An Exploration of Speech-Based Productivity Support in the Car

Nikolas Martelaro
Accenture Technology Labs
San Francisco, CA
nikolas.martelaro@accenture.com

Jaime Teevan
Microsoft Research
Redmond, WA
teevan@microsoft.com

Shamsi T. Iqbal
Microsoft Research
Redmond, WA
shamsi@microsoft.com

Abstract

In-car intelligent assistants offer the opportunity to help drivers productively use previously unclaimed time during their commute. However, engaging in secondary tasks can reduce attention on driving and thus may affect road safety. Any interface used while driving, even if speech-based, cannot consider non-driving tasks in isolation of driving—alerts for safer driving and timing of the non-driving tasks are crucial to maintaining safety. In this work, we explore experiences with a speech-based assistant that attempts to help drivers safely complete complex productivity tasks. Via a controlled simulator study, we look at how level of support and road context alerts from the assistant influence a driver's ability to drive safely while writing a document or creating slides via speech. Our results suggest ways to support speech-based productivity interactions and how speech-based road context alerts may influence driver behavior.

CCS Concepts

• **Human-centered computing → User studies**;

Keywords

Driving, Multitasking, Attention Management, Productivity

ACM Reference Format:
Nikolas Martelaro, Jaime Teevan, and Shamsi T. Iqbal. 2019. An Exploration of Speech-Based Productivity Support in the Car. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300494

1 Introduction

Time spent commuting is on the rise: in 2017 the average commute time for US drivers was 26.5 minutes [9]. Between

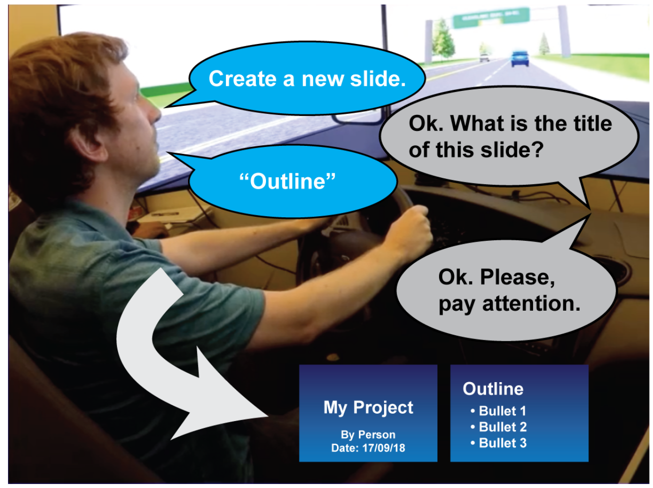

Figure 1: This paper explores how to design interfaces that safely allow drivers to be productive while commuting.

2014 and 2015, the number of hour-long commutes rose by 5% and the number of 90+ minute commutes rose by 8% [19]. To manage time in the car productively, drivers often use their mobile phones to stay connected to others [44], receive information [45], and conduct work [8, 15, 29]. Even when people are aware of the danger of using phones while driving, they often feel that this enables them to get work done [39].

While secondary tasks can distract from driving and negatively affect driving performance [17, 18, 32, 52], research suggests drivers can dual-task in some contexts, especially during easier moments of driving [32]. Speech interfaces, in particular, can be designed to reduce the impact of a secondary task on driving ability, allowing drivers to complete simple tasks such as placing a phone call or sending a message [31]. Progress in speech recognition has enabled manufacturers such as Mercedes Benz and BMW to build intelligent assistants into their cars. It is likely that these assistants will support non-driving related tasks. However, it is not clear how these interfaces can help drivers when safety must be prioritized. Thus, we are interested in exploring interfaces that support drivers' desire to use their commute productively while ensuring safety.

One potential way for in-car speech-interfaces to accommodate non-driving tasks may be *microtasks*; small tasks

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300494

typically generated by decomposing larger tasks into smaller actionable units [42]. Microtasks may be suitable for the driving environment, as they are often context free and can be done quickly and therefore can be easily interleaved during driving segments where cognitive load is low [21]. Breaking larger tasks into sub-tasks can also make them more resilient to interruptions [13] and allow them to be resumed more easily after interruptions if proper support associated to the task is provided [1]. Therefore users may be more comfortable switching away from them when the driving task needs to be prioritized.

Another way for speech interfaces to support drivers is through providing road context alerts. Previous studies have shown that alerts can improve driving performance during phone calls [21]. Furthermore, speech interfaces which provide specific information about the reason for an alert have been shown to improve driving performance and be preferred by drivers [28]. While these type of alerts do not directly assist in the driving task, they may help drivers maintain safety during a dual-task scenario.

In this paper, we present a simulator-based driving study exploring experiences with a speech-based assistant that helps drivers create documents *and* maintain safety on the road. Current technological trends such as human parity in speech recognition [49], improved tools for image segmentation and understanding of road scenes [12], and automatic task planning systems for productivity tasks [27] suggest that future speech interfaces could help drivers complete productivity tasks. Our goal in this paper is to explore how such a system might benefit a driver. Drawing from research on divided attention and microtasks, we designed a user study to explore the impact of two types of support from an in-car assistant to help drivers be productive during simple driving scenarios: 1) *level of support* and 2) *road context*. We find that while drivers are challenged by the productivity task, they are able to manage the task without significant driving performance reduction. Our level of support and road context support concepts were met with mixed reactions: drivers found support that inhibited their flow of thought to be worse for their ability to complete the task and maintain safety on a real road. The feedback reveal design issues around task and context support and suggest that people have personal differences in their thought processes that can make one type of support good for some and distracting for others. Thus, careful consideration needs to go into the type of guidance that assistants can produce in the car.

2 Related Work

Distracted Driving

Distractions from interactions with the radio or cell phones can lead to degraded driving performance across different age groups or road complexities [17, 18] and are estimated to cause 25% of vehicle crashes in the US [52]. Different distractions influence driving behavior differently with visual distractions leading to reduced speed and increased lane keeping variations while cognitive loads lead to less lane keeping variation [14]. When drivers are aware of their distraction, they will often compensate by slowing down [37] and leaving more headway distance with the car ahead [24].

Even when not interacting with other devices in the car, peoples minds can wander [34] to other thoughts unrelated to driving, especially on simple and familiar routes [50], resulting in decreased reaction time similar to more explicit distractions [51]. Moments where drivers are likely to mind wander may be good opportunities for drivers to be more explicit in their thoughts and conduct a productivity task. Given prior findings, we anticipate that drivers conducting an explicit cognitive task such as a productivity task will exhibit similar compensation strategies to other explicit cognitive distractions such as making a phone call or interacting with a speech-based e-mail client.

Speech Interfaces in Cars

Speech interfaces are becoming a more common method for interacting with in-car systems and secondary tasks such as phone calls, media players, navigation, safety alerts, and information [31, 41]. The use of speech interfaces is growing as drivers using speech interfaces have been shown to perform on par or better than with manual interfaces, although, not as well as driving alone [4, 32]. For example, Jamson et. al. [24] found that drivers using a speech-based email client had reduced reaction time to sudden braking events and less lane deviation when compared to baseline driving. The design of speech-based systems also has an impact on driver's cognitive load and subsequent driving performance. In a comparison of different speech systems, Strayer et. al. [40] found that while systems without perfect speech recognition were significantly more taxing on drivers, well functioning systems added little extra cognitive load. Overall, speech interfaces provide an opportunity for in-car interactions that can be less distracting than visual-motor interfaces [4] with careful design. In our work, we are interested in what aspects of an intelligent speech-based assistant's interactions can help the driver balance their desire to complete a productivity task while maintaining their safety.

Managing Tasks and Thoughts While Driving

Driving presents an interesting conundrum as parts of driving can often be well-learned for an experienced driver, described as 'subconscious control' [36]. Here the human control system develops specialized procedures for tasks that are relatively independent of each other and can be performed simultaneously. However, this assertion breaks when requirements of one or both of the tasks changes, requiring conflict resolution between two tasks. The complexity of both the driving and the non-driving task, does have an effect: prior work has shown driving on a simple road is more conducive to engaging in a secondary task compared to a road segment with complex driving challenges [23]. The challenge, therefore, remains in being able to redirect attention back to driving when the task requirements change.

Prior work has also explored how to compensate for reduced performance when a driver may be simultaneously engaged in a non-driving task. Alerts can help drivers redirect focus on the driving task while engaged in a secondary task [22, 43], persuade drivers to drive more economically [35] or mediate communication among passengers [33]. In studies where context is provided to remote callers, drivers have better driving performance [26].

While distracted driving is a growing problem, there is also evidence that with careful consideration of cognitive demands of the tasks, it may be possible and even beneficial to allow secondary tasks to be presented while driving without jeopardizing driving safety [2]. We explore how this experience may look by simulating a very simple driving task and a speech-based content creation task with added system support to help the driver navigate both tasks.

Task Support and Interruptions

Consideration of task structure has shown promising results in reducing the negative effects of inopportune task switching. Parts of a task where mental workload is low are typically more suitable for interruptions [36]. Such moments often occur at natural task breakpoints [20]. Many driving studies have leveraged this notion of using natural breakpoints for task switching (e.g., [7, 25, 38]). More recently, the concept of microtasks has introduced an alternate form of tasking where a complex task can be decomposed into smaller units which can be completed in short bursts of time [10, 13, 42]. Breaking larger tasks into sub-tasks can make them more resilient to interruptions [1, 13]. However, it is important to prime users with relevant information associated with the task and overall goal [1]. Bringing both threads of research together, interleaving microtasks at natural breakpoints and providing support about the task being completed may offer benefits such as reduction of information that needs to be maintained in memory [5], freeing up mental resources for other tasks [47], reduction in stress [3], and speed-accuracy trade-offs in dynamic environments such as driving [25].

Our approach builds on these existing theories in the cognitive science and HCI literature. We look at how driving performance is impacted by the systematic interleaving of microtasks that are part of a broader content creation task with the driving and how this experience compares to the baseline of free-form interaction. We identify how drivers adapt in this dual task scenario, seeing how microtasks either positively or negatively influence driving and productivity task behavior. We also look and what additional driving support helps drivers manage both tasks better.

3 Methodology

We conducted a user study to explore how a driver might complete a productivity task like writing or presentation slide creation while maintaining driving safety. In the study, participants drove in a driving simulator while interacting with a speech-based assistant that varied the *level of support* (motivated by the literature on microtasks, with the assistant providing high or low levels of support about the task) and *road context* (motivated by the literature on managing tasks while driving, with the assistant providing context alerts with high and low specificity). We aimed to address three research questions with the study:

RQ1a: How does level of task support influence driving behavior?
RQ1b: How does road context influence driving behavior?
RQ2: How do drivers behave during critical moments while driving and completing the productivity task?
RQ3a: How does level of task support influence the driver's ability to complete the productivity task?
RQ3b: How does road context influence the driver's ability to complete the productivity task?

Experimental Design

We used a 2 (level of support: low / high) by 2 (road context: low / high) within-subjects design with three baseline conditions: 1) *Driving* (no productivity task, no assistance) – to measure focused driving performance, 2) *Thinking Aloud* (productivity task only, no driving, no assistance) – to measure focused productivity task performance, and 3) *Driving + Thinking Aloud* (productivity task while driving, no assistance) – to see how well people could complete the task and drive without assistance. In total, participants experienced seven experimental conditions — three baselines and four 2 x 2 factorial conditions. The order of the seven sessions was counterbalanced based on a Latin Square design to compensate for learning effects.

Driving Task

The driving task was completed in a medium fidelity STISIM simulator using three 47-inch LCDs, shown in Figure 1. Drivers used a re-purposed Ford steering wheel and dashboard and STISIM gas and brake pedals as controls. Data were collected at 30 Hz and aligned with each simulation frame.

We created a custom driving scenario consisting of a straight four-lane road with two lanes of traffic in the driver's direction and two lanes of opposing traffic. The road included shallow hills that required the driver to adjust their throttle input to maintain a constant speed. Light traffic flowed in the opposing lanes and the left lane in the driver's direction.

Drivers completed a four-minute car-following exercise common among simulator studies [6, 26] that allowed us to measure how well the driver reacted to sudden braking events and maintained a consistent headway distance with the lead car. Drivers were instructed to stay in the right lane and maintain a two-second headway time behind a blue car traveling at 50 mph (80.5 km/h). At 50 mph a two second headway time is about 150 ft. (46 m). An orange car followed the driver and was visible in the rear-view mirror.

At four moments during the drive, we introduced a *sudden brake event* where the lead car suddenly decelerated from 50 mph to 37.5 mph and then accelerated back up to 50 mph, lasting about six seconds, as illustrated in Figure 2. These events allowed us to measure the driver's reaction time in relation to the braking of the lead car.

Each drive also had three *road context events* that did not require action, but were similar to what the Waze navigation app might tell a driver. The events include cars along the side of the road, bicyclists, motorcycles, police, crowded intersections, school zones, and debris, and were designed to measure the driver's reaction to distracting conditions.

We also created four *situation awareness events* by placing billboards, overhead signs, or buildings with signs and animal statues along the roadway, with fictional information about food, local businesses, or traffic information, similar to what may exist on an actual roadway. These events allowed us to roughly measure the driver's situation awareness and cognitive load.

The scenario was designed to be easy to drive, requiring the driver to primarily control their speed based on the hills and lead car's braking. All driving was done during dry daytime conditions with full visibility. We designed the scenario as an example situation when it might be appropriate for future systems to allow the driver to safely dual-task, as this type of low engagement driving is when drivers' minds are likely wander [50, 51]. Driving with heavy traffic, many turns, or many events on the road would most likely require more of the driver's attention and would not be suitable for incorporation of a secondary task [17].

Productivity Tasks

Participants either created written documents or presentation slides. One document or set of slides was created for each of the six study conditions that included the productivity task (the driving only baseline had no productivity task). For the writing task, participants dictated a one to two-page how-to guide on a topic of their choice. The how-to documents were intended to have a title, a goal, and a list of steps to complete

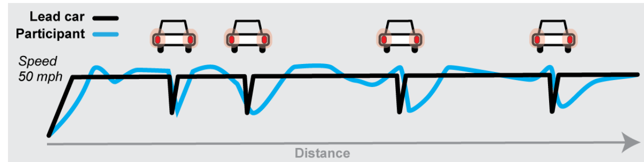

Figure 2: Participants followed a lead car trying to maintain a set distance behind. Four sudden brake events tested the driver's reaction time.

the goal. The task was primarily language based and focused on written material. For the presentation task, participants described a set of slides on a familiar topic. The presentation task included visual elements in addition to language-based content. These tasks were chosen because they are common, semi-structured, and require sufficient cognitive load. We also wanted to see if there were differences between linear (writing to-dos) and non-linear (creating presentation slides) tasks. We also aimed for the tasks to require multiple steps to meet a larger goal making them more challenging than small tasks on their own.

Before the study, participants in the writing condition were are asked to come up with six topics they were familiar enough with to write a how-to guide on, and participants in the presentation condition were asked to think of six projects. Topics could be related to work, school, home, or a hobby. We present the results of both the writing and presentation task conditions across the same set of measures. The driver's speech was recorded and the document they described was created by a crowdworker. Crowdworkers were instructed to transcribe both the content and the intent of the speech. The process took approximately one week.

Manipulations

In each condition we manipulated the level of support (low vs. high) and road context (low vs. high) that our speech-based assistant provided the participant with.

Level of Support During low support conditions, participants were asked to dictate their content, and the assistant led them through the task by asking *"OK, what's next?"* whenever they paused. This provided minimal support and did not reference any specific elements that the driver was creating. During high support conditions, the assistant used a set of microtask questions, shown in Table 1, to guide the participant. By breaking the task into subtasks, we hoped to capitalize on prior work that shows this can make a task more resilient to interruption [13] and would provide a prime for helping the driver quickly resume the task [1]. All level of support questions were triggered by the experimenter after the participant stopped speaking.

Road Context We tested two levels of road context support using speech-based alerts from the in-car assistant. During

Table 1: Task support questions for writing and presentation creation.

Writing
What is the main topic?
What is the goal of the tutorial?
What materials do you need?
What is the first step?
What is the next step? (repeated)
What other information should you look up?
Would you like anything else?

Presentation
What is the title of this slide?
What is the main point?
What layout would you like?
What text would you like?
What images would you like?
What would you like in the speaker notes?
What transitions would you like?
Would you like anything else on this slide?

the interaction, the assistant would stop any speech related to the productivity task and call out an upcoming road context event. Alerts were presented automatically 500 ft. from the road context event and occurred during and between microtasks. In low context scenarios, the assistant said, "*Please pay attention.*" In high context scenarios the assistant provided more detail, alerting the driver to event specifics and its location on the road. For example, it might say, "*Please pay attention to the construction zone on the right.*"

Users

Twenty-eight people with valid driver licenses were recruited via an ad circulated to a random sample of 2000 employees at Microsoft. Fourteen (7 F, 7 M) completed the writing task, with a mean age of 32.9 (SD = 9.6) and a mean commute length of 33 minutes (SD = 22), and 14 (3 F, 11 M) completed the presentation task, with a mean age of 33.6 (SD = 9.4) and a mean commute of 40 minutes (SD = 40).

Procedure

Upon arrival, participants filled out a consent form and completed a short questionnaire on their regular driving practices. The experimenter described the driving and productivity tasks, and the driver completed a short 1.5-minute test drive to familiarize themselves with the system. Drivers then practiced the speech-based productivity task, speaking aloud on their own and with guidance from the high and low task support questions, and drove another four-minute test drive, this one similar to the driving sessions they would experience throughout the rest of the study. They were informed that their primary task was driving and that their secondary task was the productivity task. Drivers then completed a short post-session questionnaire.

Once oriented, drivers completed the seven experimental sessions in a Latin Square order. After each session, they again completed the same post-session questionnaire. After all seven sessions, they completed a post-study questionnaire and participated in a short semi-structured interview about their experience. Approximately a week after the study they were sent the set of documents they created by crowdworkers, and asked to rank order them and provide written qualitative feedback. This time delay in receiving their documents was due to the time it took the crowdworkers to transcribe and create the document from the recorded speech.

4 Measures

In addition to the qualitative feedback we collected, we look at the following measures to understand our participants' performance on the driving and productivity tasks. The seven sessions, Driving, Thinking Aloud, Driving-Thinking Aloud (Drive-Talk), Low Context/Low Support (LC-LS), Low Context/High Support (LC-HS), High Context/Low Support (HC-LS), and High Context/High Support (HC-HS) were computed using one-way repeated measures analysis-of-variance (ANOVA). Post-hoc tests used Bonferroni correction. Effect sizes are reported using generalized eta-squared. Error bars represent standard error.

Quantitative Driving Performance

Headway distance A continuous measurement of distance between the lead car and the driver's car. Median headway distance was used to compare how well participants maintained a constant headway while driving. Minimum headway during braking was used to compare how close drivers were to colliding with the lead car.

Brake reaction time The time difference between the initial moment the lead car brakes and the moment when the driver pressed their brake pedal. Since reaction times may vary with headway distance, we also computed a headway-normalized reaction time by dividing the reaction time by the headway distance at the moment the lead car begins braking.

Speed, Acceleration, and Braking Measured continuously over the entire run. To see how well the driver matched the lead car's speed, we computed the squared correlation between the driver and the lead car speed, known as coherence. Coherence is a composite measurement describing how well the driver matched the lead car independent of the headway distance [6].

Lane position Measured continuously as the distance from the road's center median. The standard deviation of the lane position was computed as an estimate of the driver's ability to stay centered in the lane.

Post-Session Measures

Qualitative Experience Participants assessed their qualitative experience on a Likert scale (range: 1-Strongly Disagree to 7-Strongly Agree), rating their feelings of productivity (statement: “I got a lot done”), safety (statement: “I felt safe”), and ability to drive and do the task (statement: “I felt I could drive and do the productivity task well”). To measure their situation awareness, participants were also given two multiple choice questions asking which of four objects they saw on the road, as well as a fifth option of “None”. Each question included one object that was on the road and one that was not in any previous drives to avoid confusion between runs.

Task Support and Road Context Preferences Drivers rank ordered which type of task support and road context support they preferred (no, low, and high) after completing all the driving sessions.

Document Feedback Participants provided a rank order, from 1 to 7, of each document they created in terms of how well it captured their original intent and how useful it would be if they were to continue working on the document.

5 Results

We measured how driving performance and productivity were influenced by the voice assistant’s level of task support and road context support. We discuss our results in the context of our three research questions.

RQ1: How does level of support and road context support influence overall driving behavior?

Our results show that the productivity tasks did not critically change participant’s overall driving performance. We observe some increased median headway distances and decreased lane deviation. This aligns with previous distracted driving results showing that drivers will compensate when dual-tasking [18, 24] and have less steering wheel input [30]. A lack of difference in the coherence measure and brake reaction times suggests that none of the conditions significantly helped nor hindered the drivers in following the lead car and in responding to an event on the immediate roadway. This aligns with previous results showing that during easy segments of driving, driving performance while using a speech-based infotainment system was comparable to baseline driving [32].

Headway Distance To see how the task and context support influenced the driver’s following distance, we computed the median headway distance for each driving session and conducted a one-way repeated measures ANOVA across the driving session condition. For the Writing condition, Maulchy’s test indicated that the assumption of sphericity had been violated ($W = 0.11, p < .05$), therefore the degrees of freedom were corrected using Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.61$). The results show that median headway distance was significantly affected by the driving session $F(3.04, 39.5) = 4.67, p = .006, \eta^2_G = .05$. Drivers maintained a significantly higher median headway distance in the low context / low support ($M = 246, p = .035$) and low context / high support ($M = 227, p = .049$) sessions as compared the Driving baseline ($M = 185$), shown in Figure 3a.

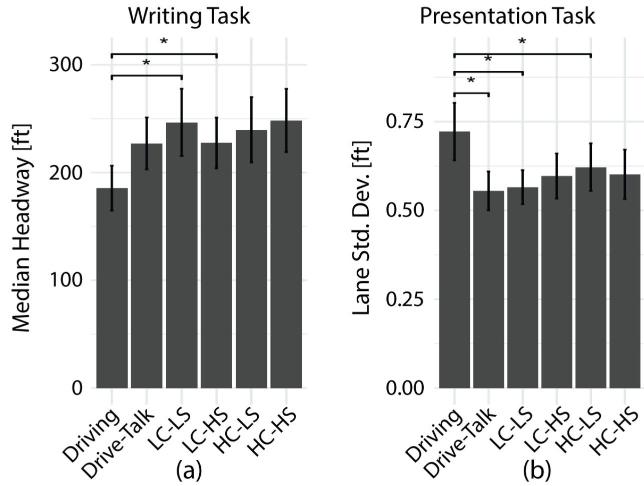

Figure 3: (a) Median headway distance for the writing task and (b) lane standard deviation for the presentation task.

For the Presentation condition, driving session did not significantly affect driver’s median headway distance $F(5, 65) = 1.32, p = .26, \eta^2_G = .05$. This suggests that for the writing task, when the system provided low context support, drivers may have compensated for the reduced safety by maintaining a greater distance with the car ahead. This was true for both high and low task support of the secondary task.

Coherence To see how task support and context support affect the driver’s ability to match the lead car’s speed, we computed the coherence of the driver’s speed to the lead car’s speed across the entire drive. Driving session did not significantly affect coherence in either the writing $F(5, 65) = .81, p = .55, \eta^2_G = .01$, or presentation $F(5, 65) = 1.02, p = .41, \eta^2_G = .04$ tasks. This suggests that neither the task support nor the road context support influenced how well the participant could follow the lead car and maintain smooth flow with traffic.

Standard Deviation of Lane Position We computed the standard deviation of driver’s lane position over each drive to estimate how well drivers could stay centered and avoid moving back and forth throughout their lane. For the Writing task, Maulchy’s test indicated that the assumption of sphericity had been violated ($W = 0.47, p < .01$), therefore the degrees of freedom were corrected using Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.5$). The results show that standard deviation of lane position was not significantly affected by the driving session $F(2.5, 32.5) = 2.67, p = .07, \eta_G^2 = .04$. For the Presentation task, driving session did significantly affect the standard deviation of lane position $F(5, 65) = 6.25, p < .001, \eta_G^2 = .05$. Drivers maintained better lane position, indicated by lower deviation in lane position, in the Drive-Talk ($M = .55, SD = .2, p = .038$), low context / low support ($M = .56, SD = .18, p = .032$), and high context / low support ($M = .62, SD = .25, p = .046$) session versus the Driving baseline ($M = .72, SD = .30$), shown in Figure 3b. This suggests that for the presentation task, drivers in low or no support scenarios could better maintain their lane position.

Situation Awareness We asked drivers to identify signs on the road. For the writing task, drivers recognition of road signs was significantly affected by session type $F(5, 65) = 15.4, p < .01, \eta_G^2 < .41$. Drivers were more likely to see signs during the driving baseline and the low support conditions. For the presentation task, driver's recognition of road signs was significantly affected by driving session $F(5, 65) = 14.0, p < .01, \eta_G^2 < .44$. Drivers recognized road signs significantly more during the driving baseline than in any other session. These results suggest that the productivity task may reduce the driver's ability to focus on objects off the road. In the case of the writing task, it appears that the low task support condition may have allowed participants to have more resources for seeing road signs, whereas the more visual presentation condition lead driver to see less.

RQ2: How do drivers behave during critical moments while driving and completing the productivity task?

To understand how drivers react to critical events, we look at average brake reaction time and minimum headway during braking, as these two measures indicate a required change of driving behavior in response to changes in the environment.

Average Brake Reaction Time For the writing task, Maulchy's test indicated that the assumption of sphericity had been violated ($W = 0.07, p < .01$), and we corrected the degrees of freedom using Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.13$). Driving session did not significantly affect driver's average brake reaction time during writing task sessions $F(2.04, 26.6) = 2.22, p = .13, \eta_G^2 = .04$. Normalizing for headway distance at the moment of the sudden brake also showed no significant difference across conditions.

For the presentation task, driving session also did not significantly affect driver's average brake reaction time $F(5, 65) = 2.46, p = .25, \eta_G^2 = .05$. Normalizing for headway distance at the moment of the sudden brake also showed no significant difference across conditions. These results suggest that even while doing the productivity tasks, drivers were able to react to sudden braking events just as well as if they were only driving.

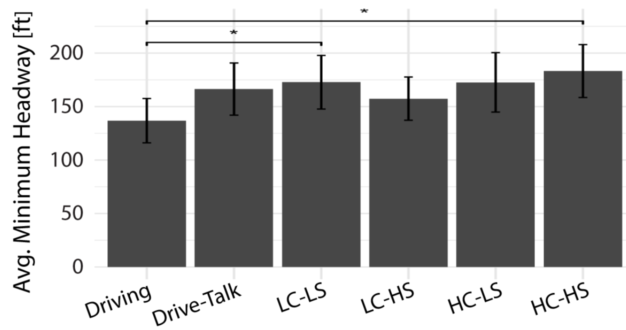

Figure 4: Average minimum headway under braking for the writing task

Minimum Headway During Braking We also looked at the average minimum headway distance during the braking events to measure how close to collision drivers would be. For the Writing condition, the results show that average minimum headway during braking events was significantly affected by driving session $F(5, 65) = 4.67, p = .005, \eta_G^2 = .03$. Post-hoc tests show that drivers on average kept significantly more distance during braking in the low context / low support ($M = 173, SD = 94, p = .035$) and the high context / high support ($M = 183, SD = 93, p = .035$) sessions as compared to the Driving baseline ($M = 136, SD = 76$).

For the Presentation condition, driving session did not significantly affect driver's average minimum headway distance during braking $F(5, 65) = 1.76, p = .13, \eta_G^2 = .05$. This suggests that at least in the writing condition, participants may have compensated for reduced safety and tried to maintain more headway distance from the lead car, however, there is no clear pattern of either task support or context support influencing this.

RQ3: How does task support and road context influence productivity task performance?

Word and Slide Counts After the documents were created using the crowdworking service, we recorded the word count for the writing task how-to guides and slide count for the presentation task. We compared the six sessions that included document creation including the thinking aloud session as a baseline (see Figure 4). For the writing task, document word count was significantly affected by driving session $F(5, 65) = 7.17, p < .001, \eta_G^2 = .13$. Average word count in the Drive-Talk condition ($M = 367, SD = 120$) was not significantly different from the talking only condition ($M = 406, SD = 151$). The high context / low support session had significantly lower word count ($M = 302, SD = 83, p = .016$) than the talking only session. In comparison with the drive-talk condition, the low context / low support

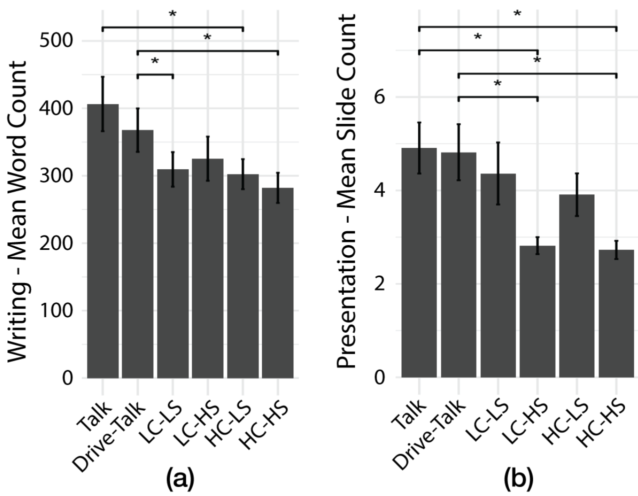

Figure 5: (a) Word count and (b) slide count across all productivity task sessions.

($M = 309, SD = 96, p = .038$) session and the high context / high support session ($M = 282, SD = 84, p = .015$) had significantly lower average word counts.

For the presentation task, three participants elected to not have their video reviewed by the crowd workers and have their documents created. Thus, only 11 participants are included in this analysis. Slide count was significantly affected by driving session $F(5, 50) = 6.62, p < .001, \eta^2_G = .25$. Average slide count was not significantly different between the talking only baseline $M = 4.9, SD = 1.8$) and the driving and talking session ($M = 4.8, SD = 2.0$). Slide count was significantly lower in the high context / high task session ($M = 2.72, SD = .65, p = .034$). In both cases, the output for Thinking Aloud and Driving+Talking were similar, but lower with high task support and road context support. This suggests that the assistant may have helped drivers better balance their attention between driving and the productivity task, the trade off being less content creation while still maintaining driving performance similar to driving only.

Subjective Feedback

While the driving behavior data showed no overall issues with driving performance, many drivers felt that the task of creating documents would be too challenging for a real-world driving situation. Drivers did feel that they would be able to complete shorter and simpler tasks such as creating a checklist or responding to an email during moments where their drive had limited on-road complications.

Safety We asked participants how safe they felt on a, 1-Not at all to 7-Very Safe, Likert scale after each driving session. For the Writing task, Maulchy's test indicated sphericity had been violated ($W = 0.043, p = .002$), therefore the degrees of freedom were corrected using Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.47$). Session type did not significantly affect feelings of safety $F(2.81, 36.4) = 1.84, p = .17, \eta^2_G = .065$.

For the Presentation task, feelings of safety were significantly affected by session type $F(5, 65) = 5.5, p < .01, \eta^2_G = .14$. The Driving+Talking session ($M = 5.3, SD = 1.14, p = .05$) and the low context / low support ($M = 4.5, SD = 1.65, p = .008$) sessions were rated as significantly less safe than the driving only baseline ($M = 6.28, SD = 1.07$). This suggests that drivers may have felt more safe when the assistant provided specific road context alerts or asked specific task questions. The decreased feelings of safety during the driving+thinking aloud and low support / low context task may also be due to drivers focusing more attention on the productivity task and less attention on the road.

Productivity To assess how productive drivers felt, we asked them to rate the statement "I felt like I got a lot done." on a 1 - Not at all to 7 - Very Much Likert scale after each drive. We included the thinking aloud session as baseline. For the writing task, drivers did not report any significant differences in their productivity across sessions. For the Presentation task, feelings of productivity were significantly affected by session type $F(5, 65) = 9.8, p < .001, \eta^2_G = .24$. Post-hoc comparisons showed that the high context / high support session was rated as significantly less productive ($M = 3.93, SD = 1.49, p = .007$) than the thinking aloud only baseline ($M = 5.5, SD = 1.34$). This suggests that drivers felt they were able to complete the writing task well no matter the task support or road context support, but drivers completing the presentation task found the high context / high support condition made them less productive. This may be because the presentation condition had more focused questions which may not have aligned with how drivers were thinking about creating their presentation. Additionally, the high context support may have led drivers to spend more time focusing on the road and less on their slides.

Driving and Working We asked drivers to rate how well they were able to drive and work at the same time. Drivers reported no significant difference in their ability to drive and complete the task based on driving session or task type.

Task and Context Support Preferences

After all the driving sessions, we asked participants to rank order their preferences for task support (No support, low, high) and road context support (No support, low, high).

Task Support For writing, 8 participants preferred the high support most, 2 preferred low support, and 3 preferred no support. For presentation, 4 participants preferred the high support most, 4 preferred low support, and 6 preferred no support. These results (Figure 6a) suggest that either something about the task support or the task itself may lead to different preferences for task support. For the writing condition, the task support could help people remember all the sections of a document they were creating. Additionally, the writing task was fairly linear, whereas the slide creation can be more non-linear, potentially explaining why many drivers preferred the no support conditions.

Context Support For writing, 7 participants preferred the high context support most, 1 preferred low support, and 6 preferred no support. For presentation, 5 participants preferred the high context support most, 4 preferred low support, and 5 preferred no support. These results (Figure 6b) suggest a split between drivers who found the high context alerts helpful and those who would have preferred not to hear them. In the post-interview, drivers would often say that they would enjoying having this feature all the time or comment that it was distracting.

Document Feedback

After reviewing the created documents, we asked drivers to rank order the documents and tell us on a 1 to 5 scale how useful the document might be to continue working. For the writing condition, there was no clear pattern with document preferences. This may be due to the fact that the how-to writing task was not very different across conditions. For the presentation condition, drivers generally rated their thinking aloud or driving+talking presentations higher and more useful. This may be due to the fact that many drivers preferred to outline many slides rather than try to create polished slides. For example, presentation participant 7 said, "*If I could do nothing more than talk and organize my thoughts in an 'outline', that would be really useful (and then use that outline to go back and manually generate slides).*"

6 Discussion

As voice-based assistants become more commonplace for in-car interactions, there may be opportunities to allow drivers to be productive, but these interactions must be designed to preserve driver safety. Any intelligent assistant must consider the non-driving task along with the driving situation in order to maintain safety *and* productivity. To our knowledge, our study is the first to explore experiences with an in-car agent that attempts to balance safety with a desire to get things done using task support and road context. We specifically aimed to see how drivers could use easy driving

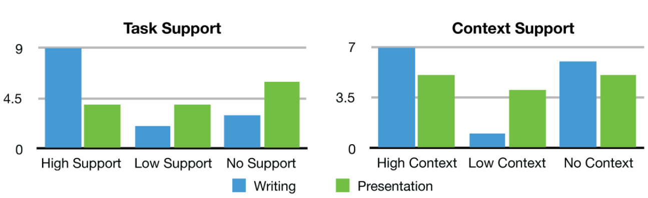

Figure 6: Preferences for task support and context support.

time when they may naturally mind wander [34] to produce content that would be of value to them. Combined, these results support the multiple resource theory [46, 47] approach we have considered and suggest that with careful design it may be possible for people to drive and engage with a productivity task via a speech interface.

Implications for Design

While the driving performance measures do not favor one level of support over another, the qualitative results suggest that the level of support may be an individual preference and that it is important to determine the right level of support for each person as the mismatch in preference can cause a distracting level of load. Some participants found the high support conditions helped their thinking and allowed them to move efficiently through a task; others felt that the high support questions took them in a different direction from where their mind was going. For example, writing task participant 1 said "*The more structured sections where I was prompted to enter the next step in a process felt too constrained and made me spend extra time thinking about how to word my idea to fit the prompt.*" This may suggest designing personalized support systems offering different workflows. More importantly, a system which overloads the driver due to causing a large shift in their thinking could have severe consequences during the wrong moment. Many of our participants said that the productivity task was too challenging and on the limit of what they could do, even with an easy driving activity in a simulator. Our tasks were designed to challenge drivers and while these exact tasks may be too tough to complete while driving, they do give us a sense of how complex productivity tasks in the car could be. Future systems could be designed to allow drivers to conduct simpler tasks and could monitor the driver's cognitive load and adjust the task support based on how well the driver is completing the task.

Our objective and subjective productivity results also suggest that other types of productivity tasks may be better suited to a driving environment. Participants in the presentation condition reported in the post-interview that many of the detailed questions about each slide such as background color, transitions, and speaker notes were unnecessary and slowed them down. Instead, participants often preferred to create a higher level outline and placeholders for slides that they could then later fill in with details. For the high support writing conditions, the activity of creating how-to guides was fairly linear and mostly consisted of the assistant asking what the next step was. This was not much different from the low task support "*Ok, what's next?*" The few additional questions for the writing task that asked drivers to provide an introduction, list materials, and add any additional information were viewed as good things to have and made for more complete documents. Thus, the majority of the participants preferred the high task support. Given these results, future task design for in-car productivity may consider tasks that are more linear or outline oriented. For example, interactions that help users prepare for or ramp down from productivity tasks may be suitable in this context [48].

In terms of the road-context support, while previous work has found evidence that specific support improves driving [21] we do not have clear evidence that specific calls to pay attention to the road scene improved driving ability over non-specific calls. However, participants completing the presentation task evenly preferred the high context, low context, and no context driving alerts. Participants completing the writing task typically preferred either high context support or no support at all.

Drivers who preferred high context saw value in it even without a secondary task. Others felt that the low context alert may have caused more distraction since identifying the cause of the alert resulted in added cognitive load. Overall, we learned that even though the interactions of the in-car assistant were grounded in prior work and hypothesized to be more suitable for a divided attention scenario, in practice there are more nuances to the design of such systems. Interactions should adapt to the cognitive load of the driver. Further, there will always be trade-offs between productivity and safety, however, safety must always be prioritized. Thus, as shown by the reduction in content creation, productivity can suffer. In a situation where cognitive resources are dynamically reallocated based on the task demands and safety must be prioritized, it is probably advisable to simplify requirements of the productivity tasks and only present tasks that do not require deep cognitive engagement.

Limitations

Our study was conducted in a simulator and although simulation based studies are crucial to the development of new driving technologies and are known to translate into real-world driving scenarios [11], they still do not provide the fully realistic assessment how how drivers will behave on the road [16]. Additionally, because we tested our productivity task in an easy driving scenario, our results cannot tell us how people will respond during very severe events on the road. Indeed, our participants recognized this with many reporting that they felt they could do the task, however thought it still may not be possible in a real car. Further work is required to understand how productivity tasks and road context support can either help or hinder drivers during more challenging moments of driving.

Another limitation with our study is the lack of control on the content that the participants created. Some participants noted the task of creating documents was more challenging when working on something harder to explain or a project that they had done in the past. It is possible that this may have influenced people's abilities to complete the task and to focus on driving and suggests that further study is required to understand how content and one's personal sense of cognitive load may influence driving performance.

Future Work

Given our findings and limitations, future work could explore better ways of breaking down larger document creation tasks using speech-only interfaces, focusing on identifying microtasks with low cognitive load that would be suitable for interleaving with driving. Furthermore, it may be useful to determine what types of microtasks may support continuing work on a document, such as outlining or taking down notes.

Beyond the level of task support, future work could explore what types of road context support can help drivers without increasing their cognitive load. For example, providing support for events further down the road could help drivers manage the event better or may lead to drivers focusing too much attention on searching for the event. Given that a driver's cognitive load influences both safety and ability to complete the productivity task, there are opportunities to use real-time cognitive load estimation to help manage both secondary tasks and safety alerts.

Finally, while our study looked at what the experience with a speech assistant for drivers who were actively driving, it would be interesting to see how drivers would perform in an autonomous vehicle. In these cases, speech-based productivity tools may be appropriate for helping keep drivers aware of the road and ready for takeover (in Level 2/3 vehicles) and to allow people to work without the risk of motion sickness caused by looking at a screen (in Level 4/5 vehicles).

7 Conclusion

The opportunities for allowing drivers to reclaim some of their commute time for productive work may be realized with the help of intelligent in-car assistants. The results of our study show that drivers were able to complete writing and presentation slide creation tasks using speech alone on very simple roads without significant reduction in driving performance. Still, many drivers felt that the task was quite challenging and found that the assistant's level of task support and road context support would either help them or lead to increased load. Given this, there may be opportunities to develop both better tasks and context alerts that better fit the way the drivers think, work, and drive.

Acknowledgments

This work was supported by Microsoft Research during the first author's summer internship. The authors thank Mary Czerwinski for providing valuable feedback during pilot testing and Ivan Tashev for technical help with the simulator.

References

[1] Erik M. Altmann and J. Gregory Trafton. 2002. Memory for goals: An activation-based model. *Cognitive Science* 26, 1 (2002), 39–83.

[2] Paul Atchley and Mark Chan. 2011. Potential benefits and costs of concurrent task engagement to maintain vigilance: A driving simulator investigation. *Human Factors* 53, 1 (2011), 3–12.

[3] Brian P. Bailey and Joseph A. Konstan. 2006. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior* 22, 4 (2006), 685–708.

[4] Adriana Barón and Paul Green. 2006. *Safety and usability of speech interfaces for in-vehicle tasks while driving: A brief literature review.* Technical Report. University of Michigan, Transportation Research Institute.

[5] Jelmer P. Borst, Niels A. Taatgen, and Hedderik van Rijn. 2015. What Makes Interruptions Disruptive?: A Process-Model Account of the Effects of the Problem State Bottleneck on Task Interruption and Resumption. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2971–2980. https://doi.org/10.1145/2702123.2702156

[6] Karel Brookhuis, Dick de Waard, and Ben Mulder. 1994. Measuring driving performance by car-following in traffic. *Ergonomics* 37, 3 (1994), 427–434.

[7] Duncan P. Brumby, Andrew Howes, and Dario D. Salvucci. 2007. A Cognitive Constraint Model of Dual-task Trade-offs in a Highly Dynamic Driving Task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 233–242. https://doi.org/10.1145/1240624.1240664

[8] Corinne Brusque and Aline Alauzet. 2008. Analysis of the individual factors affecting mobile phone use while driving in France: Socio-demographic characteristics, car and phone use in professional and private contexts. *Accident Analysis & Prevention* 40, 1 (Jan. 2008), 35–44. https://doi.org/10.1016/j.aap.2007.04.004

[9] US Census Bureau. 2017. Means of Transportation to work by selected Characteristics - S0802. https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_17_5YR_S0802&prodType=table (Accessed: 2019-01-03).

[10] Carrie J. Cai, Philip J. Guo, James R. Glass, and Robert C. Miller. 2015. Wait-Learning: Leveraging Wait Time for Second Language Education. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3701–3710. https://doi.org/10.1145/2702123.2702267

[11] Jeff K. Caird, Chelsea R. Willness, Piers Steel, and Chip Scialfa. 2008. A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention* 40, 4 (2008), 1282–1293. http://www.sciencedirect.com/science/article/pii/S0001457508000183

[12] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40, 4 (2017), 834–848.

[13] Justin Cheng, Jaime Teevan, Shamsi T. Iqbal, and Michael S. Bernstein. 2015. Break It Down: A Comparison of Macro- and Microtasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 4061–4064. https://doi.org/10.1145/2702123.2702146

[14] Johan Engström, Emma Johansson, and Joakim Östlund. 2005. Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour* 8, 2 (2005), 97–120.

[15] Charlotte Eost and Margaret Galer Flyte. 1998. An investigation into the use of the car as a mobile office. *Applied Ergonomics* 29, 5 (Oct. 1998), 383–388. https://doi.org/10.1016/S0003-6870(98)00075-1

[16] Stuart T. Godley, Thomas J. Triggs, and Brian N. Fildes. 2002. Driving simulator validation for speed research. *Accident Analysis & Prevention* 34, 5 (2002), 589–600.

[17] Peter A. Hancock, Mary Lesch, and Lisa Simmons. 2003. The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis & Prevention* 35, 4 (2003), 501–514.

[18] Tim Horberry, Janet Anderson, Michael A. Regan, Thomas J. Triggs, and John Brown. 2006. Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention* 38, 1 (2006), 185–191.

[19] Christopher Ingraham. 2017. The American commute is worse today than it's ever been. *Washington Post* (Feb. 2017). https://www.washingtonpost.com/news/wonk/wp/2017/02/22/the-american-commute-is-worse-today-than-its-ever-been/ (Accessed: 2017-09-18).

[20] Shamsi T. Iqbal, Piotr D. Adamczyk, Xianjun Sam Zheng, and Brian P. Bailey. 2005. Towards an Index of Opportunity: Understanding Changes in Mental Workload During Task Execution. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. ACM, New York, NY, USA, 311–320. https://doi.org/10.1145/1054972.1055016

[21] Shamsi T. Iqbal, Eric Horvitz, Yun-Cheng Ju, and Ella Mathews. 2011. Hang on a Sec!: Effects of Proactive Mediation of Phone Conversations While Driving. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 463–472. https://doi.org/10.1145/1978942.1979008

[22] Shamsi T. Iqbal, Eric Horvitz, Yun-Cheng Ju, and Ella Mathews. 2011. Hang on a Sec!: Effects of Proactive Mediation of Phone Conversations While Driving. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 463–472. https://doi.org/10.1145/1978942.1979008

[23] Shamsi T. Iqbal, Yun-Cheng Ju, and Eric Horvitz. 2010. Cars, Calls, and Cognition: Investigating Driving and Divided Attention. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1281–1290. https://doi.org/10.1145/1753326.1753518

[24] A. Hamish Jamson, Stephen J. Westerman, G. Robert J. Hockey, and Oliver M. J. Carsten. 2004. Speech-Based E-Mail and Driver Behavior: Effects of an In-Vehicle Message System Interface. *Human Factors* 46, 4 (2004), 625–639. https://doi.org/10.1518/hfes.46.4.625.56814 PMID: 15709325.

[25] Christian P. Janssen, Duncan P. Brumby, and Rae Garnett. 2012. Natural Break Points: The Influence of Priorities and Cognitive and Motor Cues on Dual-Task Interleaving. *Journal of Cognitive Engineering and Decision Making* 6, 1 (2012), 5–29. https://doi.org/10.1177/1555343411432339

[26] Christian P. Janssen, Shamsi T. Iqbal, and Yun-Cheng Ju. 2014. Sharing a Driver's Context with a Caller Via Continuous Audio Cues to Increase Awareness About Driver State. *Journal of Experimental Psychology: Applied* 20, 3 (2014), 270–284.

[27] Harmanpreet Kaur, Alex C. Williams, Anne Loomis Thompson, Walter S Lasecki, Shamsi T Iqbal, and Jaime Teevan. 2018. Creating Better Action Plans for Writing Tasks via Vocabulary-Based Planning. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 86.

[28] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and*

Manufacturing (IJIDeM) 9, 4 (Nov. 2015), 269–275. https://doi.org/10.1007/s12008-014-0227-2

[29] Eric Laurier. 2004. Doing Office Work on the Motorway. *Theory, Culture & Society* 21, 4-5 (Oct. 2004), 261–277. https://doi.org/10.1177/0263276404046070

[30] John D. Lee, Brent Caven, Steven Haake, and Timothy L. Brown. 2001. Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors* 43, 4 (2001), 631–640. http://journals.sagepub.com/doi/abs/10.1518/001872001775870340

[31] Victor Ei-Wen Lo and Paul A. Green. 2013. Development and evaluation of automotive speech interfaces: useful information from the human factors and the related literature. *International Journal of Vehicular Technology* 2013 (2013).

[32] Jannette Maciej and Mark Vollrath. 2009. Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accident Analysis & Prevention* 41, 5 (Sept. 2009), 924–930. https://doi.org/10.1016/j.aap.2009.05.007

[33] Angela Mahr, Margarita Pentcheva, and Christian Müller. 2009. Towards System-mediated Car Passenger Communication. In *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '09)*. ACM, New York, NY, USA, 79–80. https://doi.org/10.1145/1620509.1620525

[34] Malia F. Mason, Michael I. Norton, John D. Van Horn, Daniel M. Wegner, Scott T. Grafton, and C. Neil Macrae. 2007. Wandering Minds: The Default Network and Stimulus-Independent Thought. *Science* 315, 5810 (Jan. 2007), 393–395. https://doi.org/10.1126/science.1131295

[35] Alexander Meschtscherjakov, David Wilfinger, Thomas Scherndl, and Manfred Tscheligi. 2009. Acceptance of Future Persuasive In-car Interfaces Towards a More Economic Driving Behaviour. In *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '09)*. ACM, New York, NY, USA, 81–88. https://doi.org/10.1145/1620509.1620526

[36] Yoshiro Miyata and Donald A Norman. 1986. Psychological issues in support of multiple activities. *User Centered System Design: New Perspectives on Human-Computer Interaction* (1986), 265–284.

[37] Michael E. Rakauskas, Leo J. Gugerty, and Nicholas J. Ward. 2004. Effects of naturalistic cell phone conversations on driving performance. *Journal of safety research* 35, 4 (2004), 453–464.

[38] Dario D. Salvucci and Peter Bogunovich. 2010. Multitasking and Monotasking: The Effects of Mental Workload on Deferred Task Interruptions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 85–88. https://doi.org/10.1145/1753326.1753340

[39] David M. Sanbonmatsu, David L. Strayer, Arwen A. Behrends, Nathan Ward, and Jason M. Watson. 2016. Why drivers use cell phones and support legislation to restrict this practice. *Accident Analysis & Prevention* 92 (July 2016), 22–33. https://doi.org/10.1016/j.aap.2016.03.010

[40] David L Strayer, Jonna Turrill, James R Coleman, Emily V Ortiz, and Joel M Cooper. 2014. Measuring Cognitive Distraction in the Automobile II: Assessing In-Vehicle Voice-Based. *Accident Analysis & Prevention* 372 (2014), 379.

[41] Ivan Tashev, Mike Seltzer, Y. C. Ju, Ye-Yi Wang, and Alex Acero. 2009. Commute UX: Voice Enabled In-car Infotainment System. (September 2009). https://www.microsoft.com/en-us/research/publication/commute-ux-voice-enabled-in-car-infotainment-system/

[42] Jaime Teevan, Daniel J. Liebling, and Walter S. Lasecki. 2014. Selfsourcing Personal Tasks. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 2527–2532. https://doi.org/10.1145/2559206.2581181

[43] Remo M.A. van der Heiden, Shamsi T. Iqbal, and Christian P. Janssen. 2017. Priming Drivers Before Handover in Semi-Autonomous Cars. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 392–404. https://doi.org/10.1145/3025453.3025507

[44] Shari P. Walsh, Katherine M. White, Barry C. Watson, and Melissa K. Hyde. 2007. *Psychosocial factors influencing mobile phone use while driving*. Report. Australian Transport Safety Bureau, Canberra, ACT. https://eprints.qut.edu.au/11305/

[45] Katherine M. White, Melissa K. Hyde, Shari P. Walsh, and Barry Watson. 2010. Mobile phone use while driving: An investigation of the beliefs influencing drivers' hands-free and hand-held mobile phone use. *Transportation Research Part F: Traffic Psychology and Behaviour* 13, 1 (Jan. 2010), 9–20. https://doi.org/10.1016/j.trf.2009.09.004

[46] Christopher D. Wickens. 2002. Multiple resources and performance prediction. *Theoretical issues in ergonomics science* 3, 2 (2002), 159–177.

[47] Christopher D. Wickens. 2008. Multiple Resources and Mental Workload. *Human Factors* 50, 3 (2008), 449–455. https://doi.org/10.1518/001872008X288394 PMID: 18689052.

[48] Alex C. Williams, Harmanpreet Kaur, Gloria Mark, Anne Loomis Thompson, Shamsi T. Iqbal, and Jaime Teevan. 2018. Supporting Workplace Detachment and Reattachment with Conversational Intelligence. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 88, 13 pages. https://doi.org/10.1145/3173574.3173662

[49] Wayne Xiong, Jasha Droppo, Xuedong Huang, Frank Seide, Mike Seltzer, Andreas Stolcke, Dong Yu, and Geoffrey Zweig. 2016. Achieving human parity in conversational speech recognition. (2016). arXiv:arXiv:1610.05256 Retrieved from https://arxiv.org/abs/1610.05256v2.

[50] Matthew R. Yanko and Thomas M. Spalek. 2013. Route familiarity breeds inattention: A driving simulator study. *Accident Analysis & Prevention* 57 (Aug. 2013), 80–86. https://doi.org/10.1016/j.aap.2013.04.003

[51] Matthew R. Yanko and Thomas M. Spalek. 2014. Driving With the Wandering Mind: The Effect That Mind-Wandering Has on Driving Performance. *Human Factors* 56, 2 (March 2014), 260–269. https://doi.org/10.1177/0018720813495280

[52] Kristie Young, Michael Regan, and M Hammer. 2007. Driver distraction: A review of the literature. *Distracted Driving* 2007 (2007), 379–405.