Hierarchical Control and Driving

Nathan Medeiros-Ward
University of Illinois at Urbana–Champaign

Joel M. Cooper
University of Utah and Precision Driving Research, Salt Lake City, Utah

David L. Strayer
University of Utah

We manipulated primary task predictability and secondary task workload in the context of driving an automobile. As the driving task became less predictable (by adding wind gusts), more attention was required to maintain lane position. When drivers concurrently engaged in a secondary cognitive task in the windy driving condition, attention was diverted from driving and the ability to maintain lane position was degraded. By contrast, when the driving task was predictable (no wind), lane maintenance actually improved when a secondary cognitive task diverted attention from driving. These data provide evidence for a hierarchical control network that coordinates an interaction between automatic, encapsulated routines and limited capacity attention.

Keywords: multitasking, hierarchical control, distracted driving

In this article, we provide evidence that complex skilled behavior as diverse as playing a musical instrument, typing a manuscript, and driving an automobile are supported by a hierarchical control network that coordinates an interaction between automatic, encapsulated routines and limited capacity attention. Following Logan and Crump (2011), we assume that complex skills are goal directed. For example, musicians do not play instruments by accident, nor do drivers drive their cars by happenstance. Despite the fact that complex skills are goal directed, many times skilled performers do not know how they achieve high levels of performance, and thinking about it tends to impair performance on the task (Tapp & Logan, 2011). To help account for this puzzling characteristic of behavior, Fodor (1983) argued that separate control systems underlie skilled performance. One system is under intentional control and is easily brought into conscious awareness. The other system is automatic and operates outside of awareness.

This division was initially described in terms of a hierarchy with higher and lower levels of control. For example, Shaffer (1976) found that musicians often used two levels of control for musical performance. One level would monitor a song and notes to be played and another level would control the execution of finger movements. When a person first learns how to play the piano, the higher control level would be required for processing the song and notes as well as monitoring the hands to depress the proper keys. In these instances, the higher level is engaged and acting directly to achieve the task of playing the instrument, and performers are keenly aware of their performance. With practice, the higher level of control is not needed to accomplish the task. Instead, some of the work can be offloaded to the lower level of control, which would then directly influence performance. Finally, with extensive practice, the lower control level can become encapsulated such that it does not require higher level involvement. When this happens, performance on the task is characterized as automatic and requires minimal attention or effort (Shiffrin & Schneider, 1977). However, if the task environment changes to become unpredictable, higher level attentional control would again be required for successful performance on the task.

The notion of hierarchical control for skilled performance has advanced over the years and has been vetted using other tasks outside of musical performance. Recently, Logan and Crump (2009) used the metaphor of control loops rather than levels to describe complex typing and suggested that skilled typists rely on outer and inner loops of control for their complex performance. Specifically, the outer loop was responsible for selecting the words to be typed while the inner loop was responsible for controlling the execution of individual keystrokes. It is interesting that when participants were instructed to attend to the individual keystrokes, performance declined. The authors explained this by suggesting that the new task requirements caused the outer loop to monitor the output of the inner loop, and this additional monitoring disrupted encapsulated inner loop processing.
It is important to note that Logan and Crump’s (2009) sample consisted of expert typists. Certainly, when first learning to type, attention must be allocated to what to type (i.e., thoughts or words on a screen), as well as how to type (i.e., finger movements). This suggests a high level of involvement from the outer loop, which is resource demanding and effortful (Kahneman, 1973). With practice, typists can start to offload some of the work to the inner loop to accomplish the necessary keystrokes. Finally, with extensive practice, the task can be controlled directly by the inner loop, which is more automatic and requires minimal attention for efficient performance.

For expert performers, certain parameters of the environment must remain consistent and predictable for their performance to remain automatic and under the purview of the inner loop of control. An expert guitarist accustomed to playing one type of guitar will seem like a novice when switching to a different string instrument with very different characteristics. Likewise, exchanging a traditional QWERTY keyboard (Noyes, 1983) for a Dvorak keyboard (Cassingham, 1986) would make the typing task more resource demanding. This would require outer loop processing for successful finger movements. In these examples, the novel configurations require outer loop processing to accomplish tasks that the inner loop had previously been able to handle autonomously.

When the testing environment remains predictable but attention is nonetheless allocated to the task, complex skills can be disrupted. As previously mentioned, when expert typists pay attention to keystrokes, performance declines (Logan & Crump, 2009, 2011; Tapp & Logan, 2011). These disruptive effects have also been shown in other tasks, suggesting that they are general characteristics of hierarchical control. For example, Beilock, Carr, MacMahon, and Starkes (2002) had experienced golfers focus on swinging their clubs and experienced soccer players focus on kicking the ball and found that performance declined significantly compared with the performance of novices completing the same tasks. When athletes performed their respective skill in a predictable environment without focused attention on the mechanics of the act, they were successful. When they paid attention to the low-level mechanics, their performance declined, although for novices, focusing on the mechanics was beneficial.

According to hierarchical control theory (HCT), the aforementioned findings highlight a key difference between the outer and inner loops of control, which is the focus of the current research. The outer loop requires attention for successful task performance, whereas the inner loop suffers when attention is allocated to its processing. This can be formalized with two novel premises of HCT that lead to the predictions tested in the current research.

**Premise 1:** Performance based on the outer loop should get better with more attention allocated to the task and get worse with less attention allocated to the task.

**Premise 2:** Performance based on the inner loop should get better with less attention allocated to the task and get worse with more attention allocated to the task.

The predictions above suggest that altering task predictability and the allocation of attention to the task can provide a key test of HCT. In the following section, we suggest that the hierarchical control network is also important for key aspects of driving an automobile and that the continuous nature of steering the vehicle affords the opportunity to directly test the predictions outlined above.

### Driving and Distraction

Driving a motor vehicle is a complex, goal-directed skill that places a high demand on cognitive and motor processes (Groeger, 2000). As driving has become ubiquitous, so too has the prevalence of in-vehicle devices that divert attention from the task of driving. In fact, the National Highway Traffic Safety Administration (NHTSA, 2009) estimates that 25% of all crashes are related to distracted driving. Despite the hazards associated with distraction, more than two thirds of people surveyed reported using an in-vehicle device, such as a cell phone, while driving (AAA Foundation for Traffic Safety, 2009).

What is particularly interesting about cognitive sources of distraction is that they have produced a counterintuitive pattern across the different components of driving. On the one hand, when drivers are distracted, response times are slower and drivers are less able to detect novel or unexpected events in the driving environment (Strayer & Drews, 2007; Strayer & Johnston, 2001). This would be expected if these aspects of driving require the same resources that are needed for a secondary task, such as a cell phone conversation (Kahneman, 1973). On the other hand, examination of lane maintenance has paradoxically found improvements in performance with cognitive distraction. For example, as cognitive workload increases, lane maintenance improves (Atchley & Chan, 2011; Becic et al., 2010; Beede & Kass, 2006; Brookhuis, de Vries, & de Waard, 1991; He & McCarley, 2011; Horrey & Simons, 2007; Horrey & Wickens, 2006; Janson & Merat, 2005; Knappe, Keinath, Bengler, & Meinecke, 2007; Liang & Lee, 2010; Östlund et al., 2004; Reimer, 2009). Heretofore, there has been no adequate explanation for this complex pattern of driving behavior.

Because lane maintenance is an automatic skill for experienced drivers (Dingus et al., 2006; Michon, 1986), it is ideal for testing HCT. It is plausible that lane maintenance can be under the purview of the inner loop processing for experienced drivers in predictable driving conditions. To test the prediction in Premise 1, it is necessary to make lane maintenance depend on outer loop processing. One way to accomplish this is to make the driving environment less predictable, for example, by introducing crosswinds. Crosswinds represent an unpredictable external force pushing the vehicle out of the desired lane of travel. If drivers engage in a secondary task while driving in windy conditions, attention will be diverted from lane maintenance and performance should suffer.

To test the prediction in Premise 2, driving conditions need to be predictable. For experienced drivers, this could be brought about by driving on well-maintained roadways without crosswinds. If drivers engage in a secondary task without wind, performance should improve. Taken together, these predictions provide a critical test of HCT in the domain of driving an automobile. More important, the prediction based on HCT will help to differentiate it from another influential model of skilled performance: Adaptive Control of Thought-Rational (ACT-R; Anderson & Lebiere, 1998; Salvucci & Beltowska, 2008).

Salvucci (2006) adapted a version of ACT-R to predict driving behavior, highlighting the fact that ACT-R has built in perceptual
and motor modules that works in parallel in a way that resembles complex human behavior. In addition, there is a cognitive processor that receives all information from the perceptual module and is also in charge of all that goes into the motor module. Although these modules operate in parallel, the cognitive processor operates sequentially. As a result, when driving becomes unpredictable, the cognitive processor would be required to switch between monitoring the upcoming roadway, perceiving the strength and direction of the wind, and making adjustments to steering inputs.

In addition to predicting performance in unpredictable driving conditions, ACT-R has been used to predict dual-task driving performance. Salvucci (2006) argued that when drivers engage in secondary tasks, the cognitive processor must switch between the secondary tasks and driving, which results in suboptimal driving performance. ACT-R predicts that lane maintenance should be degraded when drivers are cognitively distracted because the cognitive processor must switch between the secondary cognitive tasks and lane maintenance in a serial fashion. This latter prediction is in direct contrast with the predictions of HCT when performance is based on inner loop processing.

We manipulated primary task predictability and secondary task workload to dissociate the outer and inner control loops. It was predicted that as driving became less predictable because of the crosswinds, the outer loop would be required for lane maintenance. If secondary cognitive tasks were added, attention would be diverted from driving, thereby impairing lane maintenance (i.e., increasing lane position variability). In contrast, it was predicted that in predictable driving conditions (i.e., without wind), the inner loop would be sufficient for performance. If secondary cognitive tasks are added, attention would be diverted from driving and this would lead to improvements in lane maintenance (i.e., decreased lane position variability).

Method

Participants

Twenty-seven participants (11 men) with normal vision and valid licenses were recruited from the University of Utah participant pool. They were between 19 and 43 years old (mean age = 25 years) and were fluent in English. Participants had their normal amount of sleep and caffeine prior to the study. They also reported having been a licensed driver for 7 years and driving an average of 10,000 miles per year.

Materials and Design

Driving performance data were collected using a fixed-base driving simulator. The roadway was a straight three-lane highway, and speed was fixed at 68 mph, simulating cruise control to eliminate any variability caused by speed fluctuation.

The levels of cognitive workload came from a delayed digit recall n-back task developed by Mehler, Reimer, and Dusek (2011). Entropy-based measures were derived from information theory to calibrate the level of unpredictability associated with the crosswind. As the wind becomes more unpredictable, staying in the lane becomes more difficult, and more attention must be allocated to the driving task.

Procedure

After providing informed consent, participants completed a warm-up scenario to allow for adaptation to the driving simulator. Participants also completed a standardized training protocol on the delayed digit recall n-back task until they achieved at least 85% accuracy on all levels. The three levels of cognitive workload were single task, 0-back, and 2-back. In the single-task condition, participants drove without a secondary cognitive task. In the 0-back and 2-back conditions, participants were presented with auditory lists of numbers ranging from 0 to 9 in four sets of 10 randomized sequences. For the 0-back condition, participants were instructed to report out loud the number they had just heard. For the 2-back condition, participants were instructed to say out loud the number that was presented two trials earlier in the sequence. For all conditions, participants were instructed to respond as accurately as possible, and their responses were recorded for later analysis.

Three levels of wind entropy were created using the sum of three sine waves (Andersen & Ni, 2005) and entropy estimates from information theory. In the low entropy condition, there was no wind. In the medium entropy condition, there was a constant lateral wind (40 mph) and a single gust (25 mph and .077 Hz). In the high entropy condition, there was a constant lateral wind (40 mph) and three gusts (all at 25 mph and .077 Hz, .059 Hz, and .032 Hz, respectively). The wind gusted laterally 175° for the first and last third of the drive and 5° for the middle third of the drive. These levels of entropy created a steady increase in uncertainty for participants trying to maintain a central lane position (Coifman & Wickerhauser, 1992; Donoho & Johnstone, 1994). Entropy was 5.91 bits in the medium and 23.61 bits in the high wind condition.

After the secondary task training, participants completed all nine driving scenarios in an order counterbalanced across participants in one 45-min session (5 min per scenario). Participants were instructed to drive in the middle lane of a three-lane highway with their hands on the steering wheel at all times.

Results

The means and standard errors for the standard deviation of lane position are presented in Figure 1. The standard deviation of lane position was analyzed using a 3 × 3 repeated-measures analysis of variance (ANOVA). There was no effect of cognitive workload, $F(2, 52) = 0.44, ns$. There was an effect of wind entropy, $F(2, 52) = 69.47, p < .05, \eta_p^2 = .73$. Most important, there was an interaction between workload and wind entropy, $F(4, 104) = 24.28, p < .05, \eta_p^2 = .48$. Pairwise comparisons indicated that as cognitive workload increased without wind, lane position variability decreased. When both cognitive workload and wind entropy increased, lane position variability increased. In the low entropy condition, lane position variability decreased from the single-task condition to the 0-back condition and finally to the 2-back condition. In the medium entropy condition, the effects of primary task predictability and secondary task workload canceled each other out. Finally, in the high entropy condition, lane position variability increased from the single-task condition to the 0-back condition to the 2-back condition. For both the low entropy condition and the high entropy condition, Bonferroni corrected post hoc comparisons were conducted at each level of cognitive workload. For low entropy, single task was significantly higher than 0-back, $t(26) = 4.10, p < .01$, and 2-back, $t(26) = 7.28, p < .01$. In addition,
0-back was significantly higher than 2-back, \(t(26) = 4.41, p < .01\).

For high entropy, single-task was significantly lower than 0-back, \(t(26) = -3.01, p < .01\), and 2-back, \(t(26) = -5.76, p < .01\). In addition, 0-back was significantly lower than 2-back, \(t(26) = -3.51, p < .01\).

Performance on the n-back task was analyzed using 2 x 3 repeated-measures ANOVA. There was no effect of wind entropy, \(F(2, 52) = .86, ns\), nor was there an interaction between entropy and cognitive workload, \(F(2, 52) = 1.85, ns\). However, there was a main effect of cognitive workload, \(F(1, 26) = 23.74, p < .05\), \(\eta_g^2 = .48\), wherein participants were nearly perfect on the 0-back task \((M = 1.0, SE = .00)\) but less accurate on the 2-back task \((M = .86, SE = .03)\). These levels of accuracy are consistent with those reported in other distracted driving studies that used the same n-back task (Mehler, Reimer, & Coughlin, 2012).

**Discussion**

Complex skilled behaviors as diverse as playing a musical instrument, typing, and driving an automobile are supported by a hierarchical control network that coordinates the interaction between automatic encapsulated routines and limited capacity attention. When the predictability of the task is manipulated, performance can be supported by either inner or outer control loops. We found that allocating attention to an encapsulated inner loop process disrupted performance, and diverting attention from an attention-demanding outer loop process also disrupted performance. We suggest that the manipulation of task predictability and the allocation of attention to the task help to provide a diagnostic signature of HCT and that this is a general characteristic of human behavior.

HCT is an integrative framework for understanding seemingly disparate (and often paradoxical) findings in the literature. Sometimes diverting attention from a task improves performance and sometimes it impairs performance. Of particular theoretical interest is the improved lane maintenance when a secondary cognitive task is added to predictable driving. In this case, diverting attention from a driving significantly improves lane maintenance. This finding is novel (i.e., aside from driving literature cited above, we are not aware of instances where adding a secondary cognitive task actually improves skilled performance) and would seem to complement the deautomatization of skills hypothesis (Beilock et al., 2002), wherein attending to the components of a skill impairs performance of experts.

HCT predicts performance that is inconsistent with ACT-R. Specifically, ACT-R predicts that as cognitive workload increases, lane maintenance should get worse. This follows because the n-back secondary task should require the same processing resources that are important for maintaining lane position (Salvucci, 2002). In contrast to these predictions, the current research found that increasing cognitive workload in predictable driving conditions reduced lane position variability. It is interesting that Salvucci and Beltowska (2008) reported that lane maintenance did get worse as cognitive workload increased, a pattern that they noted was consistent with ACT-R. This finding is unique in that most researchers have reported improvements in lane maintenance with increased cognitive workload. However, Salvucci and Beltowska (2008) placed construction cones on both sides of the roadway that may have inadvertently increased the level of difficulty of driving. Consequently, it is possible that their driving task reflected outer loop performance but did not fully capture inner loop performance. It is not immediately clear how ACT-R would predict that adding a secondary cognitive task would improve performance of an automatic procedure like lane maintenance. One possibility is that the n-back task may have flushed the contents of working memory, thereby eliminating extraneous information that could impair lane-keeping routines.

The current research also helps to explain an intriguing aspect of driving. Many drivers who are distracted or find their minds wandering arrive at their destinations without driving off the road. Previously, there was no satisfactory explanation for how this could happen. With HCT, we can now explain why distracted drivers do not drive off the road: In predictable situations, lane maintenance is an encapsulated inner loop process that does not require focused attention for success. This does not mean that distraction leads to improvements on other measures of driving. For example, if distracted drivers are required to respond to novel information, they are less likely to respond quickly and accurately (Strayer, Drews, & Johnston, 2003).

The research also helps resolve a discrepancy in the driving literature on the impact of cognitive workload (e.g., talking on a cell phone) on driving performance. Simulator-based driving studies often report significant impairments to driving with concurrent cell phone use (Strayer, Watson, & Drews, 2011). By contrast, naturalistic driving studies (e.g., Dingus et al., 2009) often suggest that there is little or no impairment when drivers talk on a cell phone. Because crashes are infrequent, naturalistic studies rely on surrogate measures of driving impairment, including sudden lane deviations and curb strikes. Given that drivers stay in their lane better under cognitive load, the surrogate measures of impairment used in the naturalistic studies are likely to underestimate the impairment from cell phone use. It is interesting that these same surrogate measures appear to be good indicators of visual and manual distractions such as sending or receiving text messages and, in this case, there is better agreement between simulator-
based (Drews et al., 2009) and naturalistic driving studies. This underscores the importance of understanding how different dependent measures affect behavior in different circumstances.

References


Received June 17, 2013
Revision received August 21, 2013
Accepted October 15, 2013