UTILITY OF THE LANE CHANGE TEST IN EXPLORING THE EFFECTS ON DRIVING PERFORMANCE OF ENGAGING IN ADDITIONAL IN-VEHICLE TASKS WHILE DRIVING

Eve Mitsopoulos-Rubens, Kristie L. Young, Michael G. Lenné
Monash University Accident Research Centre
Building 70, Monash University VIC 3800, Australia
eve.mitsopoulos@muarc.monash.edu.au

ABSTRACT: The aim of this paper is to discuss the utility of the Lane Change Test (LCT) in exploring the effects on driving performance of engaging in additional (secondary) tasks while driving, and to provide insights into some of the factors which need to be considered in interpreting the findings and making comparisons across studies. Three studies, each of which involved the LCT, are compared. Across studies, there was no difference in mean lane deviation in the baseline condition. Within studies, mean deviation was significantly higher than the baseline when the LCT was performed in conjunction with certain secondary tasks. Across studies, however, there were instances where seemingly similar tasks gave disparate findings. While the findings provide support for the LCT, they also emphasise the need to interpret LCT results in the context of the characteristics of (e.g. type, difficulty, rate), and performance on, the secondary task.

1 INTRODUCTION

The proliferation of in-vehicle information systems (IVIS) has prompted the need for research into measuring the effects of these devices on driving performance and into understanding the mechanisms underlying these effects. The Lane Change Test (LCT) is a low fidelity driving simulation that is intended to offer a common methodological platform for measuring any decrements in driving performance associated with concurrent performance of an additional task while driving [1]. The LCT requires that participants change lanes when instructed by road-side signs. The principal measure of the LCT is mean lane deviation or MDEV – the extent to which lane change performance deviates from a normative path. The greater the deviation, the more impaired the driver's performance.

From a pragmatic perspective, the LCT's appeal is in its low cost, ease of implementation and administration, and efficiency from the point-of-view of both data collection and extraction. From an applied perspective, its ultimate appeal lies in its development as a standardised tool, thus providing the basis for making comparisons across test sites. However, discrepancies across test sites in MDEV values for baseline LCT trials have been identified (e.g. [2]), thus calling into question the ultimate utility of the LCT. It follows that explaining the source of these discrepancies across studies, and adjusting for their occurrence, is important for maximising the value of the test.

A review of six published studies from among those to have used the LCT to date [2-7] revealed differences across studies in the pattern of MDEV findings across baseline and dual-task conditions, even when seemingly similar secondary tasks had been used. However, the review also uncovered several inconsistencies across studies in the methods that were adopted (e.g. participant characteristics, quantity and quality of LCT practice, data screening procedures). Of interest here are differences that relate to the characteristics of the secondary task itself including and extending beyond the high-level distinction of task type between visual-manual and audio-verbal (cognitive) tasks.

Within each task type, IVIS, such as navigation systems and music players, have been investigated as well as surrogate versions that are intended to mimic the demand of actual IVIS tasks. The relative difficulty of a given task has also been manipulated. However, in making comparisons across studies, characteristics such as the rate at which task items are presented, may also constitute important considerations. Moreover, differences exist across studies as to whether baseline secondary task data are collected and whether secondary task performance data are even analysed. These are important omissions: without such information it is not possible to ascertain whether, in the dual-task setting, performance on one task was sacrificed to preserve performance on the other. That is, it is important that the effects on driving performance associated with in-vehicle device use while driving are discussed in the context of any companion effects on secondary task performance.

The purpose of this paper is to compare three studies, each of which utilised the LCT to assess the effects on driving performance of engaging in in-vehicle activities while driving. The overall objective is to explore LCT utility and to provide further insight into some of the factors concerning the secondary task that ought to be considered when interpreting study findings. The three studies were conducted in the same laboratory with common experimental design and equipment. A key difference between the studies was in certain characteristics of the secondary tasks under examination. In Study 1, participants performed "easy" and "hard" versions of surrogate IVIS tasks: a visual-manual task and an audio-verbal (cognitive) task. In Study 2, participants performed music selection tasks with each of three different IVIS (visual) information presentation concepts. Music selection involved scrolling through the information on screen by manipulating a rotary control. Study 2 also included a condition in which participants performed a cognitive task, which was similar to the "hard" version of the cognitive task used in Study 1. In Study 3, participants performed music selection tasks using an IVIS with touch screen interface that provided different types of feedback to users accompanying a touch screen press. The same task lists for music selection were used in Study 3 as were used in Study 2. Given the purpose of the paper, the studies are discussed concurrently.

2 METHOD

2.1 Design

Each study had a repeated measures design. For the driving task, there was one independent variable (condition), which comprised one baseline (driving only) and several dual-task levels. In Study 1 these were: visual-easy, visual-hard, cognitive-easy, and cognitive-hard. In Study 2 the dual-task levels were: concept (Concept A, Concept B, Concept C), and cognitive. The dual-task levels in Study 3 were four feedback types: none, audio, visual, and audio plus visual.

For most of the secondary tasks, there were two independent variables. For music selection (Studies 2 and 3) these were condition (baseline, dual-task) and concept (Study 2) and feedback type (Study 3). In Study 1, for both the visual task and the cognitive task, the independent variables were condition (baseline, dual-task) and task difficulty (easy, difficult). The cognitive task in Study 2 was associated with only one independent variable: condition (baseline, dual-task).

2.2 Participants

Study 1 participants (14 males, 11 females) were aged between 21 and 31 years (M 24.3, SD 3.0). They had been licenced for at least three years (M 6.1, SD 3.5). Study 2 participants (16 males, 14 females) were aged between 24 and 55 years (M 31.5, SD 8.1) and had held a car driver's licence for at least five years (M 11.5, SD 7.7). Participants in Study 3 were nine males and 12 females who were aged between 22 and 54 years (M 28.2, SD 7.4). Participants had held their car driver's licence for at least three years (M 10.3, SD 7.4). Within each study there was no significant difference in the proportion of males relative to females (all $p > 0.05$). Across studies, all participants were fully licenced car drivers.

2.3 Tasks and Equipment

2.3.1 Driving Task

The LCT was implemented on a desktop PC with a 17 inch monitor and a gaming steering wheel (25 cm diameter) with accelerator and brake pedals. The simulation is a 3 km straight section of road with three lanes of travel. Participants' task is to respond to each of the 18 signs, which are positioned approximately 150 m apart on both sides of the roadway. Each sign conveys the lane into which participants must change. Participants are instructed to make the lane changes as quickly and decisively as they can from the point where the information on the signs first appears. Participants are also instructed to maintain their speed at 60 km/h. Each LCT trial lasts approximately 3 minutes.

2.3.2 Secondary Tasks

Visual Task (Study 1)

The Surrogate Reference Task (SuRT) was used in Study 1 at two levels of difficulty (easy, hard). Participants' task is to search a display for a large circle (target) from among a number of smaller circles (distractors), and then to select the portion of the display containing the target. Difficulty is manipulated by varying the number of distractors and the size of distractors relative to the target. Item presentation was self-paced – that is, following completion of one item the next item was presented. The SuRT was presented on a laptop with 14 inch monitor, which was positioned on the desktop to the left, and within easy reach, of the driver. Response times were recorded automatically by the SuRT software.

Cognitive Task (Studies 1 and 2)

The cognitive task used in Study 1 was a mental arithmetic task at two levels of difficulty (easy, hard). The hard version was also used in Study 2. Each item was an audio file of a two-digit number (e.g. 24, 46). Item administration was controlled through DirectRT software (v2006, Empirisoft) and the items were conveyed through a headset with microphone. In the easy version, participants' task was to add "5" to every number that they heard and to vocalise their response. In the difficult version, participants added "7" to every number.

Item presentation rate differed across the studies. In Study 1, items were presented in a self-paced manner. Once a response had been given, the next item was presented. In Study 2, items were presented in a quasi self-paced manner: once an item had finished playing, participants had four seconds within which to respond. Participants who did not respond to a given item within this response window were presented with the next item after a further one second. Participants who responded within the response window were presented with the next item one second following the start of their response. During this extra second, participants' responses were automatically recorded to be checked later for accuracy. Item response times were recorded automatically by the DirectRT software. A measure of accuracy was available in Study 2 only.

Music Selection Task (Studies 2 and 3)

Three display concepts for music selection were compared in Study 2. In each case, music information was arranged as a hierarchical menu: selectable options became increasingly restricted and refined as participants progressed through the menu structure. The critical difference between the three concepts was the way in which information was presented on the display. In Concept A, information was provided through a standard list. In Concept B, the presentation of items was based on the fisheye concept. Concept C was based on the cover flow concept. In every case, scrolling through menu items was achieved by turning the main rotary control. A single press of the main rotary control selected the highlighted option and returned the list in the next level down the hierarchy. The lowest level of the hierarchy was a song list.

Concept A was used in Study 3; however, in contrast with Study 2, menu item scrolling and selection was achieved through interaction with a touch screen. With a touch screen press, participants were presented with feedback. The type of feedback (none, audio, visual, audio plus visual) depended on the experimental condition. In the audio condition, participants heard a "click". In the visual condition, the "button" on the screen was highlighted. The audio plus visual condition was a combination of the audio and visual feedback conditions.

In both Studies 2 and 3, the visual display and control (Study 2) and touch screen display (Study 3) were positioned to the left of the driver and in the same approximate location relative to the steering wheel as in a current model of a popular Australian passenger vehicle. Across studies, the size of the visual display was 15 cm (width) x 8 cm (height).

The same music selection task was used in Studies 2 and 3. Each item comprised an instruction directing participants to select a song. Participants' task was to navigate through the menu structure using the rotary control (Study 2) or touch screen (Study 3) until the required song had been selected. Item administration was controlled through DirectRT software and the items were presented aurally via a headset. Item presentation was self-paced (although with a delay): for a given item, once the required song had been selected, the next item instruction was presented following a delay of three seconds. Item completion times were recorded automatically by the DirectRT software.

2.4 Procedure

The same procedure was adopted across studies. Participants first completed several training and practice exercises to familiarise themselves with the tasks and equipment. This included two practice runs of the LCT. The experimental trials followed, commencing with a baseline LCT trial. Then, participants were exposed to the dual-task conditions in counterbalanced order. For a given condition, participants completed the secondary task on its own. In turn, participants completed the secondary task with the LCT. Participants were asked to prioritise the driving task. Finally, participants completed a second LCT baseline trial. For analysis, LCT performance was averaged across the baseline trials.

3 RESULTS

For each dependent measure, the final data were analysed using analysis of variance (ANOVA) or t-tests. Statistical significance was defined as $p \leq 0.05$. Significant ANOVA results were followed with post-hoc t-tests. Bonferroni adjustment was applied to control for potential Type I errors due to multiple comparisons.

3.1 Driving Task

The dependent measure was MDEV. In preparation for analysis, LCT speed data were scrutinised to ensure that participants had performed the LCT properly. Despite instructions to maintain speed at 60 km/h, participants may have, nonetheless, reduced their speed to compensate for increases in demand associated with concurrent completion while driving of a secondary task – an effect which could vary depending on the nature of the secondary task. An implication is that any reductions in MDEV under dual-task conditions relative to the baseline may not be as pronounced as if the LCT had been performed with speed maintained at 60 km/h. In Study 1, the MDEV data of six participants were excluded from analysis because their mean speed in at least one condition was below 50 km/h. The remaining data set comprised 98% of LCT trials where the mean speed was above 55 km/h (M 59.51, SD 1.2). In Study 2, no data were excluded, as 99% of LCT trials were carried out at a minimum mean speed of 55 km/h (M 59.97, SD 1.0) and there were no cases below 50 km/h. In Study 3, the MDEV data of four participants were omitted because their mean speed in at least one condition was below 50 km/h. The remaining data set consisted of 97% of cases where the mean speed was above 55 km/h (M 59.73, SD 0.3).

In turn, the MDEV data within each study were screened for outliers and normality. There were no significant departures from normality. Any outliers were omitted. As recommended by Mattes and Hallén [1] MDEV outliers were defined as having a standard score exceeding ±2.0. Outlier omission resulted in the exclusion of data from an extra two participants in Study 1 and one additional participant in Study 3.

For each study, mean MDEV is given in Figure 1 as a function of condition. Critically, there was no significant difference in baseline MDEV values across studies ($F(2,60) = 0.38$, $p = 0.69$). Within each study a significant effect of condition was found: Study 1 ($F(4,64) = 5.66$, $p < 0.01$), Study 2 ($F(4,116) = 26.12$, $p < 0.001$), and Study 3 ($F(4,60) = 10.21$, $p < 0.001$). In Study 1, baseline mean MDEV was found to be significantly less than that in the dual-task visual-hard and also in the dual-task cognitive-hard conditions (both $p < 0.005$). In Study 2, a significantly smaller mean MDEV was observed in the baseline condition and in the cognitive condition than in each of the three music selection conditions (all $p < 0.005$). In Study 3, mean MDEV was found to be significantly smaller in the baseline condition than in each of the dual-task feedback conditions (all $p < 0.005$).

.

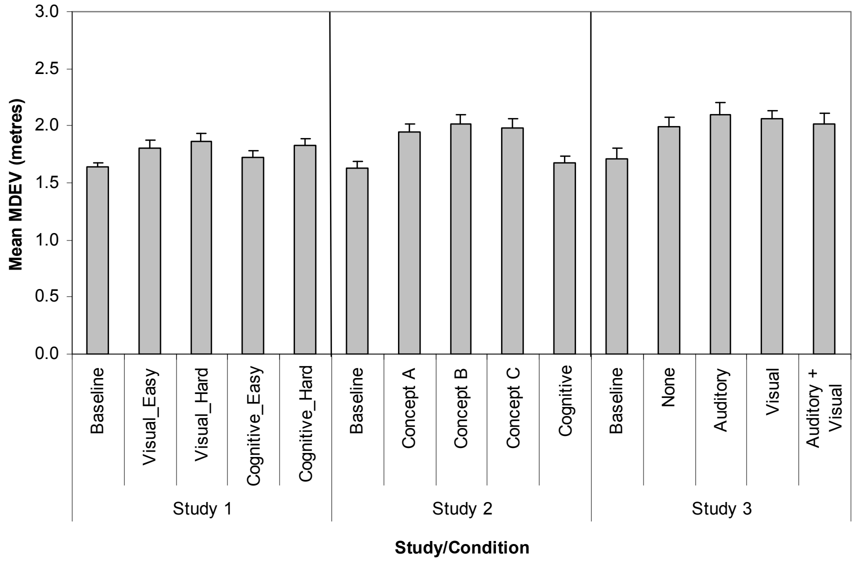

Fig.1. Mean MDEV for the baseline and dual-task conditions for each of Studies 1, 2 and 3 (Error bars represent +SEM)

3.2 Secondary Tasks

As with MDEV, all secondary task dependent measures were screened for outliers and normality prior to analysis. Significant departures from normality were treated through a transformation of the data. Any outliers were omitted.

3.2.1 Visual Task (Study 1)

Mean response time on each of the easy and hard versions of the SuRT is given in Figure 2 for each of the baseline and dual-task conditions. There was a significant effect of condition ($F(1,23) = 16.15$, $p < 0.01$), reflecting that participants took longer to complete the SuRT while also driving than under baseline conditions. There was also a significant effect of difficulty ($F(1,23) = 76.69$, $p < 0.001$): participants required more time to complete the hard version of the SuRT than the easy version. There was no condition x difficulty interaction.

3.2.2 Cognitive Task (Studies 1 and 2)

Descriptive response time data on the easy and hard versions of the cognitive task used in Study 1 are presented in Figure 2 for each of the baseline and dual-task conditions. Mean response time on the cognitive task did not differ significantly between the dual-task and baseline conditions ($F(1,22) = 1.50$, $p = 0.23$). This effect was irrespective of task difficulty ($F(1,22) = 0.05$, $p = 0.83$). There was, however, a significant main effect of difficulty ($F(1,22) = 89.19$, $p < 0.001$), with participants taking significantly more time to respond to items on the hard version of the cognitive task than on the easy version.

A different pattern of outcomes regarding response time was found for the cognitive task in Study 2. Figure 3 presents mean response time and mean accuracy data for the baseline and dual-task conditions of the task. Participants took significantly more time to respond to items in the dual-task than in the baseline condition (t (29) = 2.01, p = 0.05). However, mean accuracy did not differ significantly between the conditions (t (28) = 0.78, p = 0.44). This suggests that there was no speed-accuracy trade-off in cognitive task performance. That is, a significant decrease in response time from the baseline to the dual-task condition was not offset by a significant increase in accuracy from the baseline to the dual-task condition.

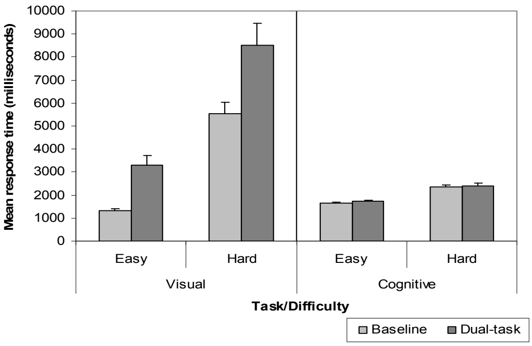

Fig 1: Mean response times on each of the visual and cognitive tasks in Study 1 as a function of difficulty and condition (Error bars represent +SEM)

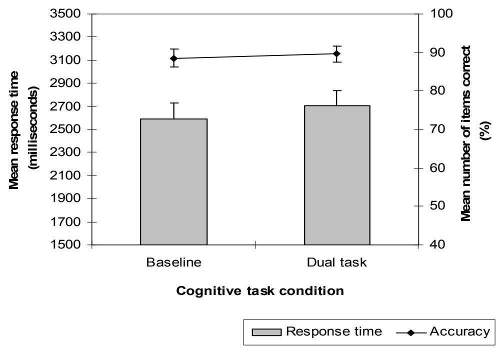

Fig 2: Mean response time and mean accuracy under baseline and dual-task conditions of the Study 2 cognitive task (Error bars represent ±SEM)

3.2.3 Music Selection Task (Studies 2 and 3)

Mean completion time is presented in Figure 4. These data are presented for each of the baseline and dual-task conditions as a function of concept (Study 2) and feedback type (Study 3). As significant departures from normality were present in the Study 2 data, these data were transformed (square root) in preparation for analysis.

There was a significant main effect of condition in Study 2 (F (1,26) = 110.18, p < 0.001) and in Study 3 (F (1,18) = 38.59, p < 0.001). Participants required more time to perform the music selection task under dual-task than under baseline conditions. In Study 2, there was also a significant main effect of concept (F (2,52) = 12.79, p < 0.01), and a trend towards a significant condition x concept interaction (F (2,52) = 2.66, p = 0.08). Under both baseline and dual-task conditions, music selection took significantly longer with Concept B than with Concept A. Also, under dual-task conditions, completion of the music selection task took significantly longer with Concept B than with Concept C (all p < 0.008). In Study 3, there was neither a significant feedback type main effect (F(3,54) = 1.74, p = 0.17) nor condition x feedback type interaction (F (3,54) = 1.97, p = 0.13).

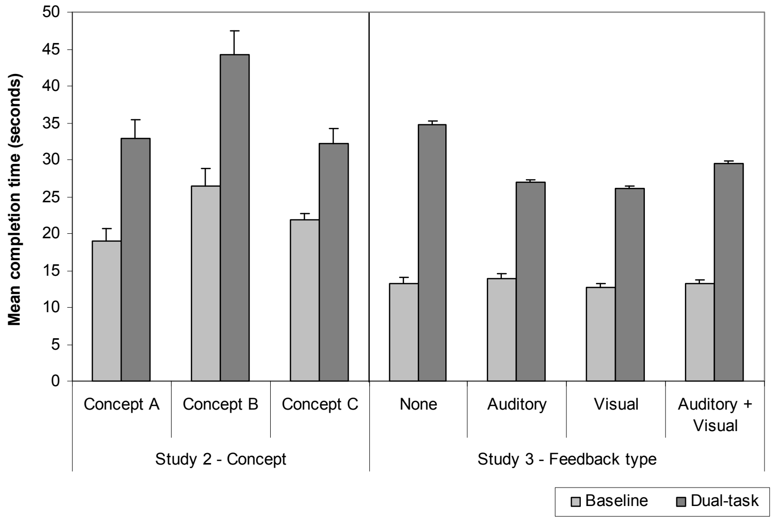

Fig.4. Mean time taken to complete music selection items under baseline and dual-task conditions in each of Studies 2 and 3 (Error bars represent +SEM)

4 DISCUSSION

The LCT was effective in highlighting impairments in driving performance associated with engaging in certain secondary tasks while driving. Across studies, relative to the baseline, a significant increase in MDEV was observed when the LCT was performed in conjunction with the visual-hard version of the SuRT (Study 1), the music selection task, irrespective of concept (Study 2) and feedback type (Study 3), and the cognitive-hard task (Study 1).

With the exception of the cognitive-hard task in Study 1, for those dual-task conditions that were associated with impaired MDEV performance (i.e. Study 1, visual-hard; Study 2, all concepts; Study 3, all feedback types), performance decrements from the baseline to the dual-task condition were also observed on the secondary tasks. Thus, in these cases at least, LCT performance (at least in terms of MDEV) was not sacrificed so that performance on the secondary task while driving could be preserved at baseline levels.

While participants took significantly longer to respond to items on the visual-easy task (Study 1) under dual-task than under baseline conditions, this effect was not accompanied by significant impairment in MDEV from the baseline to the dual-task scenario. In the case of the cognitive-easy task (Study 1) there was neither a significant impairment in MDEV nor in mean response time from the baseline to the dual-task setting. The implication of these findings is that the more demanding the secondary task (in terms of increasing competition for shared attentional resources), the more likely it is to result in impairment on the secondary task in the first instance and, ultimately, on both the driving and secondary tasks. However, this is based on the assumption that the driver prioritises the driving task.

Perhaps the most intriguing finding to emerge from this comparative exercise is the discrepancy between Studies 1 and 2 on the effect of the cognitive-hard task on MDEV. While in Study 2 there was no significant impairment in MDEV, in Study 1, a significant decrement in LCT performance was observed. A significant increase in response time from the baseline to the dual-task condition was found for the cognitive task in Study 2 - an effect which was not the result of a speed-accuracy trade-off. The effect on response time in the absence of an effect on MDEV suggests that participants were prioritising the driving task and, in so doing, were sacrificing performance on the cognitive task. In Study 1, participants' response times on the cognitive-hard task did not deteriorate under dual-task conditions. As accuracy data were not captured in Study 1, it is not known whether participants sacrificed accuracy to maintain speed. While this is a possibility, a further possibility is that participants sacrificed LCT performance to maintain performance on the cognitive-hard task at baseline levels. The primary difference between the cognitive-difficult task in Study 1 and the cognitive task in Study 2 was the rate at which items were presented. In Study 1, the task items were presented in a self-paced manner, while in Study 2 item administration was quasi self-paced. That is, despite being self-paced, participants had a maximum time window in which to respond before the next item was presented. This difference in presentation rate, coupled with the nature of the task (i.e. cognitive, relatively difficult) may have contributed to the disparate findings observed on the cognitive task across the two studies. In summary, the studies reported here provide support for the LCT as a common methodology for exploring the distracting potential of a range of secondary, in-vehicle tasks. Moreover, they highlight the need to collect and analyse secondary task performance data, both under baseline and dual-task conditions, and to interpret LCT findings in the context of the characteristics of, and performance on, the secondary task.

5 ACKNOWLEDGEMENTS

The research reported in this paper was funded through the Cooperative Research Centre for Advanced Automotive Technology (AutoCRC). We thank Amy Williamson, Margaret Trotter and Jordan Navarro for their assistance with data collection.

6 REFERENCES

1. Mattes, S., & Hallén, A. 'Surrogate distraction measurement techniques: The Lane Change Test', in Regan, M.A., Lee, J.D., & Young, K.L. (Eds.), 'Driver distraction: Theory, effects and mitigation' (CRC Press, 2008), pp. 107-122.

2. Petzoldt, T., Bär, N., & Krems, J.F. 'Gender effects on Lane Change Test (LCT) performance'. Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Big Sky, Montana, USA, June 2009, pp. 90-96.

3. Burns, P.C., Trbovich, P.L., McCurdie, T., & Harbluk, J.L. 'Measuring distraction: Task duration and the lane-change test (LCT). Proceedings of

the HFES 49th Annual Meeting, Orlando, Florida, USA, September 2005, pp. 1980-1983.

4. 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'. Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Stevenson, Washington, USA, July 2007, pp. 16-22.

5. Harbluk, J.L., Mitroi, J.S., & Burns, P.C. 'Three navigation systems with three tasks: Using the lane-change test (LCT) to assess distraction demand'. Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Big Sky, Montana, USA, June 2009, pp.24-30.

6. Maciej, J., & Vollrath, M. 'Comparison of manual vs. speech-based interaction with in-vehicle information systems', Accident Analysis and Prevention, 2009, 41, pp.924-930.

7. Wilschut, E.S., Rinkenauer, G., Brookhuis, K.A., & Falkenstein, M. 'Effects of visual search task complexity on lane change task performance'. Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, Lyon, France, April 2008, pp. 23-32.