Priming Drivers before Handover in Semi-Autonomous Cars

Remo M.A. van der Heiden
Utrecht University
Utrecht, The Netherlands
r.m.a.vanderheiden@uu.nl

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

Christian P. Janssen
Utrecht University
Utrecht, The Netherlands
c.p.janssen@uu.nl

ABSTRACT

Semi-autonomous vehicles occasionally require control to be handed over to the driver in situations where the vehicle is unable to operate safely. Currently, such handover requests require the driver to take control almost instantaneously. We investigate how auditory pre-alerts that occur well before the handover request impact the success of the handover in a dual task scenario. In a study with a driving simulator, drivers perform tasks on their phone while the car is in an autonomous mode. They receive a repeated burst audio pre-alert or an increasing pulse audio pre-alert preceding the standard warning for immediate handover. Results show that pre-alerts caused people to look more at the road before the handover occurred, and to disengage from the secondary task earlier, compared to when there was no pre-alert. This resulted in safer handover situations. Increasing pulse pre-alerts show particular promise due to their communication of urgency. Our detailed analysis informs the design and evaluation of alerts in safety-critical systems with automation.

Author Keywords

Autonomous cars; multitasking; handover; automation

ACM Classification Keywords

H.1.2 [User/Machine Systems]: Human information processing; H.5.2 [User Interfaces]: Benchmarking; I.2.9 [Robotics]: Autonomous vehicles

INTRODUCTION

With the advent of novel (in-)vehicle technologies, new challenges emerge in managing driver distraction. Research has shown that drivers perform a variety of tasks that may distract from driving including interacting visually and manually with their mobile phones [17,31]. A recent meta-review suggests that distraction is likely to increase even

further in 'self-driving' or 'autonomous' vehicles, where cars assume more of the driving responsibilities [16].

In such vehicles, gradations of automation can be identified, as defined in various standards [19,39,47]. At the lowest level of automation (e.g., no automation, or level 0 in [47]), the human driver is in full control of the car. In full automation (e.g., level 5 in [47]), the human driver is not involved in any driving task anymore. This is, for example, the vision of the Google autonomous car (e.g., [53])

For the levels in between these extremes (e.g., levels 1-4 in [47]) there is some form of shared responsibilities. For example, even if the car is driving by itself, the driver might need to intervene when the system is uncertain about what action to take. For ease of reference we will use the umbrella term "semi-autonomous vehicle" to describe this wide category. The important characteristics of these systems for our work is that the vehicle can drive relatively independently for some time, but can request the human driver to assist or take over control through an alert.

A natural question that is of relevance to the CHI community is then: what should the specifics of such an alert be? Indeed, various aspects of handover [20] and in-car alerts [44,54] have received attention in the literature. Currently alerts, such as in the Tesla model S, include a brief alert and immediately handover control to the driver. However, given that the driver may not have the proper situational awareness to be able to immediately take the proper action, immediate alerts as such can potentially result in fatal outcomes. Moreover, with higher levels of automation (e.g., level 3 in [47]) such alerts are even more crucial as the car is monitoring the environment for large time intervals and drivers might have disengaged (cf. [16]).

In this research we investigate whether providing a *pre-alert*, an additional alert that commences well before the actual handover request (in our study: 20 s) can help drivers be better prepared to safely navigate the incident for which the handover of control occurs. Designing such a pre-alert for semi-autonomous vehicles presents unique challenges, as drivers distract themselves more with other tasks as automation of the car increases [16]. If they are then asked to take-over, they should not have remnants of their preceding task (e.g., checking e-mail, reading the news) that inhibit their ability to successfully take over (cf. [52]). At the same time, tasks that drivers engage in while in an autonomous vehicle may be more important than the typical

secondary task in today's manual cars, making it more challenging to suddenly disengage [22].

In a simulator study, we study the effects of early warnings, or pre-alerts, on handover performance in dual task scenarios. We investigate three research questions: (1) how do pre-alerts affect behavior before take-over including eye-gaze and suspension of the non-driving task, (2) how do drivers perceive the pre-alerts, and (3) how do the pre-alerts affect driving performance. We present two types of pre-alerts: (A) A *repeated burst audio pre-alert*, and (B) an *increasing pulse audio pre-alert*. While the car drives itself, drivers occasionally perform a video transcription or calendar entry task on a mobile phone. Results showed that pre-alerts helped drivers prepare better for taking over control by increasing gaze on the road and earlier suspension of the phone task, followed by quicker reaction to traffic incidents compared to having no pre-alerts.

In the remainder of this paper we first provide more background on multitasking, managing multitasking, and in-car alerts. We then describe our study. Finally, we discuss the results, including their implications for theory and design, limitations, and potential for future work.

RELATED WORK

Driving and Multitasking
Multitasking is a prevalent practice while driving [17,31]. Multiple studies have documented the detrimental effects of cell phone conversations, texting, and interacting with in-vehicle systems while driving [1,9,10,51,46]. Switching from a non-driving task to driving is often challenging. For example, drivers who engage in phone conversations have slower braking reaction time [1,33], degraded steering performance [10], and a higher likelihood of accidents [45] than drivers with no distractions.

In the majority of this preceding work, the driving task has typically been considered the 'primary task' where the driver's focus is expected at all times, with other tasks as 'secondary'. However, as the automation technology matures, human drivers might be required less and less to take-over control of the car. Therefore, they might increasingly engage in other tasks (cf. meta-review in [16]), and those other tasks (e.g., checking e-mails, reading the news, having a conversation, watching a video) might even feel as a primary task with driving as a distraction (cf. [22]).

Managing Multitasking during driving
A known strategy to manage multitasking is to interleave activities, which has been well documented in conceptual frameworks (e.g., [2,6,8,50]). In the domain of driving, multiple studies have looked at how drivers interleave non-driving, secondary tasks with driving. A common strategy is to wait for 'natural breakpoints' in the task to switch attention [11,12,26,27,28,48]; for example, Iqbal et al. showed that drivers chunked a task of providing directions while driving into multiple steps and reoriented to driving at the boundaries between chunks [26]. There are many

advantages of interleaving at natural breakpoints: it reduces mental workload [3, 49] as it reduces information that needs to be maintained in memory [7], it frees mental resources such as visual attention for other tasks [50,55], it reduces stress [4], it reduces the time needed for later task resumption (cf. [2]), and it can offer speed-accuracy trade-offs in dynamic environments such as driving [28].

In autonomous vehicles, however, while people are not driving, the non-driving task might capture most of the driver's attention [22]. In a hand-over scenario, people might therefore not want to immediately let go of whatever they were working on. Priming a handover in a timely manner through pre-alerts, as we propose in this paper, has the advantage that it allows drivers to disengage from their 'primary' (non-driving) task at a pace that suits them. Moreover, gradually disengaging from the task and waiting for a natural breakpoint can also benefit the driving task. If interleaving at a more opportune task reduces workload [3,49] and stress [4], then people are in a better mental state to resume driving.

(In-car) Alerts
The idea of using alerts to gain user attention has been explored in many domains. Mediation or alerting has been proposed as one of the four interruption management methods in McFarlane's work [35]. In the driving domain, researchers have explored the effectiveness of systems for aiding driving by providing local danger alerts [14], mediating communications among car passengers [34], or persuading people to drive in a more economical manner [36]. For example, Iqbal et al. explored how alerts can gain user attention while driving and conversing on a cellphone simultaneously, and found that alerts reduced driving errors, while also reducing conversation quality [25].

In the domain of autonomous driving, the idea of using alerts before handover of control has recently gained traction, based on the concern that the current designs of immediate take over may not yield the desired outcome. Most work has focused on the design of the alert, in terms of the timing and modalities so that it conveys the required urgency [20,32,40, 44,54]. For example, Gold et al. [20] showed that alerts happening 7 seconds before the incident resulted in more successful take overs compared to alerts that were presented 5 seconds before. Walch et al. also investigated different timings and modalities [54]. While they found no difference in driving performance, drivers had a preference for the alerts reinforced through both auditory and visual means. There was no significant advantage in driving performance.

Perhaps closest to our work is Politis et al.'s study which tested language based alerts for upcoming incidents [44]. Alerts were delivered via audio, visual, or tactile means. Results showed that drivers quickly transitioned to the driving task for warnings that conveyed urgency, and performance was worst for unimodal visual alerts. Other work has looked at using auditory cues to provide drivers in autonomous vehicles awareness about the environment [5].

This work does not separate the specific scenario of handover where the driver's awareness is put to test by having to react to an incident in very short notice.

Compared to existing work, our research focuses less on the exact timing or ideal conveyance of urgency via the alerts. Rather, we draw upon designs of alerts that have been effective in conveying urgency in a timely manner [18, 21]. Our goal is to understand if a pre-alert (i.e., an early alert well before the final warning) is useful in general, and if so, why and how it supports ease of disengagement from a secondary task. Such a scenario can be crucial in autonomous vehicles.

USER STUDY

In our study of the effects of pre-alerts on preparing for handover, we aim to address three research questions:

RQ1: What do drivers do before handover? We therefore analyze eye-gaze and time on secondary task.

RQ2: How do drivers experience the handover? We therefore look at subjective ratings and physiology.

RQ3: How successful is the handover? We therefore look at the first reaction time, speed reduction, and an analysis of unsafe incidents.

To answers these questions, we conducted a user study using a driving simulator with autonomous capabilities where drivers engage in a non-driving task on a mobile device and are required to take over control in certain driving situations where the car is unable to continue.

Users

Twenty-four drivers (12 M; 12 F) with an average age of 32.5 years (*SD* = 9.6) were selected by quota sampling. Each driver had a valid driver's license and drove on average 5.4 days a week (*SD* = 2.4). All drivers provided informed consent and were compensated with a $50 gift card.

Tasks

Drivers performed two task types: driving (part manual, part autonomous), and a non-driving task on a mobile phone.

Driving task

A simulated driving task was developed in a medium fidelity simulator. Three 47" TVs projected the driving environment. Drivers sat in an adjustable car seat behind a full Ford dashboard. During manual driving, drivers used a steering wheel, gas and brake pedal (transmission was automatic).

Simulation software consisted of the STIsim simulator software. Data was recorded at a rate of 10 data points per second. An eye tracker was mounted on the dashboard to capture eye gazes on the driving scene. A custom scenario was developed, consisting of a drive on a two lane country road with bends and straight road segments without intersections. Oncoming traffic was presented on the opposite lane, not in the driver's lane.

The scenario consisted of manual driving and automated driving trajectories. The car started stationary and drivers

had to press the gas pedal to start driving manually. Drivers had to maintain the posted speed limit and remain in lane. Occasional curves were included in each scenario. The curves were subtle enough so braking was not needed but input of drivers was required to stay in their lane safely.

After 1500 feet (25-30 s after the start), an automated voice would state "Automation enabled", and the car assumed driving control. At that point, if desired, drivers could release the steering wheel and the pedals. The car continued to drive itself until it warned drivers when they were needed to take over the driving task. While the car was in auto drive there was no driver initiated way to get back control. In the event of a handover, the car would start by warning the driver and the controls were handed over back to the driver after a voice said: "Automation disabled". Optionally, such a warning could be preceded by a pre-alert, 20 s before handover. The pre-alerts are described in more detail later.

There were four handover scenario varieties, each with two instances – resulting in eight scenarios in total. Examples are shown in Figure 1. The *fog scenarios* had light or heavy fog, and required drivers to slow-down, maintain their lane, and avoid other cars. The *construction works scenario* had cones along the road and required drivers to slow down and steer accurately. One version also included a lane change. The *parked car scenario* had a car that blocked either part of the road, or the full road, requiring the driver to slow down or stop. The *dog scenario* had a dog abruptly crossing the street, requiring the driver to stop in order to avoid hitting the dog.

The scenarios required different types of responses, such as braking and accurate steering. Some scenarios allowed for multiple responses such as braking and steering away from an accident. However, although the simulator allowed for this variety of maneuvers, some were not safe. This is similar to how a driver in real traffic can sometimes respond in different ways, of which only some are safe.

Non-driving tasks

Recent reviews suggest that drivers distract themselves with non-driving tasks more when automation in the car increases, which impacts situational awareness and response times [16]. We therefore also included conditions in which drivers performed non-driving tasks. As the structure of the task can also impact when drivers look at the road [12, 27, 28], we used two tasks. Half the drivers performed a video transcription task while the other half performed a calendar task. These tasks were chosen because they represent tasks that people might prefer to do in an autonomous car – such as watching videos [44] and performing short typing tasks, as reported on a pre-survey. All secondary tasks were conducted on a Nokia Lumia 1520 phone with Windows 8.1.

Video task. A custom developed app (see Figure 2, left) showed a video screen and an input box. Drivers had to play the video (which showed elementary statistics lectures [24])

Figure 1. Four handover scenarios: (a) Fog, (b) Dog crossing, (c) Parked car on the side of the road, (d) Construction works.

and had to transcribe it in the textbox. Controls to play, pause, and forward in the video were embedded in the player. We used the standard keyboard from Windows phone 8.1, with auto-correction and -completion disabled. This allowed for more reliable measurement of writing performance. We logged the timestamp of each keypress.

Calendar task. The alternative task was a calendar task (see Figure 2, right). Drivers were asked to enter event information in a simple calendar interface. The interface had two separate screens, one showed all the upcoming events as a digital flyer, the other showed the input boxes and saved items. We again logged the timestamp of each keypress.

Warning & Pre-alerts
At each hand-over instance, in all conditions, drivers received a voice warning: "automation disabled" 1 s before the handover. Drivers were instructed to take over the driving from the system at this moment. In the *pre-alert* conditions, drivers also received either of two pre-alerts 20 s prior to this final warning. For the pre-alerts we used individual 500 Hz beeps that lasted 150 ms per beep. Beep frequency and length were recommended in the warning literature [38].

Repeated burst audio pre-alert: In this condition, bursts of 3 beeps are played 3 times, with silence in between burst sets. The burst sets started playing at 20, 10, and 1 seconds before the final warning (i.e., in Figure 4: 0, 10, and 19 seconds relative to start of pre-alert phase). Repeated burst alerts are already used in the car context, for example, to notify when a door is not closed firmly or when a car is almost out of fuel.

Increasing pulse audio pre-alert: In this condition, beeps are given throughout the 20 second pre-alert time. However, the interval between consecutive beeps is reduced gradually over time as the driver gets closer to the critical moment. The initial inter-stimulus interval is 1000 ms. The final inter-stimulus interval is 50 ms. An increasing pulse audio alert is already used at other places in the car domain to suggest increased urgency. For example, park assistant alerts decrease the inter-stimulus interval between beeps when a car gets close to another object to suggest urgency to stop. In other studies, increasing pulses (e.g., heartbeats) have also been used to successfully convey urgency [29].

No pre-alert: In this baseline condition, no pre-alert is given and drivers are only warned by the final warning voice 1 s before handover of control.

Figure 3 shows a schematic of the different stages in a single run. There are two experimental segments in each run. Each run started with a period of driving by the driver, followed by a period of auto drive during which the driver could engage in a non-driving task (depending on condition). During the auto drive there would be optionally pre-alerts (depending on experimental condition), followed by a final voice warning declaring the handover to the driver. This would be followed by the handover event during which the driver had to start driving. After a while, the second segment started, following the same procedure. The entire run was about 6 minutes long (i.e., roughly 3 minutes per segment).

EXPERIMENTAL DESIGN
We used a 3 x 2 within-subjects design. We manipulated Pre-alert type (*Repeated burst*, *Increasing pulse*, or *No pre-alert*), and number of tasks (*single-task* driving, or *dual-task* driving with secondary task). For each of the six combinations of pre-alerts and number of tasks, we developed one drive/run. Each drive consisted of two hand-over moments that had a critical incident. In sum, each participant had six drives and twelve handover situations.

Ordering of the six conditions was counterbalanced following a Latin Square design, to compensate for learning effects. As we had eight critical events, we randomly assigned these to drivers with the requirement that each run had one incident that might require braking (i.e., dog or parked car) and one incident that required accurate steering (i.e., cones or fog). Finally, we also measured single-task performance as baseline (explained in procedure).

For the secondary task, drivers either performed the video or calendar task (12 drivers per task, randomly assigned).
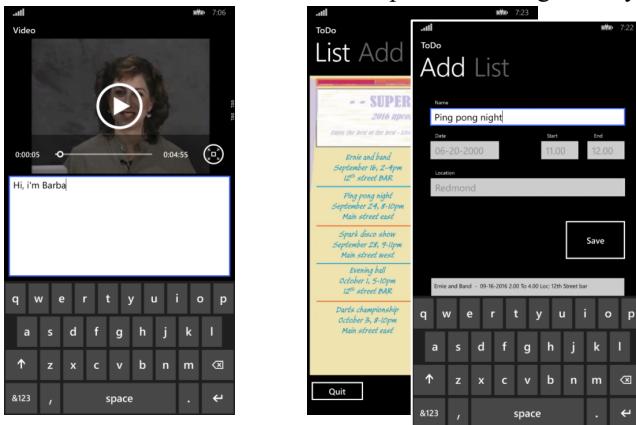

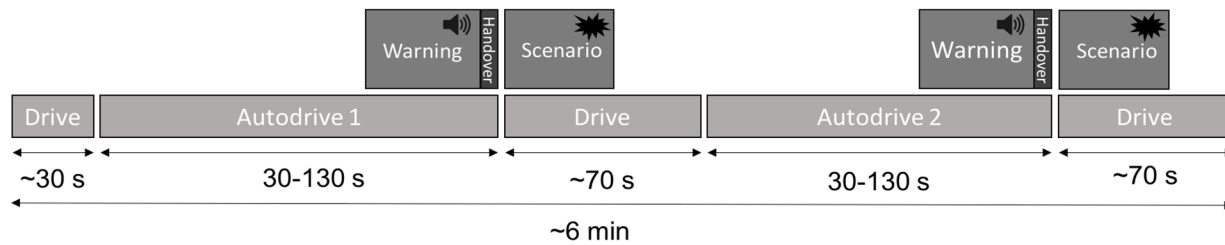**Figure 2. Layout of Video task (left) & Calendar task (right)**

Figure 3. A schematic representation of a run for one condition with two consecutive segments.

PROCEDURE

On arrival, drivers were given an overview of the study and asked to sign an informed consent form as well as to fill out a questionnaire about current driving behavior. They were then asked to make themselves comfortable in the car seat of the simulator, which was adjusted so their foot could reach the pedals and their eyes were visible for the eye tracker.

We then calibrated the eye tracker and set participants up with the Microsoft band2, which was used for measuring heart rate. This was followed by a 2-minute training session on the secondary task (video or calendar, depending on group), followed by training with the driving task. In this single-task practice drive, drivers practiced all 4 types of handover events, and were introduced to the two pre-alerts. The first event was preceded by an increasing pulse pre-alert, the second with a repeated burst pre-alert, the last two had no pre-alert.

The remainder of the experiment consisted of the six experimental drives/runs. In addition, single-task performance with either the video or calendar task (depending on assignment) was performed at a random position in between the experimental drives. After all trials ended the drivers filled out a general questionnaire on their overall experience and preferences about the pre-alerts.

MEASUREMENTS

Below we define our exact measurements, sorted by research question (RQ). Unless otherwise noted, we use a 3 (pre-alert) x 2 (number of tasks) within-subjects analysis of variance (ANOVA) with an alpha-level of .05 for significance. Where needed we use Holm-Bonferroni-corrected post-hoc tests. Error bars in all plots show standardized error of the mean.

Gaze during driving (RQ1) We used an SMI REDn eye tracker which reported tracking status (eyes tracked or not) and X,Y gaze-coordinates at 30 samples/second. The eye tracker was positioned on the dashboard just above the steering wheel. For short drivers an extra cushion was used to make sure their eyes stayed visible during the entire experiment, as tested before the experiment.

For all gaze metrics, the eye-tracker could only track the eyes when the user was looking at the simulator screen (not at the secondary task). We therefore define "looking at the road" as gaze samples where the eye tracker was at least tracking one eye, and "not looking at the road" as moments where no eyes were tracked. Given the large size of the screen and the

peripheral location of the phone, this crude metric is a good approximation of actual looking at the road. Based on this information, we calculate what *percentage of drivers* look at the road (e.g., Figure 4) and what *percentage of the time* drivers on average look at the screen. In pilot studies we crosschecked with 2 eye trackers that the eyes were consistently detected when watching the simulator screen.

Disengaging from the secondary task (RQ1). Based on logs of touchscreen keypresses, we determined the interval between the start of the pre-alert phase and the last keypress. Shorter intervals indicate a faster disengagement.

User preferences (RQ2). In a questionnaire, drivers indicated their preferences after the experiment. The questions used a five-point scale ranging from low/poor (1) to high/good (5).

Heart rate (RQ2). A Microsoft Band2 measured the number of beats per minute. One value was logged per second.

Initial reaction time (RQ3). For each handover event we measured the reaction time as the time interval between the moment automation was disabled and the first action of the driver (either a brake press or steering wheel input).

Driver speed reduction (RQ3). For each handover event, we measured at what speed the driver drove, as logged by the simulator at a rate of 10 samples/second.

Unsafe incident analysis (RQ3). For the first 10 seconds after handover, we manually labeled whether observed behavior was unsafe, following conservative but realistic pre-defined rules. In all scenarios, driving more than 10 mph over the posted speed limit, or leaving the highway was labeled unsafe (some drivers drove on the grass). In addition, in one *parked car scenarios* only a full stop avoided a collision; not doing so was labeled unsafe. Also, for one *dog scenario*, some drivers crossed into the lane of incoming traffic, despite that cars might come in. This was also labeled unsafe.

RESULTS

We measured how handover performance changes with the use of pre-alerts and the use of a secondary tasks. We discuss our results in the context of our three research questions.

RQ1: Gaze & phone task engagement before handover

Percentage of drivers looking at the road
Figure 4 shows the percentage of drivers looking at the road in the different stages of handover, relative to the start of the pre-alert (time point 0). In single-task trials (grey lines) we

found that at each time point, on average 70 to 80% of the drivers looks at the road. There is a slight increase in the phase before the handover, but this is not different between the various alert conditions.

The pattern is different for the dual-task condition (dark lines). In the phase before the pre-alert (-9 to 0 s), drivers look more at the road when they are in the no pre-alert condition (dashed line) compared to the two pre-alert conditions. This is likely because drivers know that they have no pre-alert to rely on, and want to be prepared for a later handover request.

During the period where the pre-alert is active (0 to 20 s), in both pre-alert conditions (Repeated burst, solid line; increasing pulse, dotted line), drivers look more frequently at the road compared to the no pre-alert condition (dashed line). Moreover, this accumulates over time, as drivers get closer to the handover itself. The two pre-alert conditions are hard to distinguish from one another. Qualitatively, in the repeated burst pre-alert condition drivers start looking at the road after each burst of tones (0, 10, and 19 s). The strongest bump is after the first burst. By contrast, for the increasing pulse pre-alert there is a more gradual increase over time.

In the condition where there is no pre-alert, the percentage drivers looking at the road initially is similar to before the start of the pre-alert. However, 10 seconds before handover the percentage of drivers looking at the road increases, even though they have not yet received an alert. This is because in the simulator parts of the critical event gradually become visible even before the handover request occurs. This is similar to how in real driving visible cues are sometimes available ahead of time (e.g., a traffic jam ahead). Such visual cues make non-distracted drivers look at the scene and prepare. The take-away message from this graph though is that in both pre-alert conditions, drivers look up a lot *earlier* and do *not* rely on input from the road *at the last moment*.

Gaze during auto-drive before the pre-alert period
We quantified the preceding results. First, we looked at the entire auto-drive period before the pre-alert. For each driver, we calculated the percentage of time that they looked at the road. An ANOVA showed that there is a significant effect of number of tasks on the percentage of time drivers spend watching the road $F(1,23) = 344.3$, $p < .001$, $\eta_p^2 = 0.94$. Drivers looked at the road more than ten times as much in single task condition ($M = 71\%$, $SD = 19\%$) compared to dual task condition ($M = 6\%$, $SD = 13\%$). There was also a significant effect of pre-alert, $F(2,46) = 4.441$ $p = .002$, $\eta_p^2 = 0.16$. Post-hoc tests found that gazes at the road during No pre-alert ($M = 9\%$, $SD = 4\%$) was significantly higher than Repeated burst ($M = 4\%$, $SD = 3\%$, $p = .001$) and Increasing pulse ($M = 5\%$, $SD = 4\%$, $p = .016$), as in the no pre-alert condition drivers cannot rely on a signal to warn them, they interleave the tasks more often to check for hazardous situations. The two pre-alerts did not differ from each other ($p > .1$). There was no significant interaction effect ($p > .1$).
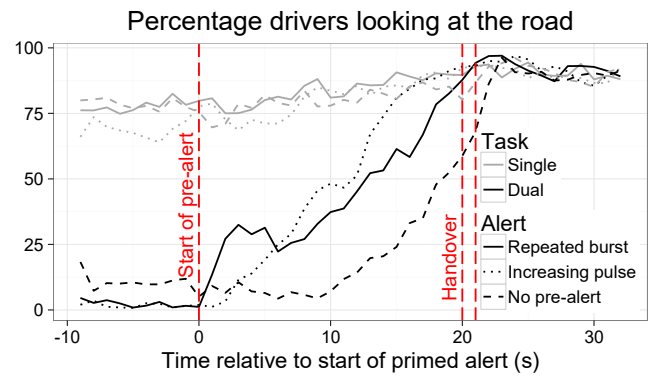

Figure 4. The percentage of drivers looking at the road relative to the start of the primed warning.

Gaze during pre-alert phase
During the pre-alert phase (i.e., 0-20 s in Figure 4), there was again a significant main effect of number of tasks on percentage of time drivers spend watching the road $F(1,23) = 168.7$, $p < .001$, $\eta_p^2 = 0.88$. In general, drivers were looking at the road about twice as much in single task condition ($M = 83\%$, $SD = 18\%$) compared to dual task condition ($M = 42\%$, $SD = 29\%$). There was also a significant main effect of pre-alert, $F(2,46) = 16.28$, $p < .001$, $\eta_p^2 = 0.41$. However, both main effects were affected by a significant interaction effect, $F(2,46) = 15.95$, $p < .001$, $\eta_p^2 = 0.41$. Post-hoc tests revealed that in single-task there was no significant difference between the three pre-alert conditions (all $ps > .1$). However, in dual-task, the percentage gaze at the road was significantly lower in the no pre-alert condition ($M = 23\%$, $SD = 15\%$) compared to repeated burst ($M = 49\%$, $SD = 19\%$) and increasing pulse ($M = 54\%$, $SD = 16\%$) all $ps < .001$. This is expected, as the pre-alerts warn drivers to look at the road and drivers therefore indeed gaze more at the road.

Disengaging from the secondary task
To test whether the alert helped drivers disengage from the secondary phone task, we tested how long they continued after the alert had started using a 3 (Pre-alert type) x 2 (Secondary task type) ANOVA. There was a significant effect of pre-alert on the time drivers continue their phone task after alert onset, $F(2,44) = 30.09$, $p < .001$, $\eta_p^2 = 0.56$. Figure 5 shows the data. Post-hoc tests showed that all three
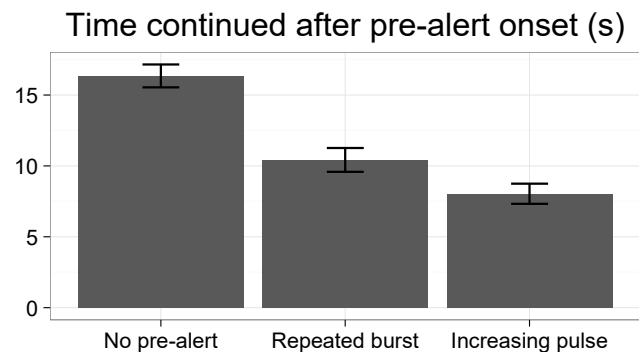

Figure 5. The average time secondary task continued after different alert onsets.

conditions differed significantly from each other (with the difference between increasing pulse and repeated burst with $p = .038$, all other $ps < .001$). As the figure shows, in the increasing pulse condition drivers quit the secondary task twice as fast compared to the no pre-alert condition. The ANOVA also revealed a marginal effect of secondary task, $F(1,22) = 3.81$, $p = .06$, $\eta_p^2 = 0.15$. Disengagement was slightly faster in the calendar task ($M = 10.5s$, $SD = 4.7s$) than the video task ($M = 12.7s$, $SD = 6.0s$). There was no interaction effect, $F(2,44) = 1.5$, $p > .1$.

RQ2: User experience

User preferences
Subjective feedback revealed overall preference for pre-alerts was divided. Twelve drivers preferred the increasing pulse pre-alert. Feedback included that these drivers liked that they could finish their task and prepare for handover. Twelve other drivers preferred the repeated burst pre-alert, as it felt less disruptive than the increasing pulse.

In the post questionnaire drivers provided various scores for the different pre-alerts on a five-point scale with anchors for low (1) and high (5). In Figure 6 we present the histograms of the score for the metrics (1) annoyance, and (2) disruptiveness of the pre-alert. The responses are again divided, as reflected in the broad distributions. Some trends are that the increasing pulse is reported more frequently as conveying high to too much urgency, and being more frequently considered as highly annoying and disruptive. However, there are also drivers who reported an inverse pattern (e.g., rated annoyance low).

Heart rate
We also tested if the pre-alerts had any effect on drivers' physiology, specifically the average heart rate. Two drivers had to be excluded from this analysis because the measurement stopped during the experiment. The ANOVA found a significant effect of pre-alert on average heart rate $F(2,40) = 4.72$, $p = .015$, $\eta_p^2 = 0.18$. Post-hoc tests found that in No pre-alert ($M = 69.8$ bpm, $SD = 6.7$ bpm) the heart rate is significantly higher than in Repeated burst ($M = 68.9$ bpm, $SD = 6.6$ bpm, $p = .034$) and Increasing pulse pre-alert ($M =$

69.2 bpm, $SD = 6.8$ bpm, $p = .046$). There was no difference between Repeated burst and Increasing pulse ($p > .1$) In addition, heart rate was significantly higher in Dual-task conditions ($M = 69.8$ bpm, $SD = 6.6$ bpm) compared to Single-task ($M = 68.8$ bpm, $SD = 6.8$ bpm), $F(2,40) = 4.719$, $p = .031$, $\eta_p^2 = 0.21$. There was no significant interaction effect ($p > .1$) In summary, heart rates are slightly increased in dual-task conditions, and when there is no pre-alert and the user needs to do the extra task of frequently checking the road. This suggests that extra workload increased heartrate (cf. [37]). Though the effect is small, the trend is consistent with subjective data.

RQ3: Success of handover
For the next two measures, we only analyzed the first segment of each driving scenario, as due to a coding error the driving speed was slightly lower on the second segment (60ft/s vs 65ft/s), which affected the time between the end of the pre-alert and the time given to hand-over. In the second segment, there was some delay between the warning that automation was turned off and the time that drivers could actually control the car. This reduces the reliability of the reaction time data on these metrics. This is not the case for the segments that we analyzed.

Initial reaction time
Previous work has mostly analyzed reaction times as a performance metric for handover. Due to the varied nature of our task, reaction to an event can be either braking or steering. We measure reaction time as time until either of these two actions occurs. We combine the data of single- and dual-task trials, resulting in three histograms for each pre-alert condition in Figure 7. Each plot shows data from 48 trials: 24 single, 24 dual. The bars cover 200 ms intervals. In general, most drivers respond within 200 ms (i.e., the first bar is the highest bar in each setting).

However, in driving analyses, we do not only care about the mean and majority of behavior, but also about extremes. This is where conditions differ. In the no pre-alert condition (left), the distribution is more right-tailed (9 trials with response longer than 600 ms, with extremes up to 2.5 s) compared to the repeated burst (5 trials) and increased pulse pre-alert (2 trials). Stated differently: in most cases, most drivers respond timely, but the trend is that more drivers respond timely when a pre-alert is given, with the number of late responses around twice as high in the no pre-alert condition.

Driver speed reduction
We also analyzed driving speed after handover. In critical situations, reducing speed is a smart strategy as this creates more time for an effective response (as less distance is covered per time interval). Figure 8 shows how the average
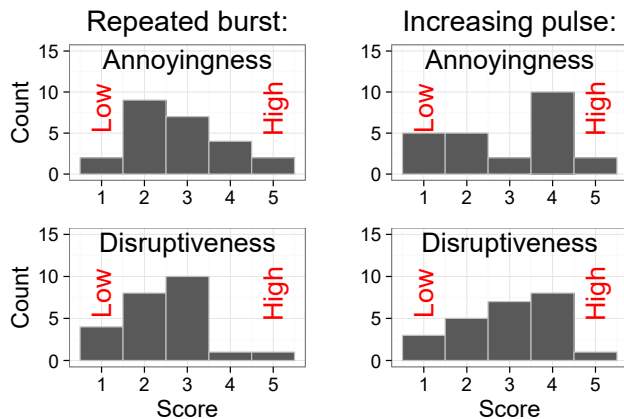

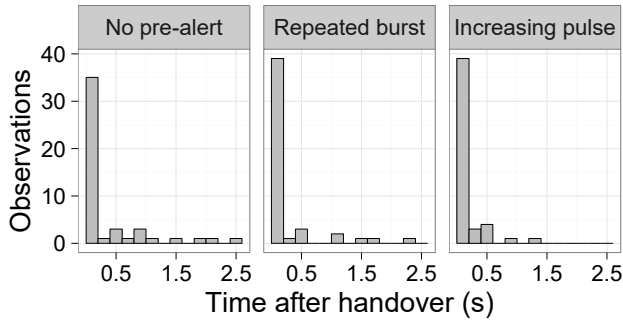Figure 6. Subjective impression of pre-alerts

Figure 7. Time until first action after handover.

Figure 8. Driving speed after handover.

speed reduces over time in single-task (top figure) and dual-task (bottom figure) for the three pre-alert conditions.

Before the handover (point 0 in the graph), the car drives automatically at a constant speed. Then, at point 0, drivers can take over. In the single-task condition, drivers reduce speed immediately in each condition, with a slight delay in the no pre-alert condition. However, if 95% CI intervals would be drawn, these would overlap between all three conditions, suggesting there is no difference. It shows that without distraction, drivers are prepared to respond.

This is different in the dual-task condition. Here, the brake response in the no pre-alert condition is delayed compared to the two pre-alert conditions. That is, drivers respond later. In fact, if 95% CIs would be drawn around the lines, the confidence intervals between the no pre-alert condition do not overlap with those of the two pre-alert conditions. A standard interpretation of a lack of overlap is that the conditions differ at a 95% confidence level (or with an alpha of .05) [15]. The intervals of the two pre-alert conditions do overlap, indicating there is no significant difference.

Unsafe incident analysis
The final question is: Did the pre-alerts lead to less unsafe behavior? If drivers had more time to look at the road before handover to build situational awareness (RQ1), do they perform better in handling the handover incident?

For this analysis, we looked at all scenario segments (288) and labeled whether unsafe behavior had occurred such as not reducing speed, going into a lane where incoming traffic might occur, or crashing into an object. Of the 288 segments, 34 segments (11.8 %) were marked as unsafe. Given the low numbers, we report the frequency of unsafe behaviors in Table 1, split up by pre-alert type (rows), and single- or dual-task (columns). There is no clear emerging pattern. Specifically, incidents still occur in the two active pre-alert
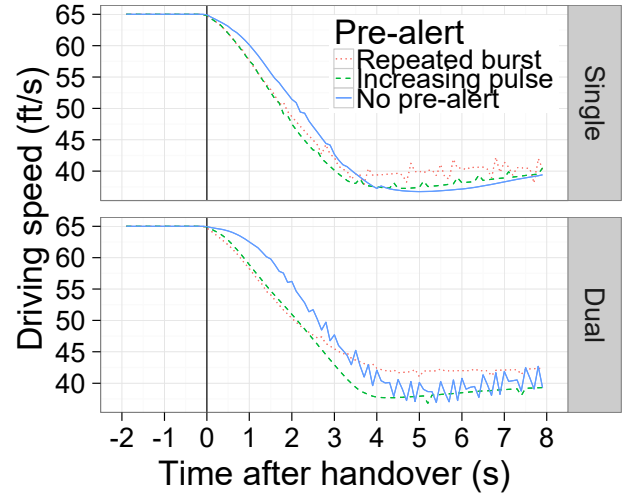
conditions. Moreover, it is not the case that there are more unsafe behaviors in the dual-task condition compared to single-task condition. If any, the trend in the data is that there were fewer unsafe behaviors in the increasing pulse condition (third row).

Due to this variety of results, we also analyzed whether unsafe behaviors were more frequent for some scenarios compared to others. This could not be done in combination with the type of pre-alert or single-/dual-task, due to the low numbers. In general, unsafe behavior was shown most in scenarios involving braking for a dog (23 segments). This was much higher compared to segments with a parked car (5 segments), construction works (6 segments), or fog (0 segments). Our interpretation of this result is that the dog scenario allowed more freedom to the user of what to do: they could either try to brake, or try to avoid it by driving past it. Moreover, as the object on the road (a dog) was relatively small compared to, for example, a parked car, drivers might not have had extra benefit from the early view of the object due to a pre-alert.

GENERAL DISCUSSION

Summary of results
We investigated how pre-alerts affect the hand-over of control from a semi-autonomous car to a human driver. A recent review showed that drivers distract themselves more with other tasks as automation in the car increases, and that this impacts their situational awareness and ability to respond correctly [16]. Based on theory, we expect that a pre-alert has four benefits (1) it allows a driver the necessary time to disengage from a secondary task [6], (2) this can reduce mental workload [3, 49] and stress [4] and leave drivers in a better state to manage the handover, (3) this allows drivers more time to reorient to the driving task and gain relevant situational awareness, and (4) with sufficient time distracting effects from the non-driving tasks may be reduced (cf. [52]). Our results demonstrate that pre-alerts are indeed beneficial. During the alerting phase (RQ1), drivers disengage from

	Single-task	Dual-task	Total
None	7	6	**13**
Repeated burst	5	8	**13**
Increasing pulse	6	2	**8**
Total	**18**	**16**	**34**

Table 1. Number of segments involving unsafe behavior

their non-driving tasks earlier when a pre-alert is given (Figure 5), and they look earlier at the road (Figure 4). In their experience (RQ2), drivers were divided in which alert they preferred, though both conveyed some urgency (Figure 6). We also found an effect on heart rate, which was slightly higher in the condition without a pre-alert. This finding needs to be replicated before conclusions can be drawn, but might indicate that a situation with no pre-alert is more stressful, as drivers cannot rely on the pre-alert to notify them. Finally, when looking at driving performance (RQ3), drivers responded faster (by braking or steering) to incidents when they were warned by a pre-alert (Figure 7), and reduced their speed more quickly (Figure 8). In effect this allows them more time to respond to an incident (i.e., at a lower speed it takes longer before an incident location is reached). Finally, there were still unsafe behaviors in all conditions, but these were the lowest with an increasing pulse pre-alert (Table 1).

Between the two pre-alert types, the preferences of drivers were divided. However, the trend in most metrics is that the increasing pulse pre-alert leads to slightly safer performance. For example, in this condition people disengage the earliest from a secondary task (Figure 5), more people experienced it as conveying high urgency (Figure 6), it had the lowest number of slow handovers (Figure 7), and the lowest number of incidents (Table 1).

Given this pattern, a general conclusion is that pre-alerts are useful compared to not getting a pre-alert. However, there is still room for improvement, as unsafe actions still occur.

Implications for theory
Our results confirm the classical result in driver distraction research (e.g., [1,9,10,28,29,33,45,46,51]) that secondary tasks distract from looking at the road (RQ1, Figure 4) and result in longer response times to incidents (RQ3, Figure 7). Our results also suggest that pre-alerts can mitigate some of these problems. This is particularly useful, as driver distraction occurs frequently in regular cars [17,31] and increases with an increase in autonomy of the car [16].

The use of alerts for hand-over situations in semi-autonomous cars is of course not new. However, in contrast to earlier work (e.g., [20]), we focus on alerts that happen ahead of time (*pre*-alerts). Sending an alert too early (e.g., minutes) in advance might not make sense, as there is no situation that the driver can notice and start to anticipate. We focused on a pre-alert of 20 seconds, as previous work has suggested that distraction of secondary tasks can continue up to 27 seconds after the task was finished, with exponential decay [52]. Twenty seconds is therefore an interval that is needed to recover from distractions and to focus on the road.

In our study, drivers still incurred incidents in some of the alert conditions. There are multiple explanations, which require further testing in future studies. First and foremost, there were incidents even in single-task situations, suggesting that some tasks were difficult to handle in general. Second, in the distraction conditions drivers might

have persisted too long with the secondary tasks even after the pre-alert and thereby not have taken enough time to react. Finally, similar to [52], even when drivers did finish secondary tasks, drivers might have had remnants of distraction. The balance that needs to be found here is between giving a just-in-time alert such that it is meaningful, while also giving sufficient time to overcome any negative effects of distraction. Further studies are needed to get this balance just right.

Although both pre-alerts that we offered were effective, the increasing pulse performed better on some metrics compared to the repeated burst (e.g., Figure 5,7). Our interpretation is that this is because increasing pulse more clearly conveyed urgency (as also confirmed by the drivers, see Figure 6). This is also in line with theory (e.g., [44]).

As our study takes place in a simulator, there are concerns about how the findings generalize. For example, there are no serious consequences of a crash. However, it also offers many advantages for our setting. First, we can test behaviors in cars that have a level of autonomy beyond those that are currently widely available. Second, we can test extremely dangerous situations such as hand-overs preceding crashes that would be unethical to test on the road. Third, we can measure behavior in-depth with multiple measures including eye-gaze, physiology, preferences, and driving performance. Finally, meta-reviews have demonstrated that situations that are shown to be dangerous in simulator conditions are also dangerous on the road [13, 23]. However, the effect size of the performance decline can differ between the two situations. A specific prediction that our work makes for the regular road is that pre-alerts can be beneficial, but also that even an interval of twenty seconds might not be enough to respond appropriately to an incident in a handover situation. This is particularly the case because everyday traffic is more diverse and perhaps less predictable than our simulator scenarios.

Implications for design
Our findings suggest that pre-alerts can be helpful in managing handover situations. This opens up a large space of future work that explores what the exact nature of pre-alerts can be. Below, we discuss some relevant parameters.

Convey urgency
Our results suggest that pre-alerts should provide a sense of urgency (cf. [29,44]), as in our increasing pulse pre-alert. However, there is still a wide design space to explore regarding exact choices. Relevant parameters include the exact length and timing of the pre-alert, the modality of the pre-alert, and the ability to perhaps also turn a pre-alert off.

Encoding more information in pre-alerts
Our pre-alert only used beeps to indicate that handover had to take place. However, it might be beneficial to also inform the user about why the pre-alert is raised (for example, is a sensor not working, does the car notice traffic), and what concrete actions they are to take (e.g., "Scan your

surroundings to see whether you can come to a stop, or can go to another lane"). Preceding work has suggested that such concrete alerts are helpful in the automotive domain [25]. We also used only one modality (audio) for the alert. Exploring multimodal alerts may result in better outcomes [54].

Balance effectiveness with less annoyance
Although the increasing pulse pre-alert trended to be slightly more effective out of the two pre-alerts, many drivers did not like it because it appeared to be annoying. Although the primary function of an alert is to increase safety, a distaste of the pre-alert might disrupt drivers too much. This would also counter two of the benefits of the pre-alert: to allow time to finish a task and to get in a low workload, low stress state. Future work should explore how design can counter this.

Timing of pre-alerts
While we selected a fixed time interval for the pre-alert (20 s before the handover) the timing could also be made dynamic depending on the type of event, level of distraction of the driver, and complexity of the required action. Dynamic timing helps address situations where a pre-alert occurring too early may diminish its urgency, or a pre-alert occurring too late may be deemed useless.

Although our focus has been on the automotive domain, our results can also be applied to other domains in which there is (A) shared control between humans and systems and (B) potential distraction. The implication there is that a pre-alert can benefit the shift of control from system to user. Our results suggest that pre-alerts that convey urgency, such as our increasing pulse pre-alert, are valuable. However, for each domain more tests are needed to determine the timing of these pre-alerts. This should take into account (1) the remnant effects of distraction [52], (2) the time needed to finish any preceding task, and (3) the time that is needed to gain situational awareness in the domain at hand.

Limitations & Future work
We conducted our work in a driving simulator to allow for an in-depth study of human behavior, in an environment where the users cannot incur harm, but that is known from meta-reviews to translate to everyday driving [13, 23]. However, this also has limitations. First, there are no real risks for drivers and they might therefore have acted slightly riskier than in normal life. Second, given the experimental set-up, they might have anticipated some incidents and hand-overs which lowered response times compared to driving on the road. In everyday life, alerts might be rarer [56], which impacts response time. Third, the study was measured over a relatively short interval (90 minutes per participant) with various incidents, whereas normal driving has incidents less frequently. These limitations do not differ from other valuable driver distraction studies that used simulators, but are to be taken into account nonetheless.

The pre-alerts that we tested were limited in scope. We have discussed relevant parameters to explore. Some specific limitations are the following. First, the use of other modalities and multi-modal pre-alerts (cf. [41,43,44]) needs to be tested. Second, we did not test voice-based commands despite their potential (e.g. [30,42,44]). Finally, our drivers were not able to turn off the pre-alert, whereas this might be a relevant option in real cars, for example, to signal to the car that you noticed the pre-alert.

In our measurements we have tried to give a detailed description of human behavior. However, there is room to go even more detailed. First, our physiological state results were subtle and need to be replicated before solid conclusions can be drawn. Second, our eye-tracking results gave insight in whether and when drivers looked at the road, but more detailed analyses regarding *where* they look would also be beneficial (i.e., to understand what information people gather, what information they might have overlooked).

Finally, in our study we tested technology based on what technology looks like today (e.g., the Tesla model S) and predictions of what future states of shared control are possible (e.g., see [19,47]). However, the history of HCI has shown that interaction between humans and technology can change when disruptive technologies are introduced (e.g., GUIs, touch screens, smartphones). Similarly, currently unanticipated developments in the automotive domain may arise that might fundamentally change the interaction between drivers and cars. This might be particularly the case if fully automated cars without handover are developed (e.g., as in the Google vision). However, until that day we benefit from a basic understanding of human capacity (e.g., when do humans pay attention, how distracting is technology?).

CONCLUSION
Our results show that semi-autonomous cars benefit from pre-alerts that warn for a future handover situation. In particular, pre-alerts that reflect urgency, such as an increasing pulse signal, show high promise.

REFERENCES
1. Hakan Alm and Lena Nilsson. 1995. The effects of a mobile telephone task on driver behavior in a car following situation. *Accident Analysis and Prevention*, 27(5), 707-715. http://doi.org/ 10.1016/0001-4575(95)00026-V

2. Erik M. Altmann and J. Gregory Trafton. 2002. Memory for goals: An activation-based model. *Cognitive Science*, 26(1), 39–83. http://doi.org/10.1016/S0364-0213(01)00058-1

3. Brian P. Bailey and Shamsi T. Iqbal. 2008. 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. http://doi.org/10.1145/1314683.1314689

4. Brian P. Bailey and John 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), 685-708. http://doi.org/10.1016/j.chb.2005.12.009

5. David Beattie, Lynne Baillie, and Martin Halvey. 2015. A comparison of artificial driving sounds for automated vehicles. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '15). New York, NY: ACM, 451-462. http://doi.org/10.1145/2750858.2807519

6. Deborah A. Boehm-Davis and Roger W. Remington. 2009. Reducing the disruptive effects of interruption: a cognitive framework for analysing the costs and benefits of intervention strategies. *Accident Analysis & Prevention*, 41(5), 1124-1129. http://doi.org/10.1016/j.aap.2009.06.029

7. Jelmer P. Borst, Niels A. Taatgen, Hedderik van Rijn. 2010. The Problem State: A Cognitive Bottleneck in Multitasking. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 36(2), 363-382. http://doi.org/10.1037/a0018106

8. Jelmer P. Borst, Niels A. Taatgen, 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 ACM Conference on Human Factors in Computing Systems,* New York, NY: ACM, 2971-2980. http://doi.org/10.1145/2702123.2702156

9. Valdimar Briem and Leif R. Hedman. 1995. Behavioral effects of mobile telephone use during simulated driving. *Ergonomics*, 38, 2536-2562. http://doi.org/10.1080/00140139508925285

10. Karel Brookhuis, Gerbrand De Vries, Dick De Waard, 1991. The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, 23, 309-316. http://doi.org/10.1016/0001-4575(91)90008-S

11. Duncan P. Brumby, Andrew Howes, 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*, New York, NY: ACM, 233-242. http://doi.org/10.1145/1240624.1240664

12. Duncan P. Brumby, Dario D. Salvucci, Andrew Howes. 2009. Focus on Driving: How Cognitive Constraints Shape the Adaptation of Strategy when Dialing while Driving, In *Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY: ACM, 1629-1638. http://doi.org/10.1145/1518701.1518950

13. Jeff K Caird, Chelsea R Willness, Piers Steel, Chip T Scialfa. 2008. A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention*, 40(4), 1282-1293. http://doi.org/10.1016/j.aap.2008.01.009

14. Yujia Cao, Sandro Castronovo, Angela Mahr, Christian Müller. 2016. On timing and modality choice with local danger warnings for drivers, In *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Essen, Germany, 75-78. http://doi.org/10.1145/1620509.1620524

15. Geoff Cumming. 2013. *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

16. Joost de Winter, Riender Happee, Marieke Martens, Neville Stanton. 2014. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27, 196-217. http://doi.org/10.1016/j.trf.2014.06.016

17. Thomas A. Dingus, Feng Guo, Suzie Lee, Jonathan F. Antin, Miguel Perez, Mindy Buchanan-King, Jonathan Hankey. 2016. Driver crash risk factors and prevalence evaluation using naturalistic driving data. In *Proceedings of the National Academy of Sciences*, 2636-2641. http://doi.org/10.1073/pnas.1513271113

18. Johan Fagerlönn, Stefan Lindberg, and Anna Sirkka. 2012. Graded auditory warnings during in-vehicle use: using sound to guide drivers without additional noise. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (AutomotiveUI '12). New York, NY: ACM, 85-91. http://doi.org/10.1145/2390256.2390269

19. Tom M. Gasser and Daniel Westhoff. 2012. BASt-study: Definitions of automation and legal issues in Germany. Presented at the Presentation at the Road Vehicle Automation Workshop, Irvine, CA.

20. Christian Gold, Daniel Dambock, Lutz Lorenz and Klaus Bengler. 2013. "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), 1938-1942. http://doi.org/10.1177/1541931213571433

21. Ellen C. Haas and John G. Casali. 1993. The perceived urgency and detection time of multi-tone and frequency modulated warning signals. In *Proceedings of the human factors and ergonomics society 37th annual meeting,* Santa Monica, CA, USA, 544-548. http://doi.org/10.1080/00140139508925270

22. Peter A. Hancock. 2013. Driven to distraction and back again. In M. A. Regan, J. D. Lee, & T. W. Victor (Eds.), *Driver distraction and inattention advances in research and countermeasures*. Chichester, UK, 9-25.

23. William J. Horrey and Chris D. Wickens. 2006. Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors,* 48(1), 196-205. http://doi.org/10.1518/001872006776412135

24. Barbara Illowsky and Susan L. Dean. 2008. *Collaborative Statistics*, Connexions.

25. Shamsi T. Iqbal, Eric Horvitz, Yun-Cheng Ju, 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*, New York, NY: ACM, 463-472. http://doi.org/10.1145/1978942.1979008

26. Shamsi T. Iqbal, Yun-Cheng Ju, Eric Horvitz. 2010. Cars, calls, and cognition: investigating driving and divided attention In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY: ACM, 1281-1290. http://doi.org/10.1145/1753326.1753518

27. Christian P. Janssen and Duncan P. Brumby. 2010. Strategic Adaptation to Performance Objectives in a Dual-Task Setting. *Cognitive Science*, 34(8), 1548-1560. http://doi.org/10.1111/j.1551-6709.2010.01124.x

28. Christian P. Janssen, Duncan P. Brumby, 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), 5-29. http://doi.org/10.1177/1555343411432339

29. Christian P. Janssen, Shamsi T. Iqbal, Yuncheng 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), 270. http://doi.org/10.1037/xap0000020

30. Myounghoon Jeon, Thomas M Gable, Benjamin K Davison, Michael A Nees, Jeff Wilson and Bruce N. Walker. 2015. Menu navigation with in-vehicle technologies: Auditory menu cues improve dual task performance, preference, and workload. *International Journal of Human-Computer Interaction*, 31(1), 1-16. http://doi.org/10.1080/10447318.2014.925774

31. Sheila G. Klauer, Feng Guo, Bruce G. Simons-Morton, Marie Claude Ouimet, Suzanne E. Lee and Thomas A. Dingus. 2014. Distracted Driving and Risk of Road Crashes among Novice and Experienced Drivers. *New England Journal of Medicine*, 370(1), 54-59. http://doi.org/10.1056/NEJMsa1204142

32. Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer and Clifford Nass. 2014. Why did my car just do that? Explaining semiautonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing*, 1-7. http://doi.org/10.1007/s12008-014-0227-2

33. 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 Driver's attention to the Roadway. *Human Factors*, 43, 631-640. http://doi.org/10.1518/001872001775870340

34. Angela Mahr, Margarita Pentcheva, Christian Müller. 2009. Towards system-mediated car passenger communication, Proceedings of the *1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Essen, Germany 79-80. http://doi.org/10.1145/1620509.1620525

35. Daniel C. McFarlane. 2002. Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Human Computer Interaction*, 17(1), 63-139. http://doi.org/10.1207/S15327051HCI1701_2

36. Alexander Meschtscherjakov, David Wilfinger, Thomas Scherndl, 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*, Essen, Germany, 81-88. http://doi.org/10.1145/1620509.1620526

37. Bruce Mehler, Bryan Reimer and Joseph F Coughlin. 2012. Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups. *Human Factors*, 54(3), 396-412.

38. Todd A. Mondor and G.A. Finley. 2013. The perceived urgency of auditory warning alarms used in the hospital operating room is inappropriate. *Candian Journal of anaesthesia*, 50, 221-228. http://doi.org/10.1007/BF03017788

39. National Highway Traffic Safety Administration. 2013. *Preliminary statement of policy concerning automated vehicles,* Washington, DC.

40. Frederik Naujoks, Christoph Mai and Alexandra Neukum. 2014.The effect of urgency of take-over requests during highly automated driving under distraction conditions. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*, 431-438.

41. Ioannis Politis, Stephen Brewster, Frank Pollick. 2014. Evaluating multimodal driver displays under varying situational urgency.

In *Proceedings of the 32nd annual SIGCHI conference on Human Factors in Computing System*, New York, NY: ACM, 2067-4076. http://doi.org/10.1145/2556288.2556988

42. Ioannis Politis, Stephen Brewster, Frank Pollick. 2014. Speech tactons improve speech warnings for drivers. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, New York, NY: ACM, 1-8. http://doi.org/10.1145/2667317.2667318

43. Ioannis Politis, Stephen Brewster, Frank Pollick. 2013. Evaluating multimodal driver displays of varying urgency. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Aplications,* New York, NY: ACM, 92-99. http://doi.org/10.1145/2516540.2516543

44. Ioannis Politis, Stephen Brewster, Frank Pollick. 2015. Language-based multimodal displays for the handover of control in autonomous cars. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (AutomotiveUI '15)*, New York, NY: ACM, 3-10. http://doi.org/10.1145/2799250.2799262

45. Donald A. Redelmeier and Robert J. Tibshirani. 2009. Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine*, 336(7), 453-458. http://doi.org/10.1056/NEJM199702133360701

46. Nick Reed and R. Robbins. 2008. *The Effect of Text Messaging on Driver Behavior: A Simulator Study Transport Research Lab Report*, PPR 367.

47. SAE International. 2014. *J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems*. Retrieved from http://standards.sae.org/j3016_201401/.

48. Dario D. Salvucci. 2005. A Multitasking General Executive for Compound Continuous Tasks. *Cognitive Science*, 29(3), 457–492. http://doi.org/10.1207/s15516709cog0000_19

49. 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,* New York, NY: ACM, 85-88. http://doi.org/ 10.1145/1753326.1753340

50. Dario D. Salvucci and Niels A. Taatgen. 2011. *The Multitasking Mind*. New York, NY: Oxford University Press.

51. Dario D. Salvucci, Daniel Markley, Mark Zuber, Duncan P. Brumby, 2007. iPod distraction: effects of portable music-player use on driver performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY: ACM, 243-250. http://doi.org/10.1145/1240624.1240665

52. David L Strayer, Joel M Cooper, Jonna Turrill, James R Coleman and Rachel J Hopman. 2015. *Measuring Cognitive Distraction in the Automobile III: A Comparison of Ten 2015 In-Vehicle Information Systems,* 1-46.

53. Chris Urmson, (2015, May 11). The View from the Front Seat of the Google Self-Driving Car. *Medium.com*. Retrieved from https://medium.com/backchannel/the-view-from-the-front-seat-of-the-google-self-driving-car-46fc9f3e6088

54. Marcel Walch, Kristin Lange, Martin Baumann and Michael Weber. 2015. Autonomous driving: investigating the feasibility of car-driver handover assistance. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, New York, NY: ACM 11-18. http://doi.org/10.1145/2799250.2799268

55. Chris D. Wickens. 2008. Multiple Resources and Mental Workload. *Human Factors*, 50(3), 449-455. http://doi.org/10.1518/001872008X288394

56. Wolfe M. Jeremy, Todd S. Horowitz, Michael J. Van Wert, Naomi M. Kenner, Skyler S. Place, and Nour Kibbi. 2007. Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General* 136(4), 623. http://doi.org/10.1037/0096-3445.136.4.623