CENTER FOR TRANSPORTATION STUDIES

ITS INSTITUTE

Ramp Meter Delays, Freeway Congestion, and Driver Acceptance

Final Report

Prepared by:
Dr. David Levinson
Dr. Kathleen A. Harder
Dr. John R. Bloomfield
Kasia Winiarczyk

Department of Civil Engineering
University of Minnesota

Center for Human Factors Systems Research and Design
College of Architecture and Landscape Architecture
University of Minnesota

Intelligent Transportation Systems Institute
University of Minnesota

CTS 05-02

HUMAN-CENTERED TECHNOLOGY TO ENHANCE SAFETY AND MOBILITY

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.
CTS 05-02		
4. Title and Subtitle **Ramp Meter Delays, Freeway Congestion, and Driver Acceptance**		5. Report Date May 2005
		6.
7. Author(s) David Levinson, Kathleen Harder, John Bloomfield, Kasia Winiarczyk		8. Performing Organization Report No.
9. Performing Organization Name and Address University of Minnesota Department of Civil Engineering 500 Pillsbury Drive S.E. Minneapolis, MN 55455		10. Project/Task/Work Unit No. CTS Project Number: 2002018
		11. Contract (C) or Grant (G) No.
12. Sponsoring Organization Name and Address Intelligent Transportation Systems Institute Center for Transportation Studies University of Minnesota 511 Washington Avenue, SE Suite 200 Minneapolis, MN 55455		13. Type of Report and Period Covered Final Report
		14. Sponsoring Agency Code
15. Supplementary Notes http://www.cts.umn.edu/pdf/CTS-05-02.pdf		

16. Abstract (Limit: 200 words)

In the current study, we conducted several experiments using both the CASP and VESP methodologies. Nominally, the same combinations of ramp meter waiting time and freeway travel time were tested in the first two parts of the CASP experiment (CASP-a and CASP-b) and in the first two VESP experiments (VESP Experiment #1 and VESP Experiment #2). The combinations of time spent waiting at ramp meters and driving on the freeway that were presented in CASP-a were the same as the combinations of desired ramp meter waiting and driving times for VESP Experiment #1. Similarly, the combinations of times presented in CASP-b were the same as the combined desired times for VESP Experiment #2. However, it should be noted that there was some variation in the actual driving times from the desired times in the VESP experiments. This variation occurred because in the VESP experiments the driving time was manipulated by varying the congestion level of the traffic in which the participants drove.

17. Document Analysis/Descriptors		18. Availability Statement
Ramp Meter Congestion Freeway	Travel Time Driver	No restrictions. Document available from: National Technical Information Services, Springfield, Virginia 22161

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	47	

Ramp Meter Delays, Freeway Congestion, and Driver Acceptance

Final Report

Prepared by:
Dr. David Levinson
Dr. Kathleen A. Harder
Dr. John R. Bloomfield
Kasia Winiarczyk

Department of Civil Engineering
University of Minnesota

Center for Human Factors Systems Research and Design
College of Architecture and Landscape Architecture
University of Minnesota

May 2005

Intelligent Transportation Systems Institute
University of Minnesota

CTS 05-02

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 the Office of Institutional Research and Reporting at the University of Minnesota in recruiting subjects. We would also 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 Kathy Carlson and Curt Olson.

Table of Contents

List of Tables

EXECUTIVE SUMMARY

The objective of this research project was to quantify the weights that drivers give to qualitatively different aspects of their travel time experience. Differences in travel time experience were produced by presenting different combinations of (a) the waiting time at ramp meters, and (b) the time drivers travel on freeways. Quantifying the weights given by drivers to different combinations of ramp meter waiting time and freeway travel time allows us to recommend ramp meter waiting times that correspond to driver perceptions.

Most previous studies exploring the value of travel time have used Stated Preference (SP) methods. With these methods, descriptions of alternative scenarios are presented to subjects who then state their preferences among the hypothetical alternatives (e.g., see Hensher, 2001a). In the current research, we presented alternative scenarios on a computer—i.e., we used a Computer Administered Stated Preference (CASP) methodology.

In addition, we developed a new approach. This approach, Virtual Experience Stated Preference (VESP) methodology, utilized an immersive driving simulator. Drivers experienced the various combinations of ramp meter waiting time and freeway travel time by driving in the simulator and, only then, stating which of the alternative combinations they preferred. Previously, immersive driving simulators have been used to investigate a wide range of issues—but we are not aware of any prior studies investigating ramp meter preferences that have been conducted using a driving simulator.

In the current study, we conducted several experiments using both the CASP and VESP methodologies. Nominally, the same combinations of ramp meter waiting time and freeway travel time were tested in the first two parts of the CASP experiment (CASP-a and CASP-b) and in the first two VESP experiments (VESP Experiment #1 and VESP Experiment #2). The combinations of time spent waiting at ramp meters and driving on the freeway that were presented in CASP-a were the same as the combinations of desired ramp meter waiting and driving times for VESP Experiment #1. Similarly, the combinations of times presented in CASP-b were the same as the combined desired times for VESP Experiment #2. However, it should be noted that there was some variation in the actual driving times from the desired times in the VESP experiments. This variation occurred because in the VESP experiments the driving time was manipulated by varying the congestion level of the traffic in which the participants drove.

For the conditions investigated in both CASP-a and VESP Experiment #1, the driving time decreased as the ramp waiting time increased. In addition, the total trip time was less for each of the three conditions in which the driver had to wait at the ramp meter than the total time for the condition in which the driver had zero ramp meter waiting time.

For the conditions investigated in both CASP-b and VESP Experiment #2, as the driving time decreased the ramp waiting time increased. In CASP-b, the *total* time spent in waiting and driving was the same for each of the four conditions; and in VESP Experiment #2, the *desired total* time spent in waiting and driving was the same for each

of the four conditions (although, in fact, there was a gradual increase in total trip time as the ramp waiting time increased).

In CASP-a and VESP Experiment #1 and in CASP-b and VESP experiment #2, four ramp meter waiting times were investigated (they were 0, 2, 4, and 6 minutes); in all these experiments, as the ramp waiting times increased, the freeway driving times decreased (while the trip *distance* was constant). We also conducted a third VESP experiment. In this experiment, the same four ramp meter waiting times were used again. However, in VESP Experiment #3, the subjects drove in free-flow traffic in all four conditions and the trip distance decreased as the ramp waiting time increased.

With both the CASP and VESP methodologies, the subjects were asked to rate each combination of ramp meter waiting time and driving time individually. They rated each combination immediately after it was presented in the CASP experiment and immediately after experiencing it in the VESP experiments. Then, after all four combinations had been presented or experienced, the subjects were asked to rank them in order of preference. We wanted to explore the hypothesis that the CASP and VESP methodologies would produce similar ratings and rankings when similar combinations of ramp meter waiting times and driving times were used.

A similar recruiting procedure was used to obtain subjects in the CASP and VESP experiments. This report presents data from 69 subjects who participated in CASP-a and CASP-b; 32 subjects who participated in VESP Experiment #1; 32 subjects who participated in VESP Experiment #2; and 25 subjects who participated in VESP Experiment #3.

The rating data (given by the subjects immediately following the presentation of each combination of ramp meter waiting time and freeway travel in the CASP experiment and in the VESP experiments) were consistent with the ranking data (given by subjects after all four combinations were presented) within each of the two methodologies.

However, the CASP methodology and the VESP methodology produced radically different results. For CASP-a, the subjects stated that they preferred the combination of the second shortest ramp meter waiting time and the second longest driving time. In contrast, in VESP Experiment #1, after actually experiencing the wait at the ramp meter and driving on the freeway, the subjects stated that they preferred the combination of the longest ramp meter waiting time and a shorter driving time.

Also when CASP-b and VESP Experiment #2 were compared, contradictory results were found; the combination of the longest ramp meter waiting time and the shortest driving time was the *least preferred* combination in CASP-b and the *most preferred* combination in VESP Experiment #2.

Not surprisingly in VESP Experiment #3, when the subjects drove in free-flow conditions in all four alternatives, the combination of conditions with the shortest ramp meter waiting time was the most preferred.

Travel choice preference models were derived from the preference data obtained in the CASP and VESP Experiments. In addition to the preference data, and the combinations of ramp meter waiting times and driving times used in the experiments, these models took into consideration personality data obtained with the NEO-FFI (Five Factor Inventory), and information obtained from a survey of socioeconomic data and travel diaries. With the CASP-a data, the following variables in the model were significant—changes in freeway travel time, changes in ramp waiting time, extraversion, openness, gender, the average age of the subject's own vehicle, and the subject's daily total travel time. In contrast, for VESP Experiment #1, only two variables in the model were significant—changes in ramp waiting time and gender. With the CASP-b data, the following variables in the model were significant—extraversion, openness, gender, and the average age of the subject's own vehicle—while for VESP Experiment #2, there were no significant variables in the model.

Unfortunately, the CASP methodology did not produce preference data that mirrors the preference data obtained with VESP methodology. The lack of agreement between the preference data obtained with the two methodologies suggests that there are problems with one, or both, methodologies. There are at least five explanations for the difference in the results obtained with the two methodologies. First, there may be a recency effect in the VESP experiments, whereby the subjects in those experiments respond to the most recent annoying part of their trip (driving in congested traffic). Second, while in the CASP study subjects were presented combinations of ramp meter times and freeway driving times, in the VESP experiments the subjects actually experienced the congestion that contributed to the freeway driving times. Third, real-world driving is a goal directed activity and drivers have a specific destination in mind when they start driving—in the VESP experiments they may not have been concerned about their destination and, as a result, could have been more annoyed by the congestion they experienced. Fourth, some subjects in the VESP experiments may have experienced more congestion than they usually do in their daily commute. Fifth, in recent years ramp meters have frequently been unfavorably mentioned in the Twin Cities media, and this may have lead to subjects in the CASP experiment indicating that they do not like ramp meters without considering how they feel about driving in heavily congested traffic.

Although we hypothesized that preference data acquired with an immersive driving simulator would be much more likely to resemble real world preference data than CASP data, the present study suggests that additional research is necessary to understand how experimental stated preference techniques are understood and interpreted by subjects, and therefore by researchers. In particular, more realistic experiments employing real-world conditions (i.e., real cars on real streets) will enable us to corroborate or refute the results from computer-based or driving simulator-based stated preference experiments.

CHAPTER 1. INTRODUCTION

1.1. Introduction

When driving, it is likely that drivers prefer to drive unhindered by other traffic. While this is possible in many cities during off-peak hours, it is often impossible during peak hours. And there are also occasions in which traffic is congested during off-peak hours as well. In an attempt to reduce congestion and produce a smooth flow of traffic on freeways, ramp meters have been implemented in a number of cities in the United States. Ramp meters are traffic signals, installed on freeway entrance ramps that are used to regulate the amount of traffic that enters the freeway. In 1995, ramp meters were in use in 23 metropolitan areas (including Minneapolis and St. Paul, Minnesota), and at that time an additional ten metropolitan areas planned to install ramp meter systems in the next five years (Piotrowicz and Robinson, 1995).

However, the use of ramp metering systems has been subject to considerable controversy as to whether or not they reduce total delay on the system and how they redistribute delay. After investigating the ramp metering system in Minneapolis and St. Paul, Levinson and Zhang (2002) suggested that the addition of ramp meters had reduced total delay on the system, but that this was at the expense of some drivers who now have to wait longer so that other drivers can proceed more easily. This would probably not be a problem if drivers perceived time in the same way irrespective of the way in which they spent it. However, time may be perceived by drivers as passing more slowly when they are stopped and waiting than when they are driving on the freeway. Because of this, although the ramp metering system may reduce total travel time, it may not reduce the drivers' total perceived travel time, or total utility to drivers who weight the waiting time more heavily than they do the time that their vehicles are in motion, even in congested traffic.

Historically, in studies exploring the value of different types of time, the most common comparison 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 (e.g., Hensher, 2001) have included more qualitative descriptions of in-vehicle time in their analyses, such as free-flow traffic travel time or slowed-down congested travel time. This approach is extended further in the research detailed in this report.

Before discussing the current research, the next two subsections of this report briefly describe previous studies. First, the findings obtained in studies of the value of time to travelers are described. Then, the transport models derived from these findings and the survey methods used to discover them are discussed.

1.2. Value of Travel Time

It can be argued that waiting at a ramp meter is analogous to waiting in other consumer-based situations. The wait at a ramp meter could be considered as a queue in which drivers wait to be "served" on the highway, in much the same way that customers wait to be served in a store or restaurant. Consumers, who generally dislike waiting for services, accept waiting under certain conditions; for example, Houston, Bettencourt, and Wenger (1998) found a strong negative correlation between waiting time and the customer's evaluation of the quality of service. Houston *et al*. also developed a field-based theory of how the concepts of consumers' wait time expectations and prior service encounters can be applied to drivers waiting at ramp meters—with the acceptability of waiting being a function of the trade-off between the negative utility of the wait (e.g. time at ramp) and the positive utility of the goal (e.g. better flowing traffic).

When they have to wait for something, people may become annoyed and sometimes stressed, because they have a sense of wasting time and because of uncertainty about the duration of the wait. However, theoretical models (e.g., Osuna, 1985) suggest that stress is alleviated to some extent when people are given information about the expected duration of a waiting period. Also, there is evidence from Hui and Zhou (1996) that providing information about the expected waiting duration has a positive effect on service quality judgments made by customers, perhaps due to the customers perceiving 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 customers when the goal of the waiting is considered to be more attractive (Meyer 1994). The use of Changeable Message Signs (CMS) at ramp meters might be usefully investigated in future studies in order to determine the degree to which wait time information ameliorates the onerousness of waiting at ramp meters. Further, drivers may be more inclined to think about the wait at ramp meters in a more positive way, if they believe the wait is a way of achieving an attractive goal—e.g., being able to drive in conditions closer to free-flow traffic levels once they get on the freeway.

When drivers encounter highly congested traffic conditions they may become more stressed and aggressive than when they are driving in less congested conditions (Hennessy and Wiesenthal, 1999). The annoyance level is likely to be greater for drivers whose commute takes place on highways with high or medium levels of traffic congestion than it is for drivers who commute on highways that usually have low traffic levels (Stokols, Novaco, Stokols, and Campbell, 1978).

Drivers may believe that highways managed by ramp meters may be less congested, and therefore less of a time risk, than unmetered highways that potentially could be congested. This reasoning may lead drivers to prefer the more certain ramp meter condition. The investigation by Katsikopoulos, Duse-Anthony, Fisher, and Duffy (2000) of the effects of average travel time and travel time variability on route choices is inconclusive, but provides an interesting discussion and review of the concepts involved.

White & Rotton (1998) found that commuting activity, in car or bus, raises both pulse rate and systolic blood pressure in commuters. When Stokols *et al.* (1978) measured blood pressure and heart rate at prearranged stations located in workplace parking lots immediately after drivers completed their commute to work, they found that drivers with longer commuting distances and times accounted for a significant proportion of the increase in driver blood pressure.

In contrast, Redmond and Mokhtarian (1999) demonstrated that the commute can offer positive utility to some commuters. Some people reported their commute is too short. And those who reported their commute was too long did not want to completely eliminate the commute—instead they just wanted to make it shorter. The positive utility of the commute may in large part be because it acts as a buffer between work and home.

1.3. Transport Models and Survey Methodologies

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. These models are probabilistic—the theoretical base for them is random utility theory. Demand models are typically discrete-choice, logit-type models. In a discrete choice model, an individual selects an option from a finite set of alternatives, and the probability of the individual choosing a particular option is a function of the relative attractiveness of the option and characteristics of that 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. 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).

Because there are many practical difficulties in measuring the value drivers give to time directly, indirect methods have been utilized. One indirect method involves collecting revealed preference data. Revealed Preference (RP) research is based on observing the actual choices made by individuals. However, there have been questions 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 RP studies may lead to inaccurate statistical estimates of the individual factors involved in choice. In addition, it can be difficult to find real world scenarios that have all of the attributes desirable for a particular study.

Studies that use the stated preferences of the subjects present an alternative to studies that use RP. In Stated Preference (SP) studies, each participant is presented with a choice between alternatives that contain predetermined levels of specific attributes. Examples of attributes that might be investigated include—running travel costs, travel time costs, or external costs such as tolls. In SP studies, the combinations of levels of the selected attributes are systematically varied in order to reveal the profile of preferences of each subject. SP experiments make it possible to disentangle the independent contributions of individual attributes.

However, SP methods also have methodological problems. They are typically used in choice experiments, which are at best only analogues to the real world. Inevitably, the alternatives presented in SP studies only account for a limited number of the real world's many dimensions. It is legitimate to question whether or not something is lost in translation in SP studies, despite the superior control that they provide.

Both RP and SP methods 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 that have been derived from the data obtained in these studies include the multinomial logit (MNL) model, the nested logit (NL) model, and mixed logit/random parameter logit (ML/RPL) models.

Recent research suggests that the popular MNL model tends to underestimate the mean VTTS. For example, this has been observed in studies, by Hensher (2000a, 2000b, 2001a, and 2001b) and by Bhat (1995), investigating urban commuting, long distance inter city travel, and urban non-commuting travel.

Hensher (2001a) examined the value of traffic conditions like free-flow time, slowed-down time, and start/stop time—rather than exploring the typical comparison of in-vehicle time and out-of-vehicle time. He conducted an experiment using Computer Administered Stated Preference (CASP) methodology. In this study, subjects had to choose between the attributes of the current trip and two alternative scenarios, each defined by the following six attributes: free-flow travel time, slowed-down travel time, stop/start travel time, uncertainty of travel time, running cost, and toll charges. Then, Hensher (2001a) developed a series of models (MNL and ML/RPL) identifying the role of each attribute. He found that VTTS for start/stop conditions is approximately 5-10 times that for free-flow traffic, while VTTS for slow traffic is 2-3 times that for free-flow traffic. This result implies that driving in start/stop conditions is less preferred than traveling in slow traffic, which in turn is less preferred than traveling in free flow conditions. This result is not surprising—it is what common sense suggests would be expected. But, the modeling process adds to this result an estimate of the approximate magnitude of these preferences in relation to each other.

1.4. Current Study

In the research reported here, an attempt was made to quantify the weights that drivers give to qualitatively different experiences of travel time. Differences in the quality of the travel time experiences of drivers were produced by varying the following—

- The waiting time at ramp meters.
- The resultant speeds at which drivers were able to drive on the freeway.

Quantifying the weights given by drivers to different combinations of ramp meter waiting time and freeway travel time will enable us to recommend ramp meter waiting times that correspond to driver perceptions.

As already mentioned, most previous studies exploring the value of travel time have simply described alternative scenarios and then asked drivers to indicate their preferences—e.g., using the CASP method (Hensher, 2001a). In the current research, we used the CASP methodology but, in addition, we developed a new methodological approach. This approach—in which we obtained Virtual Experience Stated Preferences (VESP)—utilized an immersive driving simulator. With this approach, drivers first experienced a virtual simulation of the alternative scenarios and then evaluated them—instead of making choices about hypothetical scenarios, as in studies using traditional stated preference (SP) methodologies. Immersive driving simulators have been used to investigate a wide range of issues, but we are not aware of any prior studies of ramp meter preferences that have been conducted in a simulator.

In the current study, we were able to compare preference data obtained in the more traditional way by presenting the alternative scenarios on a computer screen—i.e., CASP data—with data obtained after the subjects had driven in the simulator—i.e., VESP data.

Nominally the same driving situations were used in the first two parts of the CASP experiment (CASP-a and CASP-b) and in the first two VESP experiments (VESP Experiment #1 and VESP Experiment #2). The combinations of time spent waiting at ramp meters and driving that were investigated in CASP-a were the desired ramp meter waiting and driving times for VESP Experiment #1. Similarly, the combinations of times investigated in CASP-b were the desired times for VESP Experiment #2.

It should be noted that there was some variation in the actual driving times in the VESP experiments from the desired times. The variation occurred because in the VESP experiments the driving time was manipulated by varying the congestion level of the traffic in which the participants drove.

For the conditions investigated in both CASP-a and VESP Experiment #1, the driving time decreased as the ramp waiting time increased; the total time spent waiting and driving was less for the three conditions in which the driver had to wait at the ramp meter than the total time for the condition in which the driver had zero waiting time.

For the conditions investigated in both CASP-b and VESP Experiment #2, the driving time also decreased and the ramp waiting time also increased—however in CASP-b, the *total* time spent in waiting and driving was the same for all four conditions; and in VESP Experiment #2 the desired *total* time spent in waiting and driving was the same for all four conditions (although, in fact, there turned out to be a gradual increase in the total time as the ramp waiting time increased).

With both the CASP and VESP methodologies, the subjects were asked to rate each combination of ramp meter waiting time and driving time individually. Then, after all the combinations had been presented or experienced, the subjects were asked to rank them in order of preference. We wanted to explore the hypothesis that the CASP and VESP methodologies would produce similar ratings and rankings when similar combinations of ramp meter waiting times and driving times were used. Specifically, this meant comparing the ratings and rankings obtained for CASP-a with those obtained in VESP Experiment #1, and the ratings and rankings obtained for CASP-b with those obtained in VESP Experiment #2.

If the results were found to be similar for the two methodologies, then we could be confident that preference data could be obtained by using either traditional CASP methods or the VESP method developed in this project. Because computer-based experiments are likely to remain less expensive, they have been used by many investigators.

However, if the two experiments did not produce similar results, then the findings obtained in previous computer-based studies would need to be reconsidered and it is likely that further research would be needed.

We also conducted an additional VESP experiment. In this third VESP experiment the level of congestion (and therefore driving speed) was not varied across the conditions—instead, the distance driven by the subjects in each condition was reduced as the ramp waiting time increased.

In the CASP experiment and all three VESP experiments, in addition to preference data, we collected the following information about the subjects—

- Basic demographic and socioeconomic data
- One-day travel diary data, giving information about the last weekday before each subject participated in the study.
- Personality data, obtained by administering the NEO Five-Factor Inventory (NEO-FFI) (Costa and McCrae, 1985 and 1991).

When the rating data obtained in the CASP experiment and the three VESP experiments, were examined, we found that for some subjects there were ties between conditions. In these cases, there was little difference in terms of preferences between the alternatives. However, there were no ties when the subjects were asked to rank the alternatives—because the ranking procedure forced the subjects to chose. Preference models based on rankings force subjects to make fine judgments. In contrast, while preference models based on ratings will not break ties, they will allow for a greater spread of responses—because the rating scale ranges from one to seven, while the ranking scale is only from one to four.

CHAPTER 2. METHOD

2.1. Subjects

The same procedure was used to recruit subjects for both the CASP and VESP experiments. At the time the project was conducted, according to the Office of Institutional Research and Reporting, the University of Minnesota had 15,288 employees—exclusive of faculty. We excluded 5,404 employees who were located outside the Minneapolis-St. Paul metropolitan area. In addition, 19 Civil Engineering Department employees were omitted—as were 100 individuals who had received an e-mail inviting them to participate in the pilot study for this study. This left a potential subject pool of 9,765 University of Minnesota employees.

2.1.1. Subjects in the CASP Experiment

The first step in recruiting subjects for the CASP experiment was to select, at random, 1,308 names from the subject pool. E-mails were sent to all 1,308 people, inviting them to participate in the experiment. There were 209 replies. From this group of 209, those who identified themselves as non-driver commuters were excluded. Eighty-nine subjects were scheduled in groups of five. They were tested at one-hour intervals, over a four-day period. There were 15 no-shows. Seventy-four subjects, who were paid $10 for their participation, completed the study. However, data obtained from five subjects were not analyzed—this was because, after their arrival at the testing facility, it was discovered that these five subjects commuted primarily by bus, bike, or on foot, rather than by driving a motor vehicle. This left a total of 69 subjects whose data were analyzed.

2.1.1. Subjects in the VESP Experiments

The 1,308 people who received e-mails inviting them to participate in the CASP experiment were excluded from our subject pool—leaving a pool of 8,457 University of Minnesota employees for the VESP experiments. A similar recruiting procedure to that employed in the CASP experiment was used to recruit subjects—an e-mail was sent to a random sample taken from the subject pool. Again, respondents who do not drive when they commute to the University were excluded. In addition, phone interviews were conducted; a standard battery of questions was used to screen out people who were likely to be prone to simulator sickness.

A total of 42 subjects participated in *VESP Experiment #1*. In spite of the screening for simulator sickness, four subjects (three male and one female) had to stop driving because they exhibited symptoms of simulator sickness. In addition, there were problems with the simulator in six simulator sessions. The data from the four subjects with simulator sickness symptoms and six whose sessions in the simulator were compromised were not analyzed. This left a total of 32 subjects (16 male and 16 female) whose data were analyzed.

Seventy-two subjects took part in *VESP Experiment #2*. Twenty-three of them (nine male and 14 female) had symptoms of simulator sickness, and as a result had to stop driving. Also, there were simulator problems in 17 simulator sessions. We did not analyze the data of the 23 subjects who had symptoms of simulator sickness or the 17 whose sessions in the simulator were compromised. This left a total of 32 subjects (16 male and 16 female) whose data were analyzed.

In *VESP Experiment #3*, a total of 35 subjects participated. Six of them (two male and four female) had symptoms of simulator sickness, and there were problems in six simulator sessions. As in the first and second VESP experiments, we did not analyze the data from the six subjects who had symptoms of simulator sickness or the six whose sessions in the simulator were compromised; this left a total of 25 subjects (15 male and 10 female) whose data were analyzed.

No subject participated in more than one VESP experiment. Each subject received $50.00 for participating in a VESP experiment.

2.2. Experimental Design

In this project, four experiments were conducted—one using the CASP methodology, and three using the VESP methodology. In all four experiments subjects were asked to indicate their preferences between alternative journeys in which there were tradeoffs between the amount of time spent waiting at a ramp meter and the amount of time spent driving on the highway.

In the CASP experiment, longer ramp meter waiting times were paired with shorter travel times, and shorter ramp meter waiting times were paired with longer travel times. This relationship was a little more complicated in the three VESP experiments—in them, longer ramp meter waiting times were paired with freer flowing traffic (and *therefore* shorter travel times), and shorter ramp meter waiting times were paired with more congested traffic (and *therefore* longer travel times).

2.2.1. Design of the CASP Experiment

There were several parts to the CASP Experiment. In the first part of the experiment (which is referred to as CASP-a in the rest of this report), the combinations of ramp waiting times and driving times shown in Table 2.1 were presented to each subject.

Table 2.1: Combinations of Ramp Waiting Time and Driving Time (as well as Total Time) presented in CASP-a

Condition	Ramp Waiting Time	Driving Time	Total Time
cA	0 minutes	20 minutes	20 minutes
cB	2 minutes	15 minutes	17 minutes
cC	4 minutes	12 minutes	16 minutes
cD	6 minutes	10 minutes	16 minutes

As Table 2.1 shows, the ramp wait times for CASP Conditions cA, cB, cC, and cD increased as the driving times decreased. Also, the total travel time dropped from 20 minutes for Condition cA, to 17 minutes for Condition cB, and to 16 minutes for both Conditions cC, and cD.

At first, each subject saw each of the combinations shown in Table 2.1 individually—with the order in which they were presented counterbalanced across subjects. After seeing each combination, the subject was asked to *rate* it on a seven-point category scale, with one indicating that he or she had the least liking for the combination, and seven indicating that he or she had a strong liking for it. Then, after the subject had seen all four combinations individually and rated them, the four alternatives were presented again—but this time they were presented simultaneously, and the subject was asked to *rank* them in order of preference.

In the second part of the CASP Experiment (which is referred to as CASP-b in the rest of this report), the combinations shown in Table 2.2 were presented to each subject. In this case, the ramp wait times remained unchanged from CASP-a, but the driving times were adjusted so that the total travel times for all four combinations were equal (they were all 20 minutes).

Table 2.2: Combinations of Ramp Waiting Time and Driving Time (as well as Total Time) presented in CASP-b

Condition	Ramp Waiting Time	Driving Time	Total Time
cE	0 minutes	20 minutes	20 minutes
cF	2 minutes	18 minutes	20 minutes
cG	4 minutes	16 minutes	20 minutes
cH	6 minutes	14 minutes	20 minutes

Again, each subject saw each of the four combinations shown in Table 2.2 individually—with the order of presentation counterbalanced across subjects. After seeing each combination, the subject was asked to *rate* it on a seven-point category scale, with one indicating that he or she had the least liking for the combination, and seven that he or she had a strong liking for it. Then, when the subject had seen the four combinations separately, the four alternatives were presented simultaneously, and the subject was asked to *rank* them.

Then, each subject saw six additional presentations showing four sets of alternatives that they were asked to *rank*. These alternatives were not presented individually and the subjects were not asked to *rate* them. These alternatives were variations on CASP-a and CASP-b. The order in which they were presented was not counterbalanced across subjects and these data are not reported here.

2.2.2. Design of VESP Experiment #1

We wanted the combinations of ramp meter waiting times and driving times in VESP Experiment #1 to be similar to the combinations presented for CASP-a. However, since we wanted each subject to drive exactly the same distance and we wanted them to drive "as they normally would," we were not able to set the travel time directly. Instead, we varied the travel time indirectly—by varying the level of traffic congestion encountered by the subjects. Table 2.3 gives details of the desired travel times and the average speed at which the subjects would have to drive to achieve it.

Table 2.3: Combinations of Ramp Waiting Time and Desired Driving Time (as well as the Desired Total Time and Desired Average Driving Speed) in VESP Experiment #1

Condition	**Ramp Waiting Time**	**Desired Driving Time**	**Desired Total Time**	**Desired Average Driving Speed**
vA	0 minutes	20 minutes	20 minutes	30 mph (48.3 km/h)
vB	2 minutes	15 minutes	17 minutes	40 mph (64.4 km/h)
vC	4 minutes	12 minutes	16 minutes	50 mph (80.5 km/h)
vD	6 minutes	10 minutes	16 minutes	60 mph (96.6 km/h)

The conditions in VESP Experiment #1—vA, vB, vC, and vD—were intended to correspond to the conditions presented for CASP-a—cA, cB, cC, and cD, respectively. In VESP Experiment #1, the traffic congestion level was reduced systematically from Condition vA to Condition vB, to Condition vC, and then to Condition vD, in which the subjects were essentially driving in free flow conditions. As Table 2.3 shows, in Condition vD the subjects had to wait six minutes at the ramp meter, but then were expected to be able to drive twice as fast as they could in Condition vA and, as a result, reduce the total time (ramp waiting time plus driving time) by four minutes.

2.2.3. Design of VESP Experiment #2

We wanted the combinations of ramp meter waiting times and driving times in VESP Experiment #2 to be similar to the combinations presented for CASP-b. As with VESP Experiment #1, the travel time was varied indirectly—by varying the level of traffic congestion encountered by the subjects. Table 2.4 gives details of the desired travel times and the average speed at which the subjects would have to drive to achieve it.

In VESP Experiment #2, the traffic congestion levels were selected in order that the conditions in the experiment (vE, vF, vG, and vH) would correspond to the conditions presented for CASP-b (cE, cF, cG, and cH, respectively). The traffic congestion levels were reduced systematically from Condition vE to Condition vH. As Table 2.4 shows, in Condition vH the subjects had to wait six minutes at the ramp meter, but then were expected to reduce there drive time by six minutes.

Table 2.4: Combinations of Ramp Waiting Time and Desired Driving Time (as well as the Desired Total Time and Desired Average Driving Speed) in VESP Experiment #2

Condition	Ramp Waiting Time	Desired Driving Time	Desired Total Time	Desired Average Driving Speed
vE	0 minutes	20 minutes	20 minutes	30 mph (48.3 km/h)
vF	2 minutes	18 minutes	20 minutes	33.3 mph (53.6 km/h)
vG	4 minutes	16 minutes	20 minutes	37.5 mph (60.4 km/h)
vH	6 minutes	14 minutes	20 minutes	42.9 mph (69.0 km/h)

Table 2.4 also shows that in VESP Experiment #2 the level of traffic congestion was adjusted in each condition so that the time that the subjects would spend waiting at the ramp meter would be made up in driving—and the total time would be the same in each of the four conditions.

2.2.4. Design of VESP Experiment #3

In VESP Experiment #3, the ramp waiting times in the four conditions were unchanged from the previous experiments (i.e., they were unchanged from CASP-a and CASP-b and from VESP Experiments #1 and #2). There were, however, two changes to the conditions explored in the experiment. First, in all four conditions, the subjects drove in free-flow traffic—so that the average driving speed was expected to be similar in all four conditions. Second, the length of the journey was varied—the distance traveled in each condition increased as the ramp meter waiting time decreased. The combinations of conditions are presented in Table 2.5.

Table 2.5: Combinations of Ramp Waiting Time and Distance Traveled (as well as Expected Average Driving Speed) in VESP Experiment #3

Condition	Ramp Waiting Time	Distance Traveled	Desired Average Speed	Desired Driving Time	Desired Total Time
vI	0 minutes	10 miles	60 mph (96.6 km/h)	10 minutes	10 minutes
vJ	2 minutes	8 miles	60 mph (96.6 km/h)	8 minutes	10 minutes
vK	4 minutes	6 miles	60 mph (96.6 km/h)	6 minutes	10 minutes
vL	6 minutes	4 miles	60 mph (96.6 km/h)	4 minutes	10 minutes

2.2.5. Counterbalancing the Order of Presentation
In the first two parts of the CASP experiment (CASP-a and CASP-b) and all three VESP experiments, the order in which the alternative combinations were presented was counterbalanced across subjects, using a Latin Squares design. The six Latin Squares that were used not only counterbalanced the order, but also ensured that each condition was preceded and followed by each of the other conditions the least possible number of times. By counterbalancing the order of presentation in this way, we minimized carry-over effects, order effects, and fatigue effects.

2.3. The Driving Simulator

An advanced driving simulator was used in VESP Experiments #1, #2, and #3. The key components of this simulator were as follows:

2.3.1. Simulator Vehicle
The simulator vehicle was a full-body 2002 Saturn SC1 coupe.

2.3.2. Simulator Visuals
The driver of the simulator vehicle had a 210-degree forward field-of-view. The 210-degree forward field-of-view was provided by five flat-panel screens—each measuring 4.7-ft (1.43-m) high by 6.5-ft (1.98-m) wide. There was a central flat panel in front of the simulator vehicle, and the center of this panel was aligned with the line-of-sight of the driver of the simulator vehicle. Two intermediate panels flanked this central panel—one on the left and one on the right. Both were set at 138-degrees to the central panel. Then, there were two outer panels—again, one to the right, the other to the left. Again both were set at 138-degrees to the intermediate panels. All five flat-panel screens were elevated 16 in (40.6 cm) from the ground. Five projectors were used to project a coordinated, high-fidelity, virtual environment onto the five flat-panels comprising the 210-degree forward field-of-view. In addition, the simulator provided rear-view imagery, via a 10-ft (3.05-m) high by 7.5-ft (2.29-m) wide screen mounted behind the vehicle that the driver could see through the vehicle's rear-view mirror, and by two 5-in (12.7-cm) LCD screens that were installed in place of the simulator vehicle's side-view mirrors.

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

2.3.4. Simulator Sound System

Road and traffic noise, and the simulator vehicle's engine sounds were delivered through four speakers placed around the vehicle's exterior, near the base of the forward screen. Each speaker received independent inputs from the simulator's 3D sound generation system. Low-frequency sounds were delivered using a 10-in (25.4-cm) subwoofer placed inside the simulator vehicle's engine compartment. Recorded instructions were delivered through the four speakers at the base of the forward screen. In addition, during the experimental session, the experimenter could communicate with each subject via a dedicated intercom system that made use of four speakers installed in the simulator vehicle's factory speaker locations.

2.3.5. Simulator Vehicle Movement

A bass shaker mounted to the underside of the vehicle's frame provided additional low-frequency vibration.

2.3.6. Data Recording

The virtual position of the simulator vehicle, relative to the scenario that each subject drove, was recorded at a rate of 20 Hz. From this record, it was 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 were 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 each experimental session. A video display at the experimenter's station enabled the experimenter to monitor each subject.

There was no working clock in the vehicle. We did not inquire whether or not the subjects had watches. There was no working radio in the vehicle—the radio was not available to the subjects because it had the potential to be a confounding variable. [If different subjects had heard different programs/music on different trials, there may have been a variety of uncontrolled effects on the way in which they drove. And, identical radio program/music had been imposed on the subjects, there could also be problems—some drivers might have enjoyed the selected program/music, others might not.]

2.3.7. Driving Scenario

The subjects drove on a section of freeway that was approximately 10 miles (16.1 km) in length in all three VESP experiments. There were exit ramps at approximately two-mile (3.2 km) intervals. The level of traffic moving in the same direction as the subjects was varied from condition to condition in order to achieve the desired driving times listed in subsections 2.2.2. (for VESP Experiment #1) and 2.2.3. (for VESP Experiment #1). In VESP Experiment #3, the subjects drove in free-flow conditions, but took different exits for each of the four alternative journeys.

2.4. Procedure

A travel diary was mailed to each subject. The subject was asked to complete the diary for the day before he or she was scheduled to participate in either the CASP experiment or one of the three VESP experiments. On arriving at the testing facility, the subject read and signed the consent form, and handed in the travel diary. Any subject who forgot the travel diary was asked to mail it later. After this point, the procedure differed for the CASP and VESP experiments—the following two sub-sections of the report describe these procedures.

2.4.1. Procedure for the CASP Experiment
After signing the consent form and handing in the travel diary, each subject completed a computer-based survey. The survey was administered using a Microsoft Access Database constructed specifically for this study. The database recorded all the subject's responses. The survey included the following

- Demographic/socioeconomic questionnaire.
- Transportation attitude survey.
- Personality survey—the NEO-FFI was administered.

After, completing this three-part survey, the experimental conditions (mentioned in subsection 2.2.1.) were presented and the subject was asked to indicate his or preferences by rating and then by ranking the alternatives. First, the subject used a seven-point category scale to rate his or her preferences for each of the four combinations of ramp meter waiting time and driving times that were presented for CASP-a. Then, after the subject had *rated* each condition separately, the four conditions were presented simultaneously on a single computer screen, and the subject was asked to *rank* them in order of preference.

Next, the subject was asked to use the seven-point category scale to rate his or her preferences among the four combinations of ramp meter waiting time and driving times presented for CASP-b. Then, after *rating* each condition separately, the subject *ranked* the four combinations when they were presented simultaneously on a single computer screen.

As already mentioned in subsection 2.2.1, when CASP-a and CASP-b were completed, each subject saw six additional presentations showing four sets of alternatives that they were asked to *rank*. The subjects were not asked to *rate* them. The order in which these alternatives (variants on CASP-a and CASP-b) were presented was not counterbalanced. These rankings are not reported here.

2.4.2. Procedure for the VESP Experiments
In the VESP experiments, the subjects took the first part of the computer-based survey that was given in the CASP experiment and drove the driving simulator under a set of conditions that were different combinations of ramp waiting time and desired driving time. In each of the three VESP experiments, half of the subjects completed the

computer-based survey first and then drove in the simulator, while the other half of the subjects drove in the simulator first and then completed the survey.

Whether the survey was presented before or after the subject drove in the driving simulator, the order in which the three survey items were presented was the same in all cases. First, the subject was asked to complete the demographic/socioeconomic questionnaire; second, he or she was given the transportation attitude survey; and third, he or she completed the personality survey.

When the subject drove in the driving simulator—whether this was before or after completing the survey—he or she began with a brief practice drive. This was to familiarize the subject with driving the simulator. Then, the subject drove under each of the experimental conditions, after having been asked to "drive as you normally would." After each drive, the subject was asked to use a seven-point category scale to rate how much he or she liked that drive. After driving each of the experimental conditions, and after rating each of them separately, the subject was asked to rank the four drives in order of preference.

2.4.3. Debriefing
For the CASP experiment and all three VESP experiments, after the subject completed the experiment, he or she was debriefed and paid.

CHAPTER 3. RESULTS

The results of the CASP experiment and the three VESP experiments are presented in the following subsections.

3.1. CASP Experiment

3.1.1. CASP-a

The results obtained in the first part of the CASP experiment (CASP-a) are discussed first.

3.1.1.1. CASP-a Rating Data—For CASP-a, the four conditions were presented one at a time to each subject. After each presentation, the subject was asked to *rate* the condition using a seven-point category scale—with seven indicating that the subject had a strong liking for the condition and one that he or she had little liking for it. An analysis of variance (ANOVA) was conducted on the rating data for CASP-a. The results of this analysis are shown in Table 3.1.

Table 3.1: Summary of the ANOVA Comparing the Ratings for CASP-a

Source of Variation	Degrees of Freedom	Sum of Squares	Mean square	F-value	P-Value
Between Subjects	68	339.746	4.996		
Drive Quality Ratings	3	59.779	19.926	11.500	<0.0001
Error Term (Ratings X Subjects)	204	353.471	1.733		

As Table 3.1 shows, there was a statistically significant difference (at the $p<0.0001$ level) in the ratings obtained for the four conditions for CASP-a. The highest mean rating was obtained for Condition cB, and the lowest was for Condition cD.

3.1.1.2. CASP-a Ranking Data—After seeing each of the combinations of ramp meter waiting times and driving times and rating them individually, the four combinations were presented simultaneously on a single screen and the subject was asked to *rank* them in order of preference—with one indicating the condition that he or she preferred, and four the condition that that was least preferred.

When a Friedman Two-Way ANOVA by Ranks was conducted on these data, the resultant value of F_r was 41.32—this is statistically significant at the $p<0.001$ level (the critical value of F_r for three degrees of freedom at this significance level is 16.27).

Condition cB was the most preferred combination, while Condition cD was the least preferred.

3.1.1.3. Relationship between the CASP-a Rating and Ranking Data—If the rating and ranking data were in complete agreement, the condition that was given the highest rating should be ranked number one, the condition given the second highest rating should be ranked second, the condition given the third highest rating should be ranked third, and the condition given the fourth highest rating should be ranked fourth. The average ratings and rankings obtained for CASP-a are summarized in Table 3.2.

Table 3.2: Comparison of Mean Ratings and Mean Rankings for CASP-a

Condition	Mean Rating	Mean Ranking
cA	3.64	2.63
cB	4.13	1.69
cC	3.38	2.28
cD	2.84	3.40

As Table 3.2 shows, there was not complete agreement between the ratings and the rankings. There was agreement for Condition cB, which received the highest mean rating and the best mean ranking, and for Condition cD, which received the lowest mean rating and the worst mean ranking. However, the mean ratings and mean rankings were not in agreement for Conditions cA (which had the second highest mean rating and the third mean ranking) and Condition cC (which had the third highest mean rating and the second mean ranking). While most subjects had consistent ratings and rankings, not all subjects did.

As indicated in Table 3.3, all individuals can be internally consistent in both their ratings and rankings, and yet the mean ratings and rankings may disagree. This is because the mean rating depends upon the strength of the rating.

Table 3.3 Illustration of Internally Consistent, but Globally Inconsistent Ratings and Rankings.

Subject 1		Subject 2		Subject 3		Mean Ranking	Mean Rating
Ranking	Rating	Ranking	Rating	Ranking	Rating		
1	2.2	1	2.2	2	5.1	1.33	3.17
2	2.1	2	2.1	1	7	1.67	3.73
3	2	3	2	3	5	3.00	3.00
4	1	4	1	4	4	4.00	2.00

3.1.2 CASP-b

The results obtained in the second part of the CASP experiment (CASP-b) are discussed here.

3.1.2.1. CASP-b Rating Data—For CASP-b, the second part of the CASP experiment, the four combinations were presented one at a time to each subject. After each presentation, the subject was asked to *rate* the combinations using a seven-point scale. The results of the ANOVA conducted on the rating data are shown in Table 3.4.

Table 3.4: Summary of the ANOVA Comparing the Ratings for CASP-b

Source of Variation	**Degrees of Freedom**	**Sum of Squares**	**Mean square**	**F-value**	**P-Value**
Between Subjects	66	215.903	3.271		
Drive Quality Ratings	3	61.522	20.507	15.619	<0.0001
Error Term (Ratings X Subjects)	198	259.978	1.313		

As Table 3.4 indicates there was a statistically significant difference (at the $p<0.0001$ level) in the ratings obtained for the four combinations presented for CASP-b. The highest mean rating was obtained for Condition cE, with the lowest rating being for Condition cH.

3.1.2.2. CASP-b Ranking Data—When the four combinations of ramp meter waiting time and driving time were presented simultaneously on a single screen for CASP-b, each subject ranked them in order of preference. The Friedman Two-Way ANOVA by Ranks was used to analyze these data. The resultant value of F_r was 70.52. This is statistically significant at the $p<0.001$ level (as mentioned above, the critical value of F_r for three degrees of freedom at this significance level is 16.27). Condition cE was the most preferred combination and Condition cH was the least preferred.

3.1.2.3. Relationship between the CASP-b Rating and Ranking Data—The average ratings and rankings obtained for CASP-b are summarized in Table 3.5.

Table 3.5: Comparison of Mean Ratings and Mean Rankings for CASP-b

Condition	**Mean Rating**	**Mean Ranking**
cE	3.69	2.25
cF	3.46	1.82
cG	3.15	2.37
cH	2.42	3.56

Table 3.5 shows that there was not complete agreement between the ratings and the rankings. Condition cE received the highest mean rating and had the second best ranking. Condition cD received the second highest mean rating and the best ranking. The mean ratings and mean rankings were in agreement for Conditions cG (with the third highest mean rating and the third best mean ranking) and Condition cH (with the worst mean rating and the worst mean ranking).

3.2. VESP Experiments

In the CASP experiment the combinations of ramp meter waiting times and driving times were presented on a computer screen. In contrast, in the VESP experiments, subjects actually *experienced* the combinations—actually waiting at a ramp meter and then driving a vehicle in the simulator for each combination. The order in which the specific combinations were experienced in each of the three VESP experiments was counterbalanced across subjects.

In the VESP experiments, the subjects gave a rating immediately after experiencing each condition. As in the CASP experiment, a seven-point category scale was used—with seven indicating that the subject had a strong liking for the condition and one that he or she had little liking for it.

The results of the three VESP experiments are presented in the following subsections of the report. In each case, before presenting the preference data the actual speeds at which the subjects drove are reported.

3.2.1. VESP Experiment #1

Thirty-two subjects completed the first VESP experiment. The following subsections report the average speed at which they drove for each of the combinations of conditions as well as the rating and ranking data.

3.2.1.1 Driving Speed in VESP Experiment #1— The average speed at which the 32 subjects drove in Conditions vA, vB, vC, and vD in VESP Experiment #1 were as follows—50.58 km/h (with a standard deviation, SD, of 1.26 km/h), 63.15 km/h (SD=2.06 km/h), 80.02 km/h (SD=2.60 km/h), and 94.02 km/h (SD= 6.41 km/h), respectively. These average times (and a range of plus/minus 1.0 SD) are presented, along with the actual driving times and actual total times, in Table 3.6.

Table 3.6: Actual Average Driving Speeds, Average Driving Times and Total Times in VESP Experiment #1

Condition	Actual Average Driving Time (plus/minus one Standard Deviation)	Actual Average Driving Time	Ramp Waiting Time	Actual Average Total Time
vA	50.58 km/h (49.32—52.84 km/h)	19.1 minutes	0 minutes	19.1 minutes
vB	63.15 km/h (61.09—65.21 km/h)	15.3 minutes	2 minutes	17.3 minutes
vC	80.02 km/h (78.42—82.62 km/h)	12.1 minutes	4 minutes	16.1 minutes
cD	94.02 km/h (88.61—100.42 km/h)	10.3 minutes	6 minutes	16.3 minutes

As Table 3.6 shows the total times for Conditions vA, vB, vC, and vD, were 19.1, 17.3, 16.1, and 16.3 minutes, respectively. They were close to the total times desired, which were 20, 17, 16, and 16 minutes, respectively. The actual total time for Condition vA was 0.9 minutes slower than the desired time. The desired total times for Conditions vB, vC, and vD were slightly faster—0.3, 0.1, and 0.3 minutes faster, respectively—than desired driving times.

3.2.1.2. Rating Data in VESP Experiment #1— After experiencing each of the four combinations of ramp meter waiting time and driving time, the subjects were asked to rate the combinations. An ANOVA was used to analyze the rating data—the summary details of this ANOVA are presented in Table 3.7.

Table 3.7: Summary of the ANOVA Comparing the Ratings Obtained in VESP Experiment #1

Source of Variation	Degrees of Freedom	Sum of Squares	Mean square	F-value	P-Value
Between Subjects	31	73.535	2.340		
Drive Quality Ratings	3	93.451	31.150	36.655	<0.0001
Error Term (Ratings X Subjects)	93	79.034	0.850		

Table 3.7 shows that there was a statistically significant difference, at the $p<0.0001$ level, in the ratings obtained for the four conditions. Condition vD had the highest mean rating, while Condition vA had the lowest mean rating.

3.2.1.3. Ranking Data in VESP Experiment #1—After driving in the simulator and rating each of the four conditions individually, the subjects were then asked to rank the four drives in order of preference—again with one indicating the condition the subject preferred, and four the condition that the subject preferred least.

A Friedman Two-Way ANOVA by Ranks was used to analyze this data. The resultant value of F_r was 48.4—this is statistically significant at the $p<0.001$ level (the critical value of F_r for three degrees of freedom at this significance level is 16.27). Condition vD was the most preferred combination and Condition vA was the least preferred.

3.2.1.4. Relationship between the Rating and Ranking Data in VESP Experiment #1—The average ratings and rankings obtained VESP Experiment #1 are presented in Table 3.8. The table shows that there was complete agreement between the ratings and the rankings—Condition vD received the highest mean rating and the best mean ranking, Condition vC received the second highest mean rating and the second best mean ranking, Condition vB received the third highest mean rating and the third best mean ranking, and Condition vA had the lowest mean rating and the worst mean ranking.

Table 3.8: Comparison of Mean Ratings and Mean Rankings in VESP Experiment #1

Condition	Mean Rating	Mean Ranking
vA	3.273	3.469
vB	3.844	2.969
vC	4.625	2.188
vD	5.547	1.375

The strength of the relationship between the mean rating and mean ranking values presented in Table 3.8 were tested using the Spearman rank correlation coefficient (Spearman's "rho"). The "rho" value for the data presented in Table 3.10 is 1.00, and as Olds (1938 & 1949) shows, even though there are only four comparison points, a "rho" value of 1.00 is statistically significant at the $p<0.05$ level.

3.2.2. VESP Experiment #2

3.2.2.1. Driving Speed in VESP Experiment #2—Thirty-two subjects completed the second VESP experiment. The average driving speeds at which they drove in Conditions vE, vF, vG, and vH were 49.37 km/h (SD=1.95 km/h), 52.59 km/h (SD=2.39 km/h), 56.56 km/h (SD=2.60 km/h), and 62.08 km/h (SD= 2.54 km/h), respectively. These average times (and a range of plus/minus 1.0 SD) are presented, along with the actual driving times and actual total times, in Table 3.9.

Table 3.9: Actual Average Driving Speeds, Average Driving Times and Total Times in VESP Experiment #2

Condition	**Actual Average Driving Time (plus/minus one Standard Deviation)**	**Actual Average Driving Time**	**Ramp Waiting Time**	**Actual Average Total Time**
vE	49.37 km/h (47.42—51.32 km/h)	19.6 minutes	0 minutes	19.6 minutes
vF	52.59 km/h (50.20—54.98 km/h)	18.4 minutes	2 minutes	20.4 minutes
vG	56.56 km/h (54.05—57.07 km/h)	17.1 minutes	4 minutes	21.1 minutes
vH	62.08 km/h (59.54—64.62 km/h)	15.6 minutes	6 minutes	21.6 minutes

The desired total time for all four conditions in VESP Experiment #2 was 20 minutes. Table 3.9 shows that the actual average times gradually increased from Condition vE through Condition vF and Condition vG to Condition vH. The first two of these conditions—vE and vF were close to 20 minutes—they were 0.4 minutes slower and 0.4 minutes faster, respectively. However, the total times for Conditions vG and vH were 1.1 minutes and 1.6 minutes slower than planned, respectively.

3.2.2.2. Rating Data in VESP Experiment #2—Rating data were obtained from the 32 subjects who participated in VESP Experiment #2 directly after each subject experienced the combinations of ramp meter waiting times and driving times. The ANOVA that was conducted on these data is summarized in Table 3.10.

Table 3.10: Summary of the ANOVA Comparing the Ratings Obtained in VESP Experiment #2

Source of Variation	**Degrees of Freedom**	**Sum of Squares**	**Mean square**	**F-value**	**P-Value**
Between Subjects	30	148.177	4.939		
Drive Quality Ratings	3	11.419	3.806	5.184	0.0024
Error Term (Ratings X Subjects)	90	66.081	0.734		

Table 3.10 shows that there was a statistically significant difference (at the *p* =0.0024 level) in the ratings obtained for the four conditions in VESP Experiment #2. Condition vH received the highest mean rating and Condition vE the lowest.

3.2.2.3. Ranking Data in VESP Experiment #2—After driving in the simulator and rating each of the four conditions individually, the 32 subjects who participated in VESP Experiment #2 then ranked the four drives in order of preference.

The Friedman Two-Way ANOVA by Ranks was used to analyze this data. The value of F_r obtained in the analysis was 7.0—a value which fails to reach statistical significance (the critical value of F_r for 3 degrees of freedom and a significance level of p=0.05 is 7.82).

3.2.2.4. Relationship between the Rating and Ranking Data in VESP Experiment #2—The average ratings and rankings that were obtained in the first VESP experiment are presented in Table 3.11.

Table 3.11: Comparison of the Mean Ratings and Mean Rankings Obtained in VESP Experiment #2

Condition	Mean Rating	Mean Ranking
vE	3.77	2.81
vF	4.45	2.58
vG	4.13	2.48
vH	4.55	2.16

Inspection of Table 3:11 suggests a likely reason for the *rating* data being statistically significant (as indicated in subsection 3.2.2.2) while the *ranking* data was not significant (as indicated in subsection 3.2.2.3). The ratings were made using a seven-point category scale, which essentially expands the four-point category scale that used to rank the combinations of ramp meter waiting times and driving times. Since the range between the highest and lowest ratings for VESP Experiment #2 was small (it was only from 4.55 for Condition vH to 3.77 for Condition vE), it is not surprising that the differences on the more compressed ranking scale were not statistically significant.

Table 3.11 also shows that there was not complete agreement between the rating and ranking data. Given the small range between the highest and lowest ratings, and given that the difference between the mean rankings was not statistically significant (as mentioned in subsection 3.2.2.2 above), the lack of complete agreement is not surprising.

3.2.3. VESP Experiment #3

3.2.3.1. Driving Speed in VESP Experiment #3
Twenty-five subjects completed the third VESP experiment. The average driving speeds at which they drove in Conditions vI, vJ, vK, and vL in this experiment were—93.29

km/h (SD=6.46 km/h), 95.34 km/h (SD=7.90 km/h), 95.25 km/h (SD=6.63 km/h), and 93.63 km/h (SD= 7.55 km/h), respectively. These average times (and a range of plus/minus 1.0 SD) are presented, along with the actual driving times and actual total times, in Table 3.12.

The desired total time for all four conditions in VESP Experiment #2 was 10 minutes. Inspection of Table 3.11 shows that the actual average total times were close to this for all four conditions—Condition vI was 0.4 minutes slower, while Conditions vJ, vK and vL were all 0.1 minutes slower.

Table 3.12: Actual Average Driving Speeds, Average Driving Times and Total Times in VESP Experiment #3

Condition	**Actual Average Driving Time (plus/minus one Standard Deviation)**	**Actual Average Driving Time**	**Ramp Waiting Time**	**Actual Average Total Time**
vI	93.29 km/h (86.83—99.50 km/h)	10.4 minutes	0 minutes	10.4 minutes
vJ	95.34 km/h (87.44—103.24 km/h)	8.1 minutes	2 minutes	10.1 minutes
vK	95.25 km/h (88.62—101.88 km/h)	6.1 minutes	4 minutes	10.1 minutes
vL	93.63 km/h (86.06—101.20 km/h)	4.1 minutes	6 minutes	10.1 minutes

3.2.3.2. Rating Data in VESP Experiment #3—The results of conducting an ANOVA on the rating data that were obtained from the 25 subjects who participated in VESP Experiment #3 are summarized in Table 3.13.

Table 3.13: Summary of the ANOVA Comparing the Ratings Obtained in VESP Experiment #3

Source of Variation	Degrees of Freedom	Sum of Squares	Mean square	F-value	P-Value
Between Subjects	24	55.840	2.327		
Drive Quality Ratings	3	11.600	3.867	4.193	0.0086
Error Term (Ratings X Subjects)	72	66.400	0.922		

Table 3.13 shows there was a statistically significant difference in the ratings obtained for the four conditions. Condition vI had the highest mean rating, while Condition vL had the lowest.

3.2.2.3. Ranking Data in VESP Experiment #3—After directly experiencing the four combinations of ramp meter waiting time and driving time the 25 subjects who participated in VESP Experiment #3 ranked the four combinations in order of preference. The Friedman Two-Way ANOVA by Ranks was used to analyze this data. There were statistically significant differences in the rankings obtained—the F_r value of 12.7 that was obtained exceeds 11.34, which is the critical value of F_r for 3 degrees of freedom for the p=0.01 significance level. Condition vI had the best mean ranking and Condition vL had the poorest mean ranking.

3.2.2.4. Relationship between the Rating and Ranking Data in VESP Experiment #3—The mean ratings and rankings obtained in VESP Experiment #3 are presented in Table 3.14.

Table 3.14: Comparison of Mean Ratings and Mean Rankings Obtained in VESP Experiment #3

Condition	Mean Rating	Mean Ranking
vI	5.40	2.08
vJ	5.16	2.29
vK	5.12	2.37
vL	4.48	3.25

Table 3.14 shows that there is complete agreement between the ratings and the rankings. Condition vI received the highest mean rating and the best mean ranking, Condition vJ received the second highest mean rating and second best mean ranking, Condition vK received the third highest mean rating and third best mean ranking, and Condition vL had the lowest mean rating and the poorest mean ranking. Spearman's "rho" was used to determine the strength of the relationship between the rating and ranking values presented

in Table 3.15. The “rho” value for these data was 1.00, and as Olds (1938 & 1949) shows this value is statistically significant at the 0.05 level, when there are four comparison points.

CHAPTER 4. DISCUSSION

In the subsections that follow, first the results obtained with the CASP and VESP methodologies are compared; then the consistency of the results obtained with the CASP and VESP procedures is discussed; next travel choice preference utility models are derived using the data obtained in the CASP and VESP experiments; and finally the CASP and VESP methodologies are compared.

4.1. Comparing the Results Obtained with the CASP and VESP Methodologies

The combinations of ramp meter waiting time and driving time that were tested in the first two parts of the CASP experiment (CASP-a and CASP-b) were essentially the same as those that were used in the first two VESP experiments.

4.1.1. CASP-a and VESP Experiment #1

For the conditions tested in CASP-a and VESP Experiment #1, as the ramp meter waiting times increased, both the driving times and the total travel times decreased. Interestingly, the two methodologies used to obtain preference data produced radically different results.

For CASP-a, both the rating and ranking data showed that the subjects preferred Condition cB (which had a ramp waiting time of two minutes and a driving time of 15 minutes); while their least preferred condition was vD (which had a ramp waiting time of six minutes and a driving time of 10 minutes).

In contrast, both the rating and ranking data obtained from the subjects in VESP Experiment #1, who actually experienced the waiting and driving, preferred Condition vD (with its six-minute ramp waiting time and 10-minute driving time). Further, the order of preferences showed that they preferred longer ramp waiting times and shorter driving times to shorter ramp waiting times and longer driving times.

Given the differences between the conditions investigated in CASP-a and VESP Experiment #1, there appear to be two possible explanations as to why subjects in the VESP Experiment #1 would prefer Condition vD. First, they may have preferred it because it was the shortest in duration. Second, they may have preferred it because it allowed them to drive in less congested traffic.

4.1.2. CASP-b and VESP Experiment #2

For the conditions tested in CASP-b and VESP Experiment #2, again as the ramp meter waiting times increased the driving times decreased, but in these cases the total travel times remained constant—thus testing the first of the possible explanations offered for the differences in the results obtained in CASP-a and VESP Experiment #1. Again the CASP and VESP methodologies produced radically different preference data.

For CASP-b, both the rating and ranking data showed that the least preferred Condition was cH (which had a ramp waiting time of six minutes and a driving time of 14 minutes). This was completely contrary to both the rating and ranking data obtained from the subjects in VESP Experiment #2, who actually experienced the waiting and driving—these subjects preferred Condition vH (with its six-minute ramp waiting time and 14-minute driving time). And again, in VESP Experiment #2, the order of their preferences showed that they preferred longer ramp waiting times and shorter driving times.

In the case of this comparison, the first of the possible reasons suggested in the previous subsection to explain the differences in preferences between the subjects in CASP-a and VESP Experiment #1 could be eliminated—because the total travel time was similar for all four conditions in CASP-b and VESP Experiment #2. However the second explanation is still viable—the subjects may have preferred Condition vH because it allowed them to drive in less congested traffic.

4.2. Consistency of the Results of the CASP and VESP Experiments

Ranking data cannot be used to compare data obtained in studies involving different combinations of conditions. However, *rating* preference data can be used to make rudimentary comparisons between conditions tested in different studies.

4.1.2. Comparing Conditions in CASP-a and CASP-b
In CASP-a, the lowest average rating (2.84) was given to Condition cD, the combination presented to the subjects that involved the longest ramp meter waiting time (six minutes), and a driving time of ten minutes. Similarly in CASP-b, the lowest average rating (2.42) was given to Condition cH, which was also the combination with the longest ramp meter waiting time (again six minutes), and a driving time of 14 minutes.

4.2.2. Comparing Conditions in the Three VESP Experiments
In VESP Experiment #1, the highest average rating (5.55) was given to Condition vD, the combination in which the subjects experienced the longest ramp meter waiting time (six minutes), the longest driving time (ten minutes), *and* relatively uncongested traffic. Similarly in VESP Experiment #2, the highest average rating (5.45) was given to Condition vH, again the combination in which the subjects experienced the longest ramp meter waiting time (six minutes), the longest driving time (14 minutes), *and* the least congested traffic level in the set of four used in the experiment.

Then in VESP Experiment #3, in which the subjects drove in relatively uncongested traffic in all conditions, there were similarly high ratings (of 5.40, 5.16, and 5.12, respectively) for Conditions vI, vJ, and vK (in all of which the subjects drove in uncongested traffic for ten, eight and six minutes)—and it was only for condition vL (in which the ramp meter waiting time was six minutes and the subjects drove in uncongested traffic for *only* four minutes) that the rating dropped below 5.00.

These findings from the three VESP experiments support the explanation that subjects prefer driving in uncongested traffic—and that is the reason why they preferred the combinations of longer waiting times and shorter driving times in VESP Experiments #1 and #2.

4.3 Travel Choice Preference Utility Models

Choice models were derived from the preference data obtained in the CASP and VESP Experiments. In addition to the preference data, and the ramp meter waiting times and driving times used in the experiments, these models took into consideration information derived from the NEO-FFI, from the survey of socioeconomic data, and from the travel diaries. The Neuroticism (N), Extraversion (E), and Openness (O) scores (but not the Agreeableness or Conscientious scores) obtained with the NEO-FFI were all found to contribute to the model for at least one set of data. And information on Gender (G), Age (A), average age of all household Vehicles (V), daily one-way Commute time (C), and Total daily travel time (T), which were extracted from the survey of socioeconomic data or from the travel diaries provided by the subjects also contributed to the model. In deriving a preference utility model, it was assumed that each subject attempts to maximize his or her utility by minimizing travel time.

The preference travel time utility model from this study was:

$$U = f\,(F, R, N, E, O, G, A, V, C, T) \qquad \text{(Equation 1)}$$

where:

U = preference travel time Utility
F = Freeway travel time (in minutes)
R = Ramp wait time (in minutes)
N =Neuroticism score (from the NEO-FFI)
E = Extraversion score (from the NEO-FFI)
O = Openness score (from the NEO-FFI)
G = Gender (1 = male, 0 = female)
A = Age (years)
V = average age of household Vehicles in service (years)
C = one-way Commute time (minutes)
T = Total daily travel time (minutes)

Three of the domains explored by the NEO-FFI—Neuroticism, Extraversion and Openness—are represented in this model. The other two domains—Agreeableness and Conscientiousness—did not appear to have any connection to expressed travel preferences in either the CASP or VESP Experiments.

Also, the household income of the subject was not found to be a significant variable in the model.

In the next four subsections of this report the values of the coefficient (β's) and associated z-statistics obtained for each of these variables are presented for the CASP-a, VESP Experiment #1, CASP-b, and VESP Experiment #2.

4.3.1 Travel Choice Preference Utility Model Using Data from CASP-a

Table 4.1 presents the values of the coefficients (β's) and associated z-scores obtained for each of the variables in Equation 1 for CASP-a.

Table 4.1: Coefficient and z-scores for the Variables for CASP-a

Variable	**Coefficient**	***z*-score**	***p*-value**
Change in Freeway time (F)	-0.854	-6.29	<0.0001
Change in Ramp time (R)	-1.404	-5.80	<0.0001
Neuroticism (N)	+0.045	+1.65	NS (p=0.099)
Extraversion (E)	-0.053	-2.18	0.029
Openness (O)	+0.059	+2.10	0.036
Gender (G)	-1.410	-3.50	<0.0005
Age (A)	+0.029	+1.92	NS (p=0.055)
Average Vehicle age (V)	+0.162	+3.87	<0.0001
Commute time (C)	-0.025	-1.83	NS (p=0.067)
Total daily travel time (T)	+0.014	+2.10	0.036
Constant	-2.679	-1.49	NS (p=0.136)

The probabilities for each variable in Table 4.1 were calculated in STATA 7.0 using a binary logit model with the variables entered linearly. It was determined that 65 percent of the preferences expressed in CASP-a favored combinations with shorter ramp meter waiting times and longer driving times.

For CASP-a, as Table 4.1 indicates, change in Freeway travel time and change in Ramp wait time were both highly significant (at the p<0.0001 level).

In addition, the table shows that several other variables provided significant contributions to the model. The Extraversion score was significant, at the p=0.029 level—subjects with higher Extraversion scores were more likely to prefer combinations of shorter ramp meter waiting times and longer driving times. The Openness score was also significant, at the p=0.036 level—subjects with higher Openness scores were more likely to prefer combinations of longer ramp meter waiting times and shorter driving times. The Gender of the subject was significant, at the $p<0.0005$ level—female subjects were more likely to prefer combinations of longer ramp meter waiting times and shorter driving times than male subjects. Also, the average age of Vehicles in the subject's household was significant, at the $p<0.0001$ level—the subjects who had older vehicles in their household were more likely to prefer combinations of longer ramp meter waiting times and shorter driving times than those with newer cars. Finally, the Total daily travel time was significant, at the $p=0.036$ level—subjects who had higher Total daily travel times were more likely to prefer combinations of longer ramp meter waiting times and shorter driving times than those with lower total daily travel times

The remaining three variables—the Neuroticism score ($p=0.099$); the Age of subject ($p=0.055$); and the one-way Commute time ($p=0.067$) all approached significance.

4.3.2 Travel Choice Preference Utility Model Using Data from VESP Experiment #1
Table 4.2 presents the values of the coefficients and associated *z*-scores obtained for each of the variables in Equation 1 for VESP Experiment #1.

Table 4.2: Coefficients and *z*-Scores for the Variables in VESP Experiment #1

Variable	**Coefficient**	***z*-score**	***p*-value**
Change in Freeway time (F)	+0.405	+1.45	NS ($p=0.147$)
Change in Ramp time (R)	+1.073	+1.99	0.047
Neuroticism (N)	-0.016	-0.22	NS ($p=0.826$)
Extraversion (E)	-0.149	-1.15	NS ($p=0.250$)
Openness (O)	-0.111	-0.75	NS ($p=0.453$)
Gender (G)	-3.639	-2.62	0.0088
Age (A)	-0.077	-1.36	NS ($p=0.174$)
Average Vehicle age (V)	-0.196	-0.76	NS ($p=0.447$)
Commute time (C)	+0.111	+1.91	NS ($p=0.057$)
Total daily travel time (T)	+0.008	+0.40	NS ($p=0.689$)
Constant	+8.323	+1.66	NS ($p=0.097$)

The probabilities for each variable in Table 4.2, as with the probabilities in Table 4.1, were calculated in STATA 7.0 using a binary logit model with the variables entered linearly. In the case of VESP Experiment #1, it was determined that 81 percent of the preferences expressed were for the combination of greater ramp meter waiting time and shorter driving time—which is in opposition to the findings for the CASP-a data.

Table 4.2 indicates that for VESP Experiment #1, the change in Ramp meter waiting time was significant, at the $p=0.047$ level.

The only other variable that provided a statistically significant contribution to the model was the Gender of the subject, (at the p=0.0088)—female subjects had a greater preference for the combination of longer ramp meter waiting times and shorter driving times than the male subjects.

One other variable, the one-way Commute time (p=0.057), also approached significance. The coefficients for the remaining variables were not significant.

4.3.3. Travel Choice Preference Utility Model Using Data from CASP-b

Table 4.3 presents the values of the coefficients and associated z-scores obtained for each of the variables in Equation 1 for CASP-b

Table 4.3: Coefficients and z-Scores for the Variables in CASP-b

Variable	**Coefficient**	***z*-score**	***p*-value**
Change in Freeway time (F)	+0.054	+0.56	NS (p=0.575)
Change in Ramp time (R)	—	—	—
Neuroticism (N)	-0.024	-0.98	NS (p=0.327)
Extraversion (E)	-0.063	-2.79	0.005
Openness (O)	+0.051	+1.96	0.050
Gender (G)	-1.621	-4.45	<0.0001
Age (A)	+0.020	+1.41	NS (p=0.159)
Average Vehicle age (V)	+0.102	+2.74	0.006
Commute time (C)	-0.009	-0.77	NS (p=0.441)
Total daily travel time (T)	+0.004	+0.61	NS (p=0.541)
Constant	+0.100	+0.06	NS (p=0.952)

Again the probabilities for the variable in Table 4.3 were calculated with STATA 7.0 using a binary logit model with the variables entered linearly. For CASP-b, in which the total travel time was unchanged across the four combinations that were presented to the subjects, 68 percent of the preferences expressed favored combinations with shorter ramp meter waiting times and longer driving times.

Table 4.3 shows that four variables provided significant contributions to the model. The Extraversion score was significant, at the p=0.005 level—subjects with higher Extraversion scores were more likely to prefer combinations of shorter ramp meter waiting times and longer driving times. The Openness score was significant, at the p=0.05 level—subjects with higher Openness scores were more likely to prefer combinations of longer ramp meter waiting times and shorter driving times. Gender was significant, at the p<0.0001 level—female subjects were more likely to prefer combinations of longer ramp meter waiting times and shorter driving times than male subjects. The average age of Vehicles in the subject's household was significant, at the p=0.006 level—subjects with older vehicles in their household were more likely to prefer combinations of longer ramp meter waiting times and shorter driving times than those with newer cars.

The table also shows that the coefficients for the remaining variables were not statistically significant.

4.3.4. Travel Choice Preference Utility Model for Data from VESP Experiment #2

The values of the coefficients and associated z-scores obtained for each of the variables in Equation 1 for VESP Experiment #2 are presented in Table 4.4.

Table 4.4: Coefficients and z-Scores for the Variables in VESP Experiment #2

Variable	**Coefficient**	***z*-score**	***p*-value**
Change in Freeway time (F)	-0.096	-0.66	NS (p=0.507)
Change in Ramp time (R)	—	—	—
Neuroticism (N)	+0.103	+1.25	NS (p=0.210)
Extraversion (E)	+0.003	+0.06	NS (p=0.949)
Openness (O)	-0.007	-0.12	NS (p=0.902)
Gender (G)	-0.109	-0.17	NS (p=0.866)
Age (A)	+0.033	+1.10	NS (p=0.273)
Average Vehicle age (V)	+0.054	+0.70	NS (p=0.483)
Commute time (C)	-0.028	-1.12	NS (p=0.261)
Total daily travel time (T)	+0.012	+1.06	NS (p=0.291)
Constant	-3.525	-0.70	NS (p=0.486)

Again the probabilities for the variable in Table 4.4 were calculated with STATA 7.0 using a binary logit model with the variables entered linearly. In VESP Experiment #2, 60 percent of the preferences expressed were for the combination of greater ramp meter waiting time and shorter driving time—in opposition to the findings for the CASP-b data.

As Table 4.4 shows, none of the variables were statistically significant.

4.4. Comparing the CASP and VESP Methodologies

The CASP and VESP methodologies produce distinctly different results. To suggest reasons why this occurred, several factors should be considered.

First, there may be a recency effect. In the VESP experiments, the subjects first experienced the wait at the ramp meter before driving on the freeway. Then in the conditions in which they encountered congested traffic, the congestion occurred in the middle of the drive on the freeway. It is possible that the most recent annoying part of the trip (driving in congested traffic) might have been perceived more negatively than earlier annoying elements. This possible recency effect might partially explain the differences in the preference data obtained with the VESP and CASP methodologies. This possibility could be tested in a driving simulation experiment in which the subjects encounter delays (perhaps caused by traffic lights) at the end of the journey rather than the beginning.

Second, the subjects in the CASP and VESP experiments perceived their tasks in different ways. In the CASP study, the subjects were presented combinations of ramp meter wait times and driving times—and thus are likely to have focused on these two factors. In the VESP experiments, the subjects experienced the combinations of ramp meter wait and the drive of varying duration in traffic with varying degrees of congestion—and in this case are likely to have focused on congestion as well as the other two factors.

Third, in the real world, driving is a goal-directed activity—with the drivers having the objective of reaching a particular destination. This is particularly true of commuting. In the VESP experiments, although the subjects were driving to a particular exit in the simulated scenario, they were not driving to that exit for their own reasons (e.g., to commute to work). Because of this, the drivers in the VESP experiments may not have been particularly concerned about their destination. In the simulator, as in real-life, congested traffic may be annoying. But in real-life, the fact that you are driving to a desired destination could tend to ameliorate the annoyance. In contrast to the VESP experiments, the CASP methodology may not have tapped into the subjects' feelings about congestion.

Fourth, it is possible that some subjects in the VESP experiments may have experienced traffic that was more congested than the traffic they typically encounter on their daily commute. This may have increased their antipathy to the conditions in which they drove in congested traffic in the VESP experiments.

Fifth, ramp meters have frequently been mentioned in the media in the Twin Cities area in the past few 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. During the same time period, there has been far less media attention to congested traffic. Because of the differences in media coverage of ramp meters and traffic congestion, subjects may have responded to the CASP presentations by indicating that they do not like ramp meters—without giving consideration to how they feel about driving in heavily congested traffic. Thus, they may have expressed a built-in bias in the CASP experiment. In contrast, in the VESP experiment, the subjects actually experienced what it is like to drive in congested traffic. The latter methodology increases the likelihood that subjects actually compare the experience of waiting at a ramp meter with the experience of driving in congested traffic.

Unfortunately the CASP methodology did not produce preference data that mirrors the preference data obtained with VESP methodology. While the VESP methodology allows subjects to actually experience the conditions being investigated, it has a number of problems, including the accuracy of the simulation (are the conditions modeled realistic) and the use of an imperfect simulation tool (the number of subjects reporting simulator sickness remains a concern, though those specific subjects were excluded from the analysis). The lack of agreement between the preference data obtained with the two methodologies suggests that there are problems with one or both methodologies. Though we hypothesized that preference data acquired with an immersive driving simulator is much more likely to resemble real world preference data than CASP data, the present study suggests that additional research is necessary to understand how experimental stated preference techniques are understood and interpreted by subjects, and therefore by researchers. In particular, even more realistic experiments employing real-world conditions (real cars on real streets) will enable us to corroborate or refute the results from computer-based or driving simulator-based stated preference experiments.

REFERENCES

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

Calfee, J., Winston, C., and Stempski, R. (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*, **83**(4), 699-707.

Carson, K.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 Economics, University of Colorado at Boulder.

Christakis, N.A. and Allison, P.D. (1994). Logit Models for Sets of Ranked Items, *Sociological Methodology,* **24**, 199-228.

Colias, J. and Salazar-Velasquez, C. (1995). Integrating the Rank-Ordered Logit Model with Purchase Intent. *Canadian Journal of Marketing Research*, **14,** 46-56.

Costa, P.T. Jr. and McCrae, R.R. (1985). *The NEO-Personality Inventory Manual*. Odessa, FL.: Psychological Assessment Resources, Inc.

Costa, P.T. Jr. and McCrae, R.R. (1991). *Revised NEO-Personality Inventory and NEO Five-Factor Inventory (NEO-FFI): Professional Manual*. Lutz, FL.: Psychological Assessment Resources, Inc.

Garber, N.J., and Hoel, L.A. (1997). *Traffic and Highway Engineering*, PWS Publishing Company.

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.

Hennessy, D.A. and Wiesenthal, D.L. (1999) Traffic congestion, driver stress, and driver aggression. *Aggressive Behavior*, **25**(6), 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, D.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, M.B., Bettencourt, L.A., and Wenger, S. (1998). The relationship between waiting in a service queue and evaluations of service quality: A field theory perspective. *Psychology & Marketing*. **15**(8), 735-753.

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

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

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

Koop, G., Poirier, D.J. (1994) Rank-Ordered Logit Models: An Empirical Analysis of Ontario Voter Preferences. *Journal of Applied Econometrics*, **9**, 369-388.

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

Levinson, D. and Zhang, L. (2004). Ramp Meters. In: D. Gillen and D. Levinson (Editors) *Assessing the Benefits and Costs of ITS*. Kluwer Academic Publishers: Netherlands

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

Olds, E.G. (1938). Distribution of sums of squares of rank differences for small numbers of individuals. *Ann. Math. Statist*., **9**, 133-148.

Olds, E.G. (1949). The 5% significance levels for sums of squares of rank differences and a correction. *Ann. Math. Statist*., **20**, 117-118.

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

Piotrowicz, G. and Robinson, J. (1995) Ramp Metering Status in North America—1995 update. Virginia Department of Transportation Technical Report.

Redmond, L.S. and Mokhtarian, P.L. (1999) The positive utility of the commute: modeling ideal commute time and relative desired commute amount. *Transportation*.

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

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

Stokols, D., Novaco, R.W., Stokols, J., and Campbell, J. (1978). Traffic congestion, Type A behavior, and stress. *Journal of Applied Psychology*, **63**(4), 467-480.

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

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