Modeling Driver Distraction from Cognitive Tasks

Dario D. Salvucci (salvucci@mcs.drexel.edu)
Department of Mathematics and Computer Science, Drexel University
3141 Chestnut St., Philadelphia, PA 19104

Abstract

Driver distraction has become a critical area of study both for research in investigating human multitasking abilities and for practical purposes in developing and constraining new in-vehicle devices. This work utilizes an integrated-model approach to predict driver distraction from a primarily cognitive secondary task. It integrates existing models for a sentence-span task and driving task and investigates two methods in which the resulting model can perform multitasking. Model predictions correspond well qualitatively to two of three measures of driver performance as collected in a recent empirical study. The paper includes a discussion of the potential for building multitasking models in the framework of a production-system cognitive architecture.

Introduction

Computational cognitive modeling continues to mature rapidly as an area for both theoretical advances in understanding cognition and practical advances in developing intelligent technology. Cognitive modeling has grown from addressing only simple cognition in basic psychological tasks to capturing integrated cognitive, perceptual, and motor processes in large-scale complex, dynamic tasks (e.g., Chong, 1998; Jones et al., 1999). This paper investigates the application of cognitive models to an extremely common yet complex task: driving. Driving involves the continual multitasking of a number of subprocesses that make use of the driver's cognition, perception, and motor movements. This rich spectrum of necessary skills makes driving an ideal task in which to investigate how humans execute complex tasks and how computational models can represent and predict the multitasking behavior in these tasks.

Driver Distraction and Cognitive Modeling

One particular aspect of driver multitasking that has received enormous attention from media and researchers alike is that of "driver distraction" -- namely, the effects of multitasking while performing some secondary task. Numerous studies have now found that performing primarily perceptual-motor tasks while driving (e.g., dialing a cellular phone) can impair driver performance (e.g., Alm & Nilsson, 1995; McKnight & McKnight, 1993). These studies generally conclude, perhaps not surprisingly, that pulling a driver's visual attention from the road and/or her hand(s) off the steering wheel degrades the driver's ability to maintain a central lane position, follow a lead car with a constant headway, or react to potential hazards. Such studies have subsequently led to legislation at all government levels to restrict the use of potentially distracting secondary-task devices. While driver distraction is generally associated with effects on perceptual-motor processes, researchers have also reported that "cognitive distraction" arising from purely cognitive secondary tasks can adversely affect driver performance (e.g., Alm & Nilsson, 1995). These results are not fully conclusive and seem to depend highly on the secondary task as well as the driving situation; nevertheless, it is clear that even purely cognitive tasks can impact driver performance in certain situations.

To better understand driver behavior and the sources of driver distraction, researchers have attempted to develop integrated driver models that capture driver behavior in a computational manner (e.g., Aasman, 1995). These models provide insight into the sources of distraction by elucidating the exact processes by which a driver attends to the external environment, processes this information cognitively, and then reacts and manipulates the environment. In addition, the computational models may be used to generate predictions about the effects of distraction on driver performance; for instance, the ACT-R driver model (Salvucci, Boer, & Liu, 2001) has been integrated with various models of cell-phone dialing to predict the impact of dialing on lane-keeping performance (Salvucci, 2001; Salvucci & Macuga, 2001). However, this previous work has addressed only primarily perceptual-motor secondary tasks with little cognitive component (like cell-phone dialing); to date, no models have demonstrated the ability to represent and generate "cognitive distraction" from primarily cognitive tasks.

Modeling "Cognitive Distraction"

This paper describes the first attempt to predict cognitive distraction with a computational cognitive model. This work employs the same methodology as in previous work for perceptual-motor distraction, namely the "integrated model approach" based in a cognitive architecture (see Salvucci, 2001). Cognitive architectures are computational frameworks that incorporate built-in, well-tested parameters and constraints on human cognitive and perceptual-motor abilities. This work focuses on a particular architecture, ACT-R (Anderson & Lebiere, 1998), that represents factual knowledge as declarative chunks and procedural knowledge as condition-action "production rules". For our purposes, the ACT-R architecture has two important benefits: (1) it facilitates reuse and integration of multiple behavioral models, and (2) it provides built-in interfaces and default parameters that facilitate *a priori* predictions of real-world

metrics of human performance (e.g., reaction times, keystrokes, eye movements). The integrated model approach takes advantage of these benefits to incorporate a model of secondary-task behavior with the ACT-R driver model to predict the effects of executing the secondary task on the primary driving task.

The initial demonstrations of this approach (Salvucci, 2001; Salvucci & Macuga, 2001) examined a primarily perceptual-motor task, namely dialing a cellular phone using different modalities (e.g., manual button input versus speech input). The work showed that an integrated driving-dialing model predicted degraded steering performance for the modalities that required the driver to look at the cell phone (i.e., manual dialing), thus drawing visual attention away from the roadway. The work presented here generalizes the previous work in two important ways. First, although it utilizes the same methodology to predict driver distraction, it predicts distraction from a primarily cognitive task — namely, a sentence-span task that involves rehearsal and recall of a sequence of words. Second, unlike the previous work, it makes use of an existing model for the secondary task (with some necessary adaptation) as well as an existing model for the primary driving task, thus demonstrating the importance and benefits of model re-use.

This paper begins with a review of the driving and secondary tasks modeled here, namely those from the empirical work of Alm and Nilsson (1995) showing effects of the sentence-span task on driver car-following performance. It then provides an overview of the integrated model approach incorporating existing models of both the driving and secondary tasks, including two methods of performing explicit multitasking between the individual task models. Finally, it compares the model's predictions with Alm and Nilsson's empirical results and discusses the broader implications of the methodology to studying multitasking in the framework of a cognitive architecture.

The Sentence-Span and Driving Tasks

The task and empirical results that will be used to validate the model predictions are taken from Alm and Nilsson (1995). Their study aimed to show exactly those effects that we are attempting to model, namely effects of cognitive secondary tasks on driver performance. For the purposes of this paper, we would like to recreate this task for the integrated model as closely as possible to facilitate later comparison between model and empirical results.

Sentence-Span Task

Alm and Nilsson (1995) employed a sentence-span task that involves the processing of sentences and the recall of words in these sentences (see Daneman & Carpenter, 1980). The task comprises two stages. In the first stage, drivers listened to five sentences of the form "X does Y" — for instance, "The boy brushed his teeth" or "The train bought a newspaper." They would also report after each sentence whether the statement was generally sensible. In the second stage, drivers were asked to state the last word of each sentence in order. For instance, for the sentences "The boy brushed his teeth" and "The train bought a newspaper," the

driver would report "yes" and "no" after each sentence (respectively) and would then report the memorized list "teeth,", "newspaper," etc. The sentence-span task itself involves two concurrent activities, namely judging of sentence sensibility and memorization (and rehearsal) of final words. When combined with driving, the task puts a substantial cognitive load on drivers as they attempt to integrate the tasks.

Driving Task

As a realistic scenario in which to test interaction with the sentence-span task, Alm and Nilsson (1995) used a car-following task where the lead vehicle would sometimes perform unsafe maneuvers and leave the driver in a "safety-critical" situation. During the normal stages of the task, the lead vehicle maintained a 75 m headway distance from the driver's vehicle. Occasionally, the lead car braked suddenly with a deceleration of 4 m/s^2 until its speed reached 50 km/hr (or until a maximum of 5 s of deceleration), then accelerated at 3 m/s^2 until its speed reached 90 km/hr. The original study also included non-safety-critical situations in which the lead vehicle would indicate a right turn and eventually turn off the road; their analysis does not examine these situations in detail and they are not discussed further.

The Alm and Nilsson study provided three metrics by which they measured driver performance: (1) reaction time to the braking event, measured as the time lapse between the start of the event and the driver's initial depression of the brake; (2) lateral deviation, measured as the root-mean-squared error of the driver's vehicle position to the center of the lane; and (3) headway distance, measured as the distance between the driver's vehicle and the lead vehicle. The results section will compare the model's predictions to the empirical results from human drivers for all three metrics.

Empirical Study

Alm and Nilsson's (1995) empirical study included a total of 40 participants in two experimental groups: a *task group* that occasionally performed the sentence-span task while driving, and a *control group* that did not perform the task. In both groups, each driver negotiated four safety-critical situations in which the lead vehicle would brake suddenly. The timing of the events was randomized to either near the start or near the end of the span task (in the task group) such that drivers could not predict when the events would occur.

The driving task was performed in a high-fidelity driving simulator to give participants as realistic an impression of real-world driving as possible. The simulator included a moving-base system (based on a Saab 9000 with manual transmission), wide-angle visual system, vibration generation, and temperature regulation. The driving environment comprised a simulated 80 km two-lane highway (one lane in each direction) with oncoming traffic in the opposite lane. The highway had a very low curvature so that steering down the roadway was relatively straightforward even at high speeds. The sentence-span task was performed through an Ericsson hands-free telephone mounted on the instrument panel to the right of the steering wheel. Drivers needed only press one key to activate

(answer) the phone at the start of the task, and given practice with the phone, drivers could easily activate the phone without looking. Sentences were presented by means of a tape recorder, and driver responses were recorded on a second tape recorder.

The results of the empirical study will be discussed in a later section to facilitate comparison with model predictions. It should be noted that the original study also included both younger and older drivers to demonstrate the interactions of cognitive distraction with age. This paper only addresses the data from the younger drivers (mean age 29); the existing driver model used in this paper has been validated with data from younger drivers, and thus we expect the model to better account for the younger-driver data from the original study.

The Integrated Task Models

To model and predict the interaction of the sentence-span and driving tasks, this work utilizes the "integrated model" methodology employed in previous work (see Salvucci, 2001): Given an existing model of driver behavior, we develop or acquire a model of behavior in the secondary task, integrate the two models to perform multitasking, and finally run the integrated model to generate behavioral data. One critical element of this integration is the potential for generating *a priori* predictions — that is, rather than fitting the model to data by manipulating parameters, we carry over defaults and parameter settings from existing models and immediately use them in the integrated model. In addition, we benefit from re-use of models that have been rigorously tested in other studies. These and related issues will be discussed further in later sections. This section describes the individual task models as well as the two versions of the final integrated model.

Sentence-Span Model

The model for the sentence-span task comes from Lovett, Daily, and Reder (2000), who developed an ACT-R (Anderson & Lebiere, 1998) process model as part of their investigation of individual differences in working memory. Although the original model does not literally perform the sentence-span task, it does perform a closely related task called the MODS task in which people read strings of letters while memorizing final digits for later recall. The original model provides three critical components that are re-used in the sentence-span model: (1) the positional representation used to encode memorized items, (2) production rules that perform rehearsal of memorized items, and (3) production rules that retrieve and report the items in sequence. Interested readers can refer to Lovett, Daily, and Reder (2000) for a more detailed description of these components.

Given this core model, the sentence-span model required two modifications: (1) the addition of production rules to encode a sentence and decide whether it is sensible, and (2) the incorporation of perceptual-motor productions to hear and speak words (rather than read and type characters as in the MODS task). Table 1 shows the production rules in the final sentence-span model and indicates those rules taken from the original MODS model. While there are a number of new rules in this model, it should be noted that the first six deal with particulars of the sentence-span task involving hand movement and encoding of speech, and the final two non-MODS rules simply terminate the articulation and recall goals. The process of confirming whether or not the sentence is sensible is not modeled in any detail, but rather the model simply assumes that this process happens during the listening productions and signals a confirmation by firing the Confirm-sentence rule. In addition, the model assumes that each sentence component (subject, verb, object) requires one second of speech time.

Table 1: Sentence-span model production rules, with markings for whether they are original MODS-model rules and whether they pass control to driving in the Single-Step (SS) and Group-Step (GS) models.

Production Rule	MODS	Passes Control	
		SS	GS
Move-hand-to-phone		x	x
Press-button		x	x
Move-hand-to-wheel		x	x
Hear-sentence-subject		x	x
Hear-sentence-verb		x	x
Hear-sentence-object		x	x
Confirm-sentence		x	x
Create-memory	x	x	x
Rehearse-memory	x	x	
Done-articulate		x	x
Recall-span	x	x	
No-recall	x	x	
Say-item	x	x	
Next-item	x	x	x
Done-recall		x	x

Driver Model

The model of driver behavior is an ACT-R model that integrates control, monitoring, and decision making to navigate highway environments with traffic (Salvucci, Boer, & Liu, 2001). For control, the model employs a two-level model of steering that uses a "far point" on the road to guide predictive steering and a "near point" on the road to center the vehicle. For monitoring, the model encodes its surrounding environment using simulated eyes to maintain situation awareness. For decision making, the model checks the current situation and decides when to perform maneuvers such as lane changes. Thus, the driver model incorporates both lower-level perception and action for vehicle guidance and higher-level cognition for awareness and decision making. This driver model has been shown to account for a number of aspects of human highway driving, including nearing the inner curb during curve negotiation and switching gaze to the destination lane at the start of a lane change (see Salvucci, Boer, & Liu, 2001). Also, as mentioned earlier, the driver model has been employed to predict the effects of driver distraction from cell-phone dialing in different modalities (Salvucci, 2001a, 2001b; Salvucci & Macuga, 2001).

The complexities of the driver model make it infeasible to describe here in any level of detail. However, it is worthwhile to highlight two critical aspects of the model that are essential to the endeavor of predicting driver distraction. First, because of its implementation in the ACT-R architecture, the model is constrained to a serial line of cognitive processing. Thus, the cognitive integration of perception, action, and decision making is done in a serial loop: the model encodes relevant perceptual variables, processes these variables, then issues motor commands to manipulate the environment. When secondary tasks are added into this main loop, they naturally have some impact on the frequency with which the updating processes can occur, and thus can affect driver performance. Second, the driver model interacts with a simulated driving environment and generates a full behavioral protocol, as would a human driver navigating the same environment in a driving simulator. This faithfulness to predicting real-world data facilitates rigorous and straightforward comparison between model predictions and empirical results.

Integrated Model

In general, integration of multiple models in a production system such as ACT-R is rather straightforward: we can simply add the sentence-span memory structures to the driver model. However, two challenges arise that must be dealt with. First, the integrated model must decide how to multitask between the two component models. As in previous applications of this methodology, there does not yet exist a rigorous model of multitasking that we can employ, but we can use reasonable heuristics to guide us. Multitasking in the integrated model is performed explicitly (instead of preemptively) in that each model must pass control to the other, presumably at a fairly frequent interval. Because driving is the primary task, we are most concerned about when the secondary task model (i.e., the sentence-span model) will cede control back to driving. This paper explores two approaches for attacking this problem. The conservative approach would only allow the secondary task to fire a single production, then immediately cede control back to driving. A less conservative (though still fairly conservative) approach would allow small logical groupings of production firings to occur before passing control. These approaches were used to develop two versions of the model termed the Single-Step (SS) and Group-Step (GS) models, respectively. Table 1 indicates which productions pass control for each model. While every rule is marked for the SS model, the GS model allows certain rules to continue: the Rehearse-memory rule that rehearses memorized items in rapid succession, and the threesome of rules that combine to retrieve and report a memorized item. The choice of marked rules for the GS model is admittedly somewhat arbitrary, but at least in part guided by introspection as to how humans would perform this task. Further development on rigorous models of multitasking will help to formalize these choices in future work.

The second major challenge for model integration, not to mention model development on the whole, is the setting of parameter values. ACT-R, like similar architectures, has a number of "settable" parameters; however, all parameters have default recommended values that have withstood the test of time in modeling throughout the community. Nevertheless, the original MODS model posed an interesting problem in that it activated several learning mechanisms (e.g., learning of chunk base-level activations and production strengths) that were deactivated in the driver model. Because the MODS model had undergone more rigorous parameter testing with detailed data, it was decided to incorporate its parameter values into the integrated model, thus overriding the driver model's global settings. Fortunately (and perhaps surprisingly), this decision had no apparent adverse effects on the normal operation of the driver model, which proved rather robust to the different parameter settings and activated learning mechanisms.

Model Simulations

The driver model was made to interact with a driving simulation that mimicked the critical elements of the Alm and Nilsson (1995) car-following task. A total of 15 simulations were run: 5 runs in the *No-Task* condition without a secondary task, 5 runs in the *Task-SS* condition with the Single-Step model performing the secondary task, and 5 runs in the *Task-GS* condition with the Group-Step model performing the secondary task. Each simulation generated a detailed behavioral protocol at a rate of roughly 13 Hz including all relevant control and environmental data as well as marks for the start and end of the braking events.

Model Predictions and Empirical Results

We can now compare the model predictions with Alm and Nilsson's (1995) empirical results. It should be emphasized that the present study does not involve typical parameter estimation for fitting the model to data; rather, it centers on *a priori* predictions by simply integrating the models, running simulations, and checking the results. The goal of the study is thus to predict the effects of the secondary task on driver performance primarily in qualitative terms and, secondarily, in quantitative terms as much as possible.

Brake Reaction Time

The first and more important aspect of driver performance examined is drivers' *brake reaction time*, or the time lapse between the start of the lead vehicle's braking and the initiation of braking by the driver. Figure 1(a) shows the reaction times (means and standard deviations) predicted by the model for the No-Task, Task-SS, and Task-GS conditions. While the reaction time for the no-task condition was approximately 2.5 s, the reaction times for both task conditions were significantly higher at roughly 2.9 s, and thus the model predicts a significant impact of the secondary task on drivers' braking reaction.
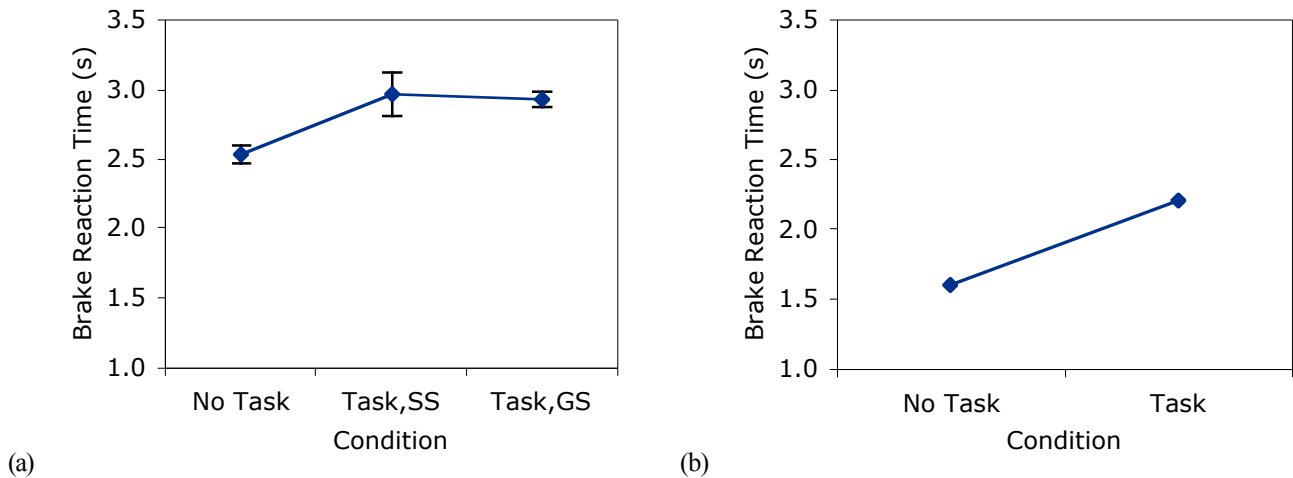
(a)

(b)

Figure 1: Brake reaction times for (a) model simulations and (b) empirical data.

Figure 1(b) shows the empirical results for brake reaction time. These results also show a clear (and significant) task effect, with an increase of reaction time from 1.6 s without the task to 2.2 s with the task. The model and empirical results therefore correspond well qualitatively. Quantitatively, the model predictions are roughly .7–.9 s too high; this discrepancy may be attributed to the fact that the model uses only distance of the lead vehicle to determines how it accelerates and decelerates, whereas the human drivers could also attend to the lead vehicle's brake lights, providing the latter with additional cues to initiate braking.

Lateral Deviation

The second aspect of performance is one of the most common measures for driver distraction studies, namely the *lateral deviation* of the driver's vehicle — defined as the root-mean-squared error of the vehicle's center with the central position in the lane. Figure 2 shows the model predictions for lateral deviation in the three conditions.
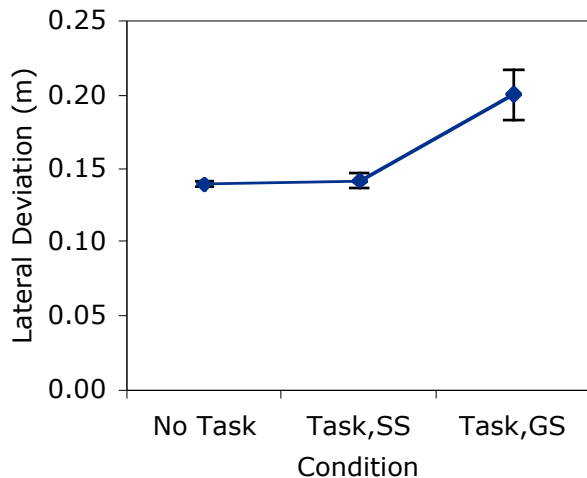

Figure 2: Lateral deviation for model simulations.

Interestingly, the SS model predicts no effect of secondary task on lateral deviation. However, the GS model predicts quite a large effect of approximately 50 cm.

Alm and Nilsson (1995) do not report specific numbers for lateral deviation; however, they do report a statistical analysis on these data that found no significant task effect on lateral deviation (against their original hypothesis). The predictions of the SS model thus support their results, demonstrating how closely interleaved multitasking can, in certain situations, have no significant effect on lateral deviation. On the other hand, the predictions of the GS model show that less conservative, "grouped" multitasking can draw cognitive attention away from the driving task enough to create a significant effect.

Headway Distance

The third aspect of performance is *headway distance*, or the distance between the driver's vehicle and the lead vehicle. While headway was maintained at 75 m during normal conditions, the lead vehicle's braking would greatly reduce this headway until the driver has a chance to react. Figure 3(a) shows the model predictions for the *minimum* headway distance, a measure of how close the driver's vehicle came to the lead vehicle. In all three conditions, the model exhibited a minimum headway of approximately 35 m.

Figure 3(b) shows the minimum distances reported in the empirical study. Here there is a clear task effect: the headway decreases significantly in the presence of the secondary task. Thus, the model predictions are not supported for the distance-headway measure. It seems that although the model clearly reacts later in the task condition, it also compensates for the late reaction by braking harder, thus eliminating any potential task effect.

Conclusions

The SS model's *a priori* predictions matched two of the three measures qualitatively, correctly predicting an effect on reaction time but no effect on lateral deviation. Given the ambiguity in the driver-distraction literature on when such
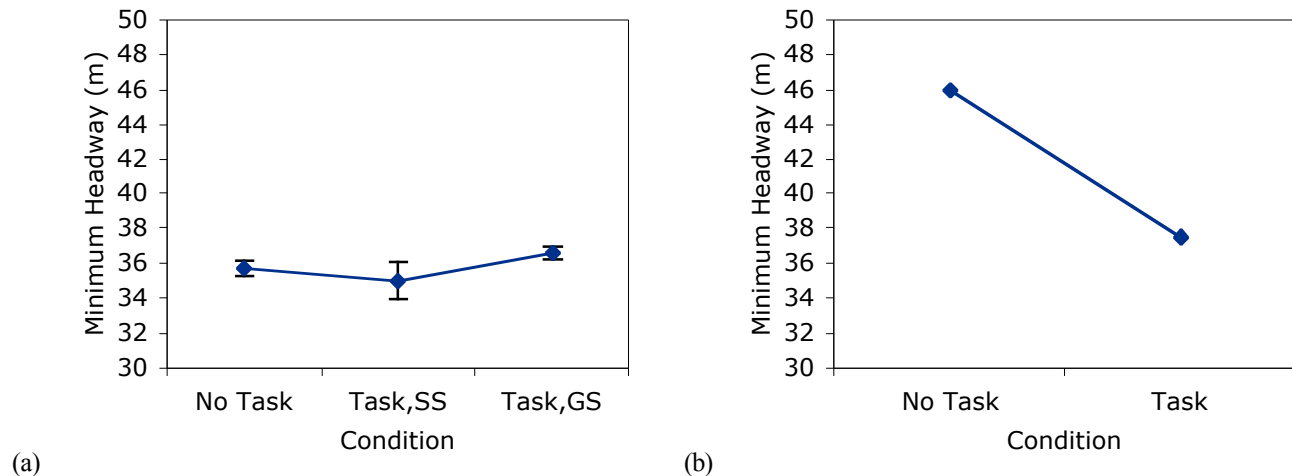
Figure 3: Minimum headway distance for (a) model simulations and (b) empirical data.

effects may occur, this is a strong result that demonstrates the model's ability to predict distraction for certain measures and not others. However, the model did not predict the effect on minimum headway, perhaps due in part to the fact that the headways were large enough that human drivers felt no dire need to closely maintain headway.

The GS model's predictions were not as good, failing to predict the absence of an effect on lateral deviation. Its failing indicates one shortcoming of this work: although the integration of models is mostly straightforward, there remain too many degrees of freedom with respect to how models can and should multitask. Combating this problem requires a more rigorous treatment of multitasking, and cognitive architectures such as ACT-R show promise in being able to account for such a process. In particular, architectures provide an opportunity to handle multitasking at the "software level" through new models implemented as production rules and/or at the "hardware level" through changes to the architecture's inner mechanisms. Recent models of complex dynamic tasks, though not yet the comprehensive models required for the long term, have already demonstrated good ability in capturing some aspects of multitasking (e.g., Chong, 1998; Jones et al., 1999).

As a related point, cognitive architectures also have the substantial benefit of facilitating re-use of models, parameters, and other components from one model to another. This study exhibits this property primarily in the re-use of two existing models for predicting distraction. However, it also opens the door to predicting numerous other aspects of behavior. For instance, Lovett et al.'s (2000) treatment of their MODS model includes an investigation of how ACT-R's *W* parameter can represent individual differences in working memory capacity. Because their work addresses mechanisms fundamental to the architecture, it can carry over directly into further investigations of the effects of capacity differences on driver distraction or even just on driving itself. This ability to share ideas and mechanisms across domains offers enormous explanatory and predictive power to architectural models in new and existing domains of study.

Acknowledgments

References

Aasman, J. (1995). *Modelling driver behaviour in Soar*. Leidschendam, The Netherlands: KPN Research.

Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis & Prevention*, 27, 707-715.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.

Chong, R. S. (1998). Modeling dual-task performance improvement: Casting executive process knowledge acquisition as strategy refinement. Doctoral Dissertation, Department of Computer Science and Engineering, University of Michigan.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.

Jones, R. M., et al. (1999). Automated intelligent pilots for combat flight simulation. *AI Magazine, 20*, 27-42.

Lovett, M. C., Daily, L. Z., & Reder, L. M. (2000). A source activation theory of working memory: Cross-task prediction of performance in ACT-R. *Journal of Cognitive Systems Research*, 1, 99-118.

McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis & Prevention*, 25, 259-265.

Salvucci, D. D. (2001). Predicting the effects of in-car interface use on driver performance: An integrated model approach. *International Journal of Human-Computer Studies*, 55, 85-107.

Salvucci, D. D., Boer, E. R., & Liu, A. (2001). Toward an integrated model of driver behavior in a cognitive architecture. *Transportation Research Record*, 1779.

Salvucci, D. D., & Macuga, K. L. (2001). Predicting the effects of cell-phone dialing on driver performance. In *Proceedings of the Fourth International Conference on Cognitive Modeling*. Veenendaal, Netherlands: Universal.