.

"Please do not touch the screen until told to do so."

<u>Radio</u>

- **"The first task is manually tuning the radio. To do this, touch "Start", then "radio" (go ahead) on the first screen menu, wait for a new menu, and then touch "tuner" (go ahead) from that menu. The radio will appear on the touch screen. The station will not immediately appear. You have to push an arrow button to make it appear."**
- **"Next, use the up and down arrows (that change the frequency by 0.2 as in a real radio) to get to the station shown on the card, and then hit the "1" button to make it the preset for button one. A beep will play the station is correct. If it is incorrect, game show buzzer will sound."**
- Guide the user through the first station selection so they become accustomed to the task.
- **"Try the next station"**: trial 1 – practice block (2-5 more tasks). Stop when they basically get it to save time.

<u>Phone</u>

- **"The second type of task is manually dialing a phone. To dial a phone number, touch "Start", then "phone" on the first screen menu and wait for the new menu, and then touch "dial" from that menu. The phone will then appear.**
- **Enter the 10-digit phone number on the card by touching the numbers on the screen. The digits entered will appear on the screen. To correct an entry, use the delete key. Once the entire phone number is entered, press talk (just like in a real cell phone). If the phone number is correct the phone will ring, otherwise a series of tones will mean the number was not correct."**
- **"Once you are told to begin you can press start whenever you wish to do so."**
- Guide the user through the first dialing task so they become accustomed to the task.

- **"Try the next number":** trial 2 – practice block (2-5 tasks). Stop when they basically get it to save time.

Navigation

- **"The third task is entering a destination address for the navigation system. To enter a destination, touch "Start", then "navigation" on the first screen menu, wait for a new menu, and then touch "dest. entry" from that menu. The navigation system will appear on the touch screen."**
- **"Enter the full address (city, street, number) on the three lines given, just as on the card, being sure not to forget the road abbreviation. After each line is entered, press return to go to the next line or to finish if you are on the last line. To correct an entry, use the backspace arrow key. A beep will sound if the address is correct. A game show buzzer will sound if the address is incorrect"**
- Guide the user through the first destination entry task so they become accustomed to the task.

"Try the next address": trial 3 – practice block (2-5 tasks). Stop when they basically get it to save time.

<u>In-Simulator: Driving Practice</u>

"We want you to drive the simulator now. Once you are accustomed to driving in the simulator, you'll be asked to perform the tasks while still driving. Just drive normally."

- **Start Practice Road**
- **Let 5 minutes pass while the subject gets used to the vehicle controls**

<u>In- Simulator: Driving While Performing Tasks</u>

"Next, practice performing the tasks while driving. We'll have you do about 4 trials of the radio, navigation system, and phone dialing." Four trials of each task type while driving should be sufficient to appreciate the challenge of using use these devices while driving.

- 4 Trials of Address entry
- 4 Trials of phone number entry
- 4 trials of radio tuning

<u>Switch to Experiment</u>

- Turn off the rear and right side projectors
- Switch projectors to "video" mode.
- Turn of the IP projector
- Start the "Anchor Displayer" program on the IP computer
- Turn off touch screen

- Recall video settings 01 on the video switcher.

<u>Rating Clips</u>

Now for the main portion of the experiment, you will be rating the demand of driving on expressways, rural roads, and urban streets as shown on video clips. Please rate the demand of actually driving the situation shown in the clip, not the demand of just watching the video. Also, state how safe you feel it is to (1) manually tune the radio, (2) manually dial a phone number, and (3) enter a navigation destination while in the situation shown.

To help rate the driving workload, reference clips will be continually shown on the left screen and you can look at them whenever you want. *play clips* **These clips have workloads of 2(on the top) and 6(on the bottom), where larger values mean more workload.**

Now we will show some practice clips for you to rate. There are 3 clips to rate at any given time. Feel free to use decimal numbers to rate the workload. For example, rating a clip as 5.6 is acceptable. Also say if it was safe for you to do each of the 3 tasks.

If you have any comments for any of the clips, please say them. If you have any questions at any time, feel free to ask them.

Give them 1 practice trial (3 clips).
- Press the "practice clips" button on the video displayer program

"Ok, you have the idea. Now let's collect some test data."

<u>Subject Wrap up</u>
Forms and payment
- Seat subject at rear table
- Complete post-test evaluation form
- Go over the form, ask for clarifications and write them in your words
- Ask for additional comments
- Payment
- Choose payment form according to affiliation
- Pay
- Document
- Walk subject to the front door
- Flip Sign

APPENDIX B –
BIOGRAPHICAL, POST-TEST, and CONSENT FORMS
Biographical Form

Personal Details

Name ____

Phone: ____________ Email address ___________________

May we contact you for future studies? No Yes: (Phone Email Both)

Born (month / day / yr) ___ / ___ / ___ in (city / state) _______________

Are you a native English speaker? (circle one) Yes No

Occupation: (e.g.,: lawyer) _______________ (if retired: occupation before retirement)

Education (circle highest level completed and fill in blank)

High-School Some-College College-Degree Graduate-School

Major ______________ (Ex: Cognitive Psychology, Micro-Biology, Accounting)

Driving

Driver's License # __________________ State:_____ Expiration Date: __________

What motor vehicle do you drive most often?

Year: ________________ Make: ______________ Model: ______________

How many miles do you drive per year? ___________

What lane of a three lane highway do you normally drive in?

Left Middle Right

Have you driven more than 30,000 miles in your lifetime? Yes No

Do you have any special driving licenses (e.g. heavy truck) and if so, what kind?

No Yes: explain -> ________________

In how many accidents have you been involved during the past 5 years? _______

In how many traffic violations have you been involved in the past 5 years? ______

Details: ____

Navigation System Use

Does your current vehicle have a navigation system? Yes No

If yes, how many times per week do you use it? ___________

If yes, do you operate the system while driving? Yes No

Radio Use

What percentage of stations you choose are preset stations (1 button press) _______

How do you change stations when not using presets (choose one)?

knob button press both

Is the radio in your car a factory or after-market system? Factory After-Market

Do you use a CD player? (Y/N) ____ What percentage of the time is it in use? _____

Do you use a portable digital music player in your car? (Y/N) _____

What percentage of time is it in use? _____

Cellular Telephone Use

Do you own a cellular telephone? Yes No

If Yes, how many calls do you make per week? ________________________

Is your cellular phone your primary phone? Yes No

What percentage of your cell calls are long distance: _______ local: ______

Have you ever used a cellular telephone while driving? Yes No

Where is your phone located normally when your are driving?

Cradle Pocket Seat Purse Other______

How often do you use a cellular telephone while driving?

Never Once in a while Once a week Once a day Constantly

Touch Screen Use

Do you use touch screens in the following places?

- Supermarkets (for example Kroger/Meijers) Yes No
- Banks/ATMs Yes No
- Other (for example wedding registry, informational displays) Yes No

Vision Circle what vision correction you use

When driving: no-correction contacts glasses: multifocal, bifocal, reading, far-vision

When reading: no-correction contacts glasses: multifocal, bifocal, reading, far-vision

For the experimenter only *12526616*

Far Acuity	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	T	R	R	L	T	B	L	R	L	B	R	B	T	R
	20/200	100	70	50	40	35	30	25	22	20	18	17	15	13
80 cm Acuity	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	T	R	R	L	T	B	L	R	L	B	R	B	T	R

Post Test Form

Estimation of Driving Workload for the Daytime

Given you have driven in many situations, you should be able to rate the factors that influence driving workload, that is the demand of driving. Assume the conditions are:

(1) daylight and sight distance/visibility is unlimited
(2) the road is flat and dry, and
(3) your vehicle is in good shape
(4) you are in good shape (e.g., not tired).

In the tables that follow, write in the workload of each cell, where:

0 = no demand, you could do it in your sleep or with your eyes closed. (Since one cannot drive while asleep (safely), no driving situation should be rated as 0.)
100 = completely requires all of your capacity to just drive (100%).

In each cell, write 3 values separated by slashes, in the following order:

(a) No or little traffic – Traffic flows freely and there is little or no restriction on speed or maneuverability caused by other vehicles.
(b) Some traffic – Traffic flows smoothly, but other vehicles are beginning to restrict drivers in their freedom to select speed, change lanes, or pass.
(c) Heavy traffic – Traffic does not flow smoothly and is sometimes below the posted speed limit, with some momentary stoppages.

Before filling in the data, you may find it helpful to think about the easiest and most difficult situation for driving on expressways. If you decide to change a rating, which is ok, just cross out the previous value.

Expressway (Evaluate: (a) no or little traffic / (b) some traffic / (c) heavy traffic)

Expressway Driving Situation	Total # Lanes (so 6=3 per direction)		
	6 (you are in left lane)	6 (in middle lane)	6 (in right lane)
Base case = straight road, 1-lane paved shoulder on each side, wide grassy median, no guardrails needed	/ /	/ /	/ /
Base case+ Curved or hilly	/ /	/ /	/ /
Base case + Interchange (entrance/exit) in view or at it	/ /	/ /	/ /
Base case + Lane drop (e.g., 3 to 2 lanes) in your or adjacent lane	/ /	/ /	/ /
Base case + Construction: Approaching or driving in lane shift or narrow lanes with concrete barriers, no shoulder	/ /	/ /	/ /
Base case + Approach or driving through crash scene	/ /	/ /	/ /
Base case but 3-foot shoulder & guardrail instead	/ /	/ /	/ /

Rural Road: (Evaluate: (a) no or little traffic / (b) some traffic only (2 values) since heavy traffic is rare on rural roads)

Rural Road Driving Situation	Total # Lanes		
	2	3 (has center pass/turn lane)	4 (in left lane)
Base case=straight road 8 foot paved shoulder + 8 foot grass beyond that	/	/	/
Base case except gentle curves or hill	/	/	/
Base case except very curved or hilly road (mountain road)	/	/	/
Base case + at or approaching intersection with traffic light	/	/	/
Base case + at or approaching intersection with a stop sign for the crossing road only	/	/	/
Base case with 1-foot shoulder, mailboxes, rocks, vegetation beyond	/	/	/

Residential/Suburban Street (Evaluate: (a) no or little traffic only, since residential roads rarely have traffic)

Residential/Suburban Street Driving Situation	Driveways (per side of the road)		
	0-<2 / block (0.1 miles)	2-5 / block	More than 5 / block
Base case, straight road, no parked cars, no intersection nearby			
Base case, but >0 - 25% of curb has parked cars			
Base case, but >25% of curb has parked cars			
Base case, but curved or hilly			
Base case, but at or approaching signed intersection, where you need to stop			

Urban Street (Evaluate: (a) no or little traffic / (b) some traffic / (c) heavy traffic)

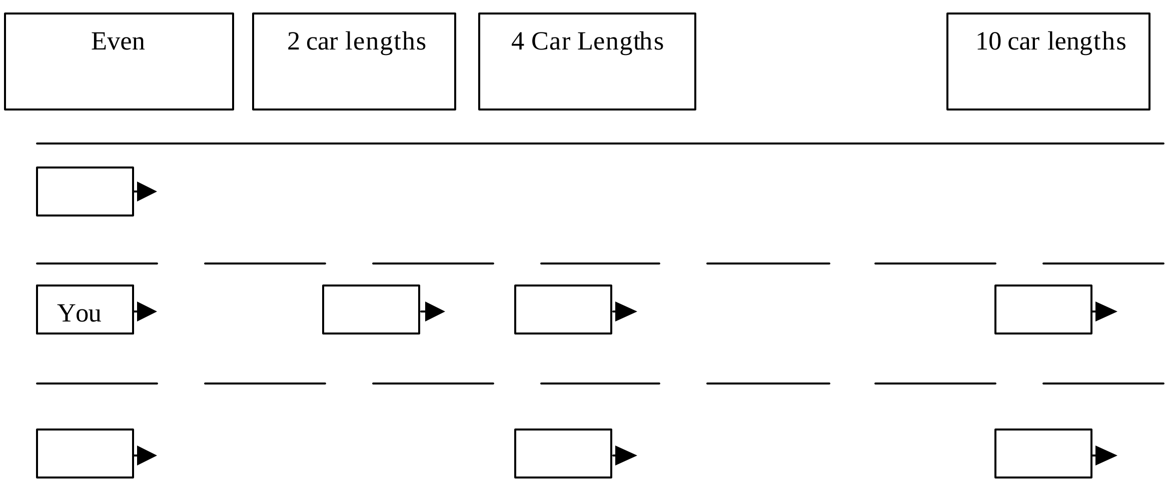
Urban Street Driving Situation	# Lanes			
	2	3 (center turn)	4 (includes turn lane)	5 or more
Base case=straight road, cars parked on side, 10 intersections/mile, most with lights, no or few pedestrians, no stores	/ /	/ /	/ /	/ /
Base case but stores or gas station on corner	/ /	/ /	/ /	/ /
Base case but numerous stores and pedestrians ("downtown"), midblock driveways, no double parking	/ /	/ /	/ /	/ /

Detailed Examination of Traffic on an Expressway

Suppose you are driving on a 6-lane expressway in the middle lane, and as before there are no visibility or traction limitations. Traffic in the left lane is moving 2 mi/hr faster than you are driving and traffic in the right lane is 2 mi/hr slower. As before, **the scale is 0 (no demand) to 100 (requires all of your capacity**).

Distances to Lead Vehicles (we need the graphical equivalent)
Repeat for left 4 car lengths ahead and 10 ahead

Lane			Rating
Left	Middle	Right	
Even or almost overlapping	10 car lengths ahead	10 car lengths ahead	
		4 car lengths ahead	
		Even or almost overlapping	
	4 car lengths ahead	10 car lengths ahead	
		4 car lengths ahead	
		Even or almost overlapping	
	2 car lengths	10 car lengths ahead	
		4 car lengths ahead	
		Even or almost overlapping	



For each of the 4 types of roads, assign points to indicate the relative importance of each factor to your impression of the workload of driving that type of road in the daytime and at night. For example, though it is unlikely, suppose that for rural roads geometry did not matter, but traffic, road surface and visibility did, and they were equally important. Then you would assign 0 points to road geometry and traffic, and 3.33 points each to traffic, road surface condition, and visibility, for a total of 10 points. Assume you are driving straight ahead and not maneuvering (stopping, accelerating, turning, changing lanes, passing)

Factors (in alphabetic order) Use counterbalanced order	Road Type			
	Express-way	Rural road	Residential street	Urban street
Road geometry-includes lane width, curvature, hills, intersections, merging & turn lanes				
Road surface condition-from dry to wet or icy, also includes road roughness, tire condition and vehicle factors that affect braking and handling				
Traffic-number of vehicles in your lane, adjacent lanes, oncoming, merging and intersecting, also includes pedestrians and bicyclists				
Visibility-how well you can see-determined by rain, snow, or fog, windshield condition, mirror design				
Total points	10	10	10	10

Consent Form

Workload of Driving: Demand of Driving As Determined from Video Clips
Investigators: Paul Green (763 3795) UMTRI Human Factors

Participan t number: ____

This experiment examines the visual demand of driving and how it is influenced by using devices such as cell phones and navigation systems. The study will be described in a detailed report for the sponsor and the public, whose results help make future vehicles that you may drive less distracting and safer.

After providing biographical data (your age, driving experience, etc.) and driving data (e.g., miles drive/year, vehicle commonly driven, crashes), you will practice driving the simulator. You cannot crash in the simulation because the car is invincible. After that, you will use devices such as cell phones and a navigation system to provide an impression of what it is like to use them and drive. Your driving performance will be recorded on videotape.

In the main part of the experiment, you will be shown clips from actual driving. You will rate the visual demand associated with driving those scenarios and say which tasks associated with in-vehicle devices can be safely completed under those conditions. This process will be videotaped.

This is an evaluation of the difficulty of driving, not your skill or ability to drive. Participation in this research is completely voluntary and you may skip any question you wish or quit at any time without consequence.

There is a possibility of motion discomfort while driving the simulator. If that occurs, please let the experimenter know immediately and we will stop the experiment. You may withdraw from this study at any time without penalty. You will be paid $70 for your time. The study should take about 3 hours.

I agree to be videotaped in this study and understand that segments from the tapes may be used in presentations to explain the results. My name will not be disclosed. The raw tapes will be erased 10 years after the project is completed.

Sign your name ____________________________

Segments from videotapes of my sessions may be used by the media (e.g., on TV) to help explain this research to the public.

[Optional]: Sign your name ____________________________

The U.S. Department of Transportation and Delco Delphi Electronics, the sponsors of this project, may ask for every piece of data collected (driving data, eye fixations, videotape, difficulty ratings) except your name, address, and phone number. I agree to release of the desired data to them for any purpose. The data will be identified only by a subject number.

[Optional]: Sign your name ____________________________

I HAVE READ AND UNDERSTAND THE INFORMATION PRESENTED ABOVE. MY PARTICIPATION IN THIS STUDY IS ENTIRELY VOLUNTARY.

____________________________ ____________________________
Print your name Date

____________________________ ____________________________
Sign your name Witness (experimenter)

Should you have questions regarding your participation in research, contact Kate Keever:
IRB Behavioral Sciences, 540 East Liberty Street, Suite 202, Ann Arbor, MI 48109-2210, Ph: 936-0933, email: IRBhsbs@umich.edu, web: http://www.irb.research.umich.edu

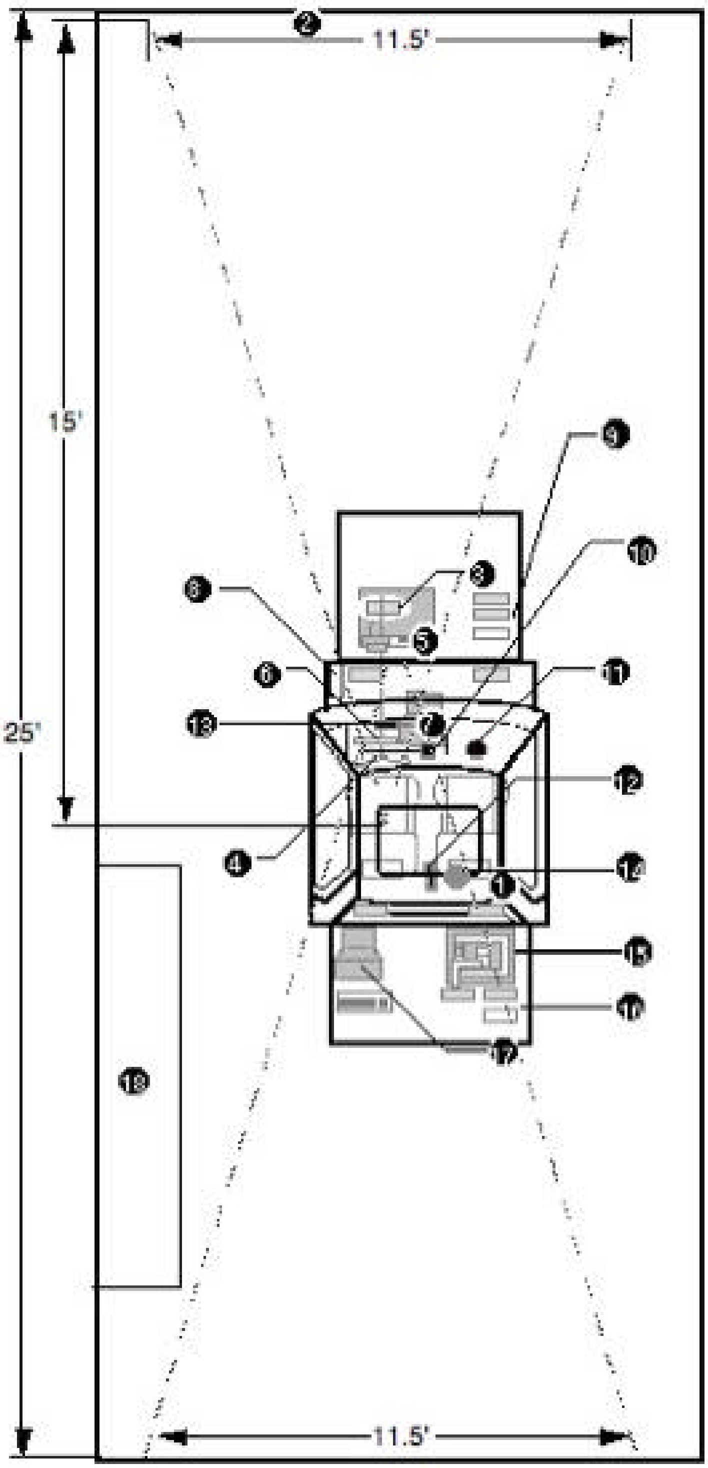
APPENDIX C – ADDITIONAL SIMULATOR INFORMATION

The DriveSafety driving simulator (www.drivesafety.com) has a full size vehicle cab with a touch screen center console, a computer-controlled, projected LCD speedometer/tachometer cluster, operating foot controls, and torque motor to provide realistic force feedback. The in-cab displays are controlled by Macintosh computers running BASIC, software that can also generate directional in-cab sounds. Those sounds are presented by a 10-speaker system from a Nissan Altima, supplemented by a 4-speaker system for road sounds.

Road scenes are projected on 3 forward screens almost 16 feet from the driver (120 degree field of view) and a rear channel 12 feet away (40 degree field of view). Each channel is 1024x768 and updates at 60 Hz. Simulated worlds are created using tiles (as in SimCity). There are about 250 tiles in the library, including scenes from rural, urban, residential, industrial, and expressway settings including intersections with programmable traffic signals. All roads comply with AASHTO and MUTCD standards.

Figure 69. View of the Inside of the Simulator Cab

Figure 70. Subject's View of the Secondary Task Screen



1. 1985 Chrysler Laser mockup with simulated hood
2. White Projection Wall for front projection screen
3. Tourque Motor - SE 808, Model #56728
4. 3-spoke steering wheel
5. Sharp color LCD projection system (model XG-E850U)
6. 4"X13" plexiglas screen
7. ELO Touch Systems Intellitouch monitor (model E284A-1345)
8. Bass Shaker - Autotek Street Machine, #sX 275 - 50 Watts
9. 2-10 Amp , 13.8 VDC Radio Shack Power Supply (1-Shaker Amp, 1-Control Loader/Torque Motor) & 1 Back UPS Pro 1400
10. Face Cam - EIA B/W CCD Camera, Model# - KPC - 5400
11. Shoulder Cam - KT&C CCD Camera - EX Vision
12. Secondary Task Cam - KT&C DSP Color, H1 - res EX Vision
13. Foot Cam - Panasonic CCU - WV - BP550
14. Dome Light - Incandescent
15. Monsoon Sound System for Simulator noise - #E203436
16. 2-10 Amp , 13.8 VDC Radio Shack Power Supply (1-Audio Power Supply, 1 for Dome, PRNDL, E stop, Seat motor) & 1 Back UPS Pro 1400
17. Secondary Task Computer - Power Mac 9500/G3 upgrade/ 350 Hz
18. Simulator Control Room - See Appendix XX

Figure 71. Equipment Layout in the Driving Simulator Buck

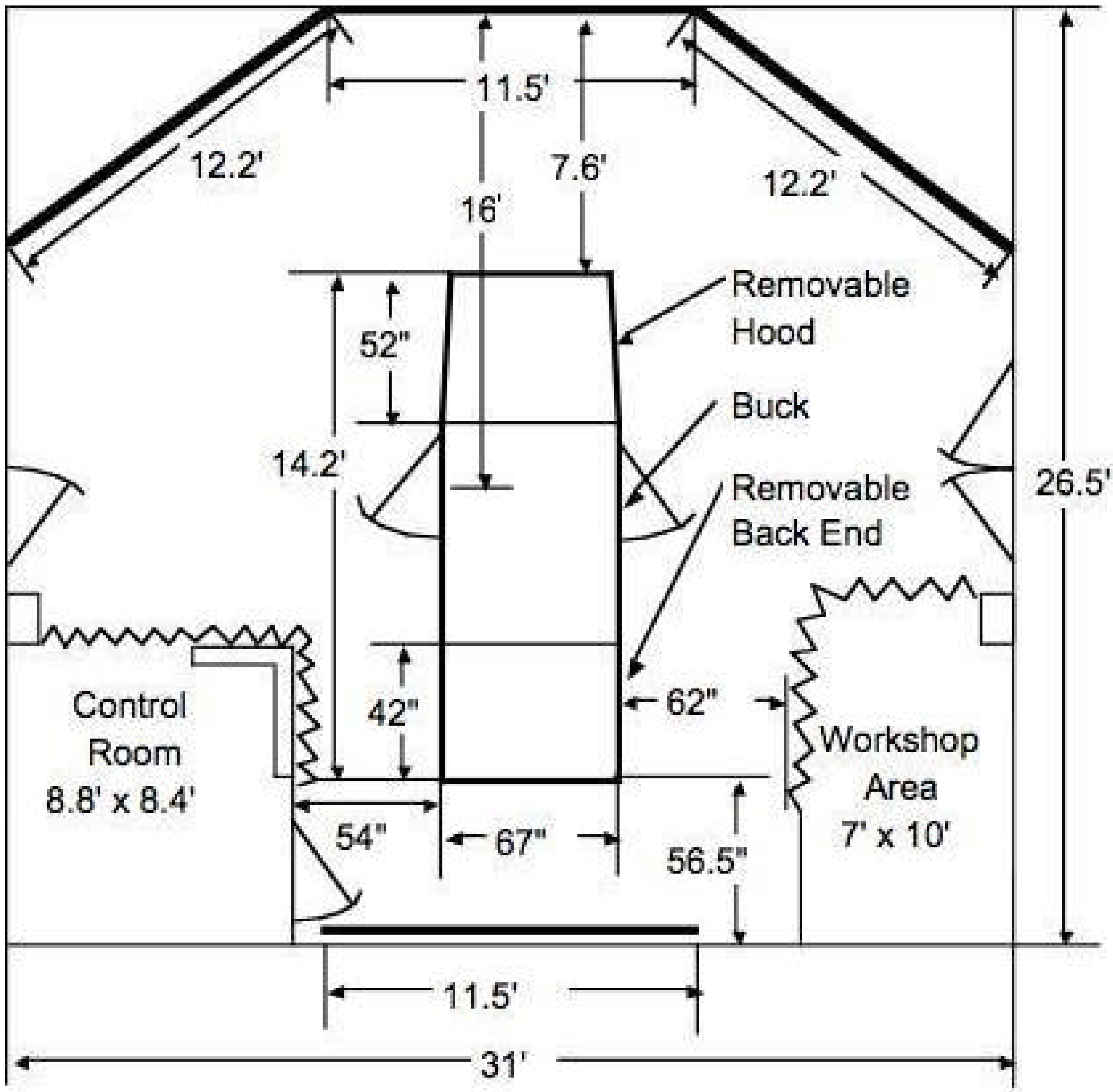


Figure 72. Dimensions of the Simulator Room

APPENDIX D - LOS VALUES FOR VARIOUS ROADS

LOS or Level of Service, is a quality measure describing operational conditions within a traffic stream, generally in terms of such service measures as speed and travel time, freedom to maneuver, traffic interruptions, and comfort and convenience (TRB (2000), 2-2).

To accurately determine LOS, much more information is needed than is available from the ACAS database, and much more time than we had to compute. Luckily, they provide ways to approximate LOS for each type of road we examined. The approximation table for Urban is located in Table 60 below. To determine LOS for the clips, the street class was determined from the speed limit signs just before the clip or during the clip, and then the travel speed of the car was pulled from the database. Using the table below, we established the clip's LOS.

Table 60. Urban LOS Approximation (TRB (2000), 15-3)

Urban Street Class	**I**	**II**	**III**	**IV**
Range of free flow speed	**55 to 45**	**45 to 35**	**35 to 30**	**30 to 25**
Typical FFS	50	40	35	30
LOS	Average Travel Speed			
A	>42	> 35	> 30	>25
B	34-42	28-35	24-30	19-25
C	27-34	22-28	18-24	13-19
D	21-27	17-22	14-18	9-13
E	12-21	13-17	10-14	7-9
F	= 16	= 13	= 10	= 7

Figure 73 shows the approximation of LOS used for rural roads. (TRB (2000), 20-4). The LOS for our clips was found by counting the number of seconds during the clip that the lead vehicle was within 100 m, then dividing by the total number of seconds in the clip (30). The vehicle speed was then pulled from the database. Combining the two measures, the LOS for the rural road was determined.

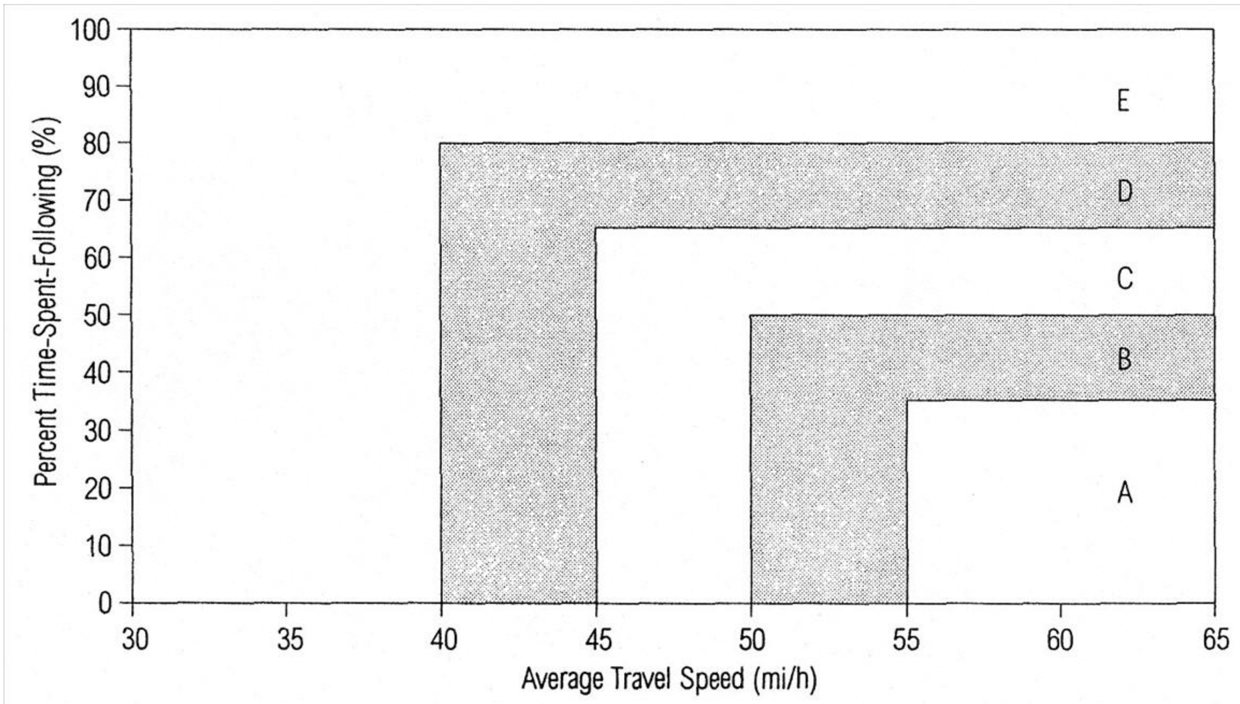


Figure 73. Rural LOS Approximation

Table 61 shows the approximation used for Highway Roads. In order to calculate the Density Range, the number of vehicles was counted in the immediate field of view, and the number of lanes was counted. We could then calculate an instantaneous density. Using Table X, the LOS for the highway clip was then found.

Table 61. Highway LOS Approximation (TRB (2000), 23-3)

LOS	Density Range (pc/mi/ln)
A	0-11
B	11-18
C	18-26
D	26-35
E	35-45
F	>45

APPENDIX E - CLIP SEQUENCE

Category
R1 – Straight Rural
R2 – Curved Rural
U1 – Urban w/o Intersection
U2 – Urban w/ Intersection
E1 – Expressway in Left Lane (3 Lanes)
E2 – Expressway in Middle Lane (3 lanes)
E3 – Expressway in Right Lane (3 Lanes)
E4 – Expressway in Right Lane w/ Merging Traffic (3 Lanes + Merging Ramp)

The number after the dash represents the set of clips used. There were two sets of clips used for each category.

Subject	Block	Trial							
		1	2	3	4	5	6	7	8
1	1	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	2	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	3	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	4	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
2	1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	3	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	4	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
3	1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	3	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	4	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
4	1	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	2	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	3	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	4	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
5	1	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	2	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	3	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	4	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
6	1	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	2	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	3	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	4	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
7	1	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	2	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	3	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	4	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
8	1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2

	3	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	4	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
9	1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	3	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	4	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
10	1	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	2	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	3	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	4	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
11	1	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	2	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	3	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	4	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
12	1	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	2	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	3	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	4	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
13	1	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	2	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	3	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	4	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
14	1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	3	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	4	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
15	1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	3	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	4	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
16	1	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	2	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	3	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	4	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
17	1	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	2	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	3	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	4	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
18	1	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	2	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	3	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	4	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
19	1	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	2	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	3	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	4	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
20	1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2

	3	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	4	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
21	1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	3	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	4	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
22	1	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	2	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
	3	R1-2	R2-2	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2
	4	R1-1	R2-1	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1
23	1	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	2	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
	3	U1-2	U2-2	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2
	4	U1-1	U2-1	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1
24	1	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	2	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1
	3	E2-1	E2-2	E3-2	E4-2	R1-2	R2-2	U1-2	U2-2
	4	E1-1	E2-1	E3-1	E4-1	R1-1	R2-1	U1-1	U2-1

APPENDIX F - EXPERIMENT RATIONALE

CLIP SELECTION

How long should clips examined be? As judged by experimenters in pilot testing, about 15 seconds of viewing time (for slow subjects for some scenes) was needed to rate the workload if the workload was stable. (This was a rough estimate, and further investigation might reveal that 12 seconds is sufficient, or 20 seconds might be needed.) However, clips were recorded at 1Hz, and when played back at that rate, road scenes were a sequence of still images, and traffic movements were difficult to follow. Playing the clips back at 2 Hz, provided continuity (and required 30 s clips). Playing back at 4 Hz (60 s actual footage) provided and even better sense of traffic movement, but the scenes had a cartoonish quality (they were described as scenes from Benny Hill) and could have led subjects to take the rating process less seriously. Also, stable workload or consistent road geometry (all straights or curves) over 60 s was rare, complicating clip selection.

Which scenes should be rated? The intent was to examine as many as possible combinations of road types, geometric factors affecting those road types, and traffic as was possible, while still have repetitions of each combination explored within subjects. Those repetitions involved having all subjects see 2 examples of each geometry-road-traffic combination and seeing each combination twice. Within the time frame of the experiment, it was not possible to examine all combinations of all factors of potential. Furthermore, even though the database of clips was extensive, some combinations of interest could not be found even after many hours of searching, or too few examples were found to support the replication goals.

TASK SELECTION

These tasks were chosen to obtain a basic understanding of the acceptability of each task while driving, and under what circumstances people would or would not use them.

Why these tasks? These tasks varied considerably in their duration, a primary factor that influences the extent to which the driver will be distracted, from often not distracting (and possibly safe/comfortable) to quite distracting (and possibly unsafe/uncomfortable). Radio tuning was selected as a short duration task as it is the benchmark in the AAM guidelines (Alliance of Automobile Manufacturers, 2003), served as the baseline for other driving studies (Tijerina, 1999, 2002), and is considered a short duration task with acceptable risk by drivers. The phone dialing was selected as the medium duration task and is commonly cited as a concern, especially in hands-free versus hand-held discussions. Also, phone use is frequently cited in the literature as a task drivers commonly perform (Glassbrenner, 2005). The navigation task, the long duration task, has been explored before as was used to provide a link to prior UMTRI research (Zylstra, Tsimhoni, Green, and Mayer, 2004). As was noted in the section describing subjects, not all drivers had performed the in-vehicle tasks of interest, certainly not recently. Therefore, to provide a consistent basis for judgment, all subjects practiced each of the 3 tasks.

APPENDIX G - P(NO) FOR VARIOUS ROAD TYPES

Urban

The P(no) graph for Urban roads by device looks very similar to the P(no) graph for all road types. There are slight shifts, but the basic pattern is the same, and the 90% thresholds are all within a half point of the thresholds for all road types.

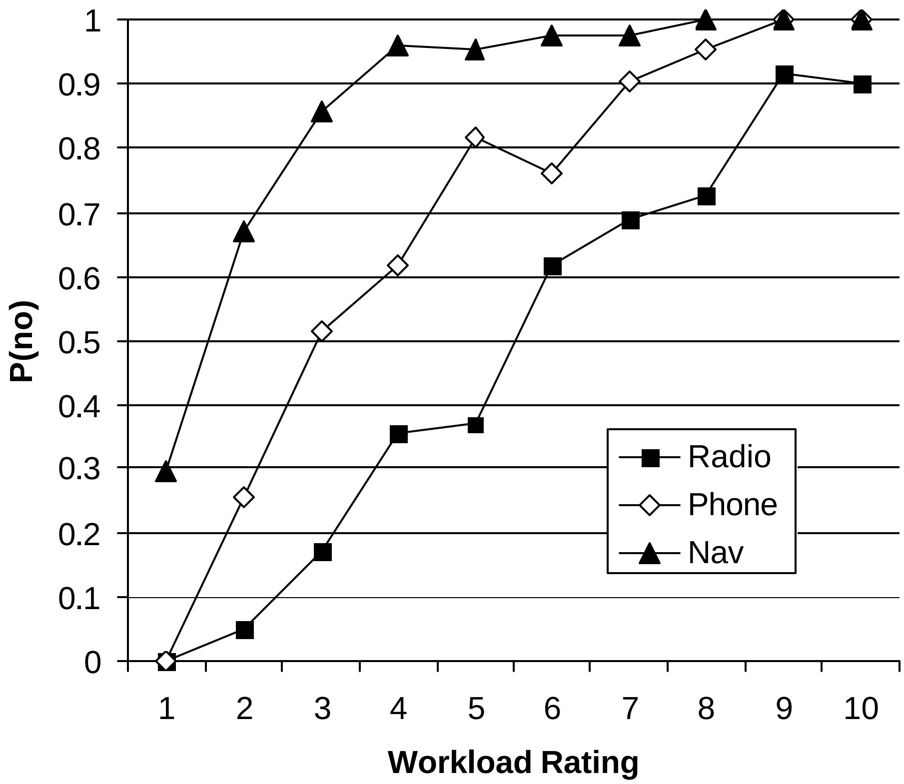


Figure 74. P(no) vs. Rating by Device Type for Urban Roads and All Ages

Age Group has a small effect on P(no) graph shape and slope. Young and Middle aged participants seem to be indistinguishable from each-other, but older participants were less likely to use devices at any workload than younger and middle aged participants

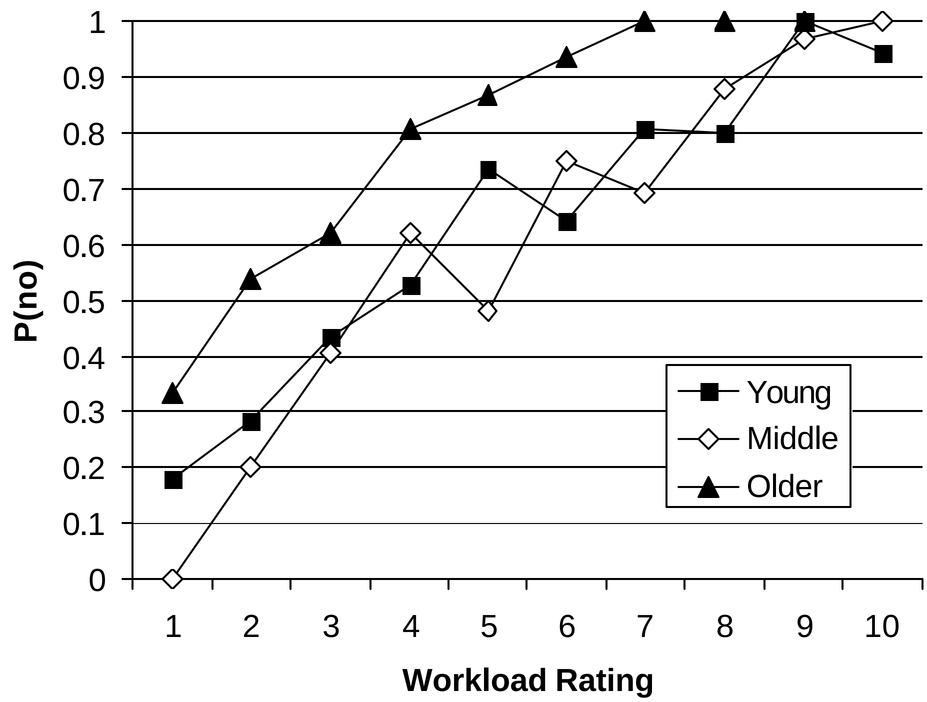


Figure 75. P(no) vs. Rating by Age Group on Urban Roads for All Device Types

Rural

The P(no) graph for Rural roads by device looks very similar to the P(no) graph for all road types. There are slight shifts, but the basic pattern is the same, and the 90% thresholds are all within about a half point of the thresholds for all road types.

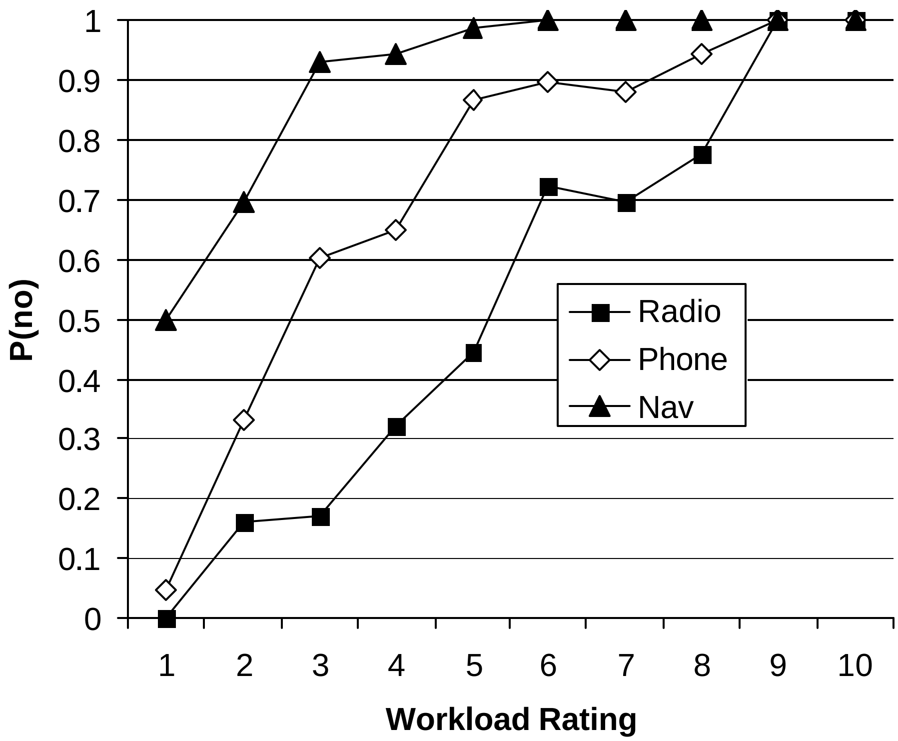


Figure 76. P(no) vs. Rating by Device Type for All Ages on Rural Roads

Age Group has a small effect on P(no) graph shape and slope. Young and Middle aged participants seem to be indistinguishable from each-other, but older participants were less likely to use devices at any workload than younger and middle aged participants.

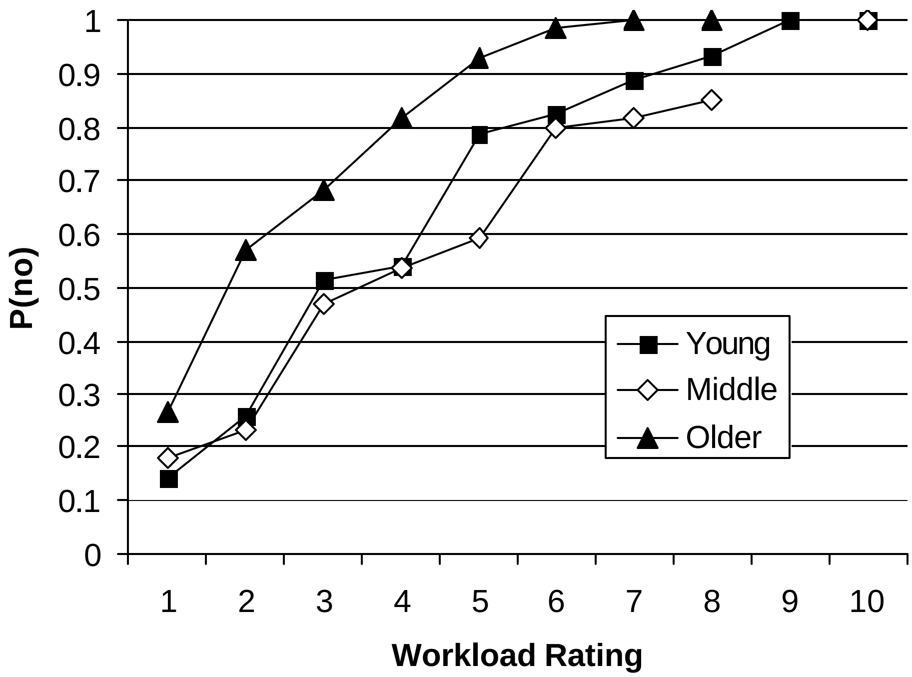


Figure 77. P(no) vs. Rating by Age Group for All Device Types on Rural Roads

Expressways

The P(no) graph for expressways by device looks very similar to the P(no) graph for all road types. There are slight shifts, but the basic pattern is the same, and the 90% thresholds are all within about a half point of the thresholds for all road types.

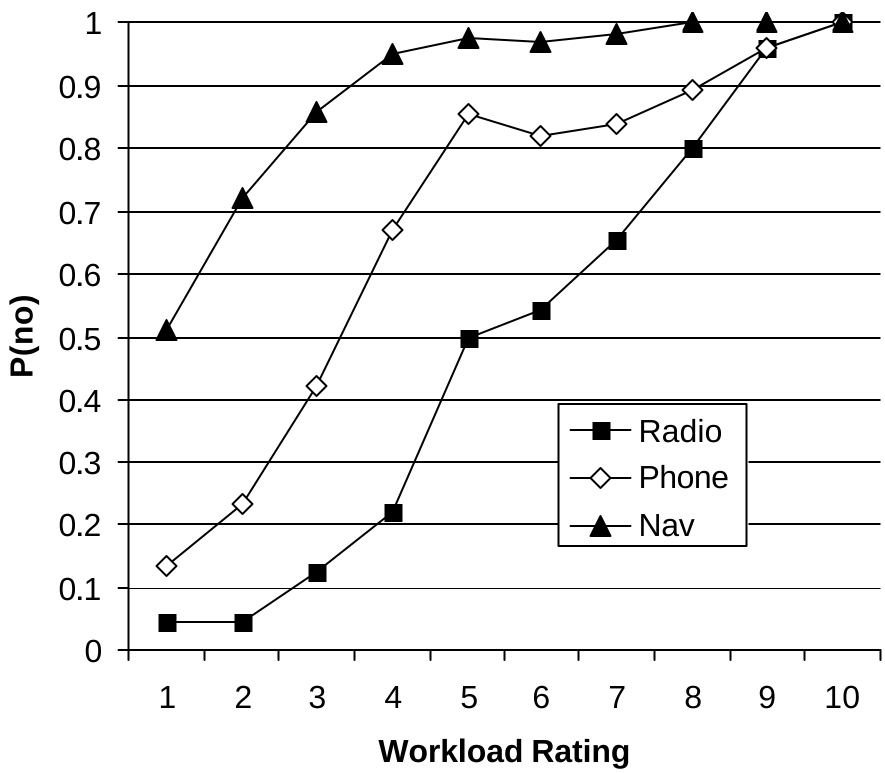


Figure 78. P(no) vs. Rating by Device Type on Expressways

Age Group has a small effect on P(no) graph shape and slope. Young and Middle aged participants seem to be indistinguishable from each-other, but older participants were less likely to use devices at any workload than younger and middle aged participants.

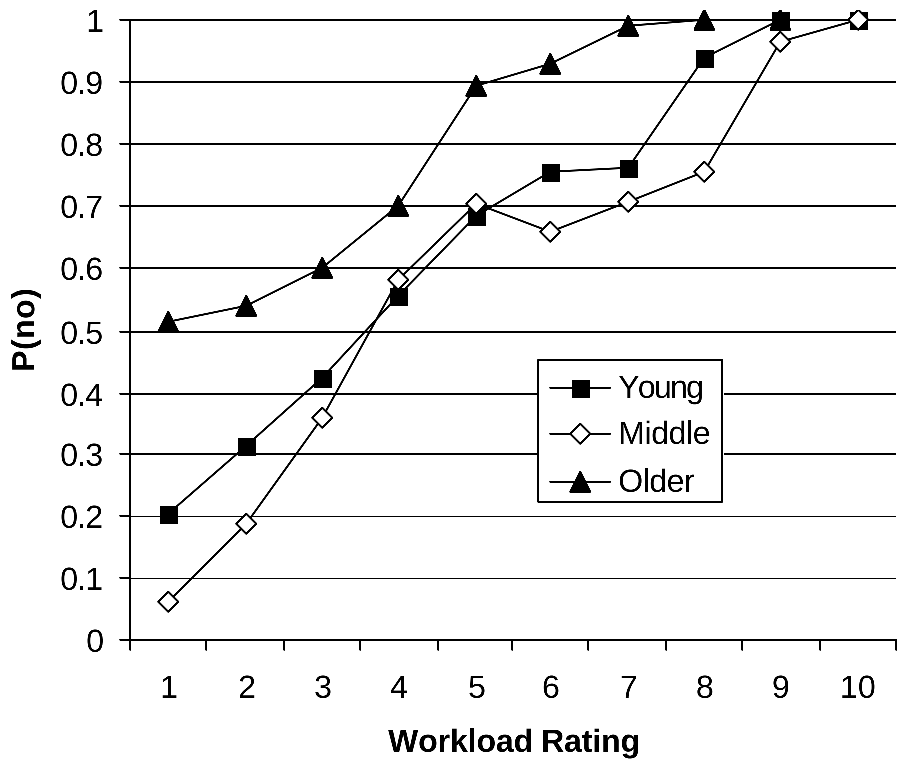


Figure 79. P(no) vs. Rating by Age Group on Expressways

APPENDIX H - P(NO) – WILLINGNESS TO ENGAGE – CALCULATIONS

Table 62. P(no) Willingness to Engage Slopes and Intercepts for Each Task

Age	Sex	Turn Radio		Dial Phone		Enter Destination	
		Intercept	Slope	Intercept	Slope	Intercept	Slope
Young	Male	-10.47	1.63	-11.42	2.9	-8.54	3.62
Young	Male	-93.67	13.28	-3.45	0.62	-12.19	13.57
Young	Male	-216.2	26.93	-7.71	3.73		
Young	Male	-12.4	3.46	-9.72	4.23		
Young	Female	-14.32	2.57	-16.39	4.03	-3.29	2.56
Young	Female	-9.21	1.68	-6.18	1.61	-5.63	2.09
Young	Female	-12.79	1.57	-5.91	1.19	-2.96	1.88
Young	Female	-11.84	1.59	-13.88	2.65	-41.46	21.54
Middle	Male	-33.8	4.04	-12.69	1.71	-5.22	1.66
Middle	Male	-12.7	3.76				
Middle	Male	-5.99	0.47	-7.73	1.82	-2.44	1.05
Middle	Male	-19.93	3.14	-5.85	1.33	-6.98	1.82
Middle	Female	9.36	1.5	-5.32	0.92	-4.84	1.37
Middle	Female	-7.52	1.13	-3.55	1.12		
Middle	Female	-7.33	0.99	-11.02	1.45	-4.99	1.28
Middle	Female	-7.63	2.89	-29.97	14.15	-24.88	12.49
Older	Male	-4.63	1.36	-1.27	0.85	-0.1	0.87
Older	Male	-9.35	2.02	-7.22	2.3		
Older	Male	-5.86	1.3	-7.29	3.19	-15.01	15.01
Older	Male	-5.96	1.44	-1.35	1.01		
Older	Female	-4.25	1.62	-24.65	13.18		
Older	Female	-1.66	0.52	-1.92	0.94	0.54	0.72
Older	Female	-5.08	0.87	-11.1	2.91	-0.87	2.02
Older	Female	-10.39	1.93	-3.26	0.9		

Note: Cell where the intercept and slope are missing device for which a particular subject said they would never do the task while driving (at any workload). In theory, the slope is infinite and the intercept is zero.

APPENDIX I – DESCRIPTION OF DRIVING STATISTICS

Table 63. List of ACAS FOT Data Analysis Factors

Category	Variable	Description	Factors Evaluated	
Number of vehicles ahead	Traffic Count	Number of Cars sensed by the vehicle's radar	Mean	
Distance to Lead Vehicle	CIPV Range	Headway Distance to Lead Vehicle (125 (max radar range) was written over any 0 (no car sensed) values, because at least 125 m was free of vehicles if the radar returned a 0)	Mean	1/ (log (mean))
	CIPV Range Rate	Headway Distance Acceleration (If there was to lead vehicle, the value was removed)	Mean	
	Headway Time	(Headway / Subject Speed) This value is blank in the data set where there is no car in the radar.	Mean	
	Time To Collision	(Lead vehicle speed)-(subject veh speed) / Headway	Mean	
	Vp	Speed of vehicle ahead	Mean	
	VpDot	Speed change of vehicle ahead	Mean	
Road	Lane	Left, Middle, Right (Right is Default)	Value	
	Curvature	Straight or Curved	Value	
	LaneWidth	Width of current lane	Mean	
	RoadClass	Rural, Urban or Expressway	Value	
Lateral position	LaneOffset	Distance from Center of Lane (StDev of Lane Position)	StDev	
	Distance to Lane Edge	Distance to lane edge that the driver is heading towards	Mean	StDev
	Steering Reversals	Count of steering wheel reversals over 2 degrees	Count	
	Lateral	Acceleration of Lane Offset	Mean	StDev

	Acceleration			
	Time To Lane Crossing	Distance to Lane Edge / (Lateral Velocity + Lateral Acceleration)	Mean	StDev
	Steering Entropy	Erwin Boer's Steering Entropy	Value	
Longitudinal Position	Throttle	Mean throttle angle	Mean	StDev
	TransSpeed	Vehicle Speed	Mean	StDev
	Ax	Vehicle Acceleration	Mean	StDev

The lateral velocity and lateral acceleration used in calculating time to lane crossing are averaged over the past .7s (7 data points). These are averaged so that the results will be more stable, as the resolution of the data was not fine enough to yield the precision needed for the time to lane crossing factor. Before filtering, values would change by more than 10 s between consecutive frames (each frame is .1 s apart). After filtering, this wild variability was greatly reduced.

Steering reversal counts were obtained from the ACAS FOT data as well.

Steering entropy was also examined. Based on the formula Eoh, Green, Schweitzer, and Hegedus, (2006), bin counts were 10 and 14. For each bin count, steering entropy was calculated using alpha values of .05, .2, and .4.