Theoretical and Methodological Issues in Driver Distraction

Dissertation

zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der Fakultät für Human- und

Sozialwissenschaften der Technischen Universität Chemnitz

am 24. Mai 2011

von Tibor Petzoldt, geboren am 30.11.1981 in Schlema

Gutachter: Prof. Dr. Josef F. Krems & Prof. Dr. Peter Sedlmeier

CHEMNITZ UNIVERSITY
OF TECHNOLOGY

Zusammenfassung

Fahrerablenkung ist ein Begriff, der in den vergangen Jahren verstärkt in das Blickfeld der Öffentlichkeit geraten ist. Dies ist im Wesentlichen zurückzuführen auf die deutlich steigende Verbreitung und Nutzung von Fahrerinformationssystemen. Gleichzeitig führt die steigende Automatisierung im Fahrzeug dazu, dass dem Fahrer in seiner subjektiven Wahrnehmung mehr Ressourcen zur Verfügung stehen, um sich anderen Aktivitäten wie etwa Essen, Rauchen oder Telefonieren zuzuwenden. Die steigende Aktualität dieser Problematik wirft viele Fragen auf. Wie häufig tritt Fahrerablenkung auf? Welche Konsequenzen hat sie? Welche kognitiven Prozesse zeichnen für diese Konsequenzen verantwortlich? Und wie kann man Fahrerablenkung messen?

Die vorliegende Dissertation besteht aus drei empirischen Beiträgen, sowie einer kurzen Einführung, die die grundlegenden Fragen und Befunde zum Thema Fahrerablenkung betrachtet. Das Augenmerk des ersten Beitrags liegt auf der Überprüfung theoretischer Annahmen zur Fahrerablenkung. Eine Vielzahl von Untersuchungen zeigt, dass sich kognitiv beanspruchende Zweitaufgaben negativ auf die Fahrleistung auswirken. Im vorliegenden Beitrag wird davon ausgegangen, dass dieser Effekt eine Folge von Interferenzen zwischen den Funktionen des Arbeitsgedächtnisses, die dazu dienen das Situationsmodell der Verkehrssituation aktuell zu halten, und den bearbeiteten Zweitaufgaben ist. Im Rahmen einer Simulatorstudie wurde diese Annahme überprüft. Es zeigte sich, dass die Probanden, die eine Zweitaufgabe ausführten, die speziell die Integration von neuen Informationen in das bestehende Situationsmodell behindern sollte, später auf antizipierbare kritische Ereignisse reagierten als Vergleichsgruppen. Im Gegensatz dazu ergaben sich für unvorhersehbare Ereignisse keine Unterschiede. Diese Ergebnisse weisen darauf hin, dass die negativen Effekte kognitiver Belastung tatsächlich auf Interferenzen mit spezifischen Arbeitsgedächtnisprozessen zurückzuführen sind.

Die beiden weiteren Beiträge befassen sich mit messmethodischen Fragen in Bezug auf Fahrerablenkung. In Beitrag zwei wird die Lane Change Task (LCT) thematisiert, eine Labormethode zur Erfassung von Ablenkung. Aufgabe der Probanden ist die Steuerung eines virtuellen Fahrzeuges mittels Lenkrad, und dabei konkret die Ausführung von Spurwechseln, bei gleichzeitiger Bearbeitung von Zweitaufgaben. Trotz eines standardisierten Versuchsaufbaus sind allerdings starke Messvarianzen zwischen verschiedenen Testreihen zu beobachten. Der Übungsgrad der Versuchsteilnehmer wurde dabei als eine mögliche Ursache identifiziert. In zwei Experimenten wurde dieser Vermutung

nachgegangen. Probanden bearbeiteten parallel zur LCT Zweitaufgaben verschiedener Schwierigkeitsstufen, nachdem sie zuvor trainiert wurden. Es konnte gezeigt werden, dass der Grad der Übung tatsächlich einen Einfluss auf die Spurwechselperformanz hat, und dass dieser Einfluss auch Monate später noch zu finden ist. Es ist jedoch zweifelhaft, dass dieser Effekt allein ursächlich für die zu beobachtenden Messvarianzen ist.

Im dritten Beitrag wird die Critical Tracking Task (CTT) betrachtet, ein Verfahren, das im Kontext Fahrerablenkung bisher kaum Beachtung fand. Die CTT ist eine einfache Trackingaufgabe, welche vom Nutzer die Stabilisierung eines dynamischen, instabilen Elementes auf einem Bildschirm fordert. Die zur Bearbeitung der Aufgabe auszuführenden Tätigkeiten der kontinuierlichen visuellen Überwachung und manuellen Kontrolle sind grundsätzlich vergleichbar mit basalen Anforderungen der Fahraufgabe. Ziel war es, das Potenzial der CTT als Messverfahren von Fahrerablenkung durch Fahrerinformationssysteme zu überprüfen. Die Ergebnisse der vier durchgeführten Experimente, in denen sowohl künstliche als auch reale Aufgaben und Systeme bearbeitet und bedient wurden, legen den Schluss nahe, dass die CTT in der Tat in der Lage ist, das Ausmaß von Ablenkung ausgelöst durch Fahrerinformationssysteme zu quantifizieren.

Contents

A few general words ...

This dissertation is composed of three manuscripts that have been submitted for publication, complemented by an introduction that shall help to understand the broader background for each of the manuscripts. One of them (paper two) has already been published, the others are currently under revision (paper one) or under review (paper three). Since all three manuscripts have been submitted to the same journal (*Transportation Research Part F: Traffic Psychology and Behaviour*), major formatting aspects (citations and references, numbering of subsections) follow the guidelines of this journal. The introduction has been formatted accordingly to provide a coherent appearance of the dissertation. In other aspects (font, alignment of text, position of figures and tables within the text) I deviated from the classical manuscript format, because I believe that this allows for a better accessibility of the document. Yet, the attentive reader will notice that the style of figures differs slightly between the manuscripts, and probably also the style of language. As this dissertation (and the respective manuscripts) evolved over the course of five years (and not necessarily in parallel), this was inevitable. Still, I hope that the final result will please the reader, or at least provide him / her with the information that he / she is looking for.

... and some thank you’s

First of all, I have to thank my family for the support and the encouragement over the years. They always made me feel that although all this might be a big step, it would be an easy one for me. Special credits go to my mom, who, a psychologist herself, somehow got me into all this. I still have not been able to figure out how she did it.

Many thanks to my collaborators Martin, Jeroen, Nina, Claudia and Hanna, who deserve credit for creating ideas, producing input, giving feedback, you name it.

I also want to thank those who directly contributed to the work described in the three manuscripts - by programming the simulator, generating material, sitting in the lab, coding video or pre-processing data - Ingmar, Chantal, Anne-Katrin, Sylvia, Sophia, Antje, Nadja, Sandy, Mirjam, Lukas, Lena and Sebastian.

Thanks as well to Prof. Josef Krems, who supervised this dissertation and played a key role not only in the work presented here, but also in my overall professional experience, be it teaching or research. Thanks to Prof. Peter Sedlmeier for agreeing to review what I have produced.

Finally - and I do not dare to give any names here for fear of leaving someone out who deserves to be mentioned - you know who you are - a big thanks to my colleagues and friends, who, at some stage or another, in one way or another, had an impact on the work presented here, and on pretty much everything else I did in the past five years.

Driver attending to cell phone hits police car

A man was apparently texting on his cell phone Tuesday when he ran into the back of a stopped St. Joseph police patrol car, knocked it into another squad car and sent two police officers to the hospital, police said. The accident occurred after police stopped southbound traffic on Interstate 29 to work a one-car accident. The man sped around stopped semi-trucks and ran into the patrol car, police said. He was arrested on an unrelated warrant. The officers suffered minor injuries.

(The Kansas City Star, May 20th, 2009)

Introduction

Driver distraction is a buzzword that has received a lot of attention inside, but also outside the scientific world recently. It has been selected word of the year 2009 by Webster's dictionary ("Distracted-driving campaign", 2010). The US Department of Transportation (2009) has planned to crack down on driver distraction, as the US Transportation Secretary announced an "administration wide effort to combat distracted driving". Even the news media caught on, and covered especially extreme cases of driver distraction with headlines like "Man charged for watching porn while driving" (2009), "Driving on razor's edge: Shaving behind the wheel" (2009) or "Police: Driver distracted by sex toy" (2010). With respect to those cases, the definition that the delegates of the International Conference on Distracted Driving agreed upon reads rather dry: "Distraction involves a diversion of attention from driving, because the driver is temporarily focusing on an object, person, task, or event not related to driving, which reduces the driver's awareness, decision-making, and / or performance, leading to an increased risk of corrective actions, near-crashes, or crashes." (Hedlund, Simpson, & Mayhew, 2006; p. 2).

There are other approaches to answer the question what exactly driver distraction is (e.g. Hoel, Jaffard, & van Elslande, 2010; Lee, Regan, & Young, 2009; Pettitt, Burnett, & Stevens, 2005; Treat, 1980). According to Regan, Hallett and Gordon (in press) all those definitions agree that "there is a diversion of attention away from driving, or safe driving; attention is diverted toward a competing activity, inside or outside the vehicle, which may or may not be driving-related; the competing activity may compel or induce the driver to divert attention toward it; and there is an implicit, or explicit, assumption that safe driving is adversely effected" (p. 1). The National Highway Traffic Safety Administration (NHTSA) defined four categories of driver distraction: "visual distraction (e.g., looking away from the roadway), auditory distraction (e.g., responding to a ringing cell phone), biomechanical distraction (e.g., manually adjusting the radio volume), and cognitive distraction (e.g., being lost in thought)" (Ranney, Mazzae, Garrott, & Goodman, 2000; p. 1). For visual distraction, Ito, Uno, Atsumi and Akamatsu (2001) further differentiated into situations in which (1) the driver's visual field is blocked and prevents the perception of relevant information, (2) the driver neglects to look at relevant areas, focusing instead on another visual target, and (3) the driver is inattentive, often described as the "looked-but-failed-to-see" phenomenon (Brown, 2005). These classifications and categorisations merely serve as examples here - there are certainly numerous others out there that might apply as well. The single aspect that renders driver distraction an important field to study is the notion that distraction has an effect on driving safety (Regan et al., in press). That alone is reason enough for researchers to engage in the investigation of driver distraction, which often results in rather applied research questions, sometimes lacking theoretical background and scientific rigour. However, many now also try approach driver distraction from a theoretical perspective by including basic psychological concepts in their studies. Ultimately, only the integration of applied and basic approaches will help to paint a complete picture of driver distraction.

The following pages shall give a very short overview on relevant issues that surround driver distraction. This overview is by no means comprehensive, it rather serves as a brief introduction for the three scientific papers that form this dissertation. Still, the most important questions will be addressed. To what extend does driver distraction occur? What are the consequences? What cognitive processes trigger those consequences? And how can distraction be measured? The answers to these questions are supposed to help understand the issues and questions raised in the three subsequent papers.

1 Prevalence of driver distraction

One of the most important questions is to what extent drivers actually engage in activities that might cause distraction. Most assessments focussed on mobile phone use, which appears to be a dominant matter of investigation in general. In a recent survey, German testing authority DEKRA (2010) found that about 22% of the drivers asked are using a handheld phone while driving, which is illegal in Germany. Questioned about the reason for not using a handsfree-kit (which would be legal), 58% responded that they just do not care about the ban. In the US, the "2007 Motor Vehicle Occupant Safety Survey" (Boyle & Lampkin, 2008) found that only 22% of the drivers who usually have a wireless phone in the vehicle reported to never talk on the phone while driving, whereas 33% admitted to talk on at least half of all trips made. In a similar survey in Canada, 37% of those who answered reported to have used a cell phone while driving in the past seven days

(Vanlaar, Simpson, Mayhew, & Robertson, 2007). In the same study, respondents were asked to indicate how often they see a variety of behaviours happening on the road, on a scale from one (never) to six (very often). The use of cell phones, with a rating of 4.85, emerged as the behaviour observed most often. Roadside observational studies point into a similar direction. In the US, data from the "National Occupant Protection Use Survey" (NOPUS) 2009, which was collected by observing 49,475 vehicles at 1,496 data collection sites, indicates that about 5% of the observed drivers were using a handheld mobile phone (Pickrell & Ye, 2010). In addition, 0.6% were observed to text-message or manipulate a handheld device in some form. In Germany, approximately 418,000 violations of the ban on the use handheld mobile phones while driving were reported to the Central Register of Traffic Offenders in 2009 (Kraftfahrt-Bundesamt, 2010).

A broader picture beyond mobile phone use might be drawn through the assessment of naturalistic driving data. As Backer-Grøndahl, Phillips, Sagberg, Touliou and Gatscha (2009) note, "naturalistic driving observation includes unobtrusively observing normal drivers in their normal driving context while driving their own vehicles" (p. 9). In an analysis of such naturalistic driving videos, Stutts, Feaganes, Rodgman, Hamlett, Meadows, Reinfurt et al. (2003) found that, even when excluding conversing with passengers, participants still spent 14.5% of the time in their vehicle engaged in some distracting activity. Notable forms of distraction where eating or drinking (including the preparation of the actual act) with 4.6%, external (i.e. outside the car) distraction with 1.6% and smoking (including lighting and extinguishing), again 1.6%. Similar results have been reported by Sayer, Devonshire and Flanagan (2007). Whereas conversing with passengers was again the single most frequent secondary behaviour (15.3%), other activities like grooming (6.5%), cell phone use (5.3%) and eating / drinking (1.9%) were also prevalent substantially in total driving time. In the most extensive study so far, Klauer, Dingus, Neale, Sudweeks and Ramsey (2006) analysed the data of about 100 drivers that drove an instrumented vehicle for approximately one year. About 20,000 so called baseline epochs (i.e. without critical incident) of 6 sec length were analysed, with the result that in about 38% of all those epochs, the driver was engaged in a secondary task (which is not further specified).

2 Effects of driver distraction

2.1 Technological driver distraction

Just like most investigations in the prevalence of distraction have focused on the use of mobile phones, so have studies on the effects of driver distraction. Already in the 1960s, it has been found that mobile phone-like secondary tasks result in poorer driving performance (Brown, Tickner, & Simmonds, 1969). In this classic study, participants had to drive a vehicle on a test track, while solving a “telephoning task” of checking the accuracy of short sentences. During the ride, they had to judge whether to drive through gaps which might be larger or smaller than the car. It was found that the “possible / impossible” judgement of the gaps was indeed affected, a result that prompted the authors to conclude that “perception and decision-making may be critically impaired” (p. 419). From the early 1990s on, loads of studies followed (e.g. Alm & Nilsson, 1994; Briem & Hedman, 1995; Brookhuis, de Vries, & de Waard, 1991; McKnight & McKnight, 1993; Violanti & Marshall, 1996 - for a complete overview see Drews & Strayer, 2009). They basically all agree that mobile phones are a risk to safe driving that should not be underestimated. And, although legislation in many countries suggests otherwise, not only the handheld use of mobile phones is a source of that risk. Drews and Strayer (2009) concluded that “the difference between handheld and handsfree conversations is minimal and potentially negligible in terms of the accident risk” (p. 185).

Text-messaging, a function of mobile phones not foreseen by the pioneers in this research area, has become a subject of scrutiny just recently. Drews, Yazdani, Godfrey, Cooper and Strayer (2009) assessed driving performance in a simulator while text-messaging. Their analysis revealed that while text-messaging, participants responded more slowly to the onset of braking lights and showed impaired longitudinal and lateral control in comparison to a control condition. Text-messaging also led to more crashes compared to the control condition. In a similar design, Hosking, Young and Regan (2009) investigated the effect on young drivers’ driving behaviour. Just as Drews et al. (2009), they found deficits in longitudinal and lateral control when text-messaging. Additionally, they reported a significant increase in eyes-off-the-road time, as well as a higher number of missed lane changes. Owens, McLaughlin and Sudweeks (2011) made the distinction between text-messaging with a handheld device versus an integrated system. Participants had to send pre-programmed messages, as well as to read (handheld) / listen to (integrated system) incoming messages while navigating a vehicle through a closed test track. Results indicated that sending and receiving text messages with the handheld phone led to impaired steering behaviour, longer and more frequent eyes-off-the-road episodes, and a higher reported workload. The integrated system went with less performance degradation, however still differed significantly from a baseline on most measures.

Navigation systems, with their rapid increase in market penetration especially in the last two decades, are another substantial source of technological driver distraction. Tijerina, Parmer and Goodman (1998) studied the distraction effects of destination entry on different navigation systems, one among them speech-based, and compared them to the operation of a mobile phone and an audio system while driving in a simulator. The systems with visual-manual entry resulted in longer imputs, longer and more frequent glances to the systems and more lane deviations than the control tasks and the speech-based system. These results were confirmed by Srinivasan and Jovanis (1997), who compared systems with visual and auditory navigation information in a simulator. They also found clear advantages for the auditory presentation. Similar results have been reported by Dingus, McGehee, Hulse, Jahns, Manakkal, Mollenbauer, et al. (1995), Gärtner, König and Wittig (2001) and Tsimhoni, Smith and Green (2004).

In-car radios have received less attention. Not in-vehicle information systems by strict definition, but still in-vehicle technology with distraction potential, radios are installed in practically any new car sold. The operation of a radio while driving is legal and therefore usually regarded as acceptable, even to the extent that it is often used as a benchmark against which to compare other systems (McKnight & McKnight, 1991). However, there is evidence that even the radio can have detrimental effects of driving performance and related variables as well. Briem and Hedman (1995) found a stronger deterioration in driving performance for manipulating radio controls compared to conversing on a mobile phone. In a field experiment, Wikman, Nieminen and Summala (1998) recorded significantly longer glances away from the road for tuning the radio than for a mobile phone task (which involved physical manipulation of the device as well). In a simulator study (Horberry, Anderson, Regan, Triggs, & Brown, 2006), participants produced higher variations in speed when tuning the radio compared to a mobile phone task, an effect that was underlined by higher workload ratings. Indeed, in an in-depth analysis of crashes, it has been found that adjusting the radio, cassette or CD-player is a major cause of crashes related to driver distraction (Stutts, Reinfurt, Staplin, & Rodgman, 2001). Jäncke, Musial, Vogt and Kalveram (1994) even found evidence for detrimental effects when simply listen-ing to a radio broadcast. Especially in complex driving situations, there was an increase in lane deviations compared to a control group.

2.2 Non-technological driver distraction

Although the largest share of research on the effects of driver distraction has focused on technological forms of distraction, the possible effects of other in-vehicle activities should not be underestimated, especially since their prevalence is often rather high. Activities like eating and drinking, smoking or talking to passengers have obviously the potential to distract drivers. However, knowledge about the effects of these forms of distraction is scarce. In a rare example, Jenness, Lattanzio, O'Toole and Taylor (2002) used "eating cheeseburger" as a control condition in an examination of voice-activated dialling of phone numbers. They found that, compared to driving only, eating a cheeseburger resulted in significantly more driving errors, decreased driving speed, and more glances away from the road. Trying to fill the gap in the literature, Young, Mahfoud, Walker, Jenkins and Stanton (2008) conducted a simulator study in which participants were instructed to eat and drink while driving. Results showed that although driving performance variables were relatively unaffected, the reported workload was higher, and more crashes occurred in a critical situation that was included in the simulator scenario.

The situation is somewhat similar for distractions outside the vehicle. Although there appears to be consensus that drivers can be distracted by roadside advertising (Wallace, 2003a), and also evidence that distraction through environmental features is at least present in a number of accidents (Stutts et al., 2001), little research has been done to get a deeper understanding of the mechanisms and specific effects of distraction from outside of the vehicle. In his review, Wallace (2003b) listed studies that investigated the effects of billboards. He concluded that the few "experimental studies suggest that billboards and signs can function as distracters" (p. 54), however acknowledged that those experiments are rather limited in terms of ecological validity. Out of what he called "statistical studies", he summarised that there appears to be a correlational relationship between billboards and accident rates. Those studies, however, are not able to establish any causal link between the two.

3 Cognitive underpinnings of driver distraction

Crucial for a deeper understanding of why distraction occurs, what effects it has, and how it can be avoided, is the development and validation of proper models that describe the cognitive processes underlying driver distraction.

Lee et al. (2009) describe distraction "as a breakdown in a multilevel control process" (p. 41). To explain the processes behind driver distraction, they define three types of control that are critical for safe driving: feedback control, feedforward control and adaptive control. Feedback control simply is the comparison of the current state and a goal state to guide behaviour. Differences between these states should lead to respective activities to minimise this difference. In contrast, feedforward control uses an anticipated future state to guide behaviour, therefore relying heavily on appropriate mental models. Finally, adaptive control reduces the difference between current state and goal state by redefining the goal state.

To explain how those different types of control influence driving, and what effects they have if they fail, Lee et al. (2009) use Michon's (1985) classical structure of the driving task, which proposes the differentiation of strategical, tactical and operational level. These three levels differ in the subtasks that are performed as well as in their timescale. The strategical level refers to the "general planning stage of a trip, including the determination of trip goals, route, and modal choice, plus an evaluation of the costs and risks involved" (Michon, 1985; p. 489), which is assumed to occur at a timescale of minutes to weeks. At the tactical level, the driver is concerned with certain manoeuvres such as obstacle avoidance, gap acceptance, turning and overtaking, which occurs on a timescale of seconds. The operational level mainly refers to longitudinal and vertical control of the vehicle, that is accelerating, braking, and steering, which occurs on a timescale of milliseconds.

Lee et al. (2009) then try to relate the three levels of the driving task, and, more importantly, their different timescales, to the different types of control, by stating challenges for each type of control for each time horizon. Those challenges, in turn, can occur as a result of driver distraction. For example, on the operational level, feedback control can only operate properly if the time constant of driver reaction is fast enough for the driving demands. If there is some activity that might lead to an unfavourable shift in this relationship (e.g. the handheld use of a mobile phone, which results in the withdrawal of one hand from the steering wheel, and obviously leads to an increase in reaction time), the probability of an incident increases. On the same level, feedforward control might be compromised when the task demands are unpredictable or unknown (e.g. an unknown route or a new in-vehicle system, which, in an unfortunate situation, might both put demands on the driver that cannot be managed in parallel). Consequently, Lee et al. (2009) state that “distraction-related incidents occur when the demands of driving and competing activities combine to undermine control”.

The concept of situation awareness provides another approach to better understand the reasons and mechanisms behind the detrimental effects of driver distraction on driving performance. Put simple, situation awareness “is knowing what is going on around you” (Endsley, 2000; p. 4). One of the accepted scientific definitions explains situation awareness as the “continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events” (Dominguez, 1994; p. 11). This definition is in agreement with Endley’s (1995) model of situation awareness, which comprises three levels - (1) the perception of the elements in the environment within a volume of time and space, (2) the comprehension of their meaning, and (3) the projection of their status in the near future.

Following this model, errors on the first level of situation awareness are mainly a result of the natural limitations of human attention. In complex, dynamic environments, attentional demands that are caused by information overload, complex decision making and multitasking can easily exceed the attentional resources of an operator, which increases the probability of not perceiving information relevant for the safe operation of the respective system. Accordingly, research from the aviation domain confirmed that a fair amount of errors stems from problems on the first level of situation awareness (Jones & Endsley, 1996). Within this category of errors, distraction elicited by other relevant tasks is the most common source of error. An analysis of fighter aircraft accidents found that lack of attention to primary instruments (the primary task) and too much attention to target planes (another relevant task) were major causes of incidents (Kuipers, Kappers, van Holten, van Bergen, & Oosterveld, 1990). The simple “not looking”, i.e. visual distraction, appears to be a decisive factor for the emergence of errors in the operation of complex systems with visual presentation of relevant information. Transferring this knowledge to the automotive context, in which the largest share of information is perceived visually (Rockwell, 1972), the additional visual-manual operation of a technological device of any form can result in an incomplete or even incorrect situation model. Given the highly dynamic nature of the road environment, it can be assumed that even minimal changes in glance behaviour to the disadvantage of the primary driving task can result in an increase of crash risk. Accordingly, the so called “eyes-off-the-road” phenomenon (Brown, 1994; Green, 2000) appears to be a major contributing factor to traffic accidents.

However, errors are not limited to the first level of situation awareness. Once information from the environment has been perceived, further processing, i.e. interpretation of the situation (level two) and anticipation of future developments (level three) occurs. It is necessary to link the acquired information with already existing mental models of the situation at hand to respond appropriately. This processing is mainly tied to working memory (Wickens & Hollands, 1999). Fracker (1987, cited in Endsley, 1995) assumed working memory to be the main bottleneck for situation awareness. Common errors on the higher levels of situation awareness are the selection of an inappropriate model, or an inappropriate confidence in a correct model. This can be caused by overloading the limited resources of working memory, sometimes to an extend that no situation model is selected at all (Jones & Endsley, 1996). Ma and Kaber (2005) confirmed that by reducing cognitive load, situation awareness improved. The overload of cognitive resources, in turn, can be caused by cognitive distraction, as “some of the resources may not be available because they are assigned to other tasks, such as entering the destination into a navigation system or talking to the passenger. The remaining working memory capacity might be too reduced to ensure that the perceived situation elements get fully connected to the relevant knowledge in long-term memory” (Baumann & Krems, 2007; p. 259). Indeed, Ma and Kaber (2005) also reported that while the reduction of cognitive load through automation led to improvements in situation awareness, the introduction of a cell phone task resulted in higher load and diminished situation awareness especially on levels two and three. It seems that “mind-off-the-road” situations, while occurring less frequently than “eyes-off-the-road”, are still a considerable source of risk on the road.

Other concepts like the SEEV-model (Salience, Effort, Expectancy, Value; Wickens, Helleberg, Goh, Xu, & Horrey, 2001), control theory (Sheridan, 2004) or threaded cognition (Salvucci & Taatgen, 2008) have been used to describe driver distraction or certain aspects of it. Most of the approaches are doing well when it comes to explain post-hoc why a certain secondary activity had the effect observed. More important, however, is to test whether those models are able to predict specific effects of different forms of driver distraction. Paper one of this dissertation (“The effect of cognitive tasks on predicting events in traffic”) tries to do that by assessing whether specifically designed secondary tasks have the impact on driving performance that would be predicted by a modified model of situation awareness.

4 Measurement of driver distraction

One of the most important issues in driver distraction is its measurement. The assessment of systems or system prototypes with regard to their distraction potential is crucial when it comes to decide whether a certain piece of technology will be allowed to be used inside a vehicle or not.

The most direct way to assess the impact of distraction on real world driving is to measure actual driving performance. Lane keeping, lane crossing, speeding, braking etc. are just a few measures that can be acquired. On-road, test track and simulator studies are capable of producing such data. However, aside from the high cost that go with all of those methods, there are crucial methodological flaws. One major drawback is the lack of standardisation. Although test tracks and driving simulators provide a somewhat controlled environment, there are still many degrees of freedom with regard to the drivers reactions as well as to environmental factors that make it hard to interpret the obtained data. Östlund, Nilsson, Carsten, Merat, Jamson, Jamson, et al. (2004) found that the influence on steering behaviour elicited by the operation of an in-vehicle device is dependent on the load of the actual driving task, and therefore can only be interpreted for the specific driving situation. It also appears that the measures acquired are often not sensitive enough to detect differences in distraction, as differences that do not directly alter driving performance are not measurable (Jahn, Oehme, Krems, & Gelau, 2005).

Physiological measures have been used mainly to assess the cognitive load that is imposed by the operation of additional in-vehicle tasks. These measures are considered quite capable of measuring global arousal or activation, as well as rather specific stages in information processing (de Waard, 1996). However, they usually require sophisticated equipment and the respective expertise for the application as well as the interpretation (Kramer, 1991). They are also comparatively sensitive to confounding variables like physiological demand, noise or emotion, whose effects might even exceed the effects of the actual distraction (Roscoe, 1987; Wilson, 1992). Generally, it is argued that physiological measures might be more suitable for “determining long-term states of the driver, such as fatigue, rather than specific reactions to particular signals." (Kantowitz in Llaneras, 2000; p. 58).

Subjective measures also mainly focus on the additional demand that is created by distracting activities. They are easy to administer, but disputed in their actual value. O'Donnell and Eggemeier (1986) list four major limitations to the use of subjective measures of load. First, there is evidence for the confoundation of physical and mental load. This is especially an issue in situations in which a differentiated assessment of physical and mental load is required. Questionnaires like the NASA-TLX (Hart & Staveland, 1988), which use different scales to assess different forms of load, might be prone to this confoundation, which should be considered in the interpretation of results. Second, raters might be unable to distinguish between external demand and actual effort, again leading to distorted results. Third, the use of subjective measures in general is based on the assumption that increased capacity expenditure will be associated with subjective feelings of effort. However, some authors (Gopher & Donchin, 1986) argue that not all cognitive processing is accessible through introspection, which would limit the sensitivity of subjective assessments of effort dramatically. Fourth, there appears to be a dissociation between subjective ratings and task performance. It seems that, especially in dual-task settings, the subjective assessment seems to be influenced by factors such as the number of tasks or task elements that have to be performed, rather than the actual overlap in resource demand between the different tasks, which in turn would lead to inaccurate assessments of the actual load imposed by the tasks. In addition to those four major limitations, O'Donnell and Eggemeier (1986) also identify methodological constraints. One issue here is that it is often poorly defined what element of the task shall actually be evaluated. Also, the inevitable practice of asking an operator about his assessment only after a task has been completed may lead to distortions, especially in situations in which multiple ratings have to be made based on the memory of the respective load. Still, despite all these limitations, subjective measures see wide use especially in applied research. Kantowitz states that "they are used because they are easy to obtain and because sometimes they can be correlated with better measures of distraction" (in Llaneras, 2000; p. 58).

To overcome these issues, researchers have taken the effort to develop simple measures of driver distraction. Their purpose is to provide an easy, standardised way of assessing the distraction that is caused by the operation of in-vehicle systems as well as other distracting activities. It has to be acknowledged that there are critical limitations that go with all those simple methods. They are hardly capable of capturing natural user behaviour, as they basically force the participant to operate the respective system at a predefined moment and frequency, totally disregarding the possibility that a driver in a real environment might opt not to operate the system in various situations. However, the fact that they provide experimental control makes their assessment of a system much more reliable than any field study can deliver. This, in addition to fundamental differences in terms of cost and effort in comparison to other methods, renders those simple measures an important factor in distraction assessment. The three most common procedures - the occlusion method, the peripheral detection task and the lane change task - will be described in detail in the following sections (for a more extensive discussion, see Krems & Petzoldt, 2011). Paper three of this dissertation ("The critical tracking task - A useful method to assess driver distraction?") tries to add a fourth promising procedure, the critical tracking task (CTT; Jex, McDonnell, & Phatak, 1966), to this canon.

4.1 Occlusion method

The so called occlusion method (Senders, Kristofferson, Levison, Dietrich, & Ward, 1967) mainly aims at the visual distraction caused by a certain task. The general approach is to obstruct a participant's sight temporarily, while he is engaged in the task that is about to be assessed. Usually, this is achieved by having the participant wear a special kind of goggles controlled by a computer, which opens and shuts the goggles at predefined intervals. This is supposed to resemble real world driving, in which the operation of a certain in-vehicle technology ("goggles open" state) is necessarily interrupted by glances outside of the car ("goggles closed" state), in which no view of the task to be completed is available. This allows for an understanding of what glance frequency and duration the task requires under realistic conditions, and whether it can be interrupted and resumed easily or not.

The task is fairly simple to employ - usually, only a standard computer system and the occlusion goggles are necessary. The computation of the most common metrics is rather easy as well - the so called chunkability index is calculated by just relating the time it takes to complete the task (so called "total task time") in a control condition to the total "goggles open" time for an experimental condition in which vision was temporarily occluded (Noy, Lemoine, Klachan, & Burns, 2004). It also appears that the method is reliable (Gelau, Henning, & Krems, 2009) and valid (e.g. Baumann, Keinath, Krems, & Bengler, 2004; Krems, Keinath, Baumann, & Jahn, 2004) as a procedure for assessing visual demand caused by various forms of in-vehicle distraction, a finding which culminated in the publication of an ISO standard on the method (ISO 16673, 2007).

However, there are also critical shortcomings to be considered. Although the method emulates glances away from the in-vehicle task as would be the case in real traffic, it does not include actual dividing of attention between different sources of information or switching between tasks. This is rather artificial, and can only be compensated by introducing additional loading tasks. It might also be the case that certain in-vehicle tasks can be split up into smaller subtasks that are shorter in duration than the predefined “goggles open” interval. While in real traffic glances to the road would be adjusted to the completion of the subtasks, in the lab, the user is forced to continue the operation of the task and interrupt it at a point which, again, is rather artificial.

4.2 The peripheral detection task

Another simple method to assess driver distraction is the so called “peripheral detection task” (PDT; van Winsum, Martens, & Herland, 1999). Participants have to respond to visual stimuli presented in different distances to their normal line of sight while driving by pressing a finger switch. Stimuli are visible for 1 to 2 sec, and are presented at variable rates. Miura’s (1986) findings on peripheral vision provide the theoretical basis for the assumptions underlying the PDT. He reported that with increasing demand of the driving task, the size of the driver’s visual field decreased, and subsequently, reaction times to stimuli presented in the peripheral field of view increased. Williams (1985) described that effect as “visual tunnelling”.

Different implementations of the task exist. Martens and van Winsum (2000) employed it in a driving simulator, where they presented a small red square on the simulator screen. Their results support the hypotheses that the more demanding the driving task (either through higher demand in the primary driving task, or through an additional secondary task), the longer the response times to the presented stimuli, and the more misses. In a similar variant, Nakayama, Futami, Nakamura and Boer (1999) attached three LED indicators to the top of the instrument panel of a simulator, which were illuminated at varying intervals. They found that reaction times to the LED stimuli correlated significantly with a steering entropy measure of workload. In an on-road study, Olsson and Burns (2000) used LED projections on the windscreen. Response times and hit rates in the PDT were impaired relative to baseline driving when additional tasks like radio tuning, changing CDs or back-ward counting were performed. Others studies which used an LED projection setup (e.g. Harms & Patten, 2003; Jahn et al., 2005; Patten, Kircher, Östlund, & Nilsson, 2004; Patten, Kircher, Östlund, Nilsson, & Svenson, 2006) have come to similar results, confirming the PDT's validity as a measure of workload.

Overall, the different forms of the PDT appear to be sensitive to driving workload and to distraction from the use of an in-vehicle information system. The task itself is fairly easy to perform, and does usually not consume resources needed for safe driving. Also, the detection of peripheral visual stimuli can be likened to the detection of objects and events relevant for driving, which makes the PDT ecologically valid to a certain degree.

4.3 The lane change task

The lane change task (sometimes also lane change test; Mattes, 2003) "is a dual-task method that is intended to estimate secondary task demand on the driver, resulting from the operation of an in-vehicle device in a laboratory setting. The method is simple and inexpensive so that it can be used by vehicle manufacturers, in-vehicle device manufacturers, and other organizations" (ISO TC 22/SC 13 WG 8, 2008, p. v). It basically uses a simple simulation of a driving scene, in which the amount of distraction is assessed through the lane change performance in response to signs demanding such a change of lanes. Participants have to control a vehicle on a 3-lane road, with no other traffic present, and are commanded to change lanes by, and according to signs appearing on both sides of this road. The task is controlled by a game steering wheel with foot pedals for throttle and break. Standard performance measures are the mean deviation (MDEV) from a nominal lane change model, or the MDEV from a participants own baseline (adaptive model).

As the task is intended to become an ISO sanctioned procedure, recent studies have mainly been conducted to support the standardisation process of the LCT. Based on the assumption that visual and cognitive tasks lead to different types of driver errors, Engström and Markkula (2007) proposed the introduction of a high pass filtered standard deviation of lateral position (SDLP) and the percentage of correct lane (PCL) choices as new performance metrics. Whereas SDLP is supposed to capture effects on path control, PCL should reflect effects of reduced sign detection / recognition. Harbluk, Burns, Lochner and Trbovich (2007) argued for lane change initiation (LCI) as a useful measure, as it incorporates the detection and response delay as a result of distraction, aspects that are part of the driving task. They also consider secondary task time, as it accounts for risk exposure. To test the task's robustness across different experimental contexts, Rognin, Alidra, Val, Lescaut and Chalandon (2007) compared LCT performance in the usual desktop setup to LCT performance in a simulator environment. They concluded that the task is transferable from the PC to a vehicle based set-up. The trends observed were similar, although a general degradation was observed for the PC set-up. Bruyas, Brusque, Aurialt, Tattegrain, Aillerie and Duraz (2008) reported comparable results, explaining the differences between the set-ups with a greater immersion in the driving scene for the simulator condition.

One of the LCT's biggest advantages over occlusion and PDT is the test's intuitive validity. The operation of this driving-like task with a steering wheel appears to be much closer to real driving. Also, it incorporates aspects of cognitive, visual and manual control, making it sensible to those kinds of workload. However, although some new performance metrics show promising results, these different aspects can hardly be separated. The MDEV value is always a result of the combination of those loading factors. It is nearly impossible to assess directly to what proportion each factor contributes to the score. Also, there appear to be issues with the standardisation of the task, as results from different test sites differ considerably. Paper two of this dissertation ("Learning effects in the lane change task (LCT) - Evidence from two experimental studies") tries to assess this problem, as it investigates learning effects in the LCT, one of the suspected reasons for the variance in results.

5 Summary

In his book "The Design of Future Things", Donald Norman (2007) writes: "Someday cars will no longer need drivers. Instead, people will all be passengers, able to gossip, read, or even sleep while the car chauffeurs them to their destination" (p. 47). Automation is getting us closer to that scenario with every new assistance system that enters the vehicle. And as more and more of those assistance systems make driving easier and more comfortable, drivers will feel increasingly inclined to attend to secondary tasks while driving (see Wilde, 1982, for an explanation). However, until we arrive at Norman's utopia, the driver is still an important component of this intelligent system that includes man and machine. And so, until man can indeed safely hand over control completely to the machine, he still needs to be aware of what is happening around him, and he still needs to be able to respond quickly and appropriately. Therefore, the study of driver distraction will be an important field of research in many years to come. As control slowly shifts from driver to vehicle, and subsequently resources are freed, we will see that drivers will invest those resources in other activities in ways that no one would ever have anticipated. As Michael Regan puts it: “People in conditions of monotony in a car automatically are going to want to keep themselves stimulated, to make life a little more difficult for themselves, so there’s going to be a natural tendency for them to want to distract themselves under such conditions because distraction is one mechanism by which we can increase arousal and increase workload.” (in Gaffney, 2011).

This introduction was supposed to shed some light onto current issues in driver distraction. The prevalence and the effects of driver distraction have been reviewed, some potentially useful frameworks to explain the processes behind distraction and its effects have been presented, and methods for distraction measurement have been assessed. The following three papers are intended to contribute further in this regard:

Manuscript I - “The effect of cognitive tasks on predicting events in traffic”

The first manuscript focuses on the validation of theoretical assumptions about driver distraction. Numerous studies have shown that cognitively demanding secondary tasks have a negative impact on driving performance. In the manuscript, it is argued that this negative impact is caused by an interference between these secondary tasks and functions of working memory that serve to update the situation model of a traffic situation. This assumption was tested in a driving simulator study. There, it appeared that participants who operated a secondary task that was supposed to disrupt the integration of new information into the situation model reacted later to critical events that could have been anticipated than comparison groups. In contrast, there were no differences between the groups for events that eluded anticipation. The results show that the negative effects of cognitive distraction are indeed the result of an interference with specific working memory processes.

Manuscript II - “Learning effects in the lane change task (LCT) - Evidence from two experimental studies”

In the second manuscript, methodological issues are addressed. The lane change task (LCT; see also 4.3), a simple driving task in which an operator has to change lanes as commanded by road signs, is supposed to provide an easy yet reliable means for the assessment of distraction caused by various forms of secondary tasks. However, different test series revealed disturbing variances in the results obtained. One possible explanation for this undesirable effect is the degree of experience participants might have with the LCT. This assumption was tested in two experiments. Participants received different forms of training, and then operated the LCT and secondary tasks of varying difficulty in a testing session. The results show that the level of experience indeed has an effect on lane change performance, and that this effect can still be found months after the initial training. However, it is unclear if this effect is the sole cause for the observed variances.

Manuscript III - "The critical tracking task - A useful method to assess driver distraction?"

In the third manuscript, another potential measurement tool is presented. The critical tracking task (CTT) is a procedure that has not received a lot of attention in the context of driver distraction so far. It is a simple tracking task, in which an operator has to stabilise a dynamic, unstable element on a computer screen. The required continuous devotion of visual attention and manual control of the task are also basic requirements of the more complex driving task. Therefore, it was assumed that the CTT might be useful as a simple tool to assess driver distraction. This assumption was tested in four experiments, in which artificial secondary tasks as well as realistic in-vehicle tasks were employed. The results show that the CTT can indeed serve as a method to assess driver distraction.

References

Alm, H., & Nilsson, L. (1994). Changes in driver behaviour as a function of handsfree mobile phones: A simulator study. *Accident Analysis and Prevention, 26*, 441-451.

Backer-Grøndahl, A., Phillips, R., Sagberg, F., Touliou, K., & Gatscha, M. (2009). *Topics and applications of previous and current naturalistic driving studies*. Deliverable D1.1, PROLOGUE.

Baumann, M., Keinath, A., Krems, J. F., & Bengler, K. (2004). Evaluation of in-vehicle HMI using occlusion techniques. *Applied Ergonomics, 35*, 197-205.

Baumann, M., & Krems, J. F. (2007). Situation awareness and driving: A cognitive model. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments. Critical issues in driver interactions with intelligent transport systems* (pp. 253-265). London: Springer.

Boyle, J. M., & Lampkin, C. (2008). *2007 Motor vehicle occupant safety survey volume 4: Crash injury and emergency medical services report*. DOT HS 810 977.

Briem, V., & Hedman, L. R. (1995). Behavioral effects of mobile telephone use during simulated driving. *Ergonomics, 38*, 2536-2562.

Brookhuis, K. A., de Vries, G., & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention, 23*, 309-316.

Brown, I. D. (1994). Driver fatigue. *Human Factors, 36*, 298-314.

Brown, I. D. (2005). *Review of the "looked but failed to see" accident causation factor*. London: Department of Transport.

Brown, I. D., Tickner, A., & Simmonds, D. (1969). Interference between concurrent tasks of driving and telephoning. *Journal of Applied Psychology, 53*, 419-424.

Bruyas, M.-P., Brusque, C., Auriault, A., Tattegrain, H., Aillerie, I., & Duraz, M. (2008). Impairment of lane change performance due to distraction: Effect of experimental contexts. In C. Brusque (Ed.), *Proceedings of the European conference on human centred design for intelligent transport systems* (pp. 89-100), Lyon: Humanist Publications.

de Waard, D. (1996). *The measurement of drivers' mental workload*. Haren, The Netherlands: The Traffic Research Centre VSC, University of Groningen.

DEKRA. (2010). Viele telefonieren ohne Freisprechanlage [Press release]. Retrieved from http://www.dekra.de/de/c/document_library/get_file?uuid=4addb529-18ac-4531-97f6-27a7586d38f7&groupId=10100.

Dingus, T., McGehee, D., Hulse, M., Jahns, S., Manakkal, N., Mollenbauer, M., et al. (1995). *TravTek evaluation task C3 - Car camera study (FHWA-RD-94-076)*. McLean, VA: Office of Safety and Traffic Operations.

Distracted-driving campaign requires attitude adjustment. (2010, January 26). *USA Today*, p. 8A.

Dominguez, C. (1994). Can SA be defined? In M. Vidulich, C. Dominguez, E. Vogel, & G. McMillan (Eds.), *Situation awareness: Papers and annotated bibliography* (pp. 5-15). Wright-Patterson Airforce Base, OH: Air Force Systems Command.

Drews, F., & Strayer, D. (2009). Cellular phones and driver distraction. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver distraction. Theory, effects, and mitigation* (pp. 191-213). Boca Raton, FL: CRC Press.

Drews, F., Yazdani, H., Godfrey, C., Cooper, J., & Strayer, D. (2009). Text messaging during simulated driving. *Human Factors, 51*, 762-770.

Driving on razor's edge: Shaving behind the wheel. (2009, March 6). *CBS4*. Retrieved from http://cbs4.com/local/driving.shaving.crash.2.1542138.html.

Endsley, M. R. (1995). Towards a theory of situation awareness. *Human Factors, 37*, 32-64.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 3-32). Mahwah, NJ: Lawrence Erlbaum Associates.

Engström, J., & Markkula, G. (2007). 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*, Stevenson, WA, pp. 199-205.

Gaffney, G. (2011, April). *Driver safety: An interview with Dr Michael Regan* [Audio podcast]. Retrieved from http://media.libsyn.com/media/uxpod/DriverSafety.mp3.

Gärtner, U., König, W., & Wittig, T. (2001). Evaluation of manual vs. speech input when using a driver information system in real traffic. In *Proceedings of the 1st international driving symposium on human factors in driver assessment, training, and vehicle design*, Aspen, CO, pp. 7-13.

Gelau, C., Henning, M. J., & Krems, J. F. (2009). On the reliability of the occlusion technique as a tool for the assessment of the HMI of in-vehicle information and communication systems. *Applied Ergonomics, 40*, 181-184.

Gopher, D., & Donchin, E. (1986). Workload - An examination of the concept. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance. Volume II, Cognitive processes and performance* (pp. 41/1-41/49). New York: Wiley.

Green, P. (2000). Crashes induced by driver information systems and what can be done to reduce them (SAE paper 2000-01-C008). In *Proceedings of the 2000 International Congress on Transportation Electronics*, Warrendale, PA: Society of Automotive Engineers.

Harbluk, J. L., Burns, P. C., Lochner, M., & Trbovich, P. L. (2007). 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*, Stevenson, WA, pp. 16-22.

Harms, L., & Patten, C. (2003). Peripheral detection as a measure of driver distraction. A study of memory-based versus system-based navigation in a built-up area. *Transportation Research Part F, 6*, 23–36.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: Elsevier.

Hedlund, J., Simpson, H., & Mayhew, D. (2006). *International conference on distracted driving. Summary of proceedings and recommendations*. Ottawa: Canadian Automobile Association & Traffic Injury Research Foundation.

Hoel, J., Jaffard, M., & van Elslande, P. (2010). Attentional competition between tasks and its implications. In J. F. Krems, T. Petzoldt, & M. Henning (Eds.), *Proceedings of the European conference on human centred design for intelligent transport systems* (pp. 575-581), Lyon: Humanist Publications.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis and Prevention, 38*, 185-191.

Hosking, S., Young, K., & Regan, M. (2009). The effects of text messaging on young drivers. *Human Factors, 51*, 582-592.

ISO 16673. (2007). *Road vehicles - Ergonomic aspects of transport information and control systems - Occlusion method to assess visual demand due to the use of in-vehicle systems*. Geneva, Switzerland: ISO.

ISO TC 22/SC 13 WG 8. (2008). *Road vehicles - Ergonomic aspects of transport and control systems - Simulated lane change test to assess in-vehicle secondary task demand*. Geneva, Switzerland: ISO.

Ito, H., Uno, H., Atsumi, B., & Akamatsu, M. (2001). Visual distraction while driving: Trends in research and standardisation. *IATSS Research, 25*, 20-28.

Jahn, G., Oehme, A., Krems, J. F., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F, 8*, 255-275.

Jäncke, L., Musial, F., Vogt, J., & Kalveram, K. T. (1994). Monitoring radio programs and time of day affect simulated car driving performance. *Perceptual Motor Skills, 79*, 484-486.

Jenness, J. W., Lattanzio, R. J., O'Toole, M., & Taylor, N. (2002). Voice-activated dialling or eating a cheeseburger: Which is more distracting during simulated driving? In *Proceedings of the Human Factors and Ergonomics Society 46th annual meeting*, Santa Monica, CA, pp. 592–596.

Jex, H. R., McDonnell, J. D., & Phatak, A. V. (1966). A "critical" tracking task for manual control research. *IEEE Transactions on Human Factors in Electronics, 7*, 138-145.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine, 67*, 507-512.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (HS 810 59).

Kraftfahrt-Bundesamt (Ed.). (2010). *Jahresbericht 2009*. Flensburg, Germany.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 279-328). Washington, DC: Taylor and Francis.

Krems, J. F., Keinath, A., Baumann, M., & Jahn, G. (2004). Die Okklusionsmethode: Ein einfaches und valides Verfahren zur Bewertung der visuellen Beanspruchung von Zweitaufgaben. In B. Schlag (Ed.), *Verkehrspsychologie: Mobilität-Sicherheit-Fahrerassistenz* (pp. 335-349). Lengerich: Pabst Science Publishers.

Krems, J. F., & Petzoldt, T. (2011). Tools and procedures for measuring safety-relevant criteria. In Y. Barnard, R. Risser, & J. F. Krems (Eds.), *The safety of intelligent driver support systems. Design, evaluation and social perspectives* (pp. 93-109). Farnham, UK: Ashgate.

Kuipers, A., Kappers, A., van Holten, C. R., van Bergen, J. H. W., & Oosterveld, W. J. (1990). Spatial disorientation incidents in the R.N.L.A.F. F16 and F5 aircraft and suggestions for prevention, in *AGARD conference proceedings No. 478*, Neuilly sur Seine, France: NATO/AGARD.

Lee, J. D., Regan, M. A., & Young, K. L. (2009). What drives distraction? Distraction as a breakdown of mulitlevel control. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver distraction. Theory, effects, and mitigation* (pp. 41-56). Boca Raton, FL: CRC Press.

Llaneras, R. E. (2000). NHTSA driver distraction internet forum: Summary and proceedings. *NHTSA Driver Distraction Internet Forum*. Retrieved from http://www-nrd.nhtsa.dot.gov/pdf/nrd-13/FinalInternetForumReport.pdf.

Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics, 35*, 939-953.

Man charged for watching porn while driving. (2009, July 20). *Northumberland News*. Retrieved from http://www.northumberlandnews.com/article/131386.

Martens, M. H., & Van Winsum, W. (2000). Measuring distraction: The peripheral detection task. *NHTSA Driver Distraction Internet Forum*. Retrieved from http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/34.PDF.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. In H. Strasser et al. (Eds.), *Quality of work and products in enterprises of the future* (pp. 57-60). Stuttgart: Ergonomia.

McKnight, A. J., & McKnight, A. S. (1991). *The effect of cellular phone use upon driver attention*. Washington, DC: National Public Services Research Institute.

McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention, 25*, 259-265.

Michon, J. A. (1985). A critical view of driver behaviour models: What do we know, what should we do? In L. Evans, & R. C. Schwing (Eds.), *Human behaviour and traffic safety* (pp. 485-524). New York: Plenum Press

Miura, T. (1986). Coping with situational demands: A study of eye movements and peripheral vision performance, in A. G. Gale et al. (Eds.), *Vision in Vehicles-II*, (pp. 205-216), Amsterdam: Elsevier.

Nakayama, O., Futami, T., Nakamura, T., & Boer, E. R. (1999). *Development of a steering entropy method for evaluating driver workload* (SAE Technical Paper No. 1999-01-0892). Michigan: Society of Automotive Engineers.

Norman, D. A. (2007). *The design of future things*. New York: Basic Books.

Noy, Y. I., Lemoine, T. L., Klachan, C., & Burns, P. C. (2004). Task interruptability and duration as measures of visual distraction. *Applied Ergonomics, 35*, 207-213.

O'Donnell, R. D. & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance. Volume II, Cognitive processes and performance* (pp. 42/1-42/49). New York: Wiley.

Olsson, S., & Burns, P. C. (2000). Measuring driver visual distraction with a peripheral detection task. *NHTSA Driver Distraction Internet Forum*. Retrieved from http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/6.PDF.

Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H., Jamson, S., et al. (2004). *Haste Deliverable 2 - HMI and Safety- Related Driver Performance*. Project GRD1/2000/25361 S12.319626.

Owens, J. M., McLaughlin, S. B., & Sudweeks, J. (2011). Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis and Prevention, 43*, 939-947.

Patten, C. J. D., Kircher, A., Östlund, J., & Nilsson, L. (2004). Using mobile telephones: Cognitive workload and attention resource allocation. *Accident Analysis and Prevention, 36*, 341–350.

Patten, C. J. D., Kircher, A., Östlund, J., Nilsson, L., & Svenson, O. (2006). Driver experience and cognitive workload in different traffic environments. *Accident Analysis and Prevention, 38*, 887-894.

Pettitt, M., Burnett, G., & Stevens, A. (2005). Defining driver distraction. In *Proceedings of the 12th ITS World Congress*, San Francisco, CA.

Pickrell, T. M., & Ye, T. J. (2010). *Driver electronic device use in 2009* (DOT HS 811 372). Washington, DC: NHTSA.

Police: Driver distracted by sex toy. (2010, August 25). Cincinnati Enquirer. Retrieved from http://news.cincinnati.com/article/20100825/NEWS010701/8260303/Police-Driver-distracted-by-sex-toy.

Ranney, T. A., Mazzae, E., Garrott, R.,& Goodman, M. J. (2000). NHTSA driver distraction research: Past, present, and future. *NHTSA Driver Distraction Internet Forum*.

Retrieved from http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/233.PDF.

Regan, M. A., Hallett, C., & Gordon, C. P. (in press). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention*.

Rockwell, T. (1972). Skills, judgement and information acquisition in driving. In T. W. Forbes (Ed.), *Human factors in highway traffic safety research* (pp. 133-164). New York: Wiley-Interscience.

Rognin, L., Alidra, S., Val, C., Lescaut, A., & Chalandon, X. (2007). Driver strategies when interacting with information and comfort systems. In *Proceedings of the 20th international technical conference on the enhanced safety of vehicles*, Lyon, France, Paper number 07-0272.

Roscoe, A. H. (1987). Introduction. In A. H. Roscoe (Ed.), *The practical assessment of pilot workload, AGARDograph No. 282* (pp. 1-10). Neuilly sur Seine, France: NATO/AGARD.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review, 115*, 101-130.

Sayer, J. R., Devonshire, J. M., & Flanagan, C. A. (2007). Naturalistic driving performance during secondary tasks. In *Proceedings of the 4th international driving symposium on human factors in driver assessment, training, and vehicle design*, Stevenson, WA, pp. 224-230.

Senders, J. W., Kristofferson, A. B., Levison, W. H., Dietrich, D. W., & Ward, J. L. (1967). The attentional demand of automobile driving.*Highway Research Record, 195*, 15-33.

Sheridan, T. (2004). Driver distraction from a control theory perspective. *Human Factors, 46*, 587-599.

Srinivasan, R., & Jovanis, P. P. (1997). Effect of in-vehicle route guidance systems on driver workload and choice of vehicle speed: Findings from a driving simulator experiment. In Y. I. Noy (Ed.), *Ergonomics and safety in intelligent driver interfaces* (pp. 97-114). Mahwah, NJ: Lawrence Erlbaum Associates.

Stutts, J., Feaganes, J., Rodgman, E., Hamlett, C., Meadows, T., Reinfurt, D., et al. (2003). *Distractions in everyday driving*. Washington, DC: AAA Foundation for Traffic Safety.

Stutts, J., Reinfurt, D., Staplin, L., & Rodgman, E. (2001). *The role of driver distraction in traffic crashes*. Washington, DC: AAA Foundation for Traffic Safety.

Tijerina, L., Parmer, E., & Goodman, M. J. (1998). Driver workload assessment of route guidance system destination entry while driving: A test track study. In *Proceedings of the 5th ITS World Congress*, Seoul, Korea.

Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: Speech recognition versus a touch-screen keyboard. *Human Factors, 46*, 600-610.

Treat, J. R. (1980). A study of precrash factors involved in traffic accidents. *The HSRI Review, 10*(1).

US Department of Transportation. (2009, October 1). *U.S. Transportation Secretary Ray LaHood announces administration wide effort to combat distracted driving* [Press release]. Retrieved from http://distraction.gov/files/for-media/10.01.09.pdf.

van Winsum, W., Martens, M. H., & Herland, L. (1999). *The effects of speech versus tactile driver support messages on workload, driver behaviour and user acceptance* (Report No. TM-99-C043). Soesterberg, Netherlands: TNO Human Factors.

Vanlaar, W., Simpson, H. M., Mayhew, D., & Robertson, R. (2007). *The road safety monitor 2006: Distracted driving*. Ottawa: Traffic Injury Research Foundation.

Violanti, J. M., & Marshall, J. R. (1996). Cellular phones and traffic accidents: An epidemiological approach. *Accident Analysis and Prevention, 28*, 265-270.

Wallace, B. (2003a). Driver distraction by advertising: Genuine risk or urban myth? *Municipal Engineer, 156*, 185–190.

Wallace, B. (2003b). *External-to-vehicle driver distraction.* Edinburgh: Scottish Executive Social Research.

Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, B. (2001). *Pilot task management: Testing an attentional expected value model of visual scanning* (ARL-01-14/NASA-01-7). Savoy, IL: University of Illinois, Aviation Research Lab.

Wickens, C. D., & Hollands, J. G. (1999). *Engineering psychology and human performance* (3rd Edition). Upper Saddle River, NJ: Prentice Hall.

Wikman, A. S., Nieminen, T., & Summala, H. (1998). Driving experience and time-sharing during in-car tasks on roads of different width. *Ergonomics, 41*, 358-372.

Wilde, G. J. S. (1982). The theory of risk homeostasis: Implications for safety and health. *Risk Analysis, 2*, 209-225.

Williams, L. J. (1985). Tunnel vision induced by a foveal load manipulation. *Human Factors, 27*, 221-227.

Wilson, G. F. (1992). Applied use of cardiac and respiration measures: Practical considerations and precautions. *Biological Psychology, 34*, 163-178.

Young, M. S., Mahfoud, J. M., Walker, G. H., Jenkins, D. P., & Stanton, N. A. (2008). Crash dieting: The effects of eating and drinking on driving performance. *Accident Analysis and Prevention, 40*, 142-148.

Man charged for watching porn while driving

A Mississauga man faces a charge of operating a motor vehicle with a TV visible to the driver, as well as speeding, after Northumberland OPP stopped his vehicle on Hwy. 401, in Port Hope, after a traffic complaint, at 12:40 a.m., on July 18.

OPP found the driver was watching a pornographic movie on a TV placed on the front seat of the vehicle. Police also noticed evidence of alcohol impairment, but the 32-year-old driver registered a low reading on a breath test, and was charged with speeding and watching TV when driving.

(Northumberland News, July 20th, 2009)

The effect of cognitive tasks on predicting events in traffic

Martin R. K. Baumann, Tibor Petzoldt, and Jeroen H. Hogema

under revision for *Transportation Research Part F*

Abstract

Numerous studies demonstrate the negative effects of cognitively loading secondary tasks on driving performance. We assume that this effect is caused by interference between these secondary tasks and central executive functions of working memory that serve to keep the driver's situation model of the current traffic situation updated. In this experiment 48 drivers had to drive in a high-fidelity driving simulator on a rural road while performing no secondary task, or a working memory task (auditive monitoring) that should not interfere with situation awareness, or a working memory task (memory updating) that should interfere with the comprehension and prediction function of situation awareness. While driving, participants had to react to events that were either announced by a warning signal or not. We hypothesised that participants would benefit least from the warning signal when they had to perform the memory updating task. The results generally support this hypothesis indicating that central executive functions of working memory are highly involved in situation awareness processes.

Keywords: Cognitive distraction; Driver distraction; Situation awareness; Event anticipation

1 Introduction

Driver distraction and driver inattention are one of the most frequent factors involved in accident causation. The Austrian Kuratorium für Verkehrssicherheit (Messner & Ohmann, 2008) reports a 10.6% fraction of all road accidents with lethal consequences to be linked to inattention. The New Zealand Ministry of Transport (2008) reports that in as much as 11% of all crashes in 2007 “diverted attention” was identified as a contributing factor. According to an analysis of crash data from the US “Crashworthiness Data System” (CDS), which is based on the thorough investigation of a subsample of all road accidents, 8.3% of involved drivers were identified as distracted, 5.4% as “looked but did not see,” and 1.8% as sleepy or asleep (Stutts, Reinfurt, Staplin, & Rodgman, 2001). Field research in the US draws a similar picture. Klauer, Dingus, Neale, Sudweeks, and Ramsey (2006) point to a large variety of sources of distraction that were present in what they define as crashes, near crashes and incidents. Stutts, Feaganes, Rodgman, Hamlett, Meadows, Reinfurt et al. (2003) report higher numbers of “adverse vehicle events” in connection to various distracting factors. Finally, there is a large body of experimental research demonstrating the negative effects of distraction, manly caused by concurrent secondary tasks, on driving performance. These studies generally find an increase in response latencies of drivers performing cognitively loading tasks (Strayer & Johnston, 2001; Patten, Kircher, Östlund, & Nilsson, 2004; Alm & Nilsson, 1995), a decrement in the breadth of visual scanning (Recarte & Nunes, 2000), an impaired anticipation of braking requirements (Jamson & Merat, 2005) or an impaired comprehension of perceived situation elements (Brown, 2005).

Whereas these studies are carefully designed to investigate possible detrimental effects of certain distracting tasks, frequently the use of cell phones, on the driving performance (e.g. Alm & Nilsson, 1995; Patten et al., 2004; McKnight & McKnight, 1993; Hancock, Lesch, & Simmons, 2003; Salvucci & Macuga, 2002; Strayer & Johnston, 2001), they are not very explicit about the causation of these effects, The aim of our study was to investigate the causation of these detrimental effects of secondary tasks on driving performance in more detail focussing on the causation of cognitive distraction effects. The basic hypothesis underlying this study was that cognitively distracting tasks interfere with the driver’s construction of a mental representation of the current traffic situation that is the basis for the driver’s action selection.

1.1 The driving task and its complexities

Despite the fact that driving is an everyday task, from a psychological point of view it is highly complex. It takes place in a dynamic environment, that is, in an environment that constantly changes even if the driver does not take action. It involves many cognitive processes ranging from perception, attention to comprehension, decision making, action selection and execution. To successfully perform this task, drivers need to perceive, identify, correctly interpret, and integrate relevant objects and elements of the current traffic situation into a coherent mental representation of the current traffic situation, the situation model. This situation model represents the driver's understanding of the current situation and how this situation will develop in the near future (e.g., Baumann & Krems, 2007; Endsley, 1995). Based on this situation model the driver is able to adapt her / his actions as soon as possible to upcoming events, such as hazardous situations.

We assume that the construction of the situation model is based on a comprehension process (Baumann & Krems, 2007; Durso, Rawson, & Girotto, 2007; Kintsch, 1998) that consists, first, of the activation of knowledge stored in long-term memory and associated with perceived elements of the current traffic situation, and, second, of the integration of this activated knowledge into a coherent network of knowledge that represents the driver's understanding of the current situation. What kind of knowledge becomes activated depends on the learnt associations between the perceived situation elements and the knowledge stored in long-term memory. So, for example, the perception of a "STOP" sign probably activates the traffic rule it stands for. But it will also activate certain actions, such as reducing the speed to stop before the sign. It will also activate expectations about the behaviour of other traffic participants in the near future, such as the expectation that a lead car will also start to decelerate in order to stop in front of the "STOP" sign. This activated knowledge will then be integrated into the coherent situation model by strengthening the activation of elements in the network being compatible with each other and with the already existing situation model and by inhibiting those elements that are incompatible with each other. The result is a network of activated knowledge held partly in working memory and partly stored in long-term memory. Those parts of the situation model that will be directly used, for example to choose actions, will be kept in working memory for further processing. Other parts of the situation model will be kept in long-term memory to be retrieved if needed, such as the destination of the trip that will be retrieved when the driver has to make a navigation decision. As the comprehension of a traffic situation also involves the activation of expectations about its future development learnt from previous encounters of similar situations, this comprehension process also serves the anticipation of upcoming events, at least in situations familiar to the driver. This assumption is supported by findings that experienced drivers are much better and much faster in identifying dangerous traffic situations than less experienced drivers (Crundall, Chapman, Phelps, & Underwood, 2003; Crundall & Underwood, 1998; Underwood, Chapman, Berger, & Crundall, 2003). In non-routine situations additional attention-demanding processes are necessary to make predictions about the further development of the situation.

The above described comprehension process relies on the availability of working memory resources to result in a coherent, complete and correct situation model. Working memory resources are necessary for the activation of knowledge from long-term memory and the integration of the activated knowledge into a coherent situation model (Baddeley & Logie, 1999; Fischer & Glanzer, 1986; Glanzer & Nolan, 1986; Kintsch, 1998). Reducing the availability of these resources by additional tasks will result in failures to integrate relevant knowledge associated with newly perceived information from the traffic situation into the situation model. As this information is not integrated into the situation model it will not affect the driver's action selection. This will also reduce the driver's anticipation capacity as in some cases this relevant knowledge that will not become integrated into the situation model consists of expectations about the future development of the traffic situation. Therefore, we assume that the detrimental effect of cognitively distracting secondary tasks is caused by an interference between these cognitively distracting tasks and the comprehension process as both tasks compete for working memory resources. The "looked but did not see" accidents described by Stutts et al. (2001), or the so called "looked-but-did-not-see phenomenon" (Brown, 2005) can easily be described with this line of argument. Relevant information is perceived but the comprehension of this information fails as the drivers' working memory resources are occupied by other tasks. Therefore, important consequences of the perceived information become not integrated into the situation model and are not considered during action selection.

1.2 Goals of the current study

The specific aim of this experiment was to test, first, whether the negative effects of cognitive distraction are at least in part caused by interfering with the anticipation of traffic events, and, second, which kind of cognitive processes are especially involved in anticipating events.

Therefore the participants in this experiment drove through a scenario that contained both predictable and non-predictable events. These events were designed to be exactly equivalent besides that in the predictable version a warning sign warned the driver of the upcoming event. The reaction to the event when the participant was warned was compared to the reaction when the driver was not warned.

While driving the participants had to perform either i) no secondary task, or ii) a cognitively loading secondary task that does not interfere with the integration of expectations into a situation model or iii) a secondary task that interferes with the integration of expectations into the situation model. For the non-interfering task an auditory monitoring task was used that was designed to load on cognitive resources by forcing the participant to attend and react to auditory stimuli as fast as possible but avoiding any early response. But this task should not interfere with the integration of information retrieved from long-term memory into the situation model. In this monitoring task participants had to react as fast as possible to an auditory signal that was presented either after a long or a short time interval after the previous signal. By using only two randomly presented interstimulus intervals, this task induces a strong tendency for rhythmic responding. If this tendency is not inhibited, errors in terms of early responses to the stimuli will occur, especially after the occurrence of two short intervals. According to Vandierendonck, de Vooght, and van der Goten (1998) inhibiting inappropriate responses is a vital cognitive function that serves the coordination of action.

As a task that should interfere with the integration of expectations into the situation model, a running memory task was used. In the running memory task (e.g., Pollack, Johnson, & Knaff, 1959) participants are presented with a constant stream of items and they have to keep in mind always just the last items of the stream, for example the last three items. That is, each time a new item is presented it has to be encoded in working memory and the “oldest” item has to be removed from working memory. Performing this task directly involves those central executive functions that control working memory content, i.e. those functions that should also be highly involved in maintaining and updating a proper situation model. Therefore, we assume that this updating of working memory is highly interfering with the comprehension of a current situation and with the anticipation of its future development.

To summarise, we assumed that participants driving the scenario without performing a secondary task should clearly benefit from the warning signs in the predictable events. The benefit from warning signs was expected to be reduced when participants have to perform an additional task while driving. And the reduction of this benefit should be greater when participants have to perform the running memory task than when they have to perform the monitoring task. The memory task was expected to interfere more with the comprehension of a traffic situation and the anticipation of its future development by impairing the integration of expectations into the situation model.

2 Method

2.1 Participants

48 participants took part in this experiment. Participants ranged in age from 21 to 58 with a mean age of 36.9 years (SD = 12.1). 29 of them were male. All participants were in possession of a valid driving licence for at least one year. Each one of them drove at least 10,000 km per year.

2.2 Driving scenario

The experiment was run in the high-fidelity driving simulator of TNO in Soesterberg, The Netherlands. A rural, two-lane road scenario was employed with curvy and straight sections, allowing for an approximate speed of 80 kph. Each participant drove the scenario only once, with a drive taking about 20 min.

In each scenario, participants encountered four critical events. In each of these events the driving lane was blocked by an obstacle. In two cases, a construction site was positioned right behind a curve, for the other two, a broken down truck of the same size as the construction site was used in a corresponding position. When participants approached the obstacle, oncoming traffic prevented an immediate passing manoeuvre. Instead, a breaking reaction was necessary to avoid colliding with the roadblock. Only after the participant had reached standstill, a gap in the oncoming traffic allowed the participant to pass the obstacle.

In two of the critical events the driver was given information that allowed the prediction of the obstacle. For the construction site, there was a sign indicating a work zone ahead. For the broken down truck, there was a warning triangle in the same position. The signs were put right before the curve that hid the obstacle. The signs themselves were hidden by trees and became only visible for the driver until the driver was 50m in front of the sign.

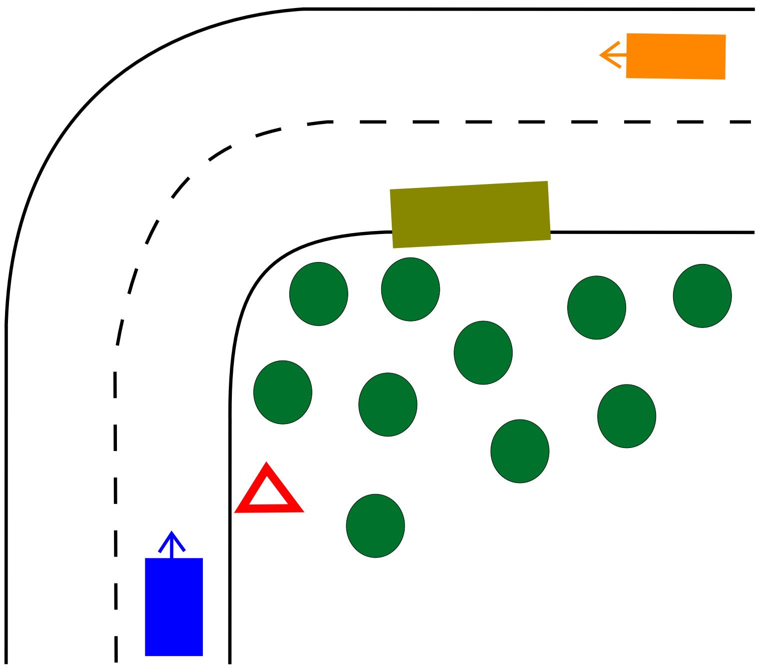


Figure 1: Sketch of the basic layout of the critical events - ego-vehicle (blue) approaching curve, roadblock behind the curve with trees blocking direct sight, oncoming traffic (red) to ensure braking manoeuvre of driver of ego-vehicle (in this example warning sign to allow for anticipation).

With this setup we were able to exactly control the point where the driver was presented with information allowing the prediction of the upcoming event. Additionally, with the sight to the roadblock after the curve being blocked by additional trees, the roadblock could only be identified after passing the sign and entering the curve. The other two critical events did not include any warning ahead of the actual obstacle. In all other aspects, however, the situations with and without warning were equivalent (see Figure 1 for an example).

2.3 Secondary tasks

In each secondary task condition participants had to perform several trials of the respective task. The trials lasted between 20 and 40 sec, and were followed by a no secondary task phase (where participants only had to drive) also lasting between 20 and 40 sec. Start and end of a secondary task trial were triggered when the participant passed certain positions on the road. Therefore it was guaranteed that the participants in the secondary task conditions encountered the critical events while performing the respective secondary task.

2.3.1 The monitoring task

In the monitoring task participants had to react as fast as possible to an acoustical signal consisting of a short, clearly audible sound. As the response device a finger switch was used that was applied to the index finger of the participant's dominant hand. The time interval between two successive signals was either 1 or 2 sec, randomly chosen. We measured the participants' response times and the numbers of errors, predominantly early responses, where the participants pressed the finger switch before the acoustical signal was presented.

2.3.2 The running memory task / updating task

In the running memory task participants were presented an audio stream of letters, presented with a fixed frequency of 1 letter per 2 sec. The participants' task was to repeat the current last three letters each time a new letter was presented. For example, assume the letters "S", "P", and "Q" were already presented and the next letter was "G", the participant had to repeat loudly "P Q G", after presentation of "G"; and after the next letter "M", the participant had to repeat "Q G M", and so on. After a variable amount of time (20 to 40 sec) an acoustical signal was presented to inform the participant about the end of the current secondary task trial. After that the last repeated triplet of letters was taken as the participant's response in this trial.

2.4 Design

Two independent variables were manipulated: the type of secondary task (no secondary task, monitoring task, running memory task) and the predictability of the critical event (predictable, non-predictable). The secondary task was manipulated as between-subjects factor. The whole sample of 48 participants was divided into three groups with 16 participants each. The participants of each group had to perform either no, the monitoring, or the running memory task. Predictability was manipulated as within-subjects factor. Each of the participants encountered both predictable and non-predictable events during his / her drive. This resulted in a 3 (secondary task) x 2 (predictability) mixed factorial design.

As dependent measures different parameters characterizing the driving performance were calculated to examine the participants' reaction to the critical events, such as time-to-collision at the time of response onset or maximum brake pressure after the obstacle became visible. Driving performance in phases without critical events was also recorded to allow for the assessment of the effects of secondary task performance on normal driving. Additionally, the participants were asked to rate their workload when performing each task while driving using the Rating Scale for Mental Effort (RSME, Zijlstra & van Doorn, 1985). Also, a short questionnaire containing items like "... indicate how much you were aware of the presence of other traffic participants..." was to be filled in. For the monitoring and the memory group some items occurred twice - once in a "driving only" context, and once in a "driving with additional task" context.

3 Results and discussion

3.1 Secondary task difficulty

The basic assumption is that the difference between the two secondary tasks regarding their effects on the prediction of traffic events is due to the fact that the two tasks interfere with different cognitive processes. One task interferes with processes highly relevant for the prediction of traffic events, the other one with less relevant ones. In situations where the prediction of traffic events is not involved, both tasks should show a comparable effect on driving performance. Therefore, seven sections of straight and seven sections of curved road were selected not involving critical situations. In four of the straight and four of the curved sections participants performed the secondary task. Within these sections, the standard deviation of lateral position was calculated as it is correlated with overall workload, especially under conditions of light workload and traffic (Green, Lin, & Bagian, 1993) to examine the overall task difficulty effect on driving performance. An ANOVA (2 x 2 x 2) for mixed designs was calculated with the within-subjects factors curvature (curved vs. straight) and load (task vs. no task) and the between-subjects factor task type (monitoring vs. updating). The results show that the standard deviation of lateral position is significantly higher in curved sections, $F(1, 30) = 6.776$, $p = .014$. Also, sections with concurrent secondary task produced significantly higher deviations, $F(1, 30) = 5.397$, $p = .027$. This is important as it shows that the tasks actually did cause workload. The interaction between the factors curvature and load was also significant, showing that especially in the curvy sections, the presence of a secondary task led to higher deviations, $F(1, 30) = 8.982$, $p = .005$. Most importantly, however, there is no significant effect of task

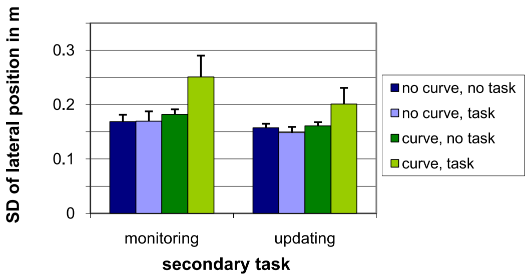


Figure 2: Standard deviation (SD) of lateral position in non-critical road sections, depending on the type of secondary task, the curvature, and the actual involvement in the secondary task; error bars indicate standard error.

type on standard deviation of lateral position, $F(1, 45) = 1.778$, $p = .192$, or any significant interaction involving the factor task type (see Figure 2).

The subjective ratings of workload obtained with the RSME were also analysed. There was no significant difference between the workload ratings of the running memory and the monitoring task, $t(29.946) = -0.3401$, $p = 0.736$. In addition, we computed the differences in ratings for those questionnaire items that where presented twice - asking for a “drive only” and for a “drive + task” rating. The analysis of those computed values revealed no significant differences for any of the “drive only” as well as the “drive + task”-items (p range from .133 to 1) between the two groups. These results further strengthen our claim of comparability of the tasks in terms of general cognitive demand.

Considering the results of the standard deviation of lateral position and the subjective workload ratings of the RSME, it can be concluded that both tasks are comparable in their global task difficulty. Therefore, any effects on the performance in the critical situations can indeed be attributed to the different structural interference of the secondary tasks due to their differences in the involvement of cognitive processes.

3.2 Reactions in critical situations

The results presented here focus mainly on two aspects of the participants’ reactions. First, there should be differences in preparatory behaviour after the warning information. That is, although the actual roadblock is not yet present, there should already be some adaptations visible for those participants who adequately perceive and process the given information, and therefore correctly predict the upcoming obstacle. Those adaptations are expected to be smaller or even absent for those participants that do not adequately predict the upcoming event due to the interference caused by secondary tasks. Second, once the roadblock is visible, “prepared” participants are expected to be reacting faster and more appropriate than those participants that are not prepared to the obstacle as they did not predict the event.

To identify any preparatory behaviour induced by the traffic signs the speed when the roadblock became visible was analysed. At this point, drivers who comprehended the information given by the sign were expected to travel at a lower speed than those who did not. Also the difference in speed between the position were the traffic sign became visible (or would have been for the unpredictable cases) and the position where the roadblock became visible was analysed to control for any adjustments in speed before the event occurs.

The analysis shows that when the roadblock became visible, there was no difference in speed among the no secondary task condition and the two secondary task conditions in case of a non-predictable obstacle, indicating the validity of the experimental design (Figure 3, left). But there was a clear difference in speed between these conditions in case of a predictable obstacle. Participants in the secondary task conditions were driving faster than participants in the no secondary task condition, indicating that the former participants were less prepared to the obstacle despite the warning sign. Comparing the speed before predictable and non-predictable obstacles, the greatest difference in speed between these predictability conditions at this position could be found in the no secondary task condition (12.3 kph), a medium difference in the monitoring condition (5.7 kph), and the smallest in the memory updating condition (4.8 kph). This indicates that participants

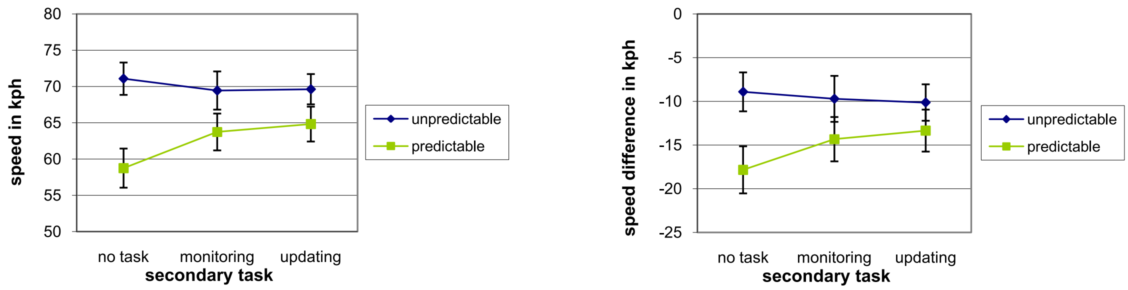


Figure 3: Preparatory behaviour in critical situations, measured as speed when the roadblock became visible (left) and speed difference between the positions where warning sign and roadblock became visible (right); error bars indicate standard error.

in the memory updating condition had the greatest difficulties to comprehend, integrate, and react to the warning sign. A 3 (secondary task) x 2 (predictability) mixed ANOVA revealed a significant main effect of predictability, $F(1, 45) = 41.439$, $p < .001$, indicating that speed was significantly higher for non-predictable events at the time when the roadblock became visible. Additionally, the interaction between secondary task condition and predictability for speed as dependent measure was significant, $F(2, 45) = 4.011$, $p = .025$, confirming the reduced benefit from the warning signal from no secondary task to monitoring task to running memory task condition. The main effect of secondary task condition did not reach significance, $F(2, 45) = .291$, $p = .749$.

For the values in speed difference between the positions where warning sign and roadblock became visible a similar picture emerges (Figure 3, right). Again, no difference in values for unpredictable events could be found, whereas the values for predictable events show a decrease in benefit from the warning sign when a concurrent task was performed. This impression is strengthened by the comparison between the tasks - the difference in values is largest for the no secondary task condition (-8.9 kph), medium for the monitoring condition (-5.6 kph), and smallest for the updating condition (-3.2 kph). The ANOVA again confirmed a main effect of predictability, $F(1, 45) = 28.491$, $p < .001$, and no main effect of secondary task type, $F(2, 45) = .331$, $p = .720$. The interaction is only significant at the 10% level, $F(2, 45) = 2.699$, $p = .078$.

Also other variables were analysed that might serve as indicators of the correct prediction of the upcoming traffic situation. More specifically, it was assessed whether participants took their foot off the throttle (Figure 4, left) or even used the brake pedal to decelerate between the position were the traffic sign became visible (or would have been for the unpredictable cases) and the position where the roadblock became visible. Figure 4 shows for how many situations per group and event category “foot-off-the-throttle” or brake events were found. There was again a difference between predictable and unpredictable events. The possibility to anticipate a critical situation yielded a much higher number of anticipatory reactions in form of deceleration by braking or just taking the foot off the accelerator pedal. Also, it appears that overall the no secondary task group scores best, whereas the updating task group scores worst. Although the tendency is clear, however, numbers are too small for a meaningful statistical analysis and can serve only as secondary indicators of preparatory behaviour.

TTC values with respect to the obstacle at the moment the driver released the throttle to decelerate after passing the location where the obstacle first became visible were also

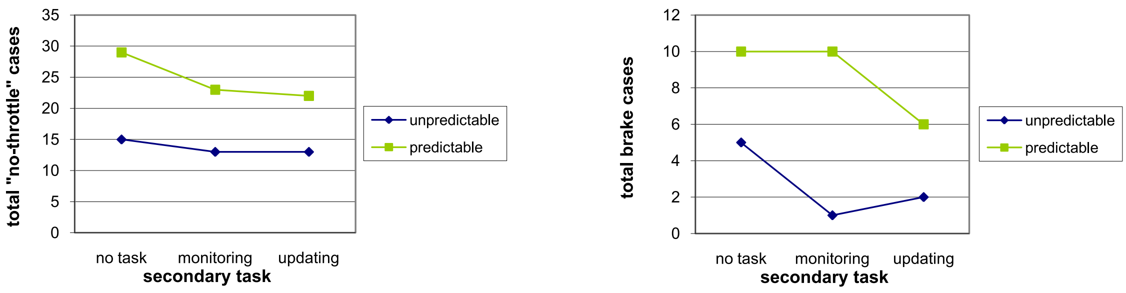


Figure 4: Preparatory behaviour in critical situations, measured as number of “no-throttle” cases until the roadblock became visible (left) and number of brake cases until the roadblock became visible (right).

analysed. As participants prepared for the obstacle should decelerate earlier than participants that did not comprehend the warning signal fully, TTC for prepared participants was expected to be higher than for unprepared. This difference should be greatest in the no secondary task condition and lowest in the running memory task condition. As shown in Figure 5 (left), the results confirm this prediction. Whereas the difference in TTC between the predictable and the non-predictable obstacle is 1.6 sec in the no secondary task condition, it is 0.8 sec in the monitoring task condition and 0.6 sec in the running memory task condition. This picture was confirmed by a 3 (secondary task) x 2 (predictability) mixed ANOVA. TTC values were significantly greater for predictable events, $F(1, 45) = 34.305$, $p < .001$. And, most importantly, the interaction between secondary task condition and predictability was significant, $F(2, 45) = 5.253$, $p = .009$, reflecting the reduction in the difference between predictable and non-predictable events from no secondary task to

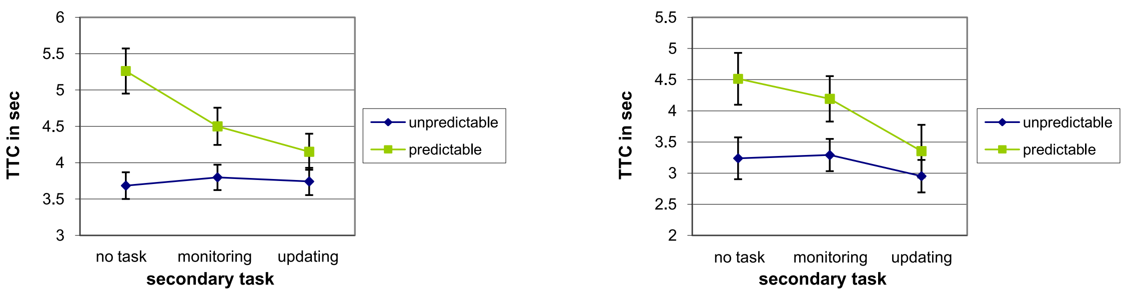


Figure 5: Direct reactions in critical situations, measured as TTC at the moment the driver released the throttle for the first time after the roadblock became visible (left) and TTC at the moment of maximum deceleration after the roadblock became visible (right); error bars indicate standard error.

monitoring task to running memory task condition. The main effect of secondary task condition did not reach significance, $F(2, 45) = 1.948$, $p = .154$.

Next, TTC values at the moment of maximum deceleration were analysed. This moment in time was chosen as it defines the point where the driver's brake reaction had completely built up. We assumed that this point characterises the driver's immediate reaction to the obstacle and her / his preparedness to the obstacle, again having in mind that an anticipation of the oncoming roadblock should lead to earlier reactions (Figure 5, right). Again, predictable events resulted in greater TTC values than unpredictable ones (main effect of predictability, $F(1, 43) = 16.911$, $p < .001$). However, no interaction between secondary task type and predictability could be found, $F(2,43) = 1.341$, $p = .272$. There was no main effect of secondary task type, $F(2,43) = 1.620$, $p = .200$.

4 Conclusions

The aim of this experiment was to examine one possible cognitive foundation of the negative effects of cognitively distracting tasks on driving performance. On the basis of a comprehension-based theory of situation awareness (Baumann & Krems, 2007; Durso et al., 2007), it was assumed that if cognitively distracting tasks interfere with processes involved in the updating and maintenance of information in working memory, then these tasks would interfere with the construction of a proper situation model. In such a case the situation model might not include all relevant implications of the perceived environmental information. The situation model is not complete and therefore the situation is not fully understood. More specifically, the situation model might lack relevant expectations about the future development of the situation that would normally be activated and integrated during an undisturbed construction process. Therefore, the driver's anticipation of traffic events should be impaired. Cognitive distraction might be especially detrimental for the anticipation of traffic events as anticipation relies heavily on the availability of working memory processes and resources. It requires both the integration of the relevant environment information itself and the integration of the expectations associated with and derived from this information. The results of the driving simulator experiment presented in this paper in general confirmed our predictions.

In accordance with the comprehension based model of situation awareness the running memory task interfered most with the anticipation of upcoming traffic events. All measures relevant to assess whether the driver anticipated the predictable event showed consistently that the drivers were least prepared to the event when simultaneously performing the running memory task. There was also a significant impairment of the drivers' anticipation performance when they performed the monitoring task and the difference between the interference by the monitoring and the running memory task was less than expected. But the monitoring task interfered always less than the running memory task.

An alternative explanation for the difference between the two cognitively demanding tasks on anticipation is that the running memory task is generally more difficult than the monitoring task and that the structural differences between the two tasks are of minor importance. In this case, one would expect to find differences between the two task conditions in driving performance also in situations where working memory updating and maintenance are not as critically involved as in the critical anticipation event but that nevertheless represent demanding driving conditions. The results show that this was not case. For example, negotiating curves was not differentially affected by the secondary tasks. Subjective measures draw a similar picture. Participants rated both tasks as of a comparable difficulty. Therefore, it can be assumed that the differential effect of both tasks on event anticipation is due to the structural differences between the tasks.

Why does the monitoring task interfere with the anticipation of events at all? One reason is that a key feature of the monitoring task is that the sequences of short and long intervals between the acoustical signals rather quickly lead to expectations about the next interval, especially in the case of two or three succeeding short intervals. In this case there is a strong tendency to respond automatically without waiting for the acoustical signal. This tendency has to be inhibited to avoid an early response. The inhibition of activated responses is an active resource consuming process (e.g., Hasher & Zacks, 1988) and might also draw on those working memory resources that are involved in the activation and inhibition of information in working memory (Baddeley, 1996; Baddeley, Emslie, Kolodny, & Duncan, 1998). Additionally, performing a secondary task in itself requires the distribution of cognitive resources among different tasks and their coordination. This definitely draws on central working memory resources thereby interfering with the comprehension processes involved in the anticipation of events.

Overall, the pattern of results confirms the assumptions of the comprehension-based model of situation awareness (Baumann & Krems, 2007; Durso et al., 2007). Furthermore, it points to one possible causal factor underlying the effects of cognitive distraction. Situation awareness is the result of a knowledge activation and integration process. During this process newly perceived information from the environment activates associated knowledge from long-term memory that is then integrated into a coherent representation of the current traffic situation. This integrated representation, the situation model, is the basis for the driver's action selection. The proper integration of activated knowledge relies strongly on the availability of working memory resources and processes. The control of working memory content, involving updating the content and removing irrelevant information from working memory, and the retrieval of information from long-term memory to comprehend encoded information in working memory are such processes (Adams, Tenney, & Pew, 1995; Baddeley, 1996; Gugerty, 1997). When this activation and integration processes are disturbed by cognitively distracting tasks, the driver's anticipation of events is impaired and one can observe the typical effects of cognitive distraction on driver performance, such as prolonged reaction times to those events. But this might be only one factor underlying cognitive distraction effects. Further experiments are necessary to investigate more deeply cognitive processes underlying the effects of cognitive distraction to better understand how to avoid or at least minimise such effects while driving.

From the findings of this study there also arise some immediate practical implications. The mere existence of differential effects of cognitive distraction challenges the idea of investigating cognitive distraction as a whole by simple lab based procedures. Various simple methods to assess driver distraction caused by the interaction with an in-vehicle system while driving, such as the lane change task (LCT; Mattes, 2003), claim to be able to reflect the impact of cognitive distraction on driving performance caused by this interaction. The results of the study indicate detrimental, but comparable effects on driving performance of both of the secondary tasks in a basic metric like lane position. But both secondary tasks clearly differed in their effect on the ability to correctly anticipate upcoming events. As simple methods for the assessment of cognitive distraction on driving performance involve only such basic metrics, the ability of simple lab based methods to assess the full spectrum of cognitive distraction by in-vehicle tasks and its effects on driving performance is at least debatable. As most simple procedures do not feature a "prediction of x" - element, it is disputable if these metrics really cover this aspect of cognitive distraction, which would result in an underestimation of possible distractive effects a certain system may cause.

It is also possible to derive some recommendations for HMI design. It appears that systems or system tasks that require drivers to cognitively manipulate working memory content should be avoided if possible. One simple example would be a congestion assistant that recommends alternative routes. If the assistant only displays the length of the alternative route, and the driver would have to compute the expected driving time and compare it to the current route himself, this would result in much higher cognitive distraction than when this information would be provided directly. The knowledge that certain cognitive tasks are even more demanding than others might also be used to additionally support the driver in situations where as of now, no support is provided. Taking away the burden of manipulation of information in working memory while driving when possible might help to free resources and can be crucial for successfully mastering a critical situation.

Overall, it appears that the nature of cognitive distraction, its diversity and its impact are not fully understood. Despite the large body of research on cognitive load and cognitive distraction in driving, there are still many questions to be answered. Still, the thorough investigation of these aspects is not futile. The presented study is just one example of how it can be possible to shed some more light on an important issue in traffic safety research.

Acknowledgments

This research received funding by the EC 6th Framework Programme under grant agreement n°507420 Network of Excellence “Humanist” (Human centred design for information society technologies). We would like to thank Ingmar Stel from TNO for programming the driving simulation scenario, preprocessing the data according to our changing requirements and supporting this study with helpful hints. We also would like to thank Marika Hoedemaeker and Wiel Janssen from TNO for their help in designing and conducting this experiment, and Chantal Grounewoud for her support in running the experiment.

References

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors, 37*, 85-104.

Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis and Prevention, 27*, 707-715.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A*, 5-28.

Baddeley, A., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of working memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 51A*, 819-852.

Baddeley A. D., & Logie, R. H. (1999). Working memory: The multi-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28-61). New York: Cambridge University Press.

Baumann, M., & Krems, J. F. (2007). Situation awareness and driving: A cognitive model. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments. Critical issues in driver interactions with intelligent transport systems* (pp. 253-265). London: Springer.

Brown, I. D. (2005). *Review of the "looked but failed to see" accident causation factor* (Road Safety Research Report No. 60). London: Department for Transport.

Crundall, D. E., Chapman, P., Phelps, N., & Underwood, G. (2003). Eye movements and hazard perception in police pursuit and emergency response driving. *Journal of Experimental Psychology: Applied, 9*, 163-174.

Crundall, D. E., & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics, 41*, 448-458.

Durso, F. T., Rawson, K. A., & Girotto, S. (2007). Situation comprehension and situation awareness. In F. T. Durso et al. (Eds.), *Handbook of applied cognition* (pp. 163-193). Chichester, UK: Wiley.

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32-64.

Fischer, B., & Glanzer, M. (1986). Short-term storage and the processing of cohesion during reading. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 38A*, 431-460.

Glanzer, M., & Nolan, S. D. (1986). Memory mechanisms in text comprehension. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 20, pp. 275-317). New York: Academic Press.

Green, P., Lin, B., & Bagian, T. (1993). *Effect of road geometry on driving workload: A pilot experiment* (Technical Report UMTRI-93-39, GLCTTR 22-91/01). Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Gugerty, L. J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied, 3*, 42-66.

Hancock, P. A., Lesch, M., & Simmons, L. (2003). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis and Prevention, 35*, 501-514.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Jamson, A. H., & Merat, N. (2005). Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. *Transportation Research Part F, 8*, 79-96.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (Report DOT HS 810-594). Washington, DC: NHTSA.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. In H. Strasser et al. (Eds.), *Quality of work and products in enterprises of the future* (pp. 57-60). Stuttgart: Ergonomia.

McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention, 25*, 259-265.

Messner, U., & Omann, D. (Eds.), (2008). *Verkehr in Oesterreich. Road accident statistics 2007*. Vienna: Kuratorium für Verkehrssicherheit KfV.

New Zealand Ministry of Transport (2008). *Diverted attention: Crash statistics for the year that ended 31 dec 2007*. Wellington, New Zealand: Strategy and Sustainability Group, Ministry of Transport.

Patten, C. J. D., Kircher, A., Östlund, J., & Nilsson, L. (2004). Using mobile telephones: Cognitive workload and attention resource allocation. *Accident Analysis and Prevention, 36*, 341-350.

Pollack, I., Johnson, L. B., & Knaff, P. R. (1959). Running memory span. *Journal of Experimental Psychology, 57*, 137-146.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied, 6*, 31-43.

Salvucci, D. D., & Macuga, K. L. (2002). Predicting the effects of cellular-phone dialing on driver performance. *Cognitive Systems Research, 3*, 95-102.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual tasks studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12*, 462-466.

Stutts, J., Feaganes, J., Rodgman, E., Hamlett, C., Meadows, T., Reinfurt, D., et al. (2003). *Distractions in everyday driving*. Washington, DC: AAA Foundation for Traffic Safety.

Stutts, J., Reinfurt, D., Staplin, L., & Rodgman, E. (2001). *The role of driver distraction in traffic crashes*. Washington, DC: AAA Foundation for Traffic Safety.

Underwood, G., Chapman, P., Berger, Z., & Crundall, D. (2003). Driving experience, attentional focusing, and the recall of recently inspected events. *Transportation Research Part F, 6*, 289-304.

Vandierendonck, A., de Vooght, G., & van der Goten, K. (1998). Interfering with the central executive by means of a random interval repetition task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 51A*, 197-218.

Zijlstra, F. R. H., & van Doorn, L. (1985). *The construction of a scale to measure perceived effort* (Technical Report). Delft, The Netherlands: Department of Philosophy and Social Sciences, Delft University of Technology.

Driving on razor’s edge: Shaving behind the wheel

(. . .) According to the Florida Highway Patrol, a two-car crash on Cudjoe Key was caused by a 37-year-old woman who was shaving her bikini area while in the driver’s seat. Her ex-husband was steering from the passenger seat.

Trooper Gary Dunick explained, “She said she was meeting her boyfriend in Key West and wanted to be ready for the visit.” (. . .)

(CBS4.com, March 6th, 2009)

Learning effects in the lane change task (LCT) - Evidence from two experimental studies

Tibor Petzoldt, Nina Bär, Claudia Ihle and Josef F. Krems

published 2011 in *Transportation Research Part F, 14*, 1-12.

Abstract

Given the ever-growing distribution of new in-vehicle information systems, the assessment of their distraction potential becomes an important issue. An accurate estimation of their impact on driver behaviour should be made in the early stages of product development. Several easy-to-use methods can be used to make this early estimate, one of them being the lane change task (LCT). As this task is being considered as an ISO standard, questions about factors that might influence or even distort the results obtained through this procedure arise. One problem, which is the focus of this paper, is the possible occurrence of learning effects. We report the results of two experiments that show that participants' performance improves significantly after just one LCT encounter, and that this improvement is rather stable.

Keywords: Lane change task; Driving; Distraction; Learning

1 Introduction

As new in-vehicle information systems have become increasingly popular (Starry, 2001), they have also become more and more the subject of thorough investigation, as it is often argued that such systems lead to an increase in driver distraction ("technology-based distraction", Young, Regan, & Hammer, 2003). As field research (Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006), official crash statistics (Kuratorium für Verkehrssicherheit, 2008; New Zealand Ministry of Transport, 2007), and in-depth crash analyses (Stutts, Reinfurt, Staplin, & Rodgman, 2001) have shown, distraction is a major issue in traffic accidents. Also, field and laboratory studies have highlighted certain aspects of the impact that driver distraction has on driving performance. Critical deterioration of driving performance has been observed concerning lateral position, speed maintenance, reaction times, and gap acceptance (see Young et al., 2003 for an overview), often as a result of changes in glance behaviour (Lansdown, 2001).

Industry leaders and government authorities are well aware of this problem. The assessment of the potential distraction caused by new in-vehicle devices is an issue that has been (and is still being) addressed in several projects, such as the Driver Workload Metrics Project from the Crash Avoidance Metrics Partnership (CAMP; Angell, Auflick, Austria, Kochhar, Tijerina, Biever, et al., 2006) and the Advanced Driver Attention Metrics (ADAM; Breuer, Bengler, Heinrich, & Reichelt, 2003) project. Within ADAM, a set of easy-to-use methods has been developed and evaluated to assess the extent of distraction imposed by performing secondary tasks while driving. The occlusion method (Senders, Kristofferson, Levison, Dietrich, & Ward, 1967; see also Baumann, Keinath, Krems, & Bengler, 2004; Keinath, Baumann, Gelau, Bengler, & Krems, 2001) targets visual distraction, especially focusing on (non-)interruptability of secondary tasks. This method has already become an ISO-standardised procedure (ISO 16673, 2007). The peripheral detection task (PDT; Jahn, Oehme, Krems, & Gelau, 2005) tries to assess cognitive and visual distraction by making use of the fact that visual and cognitive load can narrow the driver's functional field of view (Miura, 1986). A third method, the lane change test / task (LCT; Mattes, 2003), has also been developed to address the issue of visual distraction. As both occlusion and PDT are rather artificial procedures, they lack face validity, as neither bears any obvious resemblance to activities connected to actual driving. The LCT, however, employs the look of a driving simulator (see Section 1.1), and therefore, aside from its scientific grounding, it has a certain appeal when it comes to communicating results obtained in studies on driver distraction. Consequently, the LCT is now under investigation as a potential ISO standard. The major objective of this paper is to investigate whether the LCT delivers results stable enough for it to be considered a good, reliable measure of distraction.

1.1 The lane change task (LCT)

The LCT is a simple, inexpensive dual-task method intended to estimate secondary task demand on a driver as a result of the operation of an in-vehicle device in a laboratory setting. Participants have to control a simulated vehicle on a three-lane road, with no other traffic present, and are instructed to change lanes according to signs appearing on both sides of this road (Figure 1). Participants are required to maintain a constant speed of 60 kph. Exceeding this limit is not possible. The signs appear around every 150 m; duration between lane changes is therefore around 9 s.

Main performance measures are the mean deviation (MDEV) from a nominal lane change model, or the MDEV from a participant's own baseline (adaptive model). Additional measures as discussed in the ISO draft might be a modified standard deviation of lateral position, the proportion of missed or erroneous lane changes, or the mean delay in lane change initiation.

Recent studies employing the LCT have so far mainly focused on methodological aspects such as the connection between secondary task time and LCT performance (Burns, Tr-

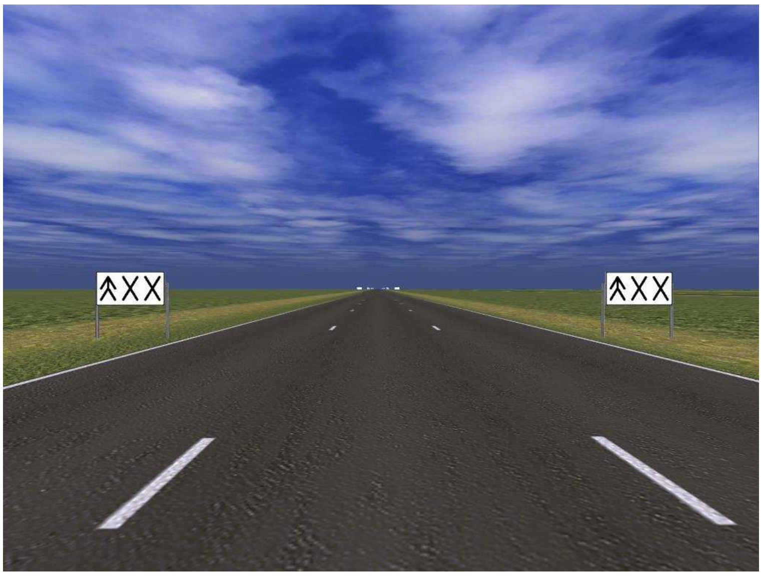


Figure 1: Lane change task (LCT) example screen.

bovich, McCurdie, & Harbluk, 2005; Harbluk, Mitroi, & Burns, 2009), the introduction of new performance metrics (Engström & Markkula, 2007; Harbluk, Burns, Lochner, & Trbovich, 2007), or the susceptibility to different experimental contexts (Bruyas, Brusque, Auriault, Tattegrain, Aillerie, & Duraz, 2008; Rognin, Alidra, Val, Lescaut, & Chalandon, 2007). However, regardless of the studies available, at the time of this writing some issues remain unresolved. Baseline values vary considerably between different test sites. For instance, Weir, Kwok, and Peak (2007) reported MDEV values of around 0.64, Rognin et al. (2007) values of 1.6. Bearing in mind that the LCT is expected to become a standardised procedure, that is, is supposed to produce comparable results under comparable circumstances, this variance in performance metrics is quite worrisome. Possible factors that might influence the metrics in such a dramatic manner have to be assessed. One first step has been taken by Petzoldt, Bär, and Krems (2009), who analysed the influence of the participant sample's composition, and more specifically, the effect of gender on LCT performance. They reported substantial differences between male and female participants in terms of LCT and secondary task performance. However, this effect cannot account for all the variance that has been observed in the aforementioned studies. Another possible factor of influence might be experience with the LCT, which is the focus of this paper.

1.2 LCT and training

One important question that needs to be answered is whether the repeated assignment of the same participants might influence LCT results. When evaluating in-vehicle information systems, manufacturers usually rely on a predefined group of possible participants who are quite regularly deployed for studies of this kind. It can be argued that this practice leads to training effects that might distort the results obtained. As long ago as the 1800s Ebbinghaus (1885 / 1971) described the "learning curve", which refers to the relation between the amount of learning and the time it takes to learn. Especially in the early stages of skill acquisition, a tremendous increase in performance quality can be observed. It is doubtful if this early stage of rapid learning is already completed after the short LCT familiarisation phase that precedes the actual test. For novices, learning, and therefore performance advancement, might still occur in experimental trials. Experts would be expected to perform at a high level right from the beginning. Shinar, Tractinsky, and Compton (2005) investigated the effect of practice on interference from a phone task while driving in a simulator. They found that after five sessions of driving and using the phone, there was a learning effect on most of the driving measures, and that this training was even sufficient to eliminate driving impairment caused by the phone task altogether in the group of more experienced drivers.

Not only for the LCT, but also for the secondary task variations in performance can be expected. Jahn, Krems, and Gelau (2009) described the course of skill acquisition in operating navigation systems, pointing out that training effects might occur for the secondary tasks, as well. Also, training effects do not have to be limited to single tasks only. It has been demonstrated that dual tasks are more than just the sum of their component tasks (Bahrick, Noble, & Fitts, 1954; Bahrick & Shelly, 1958). It appears that dual-task training not only leads to an increased automation of the respective tasks, but also helps develop the skill to optimally allocate resources between them. Damos and Wickens (1980) identified such time-sharing skills and their development in dual-task training. They also found evidence for the transfer of those skills to other task combinations. It can be argued that previous encounters with the LCT in conjunction with a secondary task might facilitate time sharing in subsequent experimental instances and might do so even with different secondary tasks. Again, considering the learning curve, LCT novices might still be acquiring the necessary skill in experimental trials, while experts start at their best and show an overall superior performance. So the first question to be answered is whether a first encounter with the LCT in an experimental setting serves as LCT training for subsequent encounters. Given the very simple nature of the initial training for the LCT and secondary task, and the practical absence of dual-task training before the actual test, we hypothesise that training effects will occur.

The second question to be examined is whether such a training effect is stable. In this context, the “forgetting curve” is a relevant concept, as it describes the course of forgetting acquired knowledge over time (Ebbinghaus, 1885 / 1971). The largest portion of forgetting appears to take place in the time directly after learning. With increasing time, forgetting happens less and less. For the LCT, it can be argued that even if training takes place, the effects will disappear very quickly. Taking into account that participants are not usually tested on a day-to-day basis, but rather with intervals of weeks and months, it is possible that the effects of training become negligible after some time, making any experimental control for training or experience redundant.

However, Ebbinghaus also provided evidence that could suggest the opposite. First, there is the effect of overlearning. Overlearning occurs when knowledge already acquired and understood perfectly is still being learned. The same applies to motor skills. In such cases, the forgetting curve does not appear. A proper and very valid example is driving a car. After learning how to drive, people of course just go on driving, making every ride another learning trial. Even if not driving for some time, an experienced driver will not forget how to drive. It is questionable whether such an effect occurs after solving the LCT once. However, regular LCT driving might indeed lead to overlearning. This would result in better performance in subsequent LCT trials even with larger inter-experiment intervals. A second aspect of relevance is the concept of savings. Ebbinghaus (1885 / 1971) stated that, even if forgetting occurs, people do not require the same amount of time to reacquire knowledge or skills once learned and then forgotten as they did to learn it the first time. Thus, although during the first LCT encounter learning is still happening in the experimental trials, this is not necessarily the case in subsequent studies. There, the simple act of training for the LCT and secondary task before starting the actual experiment might just be enough to reach the performance plateau that was reached the first time only during the experiment, and not before. This effect, in contrast to overlearning, can be expected even after only one previous LCT experience. In this paper, we present two experiments that try to shed light on the issue of the LCT and training. The first experiment assesses whether training effects occur at all. The second experiment more closely resembles the practical use of the LCT in terms of inter-experiment interval, addressing the question of whether any training effect found earlier is of practical relevance.

2 Experiment I

In this experiment, we tried to uncover any effect of experience that might occur. The LCT was used in conjunction with two different secondary tasks that resemble different aspects of in-vehicle tasks. The surrogate reference task (SuRT) is a task that is chunkable and allows for interruption. The critical tracking task (CTT), in contrast, requires continuous attention. We hypothesized that any training with the LCT would result in improved dual-task performance, either through direct learning effects or indirectly through the freeing of resources. More specifically, we expected dual-task training (LCT + secondary) to produce better performance than just LCT training.

2.1 Method

2.1.1 Participants

Fifty-two participants took part in this study; 5 had to be removed from the data set for being statistical outliers in terms of LCT performance (mean values more than 2 *SDs* different from group average). All of the remaining 47 (age *M* = 29.1 years, *SD* = 8.5 years, 25 male, 22 female) were in possession of a valid driving license and drove a mean of 16,500 km a year, *SD* = 2,250 km (outlier with 350,000 km a year excluded here). None had previous LCT experience. All of the student participants (32) received course credit; the remainder received monetary compensation. Students and nonstudents were distributed equally over the different experimental conditions.

2.1.2 Material

Lane change task (LCT) For presentation of the LCT, the desktop setup as described in the ISO draft was employed. A standard PC system with a 19” flat screen was used. To control the vehicle, a MOMO force-feedback game steering wheel with foot pedals was connected to the PC. The length of a single LCT trial corresponded to the length of one LCT track, 1,800 m, which should take roughly 3 min, provided the participants follow instructions. Any secondary task was terminated as soon as the end of the track was reached.

Surrogate reference task (SuRT) The SuRT we employed in the experiment required participants to scan stimulus displays for the one stimulus that differed from the others surrounding it (Figure 2). Target and distracters were white circles on a black background. Distracter size could be varied to create different levels of difficulty (later referred to as SuRT1, SuRT2, and SuRT3). Participants gave their response by moving a grey indicator bar to the position of the identified target and pressing the enter key for confirmation, after which the next display appeared. As a performance metric, we analysed the number of displays correctly solved per LCT trial, as this might best reflect any attentional strategy. The task was presented on an 8.37” screen to the right of the participant. The indicator bar was controlled using a standard keyboard. Position of screen and keyboard matched the requirements of the ISO draft.

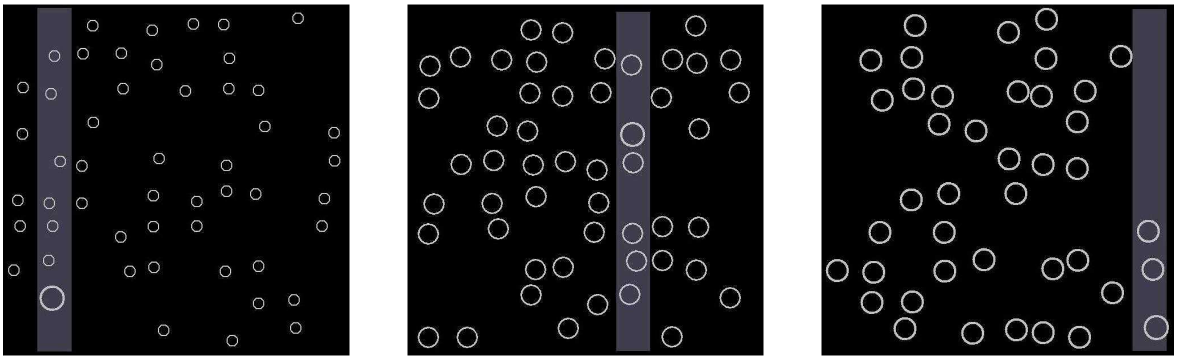


Figure 2: Surrogate reference task (SuRT) examples for low (SuRT1, left), moderate (SuRT2, centre) and high (SuRT3, right) difficulty.

Critical tracking task (CTT) The main goal in the CTT (Figure 3) is the manual control of a dynamic unstable element. This element is a simple horizontal bar that tends to leave the proposed target position at the centre of the screen. While the bar moves up or down continuously, participants try to control this deviation by using the up and down keys on a keyboard, with the ultimate goal of bringing the bar back to the centre. If the bar gets too far away from the middle of the screen, its colour changes to red to alert the participant and capture his / her attention (see Figure 3, right). The task allows for the variation of difficulty by letting the experimenter choose the level of instability. Three different levels of difficulty were used in the experiment (later referred to as CTT1, CTT2, and CTT3). The standard deviation from the central position is used as a performance indicator. Display and keyboard were the same as for the SuRT.

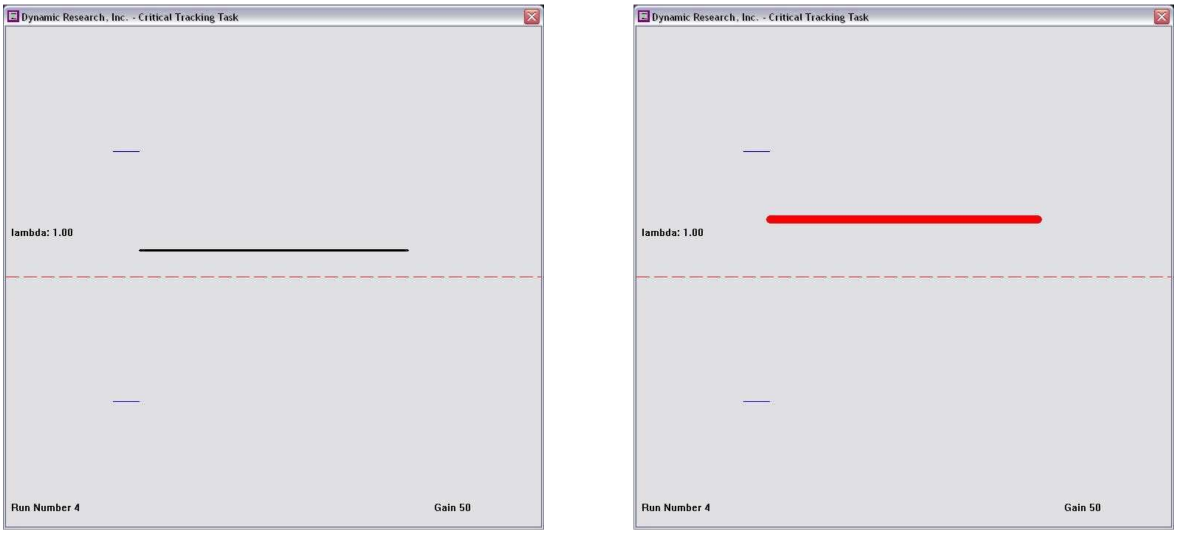


Figure 3: Critical tracking task (CTT) example screens.

Driving activity load index (DALI) For a subjective rating of workload, we employed the DALI questionnaire (Pauzié & Pachiaudi, 1997). The questionnaire is derived from the NASA-TLX (Hart & Staveland, 1988) and is intended for the assessment of workload experienced while driving with an additional task. We used five of the questionnaire's seven scales that were suitable for the experimental setup: global attention demand, visual demand, stress, temporal demand, and interference (the interference subscale was not used in baseline drives, as its purpose is to capture the interference caused by a concurrent secondary task). The auditory demand and tactile demand subscales (specific constraints induced by vibrations during the test) were omitted.

2.1.3 Procedure

Participants were divided into three groups. The control group did not receive any training and therefore had to show up only for one testing session. The second group received an "LCT only" training. In the training session, they drove the LCT with varying instructions (e.g., "drive only with your right hand") to avoid boredom and fatigue for about 20-30 min. The third group received full LCT and secondary task training. They started with a short phase of LCT driving to get familiar with the task. After that, a baseline was recorded. Then, the different secondary task conditions (blocked for task type, random for difficulty within task type) were administered in a balanced fashion. After each single trial, participants had to fill in the DALI questionnaire.

The testing session was the same for all three groups and followed the procedure of the training session for the dual-task-trained group. First, there was a familiarisation phase, then the recording of a baseline drive, and afterwards the balanced application of the secondary task conditions, each directly followed by the DALI. Training and testing session were a maximum of 1 week apart.

2.2 Results

2.2.1 LCT + SuRT, between-groups comparison

To assess possible differences between the three training groups, we analysed the mean deviation of the driven course from the normative model (MDEV) by calculating a two-way analysis of variance (ANOVA) for mixed designs. The analysis revealed a main effect of secondary task condition, $F(3, 132) = 46.61$, $p < .001$. Pairwise comparison of conditions, however, showed differences only between the baseline and all the SuRT conditions (all $p < .001$). We found no main effect of training, $F(2, 44) = 3.09$, $p = .056$. Also, there was no interaction between the secondary task condition and the type of training. As can be seen in Figure 4 (left), baseline values do not differ greatly, but the fully trained group outperformed the other two in the SuRT conditions. The control group showed the largest performance decrement with a concurrent secondary task.

To control for possible effects of variations in attention allocation, we analysed secondary task performance, measured as the number of displays correctly solved (Figure 4, right). A two-way ANOVA for mixed designs produced the anticipated main effect of secondary task difficulty, $F(2, 88) = 285.33$, $p < .001$. No effect of training, $F(2, 44) = .28$, $p = .759$, and no interaction between secondary task and training were found.

We analysed subjective measures as obtained with the DALI (for an overview, see Table 1) by calculating a two-way ANOVA for mixed designs for each subscale. Every subscale produced a significant main effect of secondary task condition (all $p < .001$). In addition, the scales global attention demand, $F(2, 44) = 4.86$, $p = .012$, and visual demand, $F(2, 44) = 7.07$, $p = .002$, showed significant main effects of training. Post hoc testing revealed a difference in global attention demand between the fully trained group and the control group ($p = .018$), as well as differences in visual demand between the fully trained group and both other groups ($p = .009$ and $p = .008$). Other scales showed no effect of training, and no interactions were found.

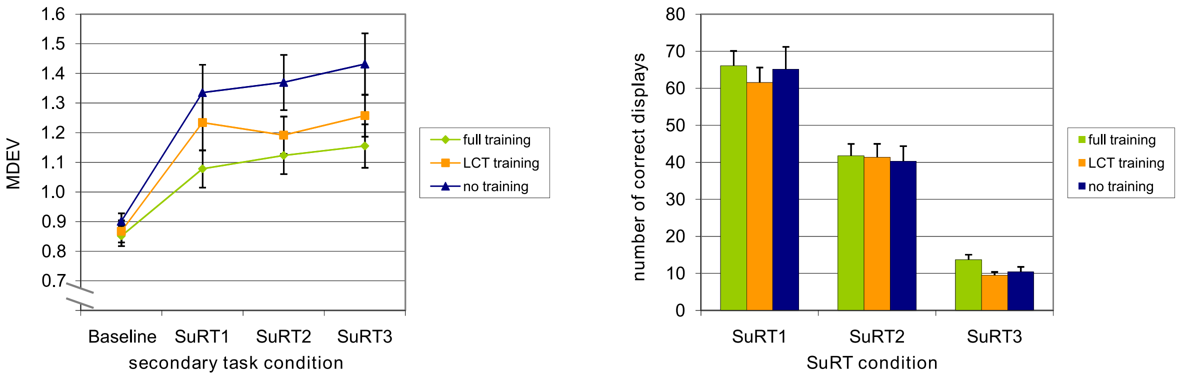


Figure 4: Mean deviation (MDEV) of the driven course from the normative model in the LCT (left) and number of correct displays in the SuRT (right) for different training groups (error bars indicate standard error).

Table 1: Subjective assessment of the lane change task (LCT) plus the surrogate reference task (SuRT) as obtained by the driving activity load index (DALI); for between-groups comparison see rows "no training" vs. "LCT only training" vs. "full training (testing session)"; for comparison of training sessions see "full training (testing session)" vs. "full training (training session)".

		baseline		LCT + SuRT1		LCT+SuRT2		LCT+SuRT3	
scale	group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
global attention demand	no training	3.24	1.48	4.71	1.05	5.41	0.94	5.88	0.33
	LCT only training	3.21	1.48	4.79	1.12	5.00	1.30	5.86	0.36
	full training (testing)	2.50	1.15	3.75	1.39	4.25	1.39	5.44	0.81
	full training (training)	3.69	1.58	4.38	1.26	4.69	1.01	5.75	0.58
visual demand	no training	3.59	1.46	4.35	1.37	5.18	0.88	5.82	0.53
	LCT only training	3.21	1.53	4.64	1.15	5.14	0.86	5.93	0.27
	full training (testing)	2.38	0.96	3.38	1.09	4.06	1.24	5.69	0.60
	full training (training)	3.69	1.62	4.25	1.29	4.88	1.09	5.94	0.25
stress	no training	1.94	1.09	3.76	1.25	3.88	1.05	5.06	0.97
	LCT only training	2.57	1.50	3.64	1.78	4.00	1.62	4.57	1.74
	full training (testing)	1.81	1.17	3.00	1.26	3.44	1.67	4.31	1.58
	full training (training)	2.94	1.48	3.81	1.22	3.69	1.25	4.94	1.18
temporal demand	no training	1.71	1.40	3.59	1.54	3.82	1.59	4.47	1.62
	LCT only training	2.00	1.36	3.21	1.53	3.93	1.77	3.93	1.44
	full training (testing)	1.81	1.22	2.94	1.57	3.19	1.42	4.06	1.73
	full training (training)	2.19	1.42	3.88	1.09	3.88	1.20	4.75	1.18

Note: SuRT1, SuRT2, and SuRT3 indicate three levels of difficulty in the task.

2.2.2 LCT + SuRT, comparison of training sessions

Since there are two complete data sets (training session and testing session) available for the fully trained group, we calculated a two-way repeated measures ANOVA on the MDEV to assess possible learning effects within this group. Again, we found a main effect of secondary task condition, $F(3, 45) = 23.46$, $p < .001$. Also, we found a main effect of training, $F(1, 15) = 5.42$, $p = .034$. There was no interaction between the secondary task condition and training level. As can be seen in Figure 5 (left), in the second session the lane keeping / changing quality increased for all secondary task conditions.

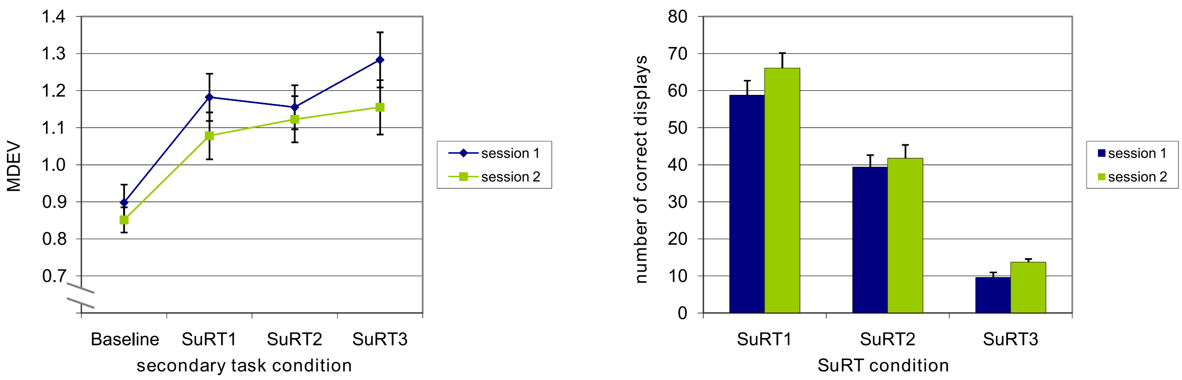


Figure 5: LCT (left) and SuRT (right) performance for different testing sessions (error bars indicate standard error).

When analysing secondary task performance (Figure 5, right), we found a main effect of secondary task difficulty, $F(2, 30) = 164.21$, $p < .001$, but no significant effect of training level, $F(1, 15) = 1.70$, $p = .213$. There was, however, a significant interaction between secondary task difficulty and training level, $F(2, 30) = 5.59$, $p = .009$.

The analysis of the DALI's (Table 1) subscales again revealed a significant main effect of secondary task condition on each scale (all $p < .001$). In addition, the scales global attention demand, $F(1, 15) = 14.60$, $p = .002$; visual demand, $F(1, 15) = 32.09$, $p < .001$; stress, $F(1, 15) = 13.33$, $p = .002$; and temporal demand, $F(1, 15) = 9.77$, $p = .007$, showed significant main effects of training. No interactions were found.

2.2.3 LCT + CTT, between-groups comparison

The analysis for assessing training effects concerning the LCT / CTT combination follows the same procedures as for the SuRT. For differences between the training groups, we calculated a two-way ANOVA for mixed designs on the MDEV values (Figure 6, left). The analysis revealed a main effect of secondary task condition, $F(3, 132) = 76.57$, $p < .001$. Pairwise comparison of conditions showed significant differences between all of them (five out of six $p < .001$). However, we found no effect of training, $F(2, 44) = 1.13$, $p = .332$. There was no interaction between the secondary task condition and the type of training.

Analysing secondary task performance, we found a main effect of secondary task difficulty, $F(2, 88) = 182.00$, $p < .001$, as well as of training, $F(2, 44) = 4.10$, $p = .023$. There was no interaction between the two. As can be seen in Figure 6 (right), the "full train-

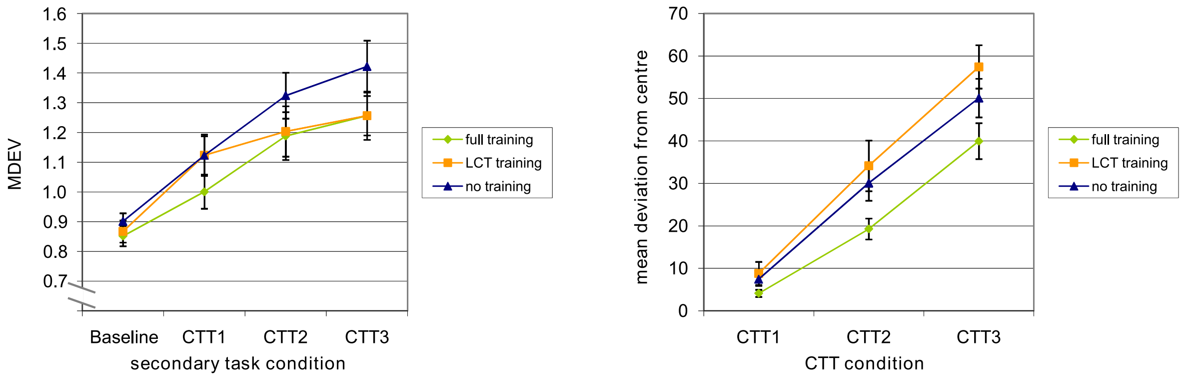


Figure 6: LCT (left) and CTT (right) performance for different training groups (error bars indicate standard error).

ing" group produced the smallest deviations, whereas the "LCT training" group performed even worse than the control group.

Subjective measures as obtained with the DALI (Table 2) showed a pattern similar to that of the SuRT. There was again a significant main effect of secondary task condition for each subscale (all $p < .001$). In addition, the scales global attention demand, $F(2, 44) = 3.32$, $p = .045$, and visual demand, $F(2, 44) = 3.72$, $p = .032$, showed significant main effects of training. Post hoc testing, however, showed no significant differences between the groups. Other scales showed no effect of training, and no interactions were found.

2.2.4 LCT + CTT, comparison of training sessions

For the fully trained group, we also assessed possible learning effects within this group from session one to session two by calculating a two-way repeated measures ANOVA on the MDEV values. Again, we found a main effect of secondary task condition, $F(3, 45) = 32.81$, $p < .001$. Also, we found a main effect of training, $F(1, 15) = 6.70$, $p = .021$. There was no interaction between the secondary task condition and training level. As can be seen in Figure 7 (left), the lane keeping / changing quality increased in all secondary task conditions.

Regarding secondary task performance (Figure 7, right), we again found substantial effects of both secondary task difficulty, $F(2, 30) = 87.02$, $p < .001$, and training, $F(2, 15) = 51.87$, $p < .001$, as well as a significant interaction between the two, $F(2, 30) = 3.69$, $p = .037$.

Table 2: Subjective assessment of LCT + CTT (critical tracking task) as obtained by the DALI; for between-groups comparison see rows “no training” vs. “LCT only training” vs. “full training (testing session)”; for comparison of training sessions see “full training (testing session)” vs. “full training (training session)”.

		baseline		LCT+CTT1		LCT+CTT2		LCT+CTT3	
scale	group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
global attention demand	no training	3.24	1.48	4.24	1.25	5.29	0.77	5.76	0.56
	LCT only training	3.21	1.48	4.36	1.08	5.21	0.80	5.93	0.27
	full training (testing)	2.50	1.15	3.88	1.20	4.38	1.02	5.31	0.87
	full training (training)	3.69	1.58	4.06	1.06	5.38	0.72	5.69	0.60
visual demand	no training	3.59	1.46	3.94	1.56	5.29	1.05	5.29	0.99
	LCT only training	3.21	1.53	4.21	1.12	5.07	1.00	5.64	0.63
	full training (testing)	2.38	0.96	3.38	1.31	4.25	1.44	4.88	1.54
	full training (training)	3.69	1.62	4.06	1.24	5.06	1.00	5.31	1.08
stress	no training	1.94	1.09	3.18	1.33	4.47	1.50	4.88	1.17
	LCT only training	2.57	1.50	3.29	1.64	4.14	1.61	4.36	1.34
	full training (testing)	1.81	1.17	2.94	1.34	4.00	1.63	4.38	1.75
	full training (training)	2.94	1.48	3.50	1.32	4.44	1.46	5.19	1.11
temporal demand	no training	1.71	1.40	3.18	1.59	4.47	1.74	4.82	1.59
	LCT only training	2.00	1.36	3.00	1.57	4.14	1.46	4.14	1.66
	full training (testing)	1.81	1.22	2.88	1.36	3.81	1.38	4.31	1.54
	full training (training)	2.19	1.42	3.06	1.48	4.19	1.83	4.56	1.71

Note: CTT1, CTT2, and CTT3 indicate three levels of difficulty in the task.

The analysis of the DALI’s subscales again produced a significant main effect of secondary task condition (all $p < .001$) on each scale. In addition, the scales global attention demand, $F(1, 15) = 27.92$, $p < .001$; visual demand, $F(1, 15) = 19.96$, $p < .001$; stress, $F(1, 15) = 39.12$, $p < .001$; and interference, $F(1, 15) = 11.04$, $p = .005$, showed significant main effects of training. For global attention demand, there also was a significant interaction, $F(1, 15) = 4.13$, $p = .027$.

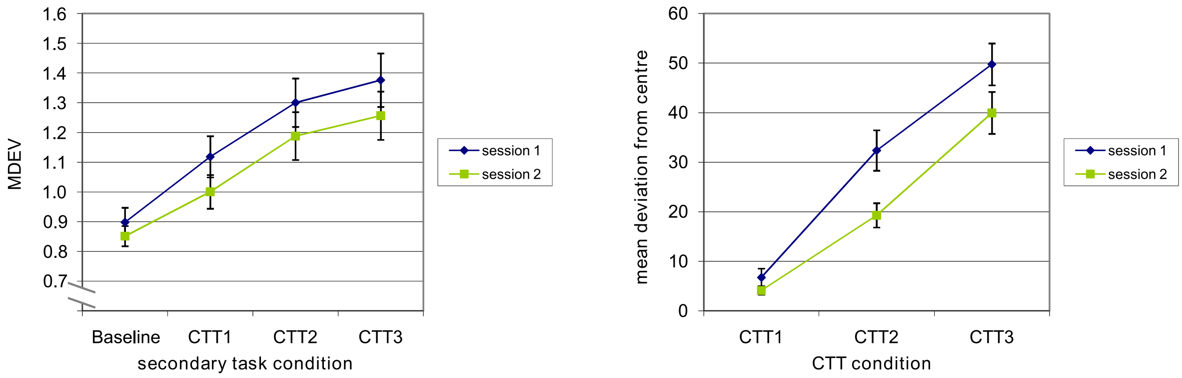


Figure 7: LCT (left) and CTT (right) performance for different testing sessions (error bars indicate standard error).

2.3 Discussion

We conducted this first experiment to find out whether any training effect occurs when repeatedly employing the same participants for LCT testing. The results clearly show that this effect exists, and that it is substantial. When comparing the different variations of training, it becomes obvious that especially dual-task training has a tremendous impact on performance. This impact, however, is not uniform. When looking at the SuRT as the secondary task, we find relevant differences between the groups in LCT performance, whereas SuRT performance does not seem to improve. In contrast, we find no distinction between the three training groups in LCT performance when the CTT is the secondary task, whereas CTT performance improves with the level of training. The comparison of the two sessions for the fully trained group, however, strongly supports the view that both primary and secondary task performance improve with the amount of training. This impression is further strengthened by subjective assessments of the participants' workload. Between groups as well as between sessions comparisons show on various scales that training lowers the level of experienced load significantly. These findings are somewhat disturbing, given that as yet there are no specifics in the ISO draft on how to deal with this issue. Although the amount of training within an experiment is allowed to vary to give each participant the chance to reach some sort of optimum performance, our results show that this just might not be enough. At the same time, it has to be acknowledged that the short training-testing interval chosen for this experiment is rather artificial. It remains to be proven that the effect of experience we found is stable over a period of time that makes it practically relevant.

3 Experiment II

After we were able to confirm short-term learning effects in LCT and secondary task performance, we tried to assess the durability of those effects. As LCT testing is usually not done on a daily basis, the assessment of long-term effects bears much more practical relevance. Therefore, we tested another group of participants a minimum of 4 months after their initial (full) training, and again compared them to a control group with no previous LCT experience. As the LCT itself was not sensitive to variations of SuRT difficulty in our first experiment, we chose to use only the CTT as the secondary task.

3.1 Method

3.1.1 Participants

Forty-eight participants took part in this experiment; three outliers had to be removed. All of the remaining 45 (age *M* = 24.8 years, *SD* = 4.2 years, 14 male, 31 female) were in possession of a valid driving license and drove a mean 10,950 km a year, *SD* 8,550 km. None had previous LCT experience. All of the student participants (39) received course credit; the remainder received monetary compensation.

3.1.2 Material

The material used was identical to the material in the first study. We again used the LCT, employing the same setup (desktop according to the ISO draft). As a secondary task we used the CTT again, with settings identical to those in the first experiment. The DALI was applied as well, although only after the testing session.

3.1.3 Procedure

Two groups of participants were employed in the experiment. One group served as a control, similar to in the first experiment. They completed one testing session with a familiarisation phase, baseline driving, and then combined LCT + CTT driving, with CTT conditions in randomised order. Each LCT drive was concluded with filling in the DALI. The second group consisted of participants who had already taken part in a small preceding study that required them to drive the LCT in conjunction with the CTT, with the respective prior familiarisation and baseline drives. In this sense, this group was comparable to the "full training" group of Experiment I. Data on these drives were available. Half of the participants were tested around 4 months, the other half approximately 7 months after their initial "training" (the reason being two different points of time for the preceding study). Testing was done in the same way as for the control group.

3.2 Results

3.2.1 LCT + CTT, between-groups comparison

To assess possible long-term learning effects, we first computed a two-way ANOVA for mixed designs on the MDEV values (Figure 8, left). The analysis revealed a main effect of secondary task condition, $F(3, 129) = 41.69$, $p < .001$, which is coherent with the results of our previous experiment. There was no significant effect of training, $F(1, 43) = 3.71$, $p = .061$, and also no interaction between the secondary task condition and the type of training.

We also calculated an ANOVA on the adaptive MDEV values. The results were close to those obtained for the normative MDEV, with a significant influence of secondary task condition, $F(3, 129) = 5.04$, $p < .001$, no significant effect of training, $F(1, 43) = 4.04$, $p = .051$, and no interaction between the secondary task condition and training.

Analysing secondary task performance (Figure 8, right), we found the anticipated main effect of secondary task difficulty, $F(2, 86) = 198.59$, $p < .001$. However, no significant

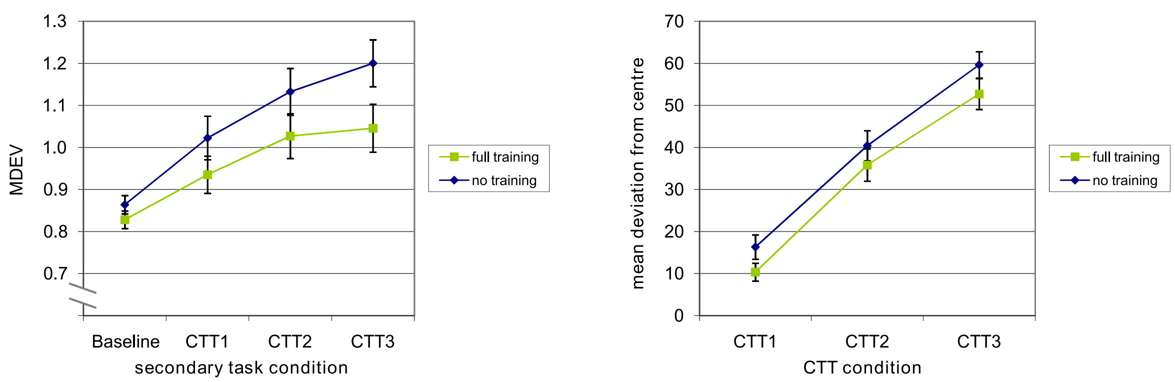


Figure 8: LCT (left) and CTT (right) performance for different training groups (error bars indicate standard error).

Table 3: Subjective assessment of LCT + CTT as obtained by the DALI (between-groups comparison)

		baseline		LCT+CTT1		LCT+CTT2		LCT+CTT3	
scale	group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
global att. demand	no training	4.86	1.04	4.95	1.17	5.68	0.48	5.95	0.21
	full training (testing)	4.00	1.10	4.62	0.97	5.24	0.70	5.76	0.44
visual demand	no training	4.26	1.42	4.83	1.19	5.48	0.73	5.78	0.52
	full training (testing)	3.77	1.23	4.82	0.91	5.23	0.81	5.68	0.57
stress	no training	2.61	1.41	3.74	1.32	4.52	1.27	4.96	1.22
	full training (testing)	2.50	1.10	3.45	1.06	4.50	1.10	4.68	1.04
temporal demand	no training	2.26	1.45	3.17	1.59	4.00	1.62	4.30	1.55
	full training (testing)	2.86	1.61	3.86	1.42	4.09	1.34	4.64	1.18

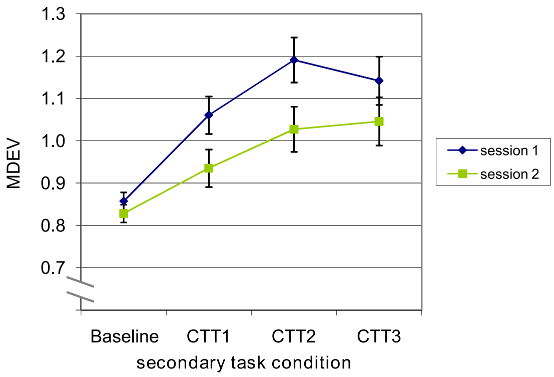
Note: CTT1, CTT2, and CTT3 indicate three levels of difficulty in the task.

effect of training was found, $F(1, 43) = 2.26$, $p = .140$. There was no interaction between the two.

The analysis of the DALI's subscales (Table 3) produced a significant main effect of secondary task condition (all $p < .001$) on each scale. In addition, for the scale global attention demand, we found a significant influence of training, $F(1, 41) = 5.79$, $p = .021$. Other scales showed no significant effect of training, and no interactions were found.

3.2.2 LCT + CTT, comparison of training sessions

Since we had two full data sets (initial training and testing) available for the fully trained group again, we compared this group's performance in the two sessions. The two-way repeated measures ANOVA on the MDEV showed a significant effect of secondary task condition, $F(3, 63) = 25.60$, $p < .001$. Also, there was a significant influence of training, $F(1, 21) = 14.84$, $p = .001$, on LCT performance (Figure 9, left). We also found a significant interaction between the secondary task condition and training, $F(3, 63) = 3.76$, $p = .015$. A similar picture emerged for the adaptive MDEV values. Again, there was a significant effect of secondary task condition, $F(3, 63) = 35.65$, $p < .001$, and a significant effect



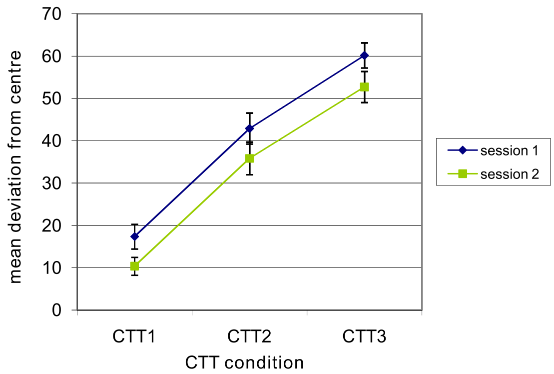


Figure 9: LCT (left) and CTT (right) performance for different testing sessions (error bars indicate standard error).

of training, $F(1, 21) = 16.21$, $p = .001$. However, there was no interaction between the secondary task condition and the type of training for the adaptive MDEV values.

Analysing secondary task performance (Figure 9, right), we found the anticipated main effect of secondary task difficulty, $F(2, 42) = 182.72$, $p < .001$. Also, a significant effect of training was found, $F(1, 21) = 20.34$, $p < .001$. There was no interaction between the two.

3.3 Discussion

With this second experiment, we tried to find out whether the effect of training found in our previous experiment is stable over a longer period of time. The results obtained clearly point in this direction. When comparing the two groups in Experiment II, we found small differences in LCT performance, using different metrics. Also the ratings in the DALI seem to support this view. Even more impressive are the results of the within-group comparison for the dual-task training group. Although first and second testing were at least 4 months apart, we still found significant improvements in primary and secondary task performance.

4 General discussion and conclusions

We carried out two experiments to find out if experience has an impact on LCT performance in the short and long run. Our first experiment shows that some form of training might indeed facilitate performance in the LCT as well as a given secondary task. Subjective ratings support this view, as trained participants reported lower levels of demand.

This finding is not too surprising, as the purpose of a familiarisation phase for any given task is to provide some sort of training to the participants so they can perform at their best in the experimental phase. What is disturbing is that the familiarisation phase in LCT studies is obviously not sufficient to reach the aspired optimum performance level. Still, this would be a minor problem if this was the case every time a participant took part in an LCT study, so comparability of these studies in general would be ensured. However, it seems that training effects are rather stable, so this optimum performance level is reached much faster in subsequent LCT encounters. Performance is better in the LCT and the secondary task, and reported demand is lower, even after longer periods of time without any exposure to those tasks. Given these results, the occurrence of this training effect might threaten the validity of LCT results.

The implications, however, are not straightforward. The findings could be used to support an “only novice participants” or “only expert participants” experimental design, regardless of any practical limitations that go with either of the two options. Using novice participants reflects a “first encounter” situation of a driver with a certain system. One might think of a rental car scenario, with a driver having to use an unfamiliar navigation system. There are no well-known, automated procedures available, and there is no a priori strategy of attention allocation that optimally fits the task at hand. An LCT assessment with novice participants would be just that, therefore providing insight into the maximum distraction or load a certain system can cause. Using expert participants, on the other hand, reflects the “everyday user” - a driver who knows how to complete a given task in a very efficient fashion, in terms of secondary task completion as well as attention allocation. Such a measurement captures more of an average distraction or load that the system will cause. Both approaches seem useful, as they reflect different use cases. Also, it is quite easy to imagine a system that is very distracting for a novice user but quite easy to handle for an expert, and another system that is easy or hard to operate for both. The difference between novice and expert would therefore be not just a matter of generally higher levels of distraction. One would rather expect some sort of interaction between the level of experience and distraction, making the assessment of both novice and expert drivers an appealing approach.

Some final remarks have to be made regarding the two different secondary tasks that were used in the experiments. It has to be acknowledged that the SuRT, which allows for interruption and easy resumption, more closely resembles possible in-vehicle tasks such as the operation of a navigation system, than the CTT does. There is hardly a task to be found that requires continuous attention like the CTT does. However, the CTT is the task that the LCT is better able to assess. One might be inclined to conclude that the LCT is not useful in assessing distraction caused by available systems, and that any learning effect found with the LCT + CTT combination probably would not apply to a LCT + regular in-vehicle task setup. Still, it is possible to argue otherwise. Although many in-vehicle tasks do not necessarily require the devotion of continuous attention, it is nevertheless possible to do so. Even though most systems allow for a strategy of interruption and resumption, there is no reason to assume a priori that users will follow such a strategy. Just as experienced and inexperienced users are equally relevant in the assessment of distraction, it is equally relevant and necessary to assess both “optimal” and “suboptimal” use of a system. If there is interest in the maximum amount of distraction a system might elicit, it would even be necessary to instruct the participants to follow a “continuous attention devotion” strategy, as there might be users in the real world doing just that. So, although there might be no system currently available that requires strategies of attention allocation comparable to the CTT, it can still be assumed that there are a substantial number of drivers on the roads who employ such deficient strategies. In this regard, the CTT is a useful secondary task to emulate this scenario, and the LCT is an important tool to assess its distractive impact.

References

Angell, L., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W., et al. (2006). *Driver workload metrics project: Final report (DOT HS 810 635)*.

Bahrick, H., Noble, M., & Fitts, P. (1954). Extra-task performance as a measure of learning a primary task. *Journal of Experimental Psychology, 48*, 298-302.

Bahrick, H., & Shelly, C. (1958). Time sharing as an index of automatization. *Journal of Experimental Psychology, 56*, 288-293.

Baumann, M., Keinath, A., Krems, J. F., & Bengler, K. (2004). Evaluation of in-vehicle HMI using occlusion techniques: Experimental results and practical implications. *Applied Ergonomics, 35*, 197-205.

Breuer, J., Bengler, K., Heinrich, C., & Reichelt, W. (2003). Development of advanced driver attention metrics (ADAM). In H. Strasser et al. (Eds.), *Quality of work and products in enterprises of the future* (pp. 37-39). Stuttgart: Ergonomia.

Bruyas, M.-P., Brusque, C., Auriault, A., Tattegrain, H., Aillerie, I., & Duraz, M. (2008). Impairment of lane change performance due to distraction: Effect of experimental contexts. In *Proceedings of the European conference on human centred design for intelligent transport systems*, Lyon, France, pp. 89-100.

Burns, P. C., Trbovich, P. L., McCurdie, T., & Harbluk, J. L. (2005). Measuring distraction: Task duration and the lane-change test (LCT). In *Proceedings of the Human Factors and Ergonomics Society 49th annual meeting*, Orlando, FL, pp. 1980-1983.

Damos, D. L., & Wickens, C. D. (1980). The identification and transfer of timesharing skills. *Acta Psychologica, 46*, 15-39.

Ebbinghaus, H. (1971). *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie* [Memory: A contribution to experimental psychology]. Darmstadt: Wissenschaftliche Buchgesellschaft. (Original work published 1885)

Engström, J., & Markkula, G. (2007). 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*, Stevenson, WA, pp. 199-205.

Harbluk, J. L., Burns, P. C., Lochner, M., & Trbovich, P. L. (2007). 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*, Stevenson, WA, pp. 16-22.

Harbluk, J. L., Mitroi, J. S ., & Burns, P. C. (2009). Three navigation systems with three tasks: Using the lane-change test (LCT) to assess distraction demand. In *Proceedings of the 5th international driving symposium on human factors in driver assessment, training, and vehicle design*, Big Sky, MT, pp. 24-30.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: Elsevier.

ISO 16673. (2007). *Road vehicles-Ergonomic aspects of transport information and control systems-Occlusion method to assess visual demand due to the use of in-vehicle systems*. Geneva, Switzerland: ISO.

Jahn, G., Krems, J. F., & Gelau, C. (2009). Skill acquisition while operating in-vehicle information systems: Interface design determines the level of safety-relevant distractions. *Human Factors, 51*, 136-151.

Jahn, G., Oehme, A., Krems, J. F., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F, 8*, 255-275.

Keinath, A., Baumann, M., Gelau, C., Bengler, K., & Krems, J. F. (2001). Occlusion as a technique for evaluating in-car displays. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics-Aerospace and transportation systems* (pp. 391-397). Aldershot, UK: Ashgate.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data (HS 810 59)*.

Kuratorium für Verkehrssicherheit. (2008). Road accident statistics 2007. In *Reihe Verkehr in Österreich. Heft 40*, Wien.

Lansdown, T. C. (2001). Causes, measures and effects of driver visual workload. In P. A. Hancock & P. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 351-369). Mahwah, NJ: Erlbaum.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. In H. Strasser et al. (Eds.), *Quality of work and products in enterprises of the future* (pp. 57-60). Stuttgart: Ergonomia.

Miura, T. (1986). Coping with situational demands: A study of eye movements and peripheral vision performance. In A.G. Gale et al. (Eds.), *Vision in vehicles-II* (pp. 126-137). Amsterdam: Elsevier.

New Zealand Ministry of Transport. (2007). *Reported injury crashes 2006*.

Pauzie, A., & Pachiaudi, G. (1997). Subjective evaluation of the mental workload in the driving context. In T. Rothengatter & E. Carbonell (Eds.), *Traffic & transport psychology: Theory and application* (pp. 173-182). Oxford: Pergamon.

Petzoldt, T., Bär, N., & Krems, J. F. (2009). Gender effects on lane change test (LCT) performance. In *Proceedings of the 5th international driving symposium on human factors in driver assessment, training, and vehicle design*, Big Sky, MT, pp. 90-96.

Rognin, L., Alidra, S., Val, C., Lescaut, A., & Chalandon, X. (2007). Driver strategies when interacting with information and comfort systems. In *Proceedings of the 20th international technical conference on the enhanced safety of vehicles*, Lyon, France, Paper number 07-0272.

Senders, J. W., Kristofferson, A. B., Levison, W. H., Dietrich, D. W., & Ward, J. L. (1967). The attentional demand of automative driving. *Highway Research Record, 195*, 15-33.

Shinar, D., Tractinsky, N., & Compton, R. (2005). Effects of practice, age, and task demands on interference from a phone task while driving. *Accident Analysis and Prevention, 37*, 315-326.

Starry, C. (2001). The emerging in-vehicle intelligent transportation systems market. *Business Economics, 32*, 49-55.

Stutts, J. C., Reinfurt, D. W., Staplin, L., & Rodgman, E. A. (2001). *The role of driver distraction in traffic crashes*. Report prepared for AAA Foundation for Traffic Safety, Washington, DC.

Weir, D. H., Kwok, A. M., & Peak, J. L. (2007). Exploratory lane change test experiments in support of ISO WG8 LCT task force. *Presentation given at the meeting of the Task Force Calibration Task*, Milano, Italy.

Young, K., Regan, M., & Hammer, M. (2003). *Driver distraction: A review of the literature*. Monash University, Accident Research Centre, Melbourne, Australia.

Police: Driver distracted by sex toy

An Elmwood Place police officer who stopped a car because it had illegally tinted windows received a bit of a shock when he looked inside. Officer Ross Gilbert said the driver, Colondra Hamilton, a 36-year-old Downtown resident, was sitting with her pants unzipped and a sex toy in her lap.

He said Hamilton told him she was using the toy while watching a sex video on a laptop computer that a passenger in the front seat held up so she could see it.

Gilbert charged her with "driving with inappropriate alertness" and having illegal tinted windows, according to the traffic ticket. (...)

(Cincinnati Enquirer, August 25th, 2010)

The critical tracking task - A useful method to assess driver distraction?

Tibor Petzoldt, Hanna Bellem and Josef F. Krems

under review for *Transportation Research Part F*

Abstract

The assessment of the distractive potential of new in-vehicle information systems has become an important issue, given their ever growing distribution in recent years. An easy-to-use method that might be a candidate to assess this distraction is the critical tracking task (CTT). The CTT requires an operator to bring a bar that is displayed on a computer screen and tends to leave a predefined target position back to this target position by pressing the respective keys on a keyboard. In this paper, we report on four experiments that try to investigate the CTT's potential as a tool to measure distraction. We employed artificial as well as realistic secondary tasks of varying demand that were to be assessed with the CTT. The results show that the CTT is able to differentiate between different levels of demand elicited by the secondary tasks, and that the results obtained correspond with the a-priori assumptions about the respective secondary tasks' distractive potential.

Keywords: Critical tracking task; Driver distraction; In-vehicle information systems; Evaluation methods

1 Introduction

Driver distraction, with its causes as well as its consequences, has received more and more attention in recent years. Webster's dictionary picked "distracted driving" as its word of the year for 2009 ("Distracted-driving campaign", 2010). The US Transportation Secretary announced an "administration wide effort to combat distracted driving" (US Department of Transportation, 2009). The term has risen to prominence for good reasons. Driver distraction has been shown to have detrimental effects on a variety of driving related variables (see Regan, Lee, & Young, 2009 for an overview), and appears to be a relevant factor in traffic accidents (Stutts, Reinfurt, Staplin, & Rodgman, 2001; Kuratorium für Verkehrssicherheit, 2008; New Zealand Ministry of Transport, 2010). In the US, driver distraction was reported to have been involved in 16% of all fatal crashes in 2008 according to data from the Fatality Analysis Reporting System (FARS), and an estimated 21% of injury crashes were reported to have involved distracted driving, according to data from the General Estimates System (GES; Ascone, 2009). Especially so called "technology-based distraction" (Young, Regan, & Hammer, 2003) has become an issue, as in-vehicle information systems (IVIS) and other related devices have become increasingly popular (Starry, 2001). Pickrell and Ye (2010) estimated that there were 672,000 vehicles driven by people using hand-held cell phones at a typical daylight moment in the US in 2009. In addition, they report that about 0.6% of drivers were text-messaging or visibly manipulating other hand-held devices.

Given the growing distribution of in-vehicle devices, it is vital to assess the distractive nature of each of these systems. Large field tests (e.g. Karlsson, Rämä, Alonso, Engelbrektsson, Franzén, Henar Vega, et al., 2009) are able to accomplish this with a high ecological validity. However, the cost associated with these tests is immense, which makes them impossible to be used on a regular basis. Such projects can provide information about the effects these systems might cause in general, but they are limited in their use as a tool to differentiate between specific systems, brands, or functions. Studies in a driving simulator are somewhat more suitable in this regard, but still, the effort is rather high. Therefore, researchers have come up with a number of easy-to-use methods to coarsely measure the distraction elicited by an IVIS. The peripheral detection task (PDT; Jahn, Oehme, Krems, & Gelau, 2005) tries to assess cognitive and visual distraction by making use of the fact that visual and cognitive load can narrow the driver's functional field of view (Miura, 1986). The occlusion method (Senders, Kristofferson, Levison, Dietrich, & Ward, 1967;

see also Baumann, Keinath, Krems, & Bengler, 2004; Keinath, Baumann, Gelau, Bengler, & Krems, 2001) targets visual distraction, especially focusing on (non-)interruptability of secondary tasks. This method has already become an ISO-standardised procedure (ISO 16673, 2007). A third method, the lane change test / task (LCT; Mattes, 2003), has also been developed to address the issue of visual distraction. Here, the task is to control a vehicle on a three lane road at a constant speed, while repeatedly performing lane changes as instructed by signs at the side of the road. The degradation of driving performance when operating an additional secondary task serves as a measure of distraction (Mattes & Hallén, 2009). A task that, in this context, only has been used as an IVIS-surrogate is the critical tracking task (CTT; Jex, McDonnell, & Phatak, 1966). In this function, the task has been subject to some criticism, as the continuous monitoring and input required are rather uncommon in most IVIS (e.g. Petzoldt, Bär, Ihle, & Krems, 2011). However, continuous monitoring and input are requirements of the driving task itself. Given this similarity to central aspects of the driving task, the CTT might function as a surrogate for actual driving. Whether it holds potential as a method to measure the distraction demand of in-vehicle tasks in this function has to be investigated.

1.1 The critical tracking task

The CTT has first been described by Jex et al. (1966). In their rather technical account, they explained the CTT as a task “in which a human operator is required to stabilize an increasingly unstable first-order controlled element up to the critical point of loss of control” (p. 138). Others have likened the task to “balancing a stick on one’s fingertip, with the stick’s length decreasing with time” (Burns & Moskowitz, 1980; p. 261). Basically, most of the CTT’s incarnations use some form of virtual bar on a computer screen that tends to leave a predefined target position. Participants use joystick or keys to bring the bar back to the target position, a task that becomes increasingly difficult over time. The level of instability is represented by a λ-value. This λ-value increases constantly. The value at which control is lost is recorded as a performance measure. The task has been found to correlate substantially with subjective ratings of workload (Rehman, Stein, & Rosenberg, 1983; Rosenberg, Rehman, & Stein, 1982), allowing for the assumption that the CTT is a valid method to assess demand on a general level.

The CTT has been used a lot in research on various kinds of cognitive and psychomotor impairment. Most of these studies have been concerned with medication, drug and alcohol related performance decrements (Barnett, Licko, & Thompson, 1985; Burns & Moskowitz, 1980; Klein & Jex, 1975; Ramaekers, Uiterwijk, & O'Hanlon, 1992; Ramaekers, Muntjewerff, van Veggel, Uiterwijk, & O'Hanlon, 1998; Ramaekers, Kauert, van Ruitenbeek, Theunissen, Schneider, & Moeller, 2006; Ramaekers, Moeller, van Ruitenbeek, Theunissen, Schneider, & Kauert, 2006). Elmenhorst, Elmenhorst, Luks, Maass, Mueller, Vejvoda et al. (2009) assessed the impairment caused by sleep deprivation. In their review on candidate procedures for test batteries to be used to assess specific operational tasks, O'Donnell, Moise and Schmidt (2005) argued that the successful operation of the CTT especially requires focussed attention and visual motor control, but also working memory, time / velocity estimation and situation awareness. It can be assumed that the CTT as a single task puts a form of demand on the operator that is fairly similar to various aspects of driving (see Groeger, 2000).

The use of the CTT is, however, not limited to the single task setup. Jex (1967) explains the task's use as a primary task in conjunction with other secondary tasks to assess the secondary task's workload. He argued that a higher level of workload should result in sooner loss of control, represented by a smaller critical λ. In addition, he also proposed the option of employing the CTT as a secondary task. The instability should be set at a constant subcritical level, demanding frequent, but not continuous attention. With this setup, he argued, the CTT might serve as a source of continuous workload that is reflected by the performance on any primary task. Following this approach, Burke, Gilson and Jagacinski (1980) studied the parallel use of two CTTs, as well as an implementation of the task that is two-dimensional, i.e. instable in two directions. In a similar account, Derrick (1988) tried to assess the validity of different measures of workload by letting participants operate various tasks (among them the CTT) in parallel. In a study reported by Wickens and Kessel (1980), the CTT was used as a source of distraction. It was found that the detection of dynamic system failures was impaired by the secondary task. Wickens, Braune and Stokes (1987) assessed age differences in the speed and capacity of information processing by using the CTT in conjunction with different versions of the Sternberg memory search task.

Although Jex (1967) clearly differentiates between primary and secondary task, this practice is not always found in the measurement of driver distraction (still, for simplicity, in this paper we use the terms "primary task" for the CTT and "secondary task" for any concurrent task). For the LCT, for example, it is explicitly required that no instructions shall be offered to participants on how to prioritize between primary and secondary tasks. The CTT is capable of that as well. Employing the "constant subcritical instability" setup, and using the bar's deviation from the target position as a performance measure, it is possible to achieve a continuous measurement of distraction. Wickens and Kessel (1980, see also Wickens et al., 1987) reported on the calculation of a tracking error as performance measure. Such a metric appears to have an advantage over the previously described critical λ. When using the critical λ, trials have to be performed as long as the participant loses control over the task. This may vary dramatically, resulting in differences in task duration and number of secondary task trials. As the secondary task is usually not a continuous one, this may give rise to a variety of problems. With the continuous measurement, on the other hand, a stop criterion can be set through the secondary task (operation of both tasks until the secondary task is finished) as well as the CTT (operation of both tasks for a predefined duration). So, the length of trials is more independent from participants' individual skills (with due acknowledgement of the fact that secondary task duration may also depend, at least partially, on individual skills).

Considering the available research on the CTT to date, it appears that the task can be a promising candidate to assess driver distraction. However, specific evidence for this claim is still missing. The four experiments described in this paper are intended to provide some deeper insight into the capabilities and limitations of this method. Goal of Experiments I and II was a general statement about the CTT's potential as a method to assess driver distraction, using valid, but simple secondary tasks. If the CTT succeeded here, a more differentiated approach (artificial tasks tapping into specific aspects of the CTT) as well as a more applied approach (real in-vehicle information systems) could be pursued in later experiments to verify the results. Experiment III followed this more differentiated approach, employing standardised artificial secondary tasks that have been described in studies that assessed the LCT. Finally, Experiment IV aimed at the practical application of the CTT by comparing the amount of distraction caused by different navigation system tasks.

2 Experiment I

In a first experiment, we tried to find some general evidence that the CTT could serve as a method to assess distraction. To accomplish this, it was necessary to find a secondary task that can claim face validity in being representative of in-vehicle tasks, but at the same time is simple to employ and manipulate.

One major aspect of distraction caused by IVIS is the legibility of information. The European statement of principles on human-machine interface (Commission of the European Communities [CEC], 2008), refers to legibility as an important design aspect of IVIS, pointing to international standards that specify further guidelines. The respective ISO standard (ISO 15008, 2009) lists numerous features of information systems that have to be considered when developing such systems, different aspects of legibility among them. Stevens, Quimby, Board, Kersloot and Burns (2002) regarded brightness, contrast, resolution, character / character spacing, font and case as relevant variables for display legibility. For navigation systems, character size, combinations of colours, background luminance, map orientation and amount of information have been identified as crucial factors (Kimura, Marunaka, & Sugiura, 1997). Pauzie (2002) reported on the influence the size of the displayed information has on glance frequency and duration. A small character size appears to force especially older users to look at the screen significantly longer in order to read the text. This age related difference disappears with larger characters. Text-background colour combinations affect legibility, but also the pleasantness of a display (Greco, Stucchi, Zavagno, & Marino, 2008). For our experiment, we chose to use character size and character-background colour combination as factors. If the CTT is useful for the assessment of driver distraction, the differences in legibility of the information should be reflected in CTT performance. More specifically, we hypothesised that larger characters and higher character-background contrasts lead to better CTT performance than small characters and low contrasts.

2.1 Method

2.1.1 Participants

Twenty-four students of Chemnitz University of Technology took part in this experiment. Eighteen of them were female, 6 male, with a mean age of 22.6 years (SD = 4.8).

2.1.2 Material

Critical tracking task (CTT) The CTT employed in this experiment (and all follow-ups) used a horizontal bar that continuously left its target position at the centre of the screen (Figure 1). Positions at the boundaries equalled deviation values of 100 and -100, the centre position was defined as 0. The participants were able to control the bar by pressing

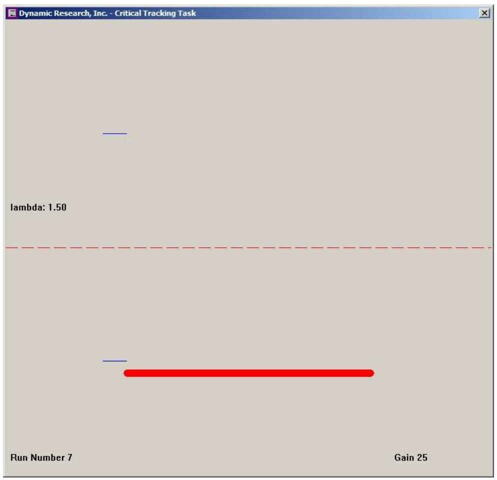


Figure 1: Critical tracking task (CTT) example screen.

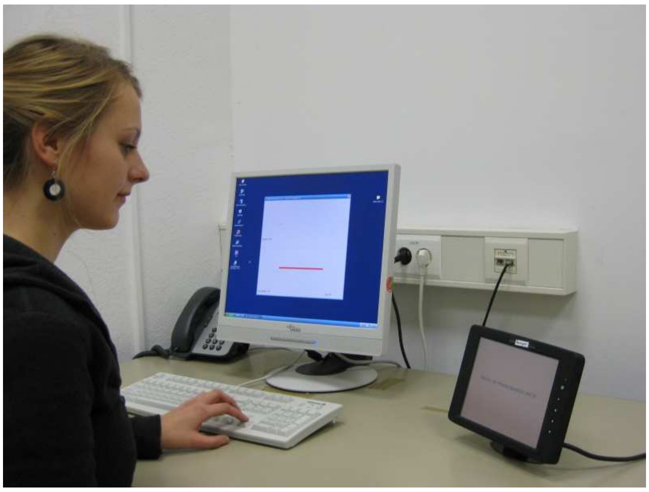


Figure 2: Experimental setup.

the up- and down-keys on the keyboard. When the bar's deviation from the centre became too large (values over 40 / under -40), it changed its colour from black to red, to indicate the need of immediate action. The position of the bar relative to the centre was recorded, and, based on that data, the mean deviation was computed as a performance metric. The task was presented on a 19" screen, positioned centred in front of the participants (Figure 2). The task was located at the centre of the screen and covered approximately 50% of it. The level of instability was set at a medium level of difficulty (as found in pre-tests, λ = 1.5).

Reading task The reading task made use of short questions that had to be read and answered by the participants. The questions were very easy to answer (e.g. "Name a country beginning with D" - with "Deutschland" being quite salient as a possible answer for Germans), as their main purpose was to assure that participants read the presented information. Their construction followed always the same pattern - "Name [something] beginning with [letter]", to ensure comparability of reading times between items. Four conditions were generated from different foreground-background colour contrasts (dark blue on yellow vs. red on violet) and different characters sizes (visual angle approx. 0.3° vs. 0.1°): high foreground-background colour contrast with large characters, high foreground-background colour contrast with small characters, low foreground-background colour contrast with large characters and low foreground-background colour contrast with small characters. The number of questions read was recorded as a reading performance metric. The questions were presented on an 8.37" screen that was placed to the right of the participants, in a position where an aftermarket navigation system would typically be put. The presentation was controlled by the experimenter. As soon as a question was answered, the next one was displayed.

2.1.3 Procedure

First, participants received instructions on the CTT and the reading task, followed by a short CTT training. After recording a baseline trial with the CTT as a single task, participants had to perform the CTT and the reading task simultaneously in four trials of 3 min each. The order of the four different legibility conditions was balanced over participants.

2.2 Results

In Figure 3, CTT standard deviation values are displayed for the five conditions. The baseline condition is only included for completeness. As can be seen, all four dual-task conditions produce higher deviations than the baseline. A two-factorial repeated measures ANOVA revealed that character size had a significant impact, $F(1, 23) = 32.87$, $p < .001$, $\eta_p^2 = .59$, with smaller characters having caused higher deviations. There was no significant main effect for colour contrast, $F(1, 23) = 1.33$, $p = .260$, $\eta_p^2 = .06$, and no interaction, $F(1, 23) = .56$, $p = .463$, $\eta_p^2 = .02$.

Following the procedure proposed by Harbluk, Mitroi and Burns (2009), we also computed the mean deviation per average task by dividing the mean deviation by the number of

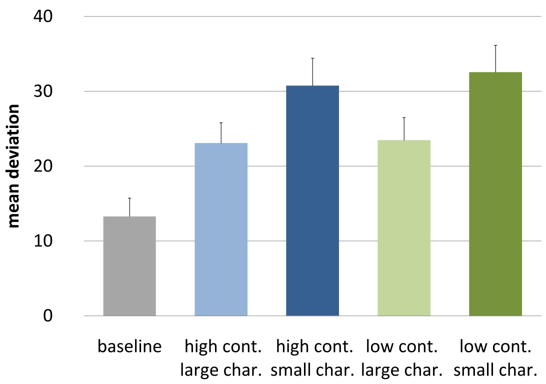


Figure 3: Mean deviation from the centre position (cont. = contrast, char. = character).

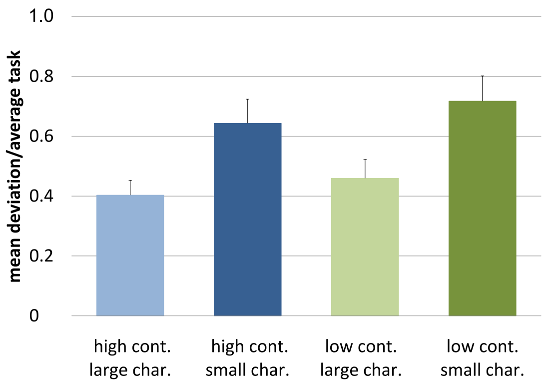


Figure 4: Mean deviation from the centre position, divided by number of completed secondary task trials during the experimental trial (cont. = contrast, char. = character).

questions read during the 3 min run (Figure 4). Again, a two-factorial repeated measures ANOVA showed a significant main effect of character size, $F(1, 23) = 38.12$, $p < .001$, $\eta_p^2 = .60$. Colour contrast appeared to influence mean deviation per average task as well, $F(1, 23) = 10.73$, $p = .003$, $\eta_p^2 = .32$. No interaction was found, $F(1, 23) = .10$, $p = .755$, $\eta_p^2 < .01$.

Since the 3 min duration of a single trial was chosen rather arbitrarily (following the LCT procedure; Mattes, 2003), we also tried to assess whether a shorter duration might produce the same results. Therefore, we calculated CTT standard deviation values for the first and second half (= 90 sec each) separately. When comparing the two segments and the overall performance, we found highly significant correlations between all three scores (first & second half: $r = .94$; first half & overall: $r = .98$; second half & overall: $r = .99$).

2.3 Discussion

Our first experiment provides support for the assumption that the CTT might be a useful method to assess driver distraction evoked by IVIS. The task was able to differentiate between small and large characters solidly. However, it failed to directly capture the expected effect of colour combination. Statements of the participants after the experiment suggest that they did not experience substantial differences between the two colour conditions during the trials. The missing impact on CTT performance might be attributed to a general absence of a meaningful effect of colour combination in the employed setup rather than an inability of the CTT to capture such an effect. The effect found when incorporating secondary task performance into the performance metric, however, suggests otherwise. It appears that the effect of colour contrast was not entirely absent, but qualitatively different from the one of character size. In any case, the inclusion of secondary task data can give additional insight into the nature of the distraction caused by different variables.

Participant's statements after completion of the experiment provided further valuable information. Many complained about the length of the trials, arguing that the dual-task situation was rather demanding. Especially at the end of trials, attention and concentration were decreasing. The three minute trial duration was chosen following the LCT procedure (Mattes, 2003). However, the LCT requires slightly different "subtasks" (lane keeping, lane change from middle to left lane, from middle to right lane etc.) to be completed within the task, so this duration is necessary to have a balanced measurement.

In context of the CTT, task duration is a result of a more or less arbitrary decision, as there are no “subtasks” involved. So, as long as the measurement can be assumed to be stable, a shorter duration would be justified. Our analysis provides evidence for this assumption, as performance within shorter time segments correlated nearly perfect with overall performance. For Experiment III, we took this into account.

3 Experiment II

In our second experiment, we wanted to complement Experiment I by getting further evidence for the CTT’s capability of reflecting distraction level differences in aspects relevant for HMI design. We again referred to the European statement of principles (CEC, 2008) to identify display position as another important factor. As one of its installation principles, the document states that “Visual displays should be positioned as close as practicable to the driver’s normal line of sight.” (p. 11). An explanation has been given by Stevens and colleagues (2002), who pointed out that “visual displays positioned close to the driver’s normal line of sight reduce the total eyes-off-the-road time relative to those that are positioned further away” (p. 23). Since driving is primarily a visual task, it appears obvious that eyes-off-the-road time is directly linked to crashes (Green, 1999). As a potential tool to assess driver distraction, the CTT should be able to distinguish between displays that are placed in different positions with regard to the line of sight. We hypothesised that a secondary task display closer to the CTT display will result in lower CTT deviation values than a display further away.

3.1 Method

3.1.1 Participants

Twenty students of Chemnitz University of Technology took part in this experiment, 15 female and 5 male, with a mean age of 22.0 years (SD = 2.1).

3.1.2 Material

Critical tracking task (CTT) The CTT employed in this experiment, as well as the overall setup (screen size & position, operation of task etc.), were identical to Experiment I.

Visual task The task that had to be dealt with on the secondary task screen was the so called Surrogate Reference Task (SuRT; Mattes & Hallén, 2009). It required the participants to scan stimulus displays for the one stimulus that differed from the others surrounding it. Target and distractors were white circles in front of a black background. Participants responded by moving a vertical grey indicator bar to the position of the identified target and pressing the enter key for confirmation, followed by the next display. For this experiment, we used a rather difficult version of the task (defined by distractor size and number of indicator bar sections). The task was presented on an 8.37” screen to the right of the participant. The indicator bar was controlled by using a standard keyboard. The position of the screen was varied between the two experimental conditions. Whereas in the “close” condition, the screen was set up directly right of the CTT screen (about 30° angle between the centre of the CTT screen and the centre of the SuRT screen), it was moved about 60 cm further to the right for the “distant” condition (about 75° angle).

3.1.3 Procedure

Participants received instructions on the CTT and the SuRT, followed by a short CTT training. A baseline trial with the CTT as a single task was recorded. Two trials (one each condition) of 3 min each followed, in which the participants operated the CTT and the SuRT in parallel. The order of the two different conditions was balanced over subjects.

3.2 Results

In Figure 5, CTT standard deviation values are displayed for the baseline and the two experimental conditions. The baseline condition is only included for completeness. It is already visible that the closer position of the SuRT screen to the CTT screen went along with lower deviation values in the CTT. A t-test for dependent measures confirmed this impression, $t(19) = 3.78$, $p = .001$, $d = .45$.

We as well analysed the mean deviation per average task, by dividing the mean deviation by the number SuRT-screens completed during the 3 min run (Figure 6). We found again significantly lower values for the close screen condition, $t(19) = 2.75$, $p = .013$, $d = .63$.

As in Experiment I, we assessed whether a shorter CTT trial duration produced the same results as the full trial. When comparing the deviation values for the first and second half (= 90 sec each) of the measurement and the overall performance, we again found highly

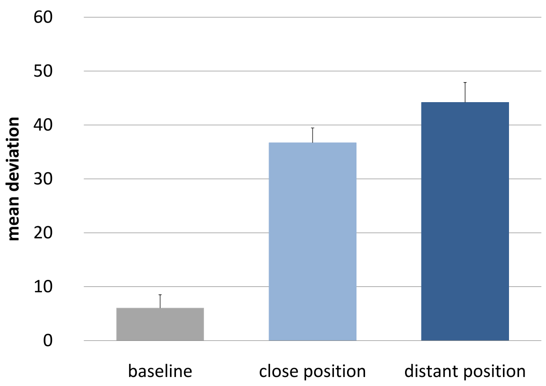


Figure 5: Mean deviation from the centre position.

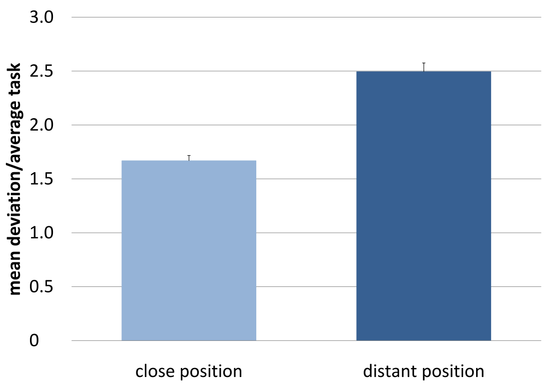


Figure 6: Mean deviation from the centre position, divided by number of completed secondary task trials during the experimental trial.

significant correlations between all three scores (first & second half: r = .90; first half & overall: r = .98; second half & overall: r = .97).

3.3 Discussion

In our second experiment, we found further evidence for the assumption that the CTT is a promising candidate for measuring driver distraction elicited by IVIS. The task was able to differentiate between different levels of distraction caused by different positions of the SuRT screen. As the compound measure of mean deviation per average task produced the same pattern of results, it seems that participants did not compensate for the distant location of the screen by solving less SuRT trials. It appears that performance decrements were only to be observed for the CTT.

With regard to the length of the trials, we again found that a shorter duration would lead to the same results. We consider the results of this and the previous experiment sufficient evidence to shorten the length of experimental trials to 90 sec in further studies.

4 Experiment III

In our third experiment, we tried to use a more standardised approach to uncover specific capabilities of the task. Visual distraction has been mentioned previously as a critical factor in accident causation. However, visual distraction is not the only influential form of inattention. Cognitive distraction has been shown to influence visual behaviour and vehicle control (e.g. Harbluk, Noy, Trbovich, & Eizenman, 2007) as well. The assessment of the CTT's ability to distinguish between different levels of visual and cognitive distraction appeared to be the logical next step towards an appropriate judgement of the task's potential. Since the operation of the CTT requires a lot of visual attentional resources, we predicted that a higher demand in visual attention in a secondary task would be reflected in degraded CTT performance. The CTT's cognitive demand, however, is rather low. Controlling the task is fairly easy, no complex cognitive processes are involved. Still, handling a dual-task situation, regardless of the respective task's nature, is a cognitive demanding process in itself. Therefore, we hypothesised that variations in cognitive demand in a secondary task would influence CTT performance as well, though most certainly less pronounced than for variations in visual demand. In our approach, we followed Mattes and Hallén (2009), who tried to validate the LCT in a similar way.

4.1 Method

4.1.1 Participants

Twenty-four students of Chemnitz University of Technology took part in this experiment. Three datasets had to be excluded from the final analysis (one dataset corrupted, two participants statistical outliers on various measures). The mean age of the remaining participants was 21.9 years (SD = 3.2), with 16 of them female, 5 male.

4.1.2 Material

Critical tracking task (CTT) The CTT employed in this experiment, as well as the overall setup (screen size & position, operation of task etc.), were identical to the previous experiments.

Cognitive task The cognitive task used in the experiment closely resembled the one that was employed by Mattes and Hallén (2009) for validating the LCT. We defined a "cognitive easy" and a "cognitive difficult" condition. For the "cognitive easy" task, participants had to count forwards in steps of two from 212 on, or in steps of five from 45 on. In the "cognitive difficult" version, the task was to count backwards in steps of six from 831 on, or in steps of seven from 581 on.

Visual task The visual task we employed in the experiment was again the SuRT. Different from Experiment II, we defined two different conditions (later referred to as "visual easy" and "visual difficult") that were identical to the one used by Mattes and Hallén (2009), by varying distractor size and number of indicator bar sections. The task was presented on an 8.37" screen to the right of the participant. The indicator bar was controlled by using a standard keyboard. Position of screen and keyboard matched the requirements of the LCT ISO draft.

Instrument for subjective ratings To confirm our a-priori assumptions about the cognitive and visual tasks, we administered a questionnaire to obtain subjective assessments of the demands the different tasks represent. We used a scale that we derived from the classical NASA-TLX (Hart & Staveland, 1988), letting participants indicate on a continuous scale a) the mental demand, b) the visual demand and c) the temporal demand represented by the task, d) the effort necessary and e) the frustration experienced. In their rating, participants were requested to assess the dual task situation (CTT + respective task).

4.1.3 Procedure

First, participants received instructions on the CTT, followed by a short CTT training. Then, we recorded two baseline trials with the CTT as a single task. After that, participants were informed about the first of the secondary tasks (either cognitive or visual). They completed four experimental trials of 90 sec, two for each difficulty level. The trials were blocked for difficulty, with the blocks in randomised order, each block preceded by a short familiarisation phase. The same procedure was then used for the remaining secondary task. The subjective rating scale was filled at the end of each respective task condition.

4.2 Results

Prior to the actual analysis, we tested whether it was justifiable to merge the two separate trials for each condition into one single score. We found no significant differences for any of the combinations, so we computed one single value for baseline, "cognitive easy", "cognitive difficult", "visual easy" and "visual difficult", respectively. Furthermore, we assessed the correlation between the two trials for each condition. With r = .70 for the two baselines, and $r > .85$ for the other pairs, we found highly significant correlations that confirm a high stability of the measurement.

4.2.1 CTT

In Figure 7, CTT standard deviation values are displayed for the five conditions. The baseline condition is again only included for completeness. A two-factorial ANOVA for repeated measures revealed that task type (cognitive vs. visual), $F(1, 20) = 86.27$, $p < .001$, $\eta_p^2 = .81$, as well as task difficulty, $F(1, 20) = 28.47$, $p < .001$, $\eta_p^2 = .59$, had a significant influence on CTT performance. Visual tasks, as well as difficult tasks, resulted in higher deviations than cognitive and easy tasks. We also found a significant interaction between task type and task difficulty, $F(1, 20) = 10.44$, $p = .004$, $\eta_p^2 = .34$. As Figure 7 shows, for the visual task, raising task difficulty had a higher impact on CTT performance than for the cognitive task.

Again following the procedure proposed by Harbluk et al. (2009), we computed the mean deviation per average task by dividing the mean deviation by the number of secondary task trials completed (SuRT screens processed or counts completed) during the 90 sec run (Figure 8). It is not surprising to again find a significant effect of task type, $F(1, 20) = 87.28$, $p < .001$, $\eta_p^2 = .81$, as task duration for processing a SuRT screen is obviously longer than a single count. Task difficulty had also a major influence on the deviation per average task, $F(1, 20) = 94.61$, $p < .001$, $\eta_p^2 = .83$. Again, there was a significant interaction

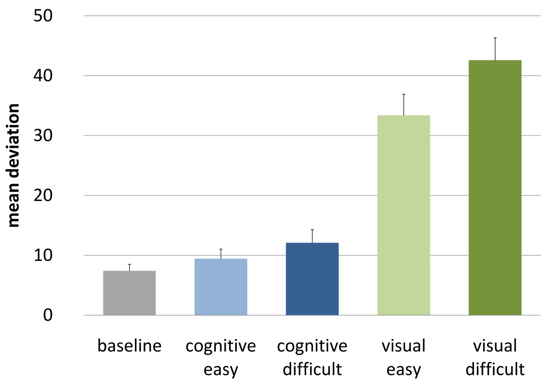


Figure 7: Mean deviation from the centre position.

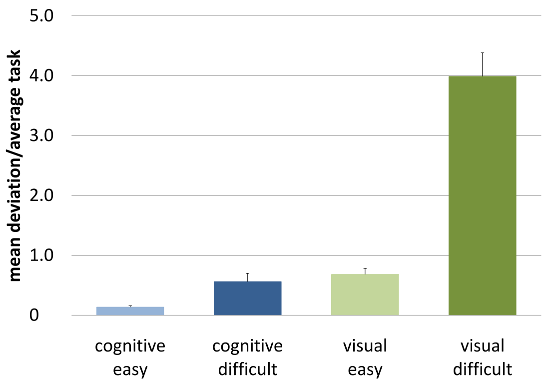


Figure 8: Mean deviation from the centre position, divided by number of completed secondary task trials during the experimental trial.

between task type and task difficulty, $F(1, 20) = 100.16$, $p < .001$, $\eta_p^2 = .83$. As can be seen in Figure 8, an increased task difficulty in the visual task had a higher impact on the performance metric compared to the cognitive task. It has to be emphasised that the absolute values found here cannot be compared to the values from Experiment I or II, since there, the duration of an experimental trial was 180 sec, which obviously results in a higher number of secondary task trials completed, and subsequently in lower mean deviation per task values. To allow for comparison, trials in Experiments I / II and III would either have to have the same length, or some transformation would have to be done (e.g. for Experiment III, calculate the number of secondary task trials that would have been completed in 180 sec before calculating the performance measure).

4.2.2 Subjective ratings

In Figure 9, the results of the subjective ratings are displayed. It is clearly visible that for nearly all dimensions, the difficult version of a task resulted in higher ratings (with the exception of “visual demand” for the cognitive tasks). A two-factorial ANOVA for repeated measures shows a significant effect of task difficulty on all five rating dimensions (see Table 1). Of special interest are also the significant interactions for “visual demand”, “mental demand” and “effort”. Figure 9 shows that an increased task difficulty in the visual task had, as expected, a much higher impact on the rating of “visual demand” compared to the cognitive task. In contrast, an increased task difficulty in the cognitive tasks had stronger impact on “mental demand” and “effort” in comparison to the visual task.

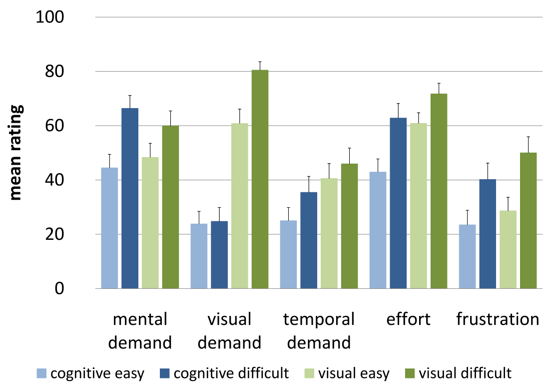


Figure 9: Subjective assessment of the secondary tasks.

Table 1: ANOVA results for subjective assessment of secondary tasks (main effects and interaction).

	task type		task difficulty		interaction	
scale	*F* (1,20)	*p*	*F* (1,20)	*p*	*F* (1,20)	*p*
mental demand	.10	.760	37.29	.000	4.71	.042
visual demand	82.68	.000	21.91	.000	14.40	.001
temporal demand	9.55	.006	6.82	.017	.88	.359
effort	9.87	.005	29.72	.000	5.76	.026
frustration	3.64	.071	31.13	.000	.60	.447

4.3 Discussion

The results of our third experiment further strengthen the claim that the CTT might be a useful method to assess driver distraction. It was able to capture variations in visual as well as in cognitive distraction. It has to be acknowledged that effects found for cognitive distraction were much smaller than for visual distraction. The interaction we found suggests that the CTT is especially suitable to assess visual distraction, whereas smaller variations in cognitive distraction would most certainly not be reflected by changes in performance. Also, distraction as measured by the CTT was in general much higher for visual than for cognitive tasks. However, as driving is primarily visual, this difference as well as the smaller effect between the cognitive tasks might just reflect the "real" relation between visual and cognitive factors in driver distraction. So, whether the limited sensitivity to variations in cognitive demand constitutes a shortcoming of the method or rather is the reflection of real world relations between cognitive and visual distraction would be a matter of further discussion.

The incorporation of secondary task duration into a combined measure again proved to be useful. Effects with and without secondary task duration included did not differ in their direction, but their magnitude, emphasising the fact that distraction is a composite of excess and duration of distraction demand. Subjective measures of the demand that the different tasks induced generally supported our a-priori assumptions about the tasks' properties. On all dimensions, the task conditions that were expected to be more difficult received higher ratings than the easy conditions. Also, as could have been anticipated, the visual tasks scored much higher on "visual demand", which is reflected as well in the

CTT performance. The results of the subjective ratings can therefore be considered an additional validation of the experimental manipulation.

As for the experimental procedure, the change from longer to shorter test trials did not appear to distort results. At the same time, complaints of participants about the heavy demand of the experimental situation decreased. We conclude that, if CTT duration is to be set at a fixed value, 90 sec trials provide enough information for meaningful calculations, while not overloading the participants.

5 Experiment IV

After confirming the general usefulness of the CTT as a method to assess driver distraction through the differentiation between artificial secondary tasks of varying distraction demand, we aimed for a more realistic, i.e. externally valid assessment of the CTT. As the purpose of the method would be to assess the distraction demand especially of driver information systems, we regarded the CTT's validation on a navigation system the next important step. The basic approach was to develop tasks that were objectively more or less distracting, and test whether the CTT is able to reflect this difference in distraction. Research has focused on the visual load of IVIS as a main issue for driver distraction. Wierwille (1993) described the need for visual sampling of the traffic environment for safe driving, which is influenced by the operation of IVIS. As the European statement of principles (CEC, 2008) states: "Increasing the frequency and / or duration of glances required to detect and acquire visually displayed information may increase the risk of potentially dangerous traffic situations caused by driver preoccupation with non-primary driving-related tasks" (p. 13). Since the interaction with IVIS very often goes beyond visual acquisition of information by requiring manual interaction as well, the document also recommends to minimise the number and length of these interactions, argues for interruptability and states that the pace of interaction should be user controlled. We therefore decided to design IVIS-tasks that have an influence on visual sampling and manual control of the CTT by varying the need to attend to the IVIS visually and manually to complete the specified task. We hypothesised that IVIS tasks that are lower in visual and motor demand result in better CTT performance compared to tasks with a higher visual and motor demand.

5.1 Method

5.1.1 Participants

Twenty-five students of Chemnitz University of Technology took part in this experiment, 18 female and 7 male, with a mean age of 21.6 years (SD = 1.5).

5.1.2 Material

Critical tracking task (CTT) The CTT employed in this experiment, as well as the overall setup (screen size & position, operation of task etc.), were identical to the previous experiments.

Easy destination entry task All tasks were performed on a nomadic navigation system which was placed to the right of the CTT, in a position where an aftermarket navigation system would typically be put (again following the LCT ISO draft). The system had a touchscreen to enter information. Destination entry had to be done letter by letter. When the desired destination appeared in a list placed above the letter block, participants were allowed to select it to complete the task (or to proceed to the next step). Participants had to navigate through menus as well, which was instructed and practiced thoroughly before the actual experiment.

The easy destination entry task required the participants to enter a single city name. An example would be “Verona, Italy”, with the systems current position (and therefore the starting point for the route calculation, set by default) always being “Chemnitz, Germany”. Once a route was calculated, participants were supposed to choose the option “avoid toll roads”, which resulted in a recalculation of the route. The destinations were chosen so that the system needed considerable time for the computation of the route and the recalculation. This allowed for a substantial total task time, with only few interactions with the system necessary to complete the task. Overall, the easy destination entry was designed to take approximately 60 sec as a single task.

Difficult destination entry task The difficult destination entry task required to use the function “extended route planning”. Here, participants had to enter a start location (city and street) as well as a destination (city and street). An example would be “Dresden, Waldstraße 1” as start and “Berlin, Hauptstraße 1” as destination. Locations were chosen to be easy to learn and remember (well known German cities, simple street names, always the same street number). Several actions (button / virtual button presses) were necessary to complete the task, whereas the calculation of the route was very fast. Again, the task was designed to take approximately 60 sec as a single task.

Mobile phone synchronisation task As a third task to be assessed through the CTT, we used the mobile phone synchronisation function of the navigation system. The participants’ task was to navigate through the different menus on the navigation system and a mobile phone (as instructed) to synchronise them. This task was designed to take approximately 60 sec as a single task as well. Different from the easy and difficult destination entry tasks, we had no specific a-priori assumptions about the task’s distraction demand in relation to the other two tasks. While the coordination of two different additional devices without doubt creates additional load, there were also phases during the synchronisation process in which no action or attention were required. Therefore, we had to rely on the subjective rating of the task’s demand (especially in relation to the easy and difficult destination entry tasks) as a criterion for validation.

Instrument for subjective ratings We tried to confirm our a-priori assumptions about the easy and difficult destination entry tasks through subjective ratings. Also, we tried to create some criterion against which to validate the CTT performance in the mobile phone synchronisation task. The instrument we used is equal to the questionnaire employed in Experiment III. Again, participants were requested to assess the dual task situation (CTT + destination entry / synchronisation) with their rating.

5.1.3 Procedure

First, participants received instructions on the CTT, followed by a short CTT training. Then, we recorded two baseline trials with the CTT as a single task. After that, participants received instructions on the first of the destination entry tasks (order balanced). They first completed one destination entry as a single task, then an additional one in parallel to the CTT as practice trials. Then, five destinations had to be entered in parallel to operating the CTT in the experimental phase. Each entry was completed as an individual trial (not five entries while continuously operating the CTT), so there were five separate measurements.

The same procedure was followed for the second of the destination entry tasks (single task training, followed by dual task training, followed by five experimental trials). After that, participants received instructions on the mobile phone synchronisation procedure, and completed one experimental trial in this task. The subjective rating scale was filled in at the end of each respective task condition.

5.2 Results

5.2.1 CTT

In Figure 10, CTT standard deviation values are displayed for the four conditions. For the two destination entry tasks, an average deviation over the five test trials was calculated. The baseline condition is again only included for completeness. As can be seen in the figure, difficult destination entry resulted in higher deviations than the two other tasks. A repeated measures ANOVA revealed significant differences between the dual-task conditions, $F(2, 48) = 67.39$, $p < .001$, $\eta_p^2 = .74$. Post-hoc comparisons (Bonferroni-corrected) confirmed a significant difference between difficult destination entry and the other two conditions (both $p < .001$), whereas there was no difference between easy destination entry and the mobile phone synchronisation ($p = .127$).

We again tried to include the task duration in one compound measure together with CTT performance. Different from the previous studies, where there was a fixed duration with a varying number of secondary task trials completed, now, one experimental trial equalled

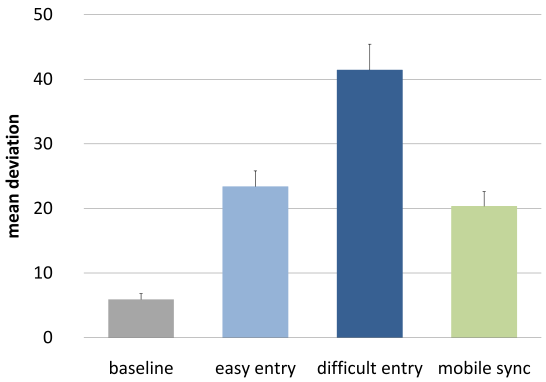


Figure 10: Mean deviation from the centre position.

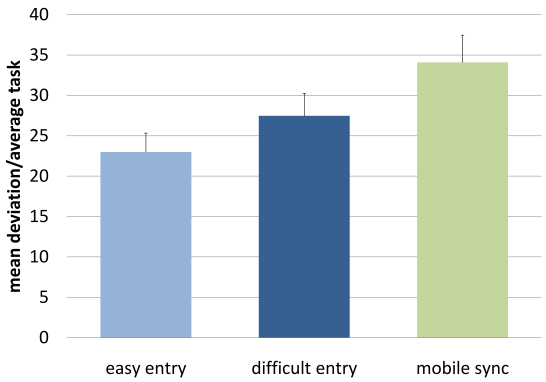


Figure 11: Mean deviation from the centre position, divided by estimated number of completed secondary task trials during a 90 sec trial.

one secondary task trial, and durations varied. We therefore estimated the number of destination entries that would have been completed within 90 sec (the length of one complete experimental trial in Experiment III). We then, as previously, divided the mean deviation by this number. Figure 11 shows a different pattern for the relative deviations compared to the absolute deviations in Figure 10. When taking task duration into account, the mobile phone synchronisation produced the worst performance, while the easy destination entry went with the best results. An ANOVA for repeated measures confirmed an effect of task condition, $F(2, 48) = 27.81$, $p < .001$, $\eta_p^2 = .54$. Post-hoc comparisons (Bonferroni-corrected) showed significant differences between all three dual-task conditions (easy vs. difficult destination entry: $p = .001$; easy and difficult destination entry vs. mobile phone synchronisation both $p < .001$).

5.2.2 Subjective ratings

In Figure 12, the results of the subjective ratings are displayed. The general pattern is the same for all five scales - difficult destination entry is scored highest, mobile phone synchronisation lowest. ANOVAs for repeated measures and post-hoc analysis (Bonferroni corrected) confirmed this picture, as nearly all comparisons reached statistical significance (Table 2).

As the subjective ratings were supposed to serve as an indicator of the CTT's accuracy in assessing the mobile phone synchronisation task's distraction potential, we calculated the correlation between CTT performance (CTT standard deviation & compound measure) and subjective rating (all scales) over all conditions to find out whether the overall pattern of results is similar. We found significant correlations between CTT standard deviation and

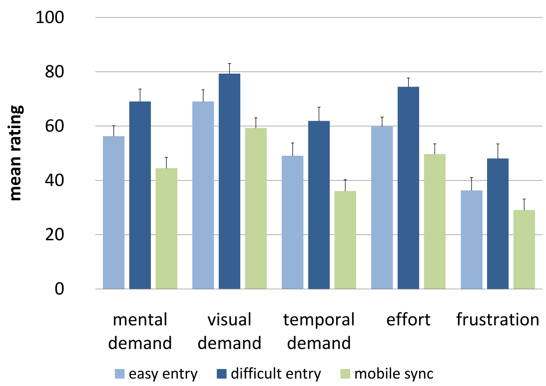


Figure 12: Subjective assessment of the secondary tasks.

Table 2: ANOVA results and post-hoc pairwise comparison for subjective assessment of secondary tasks (easy = easy destination entry, difficult = difficult destination entry, mobile = mobile phone synchronisation).

	ANOVA		easy vs. difficult	easy vs. mobile	difficult vs. mobile
scale	*F* (2,48)	p	*p*	*p*	*p*
mental demand	21.98	.000	.002	.026	.000
visual demand	16.83	.000	.004	.084	.000
temporal demand	23.46	.000	.001	.015	.000
effort	33.30	.000	.000	.005	.000
frustration	9.70	.000	.063	.235	.001

all five scales (from r = .24 to r = .52). None of the scales correlated significantly with the compound measure.

5.3 Discussion

In our fourth experiment, we found evidence for the CTT's capabilities as a measure for distraction that can serve beyond purely artificial tasks. It was able to reliably differentiate between a simple and difficult navigation system task. Subjective measures compliment these results. As for the mobile phone synchronisation task, the pattern of results found in the subjective ratings appears to be comparable to that for the CTT standard deviation, which can be interpreted as an indicator of the CTT's accuracy in the tasks assessment. The visual inspection of the results figures strengthens this claim.

Interestingly, when including task duration, the picture changes. The mobile phone synchronisation took much longer to be completed, which is reflected in the comparatively high values. Those values do not correspond with the subjective ratings. Whether this has to be interpreted as a lack of validity of the CTT, or rather a shortcoming of the subjective measure employed is a matter of discussion. Assuming that the CTT's assessment is indeed valid, it would appear that participants do not consider task duration when reporting subjective assessments of the respective tasks.

6 General discussion and conclusions

We conducted four experiments in order to find out whether the critical tracking task (CTT) has the potential to be a useful measure to assess driver distraction elicited by in-vehicle information systems (IVIS). The results of all four experiments point into this direction. The task was able to reflect simple manipulations in distraction demand of secondary tasks (Experiments I and II). It was able to assess different degrees of visual and cognitive distraction (Experiment III). And it succeeded in differentiating between actual IVIS tasks of varying demand (Experiment IV). In addition, the differences the CTT found corresponded with subjective assessments of workload in most cases. It appears that the CTT can serve as a method to assess driver distraction.

Despite the clear results, questions remain. The λ-value (which defines the level of difficulty in the CTT) used for the experiments was chosen on the basis of the impression that this value represents a medium level of difficulty. It is unclear if the pattern of results would have been identical if the task would have been easier or more difficult. Which level of difficulty most appropriately reflects the load experienced in a “normal” driving situation, or probably in an “emergency” situation still has to be assessed. Also, the choice to change the bar’s colour once it deviates too far from the centre of the screen is debatable. It is certainly true that for some driving situations, additional stimuli help to recover from driving errors (e.g. rumble strips when crossing lanes). However, there are probably even more situations where the opposite is true, situations in which no distinct stimulus assists to correct such errors. Whether not changing the bar’s colour would result in different patterns of data needs further investigation.

One major aspect of the CTT’s value is its simplicity. It is very easy to employ, requires only minimum practice, and is very easy to analyse. At the same time, it is very flexible. When comparing it to the lane change task (LCT; to which it bears some resemblance as both are tracking tasks, and the dependent measures for both are the deviations from some ideal position), it appears that the CTT has advantages in most of the aforementioned aspects. The hardware necessary for the CTT’s use (standard PC) is even simpler than for the LCT (standard PC + game steering wheel). Also, there is less need for training, as the definition of the bar’s ideal position in the CTT is much more straightforward than the normative path defined in the LCT, which requires some explanation and respective training. In addition, the possibility to use secondary tasks of different length, without the need to repeat the tasks or to stop them halfway through allows for presumably more accurate assessments of real-world tasks. This simplicity, paired with the results obtained in the reported experiments, make the CTT an interesting candidate for the assessment of driver distraction caused by in-vehicle devices.

References

Ascone, D. S. (2009). *An examination of driver distraction as recorded in NHTSA databases* (DOT HS 811 216). Washington, DC: NHTSA.

Barnett, G., Licko, V., & Thompson, T. (1985). Behavioral pharmacokinetics of marijuana. *Psychopharmacology, 85*, 51-56.

Baumann, M., Keinath, A., Krems, J. F., & Bengler, K. (2004). Evaluation of in-vehicle HMI using occlusion techniques: Experimental results and practical implications. *Applied Ergonomics, 35*, 197-205.

Burke, M., Gilson, R., & Jagacinski, R. (1980). Multi-modal information processing for visual workload relief. *Ergonomics, 23*, 961-975.

Burns, M., & Moskowitz, H. (1980). Effects of diphenhydramine and alcohol on skills performance. *European Journal of Clinical Pharmacology, 17*, 259-266.

Commission of the European Communities (2008). Commission recommendation of 26 May 2008 on safe and efficient in-vehicle information and communication systems: Update of the European statement of principles on human-machine interface. *Official Journal of the European Communities, L216*, 1-42.

Derrick, W. (1988). Dimensions of operator workload. *Human Factors, 30*, 95-110.

Distracted-driving campaign requires attitude adjustment. (2010, January 26). *USA Today*, p. 8A.

Elmenhorst, D., Elmenhorst, E., Luks, N., Maass, H., Mueller, E., Vejvoda, M., et al. (2009). Performance impairment during four days partial sleep deprivation compared with the acute effects of alcohol and hypoxia. *Sleep Medicine, 10*, 189-197.

Greco, M., Stucchi, N., Zavagno, D., & Marino, B. (2008). On the portability of computer-generated presentations: The effect of text-background color combinations on text legibility. *Human Factors, 50*, 821-833.

Green, P. (1999). *Visual and task demands of driver information systems* (UMTRI-98-16), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Groeger, J. A. (2000). *Understanding driving: Applying cognitive psychology to a complex everyday task*. Hove, East Sussex, UK: Psychology Press Ltd.

Harbluk, J. L., Mitroi, J. S., & Burns, P. C. (2009). Three navigation systems with three tasks: Using the lane-change test (LCT) to assess distraction demand. In *Proceedings of the 5th international driving symposium on human factors in driver assessment, training, and vehicle design*, Big Sky, MT, pp. 24-30.

Harbluk, J. L., Noy, Y. I., Trbovich, P. L., & Eizenman, M. (2007). An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accident Analysis and Prevention, 39*, 372-379.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: Elsevier.

ISO 15008. (2009). *Road vehicles - Ergonomic aspects of transport information and control systems - Specifications and test procedures for in-vehicle visual presentation*. Geneva, Switzerland: ISO.

ISO 16673. (2007). *Road vehicles - Ergonomic aspects of transport information and control systems - Occlusion method to assess visual demand due to the use of in-vehicle systems*. Geneva, Switzerland: ISO.

Jahn, G., Oehme, A., Krems, J. F., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F, 8*, 255-275.

Jex, H. R. (1967). Two applications of a critical-instability task to secondary work load research. *IEEE Transactions on Human Factors in Electronics, 8*, 279-282.

Jex, H. R., McDonnell, J. D., & Phatak, A. V. (1966). A "critical" tracking task for manual control research. *IEEE Transactions on Human Factors in Electronics, 7*, 138-145.

Karlsson, I. C. M., Rämä, P., Alonso, M., Engelbrektsson, P., Franzén, S., Henar Vega, M., et al. (2009). *TeleFOT D.2.2.1 Testing and Evaluation Strategy*.

Keinath, A., Baumann, M., Gelau, C., Bengler, K., & Krems, J. F. (2001). Occlusion as a technique for evaluating in-car displays. In D. Harris (Ed.), *Engineering psychology and* *cognitive ergonomics-Aerospace and transportation systems* (pp. 391-397). Aldershot, UK: Ashgate.

Kimura, K., Marunaka, K., & Sugiura, S. (1997). Human factors considerations for automotive navigation systems-Legibility, comprehension, and voice guidance. In Y. I. Noy (Ed.), *Ergonomics and safety of intelligent driver interfaces* (pp. 153-167). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Klein, R., & Jex, H. (1975). Effects of alcohol on a critical tracking task. J*ournal of Studies on Alcohol, 36*, 11-20.

Kuratorium für Verkehrssicherheit. (2008). *Road accident statistics 2007*. In Reihe Verkehr in Österreich. Heft 40, Wien: KfV.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. In H. Strasser et al. (Eds.), *Quality of work and products in enterprises of the future* (pp. 57-60). Stuttgart: Ergonomia.

Mattes, S., & Hallén, A. (2009). Surrogate distraction measurement techniques: The lane change test. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver distraction. Theory, effects, and mitigation* (pp. 107-122). Boca Raton, FL: CRC Press.

Miura, T. (1986). Coping with situational demands: A study of eye movements and peripheral vision performance. In A. G. Gale et al. (Eds.), *Vision in vehicles-II* (pp. 126-137). Amsterdam: Elsevier.

New Zealand Ministry of Transport. (2010). *Reported injury crashes 2009*.

O'Donnell, R., Moise, S., & Schmidt, R. (2005). Generating performance test batteries relevant to specific operational tasks. *Aviation, Space, and Environmental Medicine, 76(7 suppl.)*, C24-C30.

Pauzié, A. (2002). In-vehicle communication systems: The safety aspect. *Injury Prevention, 8(4 suppl.)*, 26-29.

Petzoldt, T., Bär, N., Ihle, C., & Krems, J. F. (2011). Learning effects in the lane change task (LCT) - Evidence from two experimental studies. *Transportation Research Part F, 14*, 1-12.

Pickrell, T. M., & Ye, T. J. (2010). *Driver electronic device use in 2009* (DOT HS 811 372). Washington, DC: NHTSA.

Ramaekers, J., Kauert, G., van Ruitenbeek, P., Theunissen, E., Schneider, E., & Moeller, M. (2006). High-potency marijuana impairs executive function and inhibitory motor control. *Neuropsychopharmacology, 31*, 2296-2303.

Ramaekers, J., Moeller, M., van Ruitenbeek, P., Theunissen, E., Schneider, E., & Kauert, G. (2006). Cognition and motor control as a function of Delta9-THC concentration in serum and oral fluid: Limits of impairment. *Drug and Alcohol Dependence, 85*, 114-122.

Ramaekers, J., Muntjewerff, N., van Veggel, L., Uiterwijk, M., & O'Hanlon, J. (1998). Effects of nocturnal doses of mirtazapine and mianserin on sleep and on daytime psychomotor and driving performance in young, healthy volunteers. *Human Psychopharmacology: Clinical and Experimental, 13(0 suppl.)*, S87-S97.

Ramaekers, J., Uiterwijk, M., & O'Hanlon, J. (1992). Effects of loratadine and cetirizine on actual driving and psychometric test performance, and EEG during driving. *European Journal of Clinical Pharmacology, 42*, 363-369.

Regan, M. A., Lee, J. D., & Young, K. L. (Eds.) (2009). *Driver distraction: Theory, effects and mitigation*. Boca Raton, Florida: CRC Press.

Rehmann, J., Stein, E., & Rosenberg, B. (1983). Subjective pilot workload assessment. *Human Factors, 25*, 297-301.

Rosenberg, B. L., Rehmann J., & Stein, E. S. (1982). *The relationship between effort rating and performance in a critical tracking task* (DOT/FAA/CT-82/66). Washington, DC: Federal Aviation Administration Office of Systems Engineering Management.

Senders, J. W., Kristofferson, A. B., Levison, W. H., Dietrich, D. W., & Ward, J. L. (1967). The attentional demand of automobile driving. *Highway Research Record, 195*, 15-33.

Starry, C. (2001). The emerging in-vehicle intelligent transportation systems market. *Business Economics, 32*, 49-55.

Stevens, A., Quimby, A., Board, A., Kersloot, T., & Burns, P. (2002). *Design guidelines for safety of in-vehicle information systems*. Project Report PA3721/01. Crowthorne, UK: Transport Research Laboratory TRL.

Stutts, J. C., Reinfurt, D. W., Staplin, L., & Rodgman, E. A. (2001). *The role of driver distraction in traffic crashes*. Washington, DC: AAA Foundation for Traffic Safety.

US Department of Transportation. (2009, October 1). *U.S. Transportation Secretary Ray LaHood announces administration wide effort to combat distracted driving* [Press release]. Retrieved from http://distraction.gov/files/for-media/10.01.09.pdf

Wickens, C., Braune, R., & Stokes, A. (1987). Age differences in the speed and capacity of information processing: 1. A dual-task approach. *Psychology and Aging, 2*, 70-78.

Wickens, C., & Kessel, C. (1980). Processing resource demands of failure detection in dynamic systems. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 564-577.

Wierwille, W. W. (1993). Demands on driver resources associated with introducing advanced technology into the vehicle. *Transportation Research Part C, 1*, 133-142.

Young, K., Regan, M., & Hammer, M. (2003). *Driver distraction: A review of the literature*. Melbourne: Monash University, Accident Research Centre.

Curriculum Vitae

Dipl.-Psych. Tibor Petzoldt

born: 30.11.1981 in Schlema, Germany

citizenship: German

Office address

Chemnitz University of Technology

Department of Psychology

Cognitive and Engineering Psychology

Wilhelm-Raabe-Straße 43

09120 Chemnitz, Germany

Contact

phone: +49 371 531 35568

facsimile: +49 371 531 835568

e-mail: tibor.petzoldt@psychologie.tu-chemnitz.de

Academic education

10/2001 - 09/2006	Chemnitz University of Technology, Chemnitz; Diploma Psychology, focus on occupational psychology and human factors
10/2003 - 09/2006	Student assistant at the Professorship of Cognitive and Engineering Psychology
09/2006	Thesis: "Die Erfassung kognitiver und visueller Beanspruchung durch Fahrerinformationssysteme" [engl.: "The assessment of cognitive and visual demand of driver information systems"] Degree: Diploma (final grade 1.3)

Professional experience

since 10/2006	Researcher & lecturer at the Professorship of Cognitive and Engineering Psychology, Chemnitz University of Technology, Chemnitz

Publications

Journal Papers, Book Chapters etc.

Barnard, Y., Risser, R., Kaufmann, C., Krems, J. F., & Petzoldt, T. (2011). The future of IDSS. In Y. Barnard, R. Risser, & J. F. Krems (Eds.). *The safety of intelligent driver support systems. Design, evaluation and social perspectives* (pp. 149-169). Farnham, UK: Ashgate.

Krems, J. F., & Petzoldt, T. (2011). Tools and procedures for measuring safety-relevant criteria. In Y. Barnard, R. Risser, & J. F. Krems (Eds.). *The safety of intelligent driver support systems. Design, evaluation and social perspectives* (pp. 93-109). Farnham, UK: Ashgate.

Petzoldt, T., Bär, N., Ihle, C., & Krems, J. F. (2011). Learning effects in the lane change task (LCT) - Evidence from two experimental studies. *Transportation Research Part F, 14*, 1-12. doi:10.1016/j.trf.2010.09.001

Krems, J. F., Petzoldt, T., & Henning, M. (Eds.). (2010). *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, Berlin, Germany, April 29-30 2010*. Lyon: Humanist Publications.

Lietz, H., Henning, M., Petzoldt, T., Krems, J. F., & Wanielik, G. (2010). A software tool for the visualization and annotation of naturalistic driving data stored in relational databases. In J. F. Krems, T. Petzoldt & M. Henning (Eds.). *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, Berlin, Germany, April 29-30 2010* (pp. 427-433). Lyon: Humanist Publications.

Petzoldt, T., Weiss, T., Franke, T., Krems, F., & Bannert, M. (2010). Improving driver education with multimedia applications. In K. Gustafson (Ed.), *Proceedings of the 15th International Conference "Road Safety on Four Continents"* (pp. 653-663).

Krems, J. F., Henning, M., & Petzoldt, T. (2009). Methoden zur Messung von Fahrerverhalten: Vom Labor bis zur Feldstudie. In H. Winner & R. Bruder (Eds.), *Wie realitätsnah lässt sich Fahrerverhalten messen?* (pp. 103-111). Stuttgart: ergonomia.

Mikolajetz, A., Henning, M., Tenzer, A., Zobel, R., Krems, J. F., & Petzoldt, T. (2009). Curve negotiation: Identifying driver behavior around curves with the driver performance database. In *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Big Sky, MT (pp. 391-397).

Petzoldt, T., Bär, N., & Krems, J. F. (2009). Gender effects on lane change test (LCT) performance. In *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Big Sky, MT (pp. 90-96).

Weiss, T., Bannert, M., Petzoldt, T., & Krems, J. F. (2009). Einsatz von computergestuetzten Medien und Fahrsimulatoren in Fahrausbildung, Fahrerweiterbildung und Fahrerlaubnispruefung. *Berichte der Bundesanstalt für Straßenwesen, Reihe M (Mensch und Sicherheit)*, Heft 202.

Baumann, M., Petzoldt, T., Groenewoud, C., Hogema, J., & Krems, J. F. (2008). The effect of cognitive tasks on predicting events in traffic. In C. Brusque (Ed.), *Proceedings of the European Conference on Human Interface Design for Intelligent Transport Systems* (pp. 3-11). Lyon: Humanist Publications.

Baumann, M., Petzoldt, T., & Krems, J. F. (2006). Situation Awareness beim Autofahren als Verstehensprozess. *MMI-Interaktiv, 11*, 43-57.

Presentations / Posters

Petzoldt, T., Bellem, H., & Krems, J. F. (2011). Die Critical Tracking Task - Eine geeignete Methode zur Erfassung von Fahrerablenkung?. In K. Bittrich, S. Blankenberger & J. Lukas (Eds.), *Beiträge zur 53. Tagung experimentell arbeitender Psychologen* (p. 131). Lengerich: Pabst Science Publishers.

Petzoldt, T., Weiss, T., Franke, T., Krems, J. F., & Bannert, M. (2010). Effekte eines multimedialen Trainings zur Fahranfängervorbereitung. In F. Petermann & U. Koglin (Eds.), *47. Kongress der Deutschen Gesellschaft für Psychologie, 26.-30. September 2010* (p. 434). Lengerich: Pabst Science Publishers.

Bär, N., & Petzoldt, T. (2009). Die Auswirkung von Training auf die Spurwechselperformanz - Implikationen für die Kalibrierung der LCT. In A. Eder, K. Rothermund, S.

Schweinberger, M. Steffens & H. Wiese (Eds.), *51. Tagung experimentell arbeitender Psycholog/innen in Jena* (p. 100). Lengerich: Pabst Science Publishers.

Petzoldt, T., Bär, N., Ihle, C., & Krems, J. F. (2009). The LCT - A valid measure to assess driver distraction? *First International Conference on Driver Distraction and Inattention, Gothenburg, 28.-29. September* (A60-O).

Petzoldt, T., Weiss, T., Krems, J. F., & Bannert, M. (2009). Wirksamkeit neuer Medien in der Fahranfängervorbereitung. In C. Korunka, E. Kirchler & H. Ulferts (Eds.), *Entscheidungen und Veränderungen in Arbeit, Organisation und Wirtschaft* (p. 84). Wien: Falcultas Verlangs- und Buchhandels AG.

Weiß, T., Petzoldt, T., Bannert, M., & Krems, J. F. (2009). Wirksamkeit des Einsatzes von Lernsoftware zur Optimierung der Fahranfängervorbereitung. *Presentation given at the 2. Fachtagung Pädagogische Psychologie der Deutschen Gesellschaft für Psychologie (DGPs), Saarbrücken, 7.-9. September*.

Baumann, M., Petzoldt, T., Groenewoud, C., Hogema, J., & Krems, J. F. (2008). The effect of cognitive distraction on predicting traffic events. *Poster presented at the Annual Meeting of the Human Factors and Ergonomics Society Europe Chapter, Soesterberg, NL, 15.-17. October*.

Baumann, M., Petzoldt, T., Groenewoud, C., Hogema, J., & Krems, J. F. (2008). Working memory processes as the cognitive basis of driver's situation awareness. *4th International Conference on Traffic & Transport Psychology, Washington DC, 31. August-4. September* (O68).

Petzoldt, T., & Krems, J. F. (2008). Can simple measures of driving performance be distorted by different hardware configurations? *Poster presented at the Annual Meeting of the Human Factors and Ergonomics Society Europe Chapter, Soesterberg, NL, 15.-17. October.*

Petzoldt, T., & Baumann, M. (2008). Die Auswirkung kognitiver Belastung auf die Antizipation von Verkehrsereignissen. In P. Khader, K. Jost, H. Lachnit & F. Rösler (Eds.), *Experimentelle Psychologie. Beiträge zur 50. Tagung experimentell arbeitender Psychologen*. Lengerich: Pabst Science Publishers.

Miscellaneous

Petzoldt, T., & Henning, M. (2008). 4. International conference on traffic & transport psychology (Tagungsbericht). *Zeitschrift für Verkehrssicherheit, 54*, 219-220.

Eidesstattliche Erklärung

Die vorliegende Arbeit wurde von mir selbständig verfasst.

Ich habe keine anderen als die angegeben Hilfsmittel benutzt.

Tibor Petzoldt

Chemnitz, den 24.05.2011