Weighting Waiting: Evaluating the Perception of In-Vehicle Travel Time Under Moving and Stopped Conditions

David Levinson, University of Minnesota, Department of Civil Engineering
Kathleen Harder, University of Minnesota, College of Architecture and Landscape Architecture
John Bloomfield, University of Minnesota, College of Architecture and Landscape Architecture
Kasia Winiarczyk, University of Minnesota, Department of Civil Engineering

Conference paper
Session XXX

Moving through nets:
The physical and social dimensions of travel
10th International Conference on Travel Behaviour Research
Lucerne, 10-15. August 2003

Weighting Waiting: Evaluating the Perception of In-Vehicle Travel Time Under Moving and Stopped Conditions

David Levinson[1], Kathleen Harder[2], John Bloomfield[3], and Kasia Winiarczyk[4]

1. Assistant Professor, Department of Civil Engineering, University of Minnesota
2. Human Factors Researcher, College of Architecture and Landscape Architecture, University of Minn esota
3. Human Factors Researcher, College of Architecture and Landscape Architecture, University of Minnesota and Visiting Professor of Human Factors, University of Derby, UK
4. Research Assistant, Department of Civil Engineering, University of Minnesota

David Levinson
Assistant Professor
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455 USA

Phone:: 01-612-625-6354
Fax: 01-612-626-7750
eMail: levin031@umn.edu

Abstract

This paper describes experiments comparing traditional computer administered stated preference with virtual experience stated preference to ascertain how people value stopped delay compared with stop-and-go or freeflow traffic. The virtual experience stated preference experiments were conducted using a wrap around driving simulator. The two methods produced two different results, with the traditional computer assisted stated preference suggesting that ramp delay is 1.6 – 1.7 times more onerous than freeway time, while the driving simulator based virtual experience stated preference suggested that freeway delay is more onerous than ramp delay. Several reasons are hypothesized to explain the differences, including recency, simultaneous versus sequential comparison, awareness of public opinion, the intensity of the stop-and-go traffic, and the fact that driving in the real-world is a goal directed activity. However without further research, which, if any, of these will eventually prove to be the reason is unclear. What is clear is that a comparison of the computer administered stated preference with virtual experience stated preference produces different results, even though both procedures strive to find the same answers in nominally identical sets of conditions. Because people experience the world subjectively, and make decisions based on those subjective experiences, future research should be aimed at better understanding the differences between these subjective methodologies.

Keywords

Ramp Meter, Value of Time, Stated Preference, Driving Simulator, International Conference on Travel Behaviour Research, IATBR

Preferred citation

Levinson, David, Kathleen Harder, John Bloomfield, Kasia Winiarczyk (2003) Weighting Waiting: Evaluating the Perception of In-Vehicle Travel Time Under Moving and Stopped Conditions, paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, August 2003.

1. Introduction

While a clear majority of drivers would prefer free unobstructed flow on any route they choose, at any time, this is unrealistic in many cities. Attempts to treat congestion using ramp meters (having drivers wait at freeway entrance ramps to ensure smooth flow on the freeways), have been implemented in many cities, including Minneapolis and St. Paul, Minnesota, USA. However this policy has been subject to considerable controversy over whether it reduces total delay on the system (it seems to), and how it redistributes delay (some wait more so others can proceed unimpeded) (Levinson, Zhang, Das, & Sheikh, 2002). If all time were interchangeable, this might not be a problem. However time passes slowly when stopped and waiting, so though the ramp metering system may reduce total travel time, it may not reduce total perceived travel time, or total utility to travelers who weight waiting time more heavily than time in motion, even during stop-and-go traffic.

Historically, the most common distinction in studied time value differences has been between out-of-vehicle time and in-vehicle time. The differences between the relative economic values of walking or waiting versus traveling in a vehicle have been thoroughly researched in the transportation literature. In recent years some researchers have started including more qualitative descriptions of in-vehicle time in their analyses, such as free flow traffic travel time or slowed down congested travel time (Hensher 2001). This paper further distinguishes between the qualitatively dissimilar in-vehicle travel experiences of waiting at a ramp meter, driving in free flow traffic, and driving in congested traffic, estimating the driver acceptance of each.

At the same time, this paper develops a new methodological approach to studying such questions. Most travel time value experiments use either paper and pencil or a two-dimensional computer screen presentation in the process of eliciting driver preferences. We also present the different options to drivers via runs in a state-of-the art driving simulator (AutoSim 2002). The simulation is an immersive driving experience, very much like actual driving.

The results of the simulator run, what we call a *Virtual Experience Stated Preference* (abbreviated VESP in the paper), are compared with data obtained with a more traditional methodological approach in an analogous *Computer Administered Stated Preference* (CASP) test. The CASP test is run using nominally identical conditions, in which the ramp and freeway times are represented in both a textual and a graphical manner on a computer screen. The subject rates each condition. Then after several conditions are presented, the subject ranks them. To correlate perception with other variables, each subject reports basic demographic, socioeconomic, and one-day travel diary data. Travel diaries are received from the subjects before they participate in the driving simulator run.

In brief we ask, would the typical driver prefer to a) spend time waiting at a ramp meter and experience faster moving traffic once on the highway, or b) spend no time at the ramp meter but experience the commute in relatively slow moving traffic? There are many factors that could possibly influence the drivers' preferences. People don't generally seem to like to wait in queues or remain on hold on the telephone or, for that matter, wait at ramp meters. Short wait times are experienced as annoying, long wait times are experienced as being tedious. On the other hand, highly congested traffic is uncomfortable because of the high level of sustained attention required to drive without bumping into the other cars and the frustrating nature of stop-and-go motion.

This research consists of two separately designed experiments that use similar elements and equipment, but have a different overall structure. Both of the experiments present the subject with a tradeoff between spending time at a ramp meter before entering the highway versus the wait associated with various traffic volumes while on the highway. The longer the wait experienced at the ramp meter, the less congestion is encountered by the subject once actually on the road. The shorter the wait at the ramp meter, the greater the congestion during the remainder of the trip.

Condition Set #1 exposes the subjects to four test conditions (A, B, C, and D). Condition A is the control condition; with a ramp meter wait time of zero minutes and the slowest travel speed of approximately 48 km/hr (30 mph). The other three conditions have varied ramp meter wait times, from 2 to 6 minutes, which are paired with faster travel speeds. We are not testing very long ramp meter wait times, such as 10 minutes and upwards, because it is expected (from prior experience) that these are unequivocally disliked, no matter what balancing factors might be presented. In this test, a minute of ramp delay saves more than a minute of freeway time. (We also conducted a Condition Set #2 in the CASP, which exposes the subjects to four different test conditions (E, F, G, and H), similar to A, B, C, and D above, but ensuring total travel time is the same, so that each minute of ramp wait saves exactly a minute of freeway time. However, we don't present detailed Condition Set #2 results here for brevity.)

In the experiments, the test runs are counterbalanced, using a Latin Squares design, in order to eliminate possible order and fatigue effects. These experiments provide important information for timing of ramp meters, providing for the first time explicit trade-offs between freeway delay and ramp delay, so that rather than simply minimizing total travel time on the freeway system, recommendations can be based on a utility-based approach that minimizes total perceived travel time. These recommendations should increase user satisfaction with the ramp meter system.

This paper begins with a discussion of the value and perception of time. This is followed by a review of survey and experimental methods. We then describe our methodology in detail, including the selection of subjects, survey methodology, and the driving simulator. The survey administration is then described, including a presentation of the relevant choice conditions. Our hypotheses are stated. Our results, derived from a rank order logit procedure, are presented. The CASP and VESP are compared. We conclude with some thoughts about the implications of this research for the future of stated preference analysis.

2. Value and Perception of Time

It can be argued that a wait at a ramp meter is analogous to a wait in other consumer-based situations. After all, a ramp meter is simply the queue in which drivers wait to be "served" on the highway, in the same way that customers wait to be served in a store or restaurant or post office. Consumers, while generally disliking waiting for services, accept waiting under certain conditions (Houston, Bettencourt, & Wenger 1998). Strong negative correlation between waiting time and a customer's evaluation of the quality of service is both supported by empirical evidence and used by Houston *et al.* (1998) to develop a field-based theory of how the concepts of consumers' wait time expectations and prior service encounters can be applied to waiting at ramp meters. The acceptability of waiting trades off the negative utility of the wait (e.g. time at ramp) with the positive utility of the goal (e.g. better flowing traffic).

There is evidence (Hui & Zhou 1996) that the presence of waiting duration information has a positive effect on customer service quality judgments, perhaps due to subjects perceiving greater control of the situation. There is also some evidence that time already spent waiting in queue (or sunk cost) is less negatively perceived by the customer when the goal of the waiting is considered to be more attractive (Meyer 1994). The use of Variable Message Signs (VMS) at ramp meters might be added to future studies to see to what degree information ameliorates wait at ramp meters. Further, education may matter. People may be more inclined to view the wait at the ramp meter positively if they believed that this were a way of achieving an attractive goal – better driving conditions (closer to free flow traffic levels) once they get out on the highway.

In a waiting situation, anxiety and stress may build up in an individual due to the sense of a waste of time and to the uncertainty of the duration of the wait. Theoretical models (Osuna 1985) and empirical studies both find that information about expected wait length can relieve a large amount of the stress associated with the uncertainty of the duration of the wait.

Drivers who are waiting at the ramp meter may try to predict the amount of time they will wait by performing a simple estimate of light timing duration multiplied by the amount of cars in queue in front of them — in this case, the uncertainty stress may be taken out of the equation. Then, what remains will be just the boredom of the wait itself. The boredom may be alleviated by the driver engaging in other activities (talking on cell phone) and by distractions (listening to radio, books on tape). For drivers who would engage in these boredom-reducing strategies a ramp meter could be a relatively pleasant environment. In contrast, the uncertainty stress cannot be alleviated when traveling in start and stop traffic in the absence of information about the duration of the congestion.

Both driver stress and aggression are higher in high traffic congestion conditions than in low congestion conditions (Hennessy & Wiesenthal, 1999). Annoyance is greater among high- and medium-impedance commuters than among those experiencing the low-impedance commute (Stokols Novaco, Stokols, & Campbell 1978)

Drivers may have an *a priori* perception that traffic managed by ramp meters is less variable, therefore less of a time risk than potentially more variable congested traffic. This reasoning may lead them to prefer the more certain ramp meter condition. A recent investigation of average travel time and travel time variability effects on route choices is inconclusive, but provides interesting discussion and review of concepts involved (Katsikopoulos, Duse-Anthony, Fisher, & Duffy 2000).

Commuting activity (car or bus) raises both pulse rate and systolic blood pressure in subjects (White & Rotton 1998) in the field. Stokols *et al.* (1978)., measured blood pressure and heart rate at prearranged stations located in workplace parking lots immediately after the subjects exited their car following their commute, finding increases in commute distance and time account for a significant proportion of the increase in driver blood pressure.

Redmond and Mokhtarian (1999) demonstrated that the commute itself can offer some positive utility to the commuter. Some people do report that their commute is too short, and those that report that their commute is too long typically just want shorter commute times, but not the complete elimination of their commute. The positive utility of the commute appears to be composed of enjoyment of travel for its own sake, and of allowing secondary activities to be conducted while driving, as well as acting as a buffer between the work and home zones.

The ability to conduct other activities should be negatively impacted by stop-and-go freeway conditions, when more attention needs to be focused on the road. And it should be positively impacted by waiting at a ramp meter when the driver can pursue useful secondary tasks such as listening to the radio or talking on the phone in relative comfort. If this assessment is correct, we might expect different results in a CASP (where people might imagine doing something else) than in the VESP simulator experiment, where the drivers are constrained from other activities (radio, work, etc.)

Revealed preference (RP) and SP data have been used in studies that assess the monetary equivalent of value of travel time savings (VTTS) and the relative importance of different components of travel time. The models that have been used include the multinomial logit (MNL) model, the nested logit (NL) model, and the mixed logit/random parameter logit (ML/RPL) models. A body of recent research suggests that the popular MNL model tends to under-estimate the mean VTTS. This effect has been found in studies of urban commuting, long distance inter city travel, and urban non-commuting travel by Hensher (2000a,b and 2001a,b) and Bhat (1995).

An experiment conducted by Hensher (2001a) distinguishes between the value of considering traffic conditions like free flow time, slowed down time, and start/stop time – rather than the typical grouping of in-vehicle time and out-of-vehicle time. It is a CASP experiment in which the choice is made between the attributes of the current trip and two other alternative scenarios. Travel occurs between predetermined locations, with no route switching. The design consists of two unlabeled alternatives, each defined by six attributes: free flow travel time, slowed down travel time, stop/start travel time, uncertainty of travel time, running cost, and toll charges. Hensher (2001a) develops a series of models (MNL and ML/RPL) in order to identify the role of each trip attribute, finding that VTTS for start/stop is approximately 5-10 times that for free flow, while VTTS for slow traffic is 2-3 times that for free flow.

This result implies that start/stop is less preferable than slow traffic, which in turn is less preferable than free flow. This is an expected, common sense result. What the modeling process adds to this result is an evaluation of the approximate magnitude of these preferences in relation to each other.

3. Survey Methods

Transport models which are based on observed choices made by individual travelers rather than on observed relationships for groups of travelers, are known as disaggregate demand models (DM). These models are probabilistic. The theoretical base for DM is random utility theory. DM models are typically discrete-choice, logit type models. In a discrete choice model an individual selects an option from a finite set of alternatives. The probability of individuals choosing a given option is a function of the relative attractiveness of the option and characteristics of the particular individual.

The concept of utility is used to represent the levels of attractiveness of different alternatives. A utility value can be calculated from a specific utility function. The utility of the alternative is then compared with the utility of the other options and transformed into a probability value that is between zero and one. The nature of this transformation is dictated by a transformation function, such as the logit function. The most popular and simplest of the discrete choice models is the multinomial logit model (MNL).

There are many practical difficulties in measuring the value of time in a direct manner, therefore indirect methods, such as collecting revealed preference data, have been developed and used. Revealed preference (RP) research is based on observations of actual choices made by individuals. However, there have been questions raised about the validity of the values derived from revealed preference data. These questions relate to issues of sample selection, aggregation, other sources of bias, and the ability to identify appropriate demand functions. Confounding factors in the revealed preference data can make the statistical estimation of individual factors involved in choice inaccurate. Also, it can be difficult to find real scenarios that would have all of the attributes desirable for a particular study.

Stated Preference (SP) experiments are an alternative to the RP studies. In SP experiments an individual is faced with the choice between alternatives that contain predetermined levels of specific attributes. The combinations of levels of attributes are systematically varied in order to reveal the profile of the individual preferences. Typically, a sample of travelers is assembled and then observed making choices between two different "bundles" of trip attributes. These attributes can be running travel costs, travel time costs, or external costs such as tolls. An SP experiment allows us to disentangle the independent contributions of each attribute component.

However, SP methods are not without their own methodological problems. They are typically suited to choice experiments, which are at best only analogs to the real world. After all, the world is multi-dimensional while the stated choice universe has only a few dimensions. It is legitimate to question whether or not something is lost in translation, despite the superior control.

In the current study, we use immersive driving simulation to develop what we call *Virtual Experience Stated Preference* (VESP). Unlike traditional stated preference (SP), we are not simply asking about a hypothetical scenario; rather the subject experiences (and evaluates) a virtual simulation of that scenario. We know of no prior study of ramp metering that has been conducted in such a manner. However, immersive driving simulators have been used to investigate a wide range of issues. They have been used to explore future designs of highway systems (Levitan & Bloomfield 1998; Levitan, Golimbowski, & Bloomfield 1998), rural intersections (Harder, Bloomfield, & Chihak *in preparation*), pavement markings (Bloomfield, Brown, Durlak, & Bartelme 1997), lane widths and centerline rumble strips (Harder, Carmody, & Bloomfield 2002a), in-lane rumble strips (Harder & Bloomfield 2002), vehicle displays (Harder, Bloomfield, & Chihak *in press*), and traffic calming strategies (Harder, Carmody, & Bloomfield 2002b); to investigate stopping behavior (Bloomfield, Harder & Chihak *in preparation*); to explore the effects of driving at night (Bloomfield, Bartelme, Brown, & Grant 1997), in fog (Bloomfield, Brown, Grant, & Reinach 1996; Bloomfield, Grant, Levitan, Cumming, Maddhi, Brown, & Christensen 1998), and in snow (Harder, Bloomfield, & Chihak 2003; Harder, Bloomfield, & Chihak *in press*); to examine crash scenarios (Lee, McGehee, & Brown 2000; Rizzo, McGehee, Dawson, & Anderson 2001); and to explore the effects of anti-histamine drugs and alcohol (Bloomfield, Weiler, Grant, Brown, Layton, Grant, Woodworth, McKenzie, Baker, & Young 2000; Weiler, Bloomfield, Woodworth, Grant, Layton, Brown, McKenzie, Baker, & Watson 2000), medical devices such as intraocular lens implants (Featherstone, Bloomfield, Lang, Miller-Meeks, Woodworth, and Steinert 1999), age and Alzheimer Disease (Rizzo, McGehee, Dawson, & Anderson 2001) and fatigue (Weirwillie, Wreggit, Kirn, Ellsworth, & Fairbanks 1994; Ranney, Simmons Boulos, & Macchi 1999).

4. Methodology

4.1 Subject Selection

The procedure for obtaining subjects for the CASP and VESP experiments were similar. We received a list of all University of Minnesota employees, exclusive of faculty, from the Office of Institutional Research and Reporting. The list contained a total of 15,288 employee names at all University branches. We further excluded 5,404 employees located outside Minneapolis, 19 Civil Engineering Department employees, and 100 individuals who had received an email for this study's pilot, leaving 9,765 names remaining.

For the CASP experiment, from this pool we randomly selected 1308 names and sent emails to them. We received 209 replies. From this group, self-identified non-driver commuters were excluded. Ultimately, 89 subjects were scheduled in groups of five at one-hour intervals over a four-day period. There were 15 no-shows, but the remaining 74 subjects participated in the study and completed the survey. Of these 74, five more subjects were eliminated from analysis, because while in the testing facility they said they commuted primarily by bus, biking, or walking rather than by driving a motor vehicle. This left a total of 69 subjects who gave valid observations.

For the VESP experiments, we took our list of 9,765 and subtracted out the 1,308 emails already sent, leaving 8,457 potential subjects. A similar recruiting procedure to that used in the CASP experiment was undertaken. In addition, phone interviews of those respondents were conducted to filter out unsuitable subjects. As before, we excluded self-identified non-driver commuters. We also screened out subjects likely to be prone to simulator sickness by asking them a standard battery of questions. After accounting for no-shows, there were 42 participants. Several problems arose: software crashes or bugs occurred in six of the scheduled runs; simulator sickness affected four subjects; In the end 32 simulator runs were successfully completed following the Condition Set #1 protocol. They received $50 for their participation.

4.2 Computer Administered Survey

Subjects were mailed a travel diary and asked to complete it the day before they were to take the survey. Travel diaries were handed in upon arrival at the testing facility. Those subjects who had completed but failed to bring their travel diaries were asked to send them in by campus mail. Then, the subjects completed a computer-based survey instrument. This instrument included: a demographic/socioeconomic questionnaire, a transportation attitude survey, and a personality survey (the results of the surveys are not reported in this paper). In the CASP experiment, subjects also completed a ramp meter vs. highway comparison ratings and rankings. In the VESP experiment the ratings and rankings were asked of the subject following the experience in the driving simulator.

For the CASP experiment, the surveys were administered in a group setting, with between two and five subjects taking the survey at the same time. It took approximately 15-20 minutes for the subjects to complete the survey. Compensation of $10 was given to the subjects after the successful completion of the assignment.

The surveys were administered in the form of a Microsoft Access Database constructed specifically for this study. The database recorded all subject responses. Data recording was verified after each subject completed the survey.

4.3 Driving Simulator

For the VESP experiments, each subject drove in the advanced driving simulator in the University of Minnesota's HumanFIRST laboratory. Key components of this simulator are as follows:

Simulator Vehicle—The simulator vehicle is a full-body 2002 Saturn SC1 coupe.

Simulator Visuals—The driver of the simulator vehicle has a 210-degree forward field-of-view. This 210-degree forward field of view is provided by five flat-panel screens—each of which measures 4.7-ft high and 6.5-ft-wide. There is a central flat panel in front of the simulator vehicle. The center of this panel is aligned with the line of sight of the driver of the simulator vehicle. Two intermediate panels flank this central panel, with one on the left and one on the right. They are set at 138-degrees to the central panel. Then, there are two outer panels—again, one to the right, the other to the left—set at 138-degrees to the intermediate panels. All five flat-panel screens are elevated 16 inches from the ground. Five projectors are used to project a coordinated, high fidelity, virtual environment onto the five flat-panels that comprise the 210-degree forward field-of-view. In addition, the simulator provides rear-view imagery, via a 10-ft high by 7.5-ft wide screen mounted behind the vehicle that the driver sees through the vehicle's rear-view mirror, and by two 5-inch LCD screens that are installed in place of the simulator vehicle's side-view mirrors.

Simulator Vehicle Controls—The simulator vehicle's controls are equipped with sensors that relay the subject's inputs to the steering wheel, transmission, and accelerator and brake pedals, to the driving simulator computer. This provides a real-time interface with the virtual environment. Force feedback is applied to the steering wheel, using a high-torque motor attached to the steering column. A vacuum assist pump is connected to the brake pedal in order to simulate realistic braking. The simulator vehicle is equipped with an automatic transmission interface, which is functional and is controlled by the simulator computer.

Simulator Sound System—Road and traffic noise, and the simulator vehicle's engine sounds are delivered through four speakers placed around the car's exterior near the base of the forward screen. Each speaker receives independent inputs from the simulator's 3D sound generation system. Low-frequency sounds are delivered using a ten-inch subwoofer placed inside the simulator vehicle's engine compartment. Recorded instructions can also be delivered through the four speakers at the base of the forward screen. In addition, during the experimental session, the experimenter communicates with each subject via a dedicated intercom system that makes use of four speakers that are installed in the simulator vehicle's factory speaker locations.

Simulator Vehicle Movement—A bass shaker mounted to the underside of the vehicle's frame provides additional low-frequency vibration. Servo-motors attached to the suspension components at each of the rear tires provide a partial motion base, however they were not used in this experiment.

Data Recording—The virtual position of the simulator vehicle, relative to the scenario that each subject drives, is recorded at 20 Hz throughout each experiment drive. From this record, it is possible to determine the subject's steering performance and the speed of the vehicle. In addition, three micro-video cameras positioned in the cab of the simulator vehicle are used to record (i) the subject's face, (ii) his or her foot position, and (iii) his or her steering wheel responses throughout the course of the experimental session. A video display at the experimenter's station enables the experimenter to monitor the subject throughout each session.

There was no working clock in the vehicle, although subjects were allowed to keep their watch. There was also no working radio in the vehicle. The simulator is shown in Figures 1 and 2.

4.4 Choice Conditions

Condition Set #1 consists of the four runs (labeled A, B, C, and D) that are shown in Table 1. The four conditions in Condition Set #1 involve trips that are of varying duration. All runs were 10 miles (16 km) long.

Table 1: Choice Conditions

	Ramp Wait	Drive Time	Speed
Condition	min	min	mph (km/hr)
A	0	20	30 (48)
B	2	15	40 (64)
C	4	12	50 (80)
D	6	10	60 (96)

The order of the four conditions was randomized using Latin Squares. After each condition, the subjects rated the trip on a 1-7 rating scale (with seven being best, and one being worst).

In addition, after all 4 runs were completed, the subjects ranked the conditions in order of preference. In the CASP experiment, the conditions associated with Condition Set #1 were presented to the subjects first. Then, Condition Set #2 was presented. In addition, we presented additional variations on Condition Set #1 where we doubled and tripled the times involved. Because of the time required to actually drive the simulated trips in the VESP simulations, subjects only experienced the four conditions associated with Condition Set #1.

5. Hypotheses

The data that we collected allow us to test numerous hypotheses. In this paper, we focus on two.

First concerning methodology, our hypothesis was that subjects in the CASP experiment would respond similarly to subjects in the VESP experiment. If the results are similar, then we can be confident in doing this type of research using either traditional CASP experiments or VESP simulations. Because computer-based experiments are likely to remain less expensive, if we obtain similar results that would suggest that simulators are not needed to address this question. On the other hand, if the results differ, we may need further research to determine which method provides more valid results.

The second hypothesis concerns the relationship between the effects of different types of time (waiting at a ramp meter vs. freeway travel at various speeds) on driver preference. We posit that ramp time is more onerous than freeway travel.

6. Results from CASP Experiments

Figures 3 and 4 show the ratings and rankings from the CASP experiments. By inspection, it would appear that some ramp wait is tolerable, and even preferable to no wait and more congestion. However with the longer ramp waits, the subjects prefer to spend time on the freeway.

To test that observation, we estimate a multinomial rank-ordered logit model using Stata (2003). Rank-ordered logit has been used in a number of previous studies (Katahira 1988; Koop and Poirier 1994; Colias & Salazar-Velasquez 1995; Carson 1996; Layton, Gardner, & Brown 1999; Calfee, Winston, & Stempski 2001). We compared the results of the rank-ordered logit with a binomial logit (pairwise comparing A v. B, A v. C, A v. D, B v. C, B v. D, and C v. D) as separate observations and reached very similar results, which are not shown for brevity.

After dropping observations with missing data, and using the Condition Set #1 data, we have 4 ranks each from 48 individuals for 192 observations. The dependent variable was the rank of the choice. The final utility expression took the following form

$$U = f(R, F, S, T)$$

Where:

R = ramp wait (min)
F = freeway travel time (min)
S = subject's sex (1 if male, 0 if female)
T = subject's reported daily commute time (min)

The results are presented in Table 2.

Table 2: Rank Ordered Logit Model of CASP Condition Set #1

	. I			. 2			. 3		
	β	z	P > |z|	β	z	P > |z|	β	z	P > |z|
R/F	-1.61	-3.94	0.000						

R				-1.561	-6.31	0.000	-1.404	-5.40	0.000
F				-0.816	-5.77	0.000	-0.828	-5.79	0.000
S*F				-0.121	-1.93	0.053	-0.127	-2.00	0.045
R/T							-4.681	-1.65	0.100
(Obs,groups)	(192,48)			(192,48)			(192,48)		
Log likelihood	-144.141			-125.406			-123.989		
LR chi^2	16.81			54.28			57.12		
Prob > chi^2	0.00			0.00			0.00		

First, the signs of the variables are in the expected direction, additional time, be it ramp delay or freeway travel time reduces utility. Consistent with our hypothesis, the results indicate that each minute of ramp wait is 1.6 (Model 1) to 1.7 (Models 2 and 3) times more onerous than freeway delay (in Models 2 and 3, it is determined by the ratio of the β's).

Further, these results indicate that the shorter the subject's actual commute, the more onerous is ramp wait. This observation is logical, as ramp wait becomes a larger percentage of the trip the shorter the trip. Individuals with short commutes may have internalized that assessment when comparing the alternatives. Men dislike freeway time somewhat more than women (or tolerate ramp time somewhat better).

Other variables (age, income, speed instead of time) were tested but dropped for being statistically insignificant. Other functional forms (log, quadratic) were also tested but offered no additional explanatory power over the linear model shown in Table 2. This point is interesting, and worth further experimentation, as one might anticipate threshold effects in the data (four minutes of ramp delay is more than twice as bad as two minutes of ramp delay). These results are substantively consistent with results obtained from other statistical methods (binomial logit on each choice pair).

7. Results from VESP Experiments

We estimated similar models with the VESP data again using rank-ordered logit. In contrast to the CASP, simulator users preferred almost uniformly the ramp wait rather than stop-and-go traffic on the freeway. Thus, the signs in VESP are opposite to those obtained in the CASP. In the CASP, ramp time and freeway time were both onerous, but ramp time was more onerous than freeway time. In the VESP, more ramp time (which was associated with less freeway time) added to utility. Freeway time was statistically insignificant. Gender doesn't seem to matter statistically, nor does commute time

Table 3: Rank Ordered Logit Model of VESP Condition Set #1

	. I			. 2			. 3		
	β	z	P > |z|	β	z	P > |z|	β	z	P > |z|
R/F	4.432	6.21	0.000						
R				0.753	2.58	0.010	0.759	2.29	0.022
F				0.219	1.32	0.187	0.203	1.16	0.247

S*F				-0.813	-1.33	0.182	-0.064	-1.03	0.301
R/T							-0.496	-0.21	0.835
(Obs, groups)	(124,31)			(124,31)			(112,28)		
Log likelihood	-73.833			-72.987			-66.473		
LR chi^2	49.37			51.06			45.02		
Prob > chi^2	0.0000			0.0000			0.0000		

8. Comparison of Methodologies

Contrary to our hypothesis, the CASP and VESP methodologies produce radically different results. To suggest reasons why this occurs, it is necessary to consider the differences between the two methodologies.

One difference in methodologies has to do with recency. The CASP study presents all information (ramp time and freeway time) simultaneously and asks for the subjects to rate, then rank the scenario; the VESP study, following real-world experience, presents the ramp time at the beginning of the freeway journey and then the stop-and-go (or freely flowing) freeway traffic in the following minutes. In general, the congestion occurred in the middle of the drive on the freeway, and occupied more space/time in conditions that were longer in duration. In all cases, the freeway congestion occurred after the ramp delay in all VESP scenarios. If recency matters, then the most recent annoying part of the trip may be viewed more negatively than earlier annoying elements. This might explain the different results. This hypothesis can be tested with future driving simulation studies research in which delays occur at the end of the trip rather than the beginning (such as a traffic light at the top of the exit ramp).

A second difference in methodologies relates to simultaneous versus sequential comparison In the CASP study, the trip was explicitly divided into components that were presented to the subjects simultaneously, allowing direct comparison. Also in the CASP study, it was clear to the subjects that we were interested in their response to the ramp meter wait times. In contrast, the VESP study presented the two components of the trip sequentially, making direct comparison more difficult. Also in the VESP study, it is likely that the subjects considered both the ramp meter wait and the experience of driving, which included varying amounts of free-flowing and stop-and-go traffic.

A related issue is that of awareness of public opinion. Ramp meters have been in the news in the Twin Cities over the past several years; the eight-week shutdown of the ramp meter system in Fall of 2000 was a front-page news story. It has been clear that vocal politicians and elements of the general public do not like ramp meters. At the same time, there has been little public expression of a dislike of stop-and-go traffic. Because of this awareness, the subjects may have responded to the CASP methodology by indicating that they do not like ramp meters either, without giving due thought to what driving in stop-and-go-traffic is really like. There may have thus been some built-in bias that came to the fore with the CASP experiment. In contrast, in the VESP experiment, the subjects actually experienced what it is like to drive in stop-and-go traffic. This methodology makes it more likely that the subjects actually compare the experience of waiting with the experience of stop-and-go traffic.

There is also an issue of the intensity of the stop-and-go traffic – in the simulation scenario, the stop-and-go traffic may be worse than that typically experienced by the subjects. Thus in the CASP, which asks subjects to recall freeway congestion, they may have recalled experiences which are not as severe as those portrayed in the simulator. The VESP may have been a worst-case day for many people; while for the CASP, they recalled typical days, with less intense traffic.

There is another issue. In the real world, driving is a goal-directed activity – the goal is to reach a particular destination. This is especially true of commuting. In the simulator experiment, although the subjects were driving to a particular exit, they were not going there for their own reasons (to get to work, to go shopping, to go for a meal, etc.). As a result, they may not have been particularly concerned about the destination. In the simulator, as in real-life, stop-and-go traffic may be annoying. However in real-life, the fact that you are moving towards a desired destination may outweigh the annoyance. In the simulator the annoyance may take center stage because the movement toward a destination is less important. In contrast, the CASP methodology, which asks about preferences, may evoke those real-life conditions in which the subject is always interested in arriving at his or her destination.

It should be noted that another possible explanation – that an experimental artifact relating to discomfort contributed to the VESP results—can be dismissed. In the current experiment, there was a relatively high incidence of simulator sickness (4 out of 42 subjects felt ill, and were unable to complete the experiment). This suggests that the participants who did complete it might have felt more discomfort than in other experiments. Routinely, in each simulator experiment conducted in the simulator used in this study, a questionnaire pertaining to simulator validity is administered after the experiment is completed. This questionnaire includes a question asking the subjects to rate whether driving in the simulator made them ill (on a scale of 1-7, where a response of 1 reflects that they did not at all feel ill, and 7 reflects that they felt very ill). The responses of the 31 subjects whose data are reported in the current study, were compared with those of the subjects in another recent study (Harder et al. 2003) using the same simulator. In the Harder et al. (2003) study, there were 48 subjects, none of whom had simulator sickness. We compared the distributions of responses to the question about whether driving in the simulator made the subjects feel ill. While not identical, when the Kolmogorov-Smirnov test (Siegel & Castellan 1988) was used to test the differences between the two distributions, the results were very far from statistical significance (even using a very liberal <0.10 level of significance). So, there is no evidence to support the suggestion that the participants who completed the current study might have felt more discomfort than in other experiments, and it is unlikely that discomfort contributed to the VESP results.

So, which methodology is more valid: CASP or VESP? Both methodologies involve subjective assessments, but CASP judgments are made at some temporal distance from the experience to which they pertain, while VESP judgments are made much closer to the experience. Stated preferences vary depending on how you ask about them. A third experiment, Field

Experienced Stated Preference, which puts people on the road rather than in a simulator, may be necessary to resolve this issue. Obviously these questions call for further research.

9. Conclusions

This paper describes experiments comparing traditional computer administered stated preference with virtual experience stated preference to ascertain how people value time spent waiting at ramp meters compared with time spent driving in stop-and-go traffic. The virtual experience stated preference experiments were conducted using an immersive driving simulator. The two methods produced different results, with the traditional computer assisted stated preference suggesting that time spent waiting at ramp meters is 1.6 – 1.7 times more onerous than time spent driving in stop-and-go traffic, while the driving simulator based virtual experience stated preference suggesting the opposite – that time spent driving in stop-and-go traffic is more onerous than time spent waiting at ramp meters. Several reasons were hypothesized to explain the differences, including recency, simultaneous versus sequential comparison, awareness of public opinion, the intensity of the stop-and-go traffic, and the fact that driving in the real-world is a goal directed activity. However without further research, which, if any, of these will eventually prove to be the reason is unclear. What is clear is that the CASP and VESP methodologies produce different results, even though they strive to find the same answers in nominally identical sets of conditions. Because people experience the world subjectively, and make decisions based on those subjective experiences, future research should be aimed at understanding the differences between these subjective methodologies.

10. Acknowledgements

This research was funded by the University of Minnesota's Intelligent Transportation Systems Institute under the project "Ramp Meter Delays, Freeway Congestion, and Driver Acceptance". We would like to acknowledge the assistance of Kristin Deutsch, Nova Schuler, Grace Deason, and Ben Chihak who conducted many of the simulator runs. We would also like to acknowledge the assistance of Nic Ward and Curt Olson.

11. References

AutoSim (2002) *Technical Documentation*, AutoSim, Norway.

Bhat, C. 1995. A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research*, 29B(6), 471-483.

Bloomfield, J. R., Bartelme, M. J., Brown, T. L., and Grant, A. R. (1997). A Comparison Between the Sign Recognition and Driving Performance of Elderly Drivers with Multifocal and Monofocal Intraocular Lens Implants When They Drove at Night and in Fog. Paper presented at the *Seventh International Conference on Vision in Vehicles*, Marseilles, France, 14–17 September, 1997.

Bloomfield, J. R., Brown, T. L., Durlak, C. L., & Bartelme, M.J. (1997). *The Effect of Varying the Spacing, Reflectivity and Location of Retroreflective Raised Pavement Markers (RRPMs) on Driving Performance*. Technical Report submitted to the Federal Highway Administration, under FHWA Contract Number DTFH61-94-C-00120.

Bloomfield, J. R., Brown, T. L., Grant, A. R., and Reinach, S. J. (1996). *The Effect of Fog on the Driver's Gap Acceptance when Changing Lanes.* Paper presented at Transportation Research Board National Research Council Symposium on Night Visibility and Driving Behavior, University of Iowa, Iowa, April 28-30, 1996.

Bloomfield, J. R., Grant, A. R., Levitan, L., Cumming, T. L., Maddhi, S., Brown, T. L. and Christensen, J. M. (1998). *Using an Automated Speed, Steering, and Gap Control System and a Collision Warning System When Driving in Fog*. Federal Highway Administration Technical Report No. FHWA-RD-98-050. McLean, VA: Turner-Fairbank Highway Research Center.

Bloomfield, J. R., Weiler, J. M., Grant, A. R., Brown, T. L., Layton, T. A., Grant, P. R., Woodworth, G. G., McKenzie, D. R., Baker, T. W., and Young, M. A. (1999). The Influence of First- and Second generation Anti-Histamines and Alcohol on Driving Performance. In: Hanson, M. A., Lovesey, E. J., & Robertson, S.A. (Editors) *Contemporary Ergonomics* 1999. London: Taylor and Francis, 63-67.

Bloomfield, J.R., Harder, K.A., and Chihak, B. J. (In preparation). Stopping Behavior. Paper accepted for presentation at *Vision in Vehicles X*, Granada, Spain, September, 6-10, 2003.

Calfee, John, Clifford Winston and Randolph Stempski (2001)Econometric Issues In Estimating Consumer Preferences From Stated Preference Data: A Case Study Of The Value Of Automobile Travel Time *The Review of Economics & Statistics* , vol. 83, issue 4, pages 699-707

Carson, Katherine S., (1996) "A Comparison of Multinomial and Rank-Ordered Logit Models of Voter Preferences," April 1996. Discussion Paper in Economics, The Center for Economic Analysis at the Department of Economic, University of Colorado at Boulder.

Colias, J. & C. Salazar-Velasquez (1995) Integrating the Rank-Ordered Logit Model with

Purchase Intent 46-56 *Canadian Journal of Marketing Research*, Volume 14, 1995 Contents

Featherstone, K.A., Bloomfield, J. R., Lang, A. J., Miller-Meeks M. J., Woodworth, G., and Steinert, R. F. (1999). Driving Simulation Study: Bilateral Array Multifocal versus Bilateral AMO Monofocal Intraocular Lenses. *Journal of Cataract and Refractive Surgery*, 25, 1254-1262.

Garber, Nicholas J., and Hoel, Lester A., 1997. *Traffic and Highway Engineering*, PWS Publishing Company.

Harder, K. A., Bloomfield, J.R., & Chihak, B. J. (in press). Using a Head-Up Display in Poor Visibility Conditions I: Investigating Lane Departure Warnings in Two Simulator Experiments. In: Gale, A. G. (Editor) *Vision in Vehicles* 9, Amsterdam, Holland: Elsevier Science Publishers B.V.

Harder, K. A., Bloomfield, J.R., and Chihak, B. J. (2001) "Investigating HUDs in Specialty Vehicles." Paper presented at *45th Annual Human Factors and Ergonomics Society Meeting*, Minneapolis, Minnesota, 8-12 October 2001.

Harder, K. A., Carmody, J. C., & Bloomfield (2002b). *Exploring the Effects of Traffic Calming Strategies on Driver Behavior.* Paper presented at the 13th Annual Center for Transportation Studies (CTS) Transportation Research Conference, St. Paul, Minnesota, May 22–23, 2002.

Harder, K. A., Carmody, J. M., & Bloomfield, J.R. (2002a). The Effect of Centerline Treatments on Driving Performance. Technical Report submitted to Minnesota Department of Transportation, Office of Research and Strategic Services; March 2002.

Harder, K.A. and Bloomfield, J.R. (2002). The Effect of In-Lane Rumble Strips on Stopping Behavior of Attentive Drivers. Paper presented at the 13th Annual Center for Transportation Studies (CTS) Transportation Research Conference, St. Paul, Minnesota, May 22–23, 2002.

Harder, K.A., Bloomfield, J.R., and Chihak, B. J. (2003). The Effectiveness of Auditory Side- and Forward-Collision Avoidance Warnings on Snow Covered Roads in Conditions of Poor Visibility. Draft technical report submitted, under Mn/DOT Agreement No.: 74708—Work Order No.: 156, to Minnesota Department of Transportation, March 2003.

Harder, K.A., Bloomfield, J.R., and Chihak, B. J. (In preparation). The Effect of Increasing the Saliency of a Rural Intersection on Driving Behavior. Paper accepted for presentation at *Vision in Vehicles X*, Granada, Spain, September, 6-10, 2003.

Hennessy, Dwight A. and Wiesenthal, David L. Traffic congestion, driver stress, and driver aggression. *Aggressive Behavior*. Vol 25(6) 1999, 409-423.

Hensher, D. A. (2001). The Valuation of Non-Commuting Travel Time Savings for Urban Car Drivers. World Conference on Transportation Research, Seoul, Republic of Korea.

Hensher, D.A. (2000a). Measurement of the valuation of travel time savings. *Journal of Transport Economics and Policy* (Special Issue in Honor of Michael Beesley)

Hensher, D.A. (2000b) The valuation of commuting travel time savings for car drivers: Evaluating alternative Model Specifications, *Transportation*.

Hensher, David A. (2001b). The sensitivity of the valuation of travel time savings to the specification of unobserved effects. *Transportation Research: Part E* 37, 129-142.

Houston, Mark B., Bettencourt, Lance A., and Wenger, Sutha. The relationship between waiting in a service queue and evaluations of service quality: A field theory perspective. *Psychology & Marketing*. Vol 15(8) Dec 1998, 735-753.

Hui, Michael K. and Zhou, Lianxi. How does waiting duration information influence customers' reactions to waiting for services. *Journal of Applied Social Psychology*. Vol 26(19) Oct 1996, 1702-1717.

Katahira, Hotaka (1988) On the Small Sample Properties of ML Estimators in the Rank-order Logit Analysis: Some Monte Carlo Results, *Marketing Science*, No.32, 1988

Katsikopoulos, Konstantinos V; Duse-Anthony, Yawa; Fisher, Donald L; Duffy, Susan A. The framing of drivers' route choices when travel time information is provided under varying degrees of cognitive load. *Human Factors*. Vol 42(3) Fal 2000, 470-481..

Koop, G. ; D. J. Poirier (1994) Rank-Ordered Logit Models: An Empirical Analysis of Ontario Voter Preferences *Journal of Applied Econometrics* , Vol. 9, No. 4, pp. 369-388

Layton, David, Gardner Brown and Mark Plummer. (1999). Valuing Multiple Programs to Improve Fish Populations. Dept. of Environmental Science and Policy, University of California, Davis, CA

Lee, J. D., McGehee, D. V., & Brown, T. L. (2000). Prior exposure, warning algorithm parameters and driver response to imminent rear-end collisions on a high-fidelity simulator. *Proceedings of the Human Factors and Ergonomics Society*, 3, 316-319.

Levinson, David Zhang, Lei Das, Shantanu Sheikh, Atif (2002) Ramp Meters on Trial: Evidence from the Twin Cities Ramp Meters Shut-off. Presented at 81st Transportation Research Board Conference, Washington DC January 2002 (TRB-02-2167).

Levitan, L. and Bloomfield, J. R. (1998). Human Factors Design of Automated Highway Systems. In: Barfield, W. and Dingus, T. A. (Editors) *Human Factors In Intelligent Transportation Systems*. Hillsdale, N.J.: Lawrence Erlbaum Associates 131–163.

Levitan, L., Golembiewski, G. and Bloomfield, J. R. (1998). Human Factors Issues for Automated Highway Systems. *ITS Journal*, 4, 21-47.

Mackie, P.J., Jara-Diaz S., and Fowkes, A.S., 2001. The value of travel time savings in evaluation. *Transportation Research: Part* E 37, 91-106.

Meyer, Thierry. Subjective importance of goal and reactions to waiting. *Journal of Social Psychology*. Vol 134(6) Dec 1994, 819-827.

Osuna, Edgar E. The psychological cost of waiting. *Journal of Mathematical Psychology*. Vol 29(1) Mar 1985, 82-105.

Ranney, T.A., Simmons, L.A., Boulos, Z., & Macchi, M.M. (1999). Effect of an Afternoon Nap on Nighttime Performance in a Driving Simulator. *Transportation Research Record, 1686*, Paper No. 99-1533, TRB, National Research Council, Washington, D.C., 49-55.

Redmond, Lothlorien S. and Mokhtarian, Patricia L. The positive utility of the commute: modeling ideal commute time and relative desired commute amount. *Transportation*. Dec 1999.

Rizzo, M., McGehee, D.V., Dawson, J., Anderson, S. (2001). Simulated car crashes at intersections in drivers with Alzheimer Disease. *Alzheimer Disease and Associated Disorders*. 15,10-20.

Siegel, S. and Castellan, N.J., Jr (1988) *NonParametric Satistics for the Behavioral Sciences* (2nd Edition), McGraw-Hill Inc: New York, N.Y.

Stata Corporation (2003) Stata Users Manual. Stata Corporation, College Station, Texas.

Stokols, Daniel, Novaco, Raymond, W., Stokols, Jeanette, and Campbell, Joan. Traffic congestion, Type A behavior, and stress. *Journal of Applied Psychology*. Vol 63(4) Aug 1978, 467-480.

Wardman, Mark and Waters, W.G. II, 2001. Advances in the valuation of travel time savings. *Transportation Research: Part E* 37, 85-90.

Weiler, J. M., Bloomfield, J. R., Grant, A. R., Layton, T. A., Brown, T. L., Woodworth, G. G., McKenzie, D. R., and Baker, T. W. (2000). Effects of Fexofenadine, Diphenhyramine, and Alcohol on Driving Performance: A Randomized Controlled Trial in the Iowa Driving Simulator. *Annals of Internal Medicine*, 132, 354-363.

Weirwille, W.W., Wreggit, S.S., Kirn, C.L., Ellsworth, L.A., & Fairbanks, R.J. (1994) Research on Vehicle-Based Driver Status/Performance Monitoring; Development, Validation, and Refinement of Algorithms for detection of Driver Drowsiness. Vehicle Analysis and Simulation Laboratory, Virginia Polytechnic Institute and State University, VPISU Report No. ISE 94-04, NHTSA Report No. DOT HS 808 247, December 1994.

White, Steven M. and Rotton, James. Type of commute, behavioral aftereffects, and cardiovascular activity: A field experiment. *Environment & Behavior*. Vol 30(6) Nov 1998, 763-780.

Appendix A: Figures

Figure 1: Simulator (in preparation)

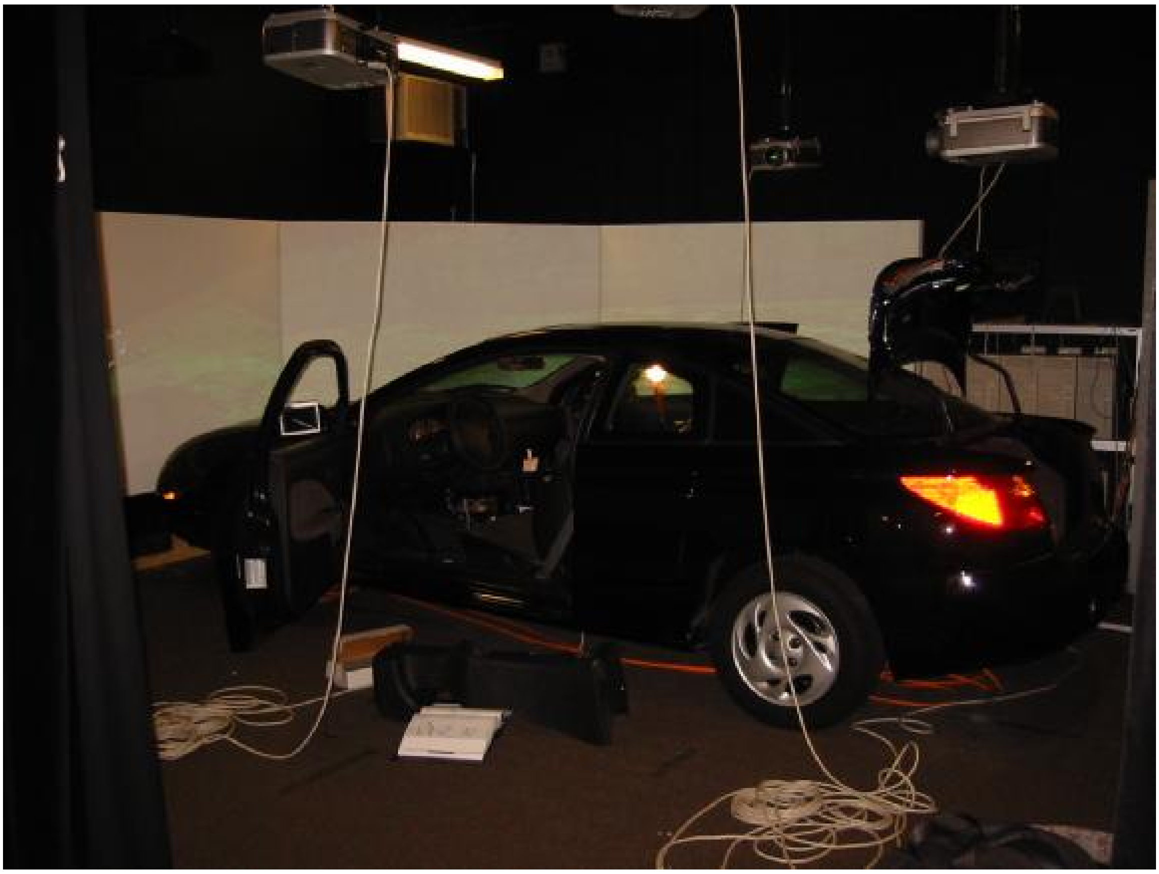

Figure 2: Simulator: Driver's Perspective

Figure 3 CASP Results: Ranks

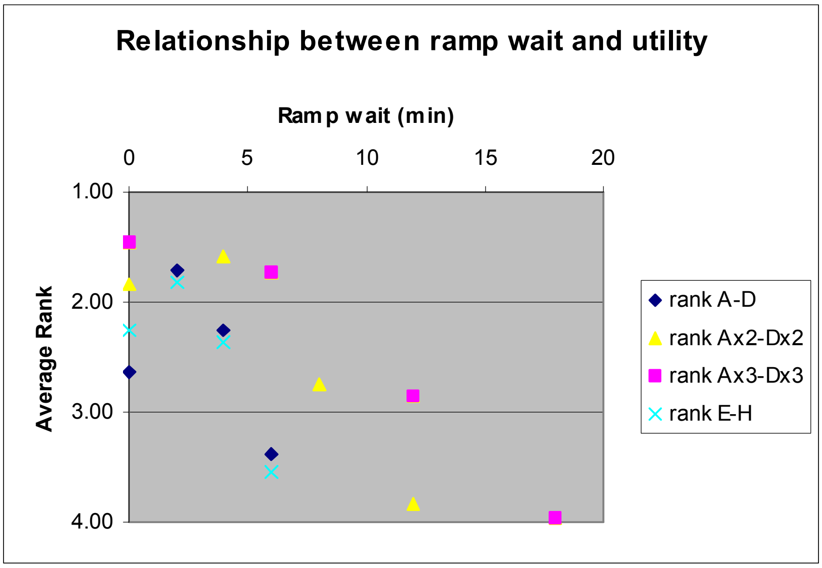

Figure 4 CASP Results (Ratings)