Predicting the Effects of In-Car Interfaces on Driver Behavior using a Cognitive Architecture

Dario D. Salvucci
Nissan Cambridge Basic Research
Four Cambridge Center
Cambridge, MA 02142
+1 617 374 9669
dario@cbr.com

ABSTRACT

When designing and evaluating "in-car" user interfaces for drivers, it is essential to determine what effects these interfaces may have on driver behavior and performance. This paper describes a novel approach to predicting effects of in-car interfaces by modeling behavior in a cognitive architecture. A cognitive architecture is a theoretical framework for building computational models of cognition and performance. The proposed approach centers on integrating a user model for the interface with an existing driver model that accounts for basic aspects of driver behavior (e.g., steering and speed control). By running the integrated model and having it interact with the interface while driving, we can generate *a priori* predictions of the effects of interface use on driver performance. The paper illustrates the approach by comparing four representative dialing interfaces for an in-car, hands-free cellular phone. It also presents an empirical study that validates several of the qualitative and quantitative predictions of the model.

Keywords

Driving, cognitive models, in-car interfaces, cognitive architectures, ACT-R, cellular phones.

INTRODUCTION

For most user interfaces on the desktop, the interface is the primary focus of a user's attention; the user (or group of users) typically sits at the interface engrossed in the interaction with few external concerns or disturbances. In contrast, for many user interfaces away from the desktop, the interface is only one of several tasks a user must manage, and sometimes is secondary to other primary tasks. Numerous examples of such interfaces arise in the domain of driving: drivers frequently interact with in-car interfaces for radios, climate controllers, navigation aids, etc., but vehicular control and safe navigation clearly take top priority. For this reason, the design and evaluation of in-car interfaces requires understanding of not only the driver's interaction with the interface but also the effects that this interaction may have on driver performance.

This paper proposes a novel approach to predicting the effects of in-car interfaces on driver behavior using a cognitive architecture. A cognitive architecture is both a theory of human cognition and performance and a framework for developing computational models of behavior [see 2, 12, 16]. Cognitive architectures have been widely employed to model human-computer interaction [e.g., 3, 8, 10] and human behavior more generally [see 2, 16]. In particular, cognitive architectures provide at least two major benefits for the purposes of the proposed approach. First, architectures incorporate well-tested parameters and constraints on cognitive and perceptual-motor processes, and any model developed in an architecture necessarily inherits these parameters and constraints. This allows architectural models to generate *a priori* predictions about behavior and performance — for instance, the time needed to type a key or the likelihood of forgetting a memorized fact. Second, architectures allow for modular models that can be combined into larger integrated models. This also lends predictive power to the models in that an architecture provides predictions of how smaller models would interact as part of an integrated model.

Our approach utilizes these aspects of cognitive architectures to predict the effects of in-car interfaces on driver behavior. The approach centers on the integration of two architectural models that we call the *user model* and the *driver model*. The user model represents the behavior needed to interact with a given in-car interface. Cognitive architectures provide a rigorous computational framework to facilitate task analysis and specification of the user model. The driver model represents the cognitive and perceptual-motor processes needed to drive. For this purpose we utilize an existing model [14] developed in the ACT-R cognitive architecture [2] that successfully captures driver behavior in naturalistic highway navigation. The cognitive architecture allows us to take the user and driver models and combine them into an integrated model that performs both tasks together. By running this integrated model, we can generate predictions concerning both the effects of interface use on driver performance and the effects of driving on interface use.

To illustrate the approach, the study in this paper examines an increasingly popular type of in-car interface, namely the cellular telephone or "cell phone." Cell phones have received attention from researchers and media alike

Salvucci, D. D. (2001). Predicting the effects of in-car interfaces on driver behavior using a cognitive architecture. To appear in *Human Factors in Computing Systems: CHI 2001 Conference Proceedings*. New York: ACM Press.

concerning the possible effects of phone use on driver behavior [e.g., 1, 5, 11, 13, 15]. This attention has also discussed the differences between hand-held cell phones and "hands-free" phones that mount in a fixed position and need not be held while conversing. In this paper we highlight four representative interfaces for dialing hands-free cell phones and compare the effects of these interfaces on driver behavior. In particular, we perform a straightforward task analysis to develop a user model for each dialing interface and then integrate the user models with the driver model to generate the desired predictions. Unlike most work concerning cell phones, our study examines the effects of dialing the phone and does not address effects of conversation on the phone.

To validate the predictions of the integrated model, the paper also describes an empirical study in which human drivers performed the exact task given to the model — namely, driving with occasional use of a cell phone. The experiment was conducted in a fixed-base driving simulator with a full car interior and a realistic front-view projection [4]. Results from the experiment confirm several of the qualitative and quantitative predictions of the integrated model.

TESTED CELL PHONE DIALING INTERFACES

We begin with a description of the cell phone interfaces to be compared with the integrated model and tested in the validation experiment. While most previous studies of cell phone use while driving have emphasized the effects of conversation or similar cognitive tasks [e.g., 1, 5, 11], some have looked more closely at the effects of dialing. Most of these studies investigated only one type of interface, namely manual dialing on a standard phone keypad [e.g., 1, 11, 13]. A few studies have also compared manual dialing with voice dialing in which the driver simply states the name of the person or place to be called [e.g., 15]. However, the studies did not explore further options for dialing interfaces and also did not compare baseline performance on the interfaces alone against performance during driving.

We extend these studies by comparing four cell phone dialing interfaces that represent plausible alternatives with different benefits and drawbacks. All four interfaces assume a hands-free cell phone that is mounted to the dashboard on or above the center console within easy reach of the driver. This position was successfully tested in an earlier study [15], which also found that the exact position of the cell phone does not have a significant effect on driver behavior assuming the same general positioning. The interfaces also assume that the phone has the phone keypad shown in Figure 1 with two special buttons: **Power**, which powers up and activates the phone for use; and **Send**, which initiates the connection for manual dialing (detailed below).

Using this basic hands-free phone, we designed four dialing interfaces with the following procedures for calling a party:

- *Full-manual* : press **Power**, key in the party's full phone number, and press **Send**
- *Speed-manual*: press **Power**, key in the party's single-digit "speed dial" number, and press **Send**

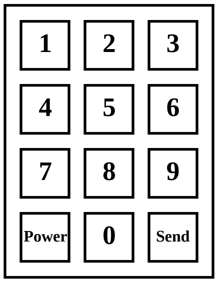


Figure 1: Dialing interface keypad.

- *Full-voice* : press **Power**, speak the party's full phone number, confirm the recognition of the number as repeated by the phone, and confirm the connection as spoken by the phone
- *Speed-voice* : press **Power**, speak the party's name, confirm the recognition of the name as repeated by the phone, and confirm the connection as spoken by the phone

Thus, the four interfaces break down according to two factors: *full* vs. *speed* dialing, indicating whether the full phone number or a shorter speed code is entered; and *manual* vs. *voice* dialing, indicating whether the dialing input is keyed in or spoken. All interfaces require initial activation of the phone through pressing of the **Power** button. The manual interfaces require that the user press **Send** to initiate connection. The voice interfaces require that that user confirm voice recognition and connection by listening to the phone repeat the desired name/number and say "Connecting…". Table 1 illustrates sample dialing sequences for each interface when calling "home" at the number 555-4283 assuming a speed dial number of "2".

Table 1: Sample dialing sequences for each tested interface when calling "home" at 555-4283.

Full-Manual	*Full-Voice*
Press **Power**	Press **Power**
Press **5, 5, 5, 4, 2, 8, 3**	Say **5, 5, 5, 4, 2, 8, 3**
Press **Send**	Listen for **5, 5, 5, 4, 2, 8, 3**
	Listen for "Connecting…"
Speed-Manual	*Speed-Voice*
Press **Power**	Press **Power**
Press **2** (speed number)	Say "**home"**
Press **Send**	Listen for "**home"**
	Listen for "Connecting…"

It is essential to note the differences among these interfaces with respect to the demands placed on the driver. The two speed-dialing interfaces require less time than the two full-dialing interfaces. According to typical desktop performance metrics, this aspect would tend to favor the speed-dialing interfaces; however, while driving, it is not clear whether the few extra seconds needed for full dialing would distract drivers any more than speed dialing if the driver is diligently attending to the road. In addition, the manual interfaces occupy more of the driver's motor processing and visual attention (assuming looks to the phone are occasionally necessary), while the voice interfaces occupy the driver's speech processing and aural attention. While previous studies have suggested that voice interfaces may sometimes be favorable [15], it is again not obvious which tradeoffs are desirable when effects on driver behavior are considered.

THE INTEGRATED MODEL

Given our four dialing interfaces, our goal is to make *a priori* predictions about behavior for the interfaces and compare them in terms of their usability while driving. One proven method of making such predictions for desktop and similar interfaces is by means of cognitive modeling. For example, Gray, John, and Atwood [8] modeled behavior in a telephone operator task using the GOMS framework [7], and Kieras, Wood, and Meyer [10] modeled the same task in the EPIC framework. However, such standalone models for our dialing interfaces would have somewhat limited usefulness: while they would predict user behavior in the interface alone, they could not predict the interactions of this behavior with the primary driving task. In other words, we require a model that integrates behavior in the dialing task with behavior in the driving task and makes predictions about the interactions between the two tasks.

This section describes how an existing model of driver behavior can be integrated with models of dialing behavior to generate the desired predictions and thus compare the interfaces in terms of their effects on driver performance. The driver model [14] is implemented in the ACT-R cognitive architecture [2] with perceptual-motor extensions [ACT-R/PM: 6] and accounts for several aspects of driving behavior including lower-level vehicular control (e.g., steering and speed control) as well as higher-level situation awareness and decision making. The interface user models are also implemented in ACT-R, facilitating the final integration of these models with the driver model.

The ACT-R cognitive architecture has been used successfully to model behavior in domains ranging from arithmetic to analogy to scientific discovery [see 2]. ACT-R posits that procedural knowledge in the form of production rules (or condition-action rules) operate on declarative knowledge represented as factual chunks. The architecture provides numerous mechanisms that account for individual variability and learning, and also for interaction with the external environment through perceptual-motor modules [6]. Space constraints preclude a detailed discussion of the architecture in this paper; please consult [2] for more information.

User Models

We begin with a description of the ACT-R user models for each of the four dialing interfaces. Our approach to creating these user models employs a straightforward task analysis to determine the cognitive, perceptual, and motor processes necessary to use the interface. The descriptions of the dialing procedures for each interface, as illustrated in Table 1, provide the basic framework for the models. However, they fail to specify the cognitive processes involved, and also do not describe aspects of visual and aural attention needed for the interfaces.

Table 2: Outline of the dialing interface user models. The marking (*) indicates where the user model passes control to the driver model (which returns control after some incremental execution).

Full-Manual	*Full-Voice*
Recall phone number	Recall phone number
Move hand to phone (*)	Move hand to phone (*)
Attend to phone	Attend to phone
Press **Power** (*)	Press **Power** (*)
Attend to phone	Move hand to wheel (*)
Recall block of numbers	Recall block of numbers
Press digit	Say digit (*)
(*repeat until last number*)	(*repeat until last number*)
Press last digit (*)	Say last digit (*)
(*repeat until last block*)	(*repeat until last block*)
Press **Send** (*)	Recall block of numbers
Move hand to wheel (*)	Listen for number (*)
	(*repeat until last number*)
	Listen for last number (*)
	(*repeat until last block*)
	Listen for "Connecting…" (*)
Speed-Manual	*Speed-Voice*
Recall speed number	Move hand to phone (*)
Move hand to phone (*)	Attend to phone
Attend to phone	Press **Power** (*)
Press **Power** (*)	Move hand to wheel (*)
Attend to phone	Say name (*)
Press speed number	Listen for name (*)
Press **Send** (*)	Listen for "Connecting…" (*)
Move hand to wheel (*)	

Augmenting the basic procedures with the necessary cognitive and attention processes, Table 2 outlines the models for each dialing interface. Each line in the table represents a separate ACT-R production rule that performs the specified action. The models utilize a blocked, or chunked, representation of a phone number as three blocks — namely, digit positions 1-2-3, 4-5, and 6-7 (assuming a seven-digit phone number) — such that blocks of digits are output together. The models assume that visual attention is drawn to the cell phone before any typing takes place (which may not be the case for very experienced users). They also assume that the right hand moves to the phone before any keying and moves back to the steering wheel when keying is completed.

One interesting aspect of the model that becomes an issue for multitasking is when to pass control such that other tasks (namely driving in the integrated model) can be interleaved if desired. We assume that the user model actively passes control rather than being interrupted. To this end, we select certain rules to pass control as indicated in the table with the (*) marking. The selected rules are rules that, after they execute, may require some period of time before the next rule executes — for instance, if the next rule uses the same modality (e.g., saying digits) or requires external input (e.g., waiting for the voice interface to repeat its recognition). This policy serves to allow the driver model to update control during times in which the user model would be in waiting. An exception arises for keying on the phone keypad: keying requires (most) drivers to look at the keypad, and thus it is more efficient to key in several numbers at one time, as facilitated by the blocked representation of phone numbers.

We should note two implementation details for these user models. First, ACT-R (or more accurately ACT-R/PM [6]) has extended mechanisms only for typing on a standard computer keyboard and not for a phone keypad. However, because the phone keypad is very similar to the numeric keypad on a keyboard, we simulate phone keying as typing on the numeric keypad. Second, we assume that the voice-dialing interfaces repeat back the spoken name or numbers at a rate of 300 ms per item (to represent non-continuous speech).

Driver Model

To test the user models in the context of driving, we require a model of driver behavior with which the user models can interact. We have recently developed such a model for navigating a naturalistic highway environment with multiple lanes and automated traffic [14]. The model interfaces with a simulated roadway, acquiring visual information through ACT-R's perceptual extensions [ACT-R/PM: 6]. This process acquires information needed for lateral and longitudinal control (i.e., steering and speed control) as well as situation awareness and decision making (e.g., speeds and positions of other cars as needed for lane changing). Based on this information, the model calculates desired control variables and outputs its control through specialized motor modules for steering, acceleration, and braking. Please refer to [14] for a full description of the model and empirical validation of its behavior.

For the purposes of this paper, the most important aspect of the driver model is the sequential nature of its processing: the model accesses information and updates control iteratively on repeating cycles. When control is the model's only task, the model updates control fairly frequently and thus produces relatively smooth behavior. However, the model sometimes has other driving-related tasks to perform, such as checking the blind spot before changing lanes. These tasks reduce the frequency with which the model can update control and thus control may not be as accurate. In this paper, we explore whether and how the model's behavior degrades when given secondary user interface tasks such as dialing a cell phone.

The complexity of a full highway environment is too rich for this initial study of cell phone use while driving; it introduces a wide variety of other factors (e.g., negotiating around other cars) that make it more difficult to assess the effects of phone use alone. Therefore we simplify the environment by having the model navigate a one-lane road without traffic and at a constant speed. Thus, the model need only be concerned with lateral control through steering, and we can focus our examination on the effects of cell phone use on driver's ability to keep the car centered on the road.

Integrated Model

The integrated model embodies both the driver model and each of the user models for the different dialing interfaces. The ACT-R architecture allows for straightforward integration of the various models because of the fact that each model implements a particular task goal: the driver model processes a goal of type *drive-car*; and the user models process goals of type *dial-full-manual*, *dial-speed-manual*, *dial-full-voice*, and *dial-speed-voice*. During normal driving, the integrated model simply processes the *drive-car* goal. When a dialing task is introduced, the integrated model runs through a single cycle of the *drive-car* goal and then decides with some probability (set arbitrarily at 0.5) whether to cede control to the dialing goal. As mentioned, the dialing goal then cedes control back to the driving goal at predetermined intervals (as noted in Table 2). To capture the fact that drivers put off secondary tasks in safety-critical situations, the integrated model does not cede control to the dialing goal if it sees that the road is too unstable; specifically, it only cedes control if both the near and far points used for steering [see 14] are moving with an angular velocity less than a certain threshold, set by exploratory observation at 2.3°/s.

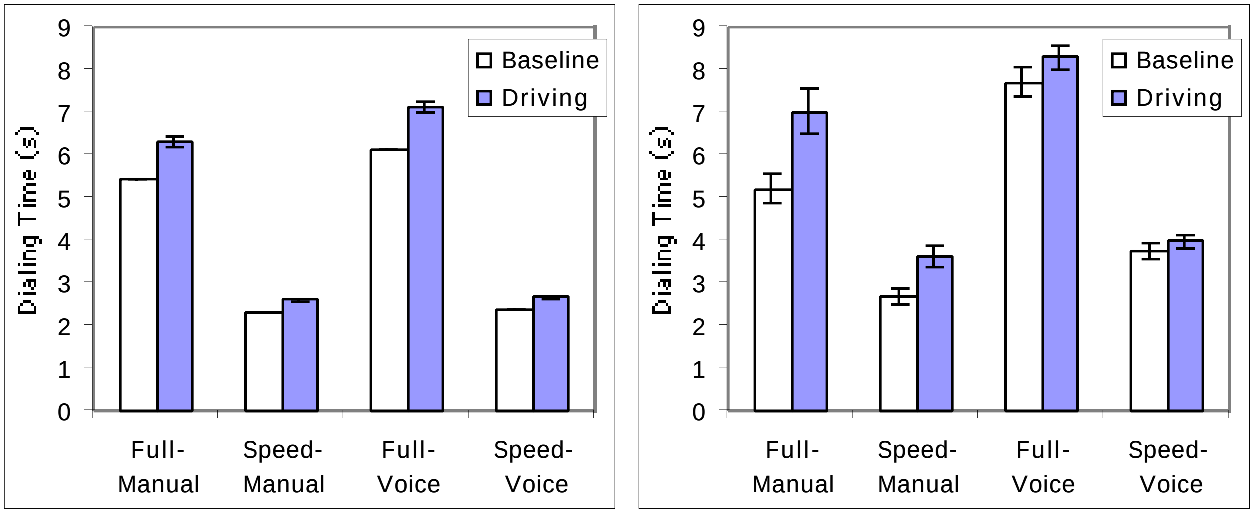
The ACT-R architecture incorporates a number of parameters that can affect model predictions. Because the primary goal of this paper is to generate *a priori* predictions of the interaction between dialing and driving, we want to avoid adjusting any parameter values to fit empirical data for the integrated model as is sometimes done for cognitive modeling applications. However, the driver model was originally developed to fit various aspects of driver behavior and thus incorporates already estimated parameter values. With this in mind, we simply incorporate the parameter values for the driver model into the integrated model, and thus the user model portions of the integrated model inherit all these parameter values. The only exception to this rule is a parameter[1] for the driver model that controls how aggressively the model centers the vehicle: whereas the original model drove on a full highway where lane centering was critical, this model drives on a single-lane road where centering is less critical. For this reason we halved the value of this parameter.

INTEGRATED MODEL PREDICTIONS

Given the integrated model, we can generate predictions about how the dialing interfaces affect driver behavior. In addition, we can generate predictions about how the interfaces differ with respect to baseline performance times in the absence of driving. This section begins with a description of the model simulation process to generate

---

[1] Namely, the parameter c_3 that scales θ_{near} [see 14].



(a) Model Predictions (b) Empirical Data

Figure 2: Mean dialing times and standard errors for baseline performance without driving and while driving.

predictions and continues with an examination of these predictions with respect to two factors, dialing time and vehicle lateral position.

Model Simulations

The predictions in the following sections were generated in two sets of three simulation runs. In the first set of simulations, the model dialed the cell phone while driving on a single-lane straight highway at a constant speed of approximately 60 miles/hour. In the second set of simulations, the integrated model dialed the cell phone alone, without driving, so that we could gauge dialing behavior with no external disruptions. For each simulation the model dialed eight numbers in each interface for a total of 32 dialing trials per simulation.

For the driving simulations, the model interacts with a simulated environment that provides it with visual information about the current scene — for instance, the near and far points in the center of the road that guide steering [see 14]. The model then produces a behavioral protocol analogous to that produced by human drivers in a driving simulator. This protocol can include steering, acceleration, braking, eye movements, and all environment information. For the current study we focus on steering and its downstream consequences for lateral position.

The driving simulations began with the car centered on the road and accelerating up to the constant velocity. After 20 seconds of normal driving, the model was given the additional goal of dialing using one of the four interfaces. When the dialing task was completed, the model reverted back to normal driving. This pattern was repeated at 20-second intervals until all 32 dialing trials were completed.

Dialing Time Predictions

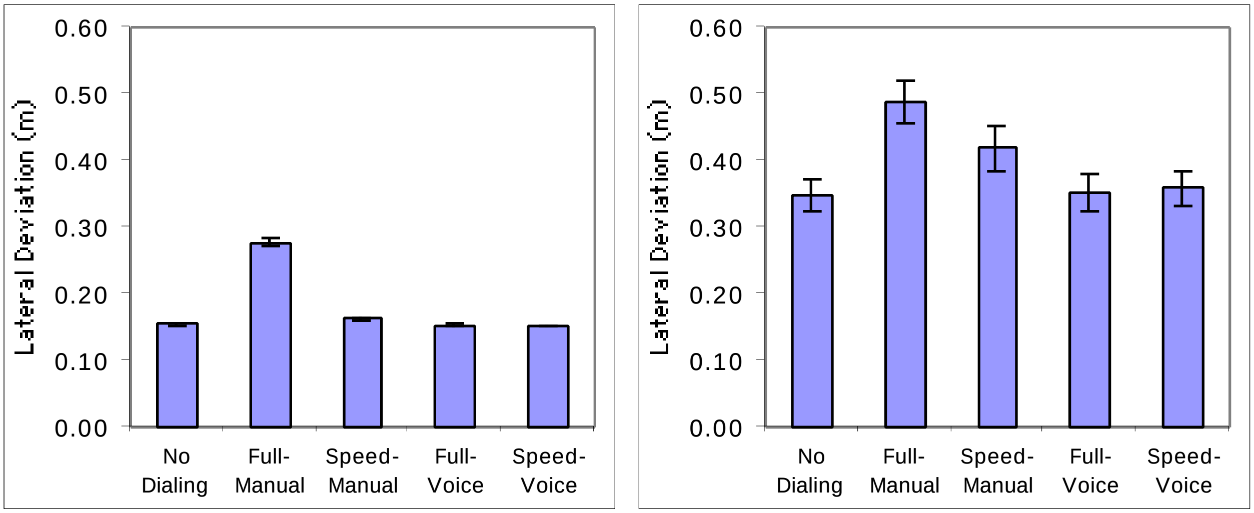
Dialing time represents the total time needed to complete a dialing sequence, from the initial request to call a specified party to the final connection (i.e., pressing **Send** for the manual interfaces, hearing "Connecting…" in the voice interfaces). Figure 2(a) shows the dialing times both alone and while driving for each of the four interfaces. (We will examine the (b) portions of the figures in the next section on empirical validation.) Not surprisingly, the full-manual and full-voice interfaces required the most time to dial. These two interfaces had fairly similar dialing times, as do the speed-manual and speed-voice interfaces. All interfaces required more time while driving since additional processing was required by the driver model to control the vehicle. However, because vehicle control usually occurs during waiting times for the user model (e.g., between initiation of spoken words), driving adds only small amounts of time to the total dialing times.

Lateral Position Predictions

Lateral position is a measure of the position of the driver's vehicle within its lane, thus measuring the performance (or accuracy) of the driver's behavior in keeping the car centered on the road. Because dialing can affect behavior not only during the dialing task but also immediately afterwards — for instance, if the driver notices a lane departure after dialing and must make a steering correction — we included a period of five seconds after dialing as part of a dialing segment. The time period of five seconds approximates the time needed to make a lane correction after deviating from the lane center [9].

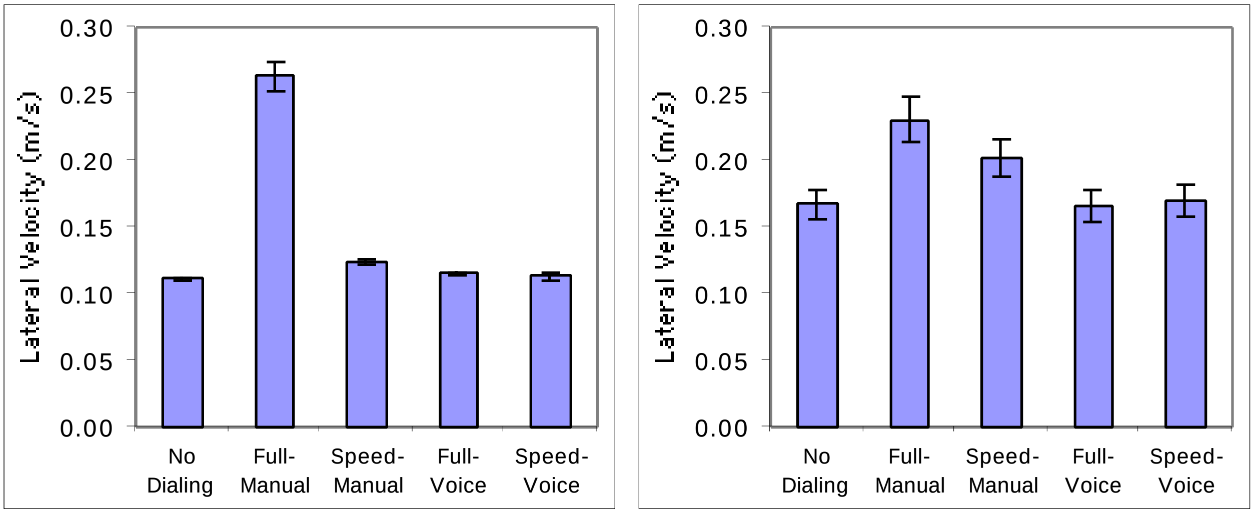
The first measure we consider is *lateral deviation* from the center of the lane, computed as the root-mean-squared deviation from the center of the lane. The mean lateral deviations for normal driving and for each interface are shown in Figure 3(a). During normal driving without a dialing task, the model exhibited a mean deviation of approximately .15 m from lane center. While dialing with the full-manual interface, the model exhibited significantly greater deviation near .28 m, $p<.01$. The deviation for the speed-manual interface was also significantly different than for normal driving, $p<.05$. The deviations for the voice interfaces were not significantly different, $p>.10$.

The other measure we examine is *lateral velocity* which serves as a measure of vehicle stability, as shown in Figure



(a) Model Predictions (b) Empirical Data

Figure 3: Mean lateral deviations and standard errors for model and data.



(a) Model Predictions (b) Empirical Data

Figure 4: Mean lateral velocities and standard errors for model and data.

4(a). As for lateral deviation, the full-manual dialing interface produced the largest effect of all tested conditions, and the speed-manual interface produced velocities marginally significantly different from normal driving, $p<.10$. The lateral velocities for the voice-dialing interfaces were not significantly different from normal driving, $p>.10$.

The model predictions suggest that the full-manual interface has a large effect on driver behavior, the speed-manual interface has a small but significant effect, and the voice interfaces have no significant effect. Two factors in the full-manual interface, namely the length of dialing and the need for visual attention, seem to combine to produce its effect. Both manual interfaces occupy more visual attention than the voice interfaces because the model needs to look at the cell phone keypad occasionally while dialing. However, the full-manual interface takes many more keypresses than the speed-manual interface (nine versus two, respectively) and thus distracts visual attention for a significantly longer time period. It is also notable that the voice interfaces do not produce significant effects. These interfaces require aural rather than visual attention (to listen to and confirm the dialed numbers), which does detract a small amount from the model's attention to the driving task; however, the model can more easily switch between the different types of attention (visual for driving, aural for dialing) than switch visual attention between the two tasks.

EMPIRICAL VALIDATION

Given the model's *a priori* predictions, we wish to validate the predictions with empirical data and determine both the validity of our study of dialing interfaces and also the validity of our general approach to predicting effects of in-car interfaces. The following empirical study aims to provide the necessary data for this validation. The study takes the environment and tasks given to the model and exposes human drivers to the same environment and tasks in our fixed-base driving simulator [4]. In this way we can compare human behavior directly to that of the model and determine the qualitative and quantitative similarities and differences in these behaviors.

Experiment

Subjects

Twelve participants with more than two years of driving experiment completed the experiment. One participant was omitted because of inexplicably erratic driving. The 11 remaining participants varied in age from 19-32.

Driving Environment

The experiment was conducted in the Nissan CBR fixed-base driving simulator [4] with a phone keypad mounted on the center console within easy reach of the driver. The simulator uses the front cab of a Nissan 240SX that has been instrumented to collect all relevant data from the driver and also provide visual and aural feedback to the driver (e.g., a working speedometer and realistic motor sounds). The simulated scene is projected in front of the cab, resulting in a field of view of approximately 70°. For this experiment, the scene comprised only a single-lane, straight textured road with a width of 3.66 m (the American standard) and a small wall 2.75 m from each road edge.

Procedure

Participants were first given an overview of the experiment and the driving simulator. With the participant in the driver seat and the experimenter in the passenger seat, the experimenter explained the four dialing interfaces and had the participant run through a sample call for each interface. To mimic a real phone's voice recognition for the voice-dialing interfaces, the experimenter repeated back the name or numbers. The participant then practiced driving for approximately one minute with instructions to keep the car centered on the road as much as possible.

The participant then performed a pretest in which s/he dialed the various interfaces in the absence of driving. For each dialing trial, the experimenter asked the participant to call one of four parties using a particular interface; the names and numbers of these parties were collected from the participant before the experiment so that they could easily recall the phone numbers. As the number was dialed, the experimenter keyed the start and end of the dialing task into the simulator computer, which then noted the events in the experiment protocol. The pretest comprised one trial for each party and each interface for a total of 16 dialing trials.

The main driving session was completely analogous to that performed by the model in the simulations. The driver tapped the accelerator to start the car accelerating up to its constant velocity. At 20-second intervals, the experimenter asked the driver to call a party with a particular interface. After calling each party twice with the same interface, driving was interrupted to allow for a short break and a review of the next interface to be used. Thus the participant performed 32 dialing trials while driving.

After the main driving session, the participant also performed a posttest in which the pretest procedure was repeated to counterbalance any effects of learning. These 16 dialing trials, combined with the 16 in the pretest, totaled 32 dialing trials in the absence of driving.

Dialing Time Results

Figure 2(b) shows the baseline and driving dialing times for the human drivers in the experiment. We conducted a repeated-measures analysis of variance (ANOVA) on each set of dialing times with three factors: category (baseline vs. driving), modality (manual vs. voice) and length (full vs. speed). All three main effects were significant, $p<.01$. The category-modality, category-length, and modality-length interactions were also significant, $p<.01$, but the three-way interaction was not, $p>.10$. The full-dialing interfaces had higher dialing times than the speed-dialing interfaces. Again the dialing times during driving were slightly higher than the baseline times because of the occasional interruptions for controlling the vehicle. These empirical results matched well with the overall pattern of model predictions in Figure 2(a), $R=.95$.

Lateral Position Results

Figure 3(b) shows the mean lateral deviations for the human drivers, as compared to those predicted by the model in Figure 3(a). The effect of modality for the empirical data was significant, $p<.001$, while the effect of length and its interaction with modality were not significant, $p>.10$. The full-manual deviations were significantly different from normal driving, $p<.05$, and the speed-manual deviations were marginally significantly different, $p<.10$. The voice-dialing interfaces did not differ from normal driving, $p>.10$. Overall the human drivers showed greater deviations than the model. The model predicted this general qualitative pattern, $R=.92$, but underpredicted the effects of speed-manual dialing. Notably, as the model successfully predicted, the human drivers showed no effects in the voice-dialing interfaces.

Figure 4(b) and 4(a) show the mean lateral velocities for the human drivers and model, respectively. The effect of modality for the human drivers was significant, $p<.001$, and the interaction was not significant, $p>.10$. In contrast with the lateral deviation results, the main effect of length was significant, $p<.05$, due to the increased velocities for full dialing. Nevertheless, as for lateral deviations, the model captured the general trend, $R=.88$, except for an underprediction for speed-manual dialing.

To summarize, these results concur with earlier results showing that voice interfaces can produce smaller effects on lateral position than manual interfaces in certain circumstances [e.g., 15]. Our study extends previous work in that it also suggests that the length of the input as measured by number of digits pressed or said aloud seems not to have a significant effect on lateral deviation.

DISCUSSION

This paper outlines a novel method for generating *a priori* predictions of the effects of in-car interface use on driver performance. The user models come from a straightforward task analysis of the dialing interfaces, and when integrated with the existing driver model, they inherit all the parameter settings and mechanisms of this model. Comparing the predictions of the integrated model with the empirical results, the model closely fits the dialing time profiles of the human drivers with no parameter fitting.

The model also reproduces some of the qualitative effects present in the empirical data, including the larger effects of the full-manual interface and the lack of significant effects for the full-voice and speed-voice interfaces. However, the model fails to capture some quantitative effects in the data, especially for the speed-manual interface. Nevertheless, we regard this result as a very encouraging starting point for illustrating the power of the proposed approach.

The main benefit of this approach lies with the predictive power that arises from integrating models of several tasks. It is noteworthy that the user models closely predict baseline dialing times when not driving, considering that the models use only default mechanisms of the ACT-R architecture (and its perceptual-motor extensions) and parameter settings inherited from a model of a very different domain (i.e., the driver model). However, the real power of an integrated-model approach is the ability to predict interactions between task behaviors, especially when one or more of the tasks are safety-critical as in driving. Using the existing driver model, a practitioner interested in evaluating a new interface can quickly devise a user model for the interface and integrate it with the driver model, which in turn generates predictions of the interface's interactions with driving.

We should note a few of the limitations of this work that could be addressed in future research. First, each user model as described here embodies only a single strategy and does not account for the strategy variability we observed in the experiment — for instance, drivers feeling the keypad to avoid looking for keys. Second, the driver model itself is relatively young and requires more testing to ensure its robustness and generality. Third, because the driver model is implemented in ACT-R, the users models must also be implemented in this architecture. Extensions in which a modeler could integrate user models from other frameworks such as GOMS [7] could allow for more flexibility and generality in using the approach for both research and practical applications.

ACKNOWLEDGMENTS

Thanks to Kristen Macuga for help in designing and running the experiment and to Ahna Girshick, Akio Kinoshita, Kristen Macuga, and several anonymous reviewers for insightful suggestions and comments.

REFERENCES

1. Alm, H., & Nilsson, L. (1994). Changes in driver behaviour as a function of hands-free mobile phones – A simulator study. *Accident Analysis & Prevention*, *26*, 441-451.
2. Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.
3. Anderson, J. R., Matessa, M., & Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction*, *12*, 439-462.
4. Beusmans, J., & Rensink, R. (Eds.) (1995). Cambridge Basic Research 1995 Annual Report (Tech. Rep. No. CBR-TR-95-7). Cambridge, MA: Nissan Cambridge Basic Research.
5. Brookhuis, K. A., De Vries, G., & De Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention*, *23*, 309-316.
6. Byrne, M. D., & Anderson, J. R. (1998). Perception and action. In J. R. Anderson & C. Lebiere (Eds.), *The Atomic Components of Thought* (pp. 167-200). Hillsdale, NJ: Lawrence Erlbaum Associates.
7. Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
8. Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: A validation of GOMS for prediction and explanation of real-world task performance. *Human-Computer Interaction*, *8*, 237-309.
9. Hildreth, E. C., Beusmans, J. M. H., Boer, E. R., & Royden, C. S. (in press). From vision to action: Experiments and models of steering control during driving. *Journal of Experimental Psychology: Human Perception and Performance,* in press.
10. Kieras, D. E., Wood, S. D., & Meyer, D. E. (1997). Predictive engineering models based on the EPIC architecture for a multimodal high-performance human-computer interaction task. *ACM Transactions on Computer-Human Interaction*, *4*, 230-275.
11. McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis & Prevention*, *25*, 259-265.
12. Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
13. Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialing task. *Ergonomics*, *42*, 1015-1037.
14. Salvucci, D. D., Boer, E. R., & Liu, A. (in press). Toward an integrated model of driver behavior in a cognitive architecture. *Transportation Research Record*.
15. Serafin, C., Wen, C., Paelke, G., & Green, P. (1993). Development and human factors tests of cell phones (Tech. Rep. No. UMTRI-93-17). Ann Arbor, MI: The University of Michigan Transportation Research Institute.
16. VanLehn, K. (Ed.) (1991). *Architectures for intelligence: The twenty-second Carnegie Mellon symposium on cognition*. Hillsdale, NJ: Erlbaum.