Cars, Calls, and Cognition: Investigating Driving and Divided Attention

Shamsi T. Iqbal, Yun-Cheng Ju, and Eric Horvitz
Microsoft Research
One Microsoft Way, Redmond, WA 98052
{shamsi, yuncj, horvitz}@microsoft.com

ABSTRACT

Conversing on cell phones while driving an automobile is a common practice. We examine the interference of the cognitive load of conversational dialog with driving tasks, with the goal of identifying better and worse times for conversations during driving. We present results from a controlled study involving 18 users using a driving simulator. The driving complexity and conversation type were manipulated in the study, and performance was measured for factors related to both the primary driving task and secondary conversation task. Results showed significant interactions between the primary and secondary tasks, where certain combinations of complexity and conversations were found especially detrimental to driving. We present the studies and analyses and relate the findings to prior work on multiple resource models of cognition. We discuss how the results can frame thinking about policies and technologies aimed at enhancing driving safety.

Author Keywords

Driving, Attention, Dual task performance, Cell phones

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Experimentation, Human Factors, Measurements

INTRODUCTION

Driving automobiles is often perceived as a fluid, nearly automatic process, and drivers often engage in secondary activities while driving [1, 7, 8, 31, 32]. Although some peripheral tasks are rapid and require only momentary shifts in attention from the primary driving task [10], other secondary activities may require more time and effort, and can lead to prolonged periods of divided attention [15, 23].

With the proliferation of mobile devices and people's desire to remain connected, talking on the phone, reviewing email, and even composing email messages and texting while driving have become commonplace. The cognitive, visual, and physical demands of such tasks can compromise the primary task of driving. Users may often overestimate their ability to divide their attention with secondary tasks because of the sense that driving is near automatic in many situations and can thus be safely shared with other tasks. However, it may be difficult to switch full attention back to driving in a timely manner so as to observe and respond appropriately when driving challenges arise, and such attentional challenges can have costly consequences [18].

The data linking the use of phones while driving to increases in accidents and fatalities has sparked legislation aimed at limiting cell phone usage during driving to hands-free configurations. However recent research has shown that using devices in a hands-free manner is no less harmful than the use of handheld devices [16]. Thus, phone use would have to be stopped entirely to avoid the challenge they present to driving safety. Unfortunately, people are unlikely to give up phone interactions while driving, and complete bans of phone use in this setting are unlikely.

Our goal is to better understand the interference between the cognition tapped for phone conversations and for driving. Insights about such potential interference would help to characterize better and worse times for phone conversations during driving, highlighting when drivers could more safely engage in phone conversations if absolutely needed. As a first step, we set out to understand how different types of cell phone conversations during varying levels of driving engagement affects driving performance and also the performance of the driver on the call itself. We sought to understand the findings in terms of interactions between cognitive resources used in driving and in handling common secondary tasks associated with phone conversations.

We conducted a controlled study with 18 participants driving within an interactive driving simulator. The participants drove on routes composed of segments that posed different types of navigation challenges. While driving, the participants would occasionally have to respond to a cell phone call, pushing a button to initiate a hands-free interaction. The cell phone calls were one of three kinds of engagement: listen to news and facts (*assimilate*), answer questions (*retrieve*), and provide directions (*generate*). In addition, for each driving trial, we asked drivers to either focus mainly on their driving, on the conversation, or do their best to both drive and handle the phone-based tasks.

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.

Not surprisingly, we found that drivers perform better on simpler routes than on complex routes when they are engaged in phone tasks. Drives on the simpler routes were associated with a lower incidence of collisions, sudden braking, and missed turns. We also found that specific types of phone conversations interfere more significantly with driving, presumably because task requirements and chunking associated with the calls interferes with cognition that is relied upon for driving. We discuss how certain tasks may steal attention from the driving task and/or make demands in other ways on the cognitive resources that are used in controlling and navigating a car.

RELATED WORK

We first discuss research on the ability of people to handle concurrent tasks. We focus on prior research and reflection within cognitive psychology on the allocation of attention in dual-task settings.

Divided Attention and Dual-Task Challenges

From the perspective of *Multiple Resource Theory* (MRT) of cognition [36], humans harness varying quantities of different kinds of cognitive resources (*e.g*., short- and long-term memory, attention, reasoning, etc.) to solve problems. In *dual-task* settings, people are challenged with completing two distinct tasks, creating potential contention for cognitive resources required to solve each one. It has been demonstrated that shifting resources from one task to another can improve the performance on the second task [25, 27], and as the difficulty of one task increases, the performance on the other decreases [38]. MRT further suggests that, for concurrent tasks, performance on both may be maintained if the tasks are in separate processing stages (e.g., response selection versus perceptual activities), or involve different processing mechanisms (e.g., spatial and analog information versus verbal and linguistic information) [36, 38]. Finally, the theory of automaticity proposed by William James suggests multiple processes can go on simultaneously, when they are habitual, involving minimal conscious control [17]. While automaticity can be obtained through training and practice [22, 33], its success also depends on consistency of the task [12].

Introduction of a second task has been shown to impact performance in many dual-task contexts. Basic research on performance in rarified dual-task settings has been done in studies of shared attention in visual search [1]. However, it is not clear if results demonstrated on low-level dual-task challenge problems, formulated and studied in psychology labs, holds for switches between higher-level and more complex tasks (*e.g.*, switching from driving to attending to a cell phone call) and how well attention can be selectively allocated, or divided across these two tasks. Pursuing such an understanding is the main thrust of this work.

The prior work on dual-task challenges provides a useful framing for research on performance tradeoffs in the setting of driving a car while talking on the phone or interacting with an in-vehicle system [16, 21]. While a seasoned driver may show an overriding automaticity in piloting a vehicle, driving a car demands resources associated with visual perception, spatial working memory, and motor responses and coordination [19, 37]. Performing secondary tasks may draw on resources used during driving [24], which may lead to performance degradations in one or both tasks.

We study commonplace dual-task situations that people face when they engage in phone conversations while driving. Different types of conversations may engage different mixes of cognitive, spatial, and verbal resources that may compete in different ways with the challenges of driving as well as the structural nature of the conversations. In this research, we study the interaction of classes of conversational tasks that we characterize broadly as *assimilation*, *retrieval*, and *generation*. Our goal is to understand how these different types of conversational tasks conflict with driving and to identify whether certain combinations of driving challenge and conversation tasks are associated with increased or diminished risk.

Effects of performing secondary tasks while driving

Dual-task scenarios of driving and performing secondary tasks such as conversing on the phone, texting, interacting with in-vehicle controls has been an area of active research. Studies have shown that dialing or answering the phone, adjusting the radio or interacting with music players have negative influences on driving, [7, 8, 32], as well as reasoning and conversing during driving [9] .

Using phones during driving has been shown to have catastrophic effects. For example, drivers with phones have slower braking reaction time [1, 20], have impaired steering control [8], and are more likely to have an accident [29, 35]. Moreover, no value of hands-free phones has been found, debunking beliefs that removing the need to physically hold phones reduces distraction during driving [29, 34]. These findings reinforce the hypothesis that cognitive demands of multitasking play a more important role in distracting drivers than manual manipulation [28, 34].

In order to understand the effects of cognitive demands on driving, researchers have looked at performance on various secondary tasks known to cause memory load in prior psychology studies. These include working memory tasks [1], mental arithmetic tasks [7] , and reasoning tasks [9]. Although not entirely representative of conversations that one may have over the phone, these tasks were used as they may replicate the cognitive demands placed on drivers while participating in more natural conversational settings. Performance on driving was reduced in all cases.

We reexamine the influence of phone conversations on driving, and also explore how phone conversations are affected by driving. We compare performance across different types of conversation while driving on courses with different levels of difficulty. Also, we investigate how varying levels of attention allocation across driving and phone conversations affect performance. While others have investigated several of these factors or partial combinations, prior research has not explored the interactions in a joint manner as we do in experiments reviewed below.

Strategies for interleaving secondary tasks with driving

Given that people can perform two tasks concurrently [36], researchers have recently looked at opportunities to interleave secondary tasks with driving. Brumby et al. investigated how interleaving a phone dialing task with driving impacted lane keeping and the dialing time under conditions of requests to prioritize either driving or dialing [10]. Results showed that when asked to prioritize the secondary task, drivers chunk components of the secondary task and switch back to driving at chunk boundaries to maintain driving performance, and while focusing on driving, the secondary task is slowed down. In a related study, Brumby et al. showed that the fastest strategy for selecting a song on a music player while driving was to scroll in one contiguous block without returning attention to the primary task of driving. For the safest strategy, more time needs to be given to the driving task, at the cost of longer response times for the secondary task, and correspondingly longer stretches of times for the dual-task scenario [11].

The prior work suggests that, for automatized tasks like driving, it may be possible to formulate strategies to perform other tasks without significantly compromising driving. Successful dual-task scenarios will depend on the availability and requirements of cognitive resources for the secondary task in light of resource consumption by the primary task and opportunities for interleaving the two tasks. We explore performance in these scenarios by generating phone calls with different cognitive demands during driving situations of different difficulties. We probe the interaction of cognitive resources for driving and handling calls via the proxy of measuring performance on both the driving and call tasks. We reflect about the timing and nature of conversations that conflict the least with driving and propose strategies for minimizing interference.

OVERVIEW OF STUDY

Our goal was to explore if and when opportunities exist when car drivers could engage in phone conversations without reducing driving performance, to understand which conversational tasks cause the most interference with driving, and the influence of increases in driving difficulty on interference between the primary and secondary tasks.

Understanding performance of driving on conversation has not been well studied. We also investigated how well users can carry out the conversation and how much they can recall afterwards. We addressed the following questions:

1. How is driving performance affected by participating in phone conversations where the driver has to interact in varying levels of engagement? How do these effects vary with changes in driving difficulty?
2. How are phone conversations influenced by concurrently driving, and how do these effects vary with changes in the levels of driving difficulty?
3. How does performance vary with requests to prioritize attention on driving, conversation, or both tasks?

To answer these questions, we conducted a controlled study using a driving simulator (see Figure 1). Using the simulator, users engaged in driving a realistic route, with a realistic steering wheel, pedals, and controls. Custom software allowed researchers to design driving scenarios and log relevant parameters during driving. To simulate a hands-free phone call environment, calls were presented through a peripheral system including a loud speaker and a microphone, and calls were accepted via a button on the driving console.

Experimental Design

The study was designed as a 3 (driving complexity) X 3 (call type) X 3 (focus) repeated measures within subjects design. Possible effects of order were countered by blocking the factors on a fully balanced Latin square design.

Users

18 people participated in our study (F=3), recruited through a call sent out to people selected randomly from the entire employee pool of our organization. The mean age of participants was 33.2 years, (S.D.=8.2) with a mean of 16.8 years of driving experience (S.D.= 9.41). All participants reported to be comfortable talking on the phone while driving.

Driving Task

Participants drove routes comprised of multiple 30s segments, each segment having either of the following three levels of complexity: *simple*, *complex,* and *unexpected occurrences*. An example of a simple segment is a single stretch of driving on a relatively empty road. Complex segments involve driving with many cars on a road, and requires changes of speed or lane changing. A segment with unexpected occurrences includes sudden, unexpected events, e.g., the car in front of the driver's car suddenly braking, a pedestrian stepping into the road, or an object rolling in front of the car. The segments with unexpected occurrences include time for the driver to recover and resume safe driving. Routes were about 10 minutes long.

Drivers were asked to follow the route straight on, unless they saw instructions to turn left or right. Instructions appeared in large banners in the frontal view of the driver and were easy to see if drivers were looking at the road.

To preserve the order of driving complexity as dictated by the Latin square design, we randomly chose segments where users would receive phone calls. Complexities were assigned to these segments according to the Latin square.

Figure 1: The STISim driving simulator. The system hosts a console with a steering wheel, turn signals, and buttons mapped to external functions. Three 47" screens placed at roughly 45° generate a convincing impression of driving a vehicle.

For consistency, segments with no phone calls were assigned difficulty levels randomly. This procedure reduced the probability that users would be able to predict phone calls on any given segment.

Phone Tasks

While driving, participants would occasionally receive phone calls, heralded by a traditional phone ring tone. Pressing the respond button on the console would mark the start of the conversation. The phone task was selected according to the design and launched. Participants had 30 seconds to perform the task.

Phone tasks belonged to one of the three categories: assimilation, retrieval, and generation (Table 1). The categories were designed based on prior studies looking at decrements of performance in driving while drivers engage in secondary tasks (see, e.g., [34]). In distinction to the prior work, we designed the tasks so as to resemble conversations that one typically may have over the phone.

For the *assimilation* task, the participants listened to a 15-20 second news headline. The choice of this task was motivated by Strayer and Johnston's [34] *book-on-tape* task, where participants were instructed to listen so that they could answer post-experiment questions. For the *retrieval* task, participants were asked to answer two questions, similar to questions one may hear from a telemarketer. For the *generation* task, participants were asked to provide driving directions between two points of interest. The first two tasks were designed to exploit verbal and linguistic/semantic resources. The third task adds more explicit spatial reasoning challenges to the mix.

Methodology

On arrival to the lab, participants were first guided through an informed consent process. They were then given an overview of the study. The experimenter then gave a demonstration of the controls of the driving simulator and the participant was provided with 10 minutes of practice driving to become accustomed to the system. Participants were also interrupted with practice phone calls during the practice session. Participants were then started with the study. To provide baseline data, users first performed *only* the phone tasks, without driving. They then went through a route of driving, without any phone interruptions.

Participants then drove 3 routes where for each route they received 9 phone interruptions, 3 for each phone call type.

Phone task type	Example
Assimilation (listen only)	A 16 year old in NY recently was texting while walking and fell into an open manhole.
Retrieval (respond)	When did you last have the oil changed in your car?
Generation (respond)	Please give directions from your home to the nearest grocery.

Table 1: Examples of phone based task categories.

Each call type was paired with each driving complexity, according to the Latin square design. To measure performance on the phone tasks, participants filled out a short questionnaire at the end of each route, designed to test recall of the content of the phone conversations. To reduce workload, we provided multiple choices when users were asked to select topics pertaining to the phone conversations.

For each route, participants were asked to distribute their attention across the two tasks according to the following focus (*focus more on driving, focus more on phone conversation, focus on both considering them to be equally important*). Focus was assigned based on the Latin square. Prior studies have shown that asking users specifically to focus on one task over another in dual-task experiments involving driving can yield different outcomes [10, 21].

The experimental session lasted about 1.5 hours. Users came back for two more sessions, where the experimental factors were varied according to the Latin square design to correct for ordering effects. In summary, each participant drove 9 routes and answered 81 phone calls over 3 sessions.

Measures

Performance for both driving and the phone tasks were measured for route segments when users were answering a phone call while driving. For driving performance, we measured the number of collisions, missed turns, and sudden braking, and driving speed while talking on the phone. We also recorded the same measures when users were driving while not engaged in a phone call, as a way of measuring baseline performance. The values were automatically recorded by the simulator.

For performance on the phone task, we measured the ability to correctly identify topics in the phone conversation. As an indicator of how users attempt to modulate the conversation to ensure driving safety, we also tracked the time participants took to respond to the ringing of an incoming phone call, and analyzed the prosodic content of their utterances through measuring the mean length of silent segments.

RESULTS

For the baseline condition, each user provided 3 data points per phone task, totaling 9 data points per user. For the experimental condition, each user provided 3 data points for each of the (Driving complexity (3) X Call type (3) X Focus (3)) conditions, totaling 81 data points per user and a grand total of 1,620 data points.

Effects on driving

The overall rates of occurrences of events characterizing driving performance (collisions, missed turns, sudden brakes etc) were low. Just as distractions do not always lead to catastrophic outcomes in real life, we saw small numbers of costly events with the simulator, similar to findings described in [34]. However, we sought to investigate the extreme cases of when such events occur, and how much they are affected by the factors (Focus, Driving complexity, and Call type) explored in this experiment. Because of the complexity of the experimental design, for each dependent factor, we will analyze first the three-way interactions (if applicable), then the two-way interactions between Call type and Driving complexity for each level of Focus, and finally the effects of Call types for each level of Driving complexity. As the length of the call varied by type, we will report on the events/minute to normalize comparisons across call types. We use α=0.0125 to control for Type 1 errors.

Collisions

A collision was recorded when the user hit a car, an object, or a pedestrian. Overall, the rate of collisions/min in the experimental conditions (with phone calls) was significantly higher than in the baseline condition (no phone calls) ($t(17)=5.19$, $p<0.0001$). For the experimental routes, number of collisions were low overall, occurring in only 53 out of 1458 phone calls (excluding the baseline phone calls). A Univariate ANOVA with Collisions/min as the dependent factor showed a significant 3-way interaction between Focus, Driving Complexity and Call Type ($F(8, 1432)=7.532$, $p<0.0001$). See Figure 2(a) for breakdowns of collisions/min across the three factors.

For cases where Focus was *driving*, Collisions/min did not vary significantly based on either Call Type or Driving Complexity. However, when Focus was *both*, significant interaction effects were found between Driving Complexity and Call Type ($F(4, 481)=10.5$ $p<0.0001$). Post hoc Bonferroni tests showed that for these cases, when the Driving Complexity was *unexpected*, significantly higher numbers of collisions/min occurred (M=0.34) compared to *simple* (M=0, $p<0.0001$). Collisions/min during *unexpected* were also higher than *complex* (M=0.1), but the differences did not reach significance ($p<0.024$). Within each type of Driving Complexity for the *both* condition, we then ran follow up test to explore effects of different Call Types. There were no significant differences in collisions/min for different Call Types for *simple* or *complex*. For *unexpected*, *retrieval* caused significantly higher number of collisions/min (M=0.998), compared to *assimilation* (M=0, $p<0.0001$) and *generation* (M=0, $p<0.0001$).

When the Focus was *conversing*, significant interaction effects were also found between Call Type and Driving Complexity. Post hoc tests showed that for these cases, *complex* resulted in significantly higher number of collisions/min (M=0.425) compared to *simple* (M=0.035, $p<0.002$). It is possible that the sustained attention required while driving on complex routes is affected when users are focusing mostly on conversing, thereby impacting the number of collisions. Follow up tests showed no significant differences across Call types for any of the Complexity levels.

Failing to follow instructions to turn

Instructions to turn left or right were provided in real time via road signs. Across the 18 users, 8.3% of the instructed turns (61/733) were missed or wrongly taken (e.g. taking a left turn instead of a right and vice versa), resulting from failing to properly notice the road signs instructing the turn. The rate of missed turns during phone calls were significantly higher than the rate of missed turns during the baseline condition of no phone calls ($t(17)=8.357$, $p<0.0001$). A univariate ANOVA with missed turns/min as the dependent factor showed a significant three-way interaction among Focus, Driving Complexity and Call Type ($F(7, 733)=10.65$. $p<0.0001$), see figure 2(b) for details.

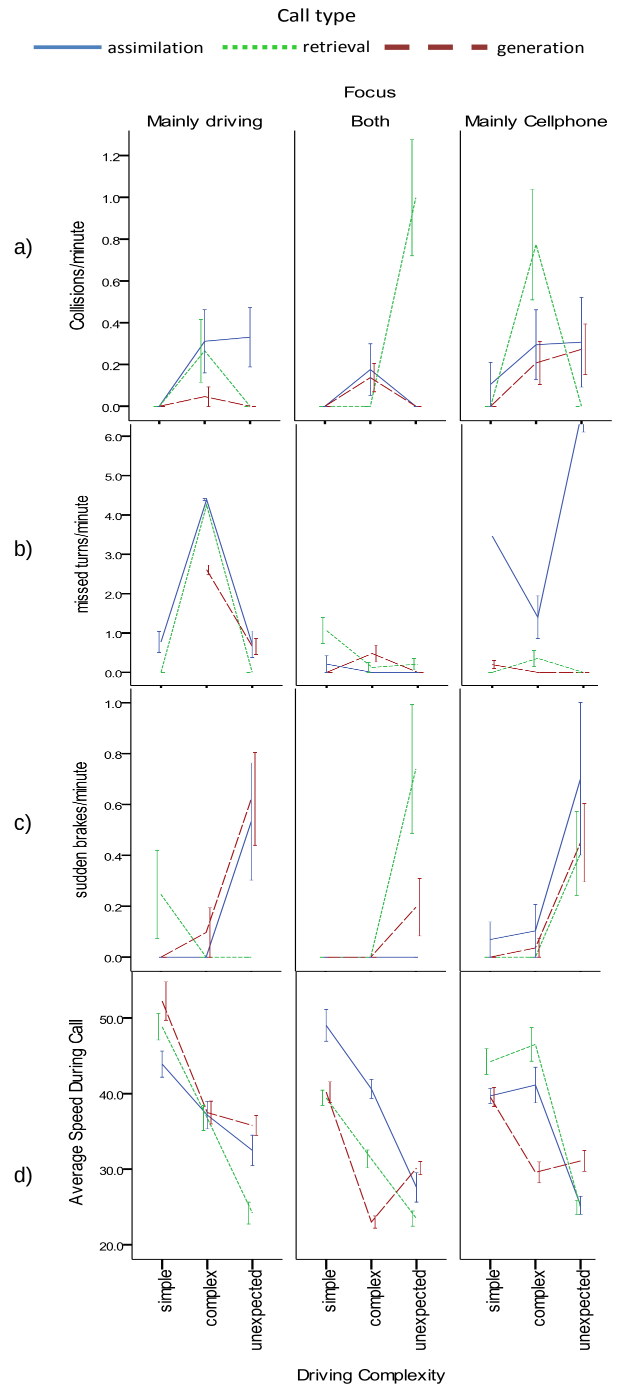

Figure 2. Mean numbers of (a) collisions, (b) missed turns, (c) sudden brakes, (d) average speed during calls, grouped by requested focus. Individual lines show missed turns by call types, separated by driving complexity (x-axis).

Overall, focusing on *both* caused users to miss significantly lower number of turns/min (M=0.23) compared to *driving* (M=1.68, $p<0.0001$) and *conversing* (M=1.357, $p<0.0001$)

only. This alludes to the overall difficulty of switching attention back and forth from one task to the other when focused primarily on either. Separating out the levels of Focus, where Focus was *driving*, there was a significant effect of only Driving Complexity ($F(2, 127)=18.35$, $p<0.0001$); *complex* caused more missed turns/min (M=3.75) compared to *simple* (M=0.39, $p<0.0001$) and *unexpected* (M=0.46, $p<0.0001$). There were no significant differences across the different call types, suggesting that when focused on driving, missed turns were affected more by road conditions rather than the act of conversing.

No differences across Call types or Complexities were found when Focus was *driving*. When Focus was *both*, there was a significant interaction between Call type and Driving Complexity ($F(4, 387) =4.3$, $p<0.002$). Followup tests showed no overall differences in missed turns/min across different Complexities. However, post hoc tests exploring the interaction revealed that for *simple*, *retrieval* caused users to missed significantly more turns/min (M=1.06) compared to *assimilation* (M=0.2, $p<0.002$) and *generation* (M=0. $p<0.0001$). No other effects were found. Focusing on *Conversing* also revealed a significant interaction between Call type and Complexity ($F(4,193) =19.3$, $p<0.0001$). Across all Complexities, *assimilation* caused more missed turns/min (M_s= 3.5, M_c=1.4, M_u=6.8, $p<0.0001$ for all) than *retrieval* (M_s=0, M_c=0.36, M_u=0) and *generation* (M_s=0.2, Mc=0, M_u=0). This finding suggests that the act of solely listening causes the user to miss written instructions.

Sudden Brakes

The rate of sudden brakes during phone calls was significantly higher than in the baseline condition ($t(17)=5.154$, $p<0.0001$). There were a total of 96 sudden brakes across the 18 users, and 55 of them happened during a phone call. A significant three-way interaction was found between Focus, Driving Complexity and Call Type ($F(8, 1432)=4.173$, $p<0.0001$). See figure 2(c).

As with Collisions, there was no significant effect of Focus on Sudden brakes/min. However, two-way interactions were found between Call type and Complexity when Focus was *driving* ($F (4, 466) =4.5$ $p<0.0014$) and when Focus was *both (* $F(4, 472) =3.8$ $p<0.005$;). Follow up tests showed that for both of these Focus levels, *unexpected* caused higher number of sudden brakes (M_d=0.38, M_b=0.31) compared to both *simple* (M_d=0.08, $p<0.006$; M_b=0, $p<0.0001$) and *complex* (M_d=0.03, $p<0.0001$; M_b=0, $p<0.0001$), accounting for 49 out of the 55 sudden brakes. Further followup tests showed significant effects of Call types on Sudden brakes/min only while focusing on *both*. *Retrieval* while focusing on *both* caused more sudden brakes/min (M=0.7) compared to assimilation (M=0.2, $p<0.005$). No other differences were found.

Driving Speed

We examined driving speeds in the phone call and no phone call conditions. Overall, average speed during phone calls were significantly lower than the average speed when there were no phone calls ($t(17)=3.45$, $p<0.003$). Again, a significant three-way interaction was found between Focus, Driving Complexity and Call Type ($F(8, 1421)=4.17$, $p<0.0001$). See figure 2(d).

Overall, users drove faster while focusing on *driving* (M=38.7 mph) compared to *conversing* (M=35.7 mph, $p<0.0001$) and *both* (M=33.8 mph, $p<0.0001$). Note that trying to focus on both resulted in the lowest speed, indicating the difficulty in maintaining continual performance on both driving and talking on the phone, and corresponding correction by the user by reducing driving speed.

Across all three levels of focus, speed was significantly higher for *simple* (M_d=48.3, M_b=42.9, M_c=41.2, $p<0.0001$ for all pairs but one) compared to *complex* (M_d=37.1; M_b=31.7, M_c=39.1 (not significantly different)) and *unexpected* (M_d=30.8, M_b=27.1, M_c=27.1). Speed for *complex* was also significantly higher than *unexpected* ($p<0.0001$ for all pairs). For each level of Focus, significant two-way interactions were found between Call type and Driving Complexity ($F_{driving}(4, 472)=5.5$, $p<0.0002$; $F_{both}(4, 474)=11.14$, $p<0.0001$; $F_{conversing}(4, 475)=14.9$, $p<0.0001$). When Focus was *driving*, *unexpected* resulted in lowering of driving speed during *retrieval* (M=24.2, $p<0.0001$ for all pairs) compared to *assimilation* (M=32.5) and *generation* (M=35.8). When Focus was *both*, both *simple* and *complex* resulted in lowering in driving speed during *retrieval* (M_s=39.4, M_c=31.3) and *generation* (M_s=40.2, M_c=23) compared to *assimilation* (M_s=49.02, M_c=40.6, $p<0.0001$ for all pairs), while *unexpected* resulted in *retrieval* causing significant lowering of speed (M_u=23.4) compared to *generation* (M_u=30.2, $p<0.0002$). When Focus was *conversing*, *complex* resulted in *generation* causing more reduction in speed (M=29.6, $p<0.0001$ for both pairs) compared to both *retrieval* (M=46.5) and *assimilation* (M=41.1). On the other hand, *unexpected* resulted in *generation* causing less reduction in speed (M=31.1) compared to both *retrieval* (M=24.9) and *assimilation* (M=25.2), $p<0.0001$ for both pairs. These numbers allude to the difficulty in dividing attention across the tasks of conversing, retrieving and generating information, the visual-spatial task of driving and the skillful maintenance of speed. However, as discussed before, being engaged in a complex secondary task that conflicts with the driving task in an obvious way may lead drivers to modulate driving speed in an attempt to ensure safe driving.

Influence of driving on conversations and related tasks

We also studied effects of driving on the performance of the conversations and actions associated with them. Since this was a dual-task scenario where the phone task would likely be foregrounded, we wanted to understand how concurrent driving influences conversational performance.

Time to respond to calls

Time to respond to calls was measured as the time between the initiation of the phone ringing and the time the user hit the button, depending on how difficult it was to switch attention from driving to the phone call. The time to respond to a phone call while driving was significantly higher

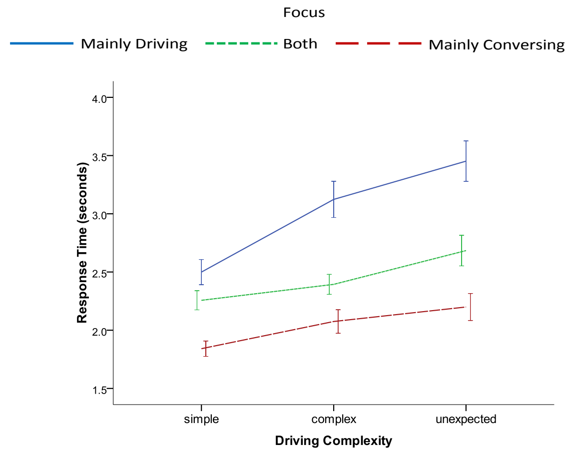

Figure 3. Mean time to respond to phone calls. Individual lines show missed turns by call types, separated by driving complexity (x-axis).

(M=2.5s,), compared to the baseline condition (M=1.52s, F(1,1608)=55.37, p<0.0001). These differences indicate the difficulty in switching attention from driving to responding to the phone call, where users often attempted to reach a safe state in driving before initiating the call.

A univariate ANOVA showed main effects of Focus (F(2, 1431)=53.3 , p<0.0001), and Driving complexity (F(2, 1431)=18.32, p< 0.0001) on Response time (Fig. 3). There were no interactions between the two. We ignore Call types for this analysis as the user did not know before responding to the call what type of conversation it would entail.

As expected, when the requested focus was on driving, response time was highest compared to the situations of focusing on both (p<0.0001) and more on phone calls (p<0.0001). Response times for situations where the requested focus was more on phone calls were lower than when the focus was on both driving and phone calls (p<0.0001). These results are in line with the expected difficulties associated with disengaging from driving to respond to the call; when focused more on driving, it took more time to switch than when prioritizing the cell phone conversation. Interestingly, prioritizing both tasks leads to response times that are not as high as the case when drivers were focused more on driving, nor as low as when focusing on the calls, indicating that there is some tradeoff in rapidity of response perhaps indicating an attempt by drivers to maintain driving safety.

For driving complexity, response times while facing an unexpected event while driving was significantly higher than when driving on *simple* (p<0.0001) and *complex* (p<0.05). Response times when driving on a complex route were also significantly higher than when driving on a simple route (p<0.0001), again demonstrating the potential difficulty of disengaging from a cognitively demanding situation (complex or unexpected event) to take a call.

Recalling information from conversations

As a measure of how effectively users were able to pay attention to the phone conversations while driving, we analyzed how much they could recall information about the conversations in a questionnaire presented to them right after finishing driving each route. The questionnaire asked them to identify from within a set of distracters keywords presented in the news headlines, topics addressed in the demographic questions, and start and end points of the directions they were asked to provide in the conversations.

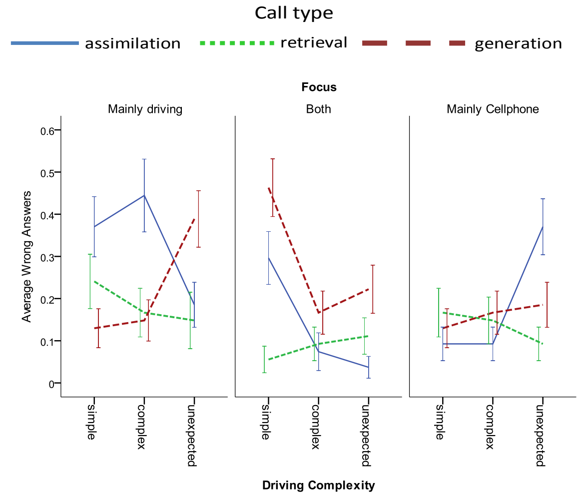

Figure 4. Mean numbers of wrong answers, grouped by requested focus. Individual lines show missed turns by call types, separated by driving complexity (x-axis).

As displayed in Figure 4, drivers had a higher number of wrong answers for questions from the experimental condition with phone calls (M=0.19) compared to the baseline condition (M=0.12, F(1, 1618)=4.83, p<0.028). As with driving performance, a significant three-way interaction on wrong answers was found among Focus, Driving Complexity, and Call Type (F(8, 1431)=6.3, p<0.0001).

Focus influenced the number of wrong answers. When Focus was *driving*, the number of wrong answers (M=0.247) was significantly higher than *both* (M=0.17, p<0.008) and *conversing* (M=0.16, p<0.003). For all focus requests, there were significant two-way interactions between Call Type and Driving Complexity ($F_{driving}$(4, 477)=6.79 p<0.0001; F_{both}(4, 477)=3.55, p<0.007; $F_{conversing}$(4, 477) =3.5, p<0.008). When Focus was *driving*, *complex* driving resulted in more wrong answers for *assimilation* (M=0.47) compared to *retrieval* and *generation* (M=0.2, p<0.01 for both). Other differences were not significant. When Focus was *both*, *simple* driving resulted in more wrong answers for *assimilation* (M=0.3, p<0.0009), and more wrong answers for *generation* (M=0.46, p<0.0001) compared to *retrieval* (M=0.06). *Unexpected* resulted in more wrong answers for *generation* (M=0.22) compared to *assimilation* (M=0.04, p<0.01). When Focus was *conversing*, *unexpected* showed more wrong answers for *assimilation* (M=0.4) compared to *retrieval* (M=0.09, p<0.001).

Overall, focusing on driving resulted in reductions in accuracy on the post-experiment test. Answering questions on *assimilation* appeared to be more difficult, and *retrieval* to

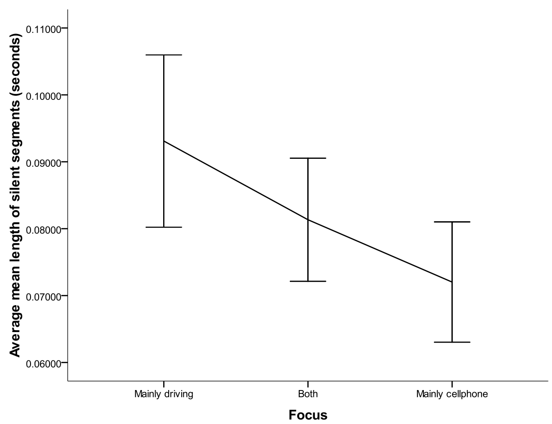

Figure 5. Average mean length of silent segments in seconds, by level of focus.

be the least problematic across the different conditions. The former could be due to the difficulties in retaining minute details of information delivered in a prosaic form. Retrieval, on the other hand, led to more detrimental effects in driving performance; users may focused more intensively on retrieval questions, sacrificing performance on driving to achieve better performance on their responses.

Prosodic analysis of drivers' speech

We analyzed attributes of the speech of drivers during phone conversations using specialized speech analysis tools. We examined the length of silent segments in their speech as a potential indicator of interleaving of attention or interference between the driving task and conversation. As the user did not speak in the assimilation tasks, we exclude corresponding phone calls for this analysis.

There was a significant effect of Focus on the mean length of silent segments for each conversation ($F(2, 938)=4.04$, $p<0.018$), but no effects of Call type or Complexity. Post hoc tests showed that silent segments were the longest ($M=0.093s$) when Focus was *driving*, significantly higher than when Focus was *conversing* ($M=0.072s$, $p<0.012$). While no significant differences were found, the length of the silent segments when Focus was *both* was higher than when Focus was *conversing*, and lower than when Focus was *driving* (see Figure 5). This finding may be interpreted as evidence that when drivers are focused on driving, it may be more difficult to disengage and switch to a call, and drivers may perform more shifts in their attention to maintain driving performance, resulting in increases in silent periods in the conversation. The cognitive load associated with difficulty in switching focus and generating utterances may also contribute to the differences in the length of silent segments [5].

Summary of results

Our findings provide compelling evidence for significant degradations in both driving performance (collision, missed turns, sudden braking, and reduction of speed) and conversation performance (response time to call, information recall, and quality of speech) while driving and conversing concurrently compared to the baseline conditions or either driving or conversing on the phone. We provide evidence of the complex interactions between varying levels of focus, driving complexity and conversation type. Overall, requested focus did not appear to have a major impact on *driving* performance, as automaticity seemed to allow users to drive and converse simultaneously. However, focus did have a stronger effect on the performance on the conversations. On receiving phone calls, users often attempted to achieve a stable driving situation before continuing with the conversation, which may explain the lengthier silent segments. The more focused the user was on driving, the more time it took them to switch to the phone call.

As expected, answering a phone call while driving on a simple route resulted in the least collisions, the least slowing down in speed, and the least sudden braking, though the rates were still higher than the baseline condition of no phone calls. We found that conversations requiring information retrieval had the most negative influence on driving performance. This type of task was associated with the most collisions during the occurrence of unexpected events while driving, the most missed turns during simple driving, the most sudden braking while focusing on both driving and conversing, and the greatest reduction in speed during unexpected events. Generation tasks caused users to slow down driving during complex routes, suggesting that there may be conflicts in visual-spatial resources that lead users to compensate by lowering driving speed in an attempt to maintain safer driving performance. Assimilation tasks generally resulted in lower negative impact on driving performance perhaps because users did not have to provide responses. However, recall of call information was poorer than for other tasks, perhaps because of the need to remember more precise details.

DISCUSSION

Our findings highlight the difficulties in determining opportune moments for engaging in calls based on the rich set of experimental conditions that we explored. Requests to users to focus on calls versus on driving was an attempt to simulate the real-world spectrum of attention to calls while driving, depending on the nature of the engagement with a call-based conversation based on its perceived importance or salience. We realize that the focus variable in the study is not necessarily reliable as the test condition hinged on peoples' voluntary intention, ability, and interpretation of the request to focus in a particular way, versus the reality of what happens in real-life settings. This was exemplified in our results where focus was generally not shown to significantly affect the performance measures.

We found that participants occasionally were able to attend to each task effectively when confronted with one task when asked to focus on the other, suggesting that drivers may be interleaving the tasks. For example, even while deeply engaged in a phone conversation, the drivers were found to make subtle adjustments to their driving (e.g., reducing their driving speed to have more control) to accommodate the conversation without compromising safety. When focusing on driving, users were found to have higher numbers of silent segments in their conversations, suggesting that they might be switching back and forth between driving and the conversation to ensure that the driving task was managed. Another potential effect is that the action of conversing may subconsciously increase drivers' awareness of their driving performance. Kubose [6] showed that talking improves driving performance compared to driving in silence, finding less variability in maintenance of lane position across easy and difficult driving. Some of our results could be interpreted by taking into account the automaticity of driving, where drivers are able to drive safely without having to exert much attention. In fact, experts can perform worse when they explicitly try to focus on the components of their skills [3, 4].

In addition to pursuing measurable outcomes of the experimental conditions, we sought insights about the cognitive mechanisms underlying the findings. We were particularly interested in understanding how cognitive demands that conversations place on drivers (*e.g.,* tasks requiring spatial resources versus requiring verbal resources) interacted with the oft automatic, visual-spatial-motor task of driving. We had hypothesized that the visual-spatial task of generating directions would result in the largest degradation in performance during driving, but found retrieval tasks had a more costly influence on driving performance. According to the multiple resource theory (MRT), a largely verbal phone conversation should not significantly interfere with a purely visual-spatial driving task [7, 34, 36]. However, for real-world tasks, it is probably important to move beyond simple notions of multiple resources and to consider the details of problem solving associated with a superficially verbal phone conversation. Tasks may require multiple stages of effort and draw upon a rich array of cognitive resources. For example, there is evidence that complex cognitive investments are made in preparation for retrieving information from memory [30]; such preparation may interfere with other cognitive resources, and the concurrent execution of a primary task such as driving.

Beyond only considering contention for similar cognitive resources, studies of interference in dual-task settings might also better leverage considerations of task decomposability and the efficiency with which people can control the sequencing of the chains of subtasks. Prior studies have examined chunking for dialing while driving, showing that drivers interleave dialing subtasks with driving so as to maintain lane keeping performance [10]. Returning to the examples at hand, it is possible that the directions-generating task, even though drawing upon resources required by the visual-spatial driving task, could be efficiently decomposed into smaller subtasks that are sequenced under user control, enabling drivers to modulate the interleaving of subtasks with the demands of driving. Drivers may consciously choose to interleave driving with call subtasks depending on the combined cognitive load and structural composition of the tasks. A retrieval task may be less decomposable, thus providing drivers with fewer opportunities to make tradeoffs in their focus of attention per the real-time requirements of the driving task. Several prior studies have explored task decomposition and opportunities for interleaving other tasks into the subtask structure [2, 26].

Overall, the findings from our study emphasize the need for more research on MRT models of dual-task performance for real-world tasks. The complexity of the results suggests that caution must be used in applying MRT to complex, real-world tasks. We also need to understand the decomposability of common tasks into subtasks that may be interleaved with subtasks of another. Better characterizations of the constellation of resources required to execute common tasks are needed. It will likely be useful to pursue models of dual-task performance that jointly consider resource contention, task decomposition and subtask sequencing.

Our general findings about the influence of phone conversations, resonate with the findings of prior studies, and highlight the potential value of designing in-vehicle systems that prioritize driving safety while being sensitive about users' communication needs [13, 14]. Evidence on conversations and driving complexity could be harnessed in automated safety services that might be engaged when attentional deficits are noted. Such systems could shift into a mode of providing aggressive warnings about road nuances and hazards, modulating their warnings based on real-time inferences about conditions that would likely surprise drivers who are handicapped with divided attention. Systems could also mediate incoming calls and calls in progress, passing callers to voicemail and even gracefully disconnecting calls in progress with an apology to callers. Intelligent systems might even one day impose a maximal driving speed or increase a minimal allowed distance to the car in front of a driver when an attention-soaking conversation is detected, thus generating more time for drivers to respond to events.

CONCLUSION AND FUTURE WORK

We conducted a study to explore how different types of phone conversations during driving result in performance degradation of both driving and conversations. Our findings show that simple routes with few or no other cars and constant speeds are safest in terms of receiving and engaging with phone calls. Despite their engagement in phone conversations, drivers appear to try to maintain safety in driving. However, problems with driving may arise when cognitive resource demands exceed resource availability, such as when drivers are engaged in conversations involving retrieving information from memory while facing a complex situation on the road. The findings underscore the complexity of interactions between different kinds of conversation-centric tasks and driving. The results also raise questions, highlighting the importance of pursuing a deeper understanding of the nature of real-world tasks and their demands on cognition in realistic dual-task settings.

ACKNOWLEDGMENTS

We thank Ed Cutrell for suggestions on the design of the driving study, Ivan Tashev for his help with the driving simulator, and the participants in our study.

REFERENCES

1. Alm, H. and L. Nilsson. The effects of a mobile telephone task on driver behavior in a car following situation. *Accident Analysis and Prevention*, *27* (5). 707-715.
2. Bailey, B.P. and S.T. Iqbal. Understanding Changes in Mental Workload During Execution of Goal-directed Tasks and Its Application for Interruption Management. *ACM Transactions on Computer Human Interaction (TOCHI)*, *14* (4). 1-28.
3. Beilock, S.L. and T.H. Carr. On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, *130*. 701-725.
4. Beilock, S.L., T.H. Carr, C. MacMahon. and J.L. Starkes. When Attention becomes counterproductive: Divided versus skill-focused attention in performance of sensorimotor skills by novices and experts. *Journal of Experimental Psychology: Applied*, *8*. 6-16.
5. Berthold, A. and A. Jameson. Interpreting Symptoms of Cognitive Load in Speech Input. in *User Modeling*, (Vienna, 1999), New York: Springer Wien New York.
6. Bock, K., G.S. Dell, S.M. Garnsey, A.F. Kramer, and T.T. Kubose. Car talk, car listen. in Antje Meyer, L.W., Andrea Krott ed. *Automaticityand Control in Language Processing*, Psychology Press, 2006, 21-42.
7. Briem, V. and L.R. Hedman. Behavioral effects of mobile telephone use during simulated driving. *Ergonomics*, *38*. 2536-2562.
8. Brookhuis, K.A., G. De Vries, and D. De Waard. The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, *23* (309-316).
9. Brown, I.D., A.H. Tickner, and D.C.V. Simmonds. Interference between concurrent tasks of driving and telephoning. *Journal of Applied Psychology*, *53* (5). 419-424.
10. Brumby, D.P., D.D. Salvucci, and A. Howes. Focus on driving: How cognitive constraints shape the adaptation of strategy while dialing while driving *Proceedings of the SIGCHI confrence on Human Factors in Computing Systems*, ACM Press, Boston, MA, 2009, 1629-1638.
11. Brumby, D.P., D.D. Salvucci, W. Mankowski, and A. Howes. A cognitive constraint model of the effects of portable music-player use on driver performance *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, HFES, Baltimore, MD, 2007, 1531-1535.
12. Fisk, A.D., P.L. Ackerman, and W. Schneider. Automatic and controlled processing theory and its applications to human factors problems. in *Human Factors Psychology*, North-Holland Publishing Co., 1987, 159-197.
13. Green, P.A. Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions, University of Michigan Transportation Institute, 2004.
14. Hoedemaeker, M. and M. Neerincx (eds.). *Attuning In-car User Interfaces to the Momentary Cognitive Load*. Springer-Verlag, Berlin, 2007.
15. Hoffman, J.D. A Dynamic Field Theory Model of Shared Visual Attention While Driving *University of Iowa*, Iowa City, 2008.
16. Horrey, W.J. and C.D. Wickens. Examining the Impact of Cell Phone Conversations on Driving Using Meta-Analytic Techniques. *Human Factors*, *48* (1). 196-205.
17. James, W. *Principles of Psychology*, New York, 1890.
18. Klauer, S.G., T.A. Dingus, V.L. Neale, J.D. Sudweeks, and D.J. Ramsey. The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data, National Highway Traffic Safety Administration, Washington DC, 2006.
19. Lee, J.D., B. Caven, S. Haake, and T. Brown. Speech-based Interaction with In-vehicle Computers: The Effect of Speech-based E-mail on Driver's attention to the Roadway. *Human Factors*, *43*. 631-640.
20. Lee, J.D., D.V. McGehee, T.L. Brown, and N.L. Ryes. Collision warning timin, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors*, *44*. 314-334.
21. Levy, J. and H. Pashler. Task Prioritisation in Multitasking during Driving: Opportunity to Abort a Concurrent Task Does not Insulate Braking Responses from Dual Task Slowing. *Applied Cognitive Psychology*, *22*. 507-525.
22. Logan, G.D. Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22* (1). 1-35.
23. McKnight, A.J. and A.S. McKnight. The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention*, *25*. 259-265.
24. Navon, D. Attention Division or Attention Sharing? *Attention and Performance*, *9*.
25. Navon, D. and D. Gopher. On the Economy of the Human Processing System: A Model of Multiple Capacity. *Psychological Review*, *86*. 254-255.
26. Newtson, D. and G. Engquist. The Perceptual Organization of Ongoing Behavior. *Journal of Experimental Social Psychology*, *12*. 436-450.
27. Norman, D.A. and D.G. Bobrow. On data-limited and resource-limited processes. *Cognitive Psychology*, *7* (44-64).
28. Nunes, L.M. and Recarte, M.A. Cognitive Demands of hands-free-phone conversation while driving. *Transportation research Part F*, *5*. 133-144.
29. Redelmeier, D.A. and R.J. Tibshirani. Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine*, *336* (7). 453-458.
30. Ruppin, E. and Y. Yeshurun. Recall and Recognition in an Attractor Neural Network Model of Memory Retrieval. *Connection Science*, *3*. 381-399.
31. Salvucci, D.D. Predicting the effects of in-car interface use on driver performance: An integrated model approach. *International Journal of Human-Computer Studies*, *55*. 85-107.
32. Salvucci, D.D., D. Markley, M. Zuber, and D.P. Brumby. iPod distraction: effects of portable music-player use on driver performance *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, San Jose, California, USA, 2007.
33. Schneider, W. and R.M. Shiffrin. Controlled and automatic human information processing ii: Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*. 127-190.
34. Strayer, D.L. and W.A. Johnston. Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science*, *12*. 462-466.
35. Violanti, J.M. Cellular phones and traffic accidents. *Public Health*, *111* (6). 423-428.
36. Wickens, C.D. Multiple Resources and Performance Prediction. *Theoretical Issues in Ergonomic Science*, *3* (2). 159-177.
37. Wickens, C.D. and J.G. Hollands. *Engineering Psychology and Human Performance*. Prentice Hall, Upper Saddle River, NJ, 1999.
38. Wickens, C.D., D. Sandry, and M. Vidulich. Compatibility and resource competition between modalities of input, output and central processing. *Human Factors*, *25*. 227-248.