

UC Berkeley

Research Reports

Title

System Fault Detection in Human-Augmented Automated Driving

Permalink

<https://escholarship.org/uc/item/3vw6w0h1>

Author

Cohn, Theodore

Publication Date

2001-09-01

CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

System Fault Detection in Human- Augmented Automated Driving

Theodore Cohn

University of California, Berkeley

California PATH Research Report

UCB-ITS-PRR-2001-26

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation; and the United States Department of Transportation, Federal Highway Administration.

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

Report for MOU 323

September 2001

ISSN 1055-1425

MOU323 Final Report
Principal Investigator: Theodore Cohn
System Fault Detection in Human-Augmented Automated Driving

I. Introduction

Lateral control of a vehicle in the Automated Highway System (AHS) has been formulated and simulated as part of the hierarchical AHS control structure. In that structure, lateral control resides in the vehicle, and is implemented as a closed-loop control system with the lateral deviation of the vehicle from a reference position as a controlled variable and the steering angle or its rate as a controlling input. The PATH AHS scenario has assumed a road reference and sensing system based on magnetic markers equally spaced along a highway lane. Freedom in the selection of the polarity of each magnet has been utilized to embed binary information along the lane, and it has been utilized to embed preview information (e.g. road curvature and superelevation) useful for the on-board controller to enhance performance both in terms of the lateral error and the occupant's riding comfort. Frequency shaped linear quadratic control (FSLQ), nonlinear sliding control, and rule-based control have been investigated for lateral vehicle control for lane keeping. Simulations of and experiments with this automated system have established that it performs quite well (keeping deviations from the lane centerline under 10 cm). This performance was verified in actual vehicle demonstrations during the 1997 NAHSC demonstration in San Diego. There is, however, still a fundamental question left open: how can the human operator (HO) best be integrated into the automated highway system?

The HO brings a unique perspective to the problem of lateral control, given his/her detailed knowledge of the driving environment, robust sensory capability, and higher level reasoning capabilities—particularly those that would be manifest in response to unanticipated situations. The problem examined in this research is how to incorporate human guidance into the lateral control system (“controller”). In particular, we concentrate on fault detection under several types of combined human/controller guidance. For each, system performance is quantified in a simulated lateral guidance task using both human cognition models derived from signal detection theory and measured performance incorporating human observers. Implementation of AHS will require a thoughtful deployment in which the human operator has been integrated into the system in ways such as this. Indeed, it is doubtful that the driving public would have the confidence to accept AHS in the absence of such participation.

This project was intended to be the first in a series of studies that explored human augmentation of AHS components. Other studies were anticipated to follow, a few of the more important being lane changing, longitudinal control, and collision avoidance.

II. Experimental Apparatus

Our experimental setup consists of a real-time, “PC”-based dynamic model of an automobile. It presents an animated view of a car driving down a road, as it would appear looking through its front windshield. This is shown in Figure 1. As indicated, the scene itself has been kept simple, incorporating only the essential roadway elements (but see below). This model (a so-called “bicycle model”) takes into account vehicle dynamics so that the lateral response of the automobile is realistic. It was developed in previous PATH projects (UCB-ITS-PRR-97-28 and “Fault Detection and Tolerant Control for Lateral Guidance of Vehicles on Automated

Highways” – Ph.D. Thesis by S.N. Patwardhan). The car can be steered either automatically (employing a simple proportional, or constant gain, controller) or by a human test subject through the use of a standard “game” steering wheel connected to the PC. The road on which the car drives is a seven-mile long loop whose initial segment replicates the Crow’s Landing test track. It is shown in Figure 2.

The model only provides for visual input. It is recognized that a human driver also receives important audio and haptic cues regarding the driving environment, but investigations into the effects of these types of inputs were beyond the scope of this project. In the discussion that follows, then, we focus exclusively on the effect of

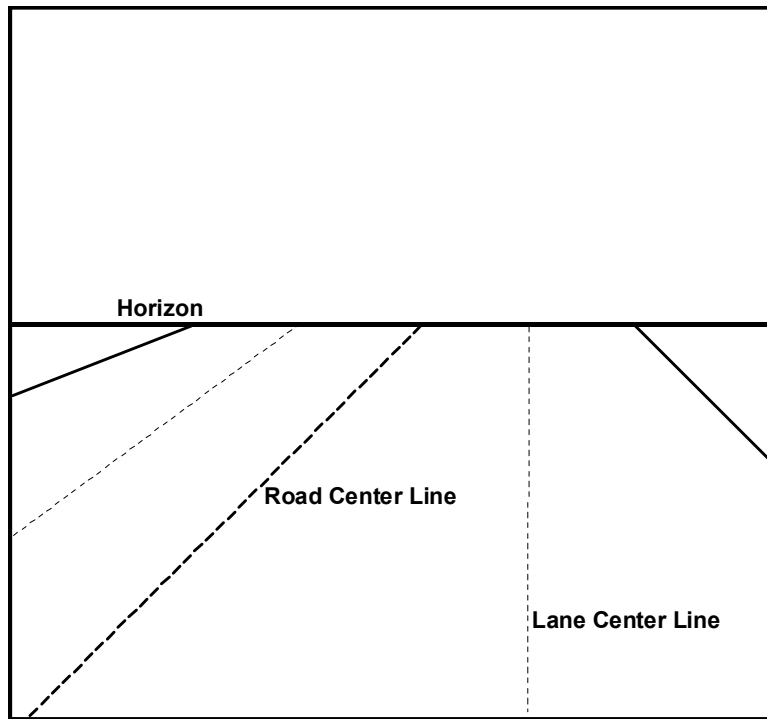


Figure 1: View Through Vehicle Windshield

human vision on the performance of the combination formed by the HO and the steering controller. This is not regarded as a severe limitation, however. First, vision is our most important sensory input, as evidenced by the fact that every state requires an individual to have certain minimum visual capabilities as a pre-requisite to receiving a driver’s license. Such stipulations don’t exist in the case of even our next-most-important form of sensory input, hearing. In view of this fact alone, then, any control strategy that might arise out of a study such as this would probably be required to rely exclusively on human vision, leaving additional benefits in performance expected from the other sensory inputs to bolster the strategy’s margin of safety.

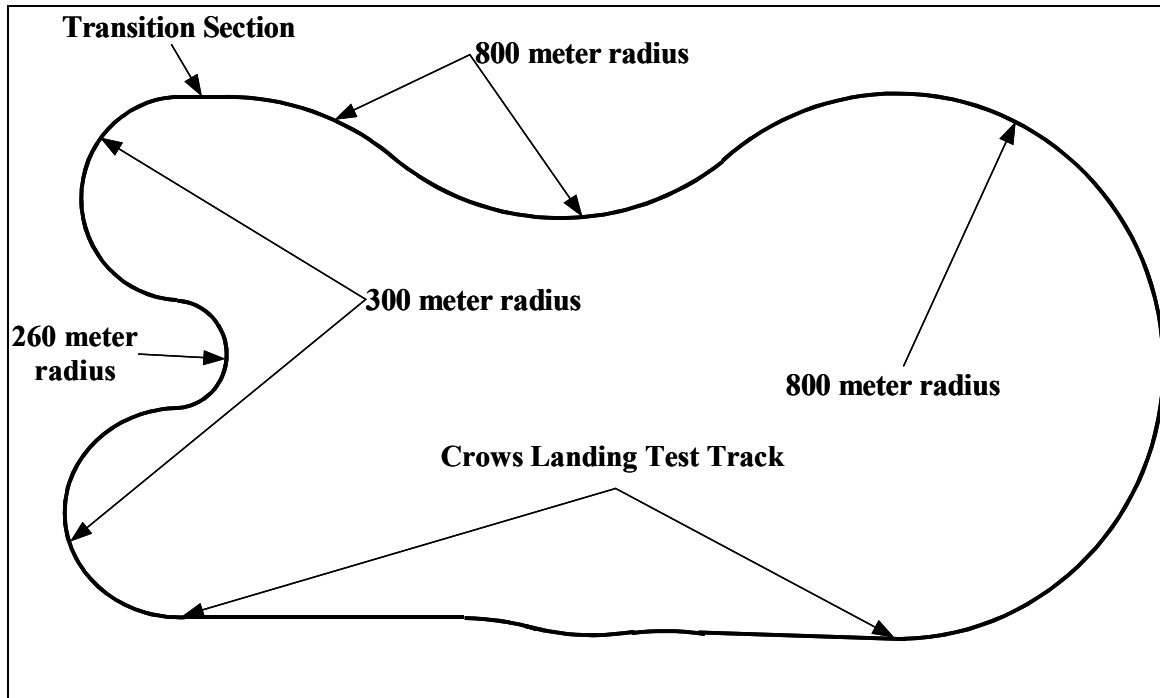


Figure 2: Vehicle Model Test Track

The PC-based model was also developed with the goal that it could easily be transported, modified, and run on any standard Windows NT computer (requiring only a standard game steering wheel that can readily be purchased at nominal cost from any computer store). It is built up from individual modules (subroutines) that can easily be augmented and replaced, allowing the user to test different roadway configurations, controller designs, fault scenarios, and different kinds of HO collaborations and interventions. The roadway scene itself can also be modified to incorporate additional visual elements and distractions. The model thus serves as a *pre-simulator*. With it, a designer can construct and simulate a particular driving scenario, and test and analyze the resulting system's performance. In those cases where the model is not sufficient to simulate a desired scenario, it can be used in the development of specifications for more sophisticated tests to be performed either on a full simulator or an actual test track.

II. Test Methodology

The project's primary goal was to assess an HO's ability to detect and respond to faults and/or hazardous driving conditions the controller may not know about. To achieve this, two different types of test methodologies have been developed:

- *Detection tests* to quantify the test subject's ability to detect a vehicle controller faults that cause near-threshold events.
- *Quality of Reaction tests* to assess the quality of the subject's reaction to faults.

Detection Test Series (Test Series #1)

The goal of Detection test series was to determine both how accurately and how quickly an HO can detect a controller fault. To ascertain this, test subjects took a “ride” in a “normally”¹, automatically controlled car as it made a circuit around the test track. (A single circuit around this track takes approximately 7 minutes to complete at 55 mph.). Gaussian noise having a zero mean and a one meter standard deviation was overlaid onto the vehicle’s position, to simulate system noise, during the tests. During the ride controller faults (the controller gain was set to zero, causing the vehicle to drift off the road) and “near faults” (the vehicle was gradually steered off the lane centerline so that it approached the threshold of a fault, and then allowed to gradually return to normal operation) were randomly initiated at the fixed locations shown in Figure 3² and the HO was asked to indicate by pushing a button on the keyboard when he or she perceived a fault. The vehicle’s lateral and angular deviation were recorded at that instant, as well as the instant that the fault actually was initiated.

A fault tolerance was also applied such that if the HO did not perceive an actual fault before the lateral/angular deviations exceeded some preset value, the HO was deemed to have “missed” the occurrence (at which point the fault was removed, and the vehicle returned to normal operation). From this we obtained the following data:

- Number of faults
- Number of “hits” (correctly perceived a fault)
- Number of “misses” (did not perceive a fault when one was present)
- Number of “False Alarms” (an indication that a fault occurred when there was no fault within a preset time interval before and after the indication was given.)
- The lateral and angular deviation from the lane centerline for each of the above.
- Point on the circuit where the instance occurred.

¹Under normal automatic control we employ a constant gain controller, which provides an accurate simulation of an actual, automatically controlled vehicle’s behavior.

² Even though the faults were applied at the same locations for each test, the test track is long enough and varied enough that test subjects did not “learn” these locations as the tests progressed.

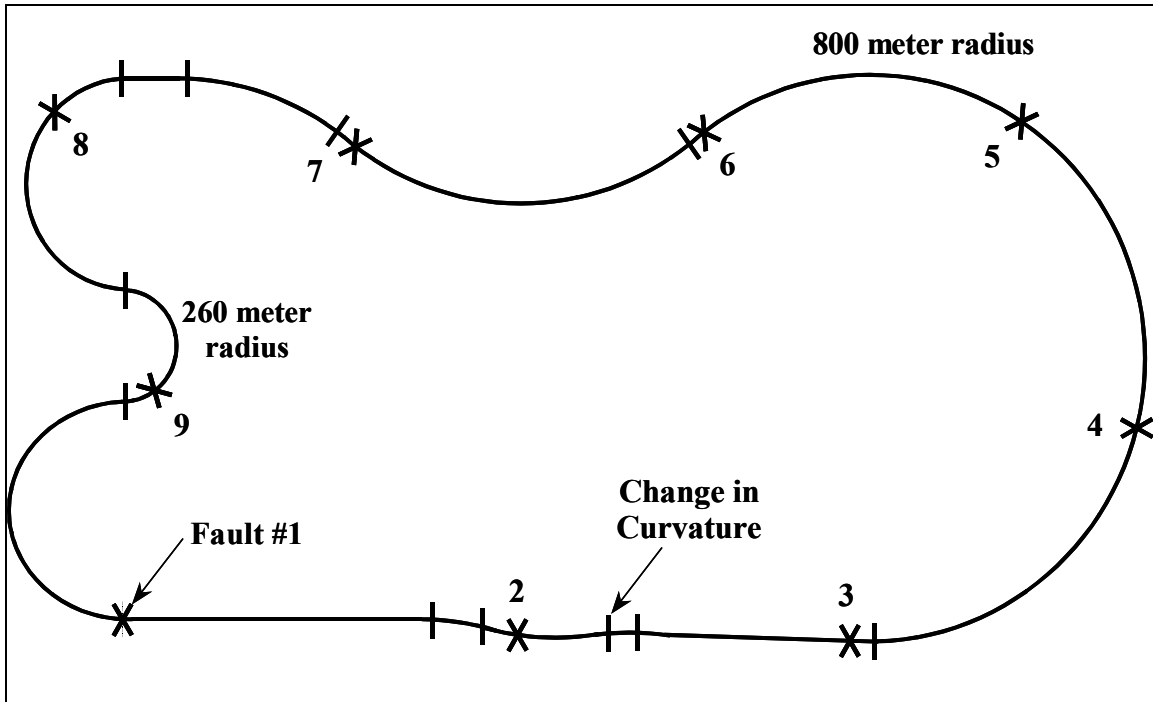


Figure 3: Fault and Curvature Change Locations

The last piece of data, event location, is of interest in that fault detection performance may be a function of road geometry (straight-away, simple curve, S-curve, etc.). Faults were therefore programmed to occur at each type of geometry, as can be seen from Figure 3. In all, 63 tests were performed with 7 different test subjects.

Quality of Reaction Test Series (Test Series #2)

- R1.** The goal of the Quality of Reaction test series was to determine how well an HO was able to compensate for a controller fault and keep the car safely in its lane. This first series of tests were baseline tests in which the subject assisted in steering a normally operating, automatically controlled car with no noise; in each of the three driving scenarios (Helper, Peer, and Crisis Handler) around the test track at 55 mph. Performance (rms lateral and angular deviation) under each of these conditions was compared to the performance of the vehicle under automatic control alone. Five such tests were conducted, the first primarily to give the test subject “practice” in “driving the car”.
- R2.** The second part of this test series was not performed. (See below.) These five tests were to be the same as the first set, except that now noise and random controller faults would be introduced. Rms lateral and angular deviation would again be recorded, as well as the points at which faults and/or noise was introduced.

III. Results

Detection Test Series

The primary results of the Detection Test Series are summarized in Tables 1 and 2. As a group, the seven test subjects correctly detected 287 of 298 “true faults” (a 96.3% “Hit Rate”), but incorrectly reacted to 115 of 269 “near faults” (a 42.8% “False Alarm Rate”). “True Negatives” and “Misses” (less important for our purposes) are also indicated. These results are

perhaps easier to grasp when presented as in Figure 4. With them it is possible to establish the criteria that the test group used to decide when a fault existed. As was mentioned previously, gaussian noise was overlaid onto the lateral position of the vehicle, so that in normal operation, its lateral position would be a normally distributed random variable with zero mean (corresponding

Table 2: Response Summary

Table 1: Event Breakdown

Event	Number
True Faults	298
Near Faults	269
Total	567

Response	Number
True Faults Correctly Detected (Hits)	287
True Faults Incorrectly Ignored (Misses)	11
Near Faults Incorrectly Detected (False Alarms)	115
Near Faults Correctly Ignored (True Negatives)	154
Total	567

		Fault Initiated=298	Fault Not Initiated=269
Detection	No Detection	Hits=287 (96.3%)	False Alarms=115 (42.8%)
		Misses=11	True Negatives =154

Figure 4: Detection Test Series Summary

to the lane centerline) and a one meter standard deviation³. This is indicated by the left-most normal curve shown in Figure 5. (Here we will only consider “positive” deviations, to the right of the lane centerline, so our normal curve becomes a half curve with its height raised such that

³ As long as the car is traveling along a constant curvature segment of roadway its mean location will be essentially on the lane centerline. At the instant it encounters a change in curvature, though, the car may deviate from the centerline by as much as 10 centimeters before (quickly) coming back to the centerline, since this is the tolerance within which the controller operates. For purposes of our analysis we ignore this.

the area under it is still one.) If the vehicle were experiencing a fault condition, its lateral position would again be a normally distributed random variable with a one meter standard deviation, but now it's mean (d in Figure 5) would drift off to the right and thus be non-zero. (In this case, since the distribution was not centered at zero, only the left-most portion of its tail is removed, and again its height is adjusted upward to keep its total area equal to one.) In deciding whether a fault condition existed or not, then, the test subject tried to establish whether the lateral deviation he experienced at any instant belonged to the normal distribution on the left in Figure 5 (no fault) or to the one on the right⁴. He would thus set some lateral deviation as his criterion, and any deviation greater than this would be judged a fault. This criterion is labeled the "Fault Criterion" in the figure. Knowing the Hit and False Alarm Rates (as provided by the tests) we can work back to obtain both the Fault Criterion and the mean for the rightmost distribution, at the point that the HO determined a fault condition existed. For the "half-normal" distribution on the left in Figure 5, it can easily be found (from a table) that the area under the curve and to the right of $z = .56511$ will be .428. With this, we can find by trial and error that an adjusted normal distribution (with its leftmost tail cut off at $z=0$, and its density function increased to keep its total area equal to one) having a mean

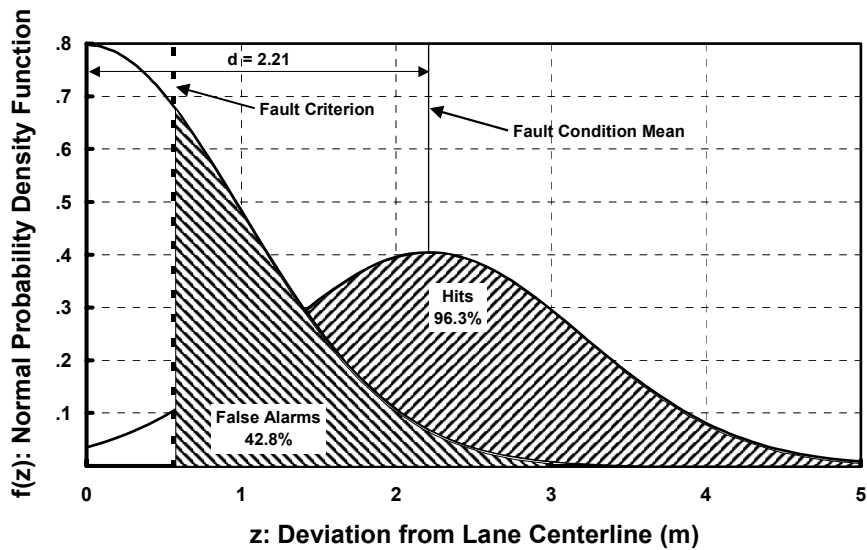


Figure 5: Criterion for Detecting Controller Faults

of 2.20923 and a standard deviation of one will have 96.3% of its total area to the right of $z = .56511$. Thus

$$\text{Fault Criterion} = .56511 \text{ m,}$$

$$\text{Fault Condition Mean} = 2.20923 \text{ m.}$$

⁴ The near fault condition was created to force all decisions to occur at fault locations, so that an unambiguous False Alarm Rate could be calculated. (Its counterpart, the number of True Negatives, is shown for completeness but otherwise is not used.) When a near fault condition is present, the vehicle's mean position will of course not be zero. Since the excursion from the lane centerline has been kept small and the duration during which it occurs kept short, we've ignored this detail to ease the analysis.

Each of the test subjects took up to 10 test runs so it is also possible (and instructive) to see how they adjusted their Fault Criteria from one run to the next, as they gained experience with the “vehicle’s” dynamics. In other words, as the test subjects become better at detecting faults, they will be able to do so for smaller values of d . We investigated this aspect of the test subjects’ behavior in the context of “Receiver Operating Characteristic (ROC) Curves”, a standard tool of Vision Science. Recall from our previous analysis that we started with the known characteristics of one of the normal distributions (the mean and standard deviation of the leftmost one) and the Hit and False Alarm Rates, and with these obtained the mean of the second normal distribution (the rightmost one). If instead we let the mean of this second distribution be the independent variable, then for any Fault Criterion we wish to specify we immediately can determine the resulting Hit and False Alarm Rates. (This is intuitively evident just by inspecting Figure 5.) So for any d we can calculate the Hit and False Alarm Rates for the full range of Fault Criteria, and then we can plot Hit Rate vs. False Alarm Rate. Such a calculation results in the smooth curves plotted in Figure 6. (To use the Figure, first select the curve associated with the particular mean d of interest. Then, for any desired False Alarm Rate on the horizontal axis, move up vertically to the selected curve and then horizontally to read the resulting Hit Rate on the vertical axis to the left. Alternatively, one can choose a Hit Rate and then read the resulting False Alarm Rate on the horizontal axis.) The plot clearly shows the tradeoff between False Alarm Rate and Hit Rate. We can’t decrease the False Alarm Rate without decreasing the Hit Rate as well. Also plotted in Figure 6 is the group’s performance on a run-by-run basis. This shows that the group started out with a high False Alarm Rate (in order to ensure a high Hit Rate) but also a relatively high value of d , and that as they became more familiar with the task they were able to reduce the False Alarm Rate and d while keeping the Hit Rate near 100%. This learning effect is further confirmed by plotting the average lateral deviation at the point that a fault was detected against test number, as is done in Figure 7. A discernible learning curve is evident here, indicating that test subjects got better at detecting faults with experience. For these particular tests no “consequences” were stipulated to the test subjects for Hits and False Alarms, so they presumably drew upon their own experiences in actual driving situations to infer that False Alarms carried little, if any penalty, while Misses (the opposite of

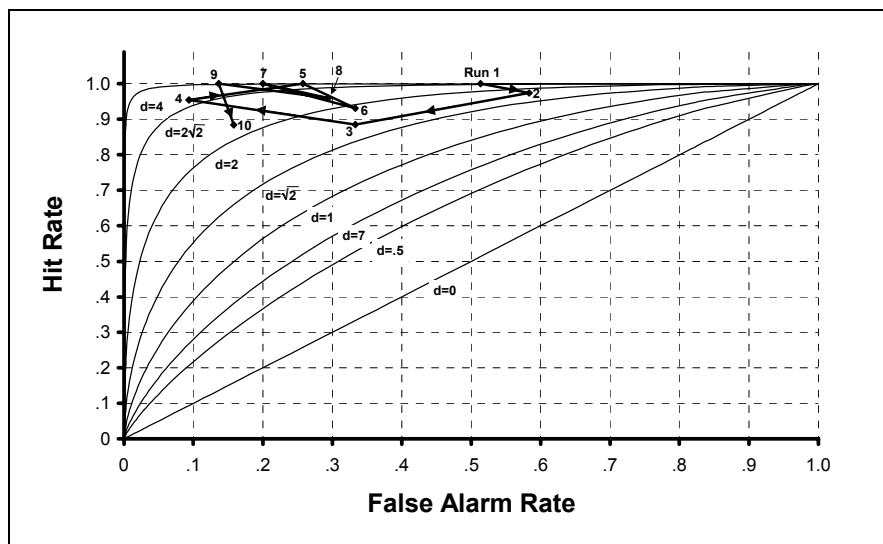


Figure 6: ROC Curves and Adjustment of Fault Criterion

Hits) carried the severe penalty of an accident. It is possible to set up different consequences by establishing explicit (say, monetary) rewards for Hits and penalties for False Alarms, and in this way influence the placement of the Fault Criterion.

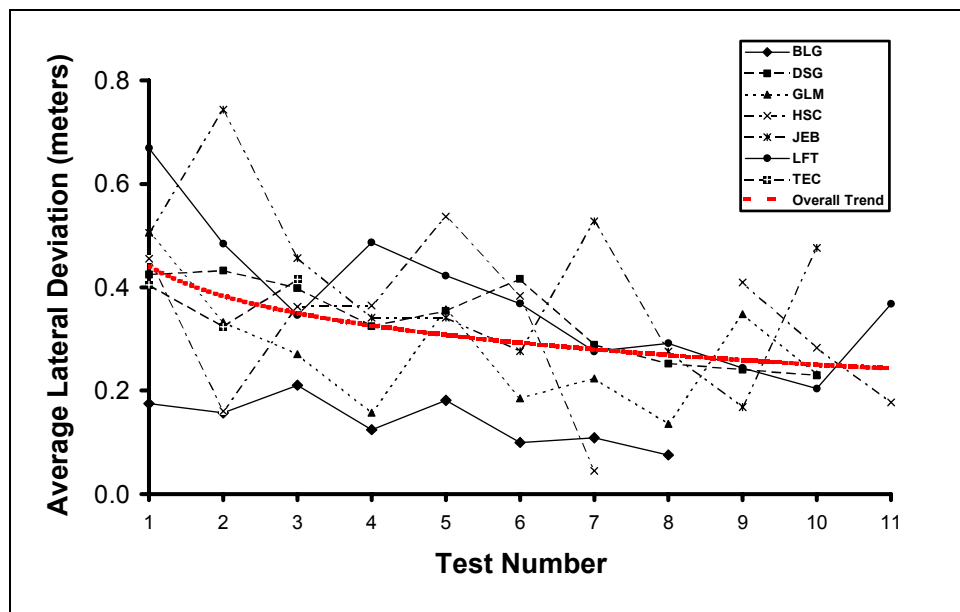


Figure 7: Test Subject Learning Effects

Quality of Reaction Test Series

Unanticipated difficulties encountered during the development of the PC-Based model at the beginning of the project introduced delays in all subsequent activities and as a result we were only able to complete the first part of this test series—the baseline tests. We found further that the simple, proportional controller employed up to this point would not permit input on the part of the HO without causing the system to become unstable. (Other, more sophisticated types of controllers can be devised that permit human interaction, and this should be the next area to be investigated.) Hence, we completed the test series for the case of HO as crisis handler (sole control of the vehicle) and compared this with the case in which the vehicle was automatically controlled. The results of these tests—mean lateral deviation and angular deviation and their respective standard deviations—are plotted in Figure 8 and listed in Table 3. In words, an HO can learn to keep the vehicle to within about 1.5 meters of the lane centerline (using two standard deviations as the appropriate measure) and headed to within 2° of the roadway’s centerline heading. In contrast to this, the automatic controller can keep the vehicle to within .015 m and $.2^\circ$, respectively. Not surprisingly, then, even this simple (proportional) automatic controller performs significantly better than an HO. (We do see some significant learning effects with respect to the human drivers, however.) One qualification to keep in mind, however, is that humans may employ a different driving strategy than the controller in guiding the vehicle. For example, there is evidence that humans “turn into” curves, thus “smoothing them out” for a more comfortable ride. Further, a human may not in general place too high a benefit on keeping

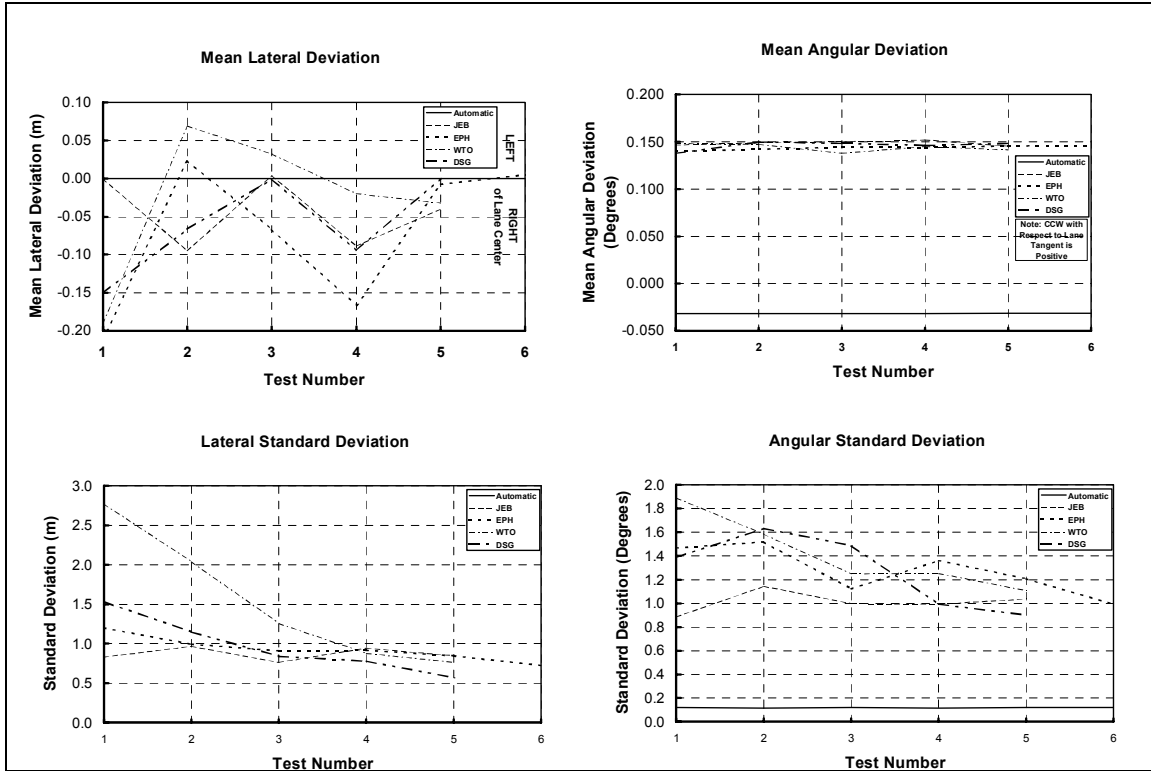


Figure 8: Quality of Reaction Test Series—Part 1

a car exactly on the lane centerline (though for these tests they were instructed to), and thus may not take corrective action if it is “a little off”, knowing that in actual driving conditions such a deviation would not be detrimental. Such factors would of course increase the deviations measured here.

Table 3: Response Summary

(data for HO is an average over the last test performed by each test subject)

	Mean		Standard Deviation	
	HO	Auto	HO	Auto
Lateral Position (m)	-.01693	.00004	.72663	.00890
Angular Deviation (deg)	.14525	-.03154	1.00862	.11797

IV. Summary, Conclusions, and Next Steps

In this project we successfully developed a versatile PC-Based simulation of an automobile driving under either automatic control or the control of a human operator (HO). Using it, we then quantified the HO’s ability to detect controller malfunctions causing the vehicle to drift out of its lane. An HO performs this task by means of establishing a “Fault Criterion”—setting a

limit for lateral deviations from the lane centerline beyond which a fault condition is assumed to have taken place. We further noted that this limit criterion is not fixed, but instead is established after weighing the cost of missing a fault condition against that of a “False Alarm”. Thus, the test administrator can influence this cost-benefit analysis by offering rewards for Hits (correctly identifying a fault) and penalties for False Alarms. We next assessed an HO’s ability to “drive” the vehicle manually, establishing his baseline performance for use in subsequent tests.

This project is necessarily a first step in the development of a more comprehensive, broader ranging study of the ways in which machine controllers can be designed to effectively utilize those human capabilities that as yet cannot be duplicated, or duplicated as well. Perhaps the more important of these include

- Image processing capabilities;
- Adaptability based on cognitive skills;
- Ability to anticipate.

(The task is made challenging by the fact that humans also respond slower, have poorer control performance—as exhibited in the Quality of Reaction Test Series—and are relatively easily distracted and fatigued.) Including the completion of the remaining tasks of this study, numerous additional questions quickly come to mind:

1. How well can the HO compensate for a controller fault after taking (partial or full) control of the vehicle when a fault is detected? What is the distribution of normal human performance in this regard?
2. With regard to the question of how well the HO can “drive” an automobile compared to an automatic controller, what is the HO’s strategy for negotiating various road geometries—for example, curves and overtaking or passing slower vehicles? What visual cues does the HO use in determining how to steer the vehicle in such situations, and how does he use these? Can such information be used to improve the operation of existing automatic controllers?
3. How is the HO’s detection/quality of reaction performance affected by
 - the manner in which he interacts with the controller
 - road geometry
 - atmospheric conditions
 - vision defects
 - the presence of other vehicles (in front and/or to the sides)

The PC-Based model can be easily modified to simulate all of these effects.

4. How best can an HO share the driving duties with an automatic controller? What type of controller/cooperative strategy works best? How would such a scheme compare with vehicle performance under full automatic control?

Finally, as was noted previously, the PC-based model was developed with the goal that it also serve as a *pre-simulator* for other types of studies. The development of “driver-assist” technologies that can be implemented on today’s automobiles before full automation takes place has received much attention recently, and constitutes one type of study for which the pre-simulator is well suited. In view of this, further development of the pre-simulator was

undertaken⁵ to provide a development environment for two of the highest priority driver-assist technologies—Adaptive Cruise Control and Automatic Lane Departure Warning systems. In essence, the goal of each is to warn an inattentive driver that that collision with the vehicle ahead or departure from the lane is imminent, but at the same time not to issue frequent false alarms that would be both annoying and cause the driver to eventually ignore the alarm. To this end, then, a “pop-up distraction screen” has now been incorporated into the visual scene, as indicated in Figure 9. Development of an Automatic Lane

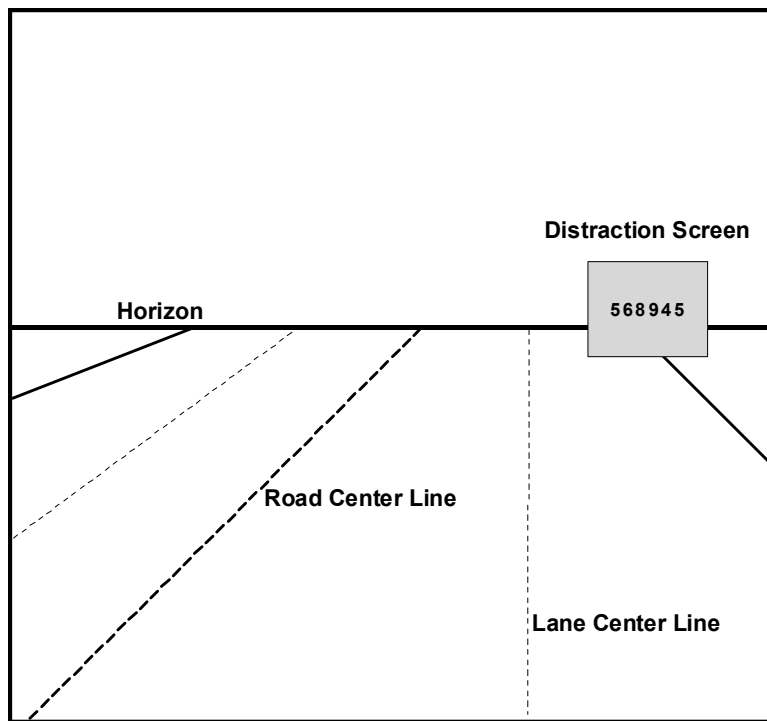


Figure 9: Visual Scene with Pop-Up Distraction Screen

Departure Warning System employing the pre-simulator could proceed as follows:

1. Without the distraction screen showing, conduct baseline driving tests (as in the Quality of Reaction baseline tests described above) in which the test subject is instructed to drive as closely as possible along the lane centerline. Record mean lateral deviation, angular deviation, their respective standard deviations, and any other pertinent measures of the test subject's ability to keep the vehicle in the middle of the lane.
2. Repeat the test with the distraction screen now visible, and instruct the test subject to monitor the screen (while continuing to drive down the middle of the lane) and indicate (by pushing a button on the steering wheel) when a particular number, word, or other symbol is shown on the screen. This symbol will be displayed at random instants, with other randomly selected symbols displayed in between. The task should be of sufficient difficulty (which can be accomplished by altering the size, location, color, texture, etc. of the screen in the visual scene; and/or the size, location, duration, color, texture, etc. of the symbols displayed) that

⁵ This work was undertaken after the project was closed, and has been performed without funding. The added capabilities described here thus go beyond the project's original deliverables.

lane departures occur with some specified frequency.

3. Algorithms to detect imminent lane departures and to warn the test subject (which can easily be incorporated into the model) can now be tested for their effectiveness and fine-tuned.

Development of an Adaptive Cruise Control system can be approached in a similar manner. Here, however, it will also be necessary to incorporate a “lead vehicle” into the visual scene, as indicated in Figure 10, as well as the ability to change vehicle speed⁶. We could then pursue development of an Adaptive Cruise Control in the following way:

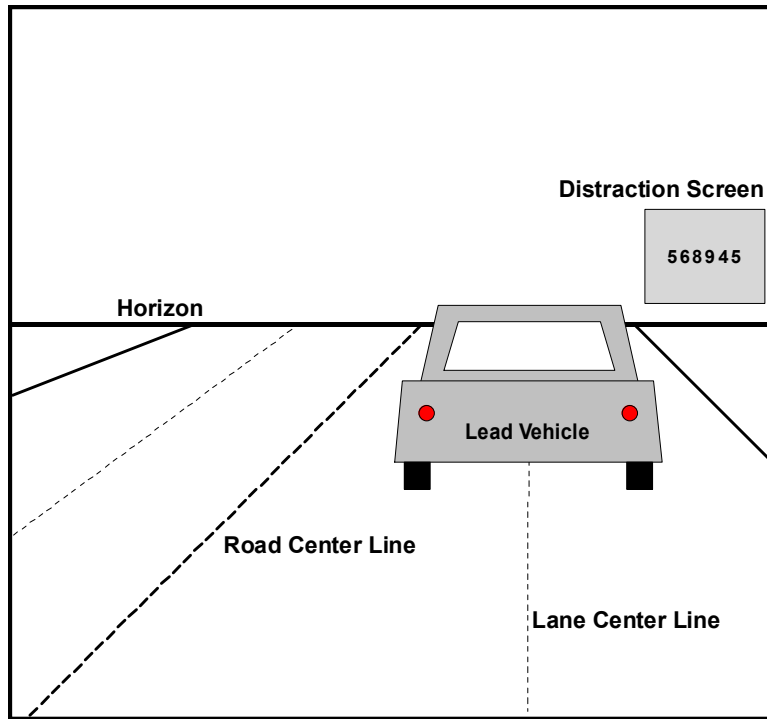


Figure 10: Visual Scene for Development of Adaptive Cruise Control

1. Without the distraction screen showing, conduct baseline driving tests (similar again to the Quality of Reaction baseline tests) where the test subject is instructed to follow the lead vehicle at a specified distance, and where his reaction to abrupt, unexpected slow downs and/or stops of the lead vehicle are initiated. Record the distance to the lead vehicle, deceleration rate, and any other pertinent measures of the test subject’s ability to react safely to the lead vehicle’s behavior.
2. Repeat the test with the distraction screen now visible, and instruct the test subject to monitor the screen (in addition to the lead vehicle) and indicate when a particular number, word, or other symbol is shown on the screen. The task should again be of sufficient difficulty that rear-end collisions occur with some specified frequency.
3. Algorithms to detect dangerous encroachment upon the lead vehicles and to warn the test subject can now be tested and fine-tuned.

⁶ These features have not been incorporated yet but could without great difficulty.

Other studies to which the pre-simulator could be applied are:

1. Simulation of new controller designs and driving procedures;
2. Development of specifications for
 - operator visual and driving capabilities,
 - road and vehicle performance capabilities;
3. Customization of cooperative strategies for individual drivers (for example, vision-impaired drivers);
4. Performance comparisons;
5. Re-creation of actual driving scenarios for further analysis;
6. Identification and quantification of the important results that more elaborate and expensive test procedures must provide
7. Development of a cognitive model combining the HO and controller into a complete system, including
 - development of the best cooperative strategy
 - incorporation of the full range of human capabilities into the driving model
 - tailoring special control strategies to fit individual drivers.