

Article

“Just One Short Voice Message”—Comparing the Effects of Text- vs. Voice-Based Answering to Text Messages via Smartphone on Young Drivers’ Driving Performances

Max Kurtz ¹, Michael Oehl ^{2,*}  and Christine Sutter ¹

¹ Institute of Traffic and Engineering Psychology, German Police University, Zum Roten Berge 18–24, 48165 Münster, Germany; maxkurtz@gmail.com (M.K.); Christine.Sutter@dhopol.de (C.S.)

² Institute of Transportation Systems, German Aerospace Center (DLR), Lilienthalplatz 7, 38108 Braunschweig, Germany

* Correspondence: Michael.Oehl@dlr.de

Abstract: Despite the well-known distracting effects, many drivers still engage in phone use, especially texting and especially among young drivers, with new emerging messaging modes. The present study aims to examine the effects of different answering modes on driving performance. Twenty-four students (12 females), aged between 19 and 25 years ($M = 20.83$, $SD = 1.53$), volunteered for the study. They accomplished the Lane Change Task (LCT) with baseline and dual-task runs in a driving simulator. In dual-task runs, participants answered text messages on a smartphone by voice or text message with varying task complexity. Driving performance was measured by lane deviation (LCT) and subjective measures (NASA-TLX). Across all trials, driving performance deteriorated during dual-task runs compared with the baseline runs, and subjective demand increased. Analysis of dual-task runs showed a benefit for voice-based answering to received text messages that leveled off in the complex task. All in all, the benefits of using voice-based answering in comparison with text-based answering were found regarding driving performance and subjective measures. Nevertheless, this benefit was mostly lost in the complex task, and both the driving performance and the demand measured in the baseline conditions could not be reached.

Keywords: driver behavior; driver distraction; texting while driving; voice messaging; driving performance; Lane Change Task



Citation: Kurtz, M.; Oehl, M.; Sutter, C. “Just One Short Voice Message”—Comparing the Effects of Text- vs. Voice-Based Answering to Text Messages via Smartphone on Young Drivers’ Driving Performances. *Safety* **2021**, *7*, 57. <https://doi.org/10.3390/safety7030057>

Academic Editor: Raphael Grzebieta

Received: 2 December 2020

Accepted: 23 July 2021

Published: 30 July 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Smartphones have become ubiquitous in our everyday lives and are used in many situations, even in situations where their use is inappropriate, like driving. Research has shown the numerous detrimental effects of phone use while driving. More than 40 years ago, the first adverse effects were found [1]. Recent meta-analyses show effects on reaction time, lane keeping, headway and speed in both simulator and field studies [2–4]. The effects of phone use on driving seem comparable to drunk driving at the legal limit when the time on task is controlled [5,6]. With advancing technology, texting gained popularity as another potential distracting feature of phone use.

Similarly, texting while driving became an important research field, with meta-analyses showing effects on reaction times, collisions, lane keeping, speed and headway [7]. Impairments are even more significant for texting than for phone calls [8,9]. Regardless of drivers recognizing the distraction of texting and adapting by choosing longer headways, risky driving behavior like increased glances off the road were still shown [10,11]. Although aware of the increased risk of road traffic crashes, many drivers engage in phone use while driving, even though manual phone use is not legal in many countries [12,13]. Phone use and distracted driving are still relevant today, and distracted driving will become even more relevant with increasing infotainment systems and advanced driver assistant systems in modern vehicles.

The openness to risk is partially age-related. Adolescents and young adults seek risks out of the powerful need for peer acceptance and approval, self-presentation and self-esteem. They seek many risks in road traffic (e.g., drunk driving [14] and speeding) and are prone to phone use while driving [15,16]. Consequently, it is not surprising that young drivers are a vulnerable group in road traffic, motivating the present study to focus on phone use and distracted driving in young drivers.

Technology continues advancing with smartphones as new means of communication and new instant messaging applications gain popularity, offering alternative ways to answer messages from photos over videos to voice messages. For answering messages while driving, voice-based input especially seems to be attractive as it offers seemingly less (visual and manual) distraction than text-based input at the human–machine interface and could be perceived as less intrusive, because instead of typing out the message, drivers record the spoken out message via a button push. Previous research already looked at the benefits of using voice-based systems while driving to control navigation systems or send emails. Although benefits of voice-based input compared with visual-manual input can be found, voice-based input still presents a critical distraction [17–20], leading to unsafe driving behaviors. Nevertheless, research comparing the effects of voice and text messaging is still very limited. Choudhary and Velaga [9] compared phone conversations and text messages with different complexities and found text messages to be more distracting. Studies further looked at the effects of listening to and answering voice messages [21–23] or showed imperfect voice recognition as a major cause of distraction and detrimental effects on driving [24]. Most of these studies compared messages read aloud or phone conversations with text messages, more or less dividing different answering modes.

In many countries, the use of handheld devices (such as mobile phones and tablets) while driving is illegal. Often, it is not a matter for legislation but for law enforcement authorities to detect and take action against the use of handheld mobile activity and the legal compliance of drivers. An increasing number of drivers still illegally use handheld communication devices while driving (e.g., [25,26]). Voice-based and hands-free devices offer an alternative and legal way of navigation, entertainment and communication while driving. Still, even those devices require a minimum of manual and visual action control, such as for pushing a button or glancing at the map of a navigation system. To investigate this kind of driver–device interaction in a controlled setting, the present study focuses on the explicit functionality of a voice-based answering mode as an alternative and maybe safer response method while driving compared with a visual-manual (i.e., text-based) answering mode. The task in this study was to respond either vocally (voice-based) or visual-manually (text-based) to a five-digit code visually presented on a display. Only a few studies have focused on comparing text and voice messages [27], showing the need for more research.

Many studies have shown the effects of driver inattention on lane keeping [28,29], offering a valid measure. The impact of phone use on lane keeping is not so pronounced, with few studies reporting smaller effects [30]. Texting while driving shows a similar picture, with substantial effects on lane keeping [31] for phones with physical keyboards and touchscreens [32,33]. As a standardized tool, the Lane Change Task (LCT) was used in previous studies to assess the influence of phone use on lane keeping. Detrimental effects of texting on driving performance in the LCT were already found [34], even for different entry methods [32]. Furthermore, voice-based entry for in-vehicle information systems [18] or navigation systems [17] showed improved performance over manual entry but still posed a significant distraction and degraded performance in the LCT.

In this study, the effect of distraction on driving performance will be looked at in lane keeping behavior as it offers a valid measure. Furthermore, earlier studies comparing phone conversation or voice messaging with text messaging, on the other hand, mainly relied on car-following tasks [22,27] or looked at driving violations [23]. Therefore, the measure of lane keeping in this study can build a clearer understanding concerning the impact of text- vs. voice-based answering on driving performance.

Additionally, in this study, the task complexity of the secondary task of text- and voice-based answering will be varied to better understand the workload on driving performance, which was rarely investigated before [9,35].

It is expected in this study that engaging in either secondary task of text- or voice-based answering will lead to increased attentional demand, leading to impaired driving performance. Voice-based answering should lead to less impaired driving performance compared with text-based answering, as both use different modalities according to Wickens' multiple resource theory [36], with voice-based answering relying more on auditory and text-based answering relying primarily on visual attention. The higher complexity of answers leads to more impaired driving performance as increased task demands compete for attention resources.

2. Materials and Methods

2.1. Design

Two independent within-person variables were realized to investigate possible detriments and compare the effects of voice- and text-based answering on driving performance. The answering mode (text vs. voice input) served as the first independent variable, whereas task complexity (simple vs. complex task) constituted the second independent variable. Both variables were manipulated in a 2×2 factorial design, combining each level of answering mode and task complexity to a sum of four conditions (text and simple task, voice and simple task, text and complex task and voice and complex task). Driving performance, as well as subjectively perceived demand, served as dependent variables. A repeated measures within-subject design was used to control for individual differences in driving performance.

2.2. Participants

In this study, a total of 24 students ($n = 24$, 12 females) with a mean age of 20.83 years ($SD = 1.53$), ranging from 19 to 25 years participated. All participants were frequent drivers, driving at least once a week, on average 5802 km ($SD = 4043$) per year and owning a driver's license for at least 2 years. All participants were frequent smartphone and instant messaging users, both using them daily and being accustomed to voice messaging. All had normal or corrected-to-normal vision, were right-handed and were novices in the LCT. No participant-reported instances of simulator sickness. All participants were university students and participated voluntarily. All participants were able to abort the study at any time and were debriefed afterward. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted under the Declaration of Helsinki and the Ethical Principles and Protocol Code of the Federation of German Psychologists Associations (adapted from the American Psychological Association (APA) Code of Ethics). The Ethik-Kommission der Deutschen Hochschule der Polizei (Ethics Committee of the German Police University) provided post hoc approval of the study (approval number DHPol-EthK.2020.Su1).

2.3. Apparatus

2.3.1. Driving Simulator and Lane Change Task

Figure 1 depicts the experimental setup. A medium-fidelity fixed base simulator was used, made up of a Golf 4 GTI dashboard equipped with a Logitech G25 racing wheel and pedals. The simulation was projected on a $1.4 \text{ m} \times 1.4 \text{ m}$ screen.

Figure 2 depicts the Lane Change Task version 1.2 [37] used in the present study.



Figure 1. Sketch of the experimental setup.



Figure 2. Screenshot in the study simulator setting of the Lane Change Task version 1.2 by Mattes [37].

2.3.2. Messaging Task

The messaging task was carried out on a Samsung Galaxy S II GT-I9100 with a 4.3" display and a resolution of 480×800 pixels running Google Android Version 4.1.2. The smartphone was fixated on the dashboard within easy reach for all participants to ensure comparability across trials. All tasks were completed within the instant messaging app WhatsApp version 4.12.449. Messages were sent to the smartphone by the experimenter. The texting task was adapted from Burge and Chaparro [38]. Participants received text messages containing letter strings made of five random letters (e.g., MZPAL). They were told to answer the text messages in four different ways, varying the answer mode and task complexity. In the simple task, participants had to copy the text messages by repeating the letters in order and sending them back. They were asked to sort the letters in alphabetical order before sending them back in the complex task. Secondly, messages were answered either via text by typing on the touch screen keyboard or via voice input by recording the message. Text messages were sent in capital letters but answered in lowercase letters, as pre-tests showed this to be more comfortable for participants. Participants were told to answer the messages as soon as they received them and respond correctly. The following message

was sent as soon as the participants had answered. This way, the tasks were completed continuously during runs with just the network communication delay in between.

2.4. Measures

2.4.1. Lane Change Task

Driving performance was measured by the mean deviation in meters. As recommended by ISO 26022 [39], the adaptive model was used. In the adaptive model, the participants' baseline is individually derived from actual baseline runs instead of a normative baseline for every participant. It is advantageous compared with the basic model because it discriminates more effectively and allows better comparability between participants by considering individual driving styles [40,41]. The LCT track between the start sign and the last lane change sign was used for analysis, resulting in a total driving time of 160 s.

2.4.2. NASA-TLX

The NASA-TLX [42] was used to measure subjective demand, with scores summed up for each dual-task run and the baseline run of the LCT (see Figure 3). Participants were asked to fill in questionnaires and to rate the subjective demand of the previous run according to their perception, after dual-task runs and after the first baseline run. The raw version of the NASA-TLX was used because it has been proven to generate comparable results to the weighted version [43] while requiring less time.

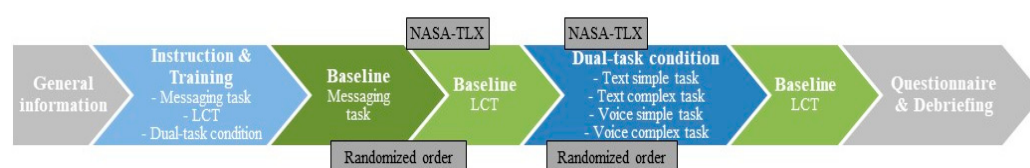


Figure 3. Procedure of the experiment.

2.5. Procedure

Figure 3 displays the procedure of the experiment. First, potential participants were informed about the experimental goals and procedures. If the participants provided their informed consent, they were familiarized with the setup and received instructions for the different tasks. Second, participants received task instructions and a training session to ensure familiarity with the setup and tasks. The training sessions contained the messaging task, the LCT and the dual-task condition (messaging task and LCT simultaneously). Third, baseline runs were conducted for the messaging task and the LCT. The order of the letter strings that had to be repeated while answering messages was standardized for each task condition. The order of blocks for the baseline runs was randomized. The baseline runs allowed comparability with performance during the dual-task condition.

The training and baseline runs were followed by the experimental dual-task condition, with participants actually answering while driving. The dual-task runs were completed in a blocked design, with one condition per run. This meant participants either answered by text or voice message and just worked on simple or complex messages while driving. According to ISO 26022 [39], participants were instructed to complete the lane changes quickly and efficiently and complete the driving and messaging tasks to the best of their ability, emphasizing performance for both. The order of dual-task runs was randomized. The order of letter strings and the order of the tracks were standardized for each condition to ensure the same requirements across participants. Finally, after completing the dual-task runs, another baseline run for the LCT was conducted to control the learning effects. To test for sufficient familiarity with the LCT, the performance in the baseline runs had to be under a mean deviation of 1.2 in the basic model according to the ISO 26022 [39] requirement. The experiment ended with a questionnaire (i.e., demographics, driving, smartphone experience and usage) and each participant's debriefing. After each of the four

dual-task runs and after the first baseline run, participants completed the NASA-TLX to measure subjective demand.

3. Results

When not stated otherwise, differences between the conditions were analyzed via repeated measures ANOVA and multiple *t*-tests with Bonferroni correction. Differences between the dual-task and baseline runs were analyzed via ANOVA, followed by analyses of the differences between the dual-task runs only with another ANOVA. Furthermore, data were analyzed for outliers via *z*-transformation. One participant's dataset was excluded from analysis for having outlying values in several runs of the LCT. The values of the remaining 23 participants were included in the analyses. Figure 4 shows an overview of the driving performance and subjective demand during dual-task runs.

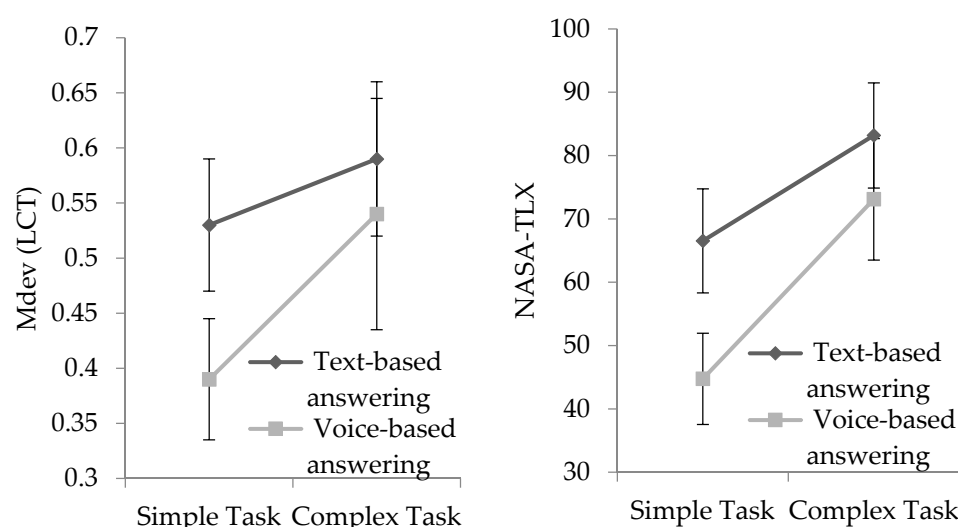


Figure 4. Main and interaction effects of mean deviations in the LCT (left) and NASA-TLX scores (right) for text- vs. voice-based answering in simple vs. complex tasks.

3.1. Driving Performance

Table 1 shows a summary of the means and standard deviations for driving performance and the subjective measures. The mean of the two baseline runs was used as the baseline in the following analyses. Differences in the mean deviation across all runs, the four dual-tasks and the baseline were analyzed via repeated measures ANOVA. Since Mauchly's test indicated a violation for the assumption of sphericity ($\chi^2(9) = 30.58$, $p < 0.001$), the degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity ($\epsilon = 0.60$ [44]). The omnibus F-test showed that the mean deviation differed significantly across runs ($F(2,38,52.45) = 38.58$, $p < 0.001$, $\eta^2 = 0.975$). In comparison with baseline driving ($M = 0.28$, $SD = 0.05$), the participants drove worse when sending text-based answers, regardless of whether the task was simple ($M = 0.53$, $SD = 0.12$) or complex ($M = 0.59$, $SD = 0.24$), and when sending voice-based answers, regardless of whether the task was simple ($M = 0.39$, $SD = 0.11$) or complex ($M = 0.54$, $SD = 0.21$). Post hoc pairwise comparisons showed these differences to be significant for all runs when writing text-based answers during the simple ($t(22) = 13.25$, $p < 0.001$, $d = -2.29$) and the complex tasks ($t(22) = 5.62$, $p < 0.001$, $d = -2.26$), as well as when recording voice-based answers during the simple ($t(22) = 11.92$, $p < 0.001$, $d = -1.01$) and the complex tasks ($t(22) = 6.63$, $p < 0.001$, $d = -1.23$).

Factorial repeated-measures ANOVA for just the dual-task runs showed a significant main effect for the answering mode ($F(1,22) = 12.84$, $p = 0.002$, $\eta^2 = 0.16$) and a significant main effect for the task complexity ($F(1,22) = 37.82$, $p < 0.01$, $\eta^2 = 0.22$). On average, the participants were able to drive better when sending voice answers ($M = 0.47$, $SD = 0.18$) than when sending text-based answers ($M = 0.56$, $SD = 0.13$) and when work-

ing on the simple task ($M = 0.46$, $SD = 0.13$) than when working on the complex task ($M = 0.57$, $SD = 0.18$). Furthermore, a significant interaction effect was found ($F(1,22) = 4.962$, $p = 0.036$, $\eta^2 = 0.04$). Post hoc comparisons showed that when working on the simple task, the mean deviation differed significantly between voice- and text-based answering ($t(22) = 5.8$, $p < 0.001$, $d = -1.26$), but when working on the complex task, the mean deviation between voice- and text-based answering did not differ significantly ($t(22) = 1.15$, $p = 0.263$, $d = -0.33$). The beneficial effect of voice-based answering in comparison with text-based answering was attenuated when working on the complex task.

Table 1. Means and standard deviations of the mean deviation in the LCT and NASA-TLX scores in the baseline and dual-task runs.

	Baseline		Text Simple Task		Text Complex Task		Voice Simple Task		Voice Complex Task	
	M	SD	M	SD	M	SD	M	SD	M	SD
Mdev (LCT)	0.28	0.05	0.53	0.12	0.59	0.14	0.39	0.11	0.54	0.21
NASA-TLX	31.09	16.72	66.52	16.43	83.17	16.62	44.74	14.39	73.09	19.18

3.2. Subjective Demand

Subjective demand, in terms of the scores in the NASA-TLX, was analyzed across all runs via repeated measures ANOVA. Mauchly's test was significant ($\chi^2(9) = 24.89$, $p = 0.003$), and therefore the Greenhouse–Geisser corrected values were used ($\epsilon = 0.73$ [44]). The omnibus F-Test showed that scores significantly differed across runs ($F(4,88) = 65.58$, $p < 0.001$, $\eta^2 = 0.99$). In comparison with baseline driving ($M = 31.01$, $SD = 16.72$), the participants perceived higher demands when sending text-based answers, regardless of whether performing a simple ($M = 66.52$, $SD = 16.42$) or complex task ($M = 83.17$, $SD = 16.61$), and when sending voice-based answers, regardless of whether performing a simple ($M = 44.74$, $SD = 14.39$) or complex task ($M = 73.09$, $SD = 19.18$). Post hoc pairwise comparisons showed this difference to be significant for all runs, both when writing text-based answers during the simple ($t(22) = 7.9$, $p < 0.001$, $d = -2.16$) and complex tasks ($t(22) = 12.03$, $p < 0.001$, $d = -3.13$) as well as when recording voice-based answers during the simple ($t(22) = 3.78$, $p = 0.001$, $d = -0.95$) and complex tasks ($t(22) = 9.90$, $p < 0.001$, $d = -2.19$).

Factorial repeated-measures ANOVA for just the dual-task runs showed a significant main effect for the answering mode ($F(1,22) = 27.87$, $p < 0.001$, $\eta^2 = 0.22$) and a significant main effect for task complexity ($F(1,22) = 92.7$, $p < 0.001$, $\eta^2 = 0.43$). On average, the participants perceived driving when sending voice-based answers ($M = 59.35$, $SD = 22.14$) to be less demanding than when sending text-based answers ($M = 75.46$, $SD = 18.18$) and working on the simple task ($M = 56.04$, $SD = 18.88$) to be less demanding than working on the complex task ($M = 78.77$, $SD = 18.34$). Furthermore, a significant interaction effect was found ($F(1,22) = 12.96$, $p = 0.002$, $\eta^2 = 0.03$). Post hoc comparisons showed that subjective demand differed significantly between voice- and text-based answering ($t(22) = 6.52$, $p < 0.001$, $d = -1.33$) when working on the simple task as well as during the complex task ($t(22) = 2.87$, $p = 0.009$, $d = -0.61$). When sending text answers, subjective demand differed significantly between the simple and complex tasks ($t(22) = -9.31$, $p < 0.001$, $d = 1.01$), as well as when sending voice answers ($t(22) = -7.86$, $p < 0.001$, $d = 1.97$).

3.3. Messaging Performance

Table 2 shows a summary of the means and standard deviations for answering performance. Differences in the number of responses between the dual-task runs and the baseline runs were tested via repeated t -tests and corrected using Bonferroni adjustment. During the dual-task runs, all messaging performance decreased significantly compared with the

baseline runs (all $p < 0.001$, except for voice-based answering during the simple task with $p = 0.002$), with participants writing fewer answers while driving.

Table 2. Means and standard deviations of the number of answers and number of erroneous answers sent by participants in the baseline and dual-task runs.

	Text Simple Task				Text Complex Task			
	Baseline		Dual Task		Baseline		Dual Task	
	M	SD	M	SD	M	SD	M	SD
Number	19.17	2.12	14.09	2.07	13.22	2.61	8.96	2.38
Erroneous	1.13	1.14	2.17	1.59	1.22	1.24	1.96	1.43
	Voice Simple Task				Voice Complex Task			
	Baseline		Dual Task		Baseline		Dual Task	
	M	SD	M	SD	M	SD	M	SD
Number	20.35	2.39	18.43	1.31	15.3	2.53	11.78	2.15
Erroneous	0.39	0.58	0.61	1.50	1.17	1.23	2.47	1.71

Differences in the number of erroneous answers between the dual-task and the baseline runs were analyzed in the same way. Similarly, the answering performance decreased during dual-task runs compared with the baseline runs, with participants writing more erroneous answers. The comparisons were significant for the text-based answers, both during simple ($p = 0.018$) and complex ($p = 0.032$) tasks, but for voice-based answers, significance was found just during complex ($p = 0.001$) and not during simple ($p = 0.487$) tasks.

4. Discussion

We conducted this study to compare the possible detrimental effects of text- and voice-based answering with text messages on driving performance. Participants answered text messages of varying complexity with different answering modes (text vs. voice) while driving in the Lane Change Task. Driving performance and subjective measures were obtained. The driving performance differed significantly from the baseline run in terms of the mean deviation in all dual-task runs. As expected, the secondary task's increased demand led to an objective decrease in performance in the primary task of driving. The increased demand could also be seen in a subjective dimension in the higher scores of the NASA-TLX for dual-task runs. Driving while sending answers was perceived as more demanding than baseline driving. This effect was found for all conditions of the answering tasks (i.e., text- and voice-based answers and simple and complex tasks).

When just analyzing the dual-task runs, it can be seen that both the answering mode and the task complexity had individual impacts on driving performance. Typing a text-based answer led to more significant impairment than recording a voice-based answer, and sorting the answer content alphabetically beforehand (complex task) led to greater impairment than copying them (simple task). When regarding the effect sizes, the influence of the task complexity was slightly larger than that of the answering mode. Nevertheless, both effects were substantial. These effects were also reflected in the subjectively experienced demands of the runs, as the NASA-TLX showed. Sending text-based answers and working on the complex task was more strenuous for participants than sending voice-based answers and working on the simple task. When looking at driving performance, the interaction of the answering mode and task also played an important role. In general, using voice input was less disadvantageous than text-based answering while still remaining inferior to baseline driving. This difference was more prominent during the simple task condition but became somewhat lost in the complex task condition. The increase in mean deviation in the LCT from simple to complex tasks was more pronounced when using voice-based answering than when using text-based answering. The voice input showed fewer disad-

vantages in driver performance, but this difference nearly leveled off when facing a more complex task.

Despite some clear evidence, the current study also faces some limitations that should be addressed in further research. The first limitation is the LCT's inability to effectively distinguish between the visual and cognitive demands of the task. The effects of the complex task could stem from longer gaze times instead of higher cognitive demand. Follow-up studies could use a more articulate analysis of the LCT data, which was not available for this study but was used by Engström and Markkula [45]. Furthermore, the LCT in a simulator setting can be seen as entirely artificial (i.e., not taking traffic or speed into account). Unlike in naturalistic driving, participants worked continuously on the task and could not freely choose when to answer. Jamson et al. [22] mentioned that choosing when to answer is a benefit of text messaging compared with a phone call, since messaging can be interrupted more quickly than a phone call.

Further research in more naturalistic settings is necessary to examine the performance and the choice of different answering modes in a complex driving environment, such as Peng, Boyle, and Hallmark [29] did by looking at driver inattention. As only young drivers participated in the presented study, exact effects on other age groups should be addressed by further research. Recent studies for texting while driving have shown that older drivers might be more negatively affected by texting while driving as young drivers seem more experienced with texting [46], while distracted driving is a higher risk factor for young drivers [47].

Regardless of the limitations, evidence shows that using voice-based answering can be less disadvantageous to driving performance compared with text-based answering. Nevertheless, it is essential to underline that the baseline performance without a secondary task cannot be achieved. As with earlier studies looking at different metrics [21,27], driving performance in lane keeping was similarly impacted. This finding is also in line with previous research findings showing that voice-based systems are helpful compared with manual control while still impairing driving performance [18,20,48]. Especially when the effect of a task's complexity is considered, it can be seen that voice-based solutions have limitations. When the verbal interaction gets too complicated, the advantage of using a voice system is mainly lost. It was shown that voice-based answering could not be seen as an alternative to text-based answering while driving. Both secondary tasks impair driving performance to a safety-critical extent. As young drivers are especially prone to texting while driving [25,49], their driving behavior will be probably impacted. As He et al. [27] stated, voice-based messaging is not a panacea. Therefore, it can be emphasized again that messaging and especially texting while driving should be avoided at all times.

5. Conclusions

The objective of this study was to compare the effects of text- and voice-based answering to text messages on young drivers' driving performance. The effects of different answering modes (voice vs. text) could be clearly distinguished in terms of driving performance and subjective demand reported by participants. While answering via voice was less disadvantageous and reported as less demanding than text-based answering, baseline driving performance and demand could still not be reached, showing again that messaging while driving, independent of answering mode, impacted the driving performance of young drivers and should be avoided.

In more general terms, the study's results might guide road safety measures in engineering, education, enforcement and public relations. Concerning engineering, the design of the driver-device interaction in manually driven vehicles is essential for increasing vehicle safety. Vehicle manufacturers should introduce fully voice-based and hands-free navigation, infotainment and communication systems as standard features in upcoming vehicles. Those systems unfold their full safety potential the simpler and more intuitive the driver-device interaction is. Furthermore, the distracting effects of manual-visual control and the negative consequences on driving performance were severe and likely underesti-

mated in the present simulator study. To further increase vehicle safety, any manual-visual driver-device interaction should be reduced to the essential to prevent visual and manual distractions. It seems worthwhile to evaluate traditional manual-visual driver-vehicle interfaces and find speech-based or other (not manually or vision-based) substitutes. Further research should be addressed toward alternative designs of driver-vehicle interfaces.

Concerning education, enforcement and public relations, the findings once more emphasize that manual-visual driver-device interaction bears severe risks, even with very simple tasks. The study's findings support the current regulation in many countries of banning the use of handheld devices while driving. The message that manual-visual distraction should be avoided at any time while driving is still valid and gains even more weight in today's technology-driven environment. Stakeholders in road traffic safety, such as public service and road safety organizations, should continue with related public awareness campaigns.

Author Contributions: Conceptualization, M.K. and M.O.; methodology, M.K. and M.O.; formal analysis, M.K.; writing—original draft preparation, M.K., M.O. and C.S.; writing—review and editing, M.K., M.O. and C.S.; visualization, M.K. and C.S.; supervision, M.O.; funding acquisition, M.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The APC was funded by the German Aerospace Center (DLR).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and the Ethical Principles and Protocol Code of the Federation of German Psychologists Associations (adapted from the American Psychological Association (APA) Code of Ethics). The Ethik-Kommission der Deutschen Hochschule der Polizei (Ethics Committee of the German Police University) provided post hoc approval of the study (approval number DHPol-EthK.2020.Su1 and date of approval 18 January 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: We wish to thank Louise Leitsch and Linn Petersdotter for valuable support in the earlier stages of the study, and Nora Cramer for continuous support throughout the project.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. Brown, I.D.; Tickner, A.H.; Simmonds, D.C. Interference between concurrent tasks of driving and telephoning. *J. Appl. Psychol.* **1969**, *53*, 419–424. [[CrossRef](#)]
2. Caird, J.K.; Willness, C.R.; Steel, P.; Scialfa, C. A meta-analysis of the effects of cell phones on driver performance. *Accid. Anal. Prev.* **2008**, *40*, 1282–1293. [[CrossRef](#)] [[PubMed](#)]
3. Horrey, W.J.; Wickens, C.D. Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Hum. Factors* **2006**, *48*, 196–205. [[CrossRef](#)]
4. Lipovac, K.; Deric, M.; Tescic, M.; Andric, Z.; Maric, B. Mobile phone use while driving-literary review. *Transp. Res. Part F* **2017**, *47*, 132–142. [[CrossRef](#)]
5. Burns, P.C.; Parkes, A.; Burton, S.; Smith, R.K.; Burch, D. *How Dangerous is Driving with a Mobile Phone? Benchmarking the Impairment to Alcohol*; Transport Research Laboratory: Berkshire, UK, 2002.
6. Strayer, D.L.; Drews, F.A.; Crouch, D.J. A comparison of the cell phone driver and the drunk driver. *Hum. Factors* **2006**, *48*, 381–391. [[CrossRef](#)] [[PubMed](#)]
7. Caird, J.K.; Johnston, K.A.; Willness, C.R.; Asbridge, M.; Steel, P.A. A meta-analysis of the effects of texting on driving. *Accid. Anal. Prev.* **2014**, *71*, 311–318. [[CrossRef](#)] [[PubMed](#)]
8. Drews, F.A.; Yazdani, H.; Godfrey, C.N.; Cooper, J.M.; Strayer, D.L. Text messaging during simulated driving. *Hum. Factors* **2009**, *51*, 762–770. [[CrossRef](#)] [[PubMed](#)]
9. Choudhary, P.; Velaga, N.R. Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour. *Accid. Anal. Prev.* **2017**, *106*, 370–378. [[CrossRef](#)] [[PubMed](#)]
10. Peng, Y.; Boyle, L.N. Driver's adaptive glance behaviour to in-vehicle information systems. *Accid. Anal. Prev.* **2015**, *85*, 93–101. [[CrossRef](#)]

11. Tsimhoni, O.; Green, P. Visual demand of driving and the execution of display-intensive in-vehicle tasks. In Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, Minneapolis, MN, USA, 8–12 October 2001; pp. 1586–1590. [\[CrossRef\]](#)
12. McDonald, C.C.; Sommers, M.S. Teen drivers' perceptions of inattention and cell phone use while driving. *Traffic Inj. Prev.* **2015**, *16*, 52–55. [\[CrossRef\]](#)
13. Oviedo-Trespalacios, O.; Truelove, V.; King, M. "It is frustrating to not have control even though I know it's not legal!": A mixed-methods investigation on applications to prevent mobile phone use while driving. *Accid. Anal. Prev.* **2020**, *137*, 105–412. [\[CrossRef\]](#)
14. Arnett, J. Drunk driving, sensation seeking, and egocentrism among adolescents. *Pers. Individ. Differ.* **1990**, *11*, 541–546. [\[CrossRef\]](#)
15. Cook, J.L.; Jones, R.M. Texting and accessing the web while driving: Traffic citations and crashes among young adult drivers. *Traffic Inj. Prev.* **2011**, *12*, 545–549. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Vollrath, M.; Huemer, A.K.; Teller, C.; Likhacheva, A.; Fricke, J. Do German drivers use their smartphone safely?—Not really! *Accid. Anal. Prev.* **2016**, *96*, 29–38. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Harbluk, J.L.; Burns, P.C.; Lochner, M.; Trbovich, P.L. Using the lane-change test (LCT) to assess distraction: Tests of visual-manual and speech-based operation of navigation system interfaces. In Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Washington, DC, USA, 9–12 September 2007; pp. 16–22.
18. Maciej, J.; Vollrath, M. Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accid. Anal. Prev.* **2009**, *41*, 924–930. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Simmons, S.; Caird, M.J.K.; Steel, P. A meta-analysis of in-vehicle and nomadic voice-recognition system interaction and driving performance. *Accid. Anal. Prev.* **2017**, *106*, 31–43. [\[CrossRef\]](#)
20. Tsimhoni, O.; Smith, D.; Green, P. Address entry while driving: Speech recognition versus a touch-screen keyboard. *Hum. Factors* **2002**, *46*, 600–610. [\[CrossRef\]](#)
21. Coleman, J.R.; Turrill, J.; Cooper, J.M.; Strayer, D.L. Cognitive workload using interactive voice messaging systems. In Proceedings of the Human Factors and Ergonomics Society 60th Annual Meeting, Washington Hilton, DC, USA, 19–23 September 2016; pp. 1894–1898. [\[CrossRef\]](#)
22. Jamson, A.H.; Westerman, S.J.; Hockey, G.R.J.; Carsten, O.M. Speech-based e-mail and driver behavior: Effects of an in-vehicle message system interface. *Hum. Factors* **2004**, *46*, 625–639. [\[CrossRef\]](#)
23. Motamedi, S.; Wang, J. The impact of text driving on driving safety. *Int. J. Traffic Transp. Eng.* **2016**, *6*, 325–338. [\[CrossRef\]](#)
24. Yager, C. *An Evaluation of the Effectiveness of Voice-to-Text Programs at Reducing Incidences of Distracted Driving*; National Technical Information Service: Alexandria, DC, USA, 2013.
25. Hill, L.; Rybar, J.; Styer, T.; Fram, E.; Merchant, G.; Eastman, A. Prevalence of and attitudes about distracted driving in college students. *Traffic Inj. Prev.* **2015**, *16*, 362–367. [\[CrossRef\]](#)
26. Prat, F.; Gras, M.E.; Planes, M.; Font-Mayolas, S.; Sullman, M.J.M. Driving distractions: An insight gained from roadside interviews on their prevalence and factors associated with driver distraction. *Transp. Res. Part F* **2017**, *45*, 194–207. [\[CrossRef\]](#)
27. He, J.; Chaparro, A.; Nguyen, B.; Burge, R.J.; Crandall, J.; Chaparro, B.; Ni, R.; Cao, S. Texting while driving: Is speech-based text entry less risky than handheld text entry? *Accid. Anal. Prev.* **2014**, *72*, 287–295. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Chisholm, S.L.; Caird, J.K.; Lockhart, J. The effects of practice with MP3 players on driving performance. *Accid. Anal. Prev.* **2008**, *40*, 704–713. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Peng, Y.; Boyle, L.N.; Hallmark, S.L. Driver's lane keeping ability with eyes off road: Insights from a naturalistic study. *Accid. Anal. Prev.* **2012**, *50*, 628–634. [\[CrossRef\]](#)
30. Shinar, D.; Tractinsky, N.; Compton, R. Effects of practice, age, and task demands, on interference from a phone task while driving. *Accid. Anal. Prev.* **2005**, *37*, 315–326. [\[CrossRef\]](#)
31. Alosco, M.L.; Spitznagel, M.B.; Fischer, K.H.; Miller, L.A.; Pillai, V.; Hughes, J.; Gunstad, J. Both texting and eating are associated with impaired simulated driving performance. *Traffic Inj. Prev.* **2012**, *13*, 468–475. [\[CrossRef\]](#)
32. Crandall, J.M.; Chaparro, A. Driver distraction: Effects of text entry methods on driving performance. In Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, Boston, MA, USA, 22–26 October 2012; pp. 1693–1697. [\[CrossRef\]](#)
33. Young, K.L.; Rudin-Brown, C.M.; Patten, C.; Ceci, R.; Lenné, M.G. Effects of phone type on driving and eye glance behaviour while text-messaging. *Saf. Sci.* **2014**, *68*, 47–54. [\[CrossRef\]](#)
34. He, J.; Chaparro, A.; Wu, X.; Crandall, J.; Ellis, J. Mutual interferences of driving and texting performance. *Comput. Hum. Behav.* **2015**, *52*, 115–123. [\[CrossRef\]](#)
35. Choudhary, P.; Velaga, N.R. Analysis of vehicle-based lateral performance measures during distracted driving due to phone use. *Transp. Res. Part F* **2017**, *44*, 120–133. [\[CrossRef\]](#)
36. Wickens, C.D. Processing resources in attention. In *Varieties of Attention*; New York Academic Press: New York, NY, USA, 1984; pp. 63–101.
37. Mattes, S. The lane-change-task as a tool for driver distraction evaluation. In *Quality of Work and Products in Enterprises of the Future*; Strasser, H., Kluth, R., Rausch, H., Bubb, H., Eds.; Ergonomia Verlag: Stuttgart, Germany, 2003; pp. 57–60.
38. Burge, R.; Chaparro, A. The effects of texting and driving on hazard perception. In Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, Boston, MA, USA, 22–26 October 2012; pp. 715–719. [\[CrossRef\]](#)

-
39. International Organization for Standardization. *Road Vehicles—Ergonomic Aspects of Transport Information and Control Systems—Simulated Lane Change Test to Assess in-Vehicle Secondary Task Demand (ISO 26022:2010)*; International Organization for Standardization: Geneva, Switzerland, 2010.
 40. Minin, L.; Benedetto, S.; Pedrotti, M.; Re, A.; Tesauri, F. Measuring the effects of visual demand on lateral deviation: A comparison among driver's performance indicators. *Appl. Ergon.* **2012**, *43*, 486–492. [[CrossRef](#)]
 41. Tattegrain, H.; Bruyas, M.P.; Karmann, N. Comparison between adaptive and basic model metrics in lane change test to assess in-vehicle secondary task demand. In Proceedings of the 21st International Technical Conference on the Enhanced Safety of Vehicles, Stuttgart, Germany, 15–18 June 2009.
 42. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv. Psychol.* **1988**, *52*, 139–183. [[CrossRef](#)]
 43. Hart, S.G. NASA-task load index (NASA-TLX): 20 years later. In Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, San Francisco, CA, USA, 16–20 October 2006; pp. 904–908. [[CrossRef](#)]
 44. Girden, E.R. *ANOVA: Repeated Measures*; Sage: Newbury Park, CA, USA, 1991.
 45. Engström, J.; Markkula, G. Effects of visual and cognitive distraction on lane change test performance. In Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Washington, DC, USA, 9–12 September 2007; pp. 199–205.
 46. Rumschlag, G.; Palumbo, T.; Martin, A.; Head, D.; George, R.; Commissaris, R.L. The effects of texting on driving performance in a driving simulator: The influence of driver age. *Accid. Anal. Prev.* **2015**, *74*, 145–149. [[CrossRef](#)]
 47. Klauer, S.G.; Dingus, T.A.; Neale, V.L.; Sudweeks, J.D.; Ramsay, D.J. *The Impact of Driver Inattention on Near-Crash/Crash-Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data*; National Highway Traffic Safety Administration: Washington, DC, USA, 2006.
 48. Lee, J.D.; Caven, B.; Haake, S.; Brown, T.L. Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Hum. Factors* **2001**, *43*, 631–640. [[CrossRef](#)]
 49. Li, L.; Shults, R.A.; Andridge, R.R.; Yellmann, M.A.; Xiang, H.; Zhu, M. Texting/emailing while driving among high school students in 35 states, United States, 2015. *J. Adolesc. Health* **2018**, *63*, 701–708. [[CrossRef](#)]