

Weighting Waiting

Evaluating Perception of In-Vehicle Travel Time Under Moving and Stopped Conditions

David Levinson, Kathleen Harder, John Bloomfield, and Kasia Winiarczyk*

Experiments are described in which traditional computer-administered stated-preference (SP) data are compared with virtual experience SP data to ascertain how people value stopped delay compared with stop-and-go or free-flow traffic. The virtual experience SP experiments were conducted by using a wraparound driving simulator. The two methods produced different results: the traditional computer-assisted SP data suggested that ramp delay is 1.6 to 1.7 times more onerous than delay on freeways, whereas the virtual experience SP data based on the driving simulator suggested that freeway delay is more onerous than ramp delay. Several factors are advanced to explain the differences, including recency, simultaneous versus sequential comparison, awareness of public opinion, intensity of the stop-and-go traffic, and the goal-directed nature of driving in the real world. However, without further research, it is unclear which, if any, of these factors will eventually prove to be the right one. What is clear is that a comparison of the computer-administered SP data with virtual experience SP data 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 on the basis of those subjective experiences, future research should be aimed at better understanding the differences between these subjective methodologies.

Although a clear majority of drivers would prefer free, unobstructed flow on any route they choose at any time, these conditions are unrealistic in many cities. Attempts to treat congestion by 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. However, this policy has been subject to considerable controversy as to whether it reduces total delay on the system (it seems to), and how it redistributes delay (some wait more so others can proceed unimpeded) (1). If all time were interchangeable, this policy might not be a problem. However, time passes slowly when one is 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 walk-

ing 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, congested travel time (2). This study further distinguishes among the qualitatively dissimilar in-vehicle travel experiences of waiting at a ramp meter, driving in free-flow traffic, and driving in congested traffic, and the driver acceptance of each is estimated.

At the same time, a new methodological approach to studying such questions is developed. Most travel time value experiments use either paper and pencil or a two-dimensional computer screen presentation for the process of eliciting driver preferences. This study also presents the different options to drivers through a state-of-the-art driving simulator (3). The simulation is an immersive driving experience and is much like actual driving.

The results of the simulator run, called a virtual experience stated-preference (VESP) test, 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 under 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, the following questions are asked: 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? Many factors could influence the drivers' preferences. People do not 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 tedious. However, 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 trade-off between spending time at a ramp meter before entering the highway versus experiencing the wait associated with various traffic volumes 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.

*Deceased. D. Levinson, Department of Civil Engineering, and K. Harder and J. Bloomfield, College of Architecture and Landscape Architecture, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN 55455.

Transportation Research Record: Journal of the Transportation Research Board, No. 1898, TRB, National Research Council, Washington, D.C., 2004, pp. 61–68.

Condition Set 1 exposes the subjects to four test conditions. Condition A is the control condition, with a ramp meter wait time of zero minutes and the slowest travel speed of approximately 48 km/h (30 mph). The other three conditions (B, C, and D) have varied ramp meter wait times, from 2 to 6 min, which are paired with faster travel speeds. Long ramp meter wait times, such as 10 min and upward, were not tested because it was 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. [Condition Set 2 was also conducted in the CASP test, which exposes the subjects to four more test conditions (E, F, G, and H), similar to A, B, C, and D but in which total travel time is the same, so that each minute of ramp wait saves exactly a minute of freeway time. However, for brevity, detailed Condition Set 2 results are not presented here.]

In the experiments, the test runs were counterbalanced by the use of a Latin squares design to eliminate possible order and fatigue effects. These experiments provided important information for the timing of ramp meters, and for the first time gave explicit trade-offs between freeway delay and ramp delay, so that rather than simply a minimization of total travel time on the freeway system, recommendations could be made for a utility-based approach that minimizes total perceived travel time. This approach should increase user satisfaction with the ramp meter system.

A discussion of the value and perception of time is presented first, followed by a review of survey and experimental methods. Then the methodology is described in detail, including the selection of subjects, survey methodology, and the driving simulator. The survey administration is also discussed, including a presentation of the relevant choice conditions. Hypotheses are stated, and the results, derived from a rank-order logit procedure, are presented. The CASP and VESP tests are compared. Some thoughts about the implications of this research for the future of SP analysis conclude the presentation.

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, the wait at 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, although generally disliking to wait for services, accept waiting under certain conditions (4). 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. (4) 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 the ramp) with the positive utility of the goal (e.g., better-flowing traffic).

There is evidence (5) that information on waiting duration has a positive effect on customer service quality judgments, perhaps because subjects perceive greater control of the situation. There is also some evidence that time already spent waiting in a queue (a sunk cost) is less negatively perceived by the customer when the goal of the waiting is considered to be more attractive (6). The use of variable message signs at ramp meters might be added to future studies to see to what degree information ameliorates the wait at ramp meters. Further, education may matter. People may be more

inclined to view the wait at the ramp meter positively if they believe that ramp metering is 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 because of the sense of a waste of time and the uncertainty of the duration of the wait. Theoretical models (7) 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 number 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's engaging in other activities (talking on a cell phone) and by distractions (listening to the radio or books on tape). For drivers who engage in these boredom-reducing strategies, a ramp meter could provide a relatively pleasant environment. In contrast, the uncertainty stress cannot be alleviated when one is 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-congestion conditions than in low-congestion conditions (8). Annoyance is greater among those on a high- and medium-impedance commute than among those experiencing the low-impedance commute (9).

Drivers may have an a priori perception that traffic managed by ramp meters is less variable and 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 choice was inconclusive but provides an interesting discussion and a review of the concepts involved (10).

Commuting activity (by car or bus) raises both pulse rate and systolic blood pressure in subjects (11) in the field. Stokols et al. (9) measured blood pressure and heart rate at prearranged stations located in workplace parking lots immediately after the subjects left their cars following their commute and found that increases in commute distance and time account for a significant proportion of the increase in driver blood pressure.

The commute itself can offer some positive utility to the commuter (12). Some people do report that their commute is too short, and those who 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 affected by stop-and-go freeway conditions, when more attention needs to be focused on the road. And it should be positively affected 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, different results might be expected in a CASP test (in which people might imagine doing something else) than in the VESP simulator experiment, in which the drivers are constrained from other activities (radio, work, etc.).

Revealed preference (RP) and stated preference (SP) data have been used in studies that assess the monetary equivalent of the value of travel time savings (VTTS) and the relative importance of different components of travel time. The models used include the

multinomial logit (MNL) model, the nested logit model, and the mixed logit–random parameter logit (ML/RPL) models. A body of recent research suggests that the popular MNL model tends to underestimate the mean VTTS. This effect has been found in studies of urban commuting, long-distance intercity travel, and urban noncommuting travel (2, 13–15).

An experiment conducted by Hensher (13) distinguishes between the value of considering traffic conditions like free-flow time, slowed-down time, and start-and-stop time rather than the typical grouping of in-vehicle time and out-of-vehicle time. His is a CASP experiment in which the choice is made between the attributes of the current trip and two 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-and-start travel time, uncertainty of travel time, running cost, and toll charges. That research develops a series of models (MNL and ML/RPL) to identify the role of each trip attribute and finds that the VTTS for start-and-stop time is approximately 5 to 10 times that for free flow, whereas the VTTS for slow traffic is 2 to 3 times that for free flow.

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

SURVEY METHODS

Transport models based on observed choices made by individual travelers rather than on observed relationships for groups of travelers are known as disaggregate demand models (DMs), which are probabilistic. The theoretical base for DMs is random utility theory. DMs 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 that individuals will choose 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 1. 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 MNL model.

There are many practical difficulties in measuring the value of time in a direct manner; therefore, indirect methods, such as collecting RP data, have been developed and used. RP research is based on observations of actual choices made by individuals. However, questions have been raised about the validity of the values derived from RP 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 RP data can make the statistical estimation of individual factors involved in choice inaccurate. Also, it can be difficult to find real scenarios that have all of the attributes desirable for a particular study.

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 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 one 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 multidimensional, whereas the SP universe has only a few dimensions. It is legitimate to question whether something is lost in translation, despite the superior control.

In the current study, an immersive driving simulation was used to develop a VESP test. Unlike traditional SP tests, it is not simply a hypothetical scenario that is involved; rather the subject experiences (and evaluates) a virtual simulation of that scenario. The authors know of no prior study of ramp metering that has been conducted in such a manner, although immersive driving simulators have been used to investigate a wide range of issues.

METHODOLOGY

Subject Selection

Similar procedures were used to obtain subjects for the CASP and VESP experiments. A list of all University of Minnesota employees, exclusive of faculty, was obtained from the Office of Institutional Research and Reporting. The list contained a total of 15,288 employee names at all university branches. Further, 5,404 employees located outside Minneapolis, 19 Civil Engineering Department employees, and 100 individuals who had received an email for this study’s pilot were excluded, leaving 9,765 names.

For the CASP experiment, 1,308 names were randomly selected from this pool and emails were sent to them; 209 replies were received. From this group, self-identified nondriver commuters were excluded. Ultimately, 89 subjects were scheduled in groups of five at 1-h intervals over a 4-day period. There were 15 no-shows, but the remaining 74 subjects participated in the study and completed the survey. Of these 74, 5 more subjects were eliminated from the analysis because while in the testing facility they said they commuted primarily by bus, bike, or walking rather than by driving a motor vehicle. Thus, 69 subjects gave valid observations. Compensation of \$10 was given to the subjects after the successful completion of the assignment. The subjects had been notified in the screening process of the level of compensation.

For the VESP experiments, from the list of 9,765 potential subjects, the 1,308 emails already sent were subtracted, with 8,457 potential subjects left. 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, self-identified nondriver commuters were excluded, and subjects likely to be prone to simulator sickness were identified by being asked a standard battery of questions. After accounting for no-shows, there were 42 participants.

Descriptive statistics of the samples are shown in Table 1. The samples were similar; the VESP sample was somewhat (but not significantly) younger with a slightly larger household size. The commuting characteristics were similar. Because the CASP sample was disproportionately female, gender was used as an explanatory variable in the models.

TABLE 1 Descriptive Statistics

		CASP		VESP	
		Female	Male	Female	Male
	Number	41	21	16	19
Age (Years)	Mean	41.4	45.1	37.7	35.3
	Std. Dev.	11.3	10.9	10.8	8.7
Auto Ownership Number	Mean	2.3	2.3	2.2	2.3
	Std. Dev.	1.2	1.3	0.8	1.2
Commute Time (Minutes)	Mean	19.6	15.5	14.0	19.9
	Std. Dev.	27.3	11.2	10.4	13.1
Housing Type	Single-Family	73%	81%	63%	79%
	Multi-Family	27%	19%	38%	21%
Household Size	Mean	1.3	1.2	2.4	2.8
	Std. Dev.	0.4	0.4	1.0	1.4

NOTE: Subjects with incomplete information are excluded from descriptive statistics. Sample includes VESP subjects who did not complete simulator run due to sickness or simulator failure.

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. VESP subjects received \$50 for their participation.

Computer-Administered Survey

Subjects were mailed a travel diary and asked to complete it on 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, which included a demographic-socioeconomic questionnaire, a transportation attitude survey, and a personality survey (the results of the surveys are not reported here). In the CASP experiment, subjects also completed ramp-meter-versus-highway comparison ratings and rankings. In the VESP experiment the ratings and rankings were asked of the subjects following their 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 to 20 min for the subjects to complete the survey.

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.

Driving Simulator

For the VESP experiments, each subject drove in the advanced driving simulator at the University of Minnesota 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, which 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 on the right, and the other on the left—set at 138 degrees to the intermediate panels. All five flat-panel screens are elevated 16 in. from the ground. Five projectors are used to project a coordinated, high-fidelity, virtual environment onto the five flat panels that make up 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-in. 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 from the steering wheel, transmission, and accelerator and brake pedals to the driving simulator computer. This input 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 three-dimensional sound generation system. Low-frequency sounds are delivered through a 10-in. 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 through a dedicated intercom system that uses four speakers 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. Servomotors 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 microvideo cameras positioned in the cab of the simulator vehicle are used throughout the course of the experimental session to record the subject's face, his or her foot position, and his or her steering wheel responses. 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 watches. There was also no working radio in the vehicle. The radio was not permitted since it had the potential to be a confounding variable (with different travelers listening to different programs or music on different trials). If the study administrators had chosen the radio program, this choice could have had other effects (the driver might not have liked the choice). If the same music was played on each run, it might have become boring. This is an area for future research.

TABLE 2 Choice Conditions

Condition Set 1	Type	Ramp Wait min	Drive Time min	Speed mph (km/hr)
A	Varying	0	20	30 (48)
B	Varying	2	15	40 (64)
C	Varying	4	12	50 (80)
D	Varying	6	10	60 (96)

Choice Conditions

Condition Set 1 consists of the four runs (A, B, C, and D) shown in Table 2. The four conditions in Condition Set 1 involve trips that are of varying duration. All runs were 10 mi (16 km) long.

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

In addition, after all four 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, variations were presented on Condition Set 1 in which the times involved were doubled and tripled. 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.

Hypotheses

The data collected allow the testing of numerous hypotheses. Here the focus is on two of those.

First, concerning methodology, the hypothesis was that subjects in the CASP experiment would respond similarly to subjects in the VESP experiment. If the results are similar, one can be confident in doing this type of research with either traditional CASP experiments or VESP simulations. Because computer-based experiments are likely to remain less expensive, similar results would suggest that simulators are not needed to address this question. However, if the results differ, further research may be needed 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 versus freeway travel at various speeds) on driver preference. It was posited that ramp time is more onerous than freeway travel because there is a trade-off between ramp time and freeway time, and previous research suggests that waiting is less preferred to the impression of making progress.

Rank-Order Logit Model

To test the models, a multinomial rank-order logit model was estimate by using Stata (16). Rank-ordered logit has been used in a number of previous studies (17–23).

The rank-order logit model is referred to as the Plackett–Lucre model in the statistics literature and as the exploded logit model in marketing. Each alternative in a set is assigned a real-valued positive parameter γ . For a group of m alternatives, the probability of ranking the choices from best to worst, w , is given by

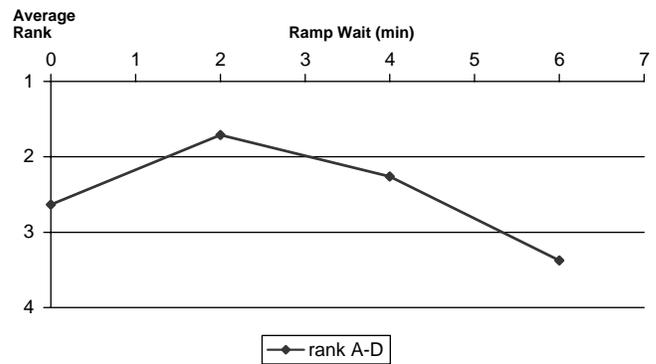


FIGURE 1 CASP results: relationship between ramp wait and rank.

$$P[W = w|\gamma] = \prod_{i=1}^m \frac{\gamma_{w_i}}{\gamma_{w_i} + \gamma_{w_{i+1}} + \dots + \gamma_{w_k}}$$

where w_i is the i th alternative in the ranking.

Given that choice i is preferred—that is, has been ranked first—the second choice is then the most preferred option among the remaining alternatives. Although the rank-order logit and traditional logit models are slightly different, the models share their density function and log-likelihood function.

The results of the rank-ordered logit model were compared with those of a binomial logit model (pairwise comparison of A versus B, A versus C, A versus D, B versus C, B versus D, and C versus D) as separate observations and similar results—which are not shown, for brevity—were reached.

RESULTS

CASP Experiments

Figures 1 and 2 show the rankings and ratings from the CASP experiments. By inspection, it would appear that some ramp wait time 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.

After observations with missing data were dropped, and with the Condition Set 1 data, there were 4 ranks each from 48 individuals

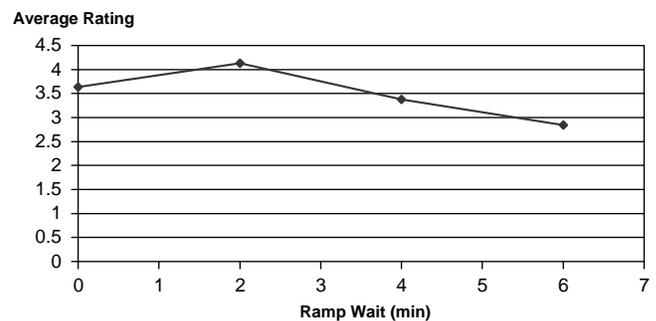


FIGURE 2 CASP results: relationship between ramp wait and rating.

TABLE 3 Rank-Ordered Logit Model of CASP Condition Set 1

	Model 0			Model 1			Model 2			Model 3		
	β	z	P> z	β	z	P> z	β	z	P> z	β	z	P> z
R+F	0.0178	0.33	0.74									
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)			(192,48)		
Log likelihood	-152.49			-144.141			-125.406			-123.989		
LR chi ²	0.11			16.81			54.28			57.12		
Prob > chi ²	0.74			0.00			0.00			0.00		

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 gender (1 if male, 0 if female), and
- T = subject's reported daily commute time (min).

The results are presented in Table 3.

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 the hypothesis, the results indicate that each minute of ramp wait is 1.6 times (Model 1) to 1.7 times (Models 2 and 3) 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 the ramp wait. This observation is logical, since ramp wait becomes a larger percentage of the trip as the trip shortens. 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, since one might anticipate threshold effects in the data (4 min of ramp delay is more than twice as bad as 2 min of ramp delay). These results are substantively consistent with results obtained from other statistical methods (binomial logit on each choice pair).

VESP Experiments

Similar models were estimated with the VESP data, again with rank-ordered logit (Table 4). In contrast to the CASP data, simulator users almost uniformly preferred the ramp wait rather than stop-and-go traffic on the freeway. Thus, the signs in the VESP data are opposite to those obtained in the CASP experiment. This finding is worth noting since economic theory predicts that travel time is a cost to be minimized, so the signs in a utility expression should be negative, which is what was found in the CASP data. Model 0, which estimates coefficients on the sum of ramp and freeway time, comes out as predicted. However, when the times are disentangled, as in Models 2 and 3, both coefficients have a positive sign, though freeway time is insignificant. The positive time for ramp wait is probably due to the correlation with the quality of the trip, in which higher total travel times (due to added ramp wait) are associated with lower freeway times and, more important, less stop-and-go conditions. These

TABLE 4 Rank-Ordered Logit Model of VESP Condition Set 1

	Model 0			Model 1			Model 2			Model 3		
	β	z	P> z	β	z	P> z	β	z	P> z	β	z	P> z
R+F	-0.416	-4.64	0.00									
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)			(124,31)			(112,28)		
Log likelihood	-84.029			-73.833			-72.987			-66.473		
LR chi ²	28.98			49.37			51.06			45.02		
Prob > chi ²	0.0000			0.0000			0.0000			0.0000		

differences are discussed further in the section on comparison of methodologies.

In the CASP experiments, ramp time and freeway time were both onerous, but ramp time was more onerous than freeway time. In the VESP experiments, 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. A significantly better model could not be obtained from the VESP data. For comparison purposes, the same models were desired from both experiments.

Comparison of Methodologies

Contrary to the 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 the subjects to rate and 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 and 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, the most recent annoying part of the trip may be viewed more negatively than earlier annoying elements, which might explain the different results. This hypothesis can be tested with future driving simulation studies in which delays occur at the end of the trip rather than at 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 the researchers were interested in their response to the ramp meter wait times. In contrast, the VESP study's presentation of the two components of the trip sequentially made direct comparison more difficult. Also in the VESP study, the subjects probably 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 8-week shutdown of the ramp meter system in fall 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 too dislike ramp meters without giving due thought to what driving in stop-and-go traffic is really like. Some built-in bias might have come to the fore with the CASP experiment. In contrast, in the VESP experiment, the subjects actually experienced what it was 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 exper-

iment, which asked subjects to recall freeway congestion, they may have recalled experiences that were not as severe as those portrayed in the simulator. The VESP experiment may have been a worst-case day for many people, whereas for the CASP experiment, they recalled typical days, with less-intense traffic.

There is yet another issue. In the real world, driving is a goal-directed activity: the goal is to reach a particular destination. This attribute 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 one is moving toward 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 the destination.

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 situation suggests that the participants who did complete the experiment 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 to 7, where a response of 1 reflects that they did not feel ill at all 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 (24) using the same simulator. In that study, there were 48 subjects, none with simulator sickness. The distributions of responses to the question about whether driving in the simulator made the subjects feel ill were compared. Although the differences were not identical, when the Kolmogorov-Smirnov test (25) was used to test the differences between the two distributions the results were far from statistical significance (even using a liberal <0.10 level of significance). 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.

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, whereas VESP judgments are made much closer to the experience. Stated preferences vary, depending on how one asks 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. These questions call for further research.

CONCLUSIONS

A traditional computer-administered SP (CASP) experiment was compared with a virtual experience SP (VESP) experiment to ascertain how people value time spent waiting at ramp meters compared with time spent driving in stop-and-go traffic. The VESP experiments were conducted with an immersive driving simulator. The

two methods produced different results, with the traditional CASP results suggesting that time spent waiting at ramp meters is 1.6 to 1.7 times more onerous than time spent driving in stop-and-go traffic, whereas the driving-simulator-based VESP results suggested the opposite—time spent driving in stop-and-go traffic is more onerous than time spent waiting at ramp meters.

There is a growing body of evidence that the design of SP experiments can have a significant influence on their outcome. Apart from the obvious differences of environment and context, the researchers attempted to ensure that the questions were asking the same thing. They believe in this case that it is the environment and the context that affect the results.

Several reasons were hypothesized to explain the different results, including recency, simultaneous versus sequential comparison, awareness of public opinion, the intensity of the stop-and-go traffic, and the goal-directed nature of driving in the real world. However, without further research, it is unclear which if any of these reasons will eventually prove to be the right one. 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. Future research should also consider the role of expected versus unexpected ramp delay and perceived versus actual travel time, the role of safety and accidents in ramp-metering schemes, and the effect of allowing radios, CD players, or books on tape in the simulator. Perhaps the most promising direction would be more on-road experiments rather than simulator experiments. Although some control of the conditions are lost, the former are likely to be much more realistic.

ACKNOWLEDGMENTS

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." The authors acknowledge the assistance of Kristin Deutsch, Nova Schuler, Grace Deason, and Ben Chihak, who conducted many of the simulator runs. They also acknowledge the assistance of Kathy Carlsson, Nic Ward, and Curt Olson.

REFERENCES

1. Levinson, D., and L. Zhang. Ramp Meters. In *Assessing the Benefits and Costs of Intelligent Transportation Systems* (D. Gillen and D. Levinson, eds.), Kluwer Academic Publishers, Boston, Mass., 2002.
2. Hensher, D. A. The Sensitivity of the Valuation of Travel Time Savings to the Specification of Unobserved Effects. *Transportation Research*, Vol. 37E, 2001, pp. 129–142.
3. *Technical Documentation*. AutoSim, Tromsø, Norway, 2002.
4. Houston, M. B., L. A. Bettencourt, and S. Wenger. The Relationship Between Waiting in a Service Queue and Evaluations of Service Qual-

- ity: A Field Theory Perspective. *Psychology and Marketing*, Vol. 15, No. 8, Dec. 1998, pp. 735–753.
5. Hui, M. K., and L. Zhou. How Does Waiting Duration Information Influence Customers' Reactions to Waiting for Services? *Journal of Applied Social Psychology*, Vol. 26, No. 19, Oct. 1996, pp. 1702–1717.
 6. Meyer, T. Subjective Importance of Goal and Reactions to Waiting. *Journal of Social Psychology*, Vol. 134, No. 6, Dec. 1994, pp. 819–827.
 7. Osuna, E. E. The Psychological Cost of Waiting. *Journal of Mathematical Psychology*, Vol. 29, No. 1, March 1985, pp. 82–105.
 8. Hennessy, D. A., and D. L. Wiesenhal. Traffic Congestion, Driver Stress, and Driver Aggression. *Aggressive Behavior*, Vol. 25, No. 6, 1999, pp. 409–423.
 9. Stokols, D., R. W. Novaco, J. Stokols, and J. Campbell. Traffic Congestion, Type A Behavior, and Stress. *Journal of Applied Psychology*, Vol. 63, No. 4, Aug. 1978, pp. 467–480.
 10. Katsikopoulos, K. V., Y. Duse-Anthony, D. L. Fisher, and S. A. Duffy. The Framing of Drivers' Route Choices When Travel Time Information Is Provided Under Varying Degrees of Cognitive Load. *Human Factors*, Vol. 42, No. 3, 2000, pp. 470–481.
 11. White, S. M., and J. Rotton. Type of Commute, Behavioral Aftereffects, and Cardiovascular Activity: A Field Experiment. *Environment and Behavior*, Vol. 30, No. 6, Nov. 1998, pp. 763–780.
 12. Redmond, L. S., and P. L. Mokhtarian. The Positive Utility of the Commute: Modeling Ideal Commute Time and Relative Desired Commute Amount. *Transportation*, Dec. 1999.
 13. Hensher, D. A. Measurement of the Valuation of Travel Time Savings. *Journal of Transport Economics and Policy*, Vol. 35, No. 1, 2000, pp. 71–98.
 14. Hensher, D. A. The Valuation of Commuting Travel Time Savings for Car Drivers: Evaluating Alternative Model Specifications. *Transportation*, 2000.
 15. Bhat, C. A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice. *Transportation Research*, Vol. 29B, No. 6, 1995, pp. 471–483.
 16. *Stata Users Manual*. Stata Corporation, College Station, Tex., 2003.
 17. Katahira, H. On the Small Sample Properties of ML Estimators in the Rank-Order Logit Analysis: Some Monte Carlo Results. *Marketing Science*, No. 32, 1988.
 18. Christakis, N. A., and P. D. Allison. Logit Models for Sets of Ranked Items. *Sociological Methodology*, Vol. 24, 1994, pp. 199–228.
 19. Koop, G., and D. J. Poirier. Rank-Ordered Logit Models: An Empirical Analysis of Ontario Voter Preferences. *Journal of Applied Econometrics*, Vol. 9, No. 4, 1994, pp. 369–388.
 20. Colias, J., and C. Salazar-Velasquez. Integrating the Rank-Ordered Logit Model with Purchase Intent. *Canadian Journal of Marketing Research*, Vol. 14, 1995, pp. 46–56.
 21. Carson, K. S. A Comparison of Multinomial and Rank-Ordered Logit Models of Voter Preferences. Discussion Paper in Economics. Center for Economic Analysis, Department of Economics, University of Colorado, Boulder, April 1996.
 22. Layton, D., G. Brown, and M. Plummer. *Valuing Multiple Programs to Improve Fish Populations*. Department of Environmental Science and Policy, University of California, Davis, 1999.
 23. Calfee, J., C. Winston, and R. Stempski. Econometric Issues in Estimating Consumer Preferences from Stated Preference Data: A Case Study of the Value of Automobile Travel Time. *Review of Economics and Statistics*, Vol. 83, Issue 4, 2001, pp. 699–707.
 24. Harder, K. A., J. R. Bloomfield, and B. J. Chihak. *The Effectiveness of Auditory Side- and Forward-Collision Avoidance Warnings on Snow Covered Roads in Conditions of Poor Visibility*. Draft Technical Report. Minnesota Department of Transportation, St. Paul, March 2003.
 25. Siegel, S., and N. J. Castellan, Jr. *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. McGraw-Hill, New York, 1988.

Publication of this paper sponsored by Passenger Travel Demand Forecasting Committee.