Report No. M-CASTL-2011-02

EVALUATION OF A METHOD TO ESTIMATE DRIVING WORKLOAD IN REAL TIME: WATCHING CLIPS VERSUS SIMULATED DRIVING

Paul Green, Brian T. W. Lin, Jason Schweitzer, Helinda Ho, and Katherine Stone

University of Michigan

University of Michigan Transportation Research Institute

September 2011

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Technical Report Documentation Page

1. Report No. M-CASTL 2011-02	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Evaluation of a Method to Estimate Driving Workload in Real Time: Watching Video Clips Versus Simulated Driving		5. Report Date September 2011 6. Performing Organization Code None
7. Author(s) Paul Green, Brian T. W. Lin, Jason Schweitzer, Helinda Ho, and Katherine Stone		8. Performing Organization Report No. UMTRI-2011-29
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. None
12. Sponsoring Agency Name and Address M-CASTL – Michigan Center for Advancing Safe Transportation throughout the Lifespan		13. Type of Report and Period Covered 9/04-9/11 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Sixteen drivers, 8 ages 18-30 and 8 over age 65, drove 53 expressway scenarios, 26 of which replicated scenarios shown to subjects as video clips in a previous experiment as part of the SAVE-IT project. In the SAVE-IT project, subjects rated the workload of driving on an open-ended scale relative to 2 video clips corresponding to light traffic (with an anchor value of 2) and moderate traffic (with a value of 6). In this follow-on experiment, subjects rated both the workload of the scenes while driving and later rated the workload of the video clips resembling them.

The mean workload ratings of video clips from this experiment were highly correlated with the mean ratings from the SAVE-IT study (r=0.97), though the overall ratings were lower. Further, the ratings of video clips from this study were highly correlated with the workload ratings for the scenes when driven (r=0.92).

However, the most important finding was that mean workload rating while driving could be estimated as 5.13 - 0.02 (mean gap), where the mean gap was measured in meters. This equation accounted for 69% of the variance of the workload equations. Also well correlated with the workload ratings while driving were the mean traffic count (r=0.65), the log10 (gap) (r=-0.83), and the inverse gap (r=0.78).

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

19. Security Classify. (of this report) (None)	20. Security Classify. (of this page) (None)	21. No. of pages 117	22. Price

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

Evaluation of a Method to Estimate Driving Workload in Real Time: Watching Video Clips Versus Simulated Driving

UMTRI Technical Report 2011-29, September 2011

Paul Green, Brian T. W. Lin, Jason Schweitzer, Helinda Ho, and Katherine Stone

University of Michigan
Transportation Research Institute
Ann Arbor, Michigan, USA

ISSUES

1. How are the workload ratings distributed?
2. How consistent are the workload ratings within subjects?
3. How consistent are the clip workload ratings across groups of subjects/experiments?
4. Are the workload ratings of various scenarios shown on video clips different from ratings obtained while driving the same scenarios in a simulator? If they differ, by how much?
5. How well do the workload equations developed from viewing of road scenes in the SAVE-IT experiment predict the workload of driving those scenes in a simulator?
6. What equations, based on factors known to be important based on the literature (inverse time to collision, inverse time gap or log gap, lead vehicle acceleration, etc.) best predict the new workload ratings.
7. What are the differences in the above between young and older drivers?

METHOD

16 subjects

Age	Women	Men
Young (18-30)	4	4
Old (>65)	4	4

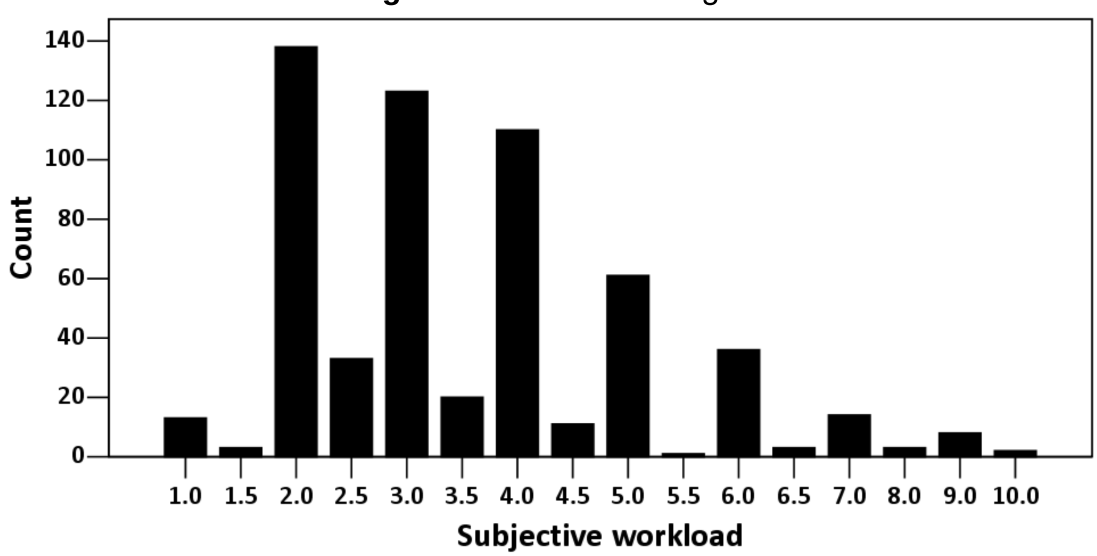
Drive 56 simulated expressway scenarios (1300 m long, 3 lanes in each direction) in UMTRI Driving Simulator; example →

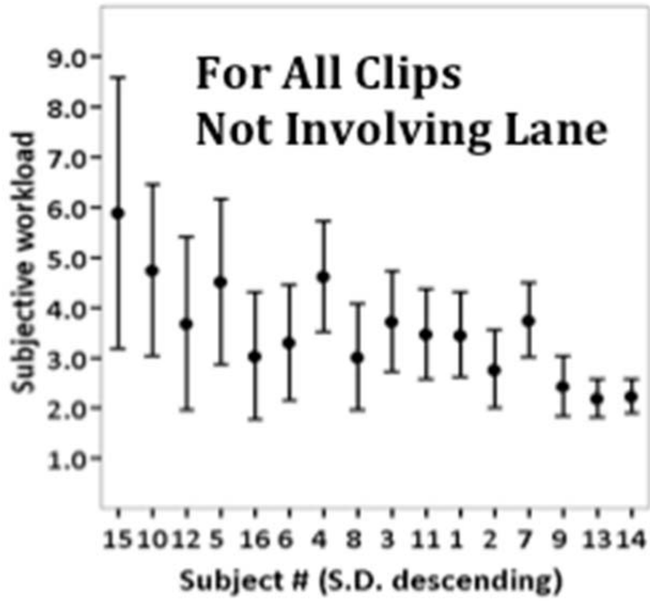
1. While driving, rate driving workload relative to scenes shown on video clips (anchors).
2. Afterward, watch clips of driven scenes and rate workload relative to the same anchor clips.

Stills from anchor clips

RESULTS AND CONCLUSIONS

1. How are the workload ratings distributed? – Integers were favored.



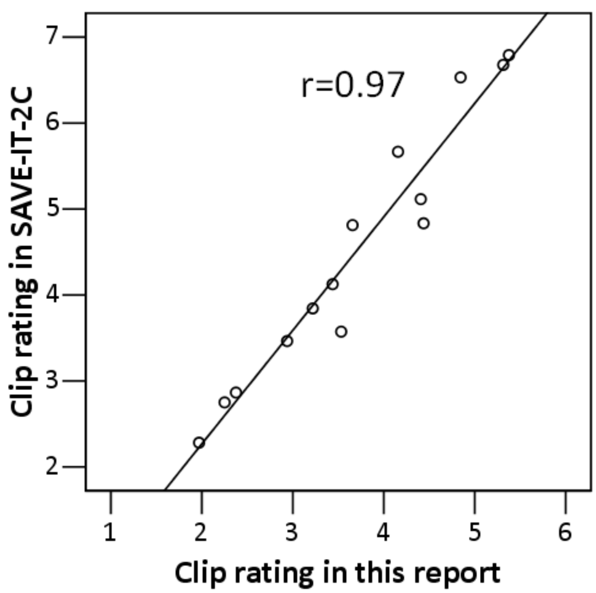


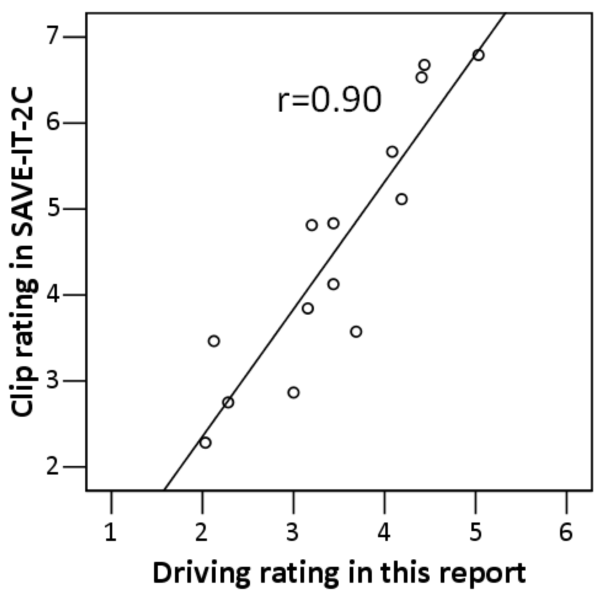
Not all subjects used the full range of ratings, and a few were very limited.

-> Instructions need to be revised.

2. How consistent are the workload ratings within subjects? - Correlations between pair of trials that were driven twice were 0.91 (2nd vs. 14th), 0.80 (14th vs. 24th), and 0.77 (2nd vs. 24th). Given these correlations are computed from individual ratings and subjects, the correlations are quite good.

3. How consistent are the video clip workload ratings across groups of subjects (across experiments)? - Overall, they were very consistent.





Correlations of workload ratings from watching clips by subject:
This study versus. SAVE-IT. - Subject 9 had lower correlations.

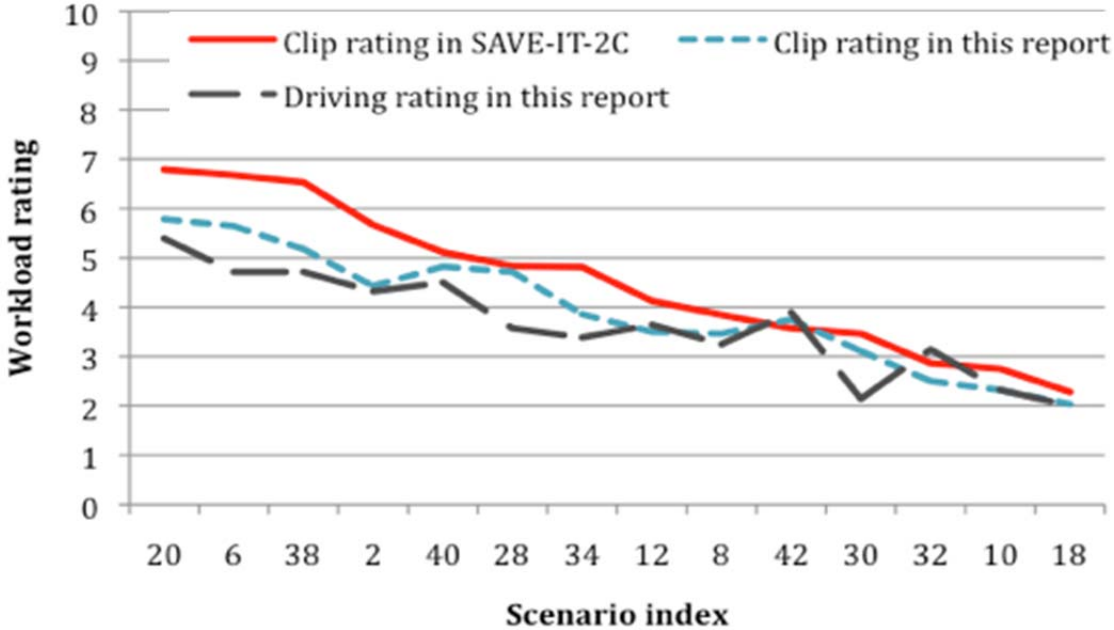
Subject	**15**	**5**	**6**	**10**	**1**	**7**	**12**	**8**
Age	Old	Young	Young	Old	Young	Young	Old	Young
Gender	Female	Female	Female	Male	Male	Female	Male	Female
r	0.93	0.93	0.90	0.87	0.87	0.86	0.84	0.83
Subject	**2**	**13**	**16**	**3**	**11**	**4**	**14**	**9**
Age	Young	Old	Old	Young	Old	Young	Old	Old
Gender	Male	Female	Female	Male	Male	Male	Female	Male
r	0.82	0.82	0.82	0.78	0.76	0.67	0.62	0.48

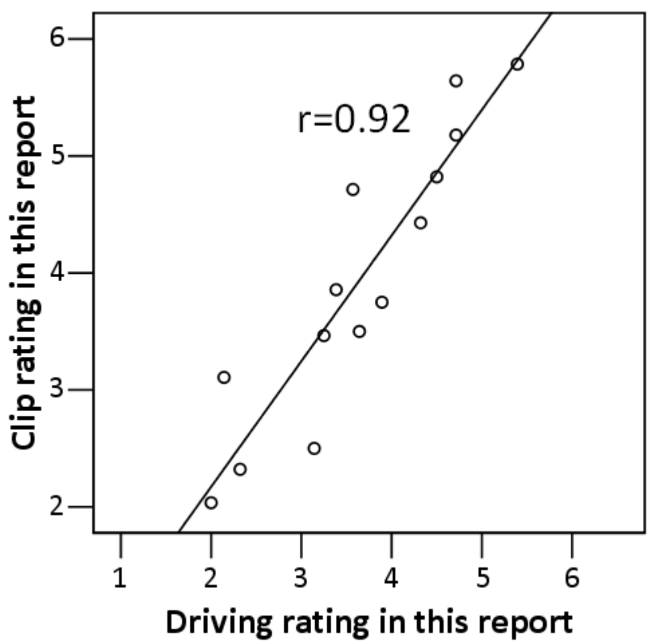
4. Are the workload ratings of various scenarios shown on video clips different from ratings obtained while driving the same scenarios in a simulator? If they differ, by how much? - Overall, they were not different. They were much lower for subjects 9 and 14.

Correlations of workload ratings by subject: Driving versus watching video clips

Subject	**5**	**3**	**1**	**2**	**10**	**11**	**13**	**15**
Age	Young	Young	Young	Young	Old	Old	Old	Old
Gender	Female	Male	Male	Male	Male	Male	Female	Female
r	*0.88*	*0.84*	*0.83*	*0.82*	*0.81*	*0.75*	*0.75*	*0.71*
Subject	**12**	**8**	**6**	**4**	**16**	**7**	**9**	**14**
Age	Old	Young	Young	Young	Old	Young	Old	Old
Gender	Male	Female	Female	Male	Female	Female	Male	Female
r	0.68	0.68	0.68	0.67	0.58	0.52	**0.28**	**0.06**

The workload ratings in this experiment were lower than SAVE-IT, even with subjects 9 and 14 removed. This probably was due to how traffic was shown and determined.





5. How well do the workload equations developed from passive viewing of road scenes in the SAVE-IT experiment predict the workload of driving those scenes in a simulator?

# factors	Original SAVE-IT equation (best fit) Workload =	Original SAVE-IT independent variables and new subjective ratings for reduced scenarios and subjects
2	8.86 -3.00*LogMeanGap +0.47*MeanTrafficCount R^2=0.82	7.90 -2.52*LogMeanGap +0.06*MeanTrafficCount R^2=0.69
3	8.87 -3.01*LogMeanGap +0.48*MeanTrafficCount +2.05*MeanLongitudinalAcceleration R^2=0.87	7.90 -2.51*LogMeanGap +0.06*MeanTrafficCount +0.51*MeanLongitudinalAcceleration R^2=0.69
4	8.07 -2.72*LogMeanGap +0.48*MeanTrafficCount +2.17*MeanLongitudinalAcceleration -0.34*MinimumLeadVehicleAcceleration R^2=0.85	8.57 -2.72*LogMeanGap +0.13*MeanTrafficCount -14.28*MeanLongitudinalAcceleration +0.20*MinimumLeadVehicleAcceleration R^2=0.74

Best predictions using new driving data (Workload =)

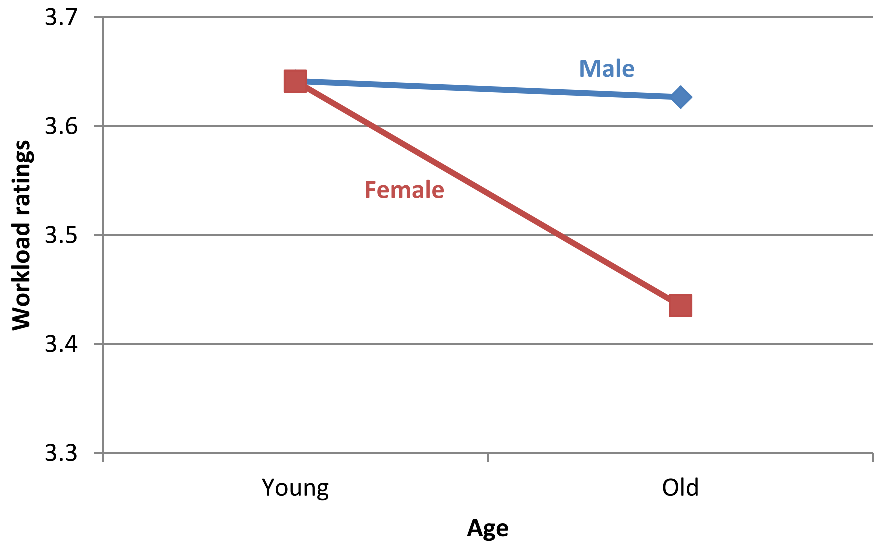
Method	New equations	Comment
Stepwise	5.13 -0.02*MeanGap R^2=0.69	only add variables whose entry was significant at $p<0.05$
Forced entry	7.80 -2.66*LogMeanGap +0.05*MeanTrafficCount -4.17*StDevLongitudinalAcceleration +0.11*StDevTLC R^2=0.69	include all variables that had the highest correlations with workload

6. What equations, based on factors known to be important based on the literature (inverse time to collision, inverse time gap or log gap, lead vehicle acceleration, etc.) best predict the new workload ratings.

Summary of Workload Correlations, r >0.40 in Bold

Category	Variables	Statistic	r	
			All Data	Reduced Set
Other Vehicles	Traffic Count	Mean	**0.65**	**0.65**
		Maximum	**0.65**	**0.53**
		Minimum	**0.55**	**0.52**
	Lead Speed	Mean	-0.37	-0.01
		Maximum	-0.34	0.06
		Minimum	-0.37	-0.04
		Standard deviation	0.01	-0.01
Subject Vehicle Longitudinal	Speed	Mean	-0.19	-0.02 (**0.43**)
		Maximum	-0.17	0.04 (**0.40**)
		Minimum	-0.22	-0.07 (0.34)
		Standard deviation	0.17	0.14 (0.21)
	Longitudinal Acceleration	Mean	-0.13	0.21
		Maximum	0.04	0.36
		Minimum	0.33	0.07
		Standard deviation	0.18	**0.43**
Subject Vehicle Lateral	Lane Position	Mean	0.15	0.02
		Maximum	0.11	-0.01
		Minimum	0.07	-0.11
		Standard deviation	0.32	0.11
	TLC	Mean	-0.07	-0.22
		Maximum	0.18	-0.13
		Minimum	-0.33	0.10
Longitudinal Relationship to Other Vehicles	Gap	Mean	**-0.76**	**-0.83**
		Maximum	**-0.67**	**-0.80**
		Minimum	**-0.74**	**-0.81**
		Standard deviation	-0.14	-0.03
	log10 (gap)	Mean	**-0.73**	**-0.83**
		Maximum	**-0.69**	**-0.82**
		Minimum	**-0.73**	**-0.83**
		Standard deviation	-0.06	0.02
	Inverse Gap	Mean	**0.66**	**0.78**
		Maximum	**0.61**	**0.78**
		Minimum	**0.70**	**0.78**
		Standard deviation	-0.16	-0.28

7. What are the differences between young and older drivers for the new subjective ratings of clips?



The differences depend on which subset of the data one considers. Overall, there were no differences, which is expected as the workload ratings were relative to anchor clips.

KEY CONCLUSIONS

The ratings of workload from watching video clips and while driving those same scenes were highly correlated with the prior SAVE-IT ratings for the same scenes, though the correlations were lower while driving. Further, the most recent workload ratings (of clips and while driving) were lower.

In spite of these difficulties, the workload ratings could be reliably predicted (R^2=0.69) using the mean gap. (Workload = 5.13 -0.02*mean gap)

Additional research is needed to examine other factors, in how traffic is counted.

TABLE OF CONTENTS

INTRODUCTION

What is the problem?

The theme of the Michigan Center for Advancing Safe Transportation throughout the Lifespan (M-CASTL), the organization that funded this project, is safety and mobility throughout the lifespan. Within that theme, the center focuses on (1) the changing perceptual, cognitive, and psychomotor abilities of older drivers to help them maintain safe driving (Eby, Shope, Molnar, Vivoda, Fordyce, 2000), (2) the transportation needs of young people and older adults when they are unable or choose not to drive themselves (Molnar and Eby, 2009), and (3) the elevated crash risk of young drivers. Within the first focus, a primary topic of interest is driver fatigue and distraction among older drivers.

Driver distraction is a prevalent problem for drivers of all ages, both young and old. Driver distraction is a situation where some activity attracts and retains driver attention, diverting attention from the primary task of controlling the vehicle. However, more commonly, situations where drivers are overloaded have been identified as driver distraction, which is technically incorrect. (See Green, 2008). According U.S. Department of Transportation, 5,474 people were killed and 448,000 injured on U.S. roads in 2009 because of driver distraction (U.S. Department of Transportation, 2010).

To determine if a driver is overloaded, one needs to quantify (1) the demands of the primary driving task, (2) the demands of the secondary distracting task (such as talking on a cell phone or entering a street address manually), and (3) the ability of the driver to carry out multiple tasks. This report concerns the demands of the primary driving task.

The primary task of driving consists of speed and path maintenance tasks combined in sequence with specific maneuvering tasks—changing lanes, turning, negotiating an intersection, and so forth. This report considers only the basic and path maintenance tasks, sometimes collectively referred to as car following.

Accordingly, by way of background, 2 questions need to be addressed.

1. What measures or statistics should be used to quantify the demands of the primary driving task?
2. What factors affect workload of the primary task and how?

How has workload been measured subjectively?

In his classic work, De Waard (1996) identified 4 categories of driving performance measures, (1) primary task (e.g., standard deviation of speed and lane position), (2) secondary (e.g., peripheral detection task time), (3) subjective (e.g., NASA Task Loading Index - TLX), and (4) physiological (e.g., Galvanic Skin Response - GSR). Each of these categories as well as measures within these categories has advantages and disadvantages. No single category of measures or specific measure or statistic predominates. However, subjective measures have the advantage of being particularly easy to implement.

Gawron (2000) compared more than 20 subjective measures of workload in terms of reliability, task time, and ease of scoring. The most popular 4 are shown in Table 1. Details follow.

Table 1. Most Common Subjective Measures of Workload

Measure	Task	Scoring
NASA Task Load Index (NASA-TLX)	6 ratings	Weighting procedure
Driving Activity Load Index (DALI)	6 ratings	Weighting procedure
Subjective Workload Assessment Technique (SWAT)	Prior card sort 3 ratings	Computer scoring
Rating Scale Mental Effort (RSME)	1-dimensional scale	0~150 mm line marking

*Source: Gawron (2000).

The Subjective Workload Assessment Technique (SWAT) (Reid, Shingledecker, Nygren, and Eggemeier, 1981) was one of the first multidimensional workload rating methods developed. SWAT assesses workload on 3 dimensions: time load, mental effort load, and psychological stress load, with levels: (1) low, (2) medium, and (3) high for each dimension (Figure 1). Determining the rating is a 2-step process that involves developing scales and then scoring tasks (Reid & Nygren, 1988).

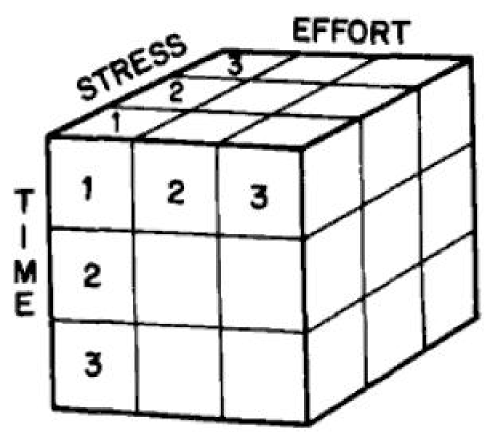


Figure 1. Three-Dimensional Structure (Reid & Nygren, 1988).

More specifically, subjects sort cards representing the 27 cells of the 3-dimensional matrix from lowest workload (1, 1, 1) to highest workload (3, 3, 3). Table 2 is an example creating scale values by subject with the order of time-effort-stress. A more complete description about the scale value calculation, data analysis and weighting scheme is given in Reid & Nygren (1988) and Reid, Potter, & Bressler (1989). During event scoring, the ratings for each subscale of a task provided by subjects will be mapped to the SWAT score from the scale development phase (1 to 100).

Table 2. Time-Effort-Stress Weighting Scheme

Rank Order	Card Label	Descriptor Combination Time	Effort	Stress	Scale Values
1	N	1	1	1	0.0
2	B	1	1	2	24.4
3	W	1	1	3	51.4
4	F	1	2	1	7.6
5	O	1	2	2	32.0
6	C	1	2	3	59.0
7	X	1	3	1	27.7
8	S	1	3	2	52.1
9	M	1	3	3	79.1
10	U	2	1	1	6.5
11	G	2	1	2	30.9
12	Z	2	1	3	57.9
13	V	2	2	1	14.1
14	Q	2	2	2	38.5
15	ZZ	2	2	3	65.5
16	K	2	3	1	34.2
17	E	2	3	2	58.6
18	R	2	3	3	85.6
19	H	3	1	1	20.9
20	P	3	1	2	45.2
21	D	3	1	3	72.3
22	Y	3	2	1	28.5
23	A	3	2	2	52.9
24	O	3	2	3	79.9
25	L	3	3	1	48.6
26	T	3	3	2	73.0
27	I	3	3	3	100.0

The NASA-TLX (Hart & Staveland, 1988) rating is computed as a weighted mean of 6 ratings: mental demands (Low-High), physical demands (Low-High), temporal demands (Low-High), individual performance (Good-Poor), effort (Low-High), and frustration level (Low-High). To compute TLX: (1) Ratings ─ subjects rate each task on each of the 6 subscales, (2) Weights ─ subjects perform 15 pair-wise comparisons of 6 subscales (to determine the weight/relative importance for each subscale), and (3) Combine ─ compute the overall workload score by multiplying each rating by the weight given to the factor. See Nygren (1991) and NASA TLX Homepage (http://human-factors.arc.nasa.gov/groups/TLX/downloads/TLX_pappen_manual.pdf) for scaling issues. NASA-TLX has successfully applied to flying (Selcon, Taylor,and Koritsas, 1991) and driving-related tasks (Alm & Nilsson, 1994). TLX is by far the most commonly used method for rating driving workload. See Hart (2006).

The Driving Activity Load Index (DALI) is an adaptation of the NASA-TLX for driving (Pauzié and Pachiaudi, 1997). DALI is a weighted average of 6 different subscales (attention demands, visual attention, auditory attention, temporal demand, interference, and situational stress) selected by experts to be particularly relevant to driving. DALI has been used to evaluate the use of hands-free phones in vehicles and navigation systems (Pauzie, 2008).

The Rating Scale Mental Effort (RSME) is the simplest of the methods described here to rate workload. RSME quantifies workload on a continuous single scale, represented as a line that runs from 0-150 mm, marked every 10 mm, with 9 subjective anchors along the scale from “absolutely no effort” to extreme effort” (Figure 2, Zijlstra, 1993). RSME has been widely used to examine the workload of completing in-vehicle secondary tasks such as dialing and answering cell phones (Rakauskas, Gugerty and Ward, 2004), using a navigation system (Van Erp & Van Veen, 2004), and using driver assistance systems (Hoedemaeker & Brookhuis, 1998).

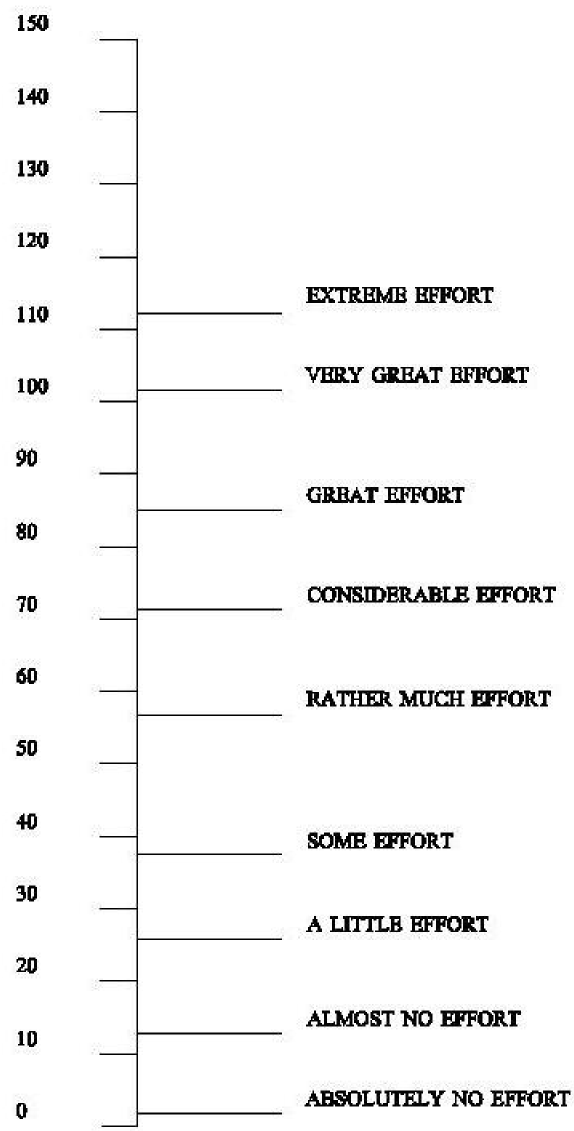


Figure 2. RSME Scale (Zijlstra, 1993).

Thus, there are a number of commonly used subjective workload measurement procedures. Unfortunately, except for RSME, the procedures are unanchored, and even for RSME, the anchors are subjective. Thus, there is no way to compare a person's ratings yesterday with today or today with tomorrow, or one person's ratings with someone else's, so study results can be compared.

Second, almost by definition, the subscales refer to abstract psychological dimensions of the task (effort), not the task characteristics (e.g., maintain speed, maintain lane position).

What factors affect the workload of the primary task and how?

There are many studies that have examined how various factors affect primary task workload. To limit this review, only those studies that include regression analysis or provide for some other relative comparison of the weight of each factor are examined. Either directly or indirectly, these studies have served to influence the design of workload managers, systems that in real time estimate driving task difficulty and then adjust what drivers can and cannot do with the driver interface at that moment.

To estimate workload, Hulse, Dingus, Fischer, and Wierwille (1989) developed the equation that follows. Subsequently subjects drove a route and rated the subjective demand of driving on a scale of 1 to 9 (1 being able to look away for 4 seconds or more, 5 being able to look away for periods of 1 to 1.5 seconds, and 9 not being able to look away at all). Correlations between the rating and workload equations were high (r=0.73).

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

where:

$A = 20 \log_2(500/Sd)$ **(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: $R = 360X / (2\pi a)$
X = arc length along the curve (m)
a = change in direction (degrees)

C = -40*So + 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 = -36.5*W + 267 **(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

Nygren (1995) examined driver workload as a function of traffic density, lighting, roadway type, visibility, and traction. He cleverly used a conjoint analysis to develop an interval scaled measure of perceived demand of workload, realizing that all combinations of factors did not need to be tested. For example, for Lighting versus Traffic Density, the only comparison needed was "which is more demanding – driving at night with light traffic or driving during the day with heavy traffic?" Based on the data from 55 heavy equipment operators, the relative importance data in Table 3 were obtained. Notice that traction was the predominant factor, though exposure to poor traction is not very common.

Table 3. Workload Combinations Examined by Nygren (1995)

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

Piechulla, Mayser, Gehrke, and Konig (2003) developed an interesting method to determine the workload of a particular road segment (Figure 3). First, the segment is coded using the 6 dimensions of the Fastenmeier coding scheme (Fastenmeier and Gstalter, 2007). These are: (1) road type (five highway classes, two rural road classes, seven city classes) (2) horizontal layout (curve versus no curve) (3) vertical layout (slope versus plane route) (4) intersections (four classes) (5) route constrictions (yes/no) (6) driving direction (straight ahead, turn left, turn right). In total, there are 186 possible combinations. Using historical driving data, Piechulla, determined the mean number of glances per second subjects devoted to a secondary task for 22 combinations of roadway. When driving, his system finds the most closely matching of the 22 segments and does a table look up to determine the estimated instantaneous workload.

However, actual workload is the current value, plus what is needed to plan ahead. Piechulla, et al. (2003) assume that is 5 s and that workload decays exponentially with time $y = 2.71866e^{-x/4.72657}$. Following is the workload model developed from this study. As shown in the figure, the presence of an intersection, hard braking (in excess of 0.1g) ACC operation, overtaking, and rapid approach are factors that affect workload.

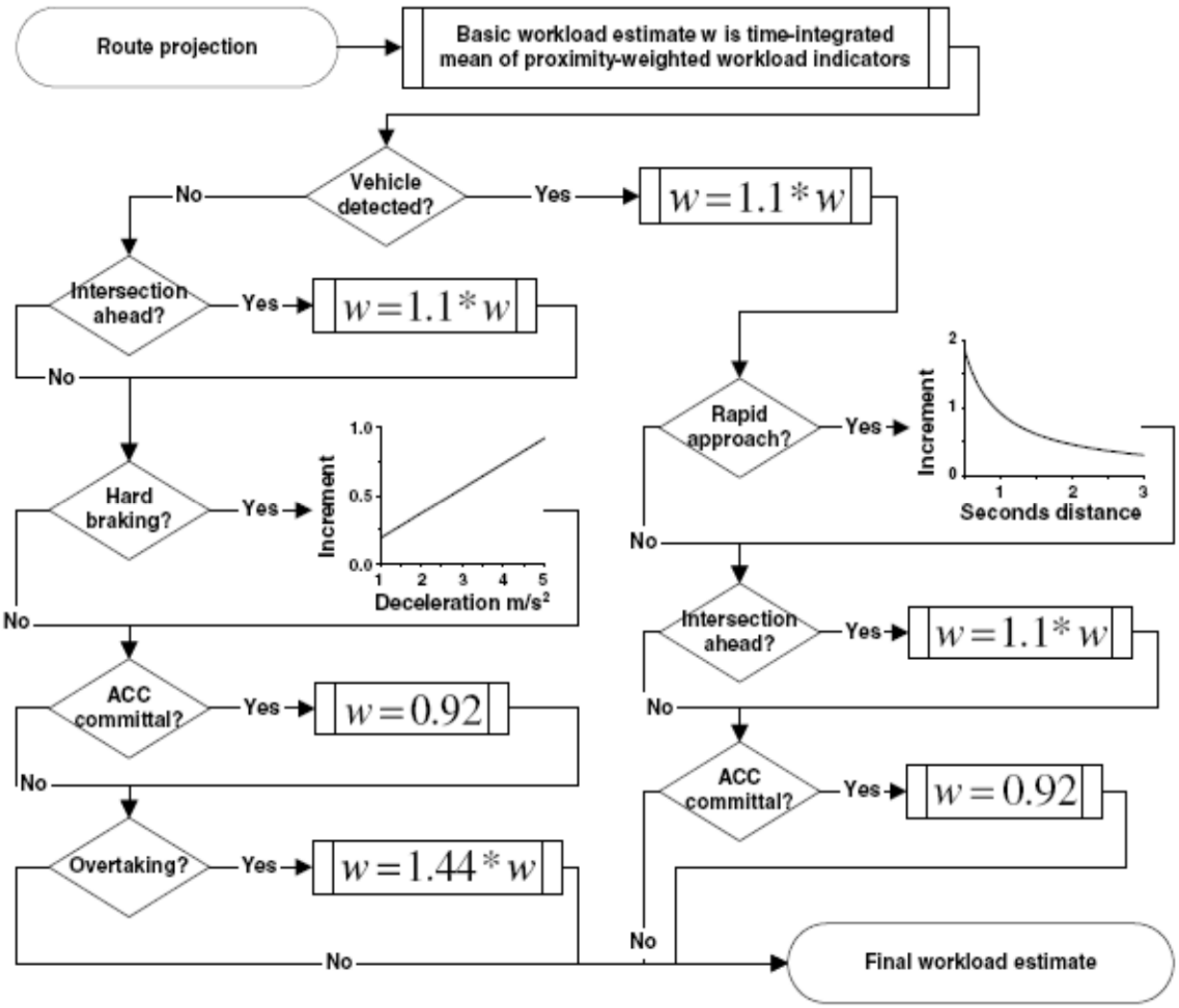


Figure 3. Piechulla et al.'s Workload Rating Scheme (2003)

Aware of these prior studies (but not the details of Fastenmeier's coding scheme), Schweitzer and Green (2007) developed a quantitative method to predict the subjective workload of drivers, in particular one that provided anchors so the method would provide results that could be used to compare different roads and drivers over time. This research was conducted under the auspices of the SAVE-IT project (www.volpe.dot.gov/hf/roadway/saveit/index.html, retrieved June 2, 20110), one of whose goals was to provide information useful to building a workload manager.

In their experiment, 24 subjects from 3 age groups were shown video clips of a wide range of driving scenarios obtained from Advanced Collision Avoidance System (ACAS) dataset (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, and Winkler, 2005). Clips included 3 classes of roads: expressways, rural, and urban roads. Two anchor video clips were shown, and subjects rated the perceived workload of each of the 40 plus scenes shown relative to the 2 anchor clips showing expressways. The traffic in those two clips, assigned ratings of 2 and 6, were Level of Service (LOS) A and C, corresponding to light and moderately heavy traffic respectively. After watching the clips, subjects filled out a post test survey, again, indicating their perceived workload for a wide range of driving situations from a scale of 0, (no demand) to 100 (high demand). The clip ratings and post-test ratings were highly correlated, so the post-test data could be used to extend the workload ratings to a wider domain as there were far more conditions in the post test.

Further, the authors used stepwise regression to examine the relationship between the mean workload ratings (across subjects) and numerous independent measures known for each clip, as the clips were collected in instrumented vehicles. Initially considered were longitudinal and lateral acceleration, lane position, Time to Line Crossing (TLC), steering wheel angle, throttle angle, steering reversals, and steering entropy, driver age and gender, traffic density, lead vehicle speed and longitudinal acceleration, gap, gap rate, Time to Collision (TTC), the number of lanes, and posted speed, measures that may vary with driving workload. However, the final analysis was restricted to factors whose correlation was greater than or equal to 0.4 to avoid spurious results. Thus, only the subject's speed, longitudinal acceleration, density and count, lead vehicle acceleration, log gap, and gap time were considered in the regression analysis. Generally, the mean, minimum and maximum of these measures yielded similar correlations, though the means tended to be higher because the values were more stable and not quantized.

For a strict entry criteria into the regression equation (p for entry <0.05), only log mean range and traffic count were significant, resulting in the equation that follows. This equation accounts for over 82% of the variance of the mean workload ratings (by scenario/video clip), which is unusually large.

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

Other equations follow, developed using more inclusive criteria. The differences between equations in variance accounted for was a few percent.

Mean Workload Rating = 8.87 - 3.01(LogMeanGap125) + 0.48(MeanTrafficCount) + 2.05(MeanSubjAxFiltered)

Mean Workload Rating = 8.07 – 2.72(LogMeanGap125) + 0.48(MeanTrafficCount) + 2.17(MeanSubjAxFiltered) - 0.34(MinLeadAx(0 removed))

where:

LogMeanGap125 = 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.

MeanSubjAxFiltered = Mean longitudinal acceleration of the subject vehicle (m/s2)

MinLeadAx(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.

Interestingly, the log of the gap was a much better predictor of the workload ratings than the gap. This makes sense. When a lead vehicle is close to a subject, the workload is high, dropping off with distance. However, once a vehicle is at a far distance, the workload difference between far and very far is negligible.

Independently, Kondoh, Yamamura, Kitazaki, Kuge, and Boer (2007), developed an equation to predict driver perception of risk. Although risk and workload are quite different, one would expect the 2 to be correlated in a car following situation, and hence predictions of risk should be insightful in developing workload predictions. In 2 driving simulator experiments, 1 involving steady state car following, another involving a closing situation, subjects compared the risk of various combinations of time headway and time to collision. There were 10 generally young male drivers. Using regression analysis, the researchers found that

Risk perception = 1/time headway + 5/time to collision

To validate the equation, the authors conducted an experiment with 15 drivers, varying in age, driving on expressways and surface streets in and around Minneapolis/St. Paul, Minnesota. They found that the line defined by 2=1/time headway + 4/time to collision distinguished between when drivers braked (below the line) and when they did not. This validates the expression developed in the simulator experiments. For an extension of these ideas and their application to last minute braking, see Wada, Doi, Tsuru, Isaji, and Kaneko (2010).

Thus, the literature has shown that the workload of driving is dependent primarily on traffic and road geometry. For traffic, the key factors are the inverse time headway (gap) to the lead vehicle and the inverse time to collision, with the inverse time to collision being more heavily weighted. In additional, also important are the log of the distance to the lead vehicle, the number of vehicles ahead, and the acceleration of the lead vehicle and the subject vehicles. For road geometry, the key factors are the log of and the inverse sight distance, the inverse curve radius, and a negative value based on the lane and road width. For recent additional information on road geometry, see Green, Diebol, Park, and Ho (2011).

Thus, the research to date is beginning to provide a solid basis for developing workload equations. In the SAVE-IT project (www.volpe.dot.gov/hf/roadway/saveit/index.html, retrieved June 2, 2011), the workload ratings were from subjects who were watching video clips of road scenes, not actually engaged in driving. This was done because it was not feasible to create driving scenarios for the large number of situations that needed to be assessed within the project budget and schedule. However, video clips for that purpose were available. Given the recent availability of software to rapidly and less expensively create driving scenarios (Schweitzer and Green, 2009), this project examined the workload of actually driving the scenarios compared to just watching them. One explanation was that passive viewing of video clips could lead to lower workload ratings than if subjects were actually engaged in driving. Also the anchor clips and clips rated in SAVE-IT were all recorded at 1 Hz and presented at 2 Hz, with the double speed playback being used to make movement continuity more apparent. However, the driving simulator image updates at 30 Hz, is in full color (making some changes more apparent), and has up to a 200-degree field of view (versus only 60 degrees for the recorded clips). The difference in the update rates could lead to increased variability in the workload ratings as the anchor clips more closely resemble the recorded clips than the simulator scenes. The wider field of view in the simulator could lead to the simulator ratings being greater than the video clips of the same scene because more of the scene is visible. Given the goal of using the video clips and equations for assessing simulator and on-road driving, the workload clip rating process needed to be examined in a driving simulator as a follow-up to the prior SAVE-IT project. Accordingly, this research addressed the following questions:

1. How are the workload ratings distributed?

2. How consistent are the workload ratings within subjects? Specifically, if a subject drives the same scenario twice, how similar are the 2 workload ratings?

3. How consistent are the clip workload ratings across groups of subjects/experiments? Specifically, how well do ratings of workload of the video clips from a new group of subjects correlate with ratings from subjects in the SAVE-IT project?

   For the video clip rating procedure to be useful, the ratings need to be stable.

4. Are the workload ratings of various scenarios shown on video clips different from ratings obtained while driving the same scenarios in a simulator? If they differ, by how much?

   Are the workload estimates from watching driving and actually driving different? When can video clips be used?

5. How well do the workload equations developed from passive viewing of road scenes in the SAVE-IT experiment predict the workload of driving those scenes in a simulator?

6. What equations, based on factors known to be important based on the literature (inverse time to collision, inverse time gap or log gap, lead vehicle acceleration, etc.) best predict the new workload ratings.

7. What are the differences in the above between young and older drivers?

TEST ACTIVITIES AND THEIR SEQUENCE

Sequence of Test Activities

This experiment aimed to validate the workload ratings given by the drivers as well as the equations developed in the previous SAVE-IT project (Schweitzer and Green, 2007). As a reminder, subjects in that experiment were shown video clips of driving scenes whose workload they rated relative to 2 anchor clips. Both the anchor and the clips rated relative to them were recorded in the Advanced Collision Avoidance System (ACAS) Field Operational Test (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, Winkler, 2005).

Subjects (recruited using the materials in Appendix A) began this experiment by completing the consent (Appendix B) and biographical forms (age, traffic violation, driving experience, etc. - Appendix C). Subjects were given a vision test to determine if they had corrected minimum eyesight of 20/40, the minimum required to drive in many states in the United States.

Then, after adjusting the driver's seat, subjects drove the UMTRI Driving Simulator for approximately 3 minutes with no traffic to become comfortable with the simulator. They changed lanes half way through that period. If the subject experienced any motion discomfort, they were instructed to inform the experimenter immediately, in which case they were excused (as was the case for 2 subjects who were replaced). If the participant did not experience motion discomfort, they drove an additional 5 minutes. During this second practice trial, the participant rated the workload of the primary driving task to verify that they understood how workload was to be rated.

To anchor the workload ratings, subjects were shown 2 monochrome looped video clips representing low and moderately high workloads, having ratings of 2 and 6 on an open ended scale (Figure 4). The original clips were recorded at only 1 Hz, a limitation of the instrumentation available when ACAS was conducted. However, to provide a sense of continuity, the clips were played back at 2 Hz, as was the case in the prior SAVE-IT experiment.

The field of view of the video clips was approximately 10 degrees vertical by 60 degrees horizontal. The anchor clips were shown on an 11-inch LCD mounted in the center console. This is approximately the size display and location that would be used if this procedure were implemented in a real vehicle for on-road testing. Subjects rated the workload of the primary task relative to the anchors, with greater values signifying greater workload and ratings typically ranging from 1 to 10 to the nearest ½ point. (See Appendix D for the exact wording.) The beginning and end of the 15-s rating interval during each scenario was indicated by voice prompts. At other times, vehicles were repositioned to desired positions and speeds for ratings.

Figure 4. Stills of the Anchor Clips

Subsequently, the 2 test blocks began. The goal was to examine simulations of 18 of the ACAS clips (Table 4). Clips varied in terms of the 2 primary characteristics that were easy to manipulate and were found to affect workload—the lane driven and the Level of Service (LOS), a measure of traffic. Details on the road scenes appear later. For those unfamiliar with Level of Service, see Appendix E.

Table 4. Clip Numbers for Lane and LOS Combinations Examined

Lane Driven	LOS		
	A	C	E
Left	144, 145	152, 153	148, 150
Middle	135, 138	140, 143	40, 139
Right	126, 129	125, 130	29, 136

There were 2 test blocks, with Table 5 showing which clips were replicated in each block. Note that 2 clips were replicated 3 times in Block 1, and a different 2 were replicated 3 times in Block 2. These replications were included to get a sense of how consistently test scenarios in the clips could be repeated. Keep in mind that the exact position of vehicles in a scenario depended upon their position in previous scenarios and how subjects drove, so exact replication of the vehicles' speeds and locations was not feasible.

Table 5. Clips Replicated in Each Scenario

ACAS Clip #	Block 1	Block 2
29	1	
40	3	
125	1	
126	1	
129		1
130		1
135	3	
136		1
138		3
139		3
140	1	
143		1
144	1	
145		1
148	1	
150		1
152	1	
153		1
Total	13	13

More specifically Block 1 was a 20-minute drive, containing 26 30-s scenarios/trials (Appendix F). Of them, 13 replicated trials from the SAVE-IT experiment (Schweitzer and Green, 2007), though in this case, they were actually driven rather than being shown on videotape. These trials had Level of Service levels of A, C, and E and were distributed throughout the block. The remaining 13 trials were transitions, in which the other cars on the road would switch lanes and adjust their speeds in order to be in the correct positions for the next replication of the SAVE-IT clips. The participant began Block 1 driving in the middle lane of the 3-lane highway, traveling 65 mph. The desired lane position and speed was displayed in red in the upper left corner of the forward channel screen, and whenever there was a lane or speed change, the experimenter verbally reminded the participant to adjust their driving to match the text displayed. The subject changed speed and switched lanes 3 times per block. At approximately the midpoint of each trial, the subject was asked to rate (aloud) the workload of a 15-s interval, with the beginning and end given by an audio recording as had been done in the practice block. The request was an audio file played back by the computer, triggered at exactly the same place for every subject.

After completion of Block 1, Block 2 was then run. Block 2 contained 27 trials, 13 of SAVE-IT replications and 14 transition trials, 1 more than the previous block (Appendix

F). The order of the Blocks was counterbalanced (see Appendix G). Each trial block (1300 m) took about 42s to complete when driving at 112 km/hr (70 mi/hr), though for many of the trials, the speed was either 60 or 65 mi/hr. The speed change was needed to properly position the vehicles and establish the desired workload.

Upon completion of Block 2, the participant was shown 18 30-second video clips from the SAVE-IT project in a counterbalanced order across subjects (Appendix G). Simulations of all of these clips had been driven in Blocks 1 and 2.

The participant was then thanked and paid. The experiment took 1.5 hours to complete on average.

Table 6 shows the estimated time for each portion of this experiment.

Table 6. Sequence of Activities

#	Activity	Duration (min)	Comment
1	Instructions	10	1. Complete Biographical Form 2. Complete Consent Form 3. Check vision
2	Familiarize with simulation	15	1. Show simulator 2. Motion discomfort warning 3. Drive 3 minutes for practice 4. Explain workload rating 5. Drive 5 minutes to practice rating workload
3	Test block 1	20	Drive 1st sequence of encounters and rate workload of each while driving
4	Test block 2	20	Drive 2nd sequence of encounters and rate workload of each while driving
5	Debrief	10	1. Rate workload of video clips 2. Thank and pay subject

Road Scenes Simulated

The test road was an expressway with 3 lanes in each direction, separated by a wide, grass median. Lanes were about 12-feet (3.6 m) wide with an outside shoulder of 10 feet (3.0 m) and an inside shoulder of 4 feet (1.2 m) as is standard practice. The road was perfectly flat and there were no crosswinds or other disturbances. There were no curves or entrances or exits (to minimize potential motion discomfort, a concern for older drivers). The road design complied with MUTCD (U.S. Department of Transportation, 2009) and AASTHO guidelines (AASHTO, 2004). Roads were constructed from tiles in the DriveSafety Library.

As was noted earlier, scenarios modeled were chosen to represent a wide range of scenarios reported by Schweitzer and Green (2007), varying in terms of the lane driven (left, middle, and right lane), and level of service (LOS).

To provide some sense of repeatability, 2 road segments were shown three times in Block 1, and a second pair was shown in Block 2. The goal was to have at least 1 clip for each lane position (left, middle, right for each of the 3 Levels of Service examined (A, C, E – Table 7).

Table 7. Estimated and Rated Workload of Test Scenarios

	Lane Driven	**Left Lane**			**Middle Lane**			**Right Lane**		
	LOS	**A**	**C**	**E**	**A**	**C**	**E**	**A**	**C**	**E**
Prior study rating from subjects	Block 1	2.8	3.8	6.7	2.4	4.1	5.7	2.3	3.7	6.8
	Block 2	3.6	5.1	6.5	3.4	4.8	5.2	2.9	3.5	4.8

Due to the capabilities of the driving simulator expressway scenario generator available when this study was conducted (Schweitzer and Green, 2009), there can only be a total of 3 vehicles in the scenario other than the subject vehicle that can be independently controlled (2 side vehicles and 1 lead vehicle). The software does allow for 2 platoons of 3 vehicles (1 ahead and 1 behind) also to be shown. The vehicles in the platoons were controlled as a group, not individually. Thus, the scenarios in the SAVE-IT videos could not be duplicated exactly, as some of them involved more than 3 vehicles serving as traffic.

As a workaround, the scenario generator allows for a platoon of lead vehicles, 1 per lane, ahead of the subject. This platoon normally is some distance from the subject and can only be encountered well above the speed limit. The vehicles in the platoon travel roughly side by side, though there is some shifting of position as would naturally occur while driving. The purpose of the platoon is to form a barrier that the subject cannot drive through or around, and thus limit the maximum speed the subject can drive, but to do so in a natural manner. In the high workload situations (LOS E), this platoon was brought close to the subject, simulating heavy traffic, though the positions did not exactly match those in the SAVE-IT clips.

Also, the clips were recorded at 1Hz, which makes it difficult to determine the exact progression of events, such as lead and side vehicle acceleration/deceleration and subject speed. In addition, the side vehicle's headway distances as well as the subject's speed were estimated as inputs to the simulator. As the reader can tell, the process of developing driving scenarios that matched the video clips was complex. When viewed side-by-side, the clips replicated matched the driving simulator representations of them. Figures 5 and 6 show some static examples. Compare the top left of the quad split image (the forward scene of the driving simulator) with the wider field of video still (from SAVE-IT) at the bottom of the figure.

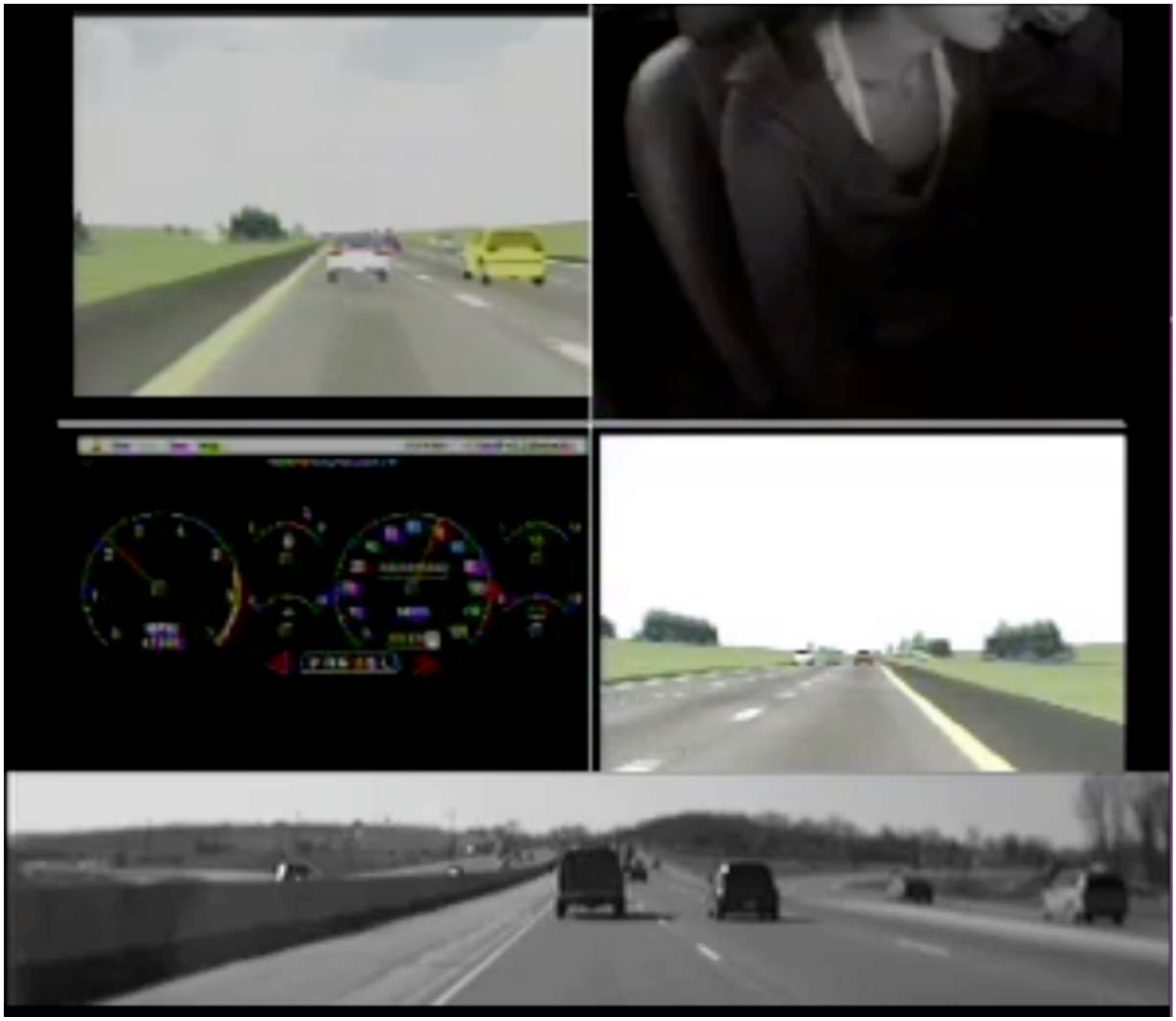


Figure 5. Scenario Example 1
Note: The 4 quadrants are forward scene (simulated, upper left), driver, speedometer cluster, and rear scene (lower right). The image at the bottom of the screen is from a forward scene clip from ACAS.

Figure 6. Scenario Example 2

Test Participants

The subjects tested consisted of 16 licensed drivers, with equal number ages 18-30 and over 65. In each age group, there were an equal number of females and males. Subjects were recruited either through an advertisement placed on Craigslist, as well as friends and acquaintances of experimenters, and through a list of past participants. They were paid $45 for 1-1/2 to 2 hours of their time.

One would expect university studies to include students as subjects, which was the case for only 5 people. Subjects drove from 500 to 20,000 miles per year, with a mean of 10,200. They were reasonably well educated, with all but 1 having at least some college. They were quite typical in their risk seeking behavior. When asked which lane they would drive in for an expressway with 3 lanes in each direction, 2 selected the left lane, 10 selected the middle lane, and 4 selected the right lane. Of those 16 subjects, 4 had been in a crash in the last 5 years, and over the last 2 years, 7 had tickets for traffic violations, mostly for minor infractions (five mi/hr over the speed limit). Except for one subject, all had corrected far acuity of 20/40 or better, and none had substantial color vision deficiencies.

UMTRI Driving Simulator

The experiment took place after the first major upgrade of the third-generation UMTRI driving simulator (www.umich.edu/~driving/sim.html). The simulator consists of a full-size cab, 10 computers, 6 video projectors, 7 cameras, audio equipment, and other items. The main functions (generating scene graphics; processing steering wheel, throttle, and brake inputs; providing steering wheel torque feedback; and saving data) were controlled by hardware and software provided by DriveSafety (Vection and HyperDrive Authoring Suite, version 1.6.2), software used at several universities and companies in the U.S.

Figure 7 shows the simulator cab and a typical forward scene from a practice drive. The simulator has a forward field of view of 200 degrees and a rear field of view of 40 degrees created by 5 forward channels and a rear channel. Each channel was 1024 x 768 and updated at 60 Hz. Depending on where the subject sat after adjusting the seat, the forward screen was 16 to 17 ft (4.9 to 5.2 m) from the driver's eyes, close to the 20-ft (6-m) distance often approximating optical infinity in accommodation studies.

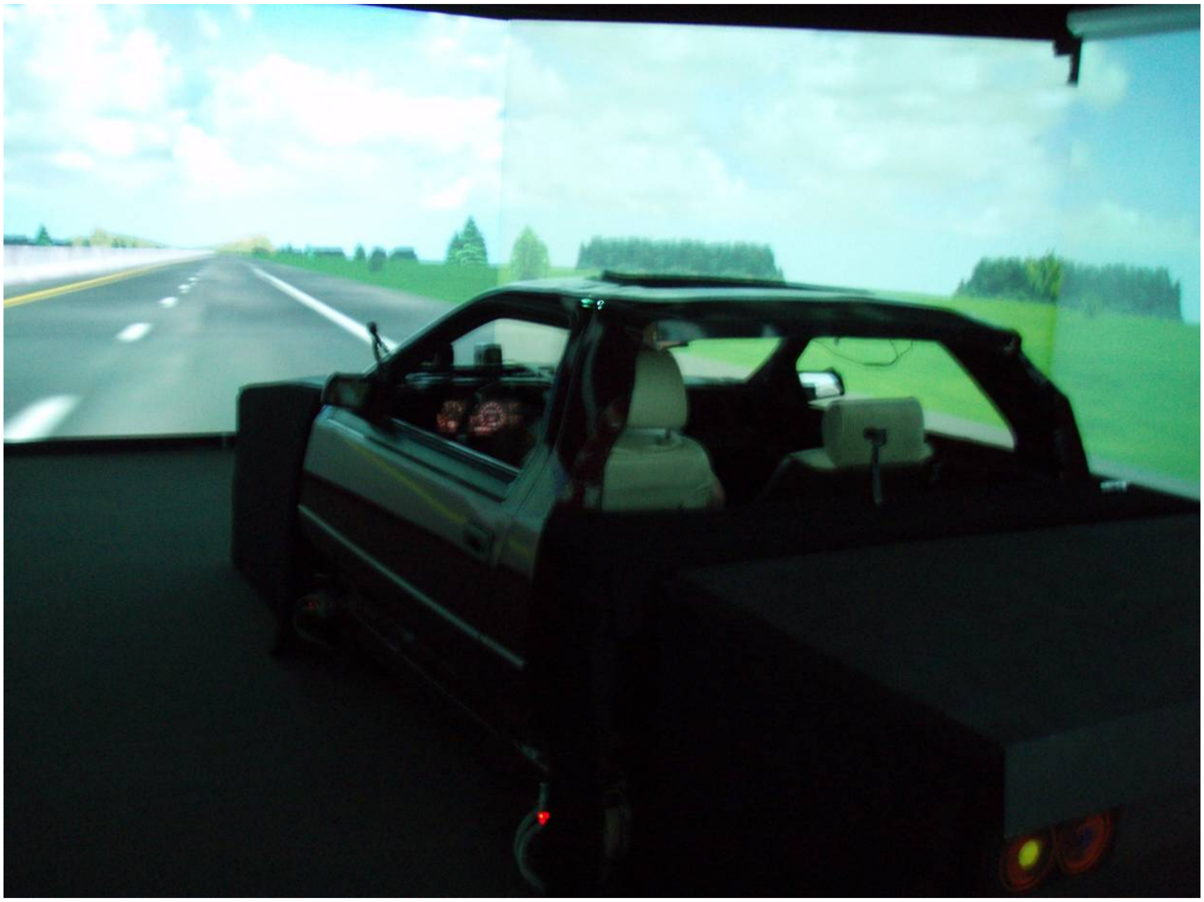


Figure 7. Simulator cab, front screen, front-right screen, and front-side screen

The simulator was controlled from an enclosure behind and to the left of the cab. The enclosure contains 4 quad-split video monitors that show the output of every camera and computer in the mockup, a display that shows the quad-split combination being recorded, a keyboard and LCD monitor for the driving simulator computers, and a second keyboard and LCD monitor to control the instrument panel and warning and scenario control software (Figure 8). Also in the enclosure was a 19-inch rack containing audio and video equipment (audio mixers, video patch panel and switchers, distribution amplifiers, DVD recorder, quad splitter, etc.) and two separate racks for the instrument panel and touch-screen computers, the simulator host computers, and the 6 simulator image generators. The instrument panel and center console computers ran the Mac OS, the user interface to the simulator ran Windows, and the simulators ran Linux.

Figure 8. Simulator operator's workstation

The vehicle cab consisted of the A-to-B pillar section of a 1985 Chrysler Laser with a custom-made hood and back end mounted on casters for easy access. Mounted in the mockup were operating foot controls, a torque motor connected to the steering wheel (to provide steering force feedback), an LCD projector under the hood (to show the speedometer-tachometer cluster), a 10-speaker sound system (for auditory warnings), a haptic seat, 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 was from a 2002 Nissan Altima and was installed in the A-pillars, lower door panels, and behind each of the two front seats. The stock amplifier (from the 2002 Nissan Altima) drove the speakers.

The speedometer-tachometer display was controlled by a Macintosh computer running REALbasic and looked similar to those in an early 1990s Honda Accord.

Mounted in and around the cab were 8 video cameras. Images included the driver's face (viewed from outside and inside the cab), 2 over-the-shoulder images (showing the instrument panel), an image from the package shelf showing the instrument panel and forward scene, an image of the feet and pedals, and an image from a "floater," a camera on a tripod that could be positioned anywhere. These images, combined with output from any of the projected images, could be recorded on videotape using a quad splitter. Real-time audio and video of simulator activity was available via a web camera mounted above the simulator control enclosure.

Figure 9 shows a close-up of the cab interior. A unique feature of the simulator is the computer-generated, back-projected speedometer-tachometer cluster.

Figure 9. View of the inside of the simulator cab
Note: The instrument panel configuration is from a prior study.

For additional information on the simulator see Green, Sullivan, Tsimhoni, Oberholtzer, Buonarosa, Devonshire, Schweitzer, Baragar, and Sayer (2008).

RESULTS

How the Rating Data Obtained While Driving Were Reduced and Analyzed

As a reminder, each of the 16 subjects provided ratings for 53 scenarios, 26 of which duplicated scenarios in the previous SAVE-IT project, and 27 of which were transitions between those scenarios. Of the 26, 4 clips were repeated 3 times to provide estimates of within subject reliability, so there were 14 unique stable driving scenarios. Further, the workload of these 14 scenarios (presented as video clips) were rated in the SAVE-IT project, and were rated again in this experiment by a different group of subjects, as well as being driven in this experiment.

As the workload model being developed was for steady state situations, transition scenarios in which the lead vehicle changed during the rating period (4, 16, 24, and 26, 15, 17, 25, 36, and 45, Table 8, were removed to provide a more appropriate data set. Keep in mind that it was important to have lane changes in the test set as the experimental goal was to convey the impression of real driving, a situation in which lane changes normally occur. There were 36 scenarios retained for analysis (including 22 transition scenarios and 14 testing scenarios).

Table 8. Scenarios Not Included in the Analysis

Scenario		Clip	Driven lane	Replicates	Reason to exclude
Type	#				
From SAVE-IT	04	135	Middle	3	Right side vehicle cut-in to the middle lane from the ramp
	16	125	Right	1	Lead vehicle merged to the middle lane
	24	139	Middle	3	Right side vehicle cut-in to be the lead one
	26	138	Middle	3	Left side vehicle cut-in to be the lead one
Transition	15	-	Right	1	No lead vehicle in this scenario
	17	-	Right	1	Right side vehicle cut-in to be the lead one
	25	-	Middle	1	Lead vehicle merged to the right lane
	36	-	Middle	1	Lead vehicle merged to the right lane
	45	-	Middle	1	Lead vehicle merged to the right lane

How Were the Workload Ratings Distributed Overall for the 44 Driving Scenarios Not Involving Lane Changes and a Lead Vehicle Is Present?

There were 44 scenarios (53 – 9) not involving lane changes. As shown in Figure 10, most workload ratings were between the anchors of 2 and 6. ,The overall distribution appears lognormal, with a mean of 3.6 (S.D.=1.6). Subjects favored integer ratings (87% = 505/579 responses). The expected value, 50%, was significantly different ($t(578)=26.8$, $p<0.001$). (Note: As clips were well distributed across the range of workload, integer and non-integer ratings should occur equally often.) This quantization increases the error the subjective estimates and how amount of variance in those estimates of which driving performance measures will account. This outcome also suggests the need to modify the instructions in future studies to encourage subjects to use non-integer ratings if they can, which may not be the case. (Subjects should be asked how precisely they are able to rate workload.)

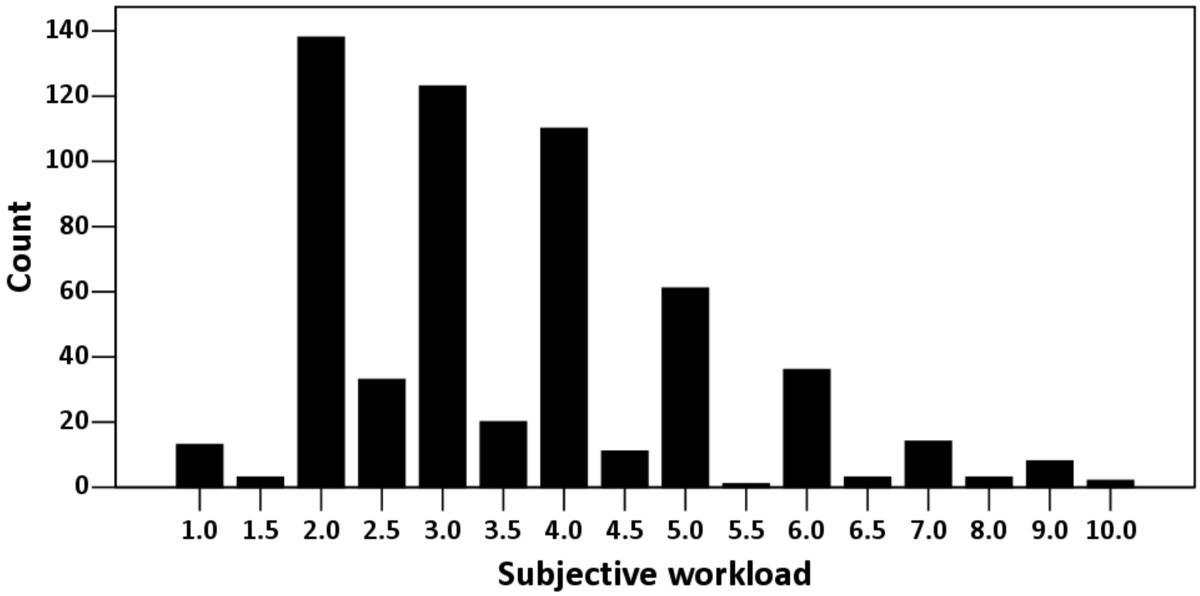


Figure 10. Distribution of Workload Ratings for All Trials Not Involving Lane Changes

Not all subjects used the full range of the rating scale (Figure 11). In particular, subjects 7 (young female), 9 (old male), 13 and 14, (both old female) only used ratings over a 2-point range (e.g. 1 to 3 or 2 to 4), possibly because they did not completely understand the instructions. This outcome was a surprise to the experimenters and suggests the need to change the instructions to emphasize use of the anchors, especially the upper anchor. There was some discussion about replacing these subjects. However, as shown later, even these subjects sometimes produced useful results.

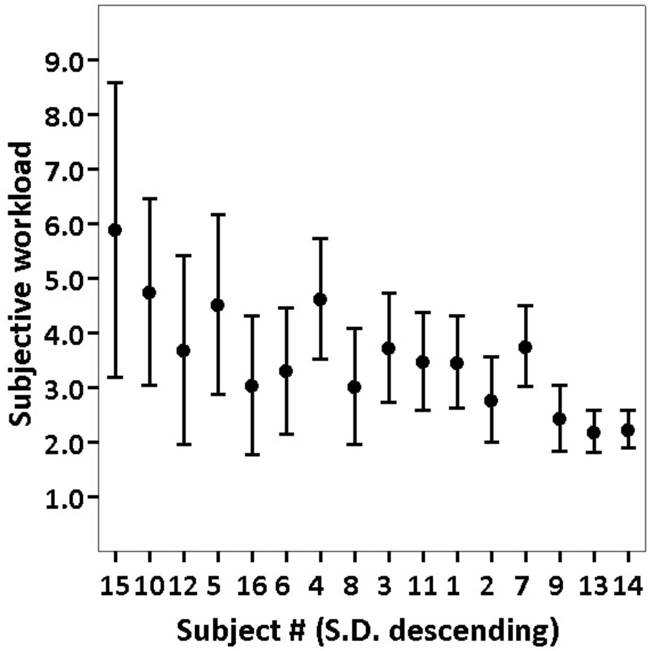


Figure 11. Workload Ratings Ordered by Standard Deviation

The absolute level of workload experienced should vary with an individual's capability to deal with it, which should vary with age and gender. However, in theory, as the workload ratings were relative to anchors, differences due to individuals, age, and gender should be minimal. In all cases, the workload rating distributions were always left-skewed. Specifically, as shown in Figure 12, ratings for men were greater than those for women (male mean 3.6, S.D. 1.3; female mean 3.5, S.D. 1.7). Young subjects (mean 3.7; S.D. 1.23) had higher workload ratings than old ones (mean 3.5; S.D. 1.9).

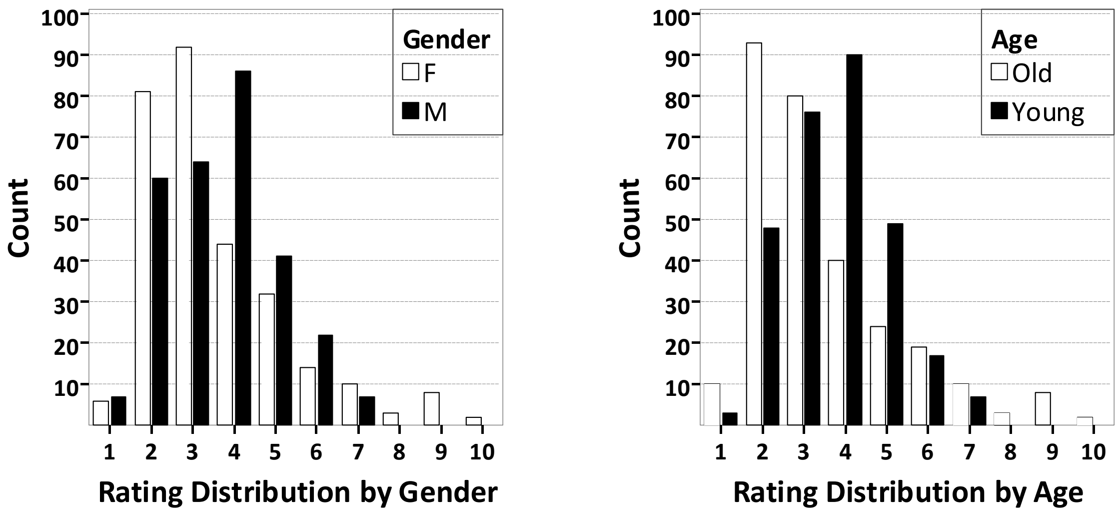


Figure 12. Distribution of Workload Ratings by Gender and Age

However, as shown in Figure 13, the major difference appears to be an interaction between age and gender, with older women having somewhat lower ratings. Considering that ratings were to the nearest 0.5, this is not a substantial difference, and an ANOVA of the ratings (with age, gender, and the age x gender interaction as main effects) shows that none of these differences were significant at the $p<0.05$ level (Gender – $F(1,575)=1.35$, $p=0.47$; Age – $F(1,575)=4.75$, $p=0.17$; Interaction – $F(1,575)=1.25$, $p=0.48$). Furthermore, as was noted earlier, there were concerns that 2 of the 4 older women may not have understood the instructions, using consistently lower ratings than other subjects.

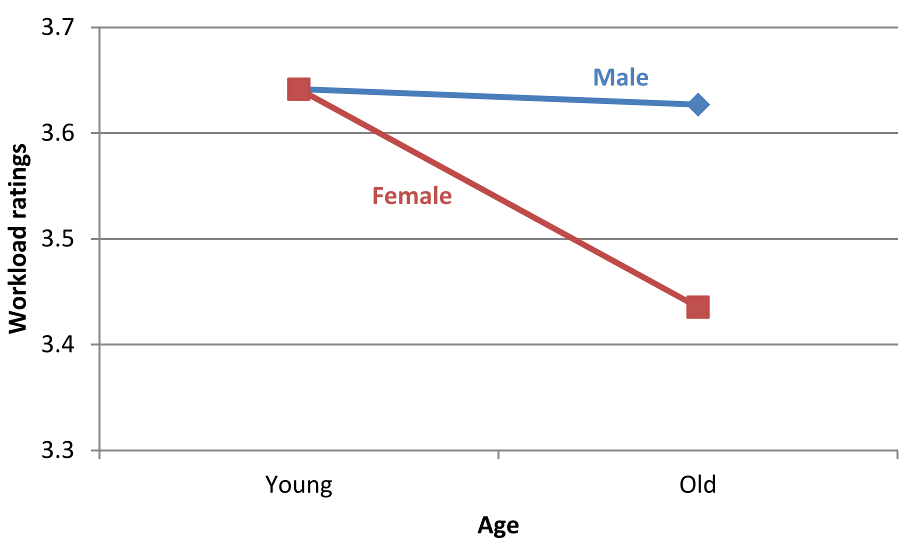


Figure 13. Mean Workload Ratings by Age and Gender

How Were the Workload Ratings Distributed from Watching the Video Clips of the 14 Driving Scenarios from SAVE-IT?

As the 14 scenarios repeated from SAVE-IT were a subset of the 44 examined in Figure 10, the distribution of the 14 repeated scenarios should be similar to Figure 10 (for the 44), which is the case. (The scenarios and their corresponding video clip numbers are shown in Appendix F.) In Figure 14, the right-skewed distribution was similar to Figure 10, with the mean workload rating of 3.7. As shown in Figure 15, values were greater for men than women as before, (F(1,220)=3,39, p=0.067). However, now the value for young subjects was significantly greater than for old subjects (F(1,220)=6.93, p=0.009). The interaction was not significant (F(1,220)=1.02, p=0.314).

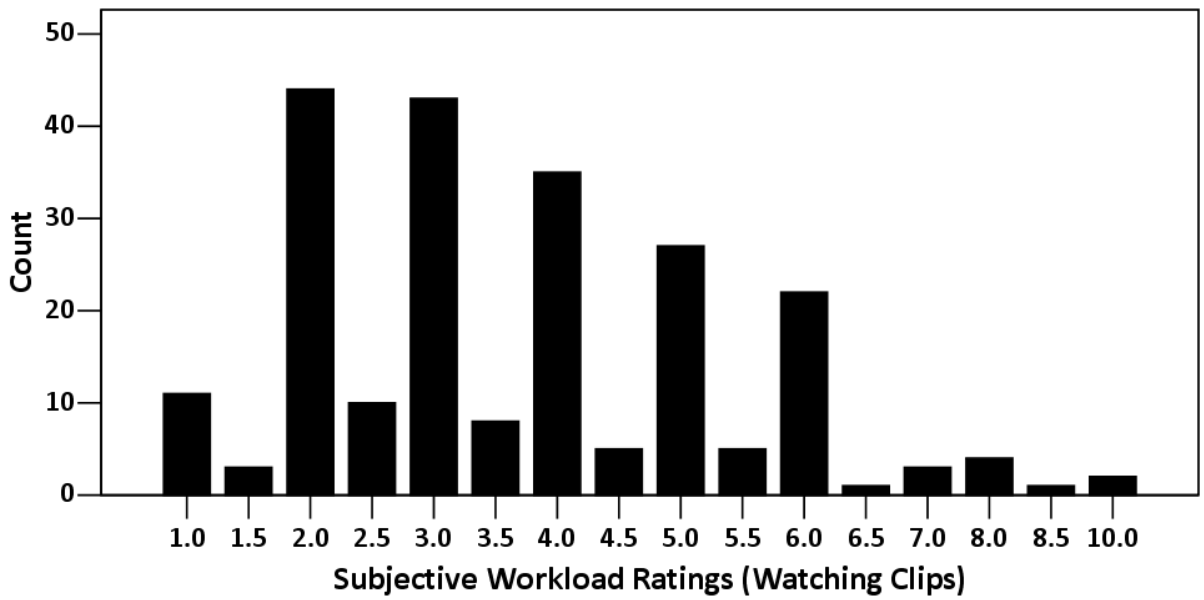


Figure 14. Distribution of Workload Ratings for Watching Video Clips: This Experiment

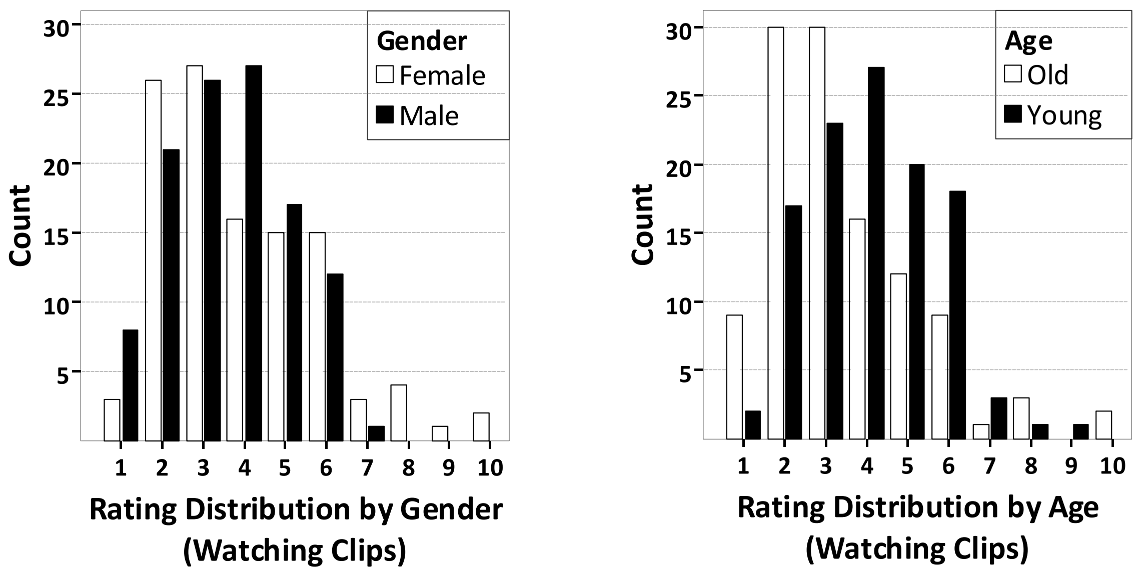


Figure 15. Distribution of Workload Ratings for Watching Video Clips by Gender and Age

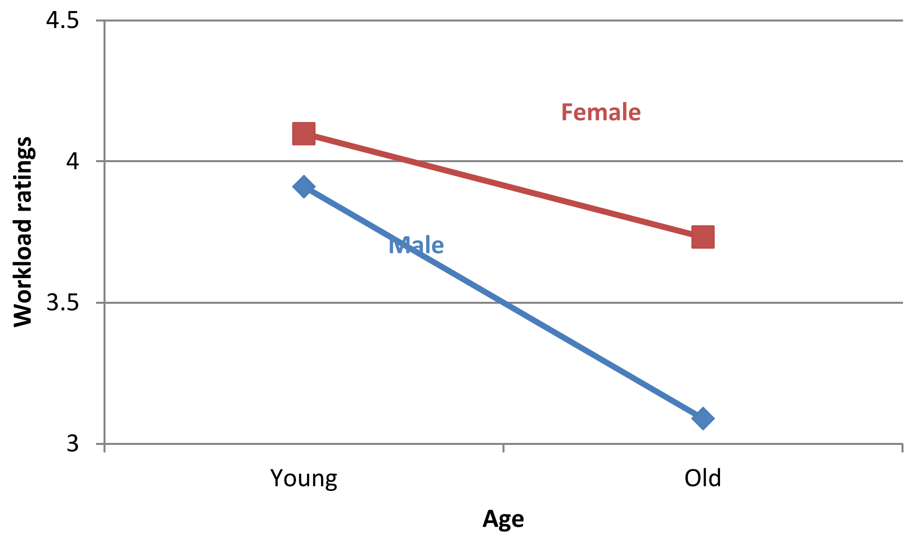


Figure 16. Mean Workload Ratings for Watching Video Clips by Age and Gender

How Were the Workload Ratings Distributed Overall from the 22 Transition Trial Ratings?

The underlying issue was whether the transition trials should be included in the equation development. As was noted previously, the unstable driving situations (4 scenarios from SAVE-IT and 5 from transition scenarios) were eliminated from the regression equation development as they represent situations for which the equation was not designed. Thus, remaining were 36 scenarios for 16 subjects, of which 22 were transition trials. As a reminder, transition trials were inserted between test trials duplicating scenarios from the SAVE-IT videos, serving to reposition vehicles/set up the next test scenario. Accelerations tended to be greater than normal and traffic flow as unstable, going beyond the bounds of the prior work (and the prior equations). Thus, the plan was to analyze the data both with and without the 22 transition trials.

Interestingly, as shown in Figure 17, eliminating those trials did not lead to a large change in the overall distribution. The mean was about the same (3.5), and the distribution was similar to the overall distribution of which it is a part (Figure 10) and the distribution of the trials from SAVE-IT (Figure 14). None of the effects were not significant (Gender – $F(1, 251)=1.58$, $p=0.44$; Age – $F(1, 251)=0.307$, $p=0.74$; Interaction – $F(1, 251)=0.256$, $p=0.76$), similar to results before removing the transition trials. See Figure 18.

To the authors, these analyses suggest there are no radical differences in the transition trials and they should be considered for inclusion in the workload equation development. As a practical matter, including these trials roughly doubles the size of the data set. Further, differences are in the opposite direction of that when the 4 subjects of interest were included.

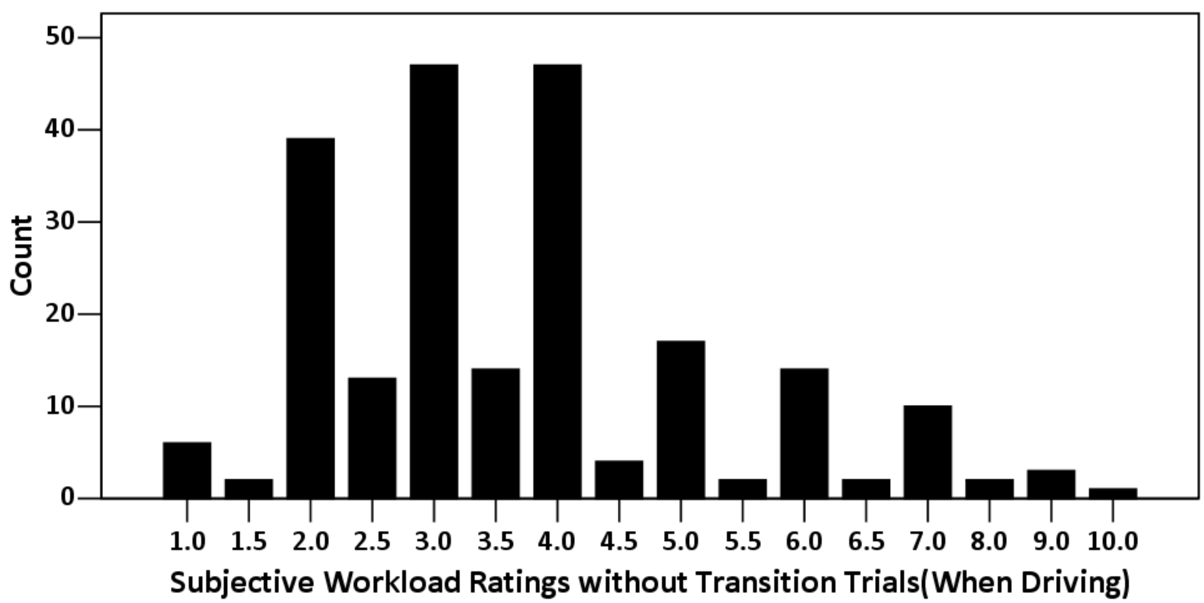


Figure 17. Distribution of Workload Ratings When Driving (22 transition trials excluded)

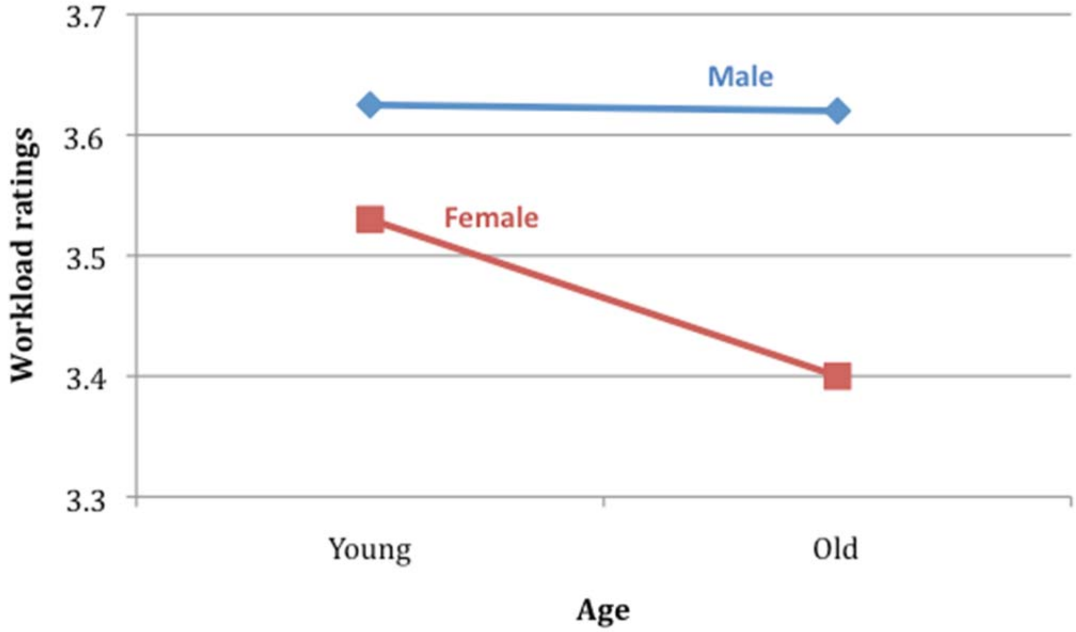


Figure 18. Mean Workload Ratings When Driving by Age and Gender
(22 Transition Trials Excluded)

Should the Data from the 4 Potentially Faulty Subjects Be Omitted?

The rating ranges for subjects 7, 9, 13, and 14 were so narrow (see Figure 11) that their ratings were largely unaffected by level of service (traffic), which is not consistent with what is known from widespread experience while driving and observations from the prior experiment.

Eliminating those 4 subjects from all driving data led to significant differences in the workload ratings as a function of age ($F(1, 433)=17.874$, $p<0.001$), as shown in Figure 19. However, the effects of gender and the age x gender interaction were not significant ($F(1, 433)=2.79$, $p=0.096$; F(1, 433=3.412, p=0.065)). Keep in mind that 3 of the subjects were old females and the fourth was a young female, leading to an unbalanced and weak analysis.

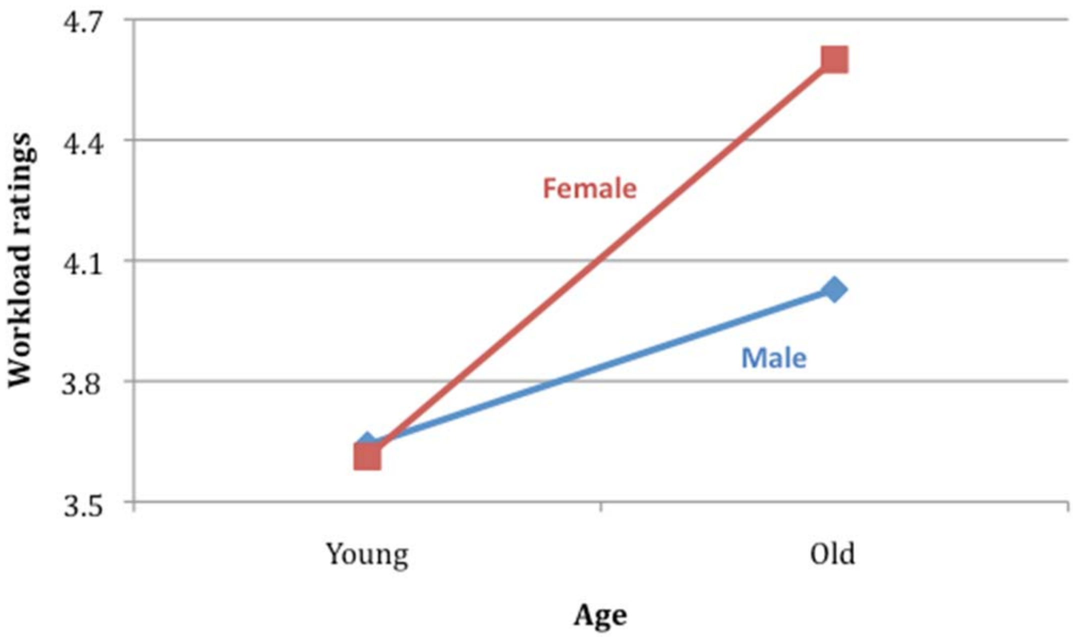


Figure 19. Mean Workload Ratings When Driving by Age and Gender
(subject 7, 9, 13, and 14 excluded)

If subjects understood the instructions, then their ratings while watching video clips should resemble those while driving. The correlations (Table 9) fall into 3 groups, (1) essentially no relationship or low correlations (0.3 or less, 2 subjects), (2) moderate positive correlations (0.5 to 0.7, 6 subjects), and (3) high positive correlation (0.7 or above, 8 subjects). For subjects 14 and to a lesser extent subject 9 (both old), there was essentially no correlation between the ratings while watching clips and while driving those same scenarios. This suggests deleting those 2 subjects.

Table 9. Correlations of Workload Ratings by Subject:
Driving Versus Watching Video Clips

Subject	**5**	**3**	**1**	**2**	**10**	**11**	**13**	**15**
Age	Young	Young	Young	Young	Old	Old	Old	Old
Gender	Female	Male	Male	Male	Male	Male	Female	Female
r	0.88	0.84	0.83	0.82	0.81	0.75	0.75	0.71
Subject	**12**	**8**	**6**	**4**	**16**	**7**	**9**	**14**
Age	Old	Young	Young	Young	Old	Young	Old	Old
Gender	Male	Female	Female	Male	Female	Female	Male	Female
r	0.68	0.68	0.68	0.67	0.58	0.52	**0.28**	**0.06**

If subjects understood the instructions and the method is repeatable, then their ratings for watching clips should agree with the prior mean SAVE-IT workload ratings for the same clips. As shown in Table 10, there were positive correlations for all subjects, though for subjects 9 and 14 they were very low, much lower than all other subjects in the sample. Notice that for many subjects, the correlation was greater than 0.75, accounting for half of the variance. As a reminder, in this experiment, subjects strongly favored integer over non-integer ratings, which leads to lower correlations.

Table 10. Correlations of Workload Ratings from Watching Clips:
This Experiment Versus SAVE-IT

Subject	**15**	**5**	**6**	**10**	**1**	**7**	**12**	**8**
Age	Old	Young	Young	Old	Young	Young	Old	Young
Gender	Female	Female	Female	Male	Male	Female	Male	Female
r	0.93	0.93	0.90	0.87	0.87	0.86	0.84	0.83
Subject	**2**	**13**	**16**	**3**	**11**	**4**	**14**	**9**
Age	Young	Old	Old	Young	Old	Young	Old	Old
Gender	Male	Female	Female	Male	Male	Male	Female	Male
r	0.82	0.82	0.82	0.78	0.76	0.67	0.62	0.48

If subjects understood the instructions, then their ratings for watching clips in this experiment should agree with other 15 subjects in the same sample, a jackknife-like procedure. As shown in Table 11, the ratings for subjects 9 and 14 were lower than those for all other subjects by 0.15 (with subject 16). The largest gap in correlation between any other pair of subjects was 0.05 (between 11 and 4).

Table 11. Correlations of Ratings from Watching Video Clips (1 vs. 15)

Subject	**6**	**10**	**5**	**1**	**15**	**7**	**3**	**2**
Age	Young	Old	Young	Young	Old	Young	Young	Young
Gender	Female	Male	Female	Male	Female	Female	Male	Male
r	0.93	0.92	0.92	0.91	0.90	0.87	0.87	0.84
Subject	**8**	**12**	**13**	**11**	**4**	**16**	**14**	**9**
Age	Young	Old	Old	Old	Young	Old	Old	Old
Gender	Female	Male	Female	Male	Male	Female	Female	Male
r	0.84	0.83	0.81	0.79	0.74	0.70	0.55	0.51

Similarly, if subjects understood the instructions, then their ratings for watching clips and driving should be similar to other subjects in the same experiment (Table 12). Again, there appears to be a gap between subjects 9 and 14, and all other subjects, but also a gap between subject 14 and 9.

Table 12. Correlations of Ratings from Watching Video Clips (1 vs. 15):
Driving Simulator

Subject	**3**	**12**	**13**	**10**	**5**	**15**	**4**	**1**
Age	Young	Old	Old	Old	Young	Old	Young	Young
Gender	Male	Male	Female	Male	Female	Female	Male	Male
r	0.92	0.89	0.84	0.83	0.82	0.81	0.80	0.79
Subject	**2**	**8**	**11**	**7**	**16**	**6**	**14**	**9**
Age	Young	Young	Old	Young	Old	Young	Old	Old
Gender	Male	Female	Male	Female	Female	Female	Female	Male
r	0.77	0.73	0.72	0.69	0.68	0.58	0.40	**0.17**

These data suggest that removing subjects 9 and 14 from further analyses is reasonable. There were very weak correlations for those subjects with other subjects in the prior sample and this sample, as well as their own ratings while driving. However, their data was clearly different from other subjects in either sample. Further, although both of these subjects are older, there do not seem to be strong age differences in the ratings, and given the ratings are anchored and therefore relative, those differences should not exist. Furthermore, comments from these subjects suggested they were rating the overall workload independent of the anchors, not the workload relative to the anchors

How Consistent within Subjects Were the Repeated Ratings of the Same Clips?

According to the trial sequence table in Appendix F, some trials were replicated 3 times. These scenario numbers are 2, 4, 24, and, 26. However, in Table 1, scenario number 4, 24, and 26 were removed before the analysis because the lead vehicle or traffic were not stable in these trials. Therefore, only scenario number 2 (clip number 40) will be analyzed in this section. Three replicates were assigned to be 2^{nd}, 14^{th}, and 24^{th} trials in the same block. Correlations between each pair of trials are 0.91 (2^{nd} vs. 14^{th}), 0.80 (14^{th} vs. 24^{th}), and 0.77 (2^{nd} vs. 24^{th}). Figure 20 shows the comparison of the 3 replicates, in which most subjects rated similar in his/her own trials. Subject 12 has a larger difference among her 3 ratings.

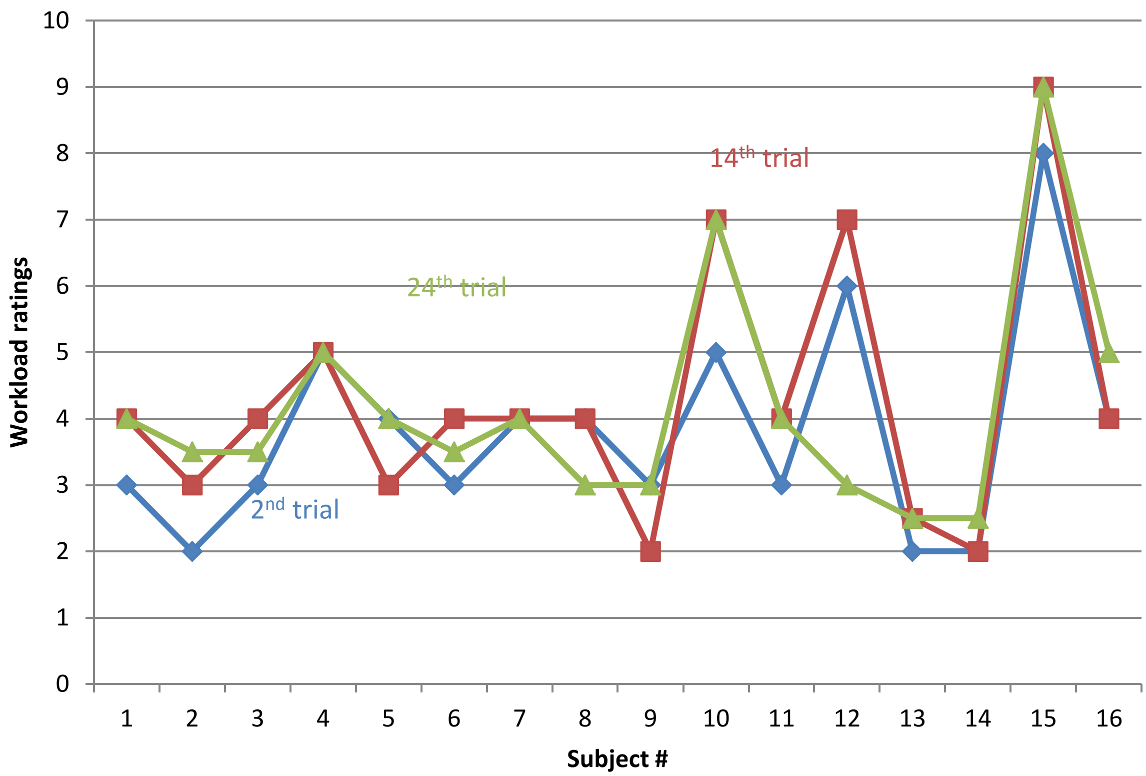


Figure 20. Comparison of 3 Replicates for Scenario 2 by Subject

How Well Did the New Workload Ratings from Watching Clips and Driving Agree with the Prior Ratings of the Same?

As the ratings process and clips used in this experiment and SAVE-IT were the same, then the ratings should agree. The analysis was first done for the 14 clips common to both experiments but without the transition trials, a close replication.

The mean ratings used in the calculations are shown in Appendix H. For the current study involving 16 subjects, driving scenarios were repeated 3 times, hence there were 48 data points for each of them, not 16 (1/subject). For scenario 34, there is 1 instance where testing was terminated too soon, and therefore, the rating from that subject for that trial was omitted.

As shown in Figure 21, there was a consistent trend for clip ratings from the prior experiment (SAVE-IT) to be greater than either the subjective ratings of the clips or of those scenarios while driven. The absolute difference between experiments declined with workload, though overall, the ratings from this experiment were about 80% of those in the SAVE-IT experiment. These differences could be due to the prior exposure to driving (which could seem more demanding) or subjects not following instructions to some degree, which was true for at least 2 subjects.

As shown in Figure 22, the correlations of mean workload ratings from watching clips and while driving those same scenes were highly correlated. The clip ratings in the SAVE-IT project (0.97 and 0.90, respectively) and the mean clip ratings with the ratings while driving this experiment (r=0.92) were all quite high.

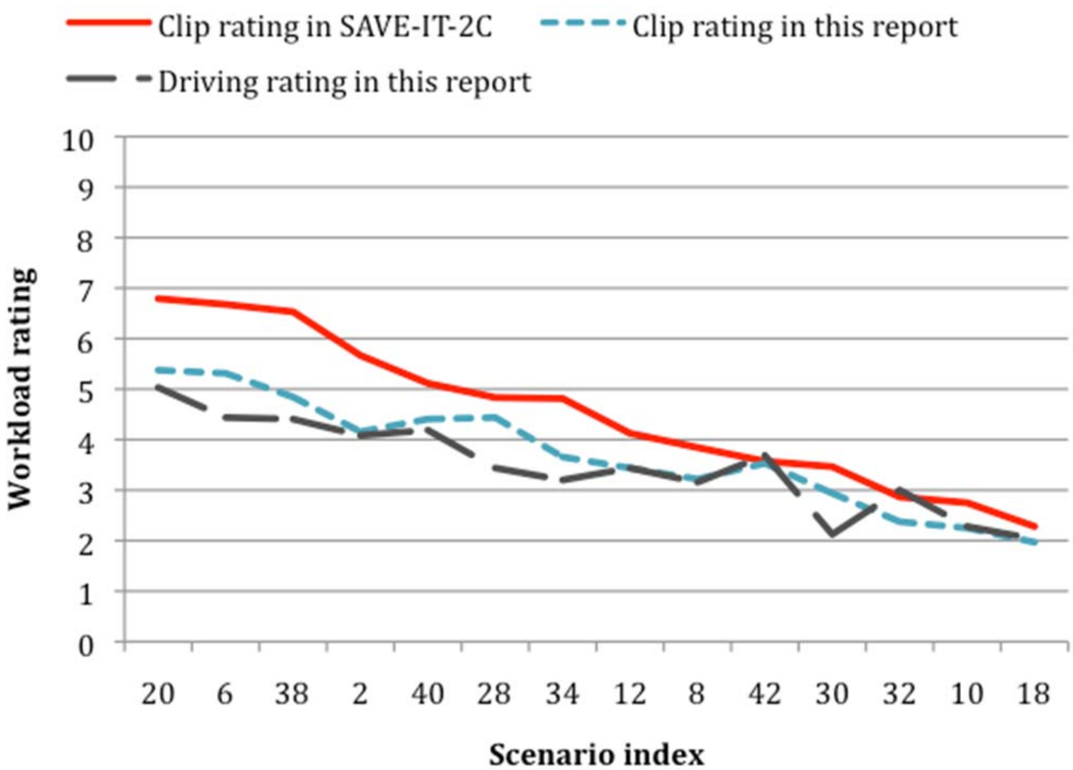


Figure 21. Mean Workload Ratings for Each Scenario (SAVE-IT vs. This Experiment)

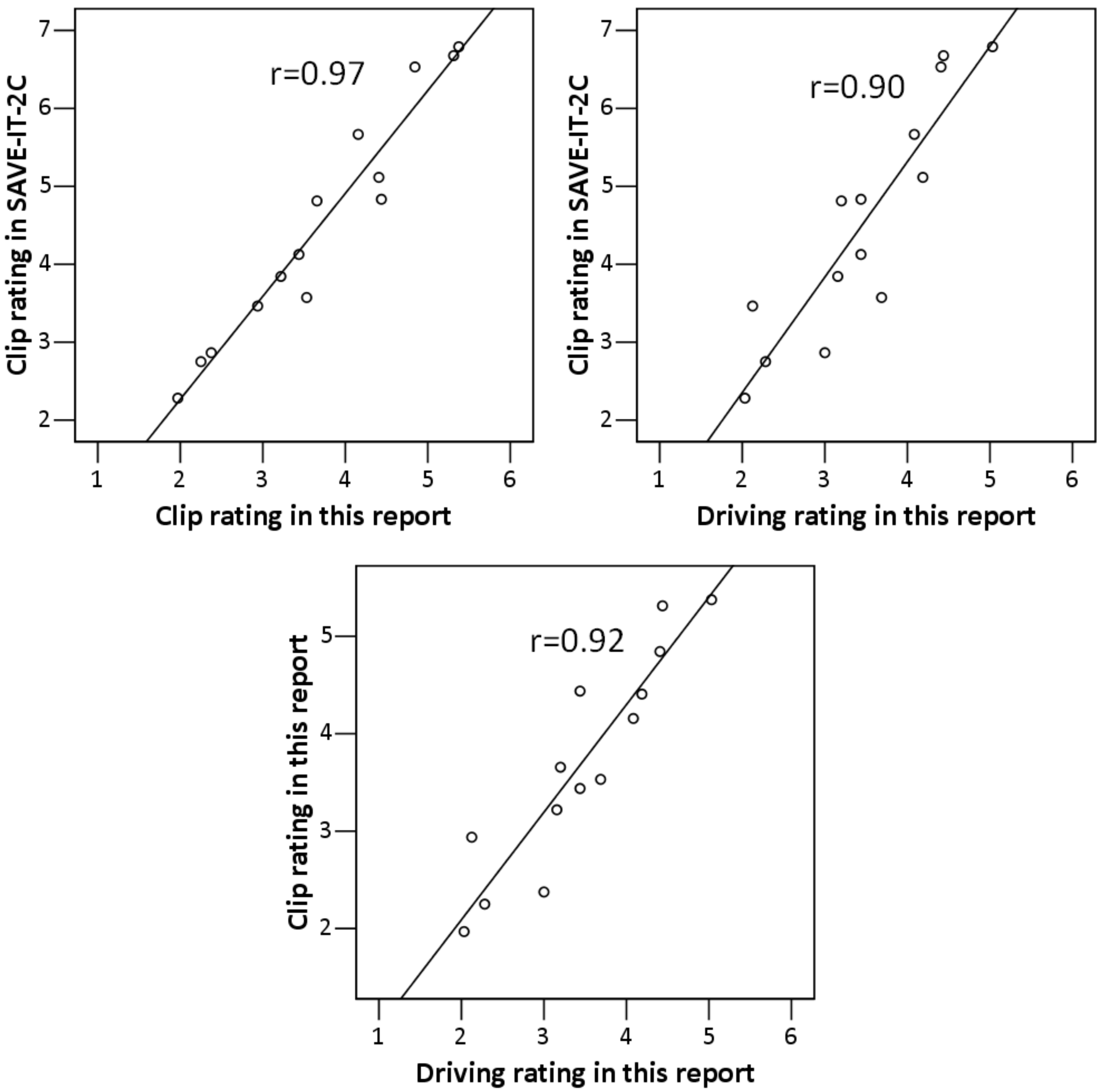


Figure 22. Correlation of Mean Workload Ratings:
SAVE-IT versus Clip and Driving Rating in This Experiment for All Subjects

Surprisingly, removing subjects 9 and 14 led to no change in the correlations of the clip workload ratings or the ratings of workload while driving (those scenes) with the previous SAVE-IT workload ratings of clips, and with the current ratings (0.97, 0.90, and 0.92, respectively), which were very high and close to the results before removing the 2 subjects. However, the absolute difference in the ratings between studies was reduced. See Figures 23 and 24. Deleting those 2 subjects was expected to improve the correlations. Thus, it is not obvious if these 2 subjects should or should not be included in further calculations.

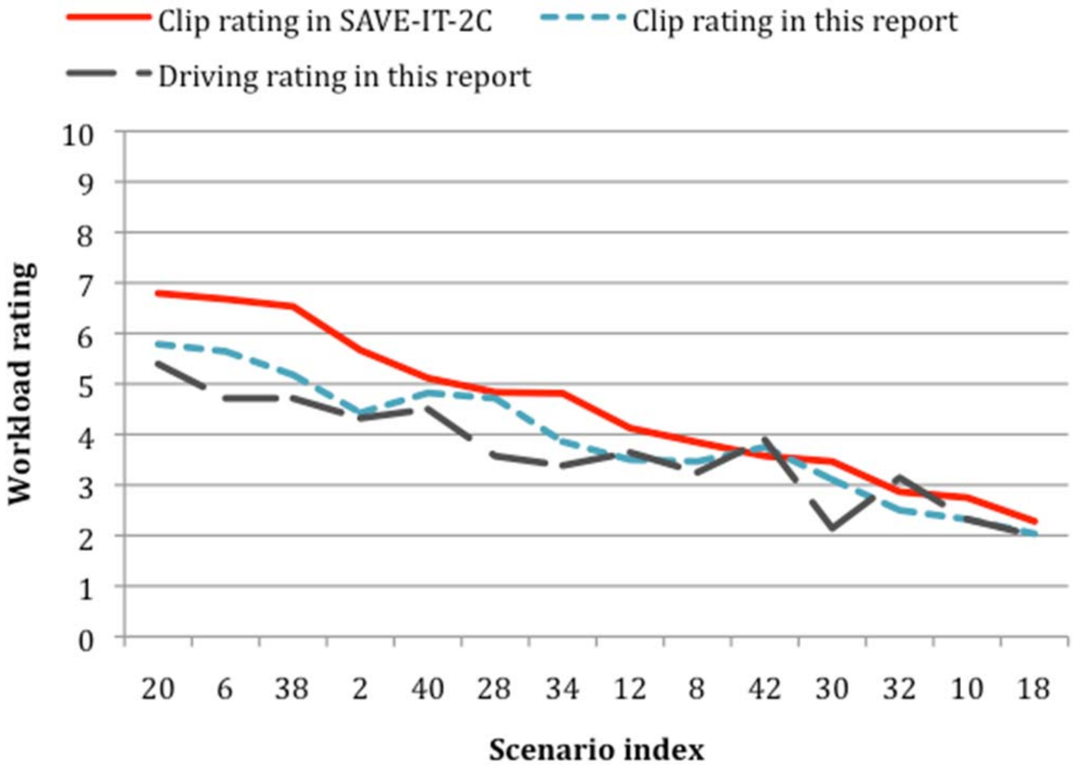


Figure 23. Mean Workload Ratings for Each Scenario:
SAVE-IT vs. Clip and Driving Rating for This Experiment, 2 Subjects, Removed

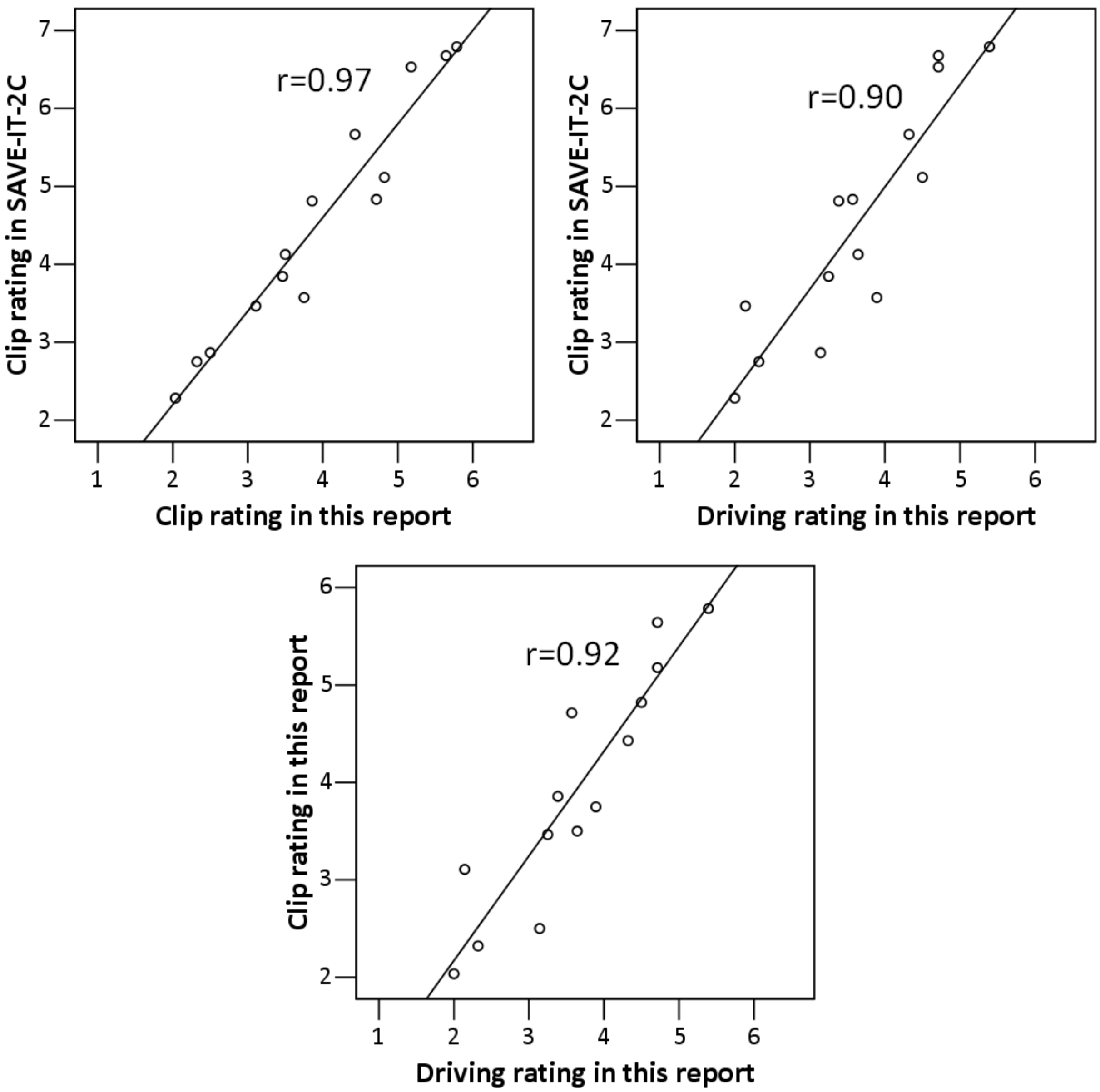


Figure 24. Correlations of Ratings for Each Scenario:
SAVE-IT vs. Clip and Driving Rating for This Experiment, 2 Subjects, Removed

How did traffic-related factors affect workload ratings while driving the scenarios?

Given the uncertainties about which data to consider, these ratings were examined in 2 ways: (1) with just the 4 scenarios with lane changes removed (including the previously rated clips and transition trials), and (2) with only the 14 test trials from before for the 14 best subjects in this experiment (subjects 9 and 14 excluded).

Generally, there were up to 6 cars with which subjects could interact (1 lead vehicle, 2 side vehicles, and 3 that formed a barrier well ahead of subjects). Only vehicles located closer than a gap of 125 m and within a 15-degree field of view (the detection constraints of the SAVE-IT radar) were counted. Correlations of the mean workload rating with the mean, maximum, and minimum of vehicles detected by radar (traffic counts) were 0.65 (0.65), 0.65 (0.53), and 0.55 (0.52) for the 2 cases described. Interestingly, the correlations were slightly greater for the full data than the “reduced” data set, and there is no evidence of different correlation patterns complete and reduced data sets (Figure 25). Of these 3 correlations for each data set, the correlations for the minimum were lower, primarily because of quantization. (The minimum must be an integer.) These data suggest using either the mean or maximum traffic count as a predictor of workload, with the mean slightly favored as was the case in SAVE-IT (and leading to more consistently higher correlations here).

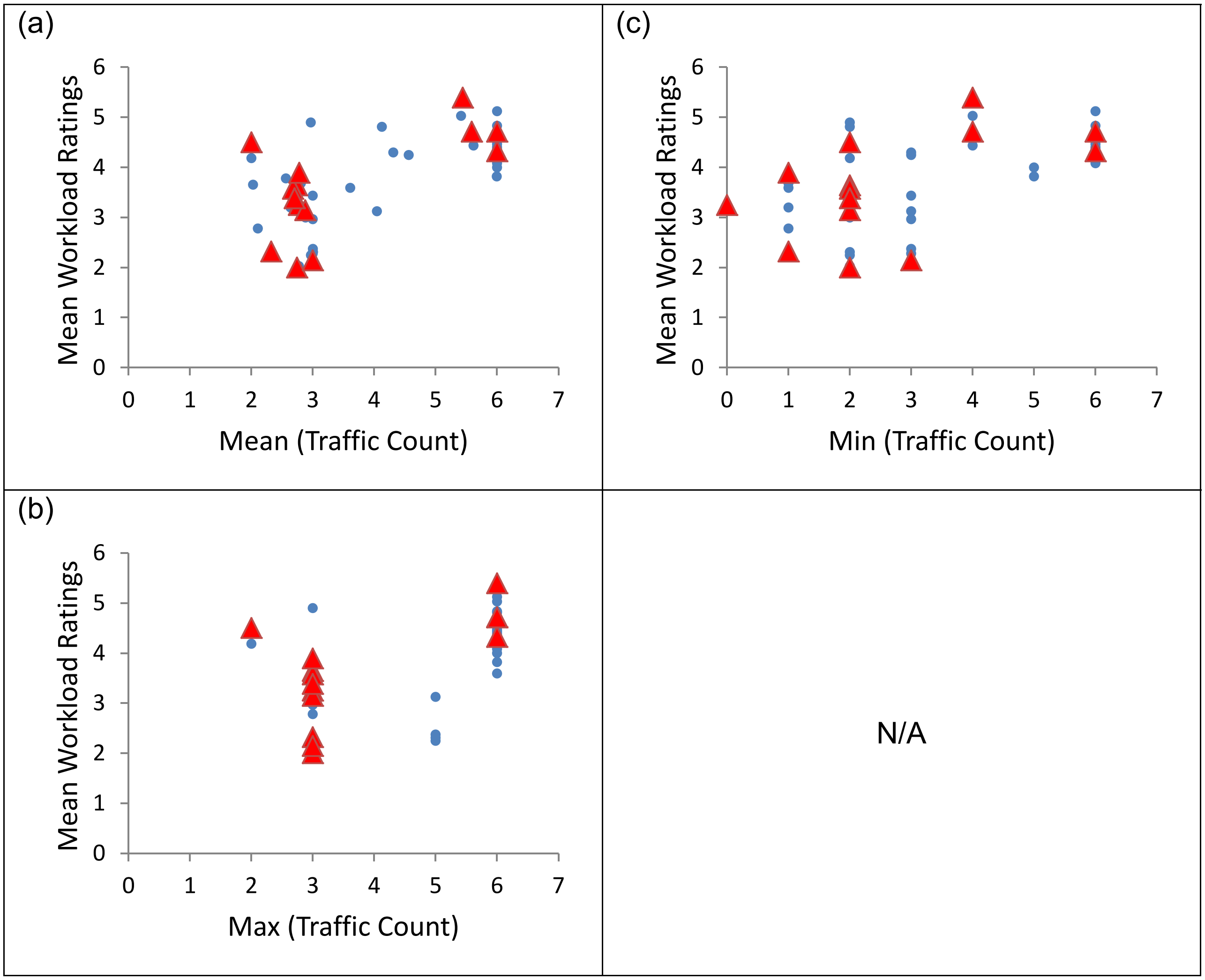


Figure 25. Mean Workload Ratings While Driving Versus (a) Mean, (b) Minimum, and (c) Maximum Traffic Count
Dots represent the mean workload ratings of the 16 subjects × 36 scenarios.
Triangles represent ratings of the 14 subject × 14 scenarios subset
(after tossing some subjects and scenarios).

Independent of other compensating factors (e.g., higher speed roads high greater radius curves), increasing speed should lead to increased workload, as the rate of information presentation is greater. For real roads, at a certain point, increasing traffic (which increases workload) decreases speed (which decreases workload).

For traffic, one could use the posted speed, which in the truest sense, is not the exposed speed. In this case, for simplicity, the lead vehicle speed was used as a surrogate for the speed of all traffic. (Subject vehicle speed is examined later.) As shown in Figure 26, the range of lead vehicle speeds is quite limited, with means of 24-34 m/s (54-76.5 mph), reasonable for an expressway with low to moderate traffic. This is consistent with the posted speeds, which were 65 or 70 mph (28.9 or 31.1 m/s, depending on the road segment.

Interestingly, for the full data set, workload decreased as speed increased. Correlations for mean, maximum, minimum and standard deviation of the lead vehicle speed were -0.37, -0.34, -0.37, and 0.01 for all the data, but there was essentially no correlation for the reduced data, (-0.01 for mean, 0.06 for maximum, 0.04 for minimum, and -0.01 for standard deviation). This suggests that lead vehicle speed (which varied over a limited range in this experiment) had no effect on driver workload.

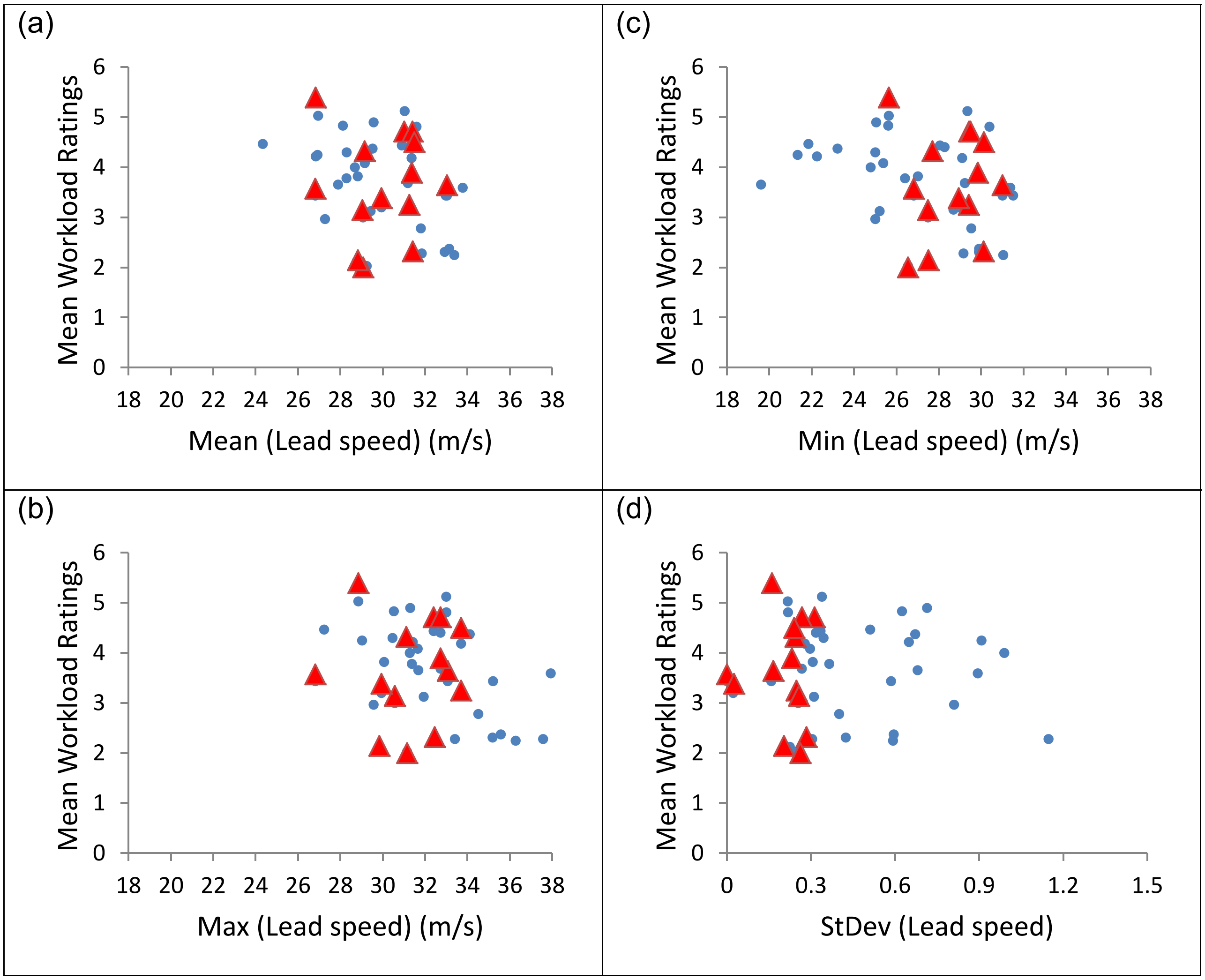


Figure 26. Mean Workload Ratings While Driving Versus (a) Mean, (b) Maximum, and (c) Minimum Lead Vehicle Speed and (d) Standard Deviation of The Speed

Instabilities in traffic flow are a source of workload. In the SAVE-IT project (Schweitzer and Green, 2007), the minimum of lead vehicle acceleration was a reasonable predictor of workload, with greater acceleration leading to greater workload. In this experiment, for the full data set (16 subjects and 36 scenarios), the correlations were 0.49 (mean), 0.16 (max), -0.01(min) and -0.21(stdev). Reducing the data set to 14 scenarios and 14 subjects led to correlations of 0.17 (mean), 0.13 (max), -0.06 (min), and -0.08 (stdev), essentially 0. The low correlations in the test scenarios partially reflect the stable and low acceleration situations examined (Figure 27). Nonetheless, that pattern is the same as that for traffic counts—the correlations with the 36 scenario – 16 subject combinations were greater, and in this case, the correlation was greatest for the mean lead vehicle acceleration.

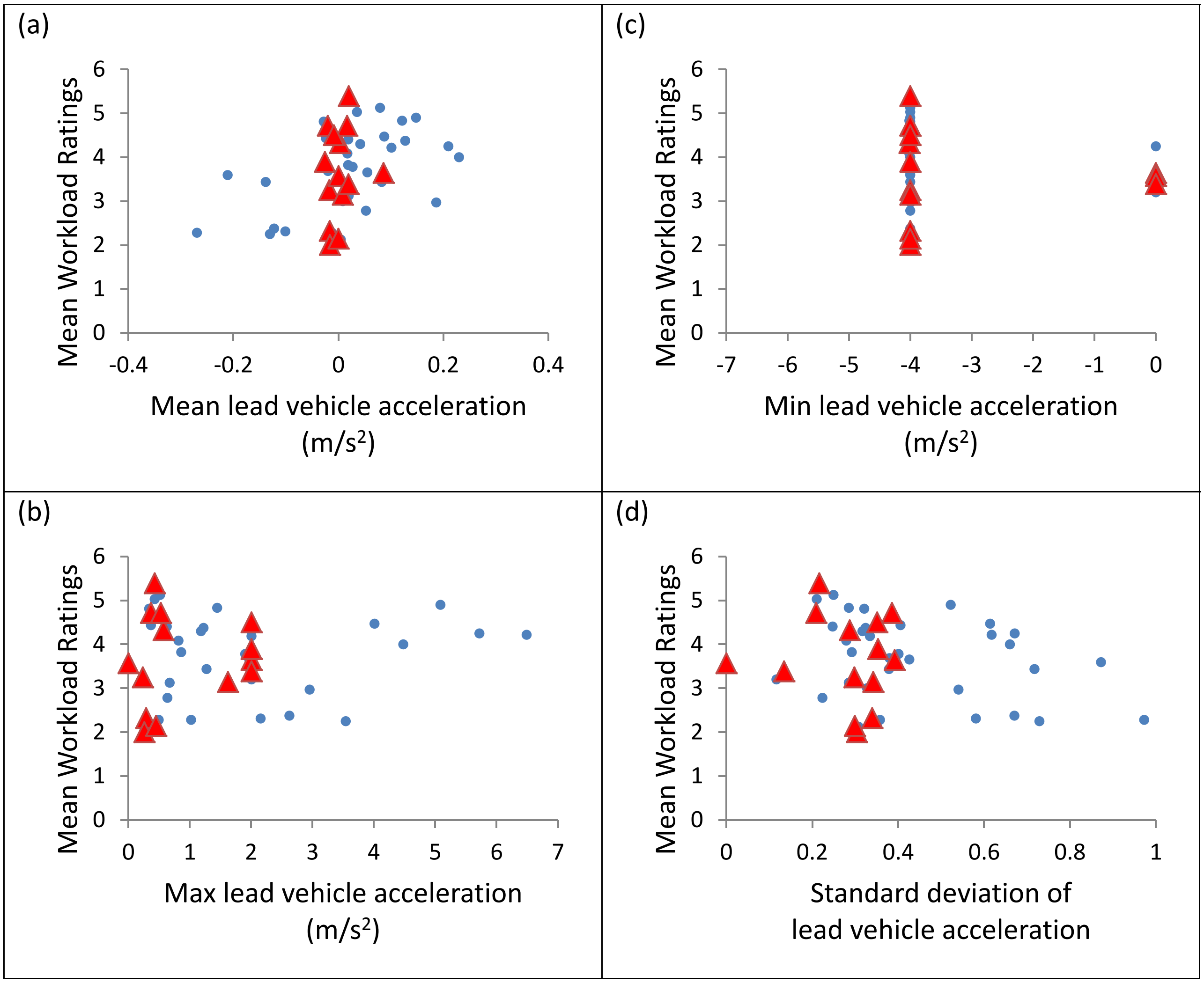


Figure 27. Mean Workload Ratings While Driving Versus (a) Mean, (b) Maximum, and (c) Minimum Lead Vehicle Acceleration and (d) Standard Deviation of The Acceleration

One could argue that workload should be greatest in the left lane because this lane moves most rapidly. One could also argue for greatest workload in the middle lane as in that lane there could be traffic on both sides of the vehicle. As shown in Figure 28, the workload was greatest for the middle lane. However, missing from this experiment were the speed differentials normally found on expressways in lighter traffic, where the left lane moves fastest (and has the greatest speed variability), and the right lane moves slowest. This, of course, ignores the effect of traffic entering and leaving at entrances and exits, which can profoundly affect workload.

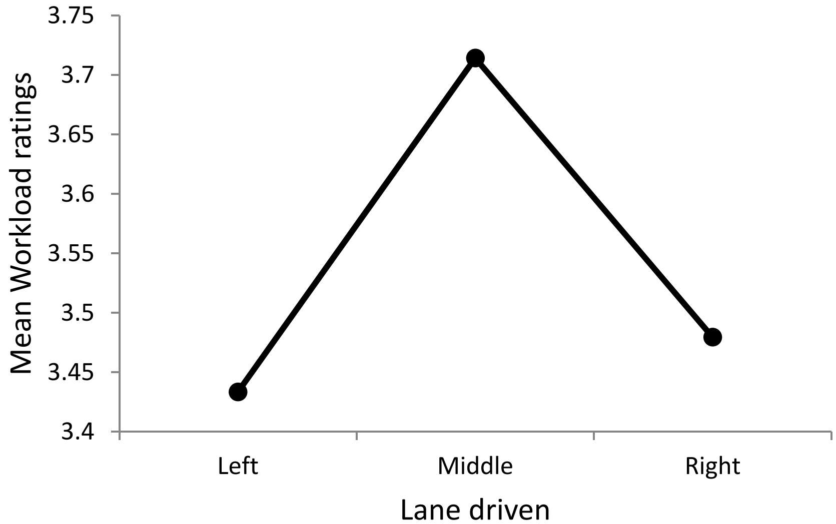


Figure 28. Mean Workload Ratings While Driving
Versus Lane Driven (36 scenarios × 16 subjects)

Besides the effect of lane driven, the level of service (LOS) will also be considered. However, the LOS of transition trials is not fixed, and only 18 scenarios with video clips have fixed LOS. Of the 18 scenarios, 4 were removed because the traffic is not stable (see Table 8). Therefore, in Table 13, the driving ratings are from 14 scenarios and 16 subjects, reduced driving ratings are from 14 scenarios and 14 subjects, and video clip ratings are from 18 scenarios and 16 subjects. Comparing the results to SAVE-IT project (see Table 7 in Test plan), ratings in this study are lower by about 1 unit.

Table 13. Mean Workload Ratings for Driving and Video Clips (This Experiment)
by Lane Driven and LOS

	Lane Driven	Left Lane			Middle Lane			Right Lane		
	LOS	A	C	E	A	C	E	A	C	E
Block 1	Driving	2.3	3.2	4.4	-	3.4	4.1	2.0	-	5.0
	Driving (Reduced)	2.3	3.3	4.7	-	3.6	4.3	2.0	-	5.4
	Watching Video Clips	2.3	3.5	5.6	1.9	3.5	4.4	2.0	2.8	5.8
Block 2	Driving	3.7	4.2	4.4	-	3.2	-	3.0	2.1	3.4
	Driving (Reduced)	3.9	4.5	4.7	-	3.4	-	3.1	2.1	3.6
	Watching Video Clips	3.8	4.8	5.2	2.8	3.9	4.3	2.5	3.1	4.7

In this experiment, subjects drove close to the posted speed 65 or 70 mph (28.9 or 31.1 m/s), depending upon the road segment. Again, ignoring traffic, higher speed should lead to greater workload. Scenario 29, appearing in the upper left corner of Figures 29a, b, and c, strongly influenced the correlations. In this scenario, subjects drove slower because the lead vehicle decelerated. For the full data set (16 subjects and 34 scenarios), the correlations of mean workload ratings for mean, maximum, minimum, and standard deviation of speed were -0.19, -0.17, -0.22, and 0.17. For the reduced data set (14 subjects × 14 scenarios), the correlations were, -0.02, 0.04, -0.07, and 0.13, essentially 0. Removing Scenario 29 (to be the 14 × 13 run), increased the correlations to 0.43, 0.40, 0.34, and 0.21 respectively. This suggests that the subject's speed did increase workload, but interactions with other factors, especially lead vehicle acceleration and gap, were extremely important.

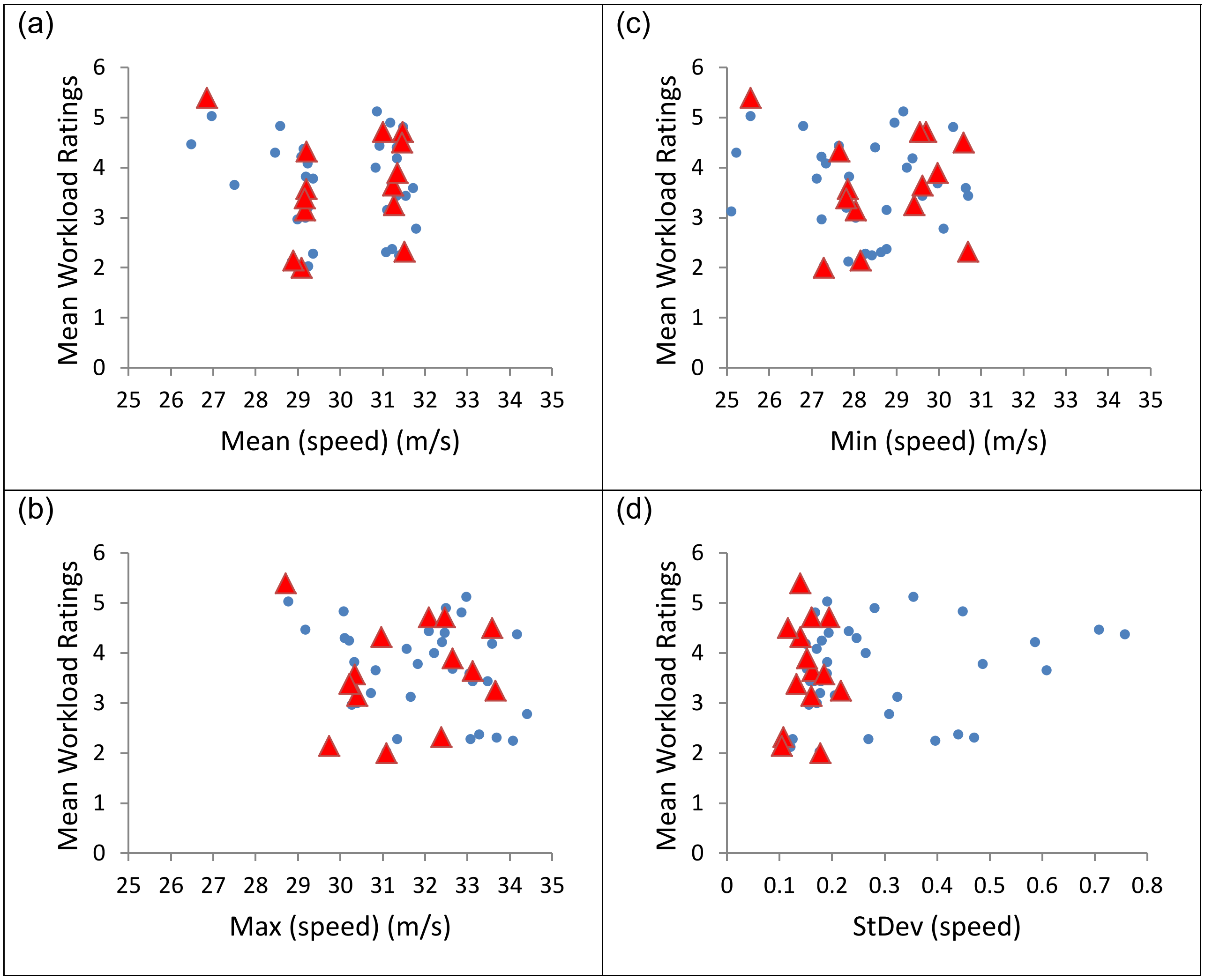


Figure 29. Mean Workload Ratings While Driving Versus (a) Mean, (b) Maximum, and (c) Minimum Subject Speed and (d) Its Standard Deviation

Logically, larger positive (true acceleration) and negative (deceleration, possibly in response to a slowing lead vehicle) values of longitudinal acceleration should be associated with greater workload. The correlations for the mean, maximum, minimum, and standard deviation of longitudinal acceleration were -0.13, 0.04, 0.33, and 0.18 for the 36 scenario – 16 subject data set and 0.21, 0.36, -0.07, and 0.43 for the reduced data set. In examining the panels in Figure 30, keep in mind that the ranges are not the same, presented in that manner here so the differences between data points would be apparent. Given the trend to favor the full data set as being most telling, there does not seem to be a strong relationship between workload and any statistic of lead vehicle acceleration, an outcome that was not expected.

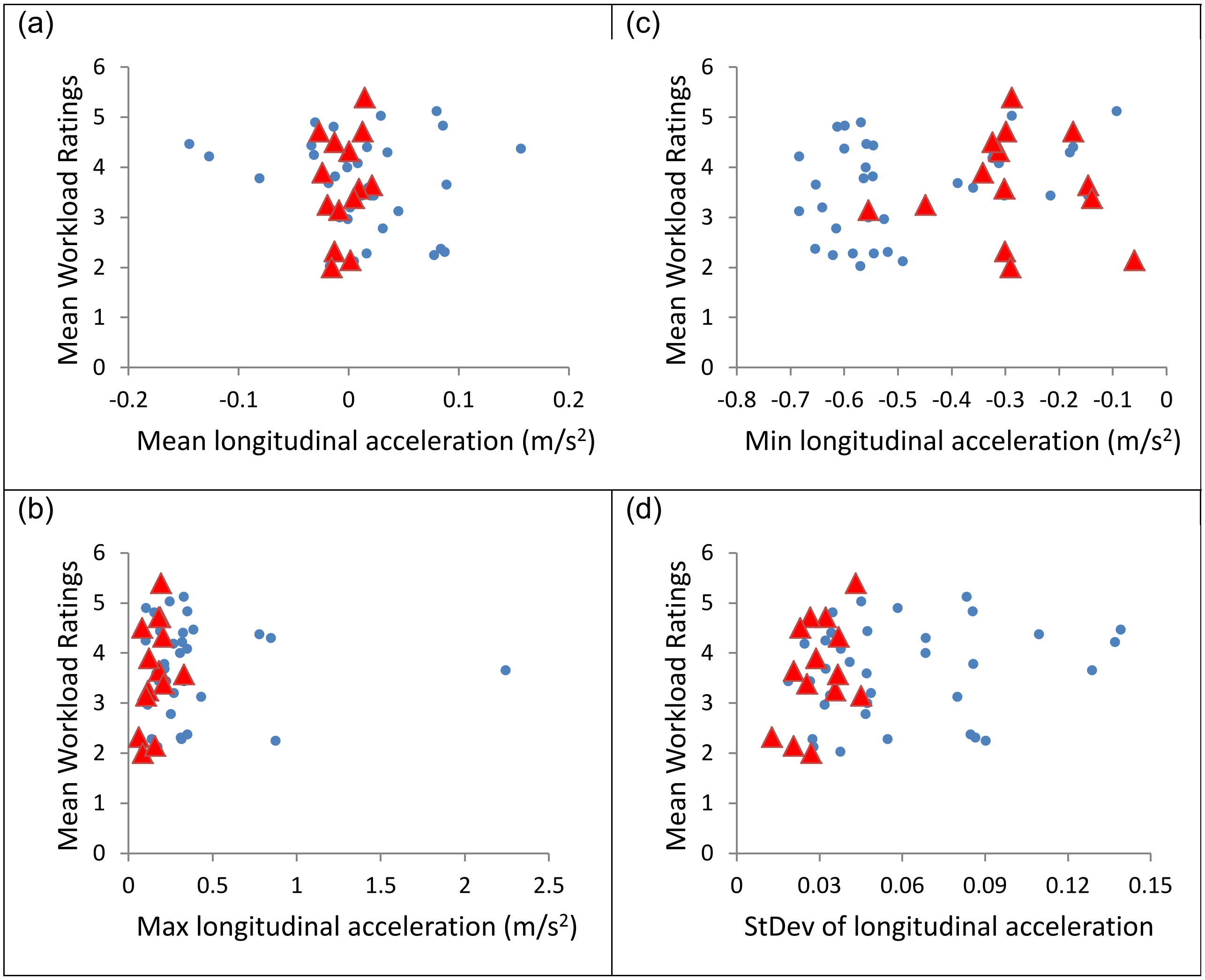


Figure 30. Mean Workload Ratings While Driving Versus (a) Mean, (b) Maximum, and (c) Minimum Subject Longitudinal Acceleration and (d) Its Standard Deviation

The gap is the distance between the front bumper of the subject vehicle and the rear bumper of the lead vehicle per SAE Recommended Practice J2944 (in progress). In the prior SAVE-IT project, this distance was referred to as range, was limited to 125 m, the range of the radar sensor. In some studies, this distance is incorrectly referred to as headway.

As a simplification, for all situations in which no vehicle was detected ahead, one was assumed to be present at 125 m. In fact, the added workload of a vehicle at that distance is quite small. In this research, the trials with the gap over 125 m were omitted, which means the gaps in the remaining data were not assumed values.

Workload should increase as the distance to the lead vehicle decreases. As was noted in Schweitzer and Green (2007), the relationship between workload and gap is probably not linear, and a log model has been proposed. In brief, if a vehicle is far away or very far away, the consequence is the same: its presence has little impact on driving, and the difference between the 2 situations is minor. As shown in Figure 31, the correlation between workload ratings and the mean, maximum, minimum, and standard deviation of gap were quite large, being -0.76, -0.67, -0.74, and -0.14 for 36 scenario-16 subject data set, and -0.83, -0.80, -0.81, and -0.03 for the reduced set. For $\log_{10}$(gap), the values were -0.73, -0.69, -0.73, and -0.06 and -0.83, -0.82, -0.83, and 0.02, respectively. Given that Kondoh, Yamamura, Kitazaki, Kuge, and Boer (2008) found TTC and inverse time headway to predict crash risk, the correlation of the workload ratings with inverse gap were examined. The correlations were 0.66, 0.61, 0.70, and -0.16 for the full data set and 0.78, 0.78, 0.78, and -0.28 for the reduced set. These correlations are comparable to the $\log_{10}$(gap) values.

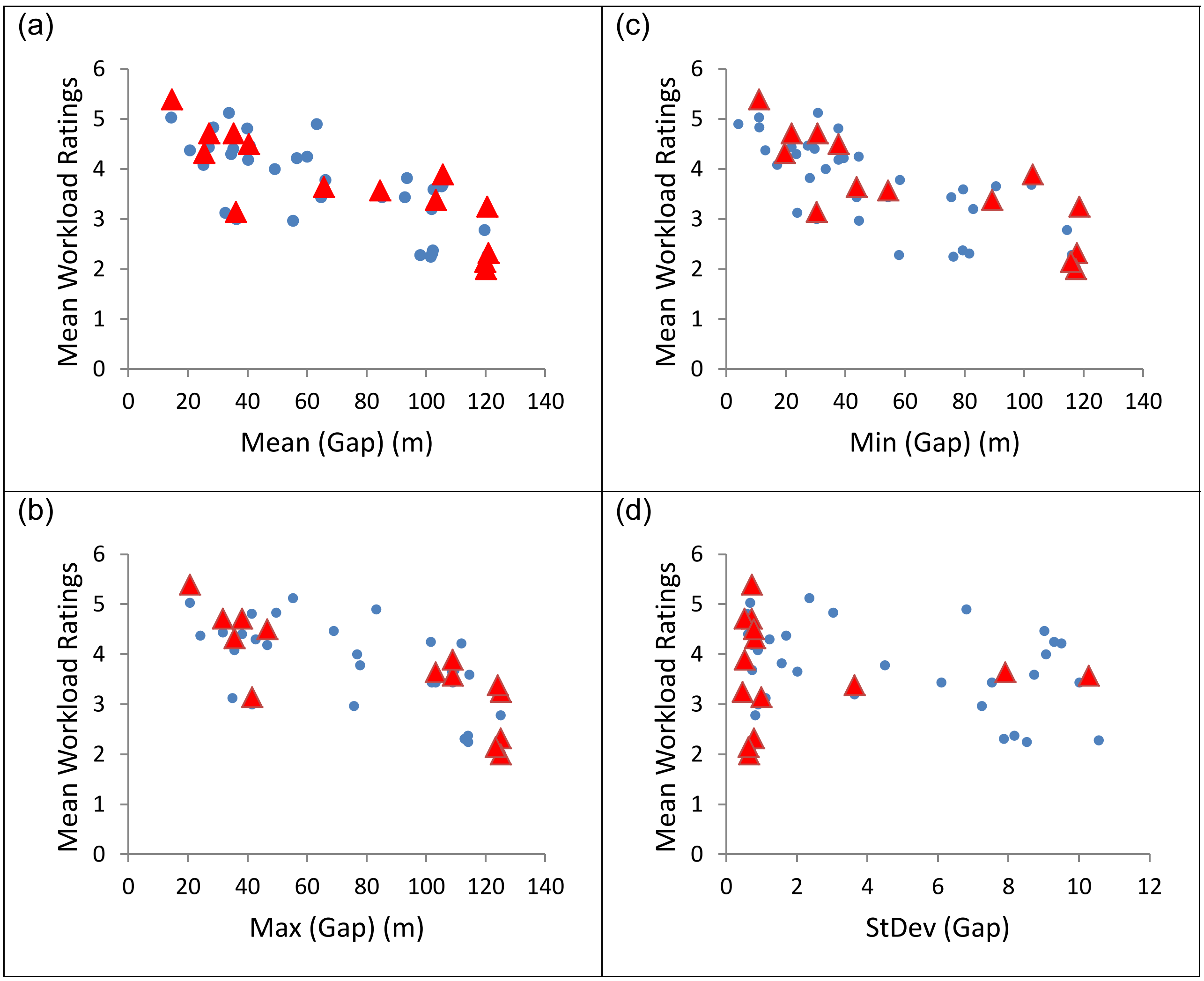


Figure 31. Mean Workload Ratings While Driving Versus (a) Mean, (b) Maximum, and (c) Minimum Gap Between Subject and the Lead Vehicle and (d) Its Standard Deviation

One could argue that a large standard deviation of lane position (lateral position) indicates high workload, as the subject is not able to keep the subject in a limited position. However, in this experiment, lateral control demands and statistics are unlikely to be linked to workload ratings; the road was straight, the subject never changed lanes, and there were no wind gusts, potholes, or any other significant lateral disturbances, making the lateral control effort consistently low. Therefore, it is no surprise that the correlation of mean, minimum, maximum, and the standard deviation of lateral position with workload ratings was 0.15, 0.11, 0.07, 0.32 for the 16 subjects × 36 scenarios data set, and 0.02, -0.01, -0.11, 0.11 for the 14 subjects × 14 scenarios data set (see Figure 32). Thus, again correlations are slightly greater for the larger sample, but the overall effect of the lateral position is primarily reflected in the standard deviation, and it is low.

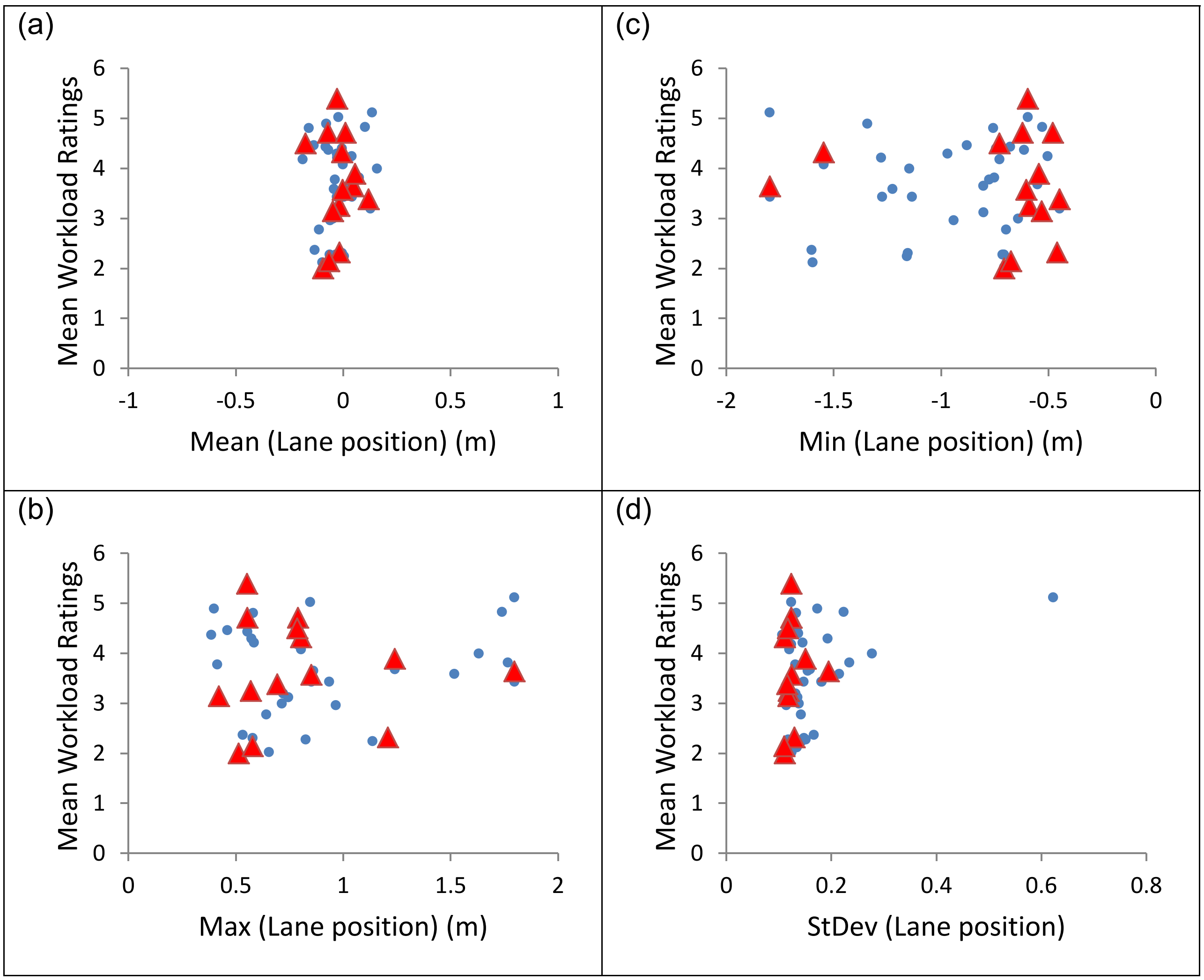


Figure 32. Mean Workload Ratings While Driving Versus (a) Mean, (b) Maximum, and (c) Minimum Lateral Lane Position of Subject and (d) Its Standard Deviation

Time-to-line crossing is a commonly cited lateral control measure. TLC can be calculated 3 ways, (1) as a function of distance and lateral velocity, (2) as a function of those 2 factors and lateral acceleration, and (3) trigonometrically, where road curvature is considered. In this instance, the distance-velocity method was used.

In determining the relationship between TLC and other factors, some thought is needed about how the TLC data are filtered. If the subject is driving stably, then lateral velocities are extremely small, in fact close to zero, so TLC values can be quite large, tens of thousands of seconds. In those situations, especially when means are computed, some filtering of the data may be needed. In cases of this study, the maximum values of subjects distributed from 20 s to 40 s, which make sense to normal driving. But for minimum TLC, all values are equal or close to 0, so this variable is not properly to be put into the prediction model.

Using that lateral velocity-distance method and filtering the data as described, the correlations of the mean, maximum, and standard deviation of TLC with all the data (16 subjects × 36 scenarios) and the reduced data set were -0.07 (-0.22), 0.18 (-0.13), and -0.33 (0.10) respectively. (No correlation for minimum TLC to workload was computed because all minimum TLC for each scenario were close to 0.) For the inverse TLC, the correlations were 0.17 (0.06), -0.27 (0.17), and -0.13 (0.23).

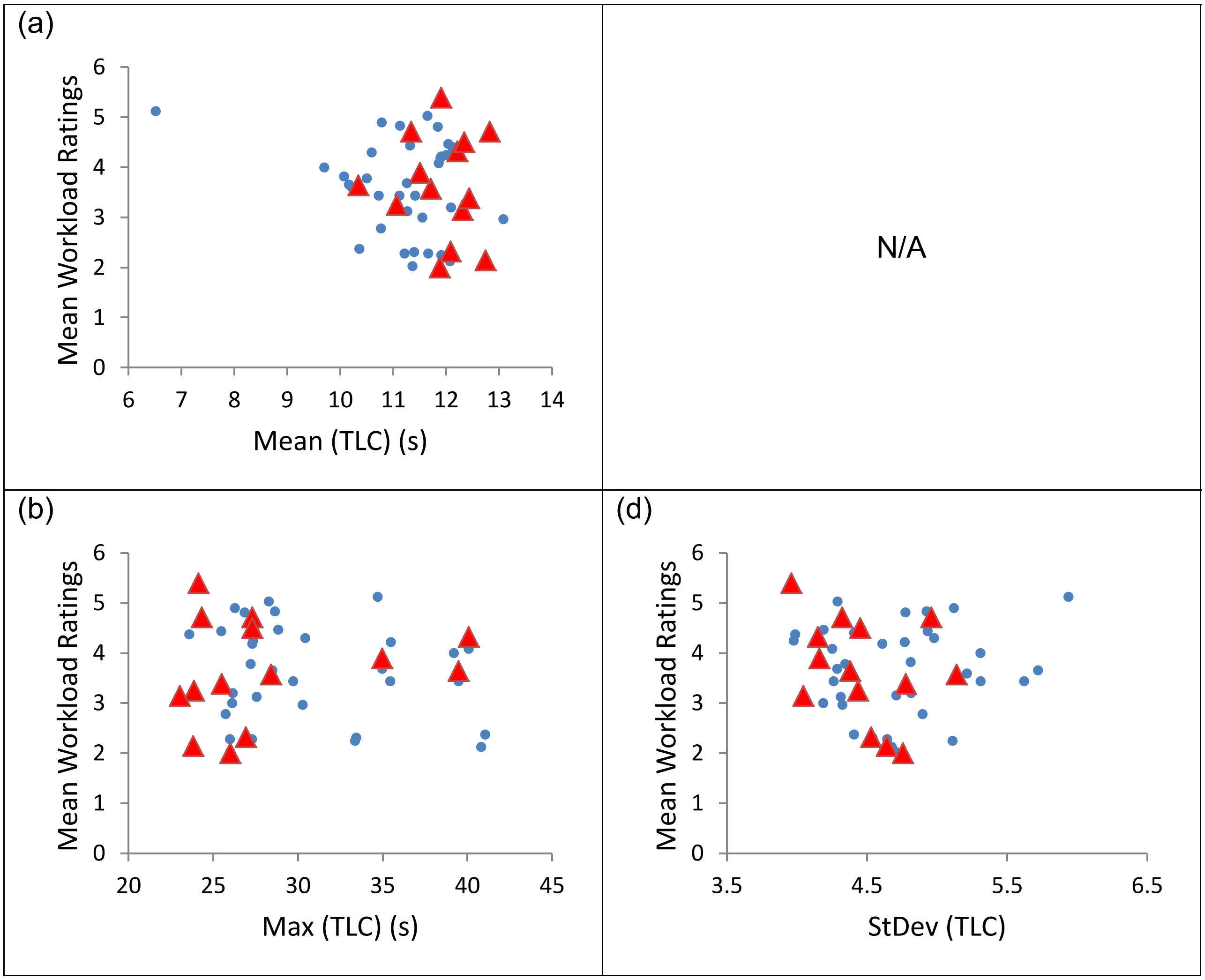


Figure 33. Mean Workload Ratings While Driving Versus (a) Mean, and
(b) Maximum Time-to-Line Crossing (TLC) and (d) Its Standard Deviation

Summary of Correlations

Table 14 shows all of the correlations of the mean workload ratings collected while driving with the associated mean driving performance statistic.

.

Table 14. Summary of Mean Workload Rating Correlations While Driving, r >0.40 in Bold

Category	Variables	Statistic	r	
			All Data	Reduced Set
Other Vehicles	Traffic Count	Mean	**0.65**	**0.65**
		Maximum	**0.65**	**0.53**
		Minimum	**0.55**	**0.52**
	Lead Speed	Mean	-0.37	-0.01
		Maximum	-0.34	0.06
		Minimum	-0.37	-0.04
		Standard deviation	0.01	-0.01
Subject Vehicle Longitudinal	Speed	Mean	-0.19	-0.02 (**0.43**)
		Maximum	-0.17	0.04 (**0.40**)
		Minimum	-0.22	-0.07 (0.34)
		Standard deviation	0.17	0.14 (0.21)
	Longitudinal Acceleration	Mean	-0.13	0.21
		Maximum	0.04	0.36
		Minimum	0.33	0.07
		Standard deviation	0.18	**0.43**
Subject Vehicle Lateral	Lane Position	Mean	0.15	0.02
		Maximum	0.11	-0.01
		Minimum	0.07	-0.11
		Standard deviation	0.32	0.11
	TLC	Mean	-0.07	-0.22
		Maximum	0.18	-0.13
		Minimum	-0.33	0.10
Longitudinal Relationship to Other Vehicles	Gap	Mean	**-0.76**	**-0.83**
		Maximum	**-0.67**	**-0.80**
		Minimum	**-0.74**	**-0.81**
		Standard deviation	-0.14	-0.03
	log10 (gap)	Mean	**-0.73**	**-0.83**
		Maximum	**-0.69**	**-0.82**
		Minimum	**-0.73**	**-0.83**
		Standard deviation	-0.06	0.02
	Inverse Gap	Mean	**0.66**	**0.78**
		Maximum	**0.61**	**0.78**
		Minimum	**0.70**	**0.78**
		Standard deviation	-0.16	-0.28

Overall, differences between the correlations for the 2 data sets were small, with correlations for the reduced data set generally greater except for traffic effects. The best predictor was gap or some gap-related statistic (log10 (gap), inverse gap), followed by traffic count. For the reduced data set, the standard deviation of lateral acceleration, mean subject vehicle speed, and maximum subject vehicle speed were all weakly correlated with workload as well.

Which Equations Estimate Workload Ratings of Clips?

Table 15 contains the original equations from SAVE-IT (also shown in Table 9) as well a new equations using the same variables but using the data from this experiment. As the 2 sets of equations fit the same scenarios, some similarities are expected. In fact, the signs and magnitudes of the new equations are close to the SAVE-IT equations, though the R2 values are slightly reduced. Given there were only 16 subjects and not 24 (which leads to more stable means), and there are concerns about ratings for 2 subjects, this outcome seems reasonable.

Table 15. Comparisons of Equations from SAVE-IT and
This Experiment with the Same Factors

# Factors	Original SAVE-IT equation (best fit)	Original SAVE-IT parameters & new subjective ratings for reduced scenarios & subjects
2	8.86 -3.00*LogMeanGap +0.47*MeanTrafficCount R^2=0.82	7.90 -2.52*LogMeanGap +0.06*MeanTrafficCount R^2=0.69
3	8.87 -3.01*LogMeanGap +0.48*MeanTrafficCount +2.05*MeanLongitudinalAccel. R^2=0.87	7.90 -2.51*LogMeanGap +0.06*MeanTrafficCount +0.51*MeanLongitudinalAccel. R^2=0.69
4	8.07 -2.72*LogMeanGap +0.48*MeanTrafficCount +2.17*MeanLongitudinalAccel/ -0.34*MinimumLeadVehicleAccel. R^2=0.85	8.57 -2.72*LogMeanGap +0.13*MeanTrafficCount -14.28*MeanLongitudinalAccel. +0.20*MinimumLeadVehicleAccel. R^2=0.74

To further examine the new reduced set of workload ratings, 2 sets of equations, one equation from stepwise regression, the other equation from force fitting variables were developed (Table 16). In the stepwise analysis, the independent variables considered were logarithmic mean gap, mean traffic count, mean longitudinal acceleration of subject's vehicle, and minimum lead vehicle acceleration. Notice the R^2 value of the 2 equations is identical, accounting for almost 70% of the variance, quite high, with the mean gap being the key variable.

Table 16. Equations Based on Workload Ratings from This Experiment

Method	New equations	Comment
Stepwise	5.13 -0.02*MeanGap R^2=0.69	only add variables whose entry was significant at $p<0.05$
Forced entry	7.80 -2.66*LogMeanGap +0.05*MeanTrafficCount -4.17*StDevLongitudinalAcceleration +0.11*StDevTLC R^2=0.69	include all variables that had the highest correlations with workload

The differences between the new and old equations in predicting the workload ratings while driving are shown in Figures 34 and 35. For the new equations, regardless if the stepwise or forced entry equation is used, the predictions are remarkably good, with only 3 of the 14 residuals exceeding 0.5, and ratings were given to the nearest 0.5. However, using the SAVE-IT equation, the residuals were quite large. Thus, although the results from the 2 studies were highly correlated, the size absolute differences deserve attention.

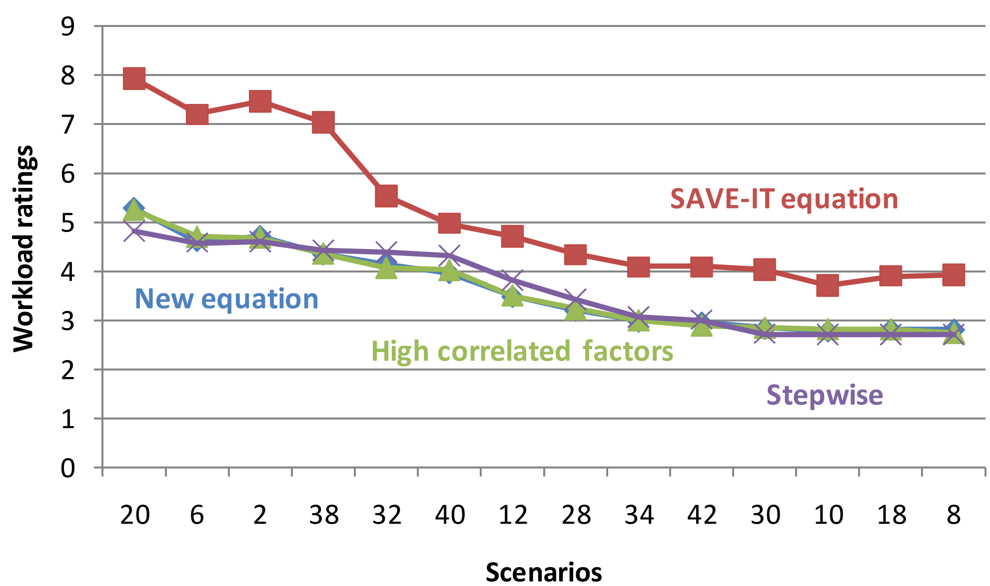


Figure 34. Prediction of Workload Ratings While Driving Using the 2-Factor SAVE-IT Equation and 2 New Equations (High p to Enter and Stepwise)

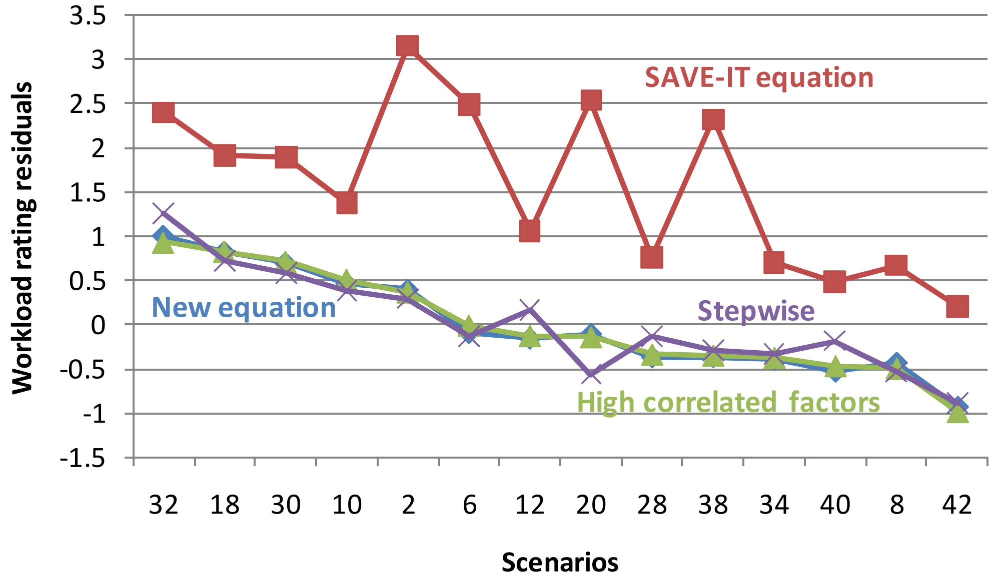


Figure 35. Residual Plot of Old 2-variable and New Equations (Predicting Workload Ratings while Driving)

How and why the new equations and those based on the SAVE-IT data differ is worth some thought. One of the key differences between the 2 experiments is how traffic was represented. In the SAVE-IT video clips, subjects were presented with a narrow field of view, and traffic close to them and to the side was not visible. Thus, the count of vehicles only considered those visible in the scene, not the total number likely to be present. In the current experiment in the simulator, the field of view was larger and more vehicles were visible, even though the density was as before. As subjects drove the simulation before watching clips in this experiment, that driving exposure may have influenced their estimates of traffic. For the SAVE-IT video clips, the number of potentially visible vehicles can be greater than 10, but the traffic counts in driving scenarios are at most 6 (a simulator software limitation), which does not lead to the workload as high as in SAVE-IT.

Comparing the previous and new equations in Table 16 with 3 and 4 equations (Figures 36-39) shows similar patterns to those with 2 variables—in general a decreased emphasis on the traffic count, as well as small residuals for the new equations but large equations for the prior SAVE-IT equation.

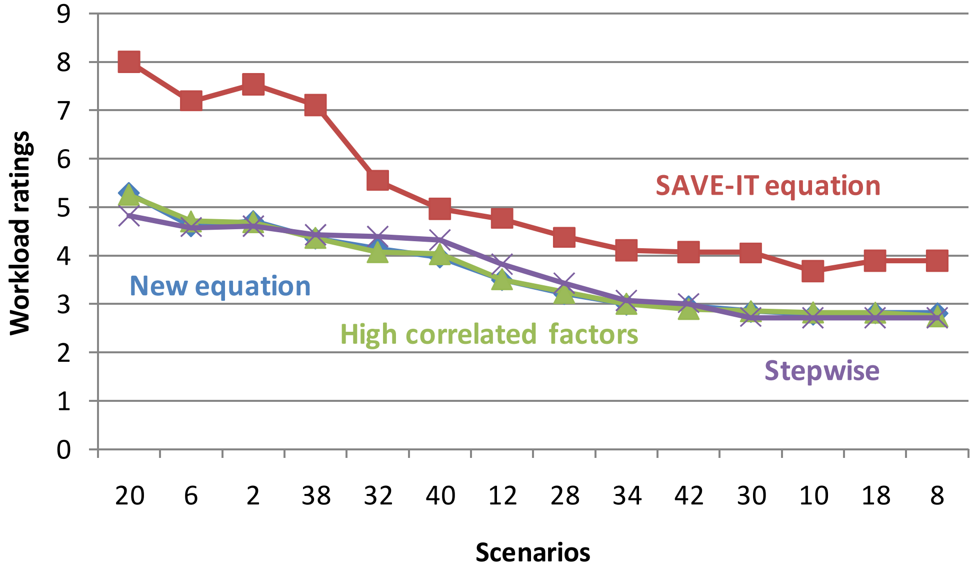


Figure 36. Prediction of Workload Ratings While Driving Using the 3-Factor SAVE-IT Equation and 2 New Equations (High p to Enter and Stepwise)

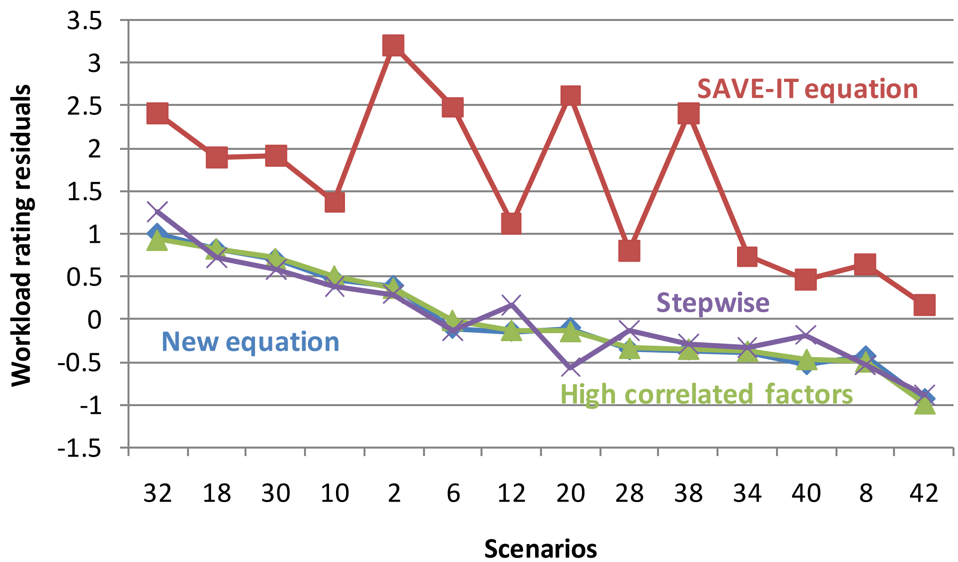


Figure 37. Residual Plot of Old 3-variable and New Equations (Predicting Workload Ratings while Driving)

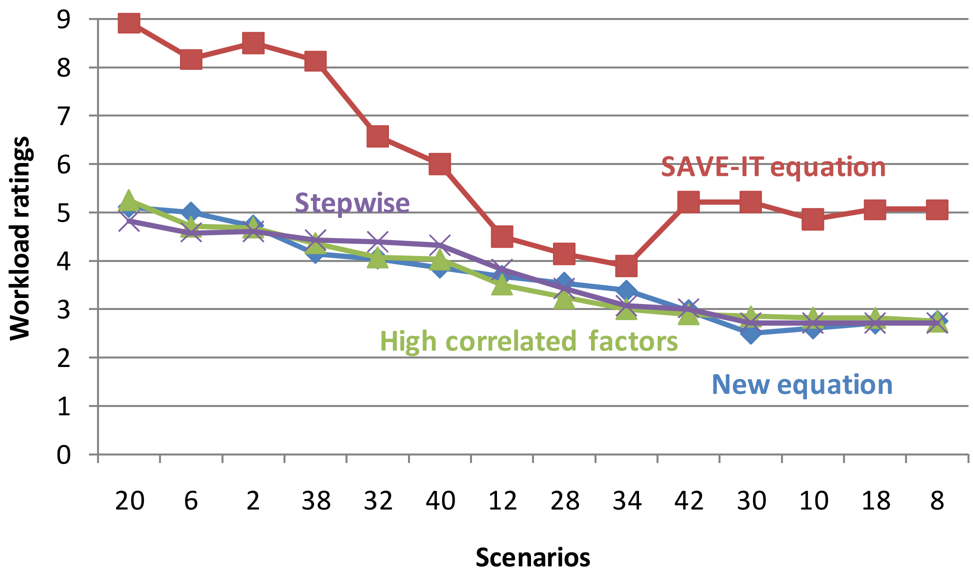


Figure 38. Prediction of Workload Ratings While Driving Using the 4-Factor SAVE-IT Equation and 2 New Equations (High p to Enter and Stepwise)

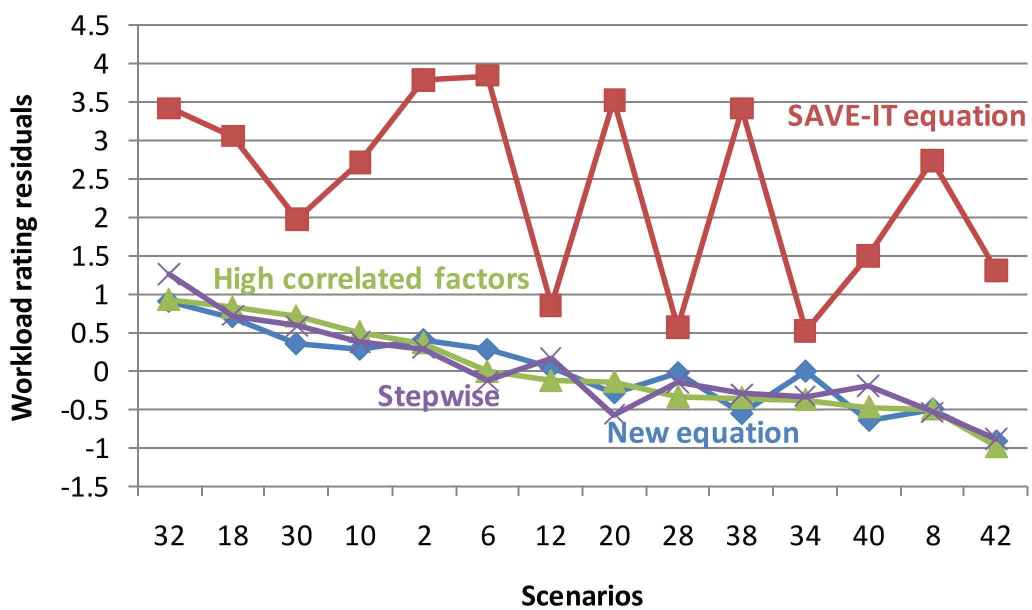


Figure 39. Residual Plot of Old 4-variable and New Equations (Predicting Workload Ratings while Driving)

CONCLUSIONS

1. How are the workload ratings distributed?

Across the range of the workload rating scale, the ratings tended to be clustered at the lower ends of the range. It may be that a few subjects, probably 2 of the 16, to some degree, did not understand the instructions, thinking that workload was low because they were not engaged in driving, even though they were to rate the workload as observed relative to the video clips. This problem was much more common in older subjects. Some modifications of the instructions could help overcome this problem.

However, the most important observation was that subjects significantly favored integer ratings over nonintegers (nearest 0.5). This could be due to limits in how precisely subjects could estimate workload or possibly some lack of emphasis in the instructions. If the ratings are imprecise, the predicting them precisely will be difficult. Here again, in future applications of this rating method, changes to the instructions will be considered. Further, when ratings are collected, subjects will be asked if they use integers because they cannot estimate workload more precisely.

2. How consistent are the workload ratings within subjects? Specifically, if a subject drives the same scenario twice, how similar are the 2 workload ratings?

One scenario was driven 3 times, with resulting correlations of 0.91 (trials 2 and 14), 0.80 (trials 14 and 24) and 0.77 (trials 2 and 24). Keep in mind that these correlations are based on the individual ratings of subjects for each trial, not means across subjects. From that perspective, they are quite good.

3. How consistent are the video clip workload ratings across groups of subjects? Specifically, how well do ratings of workload of the video clips from a new group of subjects correlate with ratings from subjects in the SAVE-IT project?

For the video clip rating procedure to be useful, the ratings need to be stable. The correlation of the mean clips ratings averaged across subjects from this experiment with the same mean values from the prior experiment was 0.97, extremely high, especially considering the concerns about several subjects underrating workload in this experiment. There was, however, a consistent trend for workload ratings from this experiment to be less than those found in SAVE-IT (about 80% of the prior ratings, on average). The underrating problems should be resolved by changes to the instructions.

4. Are the workload ratings of various scenarios shown on video clips different from ratings obtained while driving the same scenarios in a simulator? If they differ, by how much?

Are the workload estimates from watching driving and actually driving different? When can video clips be used?

The correlation of the video clip ratings from this experiment with the workload ratings of driving those same scenes was 0.92. Further, the correlation of the driving workload ratings in this experiment with ratings of clips of those scenes in the prior SAVE-IT experiment was 0.90, quite good. There were no indications of any systematic differences in the ratings from the 2 sets of data.

5. How well do the workload equations developed from passive viewing of road scenes in the SAVE-IT experiment predict the workload of driving those scenes in a simulator?

The equations developed in this experiment predicted almost 70% of the variance of the ratings of workload while driving the scenes approximated by the video clips. This is quite good.

6. What equations, based on factors known to be important based on the literature (inverse time to collision, inverse time gap or log gap, lead vehicle acceleration, etc.) best predict the new workload ratings?

Factors that were highly correlated with the workload ratings when driving in this experiment included mean traffic count (r=0.65), the mean gap (r=-0.83), the log10(gap) (r=-0.83), the mean inverse gap (r=0.78). There was also a smaller correlation with the standard deviation of longitudinal acceleration (r=0.43). Overall, the means were consistently better predictors than the maximum or minimum values due to quantization of the maximum and minimum values.

The equations found to predict the workload ratings of clips in this experiment were consistent with those in SAVE-IT. Table 17, shows the original SAVE-IT equations and the equations developed using the new workload ratings of the same clips.

Table 17. Old and New Clip Workload Rating Equations

# factors	Original SAVE-IT equation (best fit) Workload =	SAVE-IT independent variables & new subjective ratings for reduced scenarios & subjects
2	8.86 -3.00*LogMeanGap +0.47*MeanTrafficCount R^2=0.82	7.90 -2.52*LogMeanGap +0.06*MeanTrafficCount R^2=0.69
3	8.87 -3.01*LogMeanGap +0.48*MeanTrafficCount +2.05*MeanLongitudinalAcceleration R^2=0.87	7.90 -2.51*LogMeanGap +0.06*MeanTrafficCount +0.51*MeanLongitudinalAcceleration R^2=0.69
4	8.07 -2.72*LogMeanGap +0.48*MeanTrafficCount +2.17*MeanLongitudinalAcceleration -0.34*MinLeadVehicleAcceleration R^2=0.85	8.57 -2.72*LogMeanGap +0.13*MeanTrafficCount -14.28*MeanLongitudinalAcceleration +0.20*MinLeadVehicleAcceleration R^2=0.74

Notice that especially in the 2-factor equation, the relative magnitude of the LogMeanGap is about the same, but there are some differences in the effect of traffic. This is probably an experimental artifact. In the SAVE-IT clips, a roughly 60-degree field of view was visible. However, the radar that counted vehicles only had a 15–degree field of view, thus undercounting the number of vehicles in the scene.

Furthermore, in the driving simulator, the field of view was 200 degrees. In some situations, there were vehicles close to the driver but in adjacent lanes. These vehicles, especially if they are next to the subject, add considerably to workload. However, they would not be visible in any SAVE-IT clips or detected by the SAVE-IT radar. New anchor clips with a wider field of view are needed.

Another potential mismatch is that in the SAVE-IT experiments, subjects were comparing clips recorded at 1 Hz (but updating at 2 Hz) with other clips with identical recording and update rates. These rates were due to limitations as to how much data vehicles used in the ACAS field test could record.

In addition to fitting the SAVE-IT equations, workload predictions were developed using stepwise and forced entry regression methods. As shown in Table 18, forcing entry did not increase the R2 value, with both equations accounting for 69% of the variance. This is quite good, but not as good as the SAVE-IT results, primarily for the reasons just discussed. Improving the anchor clips, how traffic is counted, and the instructions should increase the R2 to the prior value.

Table 18. Workload Equations for the Driving Data

Method	Equation (Workload =)	Comment
Stepwise	5.13 -0.02*MeanGap R^2=0.69	only add variables whose entry was significant at $p<0.05$
Forced entry	7.80 -2.66*LogMeanGap +0.05*MeanTrafficCount -4.17*StDevLongitudinalAcceleration +0.11*StDevTLC R^2=0.69	include all variables that had the highest correlations with workload

Thus, these equations emphasize the importance of the lead vehicle and suggest that a rough approximation of the driving workload can be computed using a first order linear equation involving the distance to the lead vehicle.

7. What are the differences in the above between young and older drivers?

There were no indications of any systematic differences due to age or gender. This makes sense as the workload ratings were relative to anchors representing particular driving situations.

Closing Comments

One of the major weaknesses of the driving literature in general is that the workload of the primary driving task is generally not described, and where it is, the description is qualitative (e.g., moderate workload). But moderate workload while driving in Michigan’s rural upper peninsula, for example, often means encountering little traffic. That is different from driving in Ann Arbor (a small city) and certainly different from driving in Tokyo, where moderate workload is any steady movement. Without such quantification, studies in different locations or from different contexts (e.g., simulators versus public roads) cannot be compared.

There is considerable attention being given to the distraction problem. However, setting standards for how much workload can be added is difficult if one cannot quantify in addition to what.

This research is a step toward solving that problem. Building on prior research that involved rating the workload of driving scene relative to anchor clips, this experiment has subjects drive those scenes and rate the workload of each scenario. In this case, the workload ratings of clips were highly consistent with those from the prior experiment, and the workload ratings while driving could be reliably predicted from a few simple measures, the most important of which was mean gap.

As a first step, authors are recommended to report the workload from their studies using the equation that considers the mean gap.

Depending on the availability of funding, research will continue in parallel to improve the workload anchor clips and develop and validate equations for driving situations other than expressways.

RERERENCES

AASHTO (2004), A Policy on Geometric Design of Highways and Streets (5th ed.), American Association of State Highway and Transportation Officials, Washington, DC.

Alm, H. & Nilsson, L. (1994). Changes in Driver Behaviour as a Function of Handsfree Mobile Phones—a Simulator Study. Accident Analysis and Prevention, 26, 441-445.

De Waard, D. (1996). The Measurement of Driver's Mental Workload (Ph.D. thesis), Groningen, The Netherlands: Centre for Environmental and Traffic Psychology, University of Groningen (http://home.zonnet.nl/waard2/mwl.htm)

Eby, D.W., Shope, J.T., Molnar, L.J., Vivoda, J.M., Fordyce, T.A. (2000). Improvement of Older Driver Safety Through Self Evaluation: The Development of a Self-Evaluation Instrument, (technical report UMTRI-2000-04), Ann Arbor, Michigan: University of Michigan Transportation Research Institute.

Ervin, R., Sayer, J., LeBlanc, D., Bogard, S., Mefford, M., Hagan, M., Bareket, Z., and Winkler, C. (2005). Automotive Collision Avoidance System (ACAS) Field Operational Test Methodology and Results (technical Report DOT HS 809 901), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation.

Fastenmeier, W. and Gstalter, H. (2007). Driving Task Analysis as a Tool in Traffic Safety Research and Practice, Safety Science, 45, 952-979.

Gawron, V. (2000). Human Performance Measures Handbook, Mahwah, NJ: Lawrence Erlbaum Associates.

Green, P. (2008). Driver Interface/HMI Standards to Minimize Driver Distraction/Overload (SAE paper 2008-21-2002), Convergence 2008 Conference Proceedings, Detroit, Michigan.

Green, P., Diebol, J.K., Park, J.S., and Ho, H. (2011). Visual Occlusion and the Workload/Visual Demand of the Primary Task of Driving: A Review of the Literature (technical report UMTRI-2010-11), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Hart, S.G. and Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: Hancock, P.A. and Meshkati, N., (eds.), Human Mental Workload, North-Holland, Amsterdam, 139–183.

Hart, S. (2006). NASA-Task Load Index (NASA-TLX); 20 Years Later. Human Factors and Ergonomics Society Annual Meeting Proceedings, 50, 904-908.

Hoedemaeker, M. and Brookhuis, K.A. (1998). Behavioural Adaptation to Driving with Adaptive Cruise Control (ACC), Transportation Research, Part F, 1(2), 95-106.

Hulse, M. C., Dingus, T. A., Fischer, T., and Wierwille, W.W. (1989). The Influence of Roadway Parameters on Driver Perception of Attentional Demand, Advances in Industrial Ergonomics and Safety I, New York, NY: Taylor and Francis, 451-456.

Kondoh, T., Yamamura, T., Kitazaki, S., Kuge, N., and. Boer, E.R. (2007). Identification of Visual Cues and Quantification of Driver's Perception of Proximity Risk to the Lead Vehicle in Car-Following Situations, Journal of Mechanical Systems for Transportation and Logistics, 1(2), 170-180.

Molnar, L.J., and Eby, D.W. 2009. Getting Around: Meeting the Boomers' Mobility Needs, In Houston, R.B. (ed.), Boomer Bust? Economic and Political Issues of the Graying Society, vol. 2, Santa Barbara, CA: Praeger Publishing.

Nygren, T.E. (1991). Psychometric Properties of Subjective Workload Measurement Techniques: Implications for Their Use in the Assessment of Perceived Mental Workload. Human Factors, 33, 17-33.

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.

Pauzié, A. (2008). Evaluating Driver Mental Workload Using the Driving Activity Load Index (DALI), Humanist Conference Proceedings, 67-77 (http://www.conference.noehumanist.org/articles/Proceedings-HUMANIST-S2.2.pdf, retrieved June 2, 2011).

Pauzié A. & Pachiaudi G. (1997). Subjective Evaluation of the Mental Workload in the Driving Context, in Rothengatter, T. & Carbonell Vaya, E. (eds.), Traffic & Transport Psychology: Theory and Application, 173-182, Oxford, UK: Pergamon Press.

Piechula, W., Mayser, C., Gehrke, H., and König, W. (2003). Reducing Driver's Mental Workload by Means of an Adaptive Man-Machine Interface, Transportation Research: Part F, 6(4), 233-248.

Rakauskas, M.E., Gugerty, L.J., and Ward, N.J. (2004). Effects of Naturalistic Cell Phone Conversations on Driving Performance, Journal of Safety Research, 35(4), 453-464.

Reid, G.B. and Nygren, T. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload, 185-215, in Hancock, P.A., Meshketi, N (eds.), Human Mental Workload, Advances in Psychology, Oxford, England: North-Holland.

Reid, G. B., Potter, S. S., and Bressler, J. R. (1989). Subjective Workload Assessment Technique (SWAT): A User's Guide (Technical Report AAMRL-TR-89-023). Wright-Patterson Air Force Base, Ohio: Armstrong Aerospace Medical Research Laboratory

Reid, G. B., Shingledecker, C. A., Nygren, T. E., and Eggemeier, F. T. (1981). Development of Multidimensional Subjective Measures of Workload, Proceedings of the 1981 IEEE International Conference on Cybernetics and Society, 403-406

Schweitzer, J. and Green, P.A. (2007). Task Acceptability and Workload of Driving Urban Roads, Highways, and Expressway: Ratings from Video Clips (technical Report UMTRI-2006-6), Ann Arbor, MI: University of Michigan Transportation Research Institute. (May).

Schweitzer, J. and Green, P. (2009). Instructions for Using the UMTRI Expressway Scenario Creation and Control Software (technical Report UMTRI-2009-19), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Selcon, S. J., Taylor, R. M., and Koritsas, E. (1991). Workload or Situational Awareness?: TLX vs SART for Aerospace Systems Design Evaluation, Proceedings of the Human Factors Society 35th Annual Meeting, Santa Monica, CA: Human Factors Society, 62-66.

U. S. Department of Transportation (2009). Highway Statistics 2008, U.S. Department of Transportation, Federal Highway Administration, Washington, DC. Retrieved October 16, 2010 from http://www.fhwa.dot.gov/policyinformation/statistics/2008/.

U. S. Department of Transportation (2010). Distracted Driving 2009 (Traffic Safety Facts Research Note, DOT HS 811 379), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation, http://distraction.gov/research/PDF-Files/Distracted-Driving-2009.pdf).

Van Erp, J.B.F. and Van Veen, H.A.H.C. (2004). Vibrotactile In-Vehicle Navigation System, Transportation Research, Part F, 7, 247-256.

Wada, T., Doi, S., Tsuru, N., Isaji, J., and Kaneko, H. (2010). Characterization of Expert Drivers Last-Second Braking and Its Application to a Collison Avoidance System, IEEE Transactions on Intelligent Transportation Systems, 11(2), 413-422.

Zijlstra, F.R.H. (1993). Efficiency in Work Behavior. A Design Approach for Modern Tools (PhD thesis), Delft University of Technology. Delft, The Netherlands: Delft University Press.

APPENDIX A – SUBJECT RECRUITMENT MATERIALS

Recruiting advertisement

The University of Michigan Transportation Research Institute is conducting an experiment to help reduce accidents related to distracted driving. We are collecting the workload, or the demand of driving, of various road conditions. The experiment is conducted at UMTRI in a driving simulator.
We are looking for licensed drivers, ages 18-30, and over 65. The experiment lasts 1.5 to 2 hours and pays $45.
Call Katherine at (734) 763-6081, Monday through Friday, 9am – 5pm.

Recruiting Script

The University of Michigan Transportation Research Institute is conducting an experiment to help reduce problems related to distracted driving—use of cell phones, complex navigation systems, and so forth. To determine how much is too much for drivers to do, we first need to know what is the demand, the workload, of just driving a vehicle.

To determine that, we will have people, maybe you, drive our simulator in various traffic situations and rate the workload of driving using a method we developed. The experiment is quite straightforward, and the only concern is that some drivers can get motion sickness. However, either you or we will stop the experiment before that occurs.

Also, you should know that we will videotape the experiment and will show outtakes to the sponsor and the public.

The experiment pays $45 for 1.5 to 2 hours and takes place at UMTRI. We are looking for men and women, ages 18 – 30, and over 65. You must be a U.S. licensed driver. If you wear glasses when you drive, please bring them.

Please call Katherine at (734) 763 – 6081, Monday through Friday, June 8th – June 11th, 2010, 9am – 5pm, for further information or to schedule a time, or email her at kthstone@umich.edu.

APPENDIX B – CONSENT FORM

UMTRI University of Michigan Transportation Research Institute 2901 Baxter Road, Ann Arbor, MI 48109-2150	Participant # _________

Consent Form

Development of a Protocol to Assess the Workload of Driving
Investigator: Paul Green (763 3795) UMTRI Driver Interface Group

To determine how much is too much for drivers to do while driving (and when it might be unsafe), we need to measure the workload (the demand) of driving the vehicle and the added workload of other activities such as using a phone or a navigation system. In this first step, we will quantify the workload of the just driving the vehicle.

After providing biographical data (your age, driving experience, etc.) and driving data (e.g., miles drive/year), you will practice driving the simulator while rating the workload of driving on a scale we created. You cannot actually crash in the simulation because the car is invincible. Next, while being videotaped, you will drive on a simulated 3-lane expressway with varying amounts of traffic following the directions of the experimenter, again rating the workload relative to some video clips. Finally, at the end, there will be a brief questionnaire.

This is an evaluation of the workload of driving, not your skill or ability to drive. Participation in this research is 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 $20 for your time. Of course, there are no costs to you since UMTRI parking is free. The study should take about 1 hour.

Summaries of what you did (but not your name) will appear in a publicly available report whose results will make future vehicles that you may drive less distracting and safer. Records will be kept confidential to the extent provided by federal, state, and local law, though various officials can inspect them.

At any time, should you have questions regarding your rights as a participant in research, please contact the Institutional Review Board, Behavioral Sciences, 540 E. Liberty # 202, Ann Arbor, MI 48104, (734) 936-0933, email: irbhsbs@umich.edu.

> **As was stated when you were scheduled for this experiment, all participants must be "videotaped". I therefore agree to be recorded and realize my face will appear on the recording. I understand that segments from the recordings of my sessions may be used in presentations by the authors, by the sponsor, and by the media (e.g., on TV) to help explain this research. My full name will not be disclosed with the recording. The raw recordings will be discarded 10 years after the project is completed.**

I have read and understand the information presented above, and all of my questions have been answered. My participation is voluntary. I agree to participate.

______________________________ Print your name	____________________ Date
______________________________ Sign your name	____________________ Witness (experimenter)

Note: Keep one copy for the records and give the other to the participant.

APPENDIX C – BIOGRAPHICAL FORM

Workload of Driving– Biographical Form

Personal Details

Name ______________________________

Phone: ____________

Email address ____________________

May we email you for future studies? yes no

Born (month / day / yr) __ / __ / __

Occupation: _______ (if retired: main 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

What motor vehicle do you drive most often?

Year: _______ Make: _______ Model: _______

How many miles do you drive per year? _____

What lane of a 3-lane highway do you normally drive in?

Left Middle Right

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: ____________________________

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
Color-Abnormality		A	B	C	D	E	F								
		12	5	26	6	16	~								

APPENDIX D – INSTRUCTIONS TO SUBJECTS

Experiment Setup and Instructions to Subjects: MCASTL Workload Project

Advance Preparation

Turn on simulator and AV system

- Turn off the left side projectors
- Switch projector to “video” mode.
- Turn on the IP projector
- Start the “Anchor Displayer” program on the IP computer
- Turn off touch screen
- Recall video settings 01 on the video switcher.

Check audio levels

Load in DVD (make note to get DVDs if supply is low)

Make sure there are copies of all forms (consent, bio, post-test, instructions, workload ratings, payment)

Get cash to pay subjects

Turn on system to show clips

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 simulator laboratory
- Flip do not enter 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.
- Return driver’s license

- Fill out Consent Form.
  “As was noted when you were contacted earlier, we are carrying out a study of the demand of driving. While driving, people are doing all sorts of things in addition to controlling the vehicle—using phones, entering navigation information, and so forth. To determine how much is too much for drivers to do, the first step is to quantify the demand, the workload, of just driving. In this experiment, you will be rating the workload of a variety of situations in the driving simulator.

To document what we do, we videotape subjects and show outtakes of those tapes to the sponsor and the public. We want to make sure using outtakes from you is acceptable, because if it is not, you cannot participate."

Ok, given that, please read the consent form carefully, as it provides some additional details about the experiment. If you are willing to participate, then sign the consent form.

- Fill out bio form
- **"We need a few facts about you, so please fill out this biographical form."**

<u>Vision test</u>

  - 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).

<u>In-Simulator: Parked</u>

<u>Preparation</u>

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

Prepare simulator

  - Open GM Expressway Control
  - Open HyperDrive: GMExpressway
  - Load Input text file labeled "WorkloadCodePractice.txt" for the practice round
  - Then load "WorkloadCode1.txt" or WorkloadCode2.txt" depending on subject # for the actual round

- **<u>Practice Driving</u>**

"As was noted when we first contacted you and on the consent form, there is a chance that subjects can experience motion sickness. To make sure that is not a problem, there will be a short practice drive. In this drive, follow the road and drive with traffic at the posted speed limit. After about a minute or so, I will ask you to change lanes, to make sure you can do so safely. Please use your turn signal."

Make sure they do not correct excessively. If they do, intervene. If there are indications of motion sickness, say, **"Please bring the vehicle slowly to a stop."**

> After they have done so, say.
> **"It appears you are among those who is susceptible to motion and participating further is not recommended. However, we will pay you in full for coming today."** Pay them and have them sign the payment form.

Otherwise, after a minute, say, **"Ok, now would be a good time to change lanes. Please do so safely."**

After another minute

"Please bring the vehicle slowly to a stop."

<u>Practice Driving and Rating</u>

"In the driving simulator session today, you will drive a wide range of scenarios and rate the workload of each. Your ratings of workload will be relative to these 2 reference clips. Play the clips. **So, if the workload of driving a scenario was equivalent to the lower example, you might call it a "2" or equivalent to the higher example, you might call it a "6." The greater the workload, the larger the number. However, most of the situations, will not be equal to those values, but maybe in between, or higher or lower. So, ratings could be 1 or 3 or 8, or even 4-1/2. In fact, most people find they cannot rate the workload any more accurately than the nearest ½ point. Typically, values range from 1 to 10.**

In this experiment, the driving conditions are continually changing. To reduce variability in the rating process, the ratings will be the average workload over 10-15 second intervals. To help you, I will say "start" when the 15-second interval begins and "end" when it is done. Your rating should be for the average workload over that 15-second interval, not when I said start or end.

Do you have the idea of what we are trying to do?

If they say no, then explain, but do not redefine workload.

To help you understand this process, there will be a few practice trials. Are you ready?

Start the rating practice block.

"Ok, put the car in drive and drive 70 mi/hr."

When they reach the desired point.
"Start thinking about the workload now."
10-15 seconds later..
"What was the workload over the last 15 seconds?"
Note: this may be automated.
Record the rating.

After trial 3 say, **"Reduce your speed to 60 mi/hr and change to the ** lane.**
Note: The practice round (total of 5 trials consisting of 3 actual scenarios with LOS of A E and C respectively). During the first 3 trials the subject is driving 70 mi/hr. The subject will then change lanes and then drive 60 mi/hr.

Repeat this process, several times. After they get the idea, shorten the requests to "start" and "end."

"Ok, you have the idea. Bring the vehicle to a stop, put the car in park, and I will load the first of 2 experimental roads."
If their answers for the three scenarios are not anywhere close to 2, 6, and between 2 and 6 respectively, explain the workload rating process to them again, and re-do the practice trials.

<u>Main Experiment Driving and Rating</u>

Load in the first test road.

"The first road is loaded. Are you ready? Put the car into drive and accelerate to ** mi/hr and drive in the ** lane. Periodically, I will ask you to change lanes and speed in addition to rating the workload." The road takes about 20 minutes to drive.

Make sure you check their speed and lane, and get the ratings needed (start, end).

"You are near the end of the road, so pull off to the side of the road, gently bring the car to a stop, and put it in park."

Afterwards, "**Give me a moment to save the data and load in 1 more road."**

Note: Block two is in the same txt. file as block 1. All you need to do is save the data and move to block 2 in the GM expressway control screen.

Save the data and load in the second test road.

"Ready? The process is again the same. Drive the request speed, changing lanes as needed, and rate the workload when requested."
Note: Some means is need to get subjects to change lanes at the desired locations and to change speeds as well. Signs might be helpful.

Collect the ratings while driving.

Bring the vehicle to a stop, put the car in park, and I will save the data.

Save the data.

<u>Post Test</u>

"In this next to the last step, I am going to show you some clips of driving. For each one, rate the workload relative the clips from before."

They rate the clips.

"Do you have any comments?"

"You are almost done. I need to pay you for helping us. Here is $20. Please sign this payment form so we can get reimbursed.

Pay the subjects. The subject signs the form.

<u>Subject Wrap up</u>
- Walk subject to the front door
- Flip Sign

APPENDIX E – DESCRIPTION OF LEVEL OF SERVICE

Level of Service (LOS) is a measure used by traffic engineers of the extent to which traffic flows. LOS is graded A though F, where A is excellent and F is failing. LOS values can be assigned for roads and for delays at traffic signals. For this experiment, the road values are important.

Level of Service Category Descriptions

LOS	Description
A	Free-flow operation
B	Reasonably free flow Ability to maneuver is only slightly restricted Effects of minor incidents still easily absorbed
C	Speeds at or near free flow speed Freedom to maneuver is noticeably restricted Queues may form behind any significant blockage.
D	Speeds decline slightly with increasing flows Density increases more quickly Freedom to maneuver is more noticeably limited Minor incidents create queuing
E	Operation near or at capacity No usable gaps in the traffic stream Operations extremely volatile Any disruption causes queuin
F	Breakdown in flow Queues form behind breakdown points Demand > capacity

APPENDIX F – SCENARIO DETAILS

Overview

Block 1 Trials and Vehicle Maneuvers

Scenario	Trial	Subj Lane Driven	Lead Lane Driven	Lead Lane Change	S1 Lane Driven	S1 Lane Change	S2 Lane Driven	S2 Lane Change
01	01	center	center	-	left	-	right	-
02	02, 14, 24	center	center	-	left	-	right	-
03	03	center	center	-	left	-	right	-
04	04, 16, 26	center	center	-	left	-	right	center
05	05	left	center	-	left	-	right	-
06	06	left	center	-	left	-	right	-
07	07	left	center	-	left	-	right	-
08	08	left	center	-	left	-	right	-
09	09	left	center	-	left	-	right	-
10	10	left	center	-	left	-	right	-
11	11	center	center	-	left	-	right	-
12	12	center	center	-	left	-	right	-
13	13	center	center	-	left	-	right	-
14	15	center	center	-	left	-	right	-
15	17	right	center	-	left	-	center	-
16	18	right	center	-	left	-	center	-
17	19	right	center	-	left	-	center	right
18	20	right	center	-	left	-	right	-
19	21	right	center	-	left	-	right	-
20	22	right	center	-	left	-	right	-
21	23	center	center	-	left	-	right	-
22	25	center	center	-	left	-	right	-

Block 2 Trials and Vehicle Maneuvers

Scenario	Trial	Subj Lane Driven	Lead Lane Driven	Lead Lane Change	S1 Lane Driven	S1 Lane Change	S2 Lane Driven	S2 Lane Change
23	01	center	center	-	left	-	right	-
24	02, 14, 25	center	center	-	left	-	right	center
25	03	center	center	-	left	-	center	Right
26	04, 16, 27	center	center	-	left	center	right	-
27	05	right	center	-	center	left	right	-
28	06	right	center	-	left	-	right	-
29	07	right	center	-	left	-	right	-
30	08	right	center	-	left	-	right	-
31	09	right	center	-	left	-	right	-
32	10	right	center	-	left	-	right	-
33	11	center	center	-	left	-	right	-
34	12	center	center	-	left	-	right	-
35	13	center	center	-	left	-	right	-
36	15	center	center	-	left	-	center	right
37	17	left	center	-	left	-	right	-
38	18	left	center	-	left	-	center	-
39	19	left	center	-	left	-	center	-
40	20	left	center	-	left	-	center	-
41	21	left	center	-	left	-	right	-
42	22	left	center	-	left	-	right	-
43	23	left	center	-	left	-	right	-
44	24	center	center	-	left	-	right	-
45	26	center	center	-	left	-	center	right

Clips Shown in Each Block for Each Trial

Block 1			Block 2		
Clip #	**Scenario**	**Trial**	**Clip #**	**Scenario**	**Trial**
T	01	01	T	23	01
40	02	02	139	24	02
T	03	03	T	25	03
135	04	04	138	26	04
T	05	05	T	27	05
148	06	06	136	28	06
T	07	07	T	29	07
152	08	08	130	30	08
T	09	09	T	31	09
144	10	10	129	32	10
T	11	11	T	33	11
140	12	12	143	34	12
T	13	13	T	35	13
40	02	14	139	24	14
T	14	15	T	36	15
135	04	16	138	26	16
T	15	17	T	37	17
125	16	18	150	38	18
T	17	19	T	39	19
126	18	20	153	40	20
T	19	21	T	41	21
29	20	22	145	42	22
T	21	23	T	43	23
40	02	24	T	44	24
T	22	25	139	24	25
135	04	26	T	45	26
			138	26	27

T=Transition trial (to move vehicles in position)

Simulator Scenario Variables

Variable Name	Digits	Values	Units	Sample	Array Number	Text Range	Comment
SubjectNumber	2	00-99	none	01	1	1-2	(00-99) (2 Digits) The Subject Number
BlockNumber	2	00-99	none	01	2	4-5	(00-99) (2 Digits) Block Number
TrialNumber	2	00-99	none	01	3	7-8	(00-99) (2 Digits) Trial Number
Scenario	2	00-99	none	01	4	10-11	(00-99) (2 Digits) Scenario Number
SubjectSpeed	2	00-99	mph	65	5	13-14	(00-99) (2 Digits) (mph) Subject's Desired Speed, just for planning purposes or if the Scenario requires the Experimenter to tell the subject to change their speed
SubjectLane	1	1-3	none	2	6	16	(1-3) (1 Digit) Subject's Desired Lane, just for planning purposes or if the Scenario requires the Experimenter to tell the subject to change their Lane
PlatoonHeadwayBase	3	010-200	m	200	7	18-20	(010-200) (3 Digits) (m) The Average Value of the Sine Wave that controls the headway of the lead platoon vehicles
PlatoonHeadwayRangeBase	2	00-99	m	20	8	22-23	(00-99) (2 Digits) (m) The amplitude or maximum deviation from the Platoon Headway Base Value for the lead platoon vehicles
PlatoonHeadwayBaseFollow	3	010-200	m	200	9	25-27	(010-200) (3 Digits) (m) The Average Value of the Sine Wave that controls the headway of the rear platoon vehicles. This value needs to be positive if you want the rear platoon to remain behind the subject
PlatoonHeadwayRangeBaseFollow	2	00-99	m	20	10	29-30	(00-99) (2 Digits) (m) The amplitude or maximum deviation from the Platoon Headway Base Value for the rear platoon vehicles
RevealBit	1	0/1	none	0	11	32	(0 or 1) (1 Digit) Binary Bit (0/1) that controls if the reveal car will be in the trial

Variable Name	Digits	Values	Units	Sample	Array Number	Text Range	Comment
RevealDist	4	0000 or 0100-1200	m	0800	12	34-37	(0000 or 0100-1200) (4 Digits) (m) The distance down the road from the trial start point that the middle of the reveal car will be placed.
RevealLane	1	0-3	none	2	13	39	(0-3) (1 Digit) The Lane that the Reveal Car will be placed in. 0 = Nothing, 1 = Left, 2 = Middle, 3 = Right
GhostAction	1	0-1	none	1	14	41	(0 or 1) (1 Digit) 0 = Nothing (No Wind) 1 = Wind Gust
GhostDist	4	0000 or 0100-1200	m	0800	15	43-46	(0000 or 0100-1200) (4 Digits) (m) The distance down the road from the trial start point that the wind action will be performed. This is 4 digits, but only values up to 200 will be accepted. You should try to keep it under 150 m to make sure that it will trigger before the Wind Gust is canceled by the subject reaching the end of the trial
GhostForce	4	0000-9999	N	2000	16	48-51	(0000-9999) (4 Digits) (N) The force in Newtons that the wind will have.
GhostDirection	1	0/1	none	1	17	53	(0 or 1) (1 Digit) The direction that the wind will come from. Right(0) or Left(1) are the only options.
GhostForceTime	3	1.0-9.9/000	s	2.5	18	55-57	(000 or 1.0-9.9) (3 Digits) The duration of the wind gust in seconds
LeadLane	1	1-3	none	2	19	59	(1-3) (1Digit) The Lane that the Lead car is supposed to be in during the beginning of the trial. It may change lanes just after starting the trial to get into the proper lane. You can only move the lead car one lane per invocation of LeadSideLaneChangeCheckStart (which only checks at the beginning of the trial). You may also move the lead car with a Change lane action command, but again, you can only move it one lane. 1 = Left, 2 = Middle, 3 = Right

Variable Name	Digits	Values	Units	Sample	Array Number	Text Range	Comment
LeadDestLane	1	1-3	none	2	20	61	(1-3) (1 Digit) The Lane that the Lead vehicle will move to during a Lane Change action. One lane of movement only. 1 = Left, 2 = Middle, 3= Right
LeadMinHeadway	3	000-200	m	060	21	63-65	(000-200) (3 Digits) (m) The minimum headway (negative values will be tailway) for the lead vehicle. Make sure it is lower than the max headway.
LeadMaxHeadway	3	000-200	m	080	22	67-69	(000-200) (3 Digits) (m) The maximum headway (negative values will be tailway) for the lead vehicle. Make sure it is higher than the min headway.
LeadAction	1	0-2	none	1	23	71	(0-2) (1 Digit) The Action that the Lead Vehicle will take when it reaches the LeadDistanceToAction point. 0=None, 1 = Speed Change, 2 = LCM, 3 = Lane Change
LeadDecelAccel	3	0.0-9.9	m/s/s	2.0	24	73-75	(0.0-9.9) (3 Digits) (m/s/s) The acceleration rate at which the vehicle will slow or accel during a Speed Change Action
LeadDecelSpeed	2	00-99	mph	45	25	77-78	(00-99) (2 Digits) (mph) The speed that the vehicle will travel after a Speed Change Action. It will continue at this speed (ignoring min and max headway) until the next trial starts.
LeadDistanceToAction	4	0000 or 0100-1200	m	0800	26	80-83	(0000 or 0100-1200) (4 Digits) (m) The distance down the road from the trial start point that the Lead Action will be performed.

Variable Name	Digits	Values	Units	Sample	Array Number	Text Range	Comment
SideOneLane	1	1-3	none	2	27	85	(1-3) (1 Digit) The Lane that the SideOne car is supposed to be in during the beginning of the trial. It may change lanes just after starting the trial to get into the proper lane. You can only move the lead car one lane per invocation of LeadSideLaneChangeCheckStart (which only checks at the beginning of the trial). You may also move the lead car with aa Change lane action command, but again, you can only move it one lane. 1 = Left, 2 = Middle, 3 = Right
SideOneDestLane	1	1-3	none	2	28	87	(1-3) (1 Digit) The Lane that the SideOne vehicle will move to during a Lane Change action. One lane of movement only. 1 = Left, 2 = Middle, 3 = Right
SideOneMinHeadway	3 -	99-200	m	060	29	89-91	(-99-200) (3 Digits) (m) The minimum headway (negative values will be tailway) for the SideOne vehicle. Make sure it is lower than the max headway.
SideOneMaxHeadway	3 -	99-200	m	080	30	93-95	(-99-200) (3 Digits) (m) The maximum headway (negative values will be tailway) for the SideOne vehicle. Make sure it is larger than the minimum headway.
SideActionOne	1	0-5	none	1	31	97	(0-5) (1 Digit) The Action that the SideOne Vehicle will take when it reaches the SideOneDistanceToAction point. 0=None, 1 = Speed Change, 2 = LCM, 3 = Lane Change, 4 = Cutin Left, 5 = Cutin Right
SideOneDecelAccel	3	0.0-9.9	m/s/s	2.0	32	99-101	(0.0-9.9) (3 Digits) (m/s/s) The acceleration rate at which the vehicle will slow or accel during a Speed Change Action
SideOneDecelSpeed	2	00-99	mph	45	33	103-104	(00-99) (2 Digits) (mph) The speed that the vehicle will travel after a Speed Change Action. It will continue at this speed (ignoring min and max

Variable Name	Digits	Values	Units	Sample	Array Number	Text Range	Comment
							headway) until the next trial starts.
SideOneDistanceToAction	4	0000 or 0100-1200	m	0800	34	106-109	(0000 or 0100-1200) (4 Digits) (m) The distance down the road from the trial start point that the SideOne Action will be performed.
SideTwoLane	1	1-3	none	2	35	111	(1-3) (1 Digit) The Lane that the SideTwo car is supposed to be in during the beginning of the trial. It may change lanes just after starting the trial to get into the proper lane. You can only move the lead car one lane per invocation of LeadSideLaneChangeCheckStart (which only checks at the beginning of the trial). You may also move the lead car with aa Change lane action command, but again, you can only move it one lane. 1 = Left, 2 = Middle, 3 = Right
SideTwoDestLane	1	1-3	none	2	36	113	(1-3) (1 Digit) The Lane that the SideTwo vehicle will move to during a Lane Change action. One lane of movement only. 1 = Left, 2 = Middle, 3 = Right
SideTwoMinHeadway	3 -	99-200	m	060	37	115-117	(-99-200) (3 Digits) (m) The minimum headway (negative values will be tailway) for the SideTwo vehicle. Make sure it is lower than the max headway.
SideTwoMaxHeadway	3 -	99-200	m	080	38	119-121	(-99-200) (3 Digits) (m) The maximum headway (negative values will be tailway) for the SideTwo vehicle. Make sure it is larger than the minimum headway.
SideActionTwo	1	0-5	none	1	39	123	(0-5) (1 Digit) The Action that the SideTwoVehicle will take when it reaches the SideTwoDistanceToAction point. 0=None, 1 = Speed Change, 2 = LCM, 3 = Lane Change, 4 = Cutin Left, 5 = Cutin Right

Variable Name	Digits	Values	Units	Sample	Array Number	Text Range	Comment
SideTwoDecelAccel	3	0.0-9.9	m/s/s	2.0	40	125-127	(0.0-9.9) (3 Digits) (m/s/s) The acceleration rate at which the vehicle will slow or accel during a Speed Change Action
SideTwoDecelSpeed	2	00-99	mph	45	41	129-130	(00-99) (2 Digits) (mph) The speed that the vehicle will travel after a Speed Change Action. It will continue at this speed (ignoring min and max headway) until the next trial starts.
SideTwoDistanceToActi on	4	0000 or 0100-1200	m	0800	42	132-135	(0000 or 0100-1200) (4 Digits) (m) The distance down the road from the trial start point that the SideTwo Action will be performed.
FogBit	1	0/1	none	0	43	137	(0 or 1) (1 Digit) Binary Bit (0/1) that controls if the fog will be present in the current trial
FogDistance	3	000 or 010-999	m	80	44	139-141	(000 or 010-999) (3 Digits) (m) The sight distance for the subject if the fog is activated

Settings Used by Scenario Generator

Subj	Block	Trial	Scenario	Speed	SubjLane	PlatHeadBase	PlatHeadRange
01	01	01	01	65	2	100	05
01	01	02	02	65	2	100	05
01	01	03	03	70	2	175	05
01	01	04	04	70	2	175	05
01	01	05	05	70	1	110	05
01	01	06	06	70	1	100	05
01	01	07	07	70	1	175	05
01	01	08	08	70	1	175	05
01	01	09	09	70	1	175	05
01	01	10	10	70	1	175	05
01	01	11	11	70	2	150	05
01	01	12	12	70	2	150	05
01	01	13	13	65	2	120	05
01	01	14	02	65	2	100	05
01	01	15	14	70	2	175	05
01	01	16	04	70	2	175	05
01	01	17	15	65	3	175	05
01	01	18	16	65	3	175	05
01	01	19	17	65	3	175	05
01	01	20	18	65	3	175	05
01	01	21	19	60	3	090	05
01	01	22	20	60	3	090	05
01	01	23	21	65	2	120	05
01	01	24	02	65	2	100	05
01	01	25	22	70	2	175	05
01	01	26	04	70	2	175	05
01	02	01	23	65	2	125	05
01	02	02	24	65	2	125	05
01	02	03	25	65	2	175	05
01	02	04	26	65	2	175	05
01	02	05	27	65	3	150	05
01	02	06	28	65	3	150	05
01	02	07	29	65	3	150	05
01	02	08	30	65	3	150	05
01	02	09	31	65	3	150	05
01	02	10	32	65	3	150	05
01	02	11	33	65	2	123	05
01	02	12	34	65	2	175	05
01	02	13	35	65	2	125	05
01	02	14	24	65	2	125	05
01	02	15	36	65	2	175	05
01	02	16	26	65	2	175	05
01	02	17	37	70	1	100	05
01	02	18	38	70	1	100	05
01	02	19	39	70	1	175	05
01	02	20	40	70	1	175	05
01	02	21	41	70	1	150	05
01	02	22	42	70	1	150	05
01	02	23	43	65	1	150	05
01	02	24	44	65	2	125	05
01	02	25	24	65	2	125	05
01	02	26	45	65	2	175	05
01	02	27	26	65	2	175	05

PlatFollowBase	PlatFollowRange	RevealBit	RevealDist	RevealLane	GhostAction	GhostDist
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
140	10	1	0300	0	0	0000
100	10	1	0200	0	0	0000
140	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0300	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
100	10	1	0300	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	0	0000
100	10	1	0200	0	1	1000
100	10	1	0200	0	0	0000

GhostForce	GhostDirex	GhostTime	LeadLane	LeadDestLane	LeadMinHead	LaneMaxHead
0000	0	000	3	2	030	040
0000	0	000	2	2	030	040
0000	0	000	2	2	125	130
0000	0	000	2	2	125	130
1000	0	2.0	2	2	028	030
0000	0	000	2	2	028	030
0000	0	000	2	2	030	035
0000	0	000	2	2	030	035
0000	0	000	2	2	125	130
0000	0	000	2	2	125	130
0000	0	000	2	2	055	060
0000	0	000	2	2	055	060
1000	1	2.0	2	2	030	040
0000	0	000	2	2	030	040
0000	0	000	2	2	125	130
0000	0	000	2	2	125	130
1000	0	2.0	2	2	060	070
0000	0	000	2	2	060	070
0000	0	000	2	2	125	130
0000	0	000	2	2	125	130
0000	0	000	2	2	015	020
0000	0	000	2	2	015	020
0000	0	000	2	2	030	040
0000	0	000	2	2	030	040
1000	1	2.0	2	2	125	130
0000	0	000	2	2	125	130
0000	0	000	3	2	040	045
0000	0	000	2	2	040	045
0000	0	000	2	2	125	130
0000	0	000	2	2	125	130
1000	0	2.0	2	2	010	015
0000	0	000	2	2	010	015
0000	0	000	2	2	080	090
0000	0	000	2	2	080	090
0000	0	000	2	2	125	130
0000	0	000	2	2	125	130
0000	0	000	2	2	100	105
0000	0	000	2	2	100	105
1000	1	2.0	2	2	040	045
0000	0	000	2	2	040	045
0000	0	000	2	2	125	130
0000	0	000	2	2	125	130
1000	0	2.0	2	2	025	030
0000	0	000	2	2	025	030
0000	0	000	2	2	0-2	002
0000	0	000	2	2	0-2	002
0000	0	000	2	2	045	050
0000	0	000	2	2	045	050
0000	0	000	2	2	040	045
0000	0	000	2	2	040	045
0000	0	000	2	2	040	045
1000	1	2.0	2	2	125	130
0000	0	000	2	2	125	130

LeadAction	LeadAccel	LeadSpeed	LeadDistToAct	Side1Lane	Side1Dest	Side1MinHead
3	000	00	0000	1	1	023
0	000	00	0000	1	1	023
0	000	00	0000	1	1	125
0	000	00	0000	1	1	125
0	000	00	0000	1	1	030
1	2.0	65	0200	1	1	030
0	000	00	0000	1	1	125
1	2.0	68	0200	1	1	125
0	000	00	0000	1	1	125
0	000	00	0000	1	1	125
0	000	00	0000	1	1	-10
1	2.0	74	0200	1	1	-10
0	000	00	0000	1	1	023
0	000	00	0000	1	1	023
0	000	00	0000	1	1	125
0	000	00	0000	1	1	125
0	000	00	0000	1	1	060
1	2.0	68	0200	1	1	060
0	000	00	0000	1	1	125
0	000	00	0000	1	1	125
0	000	00	0000	1	1	005
1	2.0	62	0200	1	1	005
0	000	00	0000	1	1	023
0	000	00	0000	1	1	023
0	000	00	0000	1	1	125
0	000	00	0000	1	1	125
3	000	00	0100	1	1	-20
1	2.0	68	0200	1	1	-20
0	000	00	0000	1	1	030
0	000	00	0000	1	2	030
0	000	00	0000	2	1	-20
1	2.0	70	0200	1	1	-15
0	000	00	0000	1	1	040
0	000	00	0000	1	1	040
0	000	00	0000	1	1	125
0	000	00	0000	1	1	125
0	000	00	0000	1	1	-10
1	2.0	67	0200	1	1	-10
0	000	00	0000	1	1	-20
1	2.0	68	0100	1	1	-20
0	000	00	0000	1	1	030
0	000	00	0000	1	2	030
0	000	00	0000	2	1	040
0	000	00	0000	1	1	040
0	000	00	0000	1	1	045
0	000	00	0000	1	1	045
0	000	00	0000	1	1	110
1	2.0	65	0200	1	1	110
0	000	00	0000	1	1	050
0	000	00	0000	1	1	-20
1	2.0	68	0200	1	1	-20
0	000	00	0000	1	1	030
0	000	00	0000	1	2	030

Side1MaxHead	Side1Act	Side1Accel	Side1Speed	Side1DistToAct	Side2Lane	Side2Dest
026	0	000	00	0200	3	3
026	1	2.0	70	0200	3	3
130	0	000	00	0000	3	3
130	0	000	00	0000	3	2
035	0	000	00	0000	2	3
035	0	000	00	0000	3	3
130	0	000	00	0000	3	3
130	0	000	00	0000	3	3
130	0	000	00	0000	3	3
130	0	000	00	0000	3	3
0-5	0	000	00	0000	3	3
0-5	1	2.0	77	0200	3	3
026	0	000	00	0000	3	3
026	1	2.0	68	0200	3	3
130	0	000	00	0000	3	3
130	0	000	00	0000	3	2
070	0	000	00	0000	2	2
070	1	2.0	70	0200	2	2
130	0	000	00	0000	2	3
130	0	000	00	0000	3	3
010	0	000	00	0000	3	3
010	1	2.0	65	0500	3	3
026	0	000	00	0000	3	3
026	1	2.0	70	0200	3	3
130	0	000	00	0000	3	3
130	0	000	00	0000	3	2
-15	0	000	00	0200	3	3
-15	1	3.0	73	0200	3	2
035	0	000	00	0000	2	3
035	3	000	00	0440	3	3
-15	3	000	00	0200	3	3
-10	1	2.0	72	0200	3	3
045	0	000	00	0000	3	3
045	1	2.0	68	0200	3	3
130	0	000	00	0000	3	3
130	0	000	00	0000	3	3
0-5	0	000	00	0000	3	3
0-5	1	2.0	70	0200	3	3
-15	0	000	00	0000	3	3
-15	1	3.0	73	0200	3	2
035	0	000	00	0000	2	3
035	3	000	00	0440	3	3
045	3	000	00	0200	3	3
045	0	000	00	0000	3	2
050	0	000	00	0000	2	2
050	0	000	00	0000	2	2
115	0	000	00	0000	2	3
115	0	000	00	0000	3	3
060	0	000	00	0000	3	3
-15	0	000	00	0000	3	3
-15	1	3.0	73	0200	3	2
035	0	000	00	0000	2	3
035	3	000	00	0440	3	3

Side2MinHead	Side2MaxHead	Side2Act	Side2Accel	Side2Speed	Side2DistToAct	Fogbit	FogDist
020	023	0	000	00	0000	0	000
020	023	1	2.0	68	0200	0	000
100	110	0	000	00	0000	0	000
100	110	3	000	00	0400	0	000
035	040	3	000	00	0010	0	000
035	040	1	2.0	60	0500	0	000
060	070	0	000	00	0000	0	000
060	070	1	2.0	60	0200	0	000
020	025	0	000	00	0000	0	000
020	025	1	2.0	65	0200	0	000
070	075	0	000	00	0000	0	000
070	075	1	2.0	68	0200	0	000
020	023	0	000	00	0000	0	000
020	023	1	2.0	68	0200	0	000
100	110	0	000	00	0000	0	000
100	110	3	000	00	0400	0	000
010	015	0	000	00	0000	0	000
010	015	1	2.0	68	0200	0	000
125	130	3	000	00	0200	0	000
125	130	0	000	00	0100	0	000
020	025	0	000	00	0000	0	000
020	025	0	000	00	0000	0	000
020	023	0	000	00	0000	0	000
020	023	1	2.0	68	0200	0	000
100	110	0	000	00	0000	0	000
100	110	3	000	00	0400	0	000
025	030	0	000	00	0000	0	000
025	030	3	000	00	0500	0	000
090	100	3	000	00	0200	0	000
090	100	0	000	00	0000	0	000
115	120	0	000	00	0000	0	000
115	120	1	2.0	60	0200	0	000
125	130	0	000	00	0000	0	000
125	130	0	000	00	0000	0	000
040	045	0	000	00	0000	0	000
040	045	0	000	00	0000	0	000
035	040	0	000	00	0000	0	000
035	040	1	2.0	63	0200	0	000
025	030	0	000	00	0000	0	000
025	030	3	000	00	0500	0	000
090	100	3	000	00	0200	0	000
090	100	0	000	00	0000	0	000
060	070	0	000	00	0000	0	000
060	070	3	000	00	0100	0	000
075	080	0	000	00	0000	0	000
075	080	1	2.0	67	0200	0	000
040	045	3	000	00	0010	0	000
040	045	1	2.0	64	0200	0	000
025	030	0	000	00	0000	0	000
025	030	0	000	00	0000	0	000
025	030	3	000	00	0500	0	000
090	100	3	000	00	0200	0	000
090	100	0	000	00	0000	0	000

APPENDIX G – ORDER OF BLOCKS

Blocks Balanced across Subjects

Order of Blocks	Subject #
1, 2	1, 2, 5, 6, 9, 10, 13, 14
2, 1	3, 4, 7, 8, 11, 12, 15, 16

APPENDIX H – MEAN WORKLOAD RATINGS

Clip #	Scenario #	SAVE-IT 2c		This Report							
				All Subjects				2 Subjects Omitted			
		Clip ratings	N	Clip ratings	N	Driving ratings	N	Clip ratings	N	Driving ratings	N
40	2	5.7	48	4.2	16	4.1	48	4.4	14	4.3	42
148	6	6.7	48	5.3	16	4.4	16	5.6	14	4.7	14
152	8	3.8	48	3.2	16	3.2	16	3.5	14	3.3	14
144	10	2.8	48	2.3	16	2.3	16	2.3	14	2.3	14
140	12	4.1	48	3.4	16	3.4	16	3.5	14	3.6	14
126	18	2.3	48	2.0	16	2.0	16	2.0	14	2.0	14
29	20	6.8	48	5.4	16	5.0	16	5.8	14	5.4	14
136	28	4.8	48	4.4	16	3.4	16	4.7	14	3.6	14
130	30	3.5	48	2.9	16	2.1	16	3.1	14	2.1	14
129	32	2.9	48	2.4	16	3.0	16	2.5	14	3.1	14
143	34	4.8	48	3.7	15	3.2	16	3.9	14	3.4	13
150	38	6.5	48	4.8	16	4.4	16	5.2	14	4.7	14
153	40	5.1	48	4.4	16	4.2	16	4.8	14	4.5	14
145	42	3.6	48	3.5	16	3.7	16	3.8	14	3.9	14