Concurrent multitasking

Nijboer, Menno

Publication date:
2016

CHAPTER 4

Driving and Multitasking: The Good, the Bad, and the Dangerous

Menno Nijboer
Jelmer P. Borst
Hedderik van Rijn
Niels A. Taatgen

Where we examine the positive and negative effects multitasking can have on driving performance.

Driving and Multitasking: The Good, the Bad, and the Dangerous

Abstract

Previous research has shown that multitasking can have a positive or a negative influence on driving performance. The aim of this study was to determine how the interaction between driving circumstances and the cognitive requirements of secondary tasks affect a driver's ability to control a car. We created a driving simulator paradigm where participants had to perform one of two scenarios: one with no traffic in the driver's lane, and one with substantial traffic in both lanes, some of which had to be overtaken. Four different secondary task conditions were combined with these driving scenarios. In both driving scenarios, using a tablet resulted in the worst, most dangerous, performance, while passively listening to the radio or answering questions for a radio quiz led to the best driving performance. Interestingly, driving as a single task consistently led to worse performance than driving in combination with one of the radio tasks. These results suggest that drivers switch to internally focused secondary tasks when nothing else is available during monotonous or repetitive driving environments. This mind wandering appears to have a stronger interference effect with driving than non-visual secondary tasks.

Introduction

A great deal of research has been done on the effects of engaging in multitasking while driving a car. Many types of tasks have been found to reduce driving performance (Ranney, Garrott, & Goodman, 2000), ranging from phone conversations (Strayer & Johnston, 2001; Treffner & Barrett, 2004) to music listening (Brodsky, 2001). However, recent evidence has indicated that multitasking could also be beneficial for driving when the right circumstances are met (Atchley & Chan, 2010; Gershon, Ronen, Oron-Gilad, & Shinar, 2009; Ünal, Steg, & Epstude, 2012). In this work we investigate the interaction between the cognitive demands of the driving task and the secondary task. We aim to answer the following questions: How does the interaction between driving circumstances and requirements of various secondary tasks affect driving behavior? And when is the effect on driving negative or positive? In particular, we used a simulated driving experiment to investigate how perceptual and working-memory requirements of both the driving task and the secondary task affect driving performance in reasonably realistic settings and durations.

Driving and Multitasking

The findings that driving performance decreases when combined with other tasks are consistent with psychological theories of multitasking. When two tasks require the same perceptual or cognitive resource, they are said to overlap with regards to that resource. Overlap in resource use between concurrently performed tasks leads to contention for those resources (Pashler, 1994; Salvucci & Taatgen, 2008; Wickens, 2002). In turn, this contention typically leads to reduced task performance (e.g., Borst, Taatgen, & Van Rijn, 2010; Just, Keller, & Cynkar, 2008; Nijboer, Borst, Van Rijn, & Taatgen, 2014; Strayer, Cooper, & Turrill, 2013). This does not only hold for central-cognitive resources such as working memory, but for perceptual and motor resources as well. Driving requires all of these resources to some degree (Anstey, Wood, Lord, & Walker, 2005; Herbert, 1963), so according to these theories driving should be negatively affected by most secondary tasks.

Tasks that have been found to interfere with driving often contain perceptual (visual and auditory; Chaparro, Wood, & Carberry, 2005; Gherri & Eimer, 2011) or motor (manipulation of equipment; Briem & Hedman, 1995; Brookhuis, de Vries, & de Waard, 1991; Janssen, Brumby, & Rae, 2012) factors that disrupt driving performance. However, cognitive requirements of secondary tasks turned out to be at least as important. Of all tasks found to interfere with driving, cell-phone use has received much attention due to the high number of traffic accidents attributed to such devices (Redelmeier & Tibshirani, 1997). In an influential study, Strayer and Johnston (2001) showed that it is primarily the attentional component of holding a conversation that disrupts driving performance. They concluded this by ruling out explanations related to holding the phone, speaking, or listening. Several studies have shown that holding a complex conversation in particular affects driving performance

(Briem & Hedman, 1995; McKnight & McKnight, 1993).

In contrast to the findings presented so far, some studies have shown that driving improves when concurrently performing another activity. Gershon et al. (2009) showed that a multiple-choice trivia game improved driving performance under monotonous driving circumstances. Atchley and Chan (2010) had similar results with a verbal word-association task in combination with a monotonous driving task. These findings raise an interesting question: what causes performance to improve when a secondary task is introduced?

Research in other areas has also shown that a secondary task can improve performance on the primary activity. For example, doodling on a piece of paper while performing a memory task has been found to improve recognition accuracy by improving overall concentration (Andrade, 2010; Singh & Kashyap, 2015). Andrade (2010) argues that doodling improves performance because it reduces the chance to engage in daydreaming, which is also referred to as mind wandering. When mind wandering the attention is shifted away from the task at hand and instead focuses on task-irrelevant thoughts. This behavior will typically occur when tasks have low processing demands, and are thus experiences as boring or repetitive (Forster & Lavie, 2009; Giambra, 1995). This internal focus results in a decoupling of perception and environment (Cheyne, Carriere, & Smilek, 2009; Smallwood & Schooler, 2015), which can have a negative impact on performance of the main task (He, Becic, Lee, & McCarley, 2011). Killingsworth & Gilbert (2010) estimated that up to 50% of everyday life is spent mind wandering. This is in line with research that has shown that people actively seek out opportunities to multitask (Czerwinski, Horvitz, & Wilhite, 2004; González & Mark, 2004; Gould, Brumby, & Cox, 2013). Thus, a boring drive might lead to mind wandering, which has been shown to have a negative impact on visual attention during driving (He et al., 2011).

Therefore, adding a secondary task during a monotonous driving setting could have the same effect as doodling has on a boring memory task: it reduces the chance of other, more interfering tasks – such as mind wandering – to intrude the primary activity. This would imply that a secondary task does not make driving performance itself better, but is the lesser of two evils: The interference of mind wandering has a more substantial effect on the driving task than the secondary task does. Alternatively, complete focus on the driving task might lead to drivers over-regulating their driving behavior: according to the execution-focus theory, increased attentional control to highly proceduralized sensory-motor skills can disrupt execution of these skills (Baumeister, 1984; Beilock & Carr, 2001). This is typically observed in sports, but might also lead to a decrease in driving skill under circumstances where the cost of failure is significant. We will further discuss those possible explanations after we report our results.

Driving and Working Memory

Some of the additional activities that are performed during driving, such as having a (phone) conversation, will require the temporary retention of some information conveyed by the conversational partner (e.g., Van Rij, Van Rijn, & Hendriks, 2010, 2013). However, the direct effect of working-memory load during driving has received little interest in the literature. Radeborg, Briem, and Hedman (1999) found that driving had a negative effect on a concurrently performed word-recall task. Changes in driving performance itself were not reported. Gugerty (1997) concentrated on the working-memory aspect of driving by having participants remember the locations of all the cars they encountered during a driving segment. Memory load was manipulated by varying the traffic density. While performance on the memory task correlated with traffic density, no significant effects on driving performance were found. However, performance on the driving task was at ceiling, so effects may have been obscured. Alm and Nilsson (1995) asked participants to perform a word-recall task over the phone while driving, and found that this significantly decreased the distance kept to the car ahead of the participant. Other measurements such as lane deviation were not reported. In general these results are in line with other research, which has shown that multitasking situations where both tasks have a significant working-memory load lead to substantially decreased performance in one or both tasks (e.g., Borst, Taatgen, Stocco, & Van Rijn, 2010; Just et al., 2008; Nijboer et al., 2014; Nijboer, Taatgen, Brands, Borst, & Van Rijn, 2013; Strayer et al., 2013).

In determining the effect of working-memory load on driving, not only the working-memory load of the secondary task is relevant: the working-memory requirements of the driving task are very important as well. If driving does not require working memory, it will not be disrupted by a secondary task that needs working memory. The working-memory load of driving is strongly dependent on the traffic: when the road is empty the driver only has to remember information regarding the current state of the car, which can be easily retrieved from visual and aural queues that are constantly present in the environment. When there is substantial traffic, however, the driver has to keep a detailed mental model of the surrounding vehicles (Gugerty, 1997), as these will not always be visible: they might reside in the blind spot of the car, or be obscured by other vehicles. In the current study we therefore used two different driving scenarios: one with high working-memory requirements, and one with low working-memory requirements.

Current Study

We investigated the contrast between beneficial and detrimental multitasking during driving by exploring a diverse set of combinations of secondary tasks and driving scenarios. We varied the working-memory, perceptual, and motor load of both the secondary task and the driving task to determine under what conditions a particular type of secondary task has a positive or negative influence on driving performance. In particular, we performed a driving experiment where we tested two groups of participants in one of two driving conditions that differed in traffic density and associated working-memory load. To test the effects of secondary tasks on these driving conditions, the drivers in each group performed four different secondary-task conditions: no secondary task, passive radio listening, a radio quiz, and a tablet-based quiz. This resulted in a detailed evaluation of the effect of secondary tasks with different cognitive requirements on driving performance under contrasting traffic circumstances.

In the remainder of this paper we will first explain the driving paradigm in detail, followed by the results of the experiment. We finish with a discussion of how the results yield new insight into the interaction between driving conditions and concurrent execution of a secondary task.

Method

Paradigm

We created a paradigm that tested the effect of four different secondary activities on driving performance during two different driving scenarios, referred to as the No-Traffic and Traffic scenarios. Driving scenario was a between-subject variable, while the four different secondary tasks are within-subject variables. For each condition we record a number of measurements during driving (lane keeping, speed, secondary-task performance, and steering). Both scenarios used a two-lane highway in a desert environment. The road contained two five-meter-wide lanes (cf. the minimum highway lane width in the US is 3.7m[1], while the standard highway lane width in the Netherlands is 3.5m) and had a subtle curvature, approximately 3.5 cm of lateral displacement per meter of road, to ensure that minor changes to the car heading needed to be made on a regular basis. The car that was driven was a 1966 Ford Mustang, with a width of approximately 1.7 meters and a length of 4.6 meters. In both scenarios participants were instructed to drive 80 km/h (50 mph), and to not exceed this speed.

The No-Traffic scenario was constructed to test the effects of secondary tasks with different cognitive loads during situations where the driving itself was easy. In the No-Traffic scenario the highway had no traffic in the right lane: participants were occasionally overtaken by other cars, but did not have to overtake any cars themselves. This is illustrated in Figure 4-1A. With no relevant traffic to keep track of, the working-memory load of the No-Traffic scenario was low.

The Traffic scenario was designed to test how multitasking affects typical highway driving when the road situation has to be monitored constantly. This was achieved by introducing traffic in both driving lanes. Participants were often overtaken by other cars, and also had to overtake slower cars in the right lane as shown in Figure 4-1B. The slow right lane traffic was distributed such that at 80 km/h participants would overtake approximately 60 cars during a 30-minute block, or 2 cars per minute. The

[1] AASHTO Green: A Policy on Geometric Design of Highways and Streets (2001). Fourth edition, American Association of State Highway and Transportation Officials

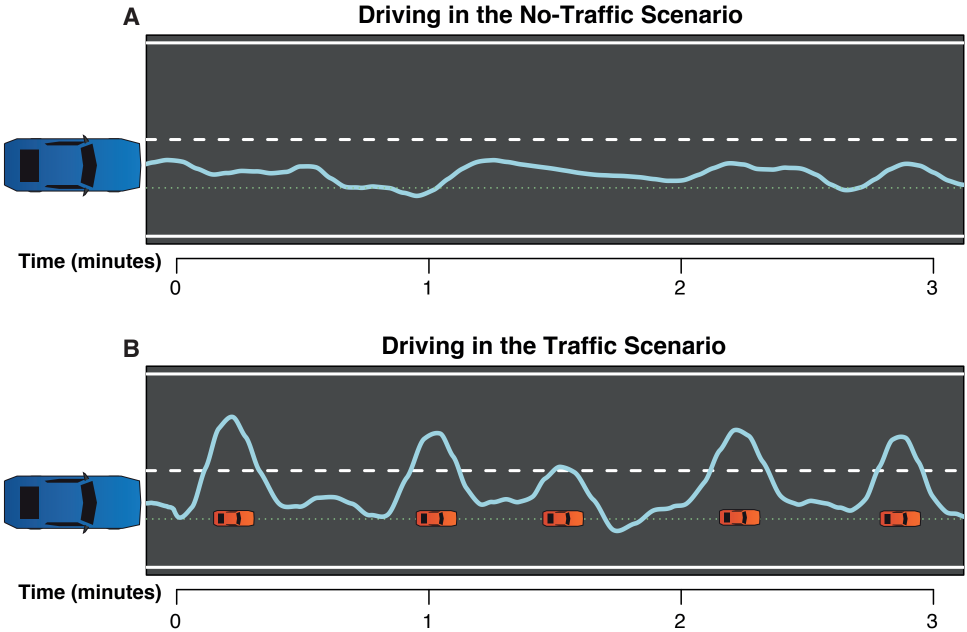

Figure 4-1. Two examples of typical driving paths for the two different scenarios. The blue line is the path taken by the participant. The green dotted line is the center of the right lane, and the white dashed line is the division between the two lanes. Red cars represent slower left-lane traffic that the participant needs to overtake.

distance between right-lane cars varied between 60 and 95 meters. The left-lane traffic would overtake the participant at set points, distributed over the span between the previously overtaken right-lane car and the next right-lane car: at approximately 25, 40, 55, and 75% of the total inter-car distance these cars could appear, with a 50% probability for each possibility – so on average participants would be overtaken by two of these cars before having to overtake again themselves.

While most cars in the left lane would overtake the participant at random moments, there were two special types of cars. The first type overtook participants at the time they needed to overtake a slow car in the right lane themselves, forcing them to wait until the faster car had passed. This encouraged participants to keep a mental model of the traffic around them. The second special type of car would drive behind the participant in the left lane at a reasonable distance until the participant overtook a slower car in the right lane. The car would then change lanes to stick behind the slow car. These cars were added to reduce the predictability of left-lane traffic: if all left-lane traffic overtakes the participant, then it would always be optimal to just wait in the right lane until no more traffic can be seen in the rear-view mirror. However, if some cars never pass, this requires a more active role of the participant in anticipating the best time to overtake the leading car. The two special types of cars (who also have a 50% probability to appear, like all other left-lane traffic), together with the random left-lane traffic and randomly spaced slower right-lane cars, created a dynamic highway situation.

To determine how the working-memory, perceptual, and motor requirements of the secondary task affect driving performance we created four different secondary task conditions. (1) No Secondary Task Condition (Single): In the Single condition there is no secondary task. (2) Listening Condition: A radio talk show would play during the entire block. Because participants were informed that no information presented in the talk show would need to be recalled later, we assumed that this condition had a low cognitive load. The perceptual load was low as well, given that the show was presented aurally. (3) Radio-Quiz Condition: In the Radio-Quiz condition, fragments of a radio talk show, similar to shows in the previous condition, were played split into multiple fragments. A multiple-choice question followed each audio fragment, and participants had to choose between three answers using buttons on the steering wheel. In this condition we assume that the working-memory load is higher than in the Listening condition, as the information presented in the radio show needs to be retained in memory to answer the questions. The motor load is also slightly higher as a button needed to be pressed to respond to the questions. (4) Tablet-Quiz Condition: A variation of the Radio Quiz where all information, both the text of the talk show and the questions, was presented on a tablet in the lower-left corner of the screen instead of aurally. The working-memory and motor loads are expected to be similar to the Radio Quiz, but the perceptual load is much higher as participants has to shift their gaze from the road to the tablet.

To motivate participants to perform well, they could increase their financial reward by collecting bonus points. Each bonus point was worth 10 cents, and the starting bonus was 40 points. The maximum bonus was 100 points. Points could be earned during the radio and tablet quiz. Each correct answer was worth 1 point. However, participants could lose points by either driving off the road (-1 point per second off-road), hitting other cars (-2 points per hit), or not signaling properly when changing lanes (-1.5 points per offense).

Participants

We recruited 48 native Dutch speakers that were divided over two experimental groups of 24 participants each. The first group drove in the No-Traffic scenario (16 female, M_{age} = 24.6, age range: 20-36), while the second group drove in the Traffic scenario (13 female, M_{age} = 23.6, age range: 20-32). Both groups contained experienced drivers (No-Traffic: $M_{license}$ = 5.0 years, M_{driven} = ~65000 km. Traffic: $M_{license}$ = 5.4 years, M_{driven} = ~60000 km). The two groups can be considered comparable, as the Bayes Factors[2] of the difference between groups for license years and kilometers driven were 0.31 and 0.29 respectively. Participants received a minimum of €20 upon completion, and could earn up to an additional €10 depending on task performance. The average

[2] According to Jeffreys (1961, p. 432), a Bayes Factor <1/3 qualifies as substantial evidence against the alternative hypothesis (in this case that the groups are different).

Figure 4-2. The simulated driving environment. *Left:* The environment during the Single, Passive Listening, and Radio-Quiz conditions. *Right:* The environment during the Tablet-Quiz condition.

received bonus was €6,40. All participants had normal or corrected-to-normal vision.

Materials and Procedure

Apparatus. The driving scenarios were built and executed in a driving simulator designed and programmed specifically for our paradigm. A steering wheel (Logitech driving force GT) with feet pedals was used to control the car, which had an automatic transmission. The center of the steering wheel contained three buttons on the right side that were used to answer the quiz questions. On the back of the wheel two buttons (one on each side) could be used to activate the left or right turn signal. Participants wore headphones for the auditory stimuli. The simulation was viewed on a 23-inch LCD display at 120 Hz, at a distance of approximately 70 cm from the participant. The simulation environment can be seen in Figure 4-2. Visible are the hood of the car, the windscreen wipers, a speedometer, turn signal indicators, a rear-view mirror, and the current bonus score. The hood of the car is shown to better judge the road position while driving and to give a sense of size to the information presented in the outside world. Continuous data from the simulation was recorded at 50 Hz. We recorded the car position, pedal pressure, wheel angle, speed, direction indicators, and contact with other cars.

Listening stimuli. For the passive radio-listening condition we selected 30-minute segments of two episodes of a popular-science public-radio talk show. The topics of the shows were addiction and music perception. The two shows were balanced across participants within each scenario, with half of the scenario group listening to the first show, and the other half listening to the second.

Radio-Quiz stimuli. For the quiz we used two episodes of the same science talk show, but that were different from the ones used in the Listening condition. This time the topics were depression and improving mental health. We generated 30 questions for each talk show, with 3 possible answers per question. The Apple OSX text-to-speech function was used to create the audio files. As all stimuli were in Dutch we used the 'Xander' voice to ensure intelligibility. Like the listen-only radio shows, the

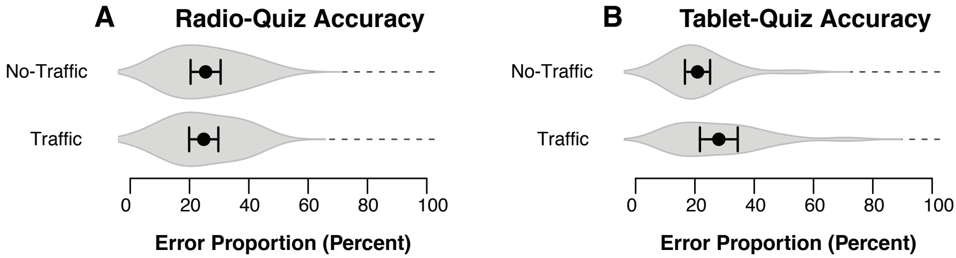

Figure 4-3. The performance on the quiz tasks. Black dots represent the mean across subjects, and bars denote 95% CI. *Panel A:* Performance on the Radio Quiz in both driving scenarios. *Panel B:* Performance on the Tablet Quiz in both driving scenarios.

presented show was balanced across participants within each scenario group.

Tablet-Quiz stimuli. The stimuli for the Tablet Quiz were transcripts of the talk show fragments and questions used in the radio quiz. The rate of sentence presentation was matched to the length of the original audio fragments, and each presentation of a new sentence was accompanied with a tone sound. The display accommodated a maximum of 10 lines at a time, which covered around 30% of the width of the entire screen. A sentence was on screen for at least 10 seconds. The radio and tablet quizzes were paired in such a way that participants were not presented the same topic twice.

Procedure. The experiment lasted slightly under 2.5 hours. Participants started with a 5-minute training session to familiarize themselves with the driving task and how to overtake other cars. To become accustomed to handling a secondary task while driving, participants performed a second 5-minute driving session during which they also carried out the tablet task. The actual experiment consisted of four 30-minute blocks, resulting in a drive length of 120 minutes in total. Therefore each block was slightly longer than the average commute time in the United States (25 minutes; McKenzie & Rapino, 2011). Each block corresponded to one of the four secondary task conditions, which was performed in the driving scenario that the participant was assigned to. The order of conditions was counter-balanced across participants using a Latin square to avoid order effects. After each block there was an opportunity for participants to take as long a break as they required before continuing on to the next block.

Results

Unless mentioned otherwise, all *p*-values of the main effects are from analyses of variance performed on linear mixed-effects models (LME). Accuracy data were modeled using binomial LMEs. The *p*-values of individual comparisons between conditions were computed by performing a Tukey honest significant difference test on each LME. All models were constructed and analyzed in R (3.0.2) with the lme4 package (1.0-5). All error bars in figures depict the upper half of 95% confidence

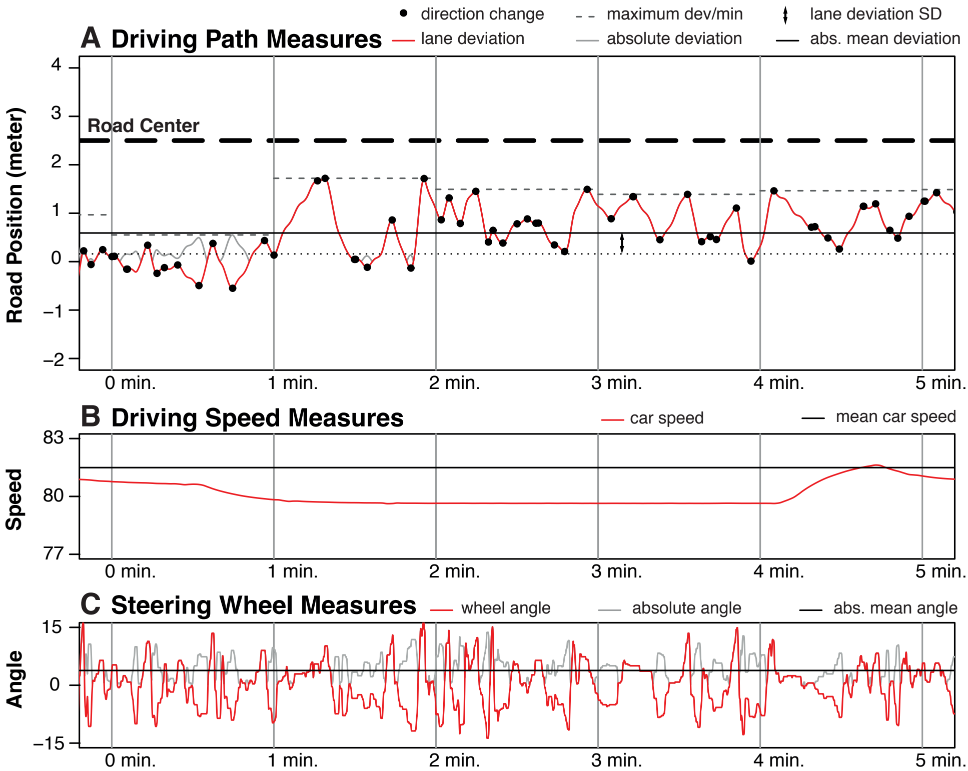

Figure 4-4. Measurements taken from a random participant in the No-Traffic condition. Red lines are the raw data, light-grey lines are the absolute values of that data. The segment between two vertical grey lines denotes one minute of driving. *Panel A:* Measurements taken from the position of the car. The red line is the deviation from the center of the lane, while the light grey line is the absolute deviation from the lane. The black line with thick dashes is the center of the road, consisting of two lanes. The center of the right lane is denoted by 0 on the y-axis. The solid-black line is the average car position. The distance between the solid-black line and the dotted-black line is the SD of the car position. The black circles located on the dark-grey line are direction changes, where the heading of the car shifted from left to right, or vice versa. The dashed-grey line segments are the observed maximum absolute deviation from the center of the lane for each minute of driving time. *Panel B:* Measurements taken from driving speed. The red line shows the driving speed, while the solid-black line is the average observed speed over the block. *Panel C:* Measurements taken from the steering wheel position. The red line is the steering angle plotted over time, while the light-grey line is the absolute steering angle. The solid-black line is the average absolute position of the wheel over the block.

intervals for the mean, corrected for a within-subject design. Gray volumes behind the means are (the smoothed estimates of) the underlying distribution of the data (Sheather & Jones, 1991). Bars along the side of the figures indicate a significant difference between the two indicated conditions. Stars without any bars indicate that the condition was significantly different from all other conditions. The statistics are reported in Tables 4-1 through 4-4.

Overall, the Tablet-Quiz condition led to the worst driving performance, followed by the Single (no secondary task) condition. The Listening and Radio-Quiz conditions resulted in the best driving performance. This pattern appears in most of the variables that were measured, over both driving scenarios, and therefore appears robust. We will now discuss the separate measures in more detail.

Secondary Tasks

Before examining the driving itself, we will look at the performance on the two quiz tasks. Figure 4-3A shows that the error proportion of the Radio Quiz was low for both driving scenarios, indicating that participants did perform the secondary task while driving. The performance on the Tablet Quiz in Figure 4-3B has a similar range to that of the Radio Quiz. The seems to indicate that performance was better for the combination of a No-Traffic road with the Tablet Quiz when compared to the Radio-Quiz performance. However, this difference did not reach significance.

Lane Deviation and Swerving

The distance to the center of the driving lane, or lane deviation, is a standard measure to investigate driving performance: a large standard deviation indicates a large degree of "swerving" across the road (see Figure 4-4A). The average and standard deviation of the car position were plotted for both driving scenarios in Figure 4-5. The values of each participant were demeaned using the grand mean of the participant over all conditions, in order to remove any inherent bias of a participant for a specific position in the lane. For the Traffic scenario all driving segments where participants were overtaking other cars were discarded: These were defined as all data points ranging from 3 seconds before signaling a lane change to the left lane (using the blinkers) until 3 seconds after the center of the car crosses the center of the road from the left lane to the right lane, after overtaking a car. The differences between conditions are similar for both driving scenarios: The worst lane-keeping performance occurred when participants had to perform the Tablet Quiz while driving: the degree of swerving was larger, and the average distance to the ideal lane position was also larger. The best performance was obtained when participants were in either the Listening or the Radio-Quiz condition. Consequently, the Single condition results were ranked in the middle of all conditions. However, only the difference between the Tablet-Quiz condition and all other conditions was significant.

Lane deviation by itself is a limited means of evaluating the consistency and safety of a driver's lane-keeping behavior, as it reduces all the complexities of driving into a single value. In order to study lane keeping in more detail we devised two variables that characterize lane-keeping behavior: the number of changes in car heading, and the maximum observed distance to the ideal lane position (per minute; see Figure 4-4A). Essentially this divides swerving as calculated by the standard deviation into

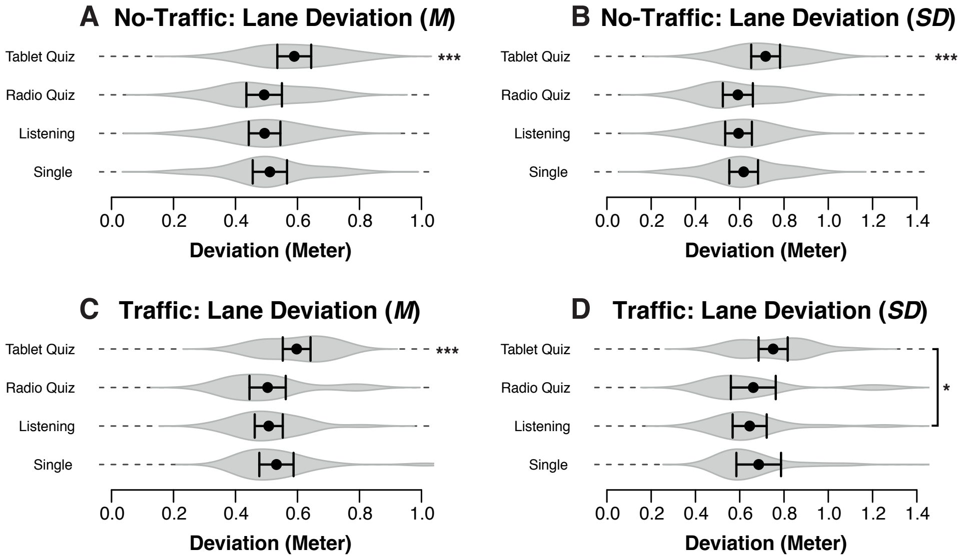

Figure 4-5. Lane deviation during the No-Traffic driving scenario and the non-overtaking sections of the Traffic driving scenario. Black dots represent the mean across subjects, and bars denote 95% CI. All data has been demeaned for each participant using the grand mean over all conditions. *Panel A and C:* The mean deviation from the center of the right lane. *Panel B and D:* The standard deviation of the car position.

Comparisons	Lane Deviation (M)			Lane Deviation (SD)		
No-Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	1.21	.017	.611	1.12	.023	.668
Single vs. Radio-Quiz	1.08	.018	.694	1.35	.026	.519
Radio-Quiz vs. Listening	-.105	-.001	.999	-.205	-.003	.997
Tablet-Quiz vs. Single	**3.64**	**.079**	**< .01**	**4.07**	**.099**	**< .001**
Tablet-Quiz vs. Listening	**4.93**	**.096**	**< .001**	**4.78**	**.122**	**< .001**
Tablet-Quiz vs. Radio-Quiz	**4.78**	**.097**	**< .001**	**5.27**	**.125**	**< .001**
Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	1.50	.025	.133	1.60	.041	.111
Single vs. Radio-Quiz	1.40	.028	.161	.667	.024	.505
Radio-Quiz vs. Listening	-.240	-.004	.810	.604	.016	.546
Tablet-Quiz vs. Single	**2.88**	**.065**	**< .01**	1.41	.065	.158
Tablet-Quiz vs. Listening	**4.55**	**.090**	**< .001**	**2.92**	**.106**	**< .01**
Tablet-Quiz vs. Radio-Quiz	**4.16**	**.094**	**< .001**	**2.06**	**.090**	**.040**

Table 4-1. Between-conditions comparisons of measurements related to lane deviation in both driving scenarios. Comparisons were computed by applying a Tukey honest significant difference on the linear mixed-effects models. The resulting z values, p values, and estimates (β) are reported. Bold numbers indicate significance at the 0.05 level.

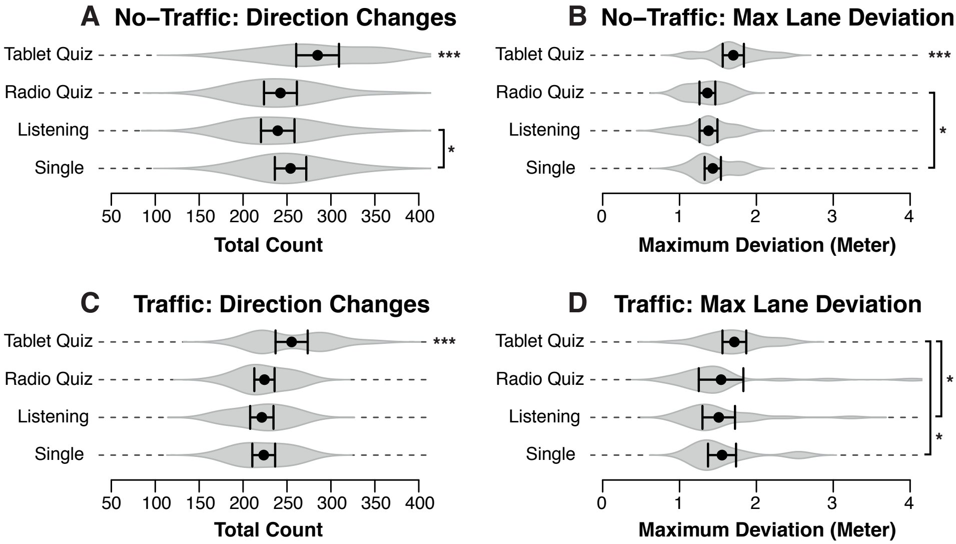

Figure 4-6. Performance variables related to lane-keeping consistency. Data for the No-Traffic group were recorded over the entire block, while data of the Traffic group were filtered to remove sections where participants were overtaking other cars. Black dots represent the mean across subjects, and bars denote 95% CI. *Panels A and C:* The number of heading changes made during a block. *Panels B and D:* The mean of the maximum absolute deviation from the center of the lane, computed for every minute of driving.

Comparisons	Direction Changes			Max Lane Deviation		
No-Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	2.01	14.6	.045	1.19	.055	.232
Single vs. Radio-Quiz	1.57	11.5	.116	**2.19**	**.070**	**.029**
Radio-Quiz vs. Listening	.467	3.08	.641	-0.38	-.015	.706
Tablet-Quiz vs. Single	**1.57**	**11.5**	**.012**	**3.68**	**.267**	**< .001**
Tablet-Quiz vs. Listening	**4.55**	**45.4**	**< .001**	**4.58**	**.322**	**< .001**
Tablet-Quiz vs. Radio-Quiz	**4.24**	**42.3**	**< .001**	**5.54**	**.337**	**< .001**
Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	.343	2.17	.732	.599	.043	.549
Single vs. Radio-Quiz	-.136	-.833	.892	.147	.012	.883
Radio-Quiz vs. Listening	.384	3.00	.701	.307	.031	.759
Tablet-Quiz vs. Single	**31.8**	**5.49**	**< .001**	**2.05**	**.160**	**.040**
Tablet-Quiz vs. Listening	**33.9**	**3.73**	**< .001**	**2.52**	**.203**	**.012**
Tablet-Quiz vs. Radio-Quiz	**30.9**	**4.12**	**< .001**	1.46	.173	.145

Table 4-2. Between-conditions comparisons of measurements related to lane keeping in both driving scenarios. Comparisons were computed by applying a Tukey honest significant difference on the linear mixed-effects models. The resulting *z* values, *p* values, and estimates (β) are reported. Bold numbers indicate significance at the 0.05 level.

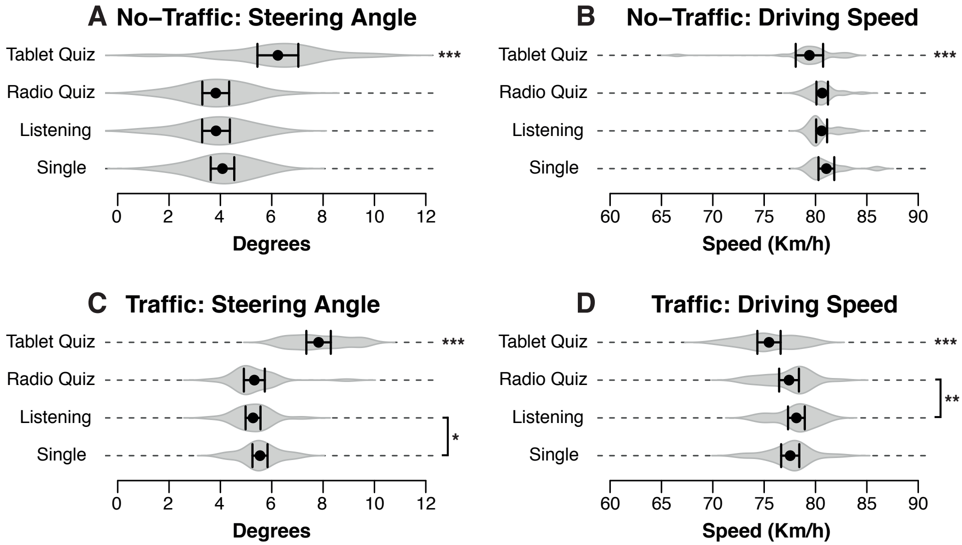

Figure 4-7. Driving performance measured for steering and speed variables. Data for the No-Traffic group were recorded over the entire block, while data of the Traffic group were filtered to remove sections where participants were overtaking other cars. Black dots represent the mean across subjects, and bars denote 95% CI. *Panels A and C:* The mean absolute angle of the steering wheel position. *Panels B and D:* The mean driving speed.

two separate measures that quantify driving consistency and safety. The directional changes (Figure 4-6A and 6C) and maximum deviation (Figure 4-6B and 6D) are consistent with the u-shaped lane deviation results of Figure 4-5: in both graphs the Tablet Quiz clearly leads to the worst lane-keeping performance, while the Radio-Quiz and Listening conditions result in the best performance. This pattern is most pronounced in the No-Traffic condition. The number of direction changes is higher across conditions in the No-Traffic scenario. This is because there is less data available for the Traffic scenario as the overtake sections have been taken out. Again, the Tablet-Quiz condition is significantly different from all but one of the other conditions. In terms of direction changes the Single condition performs significantly worse than Listening condition in the No-Traffic scenario. Examining the maximum lane deviation the Single condition performs significantly worse than the Radio-Quiz condition in the No-Traffic scenario.

Steering and Speed

Additionally, driving performance was measured using the steering-wheel and car-speed data (recorded as shown in Figure 4-4B and 4C). Figure 4-7A and 7C show that these data are consistent with the lane-keeping data: The significantly larger angle in the Tablet-Quiz condition indicates that participants made sharper steering corrections compared to the other conditions. While steering is related to lane deviation, the steering angle does give different information: the deviation shows the magnitude of

Comparisons	Steering Angle			Driving Speed		
No-Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	1.41	.252	.158	1.78	.463	.075
Single vs. Radio-Quiz	1.78	.263	.076	1.58	.407	.113
Radio-Quiz vs. Listening	-.069	-.011	.945	.325	.055	.745
Tablet-Quiz vs. Single	**6.62**	**2.16**	**< .001**	**-2.55**	**-1.65**	**.011**
Tablet-Quiz vs. Listening	**7.63**	**2.41**	**< .001**	**-2.07**	**-1.19**	**.038**
Tablet-Quiz vs. Radio-Quiz	**9.70**	**2.42**	**< .001**	**-2.17**	**-1.25**	**.030**
Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	**2.37**	**.271**	**.018**	-1.42	-.601	.157
Single vs. Radio-Quiz	1.48	.221	.139	.314	.116	.753
Radio-Quiz vs. Listening	.408	.050	.683	-2.83	**-.716**	**< .01**
Tablet-Quiz vs. Single	**15.5**	**2.28**	**< .001**	**-4.48**	**-2.07**	**< .001**
Tablet-Quiz vs. Listening	**15.0**	**2.55**	**< .001**	**-7.36**	**-2.67**	**< .001**
Tablet-Quiz vs. Radio-Quiz	**13.7**	**2.50**	**< .001**	**-5.83**	**-1.95**	**< .001**

Table 4-3. Between-conditions comparisons of measurements related to driving performance in both driving scenarios. Comparisons were computed by applying a Tukey honest significant difference on the linear mixed-effects models. The resulting z values, p values, and estimates (β) are reported. Bold numbers indicate significance at the 0.05 level.

swerving across the lane, while the steering angle gives more information regarding how fast corrections were made. Again, the Radio-Quiz and Listening conditions resulted in the best performance, while in the Traffic scenario the Single condition was significantly worse than the Listening condition, but better than the Tablet-Quiz. Across all conditions the steering angle was higher in the Traffic scenario when compared to the No-Traffic scenario. Finally, the average speed shown in Figure 4-7B and 7D shows that while participants were able to keep to the instructed speed quite well, the Tablet-Quiz condition consistently led to the slowest driving speed, and was significantly different from all other conditions in both scenarios. In addition, the Radio-Quiz and Listening conditions differed significantly in the Traffic scenario.

The most immediate difference between the No-Traffic and Traffic scenarios is seen in the average speed patterns of Figure 4-7: while the Tablet Quiz condition still deviates the most from the 80 km/h target speed, none of the conditions result in average speeds that exceeded the target. The Listening condition came closest to the target speed. The patterns of the directional changes and maximum deviations are less pronounced when compared to the data of the No-Traffic scenario. This is likely because less data were available for the Traffic scenario, as all data where participants overtook other cars were removed for this analysis. A difference from the No-Traffic data is that here the number of directional changes indicate that the Single condition did not lead to more frequent swerving compared to the Listening and Radio-Quiz conditions.

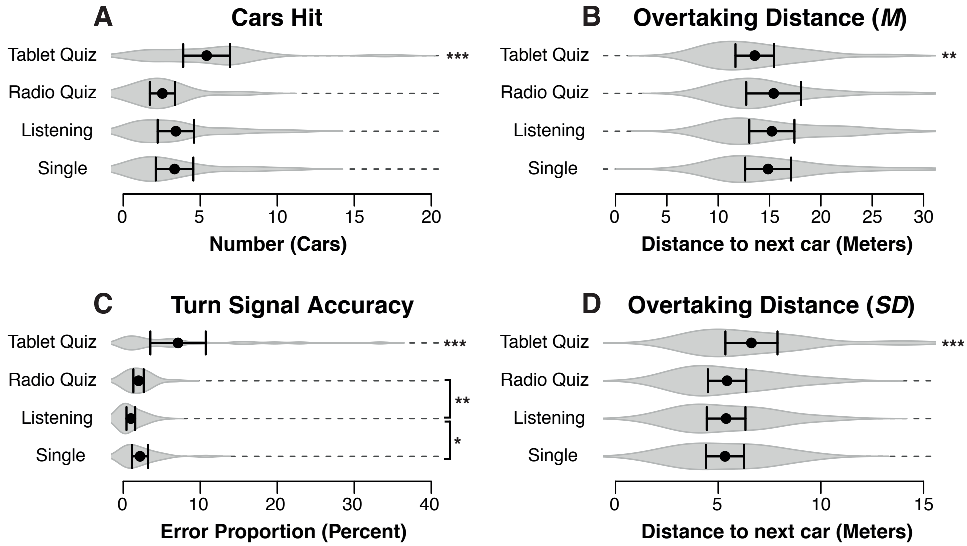

Figure 4-8. Measurements of overtaking performance. Black dots represent the mean across subjects, and bars denote 95% CI. *Panel A:* The number of cars the participant connected with. *Panel B:* The mean distance to the leading car before the participant initiated an overtake maneuver. *Panel C:* The accuracy of the turn-signal use when changing lanes. *Panel D:* The standard deviation of the distance to the leading car when initiating an overtake maneuver.

Overtaking

To evaluate how the overtaking of other cars was affected by secondary tasks we considered three variables. The first is the number of cars a participant hit, collapsing over hits to left-lane and right-lane traffic. The second variable is accurate turn signal use. Accurate use was defined as using the left turn signal when moving to the left lane and the right turn signal when moving to the right lane. Any other combination, or not using the turn signal at all, was registered as an error. Finally, the overtake-distance was the distance between the participant's car and the leading car at the moment the participant's car crosses the center of the road to overtake that car by switching to the left lane.

In accordance with the lane-keeping measurements, all three variables in Figure 4-8 present the Tablet Quiz as the secondary task that resulted in the worst overtaking performance. The Radio Quiz led to the least number of cars hit (Figure 4-8A), while the Listening and Single conditions performed similarly to each other. Approximately 89% of the cars that were hit were faster cars in the left lane, while the remaining 11% were slower cars in the right lane: left-lane cars are the cars that will overtake the participant, and hitting them indicates that either the participant did not see that car, or misjudged when that car would overtake the participant. The right-lane cars are the cars the participant had to overtake. Hitting them indicates that the participant

Comparisons	Cars Hit			Overtake Distance (M)		
No-Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	-.134	-.083	.893	.458	.365	.647
Single vs. Radio-Quiz	1.34	.792	.173	.759	.542	.448
Radio-Quiz vs. Listening	-1.43	-.875	.153	-.205	-.177	.838
Tablet-Quiz vs. Single	**2.20**	**2.08**	**.028**	**2.04**	**1.30**	**.042**
Tablet-Quiz vs. Listening	**2.22**	**2.00**	**.027**	**2.60**	**1.66**	**< .01**
Tablet-Quiz vs. Radio-Quiz	**3.70**	**2.88**	**< .001**	**2.61**	**1.84**	**< .01**
	Turn Signal			Overtake Distance (SD)		
Traffic Scenario	*z*	*β*	*p*	*z*	*β*	*p*
Single vs. Listening	**-2.34**	**-.924**	**.019**	-.144	-.054	.886
Single vs. Radio-Quiz	.529	.145	.596	-.262	-.101	.793
Radio-Quiz vs. Listening	**-3.20**	**-1.07**	**< .01**	.148	.047	.882
Tablet-Quiz vs. Single	**-2.79**	**-.889**	**< .01**	**2.20**	**1.29**	**.028**
Tablet-Quiz vs. Listening	**-4.60**	**-1.81**	**< .001**	**3.39**	**1.23**	**< .001**
Tablet-Quiz vs. Radio-Quiz	**-2.69**	**-.743**	**< .01**	**2.50**	**1.19**	**. 012**

Table 4-4. Between-conditions comparisons of measurements related to overtaking actions. Comparisons were computed by applying a Tukey honest significant difference on the linear mixed-effects models. The resulting z values, p values, and estimates (β) are reported. Bold numbers indicate significance at the 0.05 level.

did not steer accurately (while overtaking), or misjudged the speed at which the other car was moving, as this speed varied over time. In terms of turn-signal use (Figure 4-8C) the Listening condition outperformed all others, with the Radio-Quiz and Single conditions sitting in the middle. Figure 4-8B and Figure 4-8D present the differences in overtake distance. The only condition that stands out is the Tablet Quiz: performance of the other conditions was similar for both the average and the standard deviation.

The act of overtaking might reduce secondary-task performance if the driving task is prioritized. For both the Radio-Quiz and Tablet-Quiz tasks we investigate how likely drivers would answer questions during overtaking when that question was prompted during the first half of an overtake maneuver. If the secondary tasks were not affected by overtaking, we would expect to find a ratio around 1:1 between responding during overtaking and responding afterwards. However, we found that drivers were twice as likely to respond to questions after the overtake maneuver was completed for both the Radio Quiz (65%; $p < .001$ given P(H) = 50%) and Tablet Quiz (68%; $p < .001$ given P(H) = 50%). When examining the accuracy on questions answered during overtaking we found a reduction in accuracy for the Tablet Quiz. However, neither the difference in the Tablet Quiz (74% vs. 67%; $\beta = -.640$, $z = -1.91$, $p = .057$) nor the difference for the Radio Quiz (74% vs. 78%; $\beta = .120$, $z = .439$, $p = .660$) reached significance.

Measurement	Single	Listening	Radio Quiz	Tablet Quiz
No-Traffic				
Lane Deviation (*M*)	-	+	++	--
Lane Deviation (*SD*)	-	+	++	--
Direction Changes	-	++	+	--
Max Lane Deviation	-	+	++	--
Steering Wheel Angle	-	++	+	--
Driving Speed	--	++	+	-
Average	-	+/++	+/++	--
Traffic				
Lane Deviation (*M*)	-	+	++	--
Lane Deviation (*SD*)	-	++	+	--
Direction Changes	+	++	-	--
Max Lane Deviation	-	++	+	--
Steering Wheel Angle	-	++	+	--
Driving Speed	+	++	-	--
Average	-	++	+	--
Car Hits	+	-	++	--
Overtake Distance (*M*)	-	+	++	--
Overtake Distance (*SD*)	++	+	-	--
Turn Signal Accuracy	-	++	+	--
Average	-	+/++	+/++	--

Table 4-5. A ranking of all measurements made for the No-Traffic and Traffic scenarios, as well as the measurements related to overtaking. The ranking of best performance to worst performance is ++, +, -, and finally --.

Summary of Empirical Results

When taken together, the measurements we computed using the driving data present a comprehensive evaluation of driving consistency and safety under varying different loads for working-memory, perception, and motor actions. The influence of secondary tasks on driving performance was compared by ranking the performance of each secondary task, per variable, for each of the scenarios as presented in Table 4-5: ranks were assigned from worst performance (--) to best performance (++) according to the averages presented in earlier plots. For a comprehensive analysis of the ranked data we refer to the Appendix.

The Tablet Quiz is the worst scoring secondary task, ranking lowest in almost all variables across both scenarios. The remaining three conditions see some variation across scenarios: In the No-Traffic scenario, the Single condition resulted in the lowest performance after the Tablet-Quiz condition, while the Radio-Quiz and Listening conditions scored equally well. In the Traffic scenario the Single condition remains the second-lowest performing. However, there is a difference between Radio Quiz and Listening, with the Radio Quiz leading to slightly lower driving performance overall. Looking at the variables related to overtaking, the order of the conditions is much less defined (with the exception of the Tablet Quiz): The Single condition leads to performance that is only slightly worse than the two aural conditions. The two remaining conditions, Listening and Radio Quiz, showed performance similar to each other.

Discussion

In this study we expanded on previous research regarding the positive and negative effects of multitasking during driving. We compared different secondary tasks and driving scenarios within a single paradigm. Furthermore, we used typical commute durations of 30 minutes for each condition to approximate realistic driving circumstances. This paradigm allowed us to compare secondary tasks and driving scenarios based on the cognitive requirements placed on the driver. To summarize, we found that when ordering the different secondary-task conditions based on the expected interference with driving (i.e., visual working-memory task, aural working-memory task, aural task, and finally no secondary task), a u-shaped pattern appears that was consistent across measurements. This pattern indicated that a visual working-memory secondary task resulted in the worst driving performance, while aural secondary tasks led to the best driving performance – better than not having a secondary task. The overtaking data did not show a consistent pattern, except for the visual tablet task, which led to significantly lower driving performance across measurements.

Thus, all results clearly show that the visual secondary task led to the worst driving performance across measures. However, the result that stands out was that driving without a secondary task did not lead to the best performance; instead listening to the radio resulted in better driving. This is in line with earlier research on monotonous driving conditions with very sparse traffic (Gershon et al., 2009; Atchley and Chan, 2010). We extended these results to a driving scenario with a substantial number of vehicles. Under these circumstances the driving task is more engaging because the driver must monitor, and react to, traffic.

As discussed in the introduction, at least two theories can explain why driving without a secondary task might lead to worse performance under both circumstances. Execution-focus theories imply that increasing the step-by-step attentional control of skilled processes – which is the case with driving as a single task – might disrupt proceduralized processes in sensorimotor tasks (Baumeister, 1984; Beilock & Carr, 2001). Actions that are normally executed as a single uninterrupted unit are divided up it to smaller units that are executed separately, leading to slower actions. Thus, over-regularization of the driving task can lead to a decrease in performance because the actions that must be performed are slowed down. This is similar to explicitly thinking of how each foot is placed while walking. The issue with these theories is that they assume that the driver is under significant performance pressure, which is unlikely in our paradigm for both of the traffic scenarios.

Another possibility is that while the traffic scenario that we used is not quite as monotonous as the sparse-traffic scenarios in previous studies, it is still quite a repetitive sequence of lane keeping and overtaking, which likely leads to boredom over a 30-minute drive. This might cause people to shift focus towards internal processes, resulting in a decoupling from the external environment (Cheyne et al., 2009; Smallwood & Schooler, 2015). This type of mind wandering can have a significant negative impact on driving behavior because it affects how well a driver observes the surroundings (He et al. 2011). To us, this seems the more likely explanation. To test this possibility, it would be interesting to measure mind wandering during driving directly, for example with EEG or pupil dilation (e.g., Mittner et al., 2014).

In general, the effects we found were small, with differences typically in the order of 10%. People seem to adapt their behavior to the driving circumstances. This is most evident in the average driving speed presented in Figures 7B and 7D: The difficulty of the visual tablet task leads drivers to slow down. In addition, there are indications that drivers prioritize the driving task during overtaking, given that they are inclined to postpone a response to a task until the maneuver has been completed. Despite this, concurrent performance of the task involving a tablet did clearly lead to the worst overtaking performance. This result could be explained if we assume that even during overtaking an occasional switch of attention to the secondary task would occur: a glance at the tablet is more costly compared to the other tasks, as visual interference has a larger impact on driving performance than aural interference because the environment can no longer be monitored.

The results of this study provide a more complete understanding regarding the interaction between secondary tasks and driving circumstances and the resulting driving performance. Essentially, the observations are in line with current theories of multitasking (Salvucci & Taatgen, 2008; Wickens, 2002): performance is primarily reduced when there is a resource conflict. A driving scenario with no right-lane traffic has low working-memory and motor load. Under these circumstances the driving task is complemented by tasks that require aural and working-memory facilities, as there is no resource overlap. While driving without a secondary task is expected to lead to the highest driving performance, it is also the condition with the highest risk of mind wandering. The effects of mind wandering are consistent with the observed driving performance: mind wandering can lead to a narrowed visual focus, which could cause insufficient monitoring of the environment (He et al., 2011). Finally, a visual task interferes strongly with driving task even when there is no traffic to account for, as lane keeping still requires constant visual attention.

When there is other traffic that needs to be reacted to, driving has a much higher working-memory and motor load: the location of surrounding vehicles has to be monitored, and the environment has to be navigated (Gugerty, 1997). As a consequence, driving performance was affected when a secondary task required additional information to be maintained in working memory: a concurrently performed aural working-memory task led to worse driving performance than an aural listening task, as primarily seen in the average driving speed. This is not observed when there is no surrounding traffic, thus it is likely due to the overlap in working-memory requirements. The effect is only minor, however, as driving without a secondary task still leads to comparatively worse driving performance. A visual working-memory task again leads to the lowest driving performance, as there is overlap in two crucial resources.

The study we performed shares similarities with the work by Gershon et al. (2009) and Atchley and Chan (2010), who both showed that a secondary task during driving could have a positive effect. Furthermore, Brookhuis et al. (1991) found that on monotonous stretches of highway a telephone conversation would improve a driver's swerving behavior. However, whereas these previous investigations tested the effect of either a single secondary task on driving performance or a single driving scenario (or both), we tested a range of different tasks with different resource requirements, under different driving conditions. The paradigm we used made it possible to directly compare secondary tasks, as well as investigate the interaction between secondary task and driving circumstances. This allowed us to determine that a simple listening task complements driving more consistently under varying traffic circumstances than the relatively complex tasks used by Gershon et al. (2009) and Atchley and Chan (2010). Furthermore, we found no evidence that a more involved secondary task can have a stronger positive effect on driving than the passive listening task.

To conclude, safe multitasking during driving depends on engaging in tasks that complement the requirements of driving at that particular time. When the driver is fully engaged, such as when driving through city traffic, it is best to focus fully on driving as indicated by previous research (Neyens & Boyle, 2007; Stein, Parseghian, & Allen, 1987). However, on roads with low traffic density the driving task is much less demanding, and may lead to mind wandering. Such an internally focused distraction can lead to bad driving performance because the environment is no longer monitored sufficiently. While mind wandering is not as dangerous as a visual distraction during driving, this work shows that it might be sensible to engage in mildly distracting activities such as listening to the radio. These can prove beneficial to driving performance by providing a less interfering task alternative.

Appendix: Integrated Driving Performance

Here we report the performance differences between conditions and driving scenarios when integrating the different measurements. When combined, the measurements we computed using the driving data present a comprehensive evaluation of driving consistency and safety under varying cognitive and peripheral loads. For each dependent variable an order can be determined per participant that ranks the four different conditions from worst to best performance. From best to worst, conditions

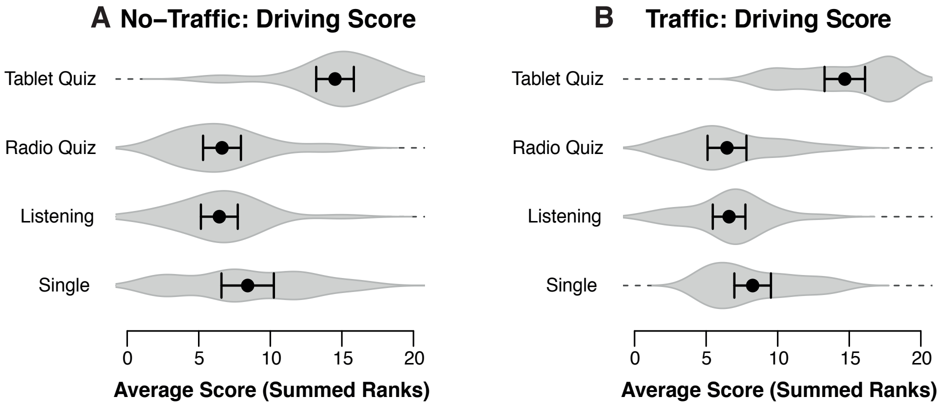

Figure 4-9. Aggregated performance scores for each condition. Black dots represent the mean across subjects, and bars denote 95% CI. *Panel A:* The mean penalty for driving performance during the No-Traffic scenario. *Panel B:* The mean penalty for driving performance during the non-overtaking sections of the Traffic scenario.

Comparisons	No-Traffic			Traffic		
	z	*β*	*p*	*z*	*β*	*p*
Single vs. Passive	**-2.61**	**-.562**	**.014**	**-2.21**	**-.463**	**.033**
Single vs. Radio-Quiz	**-2.27**	**-.484**	**.028**	**-2.47**	**-.524**	**.020**
Radio-Quiz vs. Passive	-.371	-.078	.711	.289	.061	.773
Tablet-Quiz vs. Single	**-8.89**	**-2.19**	**< .001**	**-9.48**	**-2.40**	**< .001**
Tablet-Quiz vs. Passive	**-11.0**	**-2.75**	**< .001**	**-11.1**	**-2.87**	**< .001**
Tablet-Quiz vs. Radio-Quiz	**-10.8**	**-2.67**	**< .001**	**-11.3**	**-2.93**	**< .001**

Table 4-6. Between-conditions comparisons of ranked measurements related to lane keeping in both driving scenarios. The resulting *z* values, *p* values, and estimates (β) are reported. The *p*-values were corrected for multiple comparisons using a False Discovery Rate correction. Bold numbers indicate significance at the 0.05 level.

were given a rank from 0 to 3 depending on the performance of the particular participant. To visualize these rankings they were summed to determine a total score per condition, the averages of these scores can be seen in Figure 4-9. A higher score indicates worse overall performance for that condition. To score the lane keeping performance we summed the ranks of average lane deviation, directional changes, max deviation, wheel angle, and average speed. The limitation of this scoring approach is that it is based on an ordinal scale, and therefore not continuous. However, it is the best method we have found to integrate the dissimilar measurements of driving performance.

To determine if the ranks achieved by participants were significantly different between conditions, we computed cumulative-link models (Agresti, 2002) for both traffic scenarios. The output of these models can be seen in Table 4-6. Integrated driving performance scores for the No-Traffic (Figure 4-9A) and Traffic (Figure 4-9B) scenarios followed the same u-shaped pattern seen in plots of individual measurements in the main text. The Tablet-Quiz has the highest average aggregated score, which was significantly different from the remaining three conditions. The scores of the Listening and Radio-Quiz conditions were significantly lower than the Single condition in both driving scenarios, and thus the two auditory conditions led to the best performance on average.